Draft – Please do not cite
August 2008
Development Economics: Theory, Empirical Research and Policy Analysis
Julie Schaffner

Chapter 4
Economic Growth: Theories in Historical Perspective

The exposition in Chapter 3 provides the reader with a contemporary and eclectic perspective on the sources of economic growth, and on the roles that development actors might play in speeding or slowing growth. It reflects the contemporary understanding that

- growth involves a diverse mix of increases in physical and human capital, improvements in technology and efficiency, and reductions in waste;
- markets play important roles in providing people with incentives to make choices that set these processes into motion, though markets may sometimes fail to provide adequate incentive;
- policies of governments and other development actors have the potential to improve some of these choices, but also the potential to inhibit good choices; and
- good growth policy design requires careful, disaggregated and context-specific study that is sensitive to the role of institutions in shaping policy impacts.

Growth has not always been understood in these broad terms. In the early years after World War II, when the field of economic development was born, growth was thought to arise primarily out of increases in physical capital, and the world was thought to be filled with rigidities that caused markets to fail almost without exception, leaving successful growth entirely in the hands of governments. This chapter describes the evolution of economic thought regarding growth in developing countries from then until now, paying special attention to the evolution of theoretical growth models.

The 1940s and 1950s: Capital fundamentalism, structuralism and dualism

Economic thought regarding growth in developing countries in the 1940s and 1950s was characterized by capital fundamentalism, structuralism and dualism. **Capital fundamentalism** refers to the belief that the most important proximate source of economic growth is the accumulation of physical capital. The emphasis economists placed on capital accumulation in the 1940s is understandable, given the perception at the time that the Soviet Union's recent industrialization through rapid (and forced) saving and physical capital investment had been a great success.

Structuralism refers to the belief that the world is a rigid place, in which economic actors – especially the poor and tradition-bound peasants that were thought to populate developing countries – have little ability or inclination to alter what they do in response to economic incentives. When, for example, labor is plentiful and capital is scarce, the price of capital might rise, but such price increases would not set into motion self-equilibrating market processes, in which people decide to invest in the creation of more capital or to shift to production methods that make greater use of cheap labor and rely less on costly capital. In such a world unemployment can remain rampant for long periods without giving rise to any market forces that might tend to absorb the unemployed. This view was rendered compelling by the recent experience of the Great Depression.

Dualism refers to the belief that developing economies must be understood as having two sectors, traditional and modern, each with peculiar characteristics and very different growth prospects. The large traditional sector, often identified with agriculture, was assumed to be "backward," in the sense that production was overseen by poor, irrational peasants who were not driven by modern profit-maximizing motivations, who operated at very low levels of productivity, and who would have little prospect for improving their productivity even if they were inclined to try. The modern sector, often identified with manufacturing, was by contrast seen as a potentially dynamic sector, though still small in most developing countries. It was in modern manufacturing that physical capital accumulation was thought to hold tremendous promise of increasing productivity and incomes, and where producers were more inclined to plow growing profits back into more saving and investment.

The predominant perspective on economic growth during this period had profound policy implications. Capital fundamentalism dictated the importance of raising rates of saving and physical capital accumulation. Structuralism dictated that the government, perhaps with foreign assistance, would have to play an important role in bringing these increases about. Dualism dictated that the new capital be concentrated in the modern manufacturing sector, and that such "industrialization" be encouraged through subsidies and protection of manufacturing enterprises from international trade, because efforts to increase manufacturing profits would redound in increased rates of saving, investment and future growth.

The following sections offer a general introduction to theoretical growth models and then exposit two influential theoretical growth models of the 1940s and 1950s. The Harrod-Domar Model, which influenced thinking about growth in developed as well as developing countries, exhibits both capital fundamentalist and structuralist thinking. The Lewis Model incorporates dualism as well, and was considered useful primarily for understanding developing countries.

An Introduction to Theoretical Growth Models

Before examining our first theoretical model of economic growth, it is useful to introduce the reader to the general structure and purpose of theoretical growth modeling exercises. A **theoretical growth model** is defined by a set of assumptions regarding the way economies work, which the model builder considers the most salient features of the economy for understanding economic growth. The purpose of the modeling exercise is to figure out what the assumptions imply about how and why growth rates might change over time or differ across countries, often with the ultimate aim of providing guidance to policymakers interested in encouraging more rapid economic growth.

The assumptions on which growth models are based fall into two categories, which we may label "technological" and "behavioral". **Technological assumptions** describe physical constraints on the maximum quantities of goods and services that an economy's firms may produce from various combinations of inputs. Simple growth models treat the entire economy as if it were a single firm, and formalize the technological assumptions by writing down a single "aggregate production function", which describes the maximum quantity of total GDP that may be achieved at any point in time, as a function of the national aggregate quantities of various inputs that are employed in production and the current state of technical knowledge and expertise.

The technological assumptions underlying growth models have evolved over time in several important ways. First, early models assumed that GDP was a function of only two inputs: labor and physical capital. Only later did model builders recognize human capital as an additional important input to production. Second, while the earliest models assumed a very rigid technology, in which producers faced only a single option for combining labor and capital to produce output, later models assumed that a wider array of technological options give producers flexibility to combine labor and other inputs in varying proportions. Third, while the earliest models assumed that production was characterized by constant returns to scale and constant marginal returns to capital, these were replaced in the 1960s by models assuming constant returns to scale but diminishing marginal returns to capital, which were in turn replaced in the 1980s by models in which technologies are assumed to be characterized by increasing returns to scale and constant or increasing marginal returns to capital. We will see the significance of these changing assumptions for our understanding of economic growth below.

Growth models' **behavioral assumptions** describe how the inputs available at any point in time are allocated across activities (e.g. production of goods and services, research and development activities, or unemployment), and how the quantities of inputs available in the economy and the economy's technological options tend to evolve over time. We call these "behavioral" assumptions, because they describe the ways in which the economy's firms choose what to produce, how to produce it and what investments to make, the ways in which the economy's households choose where to work, what to consume and what to save, and the ways in which markets work (or fail to work) when producers and consumers interact with each other. The simplest models assume that firms and households follow simple rules of thumb, such as the rule that households always save a certain percentage of their income, and assume that key features of the economy (such as the "level of technology") always grow at exogenously fixed rates. More elaborate models build up assumptions from more detailed and thorough-going micro-economic foundations, describing producers and households as seeking to maximize profits or utility subject to the constraints they face.

Behavioral assumptions, too, have evolved over time in several important ways. First, the earliest models assumed that producers and households had little ability or inclination to modify their choices in response to economic conditions, and that, as a result, markets would fail to guide all the economy's inputs into their most productive uses. These were replaced by models making almost diametrically opposed assumptions, in which producers and households were assumed to face great flexibility and markets were assumed to function perfectly. More recently yet, these have been replaced by models staking out a more nuanced, intermediate position, in which producers and households are flexible and responsive to economic conditions, but in which markets nonetheless might fail to direct all inputs into their best possible use. Second, the earliest models largely took technological conditions as fixed and unchanging. These were replaced by models in which technology was assumed to improve over time, for reasons that were unexplained within the model. More recently, model builders have tried to identify actors within their models that are responsible for generating technological improvements and to formulate assumptions about how these actors make the decisions that determine the rate of **technical change,** or the rate of technological improvement.

Model builders make their assumptions specific and clear by writing them down as a system of mathematical equations. These equations involve both **exogenous** parameters, whose values are taken as given from outside the model, and **endogenous** variables, whose values are determined (as functions of the exogenous parameters) by solving the system of equations. The most important endogenous variable in all these models is, of course, the rate of economic growth. While early models treated only the growth rate as endogenous, later models have treated a growing array of variables – such as rates of investment and technical change – as jointly endogenous with the growth rate. For each model we examine below, we will describe the technological and behavioral assumptions, and their implications for: (a) how the growth rate would evolve over time if the parameters of the model remained constant; and (b) how a country's growth trajectory would change as a result of changes in key exogenous parameters. In what follows we elaborate on this evolution of theoretical assumptions regarding economic growth, and describe the intellectual process of interplay between theoretical predictions and empirical observations that has driven this evolution.

The Harrod-Domar Growth Model

The ideas underlying what is now known as the Harrod-Domar Growth Model were introduced independently in papers by Roy Harrod (1939) and Evsey Domar (1946). At the center of the Harrod-Domar Model is a very simple and extreme assumption regarding the economy's production technology: a fixed quantity (u) of labor and a fixed quantity (v) of physical capital are required to produce each unit of GDP. That is, the aggregate production function is assumed to take the following extreme form:

$$(1)\ Y = min[(\frac{1}{u})L, (\frac{1}{v})K],$$

where Y is total GDP, L is the total quantity of labor, and K is the total quantity of physical capital. The $min[.,.]$ function indicates that total output is equal to the smaller of the two arguments. The parameters u and v are technological parameters, the "labor-output ratio" and the "capital-output ratio", which are unchanging within this inflexible technology. Notice several things about the technological assumptions embodied in (1).

First, it assumes that capital and labor are the only important inputs to production. Second, it assumes that technological options are few and rigid. Indeed, labor and capital must always be combined in exactly the same proportion (u/v units of labor per unit of capital) to produce a unit of output; for this reason the technology is labeled a **fixed proportions production technology**. If the ratio of labor to capital in the economy is greater than u/v, then some labor has no use in production and will remain unemployed, while if the ratio of labor to capital is less than u/v, some capital will remain unemployed. Third, it assumes that production is characterized by constant returns to scale. If the quantities of labor and capital were both doubled, output would exactly double.

The model's behavioral assumptions offer simple descriptions of how K and L evolve over time, and how given stocks of K and L are utilized at any point in time. The capital stock, K, is increased through investment, which (in a closed economy) must be financed by saving. The critical behavioral assumption driving growth in K then, is the assumption that people save a fixed proportion (s) of their total income. Capital accumulates when this saving is greater than the rate at which old capital depreciates. The labor force, L, is assumed to grow at an exogenously fixed rate, n.[1] Significantly, the technological parameters u and v are taken as exogenously fixed, with no built-in tendency to evolve over time. Labor is assumed (at most times) to be "abundant" relative to capital, in the sense that the current stock of capital is too small to provide productive employment for all workers. (This notion of labor abundance makes sense only as a result of the fixed proportions production technology, which implies that the current stock of capital, K, can provide productive employment for only $(u/v)K$ units of labor. Labor in excess of this amount remains unemployed, or at least unproductive.) In practice, then, output can be treated as a function of K alone, because the economy's stock of labor never represents a constraint on production.

The assumptions of the previous two paragraphs may be summarized by a simple set of mathematical equations. Combining the technological assumptions with the assumption that labor never represents a constraint on production, we can state the economy's effective production technology as:

$$(2)\ Y = \left(\frac{1}{v}\right) K$$

This says that total GDP is just a constant multiple of the size of the capital stock. Denoting the derivative of a variable x with respect to time by $\dot{x}$, we can re-state (2) in a way that illustrates the link between increases in Y and increases in K:

$$(3)\ \dot{Y} = \left(\frac{1}{v}\right) \dot{K}\ .$$

[1] Throughout this chapter we will essentially treat the labor force as a fixed proportion of the total population, in which case the growth rate of the labor force is just the population growth rate, and growth in output per person is identical to growth in output per worker.

That is, Y rises and falls strictly in proportion to K. Notice that as long as labor remains abundant relative to capital (so that this equation is a good description of aggregate production), the economy is characterized by constant marginal returns to capital: each additional unit of capital increases Y by the same amount.

In light of (3), if we are to understand growth in Y, we must understand growth in K. Increases in the capital stock ($\dot{K}$) are achieved through investment in excess of the rate of depreciation of the old capital stock. Invoking the assumption that people always save a fixed proportion of their income, and letting d represent the fraction of the old capital stock that decays or becomes obsolete each year, we can describe the change in size of the capital stock in a short interval of time as

$$(4)\ \dot{K} = sY - dK$$

Equations (3) and (4) fully describe the assumptions of the model. We "solve" the model, seeking to understand its implications for growth, by solving this set of equations for a statement describing how the rate of economic growth ($\dot{Y}/Y$) is determined as a function of the key exogenous parameters of the model. Substituting (4) into (3), we find that

$$(4)\ \dot{Y} = (\frac{1}{v})(sY - dK)$$

or (dividing both sides by Y, and recognizing that $v=K/Y$)

$$(5)\ \frac{\dot{Y}}{Y} = \frac{s}{v} - (\frac{d}{v})(\frac{K}{Y}) = \frac{s}{v} - d$$

Equation (5) describes how the model's main endogenous variable (the growth rate of total GDP) is determined as a function of the model's parameters.[2] We conclude from this expression that as long as the parameters remain fixed, growth will proceed at a constant rate. Increasing the growth rate would require increasing the saving rate s or reducing the capital-output ratio v (assuming the rate of depreciation is unalterable). In the model's world of rigidities and poorly functioning markets, increases in s are unlikely in the absence of intervention by developing country governments. And these governments might in turn require assistance from the world's rich countries. Reductions in the technological parameter v, which require improvements in efficiency or technology allowing producers to achieve more Y from a given quantity of K, were thought even more difficult to engineer even by government. The main "conclusion" of the model thus comes as little surprise given the assumptions on which it is built: to increase the growth

[2] Notice that while subsequent growth models focus on the determinants of the rate of growth of GDP per capita, the Harrod-Domar model focuses primarily on the rate of growth of total GDP. Only if the rate of growth of total GDP is faster than the rate of growth of the population is GDP per capita growing.

rate a country government must increase the rate of investment by the public sector, financing that investment through taxes and foreign aid.

The model has additional implications that are perhaps less obvious, and certainly less compelling, especially in light of historical experience. In this model the growth of K (and thus the rate of growth in the number of productive jobs) is governed by the saving rate, while growth in the number of workers (L) is governed by unrelated forces. It can be shown (see problem 1 below) that K (like Y) grows at the rate $(s/v)-d$, while the labor force grows at the rate n, which need not be the same. If K grows more rapidly than L then the number of jobs in the economy grows more rapidly than the number of workers, and unemployment falls. Under such circumstances, GDP per capita rises, even though the productivity of each employed worker remains fixed, simply because the share of workers who are in unproductive unemployment falls. If capital continues to grow more rapidly than the labor force, production might eventually "use up" all available labor. In this case further increases in capital accumulation bid up wages and prices, generating inflation without increasing output, and growth stops. If, however, K grows less rapidly than L, then unemployment must rise and GDP per capita falls. Unless $s/v-d$ just happens to equal n, the economy will be marked by rising or falling unemployment. This **knife-edge** property of the model implies perpetual instability. Furthermore, if $s/v-d$ did just equal n, total GDP would rise but GDP per capita would remain constant. As we will see, the mismatch between these implications and real world experience contributed to the eventual rejection of the Harrod-Domar model.

The basic ideas underlying the Harrod-Domar model have been built into a variety of more elaborate models over the years. The so-called "Two-Gap Model", like the Harrod-Domar model, pointed to the possibility of a "gap" between the amount of capital accumulation required to keep a growing population fully employed and the amount of financing forthcoming through private savings, and raised the possibility of a second gap between the amount of foreign exchange required to purchase the required capital equipment from abroad and the amount of foreign exchange forthcoming from exports (Chenery and Strout, 1966). Multi-sector versions of these models, which distinguished multiple sectors by their requirements for capital, foreign exchange, and the outputs of other sectors, allowed policymakers to calculate the gaps that must be filled by policy on a sector-by-sector basis. For a provocative discussion of the genesis and on-going legacy of these models in international development institutions, see Easterly (2002), especially Chapter 2.

The Lewis Model of "Labor-Surplus Dualistic Development"

In his famous article, "Economic Development with Unlimited Supplies of Labour," W.A. Lewis (1954) embedded capital fundamentalist notions in a model of an economy characterized by profound dualism, and described economic growth as a process in which physical capital accumulation and structural change are intimately inter-related. The dualistic economy is composed of two sectors, the **subsistence sector** and the **capitalist sector**, which differ from each other in both production technology and behavioral relationships. The **structural change** that accompanies growth in this model is a shifting of labor out of the traditional sector, where some labor is largely unproductive, into the modern sector, where that same labor can be brought into productive employment as physical capital accumulation creates new jobs.

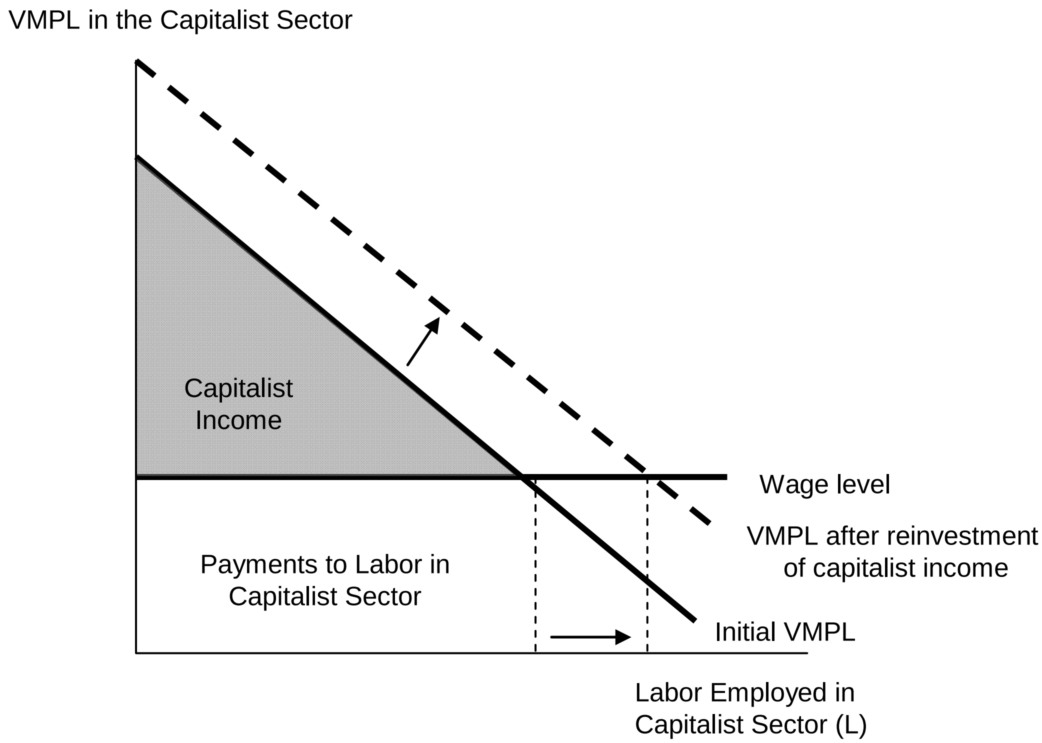
In the large subsistence sector, which Lewis associated with traditional agriculture as well as some traditional handicrafts and services, production is a function only of labor and land. Physical capital plays no role, thus the sector has no potential to grow through capital accumulation. While in principle subsistence sector producers can produce output with varying combinations of land and labor, in practice the ratio of labor to land is so high that the marginal product of labor is virtually zero. Labor is "in surplus" or "redundant", in the sense that some workers can be withdrawn from the sector without reducing the sector's total output. Workers in agriculture are paid a wage that is fixed by tradition near the subsistence level, rather than by market forces, and this wage is greater than their marginal product of zero. Neither producers nor workers in this sector save any of what they earn. Subsistence sector workers are willing to move into the modern sector if offered a wage equal to the subsistence wage plus some premium to compensate them for the difficulties and psychological costs of making such a move. Because their marginal product in agriculture is zero, if they respond to such a premium and leave the sector, subsistence sector production remains the same. Thus from a social perspective they are "free labor" (similar to the unemployed workers in the Harrod-Domar model) that can be moved into productive use in the modern sector without sacrificing any output elsewhere in the economy.

In the small capitalist sector, production is a function of labor and capital. Here is where capital accumulation has the potential to drive growth. The production function in this sector is of the sort described in Chapter 3, in which labor and capital may be combined in varying proportions to produce output, with each input subject to diminishing marginal returns. Capitalists (i.e. the owners of capital, who run firms in this sector) seek to maximize the profits they derive from their current holdings of capital and are motivated to save and re-invest those profits. In fact, they are assumed to save and reinvest all of their profits. To attract workers from the subsistence sector, they pay a wage equal to the subsistence wage plus the required premium. Profit-maximization requires that they hire the quantity of labor at which the value of the marginal product of labor (VMPL) equals this wage. (For a review of why this is the profit-maximizing quantity of labor, see Chapter 5.)

A convenient way to illustrate the implications of this set of assumptions is with the graph in Figure 3A.1. The horizontal axis measures the quantity of labor employed in the capitalist sector. Any labor not employed in the capitalist sector remains in the subsistence sector. Thus a move to the right in this graph signifies a "structural change", in which the share of labor employed in the subsistence sector declines and the share employed in the capitalist sector rises, as workers are drawn into capitalist sector employment. A move to the right also signifies an increase in the overall productivity of labor in the economy, as labor is drawn out of a sector in which its marginal product is zero, into a sector in which it has a positive marginal product.

The vertical axis in Figure 3A.1 measures the value of the marginal product of labor (VMPL) in the capitalist sector. As described in Chapter 3, given any quantity of capital employed in the capitalist sector, we can trace out a downward-sloping VMPL curve. Each additional unit of labor put to work with a given capital stock increases output (hence the VMPL is positive), but the more labor employed the lower the VMPL (hence the downward slope). The solid downward-sloping curve is the VMPL curve associated with an initial stock of capital. The horizontal line is drawn at the height of the wage that must be paid to draw workers out of the subsistence sector. What is important is that this wage is rigidly determined by tradition and does not increase as workers are drawn out of agriculture. The intersection of the solid VMPL curve with the horizontal wage line identifies the initial level of employment in the capitalist sector.

Figure 3A.1



What causes this figure to yield implications for economic growth is the assumption that capitalists save and reinvest their profits, while workers consume all their earnings. The total value of production in this economy, and thus total income, is given by the total value of production in the capitalist sector plus the unchanging total value of production in the agricultural sector. The initial total value of production in the capitalist sector is the area under the solid VMPL curve up to the point at which it intersects the wage level line.[3] Of this total area, the lower rectangle, which has width equal to the number of

[3] To see why this is so, remember that the height of the VMPL curve traces out the amount by which the value of output increases when each additional unit of labor is added to production (holding other inputs fixed at their current levels). Each unit of labor (up to the total amount employed, which is associated with

workers in the capitalist sector and height equal to the wage, has area equal to the total payments to labor in the capitalist sector. The remaining capitalist sector output, represented by the area of the shaded triangle, indicates the earnings of the capitalists (i.e. owners of capital) and thus the portion of current income that is reinvested. As this investment creates new capital stock, the VMPL curve shifts up and to the right. As the VMPL shifts, the profit-maximizing quantity of labor to employ in the capitalist sector rises, and more workers are drawn out of the subsistence sector into the capitalist sector. The total value of output in the economy also grows. When workers move from subsistence to capitalist sectors, their wages may increase by the size of the premium paid in the capitalist sector, but the level of the wage in either sector remains fixed, and total payments to labor throughout the economy rise little. Most of the increased value of production arising out of investment in capital translates into increased capitalist income, leading to an increase in the capitalist share of income, and an increase in the rate of saving and re-investment. The existence of surplus labor, which causes the wage to remain low despite the new pull of a rising capitalist sector, allows profits and the rate of saving and investment to increase. Surplus labor thus fuels accelerated growth and structural change.

Like the Harrod-Domar model, the Lewis model places great importance on saving and investment in physical capital as the route to economic growth. It differs, however, in its implications for the best way to promote that saving and investment. It describes a world in which capitalists are willing and able to undertake investment, provided they have profits to re-invest. From a social perspective, however, capitalists perceive too little incentive to produce and invest, because they must pay a wage that is higher than their workers' marginal product in agriculture. This means that at the private profit-maximizing level of capitalist sector employment (where the marginal product of labor equals the wage), the marginal product of labor is higher in the capitalist sector than in agriculture. If capitalists could be encouraged to draw even more workers out of the subsistence sector, the total value of production in the economy would rise. Thus the model suggests that capitalists should be provided with added incentives to expand their production and profits. For analysts who took the capitalist sector to be synonymous with manufacturing, and who observed that most manufactured goods sold in developing countries at the time were goods imported from the developed countries, the natural policy prescription of the model was to encourage the expansion of import-competing manufacturing enterprises by using tariff policies to protect them from international competition.

Notice that the model treats the large subsistence agricultural sector as playing a purely passive role in growth and development, that of releasing surplus workers to the modern manufacturing sector. Agriculture was assumed to have no growth potential of its own. Indeed, the model suggests that growth proceeds most successfully when agricultural productivity remains low, so that workers may be attracted out of agriculture at low wages that show no tendency to rise. If agricultural productivity and wages were to rise, the rate of profits in the manufacturing sector would fall, reducing the rate of saving and

---

the intersection of the VMPL and wage line) adds to the value of output approximately a value equal to the area of a rectangle with base width equal to one unit of labor and height equal to the height of the VMPL.

re-investment. Thus the world view embodied in the Lewis Model encouraged a neglect of agriculture in the pursuit of growth and development.

The 1960s and 1970s: Neoclassical Perspectives, Technology and Human Capital

Research in the 1960s and 1970s chipped away at the theoretical and empirical underpinnings of capital fundamentalism, structuralism and dualism. Capital fundamentalism gave way to interest in the combined contributions of physical capital accumulation, technical change and human capital accumulation. The introduction of the "Neoclassical Growth Model" (described below) raised theoretical questions about the potential for physical capital accumulation alone to generate the sustained growth observed in the United States and other countries, and pointed toward technical change as an additional source of growth, while simultaneous developments in the study of labor markets and wage formation pointed to the importance of human capital accumulation.

Structuralist views of developing country labor markets diminished as collection and analysis of data from farms and households in developing countries demonstrated that even poor rural families appear to respond to economic incentives in ways predicted by models of rational economic behavior. The structuralist assumption that poorly functioning markets could be over-ridden by extensive government intervention was also called into question by highly disappointing experiences with government planning. Indeed, the intellectual pendulum swung to the opposite extreme in this period, as a growing number of development analysts argued for withdrawal of government intervention and greater reliance on markets for allocating resources and motivating growth.

Dualism of the sort exhibited in the Lewis model lost favor in this period, too, for a variety of reasons. Research suggested that the marginal product of labor is not zero in agriculture, that agriculture can be a dynamic source of saving and physical capital accumulation, and that investment in physical capital can induce productivity increases in agriculture as well as manufacturing. The experience of the Green Revolution (see Chapter 17), moreover, demonstrated the great potential for new technologies to increase agricultural productivity. On top of this, it was becoming apparent that many of the poor in developing countries earned their livelihoods in agriculture, and would not begin to share in the benefits of growth and development for a very long time, unless new efforts were made to increase agricultural productivity and income. For a more detailed discussion of changing economic thought regarding growth and development in this period, see Little (1982).

The Simple Neoclassical Growth Model

In his landmark paper "A Contribution to the Theory of Economic Growth" (1956), Robert Solow introduced what we now know as the Neoclassical Growth Model, which served as a unifying framework for thought about economic growth for several decades. He motivated the new model by pointing out two reasons to be dissatisfied with the thenpopular Harrod-Domar Growth Model. First, that model rests on stark technological assumptions that contemporary economists found difficult to swallow. Second, the knife-edge property of that model implies a kind of perpetual macroeconomic instability that seemed counter to observed experience, at least in the United States.

Solow pointed out that the rigid assumptions and the unsatisfying implications of the Harrod-Domar model are closely related to each other. By assuming a production technology in which labor and capital can only be productively employed in exactly fixed proportions, the model assumed away any potential for producers to respond to conditions of capital scarcity by using more labor-intensive methods of production. If some workers are unemployed, we might expect them to offer to work at lower wages, rendering labor cheaper relative to capital. Producers who can choose from multiple production methods – some involving higher ratios of labor to capital than others – might be encouraged by such factor price changes to switch to production methods that require more labor and less capital per unit of output. By so doing, they would increase the number of workers who can find employment with the fixed capital stock. Indeed, if markets work perfectly, this process should continue until unemployment disappears.

Solow thus decided to examine what would happen to the predictions of the Harrod-Domar model as a result of just two changes in its assumption (one technological and one behavioral): (1) the replacement of the rigid fixed proportions production technology by a **variable proportions** or **neoclassical** production function, in which output can be produced with varying combinations of labor and capital and each input is subject to diminishing marginal returns (assumptions that had long been employed in much economic analysis outside the areas of growth and development); and (2) the replacement of the assumption of extensive unemployed labor by the assumption that rational producers respond to signals in perfectly functioning markets in ways that drive the economy to full employment of both labor and capital stocks at all times. With the theoretical tools available to Solow, this was most easily accomplished by assuming that all producers in the economy have access to the same production technology, and that the technology exhibits constant returns to scale. Solow continued to assume, along with Harrod and Domar, that saving is a fixed proportion of income and that the labor force grows at an exogenously determined rate n.

Solow showed that introducing such changes to the model's assumptions eliminated its undesirable predictions of instability. Unemployment is eliminated from the model, and capital can accumulate relative to labor with no potential for a sudden deceleration of growth when unemployed labor is used up. He also demonstrated, however, that making these changes leads to important changes in some of the model's predictions that had not previously been questioned. In the Harrod-Domar model, a constant saving rate can serve as the source of steady, sustained growth (as long as the economy doesn't run out of unemployed labor), and increases in the saving rate lead to sustained increases in the growth rate. We will see below that in Solow's neoclassical model, a constant rate of saving by itself cannot sustain growth in the long term, and permanent increases in the saving rate produce only temporary boosts in the rate of growth. The source of this dramatic change in theoretical prediction is the assumption of neoclassical production technology, characterized by the diminishing marginal returns to capital. In the presence of such diminishing returns, a steady rate of saving implies steady additions to the capital stock as a proportion of output, but this translates into a declining rate of increase in capital per worker and a declining rate of economic growth, because as the quantity of capital per worker increases, diminishing returns dictate that additional units of capital lead to smaller and smaller increases in output per worker. In fact, growth in output per worker based entirely on saving and capital accumulation is bound eventually to grind to a halt in the simple neoclassical model. Increases in savings rates might boost growth in the short-run, and allow the economy to achieve a higher *level* of per capita GDP before economic growth halts, but such growth would ultimately cease. We demonstrate this formally below.

Solow did not stop there, however. He pointed out that the revised model, too, failed to match historical experience. The U.S., for example, *had* experienced a nearly constant saving rate *and* steady growth over a very long period of time. He concluded that if growth based on steady capital accumulation alone is bound to diminish over time, then U.S. growth must have been based on something more than just steady capital accumulation. The "something more" to which he drew attention was technical change.

In what follows we first exposit what happens when we introduce Solow's neoclassical modifications of the Harrod-Domar model, without introducing the possibility of technical change. We will examine this simple model's implications, including the puzzling prediction that growth must eventually grind to a halt. We will then demonstrate how introducing the notion of technical change can solve the puzzle by rendering the model capable of explaining sustained positive growth together with a constant rate of saving over a long period of time.

We begin with the assumptions that the economy's aggregate production function is "neoclassical" and that the economy's stocks of labor and capital are fully employed at all times. These assumptions together underlie the assumption that

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

where Y is GDP, and K and L are the economy's stocks of capital and labor inputs. The production function $F(.,.)$ is assumed to be characterized by diminishing marginal returns to either input. The production function is furthermore assumed to exhibit **constant returns to scale**, which means that doubling both inputs leads to a doubling of output. (The standard justification for the constant returns to scale assumption at the time was that when doubling all inputs, one could simply create a duplicate of the original production establishment, and this ought to double output.) The assumption of full employment rests on the assumption that the economy's many owners of capital and labor resources can sell capital and labor services to producers in perfect markets. Under such assumptions no resources remain unemployed, and labor and capital are allocated across production establishments in ways that maximize their productivity. All this happens regardless of how ownership of the factors is distributed, thus total production in the economy depends only on the total quantities of K and L, and not on how their ownership is distributed. The rest of the model is the same as in Harrod-Domar: L is assumed to grow at a fixed rate n, K grows through capital accumulation financed by saving (in excess of the rate of depreciation d), and saving is a constant proportion s of income.

Notice that the only way for output per worker ($y=Y/L$) to grow in this model is through increases in the capital labor ratio ($k=K/L$), which indicates the quantity of capital available per worker in the economy. Given our interest in the growth of per capita GDP, it will be useful to recast our technological assumptions in per-worker terms. Dividing both sides of (6) by L, and invoking the assumption of constant returns to scale, we find that

(7)) $y = \frac{Y}{L} = (\frac{1}{L})F(K,L) = F(\frac{K}{L}, 1) = f(k),$

where the last equality simply assigns a new name to the function describing the relationship between capital per worker and output per worker. This new function $f(.)$ relates the average product of labor (y) to the level of capital per worker (k). As illustrated in Figure 3.1b, y must increase with k, but must increase at a decreasing rate, as a result of diminishing marginal returns to capital.

Given the close link between the capital-labor ratio (k) and per capita production (y), we can understand the model's implications for the rate of economic growth by understanding what governs the rate of change in k, which we denote by $\dot{k}/k$. It may be shown that the rate of growth of the capital-labor ratio is just the rate of growth of capital minus the rate of growth of labor, as in

(8) $\frac{\dot{k}}{k} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L}$

As in the Harrod-Domar Model, $\dot{K}$ is equal to $sY\text{-}dK$, where Y now equals $F(K,L)$, and the rate of growth of the labor force is n. After making these substitutions, (8) becomes

(9) $\frac{\dot{k}}{k} = \frac{sF(K,L)}{K} - d - n$

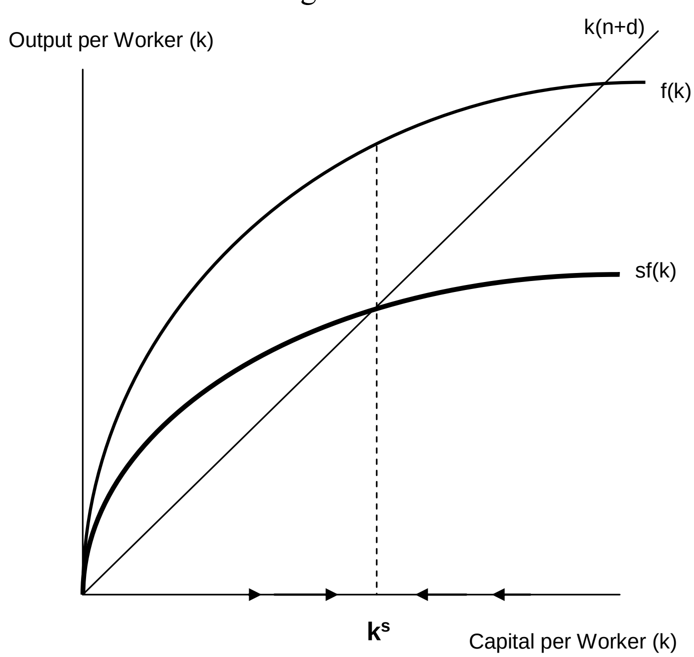
Noting that F(K,L) may be written as $LF(k,1)=Lf(k)$ (making use of the constant returns to scale assumption), and multiplying both sides of the equation by k, we find the equation embodying the core implications for growth of the simple neoclassical model:

(10) $\dot{k} = sf(k) - k\,(n+d)$

The left hand side is the change in k experienced in any period of time. If this is positive, k is rising; if it is negative, k is falling. The right hand side relates this change to current economic circumstances. The first term on the right hand side describes the saving and investment forthcoming per person at any current level of *k*. The second term indicates the investment per person required simply to maintain the entire labor force at the current level of *k,* given the need to replace depreciating capital and the need to equip additional workers entering the market through labor force growth. The equation tells us that if per-worker investment forthcoming is greater than per-worker investment required to maintain a constant level of capital per worker, then the capital-labor ratio (and output per worker) will rise. If the per-worker investment forthcoming is less than $k(n+d)$, however, then the capital-labor ratio will fall.

To see the long-run implications of this equation, it is useful to employ the diagram found in Figure 3A.2. The horizontal axis measures the capital-labor ratio, *k*. The vertical axis is measured in units of output per worker, which are also the units in which we measure capital, saving and investment. As indicated by the $f(k)$ graph, output per worker rises at a decreasing rate as *k* increases, because capital's marginal product is positive but diminishing. The figure describes what happens to the two components of the right hand side of (10) as *k* increases. Since $sf(k)$ is just a constant multiple (between zero and one) of $f(k)$, diminishing marginal returns to capital implies that the $sf(k)$ graph, too has positive but decreasing slope. Since *n* and *d* are constants, $k(n+d)$ is a straight line with slope $n+d$.

Figure 3A.2



We can use this graph to demonstrate a tendency for the economy to progress toward a long-run or "steady state" equilibrium in which the capital-labor ratio, and output per worker, is unchanging.[4] That is, the economy will progress toward an equilibrium in

[4] This will be true as long as the $sf(k)$ curve is not so low that it lies entirely beneath the $k(n+d)$ locus, and that it is sufficiently curved that its slope does eventually fall below the slope of the $k(n+d)$ locus.

which there is no (per capita) economic growth! To see this, notice first that if the economy ever reached the level of capital per worker k^s associated with the intersection of the $sf(k)$ and $k(n+d)$ graphs, the economy would remain at that level of k forever. At that level, the amount of investment forthcoming in the economy, $sf(k^s)$, is just enough ($k^s(n+d)$) to replace depreciated capital and to provide new workers with the same level of capital per person as the old workers already have. According to (10), $\dot{k}$ would be zero, and there would be no impetus for k to change. Notice second that if we observe an economy at a level of k other than k^s, economic forces will be driving k toward k^s. For example, consider an economy with a level of k below k^s. For such an economy $sf(k)$ would be greater than $k(n+d)$, and $\dot{k}$ would be positive. This means that k is growing, and the economy is moving to the right in the diagram. It will continue to move to the right until it reaches the level k^s, at which point k ceases to grow. Similarly, if an economy started out with a level of k to the right of k^s the capital labor ratio would be falling, and it would continue to fall until the economy finds itself in steady state or long-run equilibrium at k^s.

We can use this graph to examine what would happen if we took an economy in initial long-run equilibrium and increased its saving rate s. The $sf(k)$ function would pivot up, remaining connected to the origin, as the unchanging $f(k)$ function is multiplied by a larger fraction. With an initial capital-labor ratio of k^s and a new higher saving rate s', $s'f(k^s)$ would exceed $k^s(n+d)$, $\dot{k}$ (and the rate of growth in output per worker) would become positive, but this new growth in k and y would eventually die out, as the economy tends toward a new long-run equilibrium with zero growth in per capita income. The increased frugality would not be without benefit, because the economy would eventually find itself in a new steady-state equilibrium characterized by higher k and higher income per capita (at the new intersection between the $k(n+d)$ line and the higher $s'f(k)$ curve), but the sustained increase in s would produce a period of growth in output per worker that lasts only as long as it takes to achieve the new long-run equilibrium level of k.

The Neoclassical Growth Model with Technical Change

The implication that growth driven only by a constant saving rate should grind to a halt flies in the face of real world experiences, such as that of the United States, which has managed to save and experience per capita income growth at fairly steady rates over long periods of time. Is this the sort of unfortunate theoretical implication that should lead us to discard the model? Solow's answer was that we should modify rather than discard the model. In particular, he suggested replacing the aggregate production function (1) by the function

(11) $Y = F(K,AL)$

where A is a parameter describing the level of technology. More specifically, A is treated as a parameter that enters the production function as a multiple of the labor input, and whose increase signifies **labor augmenting technical change**. We can think of AL as the quantity of "effective labor" in the economy, and technical change (denoted by an increase in A) as change that allows each physical unit of labor to supply more effective labor. As we will see, in the presence of labor-augmenting technical change, the ratio of capital to effective labor, and thus the marginal product of capital, can remain constant, even while the ratio of capital to physical labor is rising. Such technical change serves to counteract the diminishing returns to capital that caused growth to grind to a halt in the simplest neoclassical model.

To understand growth and long-run equilibrium in this economy, it is useful to define and understand the evolution over time of the ratio of capital to effective labor, *K/AL*, which we will call k^*. It can be shown that the growth rate of a ratio like k^* is equal to

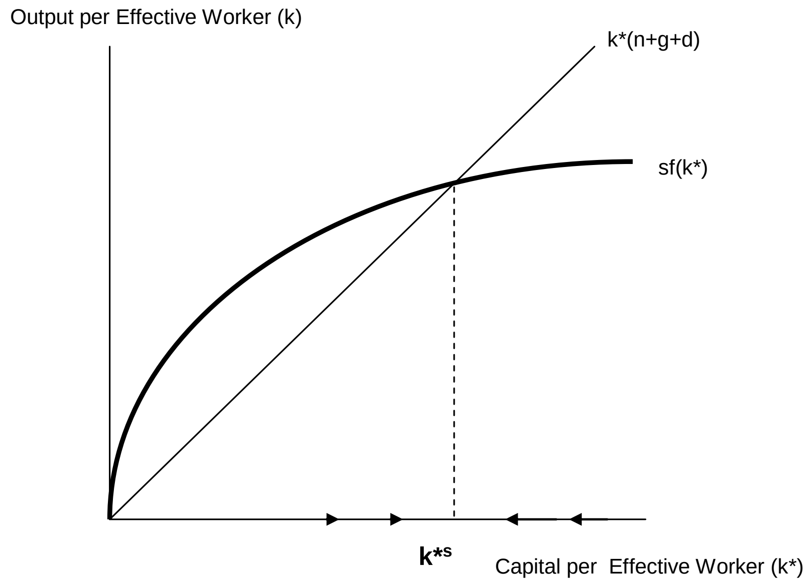
$$(12)\quad \frac{\dot{k}^*}{k^*} = \frac{\dot{K}}{K} - \frac{\dot{L}}{L} - \frac{\dot{A}}{A}$$

Making the same assumptions as before regarding the growth of K and L, assuming that A increases at the rate g, and multiplying both sides by k^*, this becomes

$$(13)\quad \overset{*}{k}{}^{*} = sf(k^*) - k^*(n+d+g)$$

To understand the growth performance of the economy described by this equation, we turn to Figure 3A.3. Here the horizontal axis measures the level of k^*, while the vertical axis is measured in units of output and investment as in the previous diagram. Using the same logic as we applied to the previous diagram, we find that this economy will converge toward a steady state equilibrium in which k^* is constant at the level k^{*s}, associated with the intersection of the two curves. Notice, however, that a steady state in which $k^* = K/AL$ is constant is a steady state in which K/L and GDP per capita must be rising! In fact, with A rising at the rate g, K/L must be rising at the rate g as well to hold k^* constant, and output per capita must be growing at the rate g, too. This model thus predicts that the economy will achieve a steady-state in which output per worker grows at a constant rate, equal to the rate of technical change, g. The introduction of technical change thus allows the model to explain sustained growth in the long run. While the rate of saving, s, does not influence the rate of growth in steady state, it continues to play a role in determining the steady state level of $\overset{*}{k}{}^{*}$, and thus the steady-state path of GDP per capita (as can be seen by noting the impact on the intersection between the curves in Figure 3A.3 of increasing *s)*.

Figure 3A.3



Solow's main purpose in constructing the neoclassical model with technical change was to construct a model that can explain the coexistence in a single country of a stable, positive growth rate and a stable rate of saving over a long period of time. But many subsequent researchers have pointed out the model's implications for differences across countries in rates of economic growth, and have subjected these implications to empirical testing. According to the model, if all countries were already in steady state and had access to the same technology, then all countries should be growing at the same rate, g. As this is clearly not the case, we must consider what the model implies for differences in growth rates among countries that are heading toward steady state but have not yet achieved steady state. Notice, first, that among countries heading for the same steady state (because they are characterized by the same values for s and $n+g+d$), countries that start out poorer should grow faster. (Diagrams like Figure 3A.3 would be identical for such countries, except that poorer countries would be starting out further to the left of the steady state level of k^*.) If all countries were heading toward the same steady state, we should see evidence of "**unconditional convergence**", in which poorer countries (i.e. countries starting at lower levels of k^* and GDP per capita) simply grow faster than richer countries across the board. Though they may start out poorer, the model predicts that they will ultimately catch up, erasing those initial differences and enjoying the same level of growth and per capita income in steady state. Unfortunately, the data graphed in Figure 3.2 contradict this rather optimistic prediction. In the world as a whole there has been no systematic tendency for poorer countries to grow faster than richer countries. Indeed, some of the poorest countries have also experienced the worst growth performance.

A second observation regarding the model's implications offers a possible reconciliation between the model and the data of Figure 3.2. If countries are heading toward different steady states (because they differ in their values of s and $n+g+d$), then among countries starting at the same level of k^* and income per capita, those with lower s and higher

$n+g+d$ should be heading toward lower steady state levels of k^*. Since they are closer to their ultimate steady states, they should be growing slower. Once we allow for these differences in the steady states toward which countries are heading, the model predicts "**conditional convergence**." This means that even though poor countries don't always grow faster than richer countries, it may still be the case that, *after controlling for differences across countries in* s *and* $n+g+d$, poorer countries are growing faster than richer countries. That is, if you run a regression of growth rates on measures of s, $n+g+d$, and initial income per capita, the coefficient on income per capita should be negative (and the coefficient on s should be positive and the coefficient on $n+g+d$ negative). Regressions employing data from the last four decades often reveal such patterns. While such patterns have sometimes been taken as evidence in favor of the neoclassical model and against alternative models proposed later (which we discuss below), it is now broadly recognized that empirical conditional convergence results can be reconciled with a wide range of theoretical models, and thus do not, in fact, help distinguish between models (Durlauf, et al., 2005; Romer, 1994).

More elaborate versions of the neoclassical growth model, which pay more explicit attention to the choices of the economy's many producers and owners of labor and capital, reveal that as long as the technologies faced by individual producers are characterized by constant returns to scale, and as long as they operate in perfectly competitive output and input markets, then the economy will deliver optimal growth and continuous full employment without the need for government intervention. People who care about the future, and who operate in perfect financial markets, will correctly perceive the incentive to save and invest. There is thus no need to use government policy to encourage more investment or production. In fact, by distorting incentives for factor use and investment, government intervention is more likely to hurt growth than to help it. The model is thus consistent with a world view of rising importance in the 1960s and 1970s, in which the key to speeding growth was to pull government out of production, reduce government intervention in markets and reduce government efforts to protect specific sectors of the economy from international competition.

Neoclassical Growth with Human Capital

Though the notion of **human capital** is often traced back to Adam Smith, it was not until the 1960s that research into the role of human capital in determining productivity and incomes really took off, encouraged by the rising availability of large individual-level datasets on schooling and earnings (at least in the United States), as well as the introduction of computing power adequate for econometric analysis of those new datasets (Mincer, 1970). The emerging literature quickly demonstrated an important role for schooling and experience in raising individual incomes, presumably by increasing workers' productivity. It was natural, then, for researchers whose interest in economic growth was revived by Solow's theoretical contribution to consider augmenting the model to incorporate human capital as an additional determinant of aggregate production. One example of how human capital may be introduced formally into the Solow framework is found in Mankiw, Romer and Weil (1992), who replace Solow's aggregate production function by

(14) $Y = F(K,H,AL)$

where H is total human capital, A continues to represent labor-augmenting technical change, and $F(.,.,.)$ exhibits diminishing marginal returns in each of its three arguments.[5] The authors assume that the creation of both physical and human capital are financed by saving out of income, and assume that a fraction s of income is devoted to physical capital investment, while a fraction b is devoted to human capital investment. As the reader might guess, in this augmented Solow model, the economy converges to a steady state in which the values of physical capital per effective worker ($k^*=K/AL$) and human capital per effective worker ($h^*=H/AL$) are constant, the steady state levels of these variables depend on the investment rates (s and b) as well as $n+g+d$, and steady state income per worker grows at the rate of technical progress.

While interest in human capital introduced only minor modifications to economic growth theory in the 1960s and 1970s, it was associated with dramatic changes in the practice of development work by the large international organizations, for whom investment in education became the centerpiece of many efforts to promote growth and development. The popularity of investment in education can be explained in part by strong suspicions that such investments would be simultaneously beneficial for promoting growth, reducing inequality and satisfying "basic needs" of the poor.

The 1980s to the present: Market imperfections, increasing returns, and poverty traps

The late 1980s brought a resurgence of interest in theoretical study of economic growth, as new data brought to light new patterns of growth experiences across countries, and as developments in theoretical economics provided new analytical tools (Romer, 1994). While study of growth during the 1960s and 1970s was shaped by a single theoretical framework (the Neoclassical Model), theorists in more recent decades have developed a plethora of new models as they seek to wrestle with three sets of questions that were raised but not answered by the Neoclassical Model:

1. What drives technical change?
2. If investment in physical capital, human capital and technical change is so important for growth and future well-being, and if people care about the future, then why don't we see more investment in poor countries? That is, why are some countries stagnating with low rates of saving, investment, productivity and growth? and
3. If policies have a role to play in improving incentives for investment and growth, then why don't we see more good policies?

[5] In fact, they assumed a particular functional form, $Y=K^{\alpha}H^{\beta}(AL)^{1-\alpha-\beta}$, which allowed them to derive equations defining steady state per capita income as an explicit function of (n+g+d), s, b, α and β, as well as a term capturing the fact that in steady state income per capita is continuously rising at the rate of technical change, g. They show that this equation does a surprisingly good job of explaining differences across countries in per capita income.

Altogether the new models of this period have brought thinking about economic growth even further from capital fundamentalism, reinforcing interest in human capital, expanding interest in the creation and adoption of improved technologies, and introducing new interest in reducing inefficiency and waste. The new models also offer more nuanced alternatives to the structuralism and dualism of the 1940s and 1950s than was offered by the extreme neoclassical vision of perfect markets and perfectly integrated sectors. The models highlight a wide variety of reasons why reasonably well-functioning markets might fail to deliver ideal investment and growth, and some models raise the possibility of a new sort of dualism, in which entire poor countries, or poor populations within countries, might find themselves "trapped" in traditional, low-productivity production methods, even while their higher income neighbors pursue modern, higher-productivity methods and enjoy greater growth and prosperity. The models highlight the potential importance of a variety of departures from the neoclassical assumptions of constant returns to scale production and perfect competition.

Endogenizing technical change

The primary objective of a first subset of "new growth theory" was to "endogenize" technical change. Theorists contributing to this literature agreed with Solow that technical change is important for long-run growth, but were unsatisfied with Solow's theoretical treatment of technical change in the neoclassical model, where technical change is simply assumed to take place at an exogenously given rate. That is, technical change was treated by the Solow model as "manna from heaven," dropped into the economy from outside the model. **Endogenizing technical change** requires identifying specific choices people make that might give rise to technical change, and making assumptions about the ways in which current economic circumstances affect peoples' incentives toward these choices, thereby allowing the rate of technical progress to be determined within the model.

Endogenizing technical change proved impossible to do within neoclassical theoretical models that maintained the assumptions of constant returns to scale and perfect competition. These two assumptions together imply that payments to capital and labor must fully exhaust output, leaving nothing left over to serve as a reward for purposeful improvements in technology.[6] New growth theorists "solved" this problem in two quite different ways, both of which raised the possibility that markets might not lead the

[6] We may demonstrate this formally as follows. Let $F(K,L;A)$ be a production function, w be the wage rate paid and r be the rental rate paid for a unit of capital. The production function exhibits constant returns to scale in K and L if and only if, for any value of λ, $F(\lambda K, \lambda L; A) = \lambda F(K,L; A)$. Taking derivatives of both sides of this equation with respect to λ , and letting F_K and F_L denote the partial derivatives of the production function with respect to its first and second arguments, we get $F_K(\lambda K, \lambda L; A) * K + F_L(\lambda K, \lambda L; A) * L = F(K,L;A)$. Since this must be true for any value of λ, it must be true for λ=1, indicating that $F_K(K, L; A) * K + F_L(K, L; A) * L = F(K,L;A) = Y$. If producers behave in a perfectly competitive manner, then they should hire factors until marginal product of capital (F_K) is equal to the rental rate paid per unit of capital, r, and the marginal product of labor (F_L) is equal to the wage paid per unit of capital, w. Thus constant returns to scale and perfect competition together imply that $rK + wL = Y$. This tells us that with constant returns to scale (in K and L) and perfect competition, total payments to K and L indeed fully exhaust total output Y, leaving nothing left over to compensate anyone for purposeful efforts to increase A.

economy to ideal growth performance. The first solution was to treat productivity advance as an unrewarded by-product of decisions made without regard to their productivity impacts. That is, they treat technical change as an externality rather than an outcome that people seek to achieve "on purpose." This eliminates the need to find any portion of total product left over after payments to capital and labor, which can be held out as a reward for purposeful efforts to improve technology. The second solution was to treat technical advance as the output of a purposeful, reward-seeking activity, and to build rewards for these activities into the model by abandoning the assumption of perfect competition. The first solution seems more suited to explaining productivity advance that arises through informal learning by doing in production establishments, while the second solution seems more suited to understanding productivity advance that arises through formal, purposeful research and development activities. Both solutions led to visions of economies that are characterized by "increasing returns to scale", as described below. The class of models these efforts gave rise to are known as **endogenous growth models**, because they sought to endogenize the rate of technical change, which was the sole determinant of the long-run rate of economic growth in the Solow framework.[7]

Productivity advance as a by-product of investment in physical or human capital. The first approach to endogenizing productivity advance was to think of the productivity level (A) as an outcome that grew as an unintended by-product of choices to increase the stock of physical or human capital. This might happen if the investments people make to improve their own human capital have positive externalities on the productivity of other workers and producers throughout the economy. It might be the case, for example, that producers and workers have more scope to learn from creative experimentation when they are surrounded on a day-to-day basis by more educated people. A seminal paper introducing this possibility is Lucas (1988).

To introduce this notion formally into a model of economic growth, we might re-write the aggregate production function this way:

(15) $Y = A(H)\, F(K,H,L)$

The term $A(H)$ indicates that the level of productivity depends on the aggregate stock of human capital in the economy. Individual producers are assumed to be unaware of this link between their individual human capital decisions and the economy's level of productivity, and simply take the current level of A as given. When holding A constant, the production function continues to exhibit constant returns to scale in K,H and L. From the perspective of the economy as a whole, however, investments in human capital not only add to the economy's human capital stock, but also increase the level of productivity.

---

[7] Originally the label "endogenous growth model" also connoted that the model yielded a prediction that steady state growth rates would be a function of policies shaping rates of investment in physical and human capital. This contrasted with the neoclassical model, in which such policies affect only the level of income per capita in steady state, and rates of growth during the transition to steady state, while the rate of growth in steady state is fixed exogenously and unaltered by policy. This connotation has diminished in importance, for reasons discussed in Jones (1998).

While this might look like a very small change to the formal structure of a growth model, it can in fact have profound implications. The key reason is this: If the production function exhibits constant returns to scale in K, H and L while holding A constant, then when we acknowledge that increases in H also increase A, we discover that the economy as a whole is characterized by **increasing returns to scale**. That is, if we doubled K, H and L we would more than double Y. More to the point, while H is (according to standard and sensible assumptions) subject to diminishing marginal returns when we hold A constant, it might be subject to constant or even increasing marginal returns when we acknowledge the externality: the boost to output associated with increases in H might remain constant or even rise (depending on the strength of the effect of H on A) as the level of H increases, contrary to our expectations in the presence of diminishing marginal returns to capital. In a diagram like Figure 3A.3 above, the $sf(k^*)$ curve might be straight or even convex, rather than concave. This raises the possibility that the return to human capital investment might be larger in richer countries, because the external productivity effects of having more human capital per worker outweigh the force of diminishing marginal returns to human capital. Such a modification thus eliminates the Solow model's logic regarding the tendency for growth based on accumulation of physical or human capital to grind to a halt, the tendency for poor countries to grow faster than rich countries, and the tendency for poor countries to hold the expectation of eventually converging to the same level of income enjoyed in initially richer countries. This seemingly small modification to the model thus suggests the possibility of a world in which poor countries fail to catch up. Because the modification is built on the assumption of important externalities, it also raises the possibility that private human capital investors, who do not take the external benefits of their choices into account, will invest less than is socially desirable, and that intervention by governments and other development actors might be useful for improving growth performance.

Productivity advance as the outcome of purposeful research, development and technology transfer activities. The second approach to endogenizing A was to treat increases in A as the outcome of intentional efforts to develop or adopt new technologies, and to "make theoretical room" for some of output to be used as reward for those activities by abandoning the assumption of perfect competition. These models distinguish between the currently developed countries that are **technology leaders**, in which the intentional activities of interest are the research and development activities that produce cutting edge technologies, and the currently less developed countries that are **technology followers**, in which the productivity-enhancing activities are efforts to acquire and adopt technologies already developed by the technology leaders. These papers highlight a variety of reasons why private investment decisions might not deliver the socially most desirable levels of investment in the creation and acquisition of new technology. We briefly mention these issues here, and discuss them in more detail in Chapter 16. The models also demonstrate how explicit attention to the peculiar features of technological ideas can lead to models characterized by increasing returns to scale, and thus by the possibility that poor countries will fail to catch up.

Consider first the models, such as that in Romer's (1990) seminal paper, describing technical change and growth in the technology leaders. They point out that the purposeful research and development activities that often underlie technical change (and cause A to grow in formal models) require inputs of labor and other factors, and thus are costly to undertake. They will thus be undertaken only if investors expect to receive at least enough future reward for new technologies they produce that they can cover their costs. If the new technological "ideas" produced through research and development could be freely copied, then potential investors would see no way to reap returns on their investments, because they would not be able to charge for the use of their ideas. If they are to perceive an incentive to invest, a system of patent protection (or effective trade secrets) must give them temporary monopoly power over the use of their new ideas.

In Romer's (1990) model, the costliness of research and development activities is captured by assuming that some of the economy's labor resources must be allocated to research and development activities if technology is to advance. This creates a trade-off between the two uses of labor in the economy: final goods production and efforts to improve technology. Total final goods production is shaped by the production function:

(16) $Y = F(K,AL_Y)$

where K represents physical and human capital, and L_Y represents the quantity of labor devoted to final goods production. The level of technology, A, is now assumed to be the cumulative result of past research and development activity. More specifically, A is interpreted as the total number of "ideas" generated by research and development activities to date, and the ideas are blueprints for the production of new intermediate capital goods (such as new digital devices and software) that can be used in the production of final goods and services. Producers are assumed to achieve higher productivity (in their transformation of labor and capital into final goods) through the use of a richer array of intermediate capital goods. Some clever math, beyond the scope of the current chapter, describes a more detailed production function involving multiple capital goods, and shows that total output can ultimately be summarized this way, as a function only of the total value of all the many capital goods, K, and the number of capital goods, A, as well as the labor input, L.

A key innovation in the Romer model over neoclassical models is in the modeling of how the behavior of profit-seeking producers and research investors might determine how much labor is invested in research and development activities, and thus how rapidly A grows. The assumption of an exogenously given rate of growth of A is replaced by a "production function" for changes in A, together with behavioral assumptions describing how potential investors in research and development determine how much current labor to devote to the production of those changes. More specifically, the production function for changes in A is given by

(17) $\dot{A} = v(L_A,A)L_A$

where L_A is the quantity of labor devoted to research and development activities (and must be equal to the difference between the total labor stock and L_Y). The quantity *v(.)* represents the productivity of labor in producing new ideas. It may be a declining function of L_A if larger numbers of researchers are more likely to duplicate each others' efforts, so that higher L_A results in a smaller output of ideas per research worker. Per-researcher productivity, *v(.),* may rise or fall with the current level of technological attainment, *A*. If researchers who start with a richer body previous research are equipped to be more productive in their current research, then *v(.)* is an increasing function of *A*. In principle, *v(.)* might instead be a decreasing function of *A*, if the easiest-to-discover ideas tend to be discovered first, so that as *A* rises additional research tends to produce fewer new ideas. In the model these effects of L_A and *A* on *v(.)* are assumed to represent externalities, which are true at the level of the economy as a whole, but are ignored by individual investors in research and development, who take the current level of *v(.)* as unaffected by their choices.

Inventors are assumed to have exclusive rights to the use of their new designs. They may either use the designs to produce and sell the new intermediate goods themselves, or sell the use of the ideas to independent intermediate goods producers. Either way, the exclusive design use rights give intermediate goods producers monopoly power in their sales to final goods producers. It is these monopoly profits in intermediate goods production that provide the motivation for investment in research to produce new designs. Final producers, who sell their goods in perfectly competitive markets, find all currently available intermediate goods useful and imperfectly substitutable. As the number of intermediate goods available in the economy rises, the profits earned by individual intermediate goods producers fall. Research investors produce designs to the point at which the expected monopoly profits received from a new design just cover the cost of producing the design.

This model points to several reasons why the behavior of private individuals in markets might fail to deliver the socially most desirable investments in the development of new technology, and thus offer additional reasons to think that well-designed government intervention might enhance growth behavior. First, it seems likely that today's research discoveries make tomorrow's research sector more productive, but unlikely that today's research investors take this positive future impact of their work into account when making their current investment decisions. This provides a reason to suspect that people will under-invest in research. Second, it is possible that when the number of current researchers rises (as the rate of investment in the research sector increases), the tendency for researchers to duplicate each others' efforts may rise, implying a reduction in the per-researcher production of distinct new ideas. Failing to take this effect into account, research investors might tend to over-invest in research. Finally, it can be shown in these models that while monopoly profits are crucial as a source of return to research and development activities, they actually understate the benefits to society of the new ideas, suggesting another reason why private investors might under-invest in research. The net impact of all these imperfections is theoretically ambiguous, but they suggest the value of paying careful attention to the context-specific micro-economics of productivity advance.

Now consider models that analyze the rate of diffusion of new technologies to technology followers (see, for example, Jones, 1998, Chapter 6). The world technology frontier is assumed to grow through purposeful research and development efforts in the technology leaders, while the level of A in technology follower countries is assumed to grow through efforts to adopt new technologies that are closer and closer to the technology frontier. The challenges of adopting new technologies may include the need to achieve high enough levels of human capital for operating the technologies successfully, the need to find funds to purchase the technology from a foreign inventor, and the need to convince the foreign inventor that its patent-based monopoly power over the idea will be protected within the technology follower country. Such models thus reinforce interest in human capital investments as a source of growth, and raise questions about protection of intellectual property rights in developing countries (a subject to which we return in later chapters).

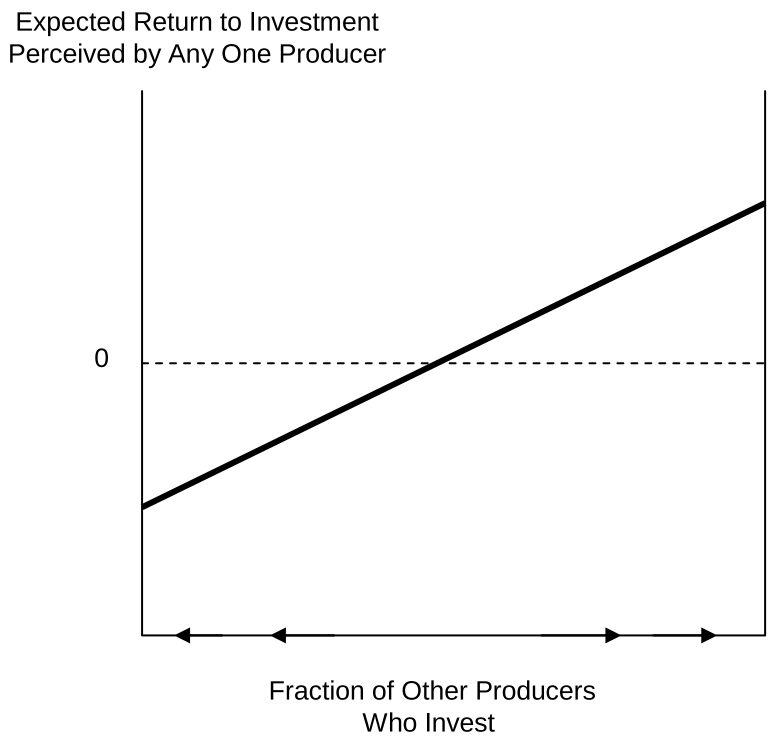
Investigating poverty traps

At the same time as endogenous growth theorists took the apparent importance of technical change to long-term growth as their starting point, and sought to understand its theoretical underpinnings, other theorists took the apparent failure of some poor countries to grow as their starting point, and sought to illuminate a wide variety of reasons why poverty might be **self-reinforcing.** That is, they examined reasons why in poor economies, or for poor groups within economies, poverty itself might in some sense be the barrier that prevents potential investors from investing (whether in physical capital, human capital or modern technology), thus preventing the countries from growing out of poverty.

Macro poverty traps. One class of models examined reasons for what we might call "macro poverty traps." Simple versions of these models conceive of economies as populated by large numbers of producers, each of whom faces a choice between two alternatives: (1) continue producing with low productivity methods, or (2) invest in order to produce using higher productivity methods. Murphy, Shleifer and Vishny (1989), which formalizes an idea first suggested by Rosenstein-Rodan (1943), construes the high productivity option to be modern manufacturing production and the low productivity option to be traditional handicraft production. Models of this sort seek to illuminate reasons why countries might end up in "bad equilibria", in which all producers use low-productivity methods (dooming the economy to low income per capita) and no one perceives any incentive to increase productivity by investing. Countries may end up in such equilibria even when "good equilibria", in which everyone invests and achieves high productivity, are also possible.

Central to such models are relationships of the sort stylized in Figure 3A.4. The horizontal axis measures the fraction of producers choosing to invest, ranging from zero at the left end to one at the right. Movements to the right along this axis signify increases in the economy's level of investment and in the level of income per capita achieved after investment. The left end is thus associated with poverty and stagnation, while the right end is associated with prosperity and growth. The vertical axis measures the returns that individual producers anticipate receiving if they choose to invest in setting up a modern manufacturing enterprise. If they are to perceive an incentive to invest, this return must be greater than zero. The economy faces the potential of falling into a macro poverty trap if the investment returns perceived by any one producer are related to the prevalence of high productivity choices among other producers in the economy in the way illustrated in the diagram. The *slope* of the investment returns schedule is positive, indicating **complementarity** among investment decisions: as more producers decide to invest, the return anticipated by any one producer rises. The *height* of the curve is such that when no one invests (at the left hand end of the diagram), the anticipated return on investment is negative, while if all producers invest (at the right end of the diagram), the anticipated return is positive. If this is the case, then the economy is characterized by **dual equilibria.** This means that the economy could settle down into two quite different equilibrium levels of investment and prosperity, in which no forces create any tendency for change. The "good equilibrium" is found at the right hand end of the diagram. If all producers choose the high productivity option, then all producers perceive strong incentives to continue pursuing the high productivity option. Unfortunately, this economy also has the potential to end up in a "bad equilibrium" at the left-most end of the diagram. If no producers invest, then no producer has any incentive to invest. Even if a few bold producers happened to invest, returns would remain sufficiently low as to discourage investment, tending to return the economy to the bad equilibrium.

Figure 3A.4



Whether the dual equilibrium vision of Figure 3A.4 is relevant to the real world depends on whether the investment decisions of many producers really are interdependent in a way that causes the anticipated returns schedule to rise as it does in the figure. Authors in the macro poverty traps literature offer diverse reasons why this might be the case. Murphy, et al. (1989) offer several reasons, all related to the role that individual investments play in increasing the "size of the market" for the goods produced by other investment projects. If, for example, producers must pay wage premiums to attract workers out of traditional handicraft activities into modern manufacturing, then when they set up a new modern enterprise they increase total wage income in the economy. If workers want to spend their income on the wide variety of goods produced by the economy's many potential investors, then this addition to wage income represents an increase in the demand for the outputs that might be produced by the economy's many other potential investors. This increase in the "size of the market" faced by other potential producers may mean that an investment by one producer increases the expected investment returns perceived by other investors, motivating the upward-sloping expected return schedule in Figure 3A.4. As the authors point out, increasing market size only increases the potential return on investment (and yields the positive slope in the figure) if modern manufacturing production is subject to increasing returns to scale, so that unit costs of production fall as the scale of production increases and costs thus fall below price only if the scale of production is sufficiently large. Thus assumptions of increasing returns to scale at the micro level of individual investment projects are crucial for constructing macro poverty trap models of this sort.

As the literature has demonstrated, macro level poverty traps can arise for quite different reasons. Rather than focusing on the choice between traditional and modern manufacturing production, Murphy, et al (1993) focus on choices between productive investment and rent-seeking. As defined in Chapter 3, "rent-seeking" is a label applied to activities in which actors attempt to acquire for themselves wealth produced by others, including theft, demands for bribes, and more. The authors suggest that any one producer's incentive to engage in productive investment rather than rent-seeking will rise as more producers around them also choose investment over rent-seeking. Turning this around: the greater the share of potential investors who choose rent-seeking, the greater the incentive to become another rent-seeker. Why might this be? As larger fractions of the population engage in fraud, theft, and other efforts to gain by taking rather than producing, the probability of being caught and punished for such activities may fall, and the social penalties for those who are caught may also fall, tending to reduce the costs relative to the benefits of rent-seeking. At the same time as rates of fraud, theft and other forms of rent-seeking rise, investors expect to enjoy fewer of the rewards of their investment. Thus it is possible for societies to become stuck in bad equilibria characterized by much rent-seeking and little investment.

Models of macro poverty traps construe poor economies' failure to grow as the result of a **coordination problem.** They are poor because no one individual has an incentive to undertake the kind of investment that can make the economy prosperous. Only if many potential investors could be encouraged to invest in more productive activities simultaneously would they find that their investments pay off, leading the economy to a better equilibrium. The models thus point to a coordinating role for government similar to that suggested by the collective action problems that prevent investments in infrastructure assets with public good qualities.

Micro poverty traps. Another class of models (such as those found in Galor and Zeira, 1993, and Banerjee and Newman, 1993) points to the potential for **micro poverty traps**, illuminating reasons why poor households might be unable to undertake profitable investment, and thus remain mired in poverty, even while more prosperous households within the same economy undertake profitable investment and experience the benefits of growth. Again, increasing returns to scale at the micro level of individual investment projects are important to the theoretical arguments, which also point to imperfections in credit markets as an integral part of the story. If the investments that can raise productivity and income are profitable only when undertaken at a scale above some minimum threshold, then only investors with adequate financing for undertaking sufficiently large investments may hope to profit from investment. Potential investors for whom only smaller investments are feasible may rationally choose not to invest at all. "Information problems" in credit markets (relating to lenders' inability fully assess borrowers' abilities and motivations or to fully monitor and supervise their use of credit, as discussed in Chapter 9) may cause lenders (such as banks) to offer credit only in proportion to the collateral that a borrower has to offer. The poor, who lack collateral (and savings, with which they could avoid the need to borrow for investment), are likely to find investments of adequate size infeasible, and thus choose not to invest. Their incomes remain stagnant, even while the wealthy around them invest profitably and enjoy growing incomes. In such models, re-distributions of wealth or income from wealthy to poor households can speed rates of growth, by making it feasible for larger fractions of the population to undertake investments of adequate size.

Political economy of growth policy

A final growth literature that burgeoned in the 1990s asked the question: If good policies are useful for promoting growth, then why do some countries persist in implementing bad policies? This large and wide-ranging literature on the "political economy of growth" acknowledges that policies are not put into place by selfless politicians seeking to maximize social welfare. Rather, policies are put into place through the interplay of diverse political actors who are guided by a mix of personal and social motivations, whose interests conflict, whose political power differ, and who are constrained by their countries' political institutions. "Political institutions" are the formal rules and informal norms that govern the way policies are selected, and are differentiated along such lines as whether they are democratic or nondemocratic institutions. The literature examines three layers of questions (Acemoglu, 2009). First, within a given set of political institutions (e.g. within a democracy), what determines the growth-relevant policies that are put into place? For example, Alesina and Rodrik (1994) describe how greater wealth inequality, which (holding the average level of wealth constant) implies a larger fraction of the population living at low levels of wealth, might tip the scales of democratic political institutions in favor of policies that discourage investment and growth by imposing higher taxes on returns to investment, in order to supply public goods of more immediate value to those of low and medium incomes. In a similar vein, Alesina, et al. (1999) describe how, in societies characterized by greater ethnic polarization, democratic processes might lead to less investment in infrastructure and other public goods. Second, how do different types of political institutions compare in their facilitation of good growth policies? For example, Przeworski and Limongi (1993) compare the performance of democracies and nondemocracies, concluding that there is little evidence of a systematic tendency for democracies to perform better, and suggesting that differences in political institutions among democracies and among non-democracies are in fact more important in shaping growth outcomes. Finally, the deepest level of question: what determines the type of political institutions in place in a country? For example, Acemoglu (2009) discusses the evolution of political institutions over time, pointing out the interplay between the inertia of political institutions already in place and the role of economic and social change in creating groups with an interest in creating new political institutions.

References and Suggested Readings

Acemoglu, D., 2009, *Introduction to Modern Economic Growth*, Princeton University Press.

Chenery, H.B. and A.M. Strout, 1966, "Foreign Assistance and Economic Development," *American Economic Review 56(4),* pp. 679-733.

Banerjee, A.V. and A.F. Newman, 1993, "Occupational Choice and the Process of Development," *The Journal of Political Economy 101(2),* pp. 274-298.

Durlauf, S.N., A. Kourtellos, and C. M. Tan, forthcoming, "Empirics of Growth and Development", in A.K. Dutt and J. Ros, eds., *International Handbook of Development Economics*, Edward Elgar.

Galor, O. and J. Zeira, 1993, "Income Distribution and Macroeconomics," *The Review of Economic Studies 60(1),* pp. 35-52.

Alesina, A., R. Baqir and W. Easterly, 1999, "Public Goods and Ethnic Divisions," *The Quarterly Journal of Economics 114(4)*, pp. 1243-1284.

Alesina, A. and D. Rodrik, 1994, "Distributive Politics and Economic Growth," *The Quarterly Journal of Economics 109(2),* pp. 465-490.

Domar, E., 1946, "Capital Expansion, Rate of Growth, and Employment," *Econometrica 14*, pp. 137-147.

Easterly, W., 2002, *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics,* MIT Press.

Grossman, G.M. and E. Helpman, 1994, "Endogenous Innovation in the Theory of Growth," *The Journal of Economic Perspectives 8(1)*, pp. 23-44.

Harrod, R.F., 1939, "An Essay in Dynamic Theory," *Economic Journal 49(193)*, pp. 14-33.

Jones, C.I., 1998, *Introduction to Economic Growth,* W.W. Norton.

Lewis, W.A. (1954), "Economic Development with Unlimited Supplies of Labour," *Manchester School 22*, pp. 139-192.

Little, I.M.D.,1982, *Economic Development: Theory, Policy, and International Relations,* Basic Books, Inc.

Lucas, R.E., Jr., 1988, "On the Mechanics of Economic Development," *Journal of Monetary Economics 22,* pp. 3-42.

Mincer, J., 1970, "The Distribution of Labor Incomes: A Survey with Special Reference to the Human Capital Approach," *Journal of Economic Literature 8(1),* pp. 1-26.

Mankiw, G.N., D. Romer and D.N. Weil, 1992, "A Contribution to the Empirics of Economic Growth," *The Quarterly Journal of Economics 107(2)*, pp. 407-437.

Murphy, K.M., A. Shleifer and R.W. Vishny, 1989, "Industrialization and the Big Push," *The Journal of Political Economy 97(5),* pp. 1003-1026.

Murphy, K.M., A. Shleifer, R.W. Vishny, 1993, "Why Is Rent Seeking So Costly to Growth?" *The American Economic Review 83(2)*, Papers and Proceedings of the Hundred and Fifth Annual Meeting of the American Economic Association, pp. 409-414.

Przeworski, A. and F. Limongi, 1993, "Political Regimes and Economic Growth," *Journal of Economic Perspectives* 7(3), p. 51-69.

Romer, P.M., 1990, "Endogenous Technological Change," *The Journal of Political Economy 98(5)*, Part 2: The Problem of Development: A Conference on the Institute for the Study of Free Enterprise Systems, pp. S71-S102.

Romer, P.M., 1994, "The Origins of Endogenous Growth," *The Journal of Economic Perspectives 18(1),* pp. 3-22.

Rosenstein-Rodan, P.N., 1943, "Problems of Industrialisation of Eastern and South-eastern Europe," *The Economic Journal 53 (210/211)*, pp. 202-211.

Solow, R.M., 1956, "A Contribution to the Theory of Economic Growth," *The Quarterly Journal of Economics 70(1),* pp. 65-94.

Solow, R.M., 1994, "Perspectives on Growth Theory," *The Journal of Economic Perspectives 18(1),* pp. 45-54.

Questions for Review

1. Define the terms **capital fundamentalism, structuralism** and **dualism**. Discuss the implications of these beliefs for growth policy.
2. Theoretical growth modeling exercises begin with the specification of **assumptions describing the economy's production technology** and **behavioral assumptions**, and then work out the implications of these assumptions for the economy's growth trajectory and how this trajectory will be influenced by changes in key **exogenous** parameters of the model. What is described by a model's "assumptions describing the economy's production technology"? What is described by a model's "behavioral assumptions"? What is the distinction between **exogenous** parameters and **endogenous** variables?
3. For the Harrod-Domar Growth Model, (a) describe the technological assumptions, (b) describe the behavioral assumptions, (c) identify the key exogenous parameters that shape the economy's growth performance, (d) discuss what happens to the economy's growth performance when each parameter is increased, and (e) discuss the implications the model was thought to yield for policy.
4. What is meant by a **fixed proportions production technology**?
5. Describe the **knife-edge** property of this model, which eventually contributed to its replacement by new models.
6. For the Lewis model, (a) describe the technological assumptions relevant to the **subsistence sector** and the **capitalist sector,** (b) describe the behavioral assumptions relevant to the subsistence sector and the capitalist sector, (c) define the notion of **structural change** that inevitably accompanies economic growth in this model, (d) employing Figure 3.1,discuss what would happen to the initial rate of growth if the wage required to draw workers into the capitalist sector fell or if the value of the marginal

product of labor in the capitalist sector were shifted up through policies that raise the capitalist sector output price, and (e) discuss the implications the model was thought to yield for policy.

7. For the Simple Neoclassical Growth Model, and then again for the Neoclassical Growth Model with Technical Change: (a) describe the technological assumptions, (b) describe the behavioral assumptions, (c) identify the key exogenous parameters that shape the economy's growth performance, (d) discuss what happens to the economy's growth performance when each parameter is increased, and (e) discuss the implications the model was thought to yield for policy.
8. What is meant by a **variable proportions production technology**?
9. What is **labor augmenting technical change**?
10. Define **human capital** and discuss how it can be incorporated into the Neoclassical growth model.
11. What does it mean to "endogenize technical change" in a formal growth model? Why was it difficult to endogenize technical change in growth model assuming constant returns to scale and perfect competition?
12. What is meant by references to **endogenous growth models**?
13. What are **increasing returns to scale** and what is the significance of possible increasing returns to scale for our understanding of economic growth?
14. Describe the two approaches theorists took to endogenizing technical change, and some of the key implications of each approach.
15. What is the distinction between **technology leaders** and **technology followers**?
16. What does it mean for poverty to be **self-reinforcing**?
17. Use Figure 3A.4 to describe the conditions under which economies might have the potential to settle into two very different equilibria with regard to investment and prosperity. In such a model, in what sense is poverty the result of a **coordination problem?**
18. What is meant by the term **micro poverty trap**?
19. What role do increasing returns to scale play in generating the potential for macro and micro poverty traps?
20. What sorts of research questions are asked in the literature on "the political economy of growth"?

Questions for Discussion

1. The Harrod-Domar model exhibits the "knife-edge" property that growth is consistent with continuously full employment only if *(s/k)-d* just happens to equal *n*. This knife-edge property is eliminated in the Simple Neoclassical Model, which is characterized by continuously full employment. Discuss the importance of assuming variable rather than fixed proportions production technology for working this change.
2. For each of the bodies of growth theory described in this chapter, state whether production technologies are characterized by constant or increasing returns to scale. Where production technologies are characterized by increasing returns to scale, state whether the increasing returns to scale are present only at the aggregate level (while individual producers continue to work under the assumption that their technology is characterized by constant returns to scale), or whether increasing returns to scale are also present and important at the level of the individual firm. What roles do these assumptions play in shaping the predictions of the models?
3. Discuss the role of empirical research (including simple empirical observations) in driving the evolution of growth theory.

Problems

1. As discussed in the text, the assumptions of the Harrod-Domar model may be summarized by the equations

(i) $$Y = \left(\frac{1}{v}\right) K$$

(ii) $\dot{Y} = (\frac{1}{v})\dot{K}$.

and

(iii) $\dot{K} = sY - dK$

where the notation is as defined in the text. (Equations (i) and (ii) are two ways of stating the same assumption, but both expressions are useful to remember in the derivations you will be required to do below.)

(a) Demonstrate that equations (ii) and (iii) together imply the following result regarding the growth rate of GDP. (Notice that the text offers guidance about how to derive this equation.)

$$\frac{\dot{Y}}{Y} = \frac{s}{v} - d$$

(b) Show that equations (i) and (iii) together imply the following result regarding the growth rate of K.

$$\frac{\dot{K}}{K} = \frac{s}{v} - d$$

2. Consider the Neoclassical Growth Model with Technical Change, and its diagrammatic summary in Figure 3A.3. Suppose the rate of population growth n increased. Which element of the graph (i.e. the $k^*(n+d+g)$ line or the $sf(k^*)$ curve) would change and it what way? Draw such a change into a graph like Figure 3A.3. When the rate of population growth increases like this, what happens to the steady state level of income per capita? What happens to the steady state rate of growth in income per capita? What is the immediate impact on the rate of growth in k*? What is the immediate impact on the rate of growth of GDP per capita? Using intuitive, plain language, explain the intuition behind why the increase in the population growth rate has the short run impact on growth that you just described, and why that short-run impact eventually fades away.

3. Suppose the aggregate production function takes the form

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

where $A(K)=K^{\beta}$ describes an external, economy-wide effect of K on A, $F(K,L)=L^{\alpha}K^{1-\alpha}$ and $0<\alpha<1$.

a. Demonstrate that if you double both K and L while holding the initial value of A constant (i.e. ignoring the external effect of K on A) Y doubles.

b. Demonstrate that if you double both K and L, taking into account the external effect of K on A, Y more than doubles.

c. Derive an expression for the marginal product of capital while ignoring the external effect of K on A. That is, holding A constant (rather than treating it as a function of K), take the derivative of the aggregate production function with respect to K. Show that if K increases while L holds constant, this marginal product of capital falls.

d. Derive an expression for the marginal product of capital taking into account the external effect of K on A. Show that if K increases while L holds constant, the marginal product of capital may fall or rise, depending on the values of α and β.

4. Critical to the construction of some "macro poverty trap" models is the assumption that the profitability of setting up a modern, high productivity establishment in any one sector depends positively on the "size of the market" the establishment will face (which in turn is taken to depend positively on number of other sectors in which modern, high productivity establishments have set up). In this problem you will examine a very simple technology for modern production, involving a "fixed cost of set-up", in which profitability of setting up indeed depends positively on the number of units of output the firm anticipates being able to sell. Suppose that modern production can take place only after incurring a fixed cost of F units of labor. Once that cost is incurred, each unit of additional labor produces $\alpha>1$ units of output. The price of a unit of

labor is 1. Suppose the price of a unit of output is 1 also (perhaps because lower-productivity traditional establishments can produce a unit of output using 1 unit of labor). Let Q be the quantity of output the potential investor anticipates selling.

a. Derive an expression for the producer's profits (i.e. revenue minus labor costs) as a function of Q, F and α.

b. Making use of this expression, show that if F=0 (meaning that there are no fixed costs) then setting up is profitable regardless of the level of Q.

c. Now assume F>0. Derive an expression for the minimum level of Q at which production is profitable. How does this minimum profitable scale change as F increases? As α increases?