Introduction

I.1 The Importance of Growth

To think about the importance of economic growth, we begin by assessing the long-term performance of the U.S. economy. The real per capita gross domestic product (GDP) in the United States grew by a factor of 10 from $3340 in 1870 to $33,330 in 2000, all measured in 1996 dollars. This increase in per capita GDP corresponds to a growth rate of 1.8 percent per year. This performance gave the United States the second-highest level of per capita GDP in the world in 2000 (after Luxembourg, a country with a population of only about 400,000).[1]

To appreciate the consequences of apparently small differentials in growth rates when compounded over long periods of time, we can calculate where the United States would have been in 2000 if it had grown since 1870 at 0.8 percent per year, one percentage point per year below its actual rate. A growth rate of 0.8 percent per year is close to the rate experienced in the long run—from 1900 to 1987—by India (0.64 percent per year), Pakistan (0.88 percent per year), and the Philippines (0.86 percent per year). If the United States had begun in 1870 at a real per capita GDP of $3340 and had then grown at 0.8 percent per year over the next 130 years, its per capita GDP in 2000 would have been $9450, only 2.8 times the value in 1870 and 28 percent of the actual value in 2000 of $33,330. Then, instead of ranking second in the world in 2000, the United States would have ranked 45th out of 150 countries with data. To put it another way, if the growth rate had been lower by just 1 percentage point per year, the U.S. per capita GDP in 2000 would have been close to that in Mexico and Poland.

Suppose, alternatively, that the U.S. real per capita GDP had grown since 1870 at 2.8 percent per year, 1 percentage point per year greater than the actual value. This higher growth rate is close to those experienced in the long run by Japan (2.95 percent per year from 1890 to 1990) and Taiwan (2.75 percent per year from 1900 to 1987). If the United States had still begun in 1870 at a per capita GDP of $3340 and had then grown at 2.8 percent per year over the next 130 years, its per capita GDP in 2000 would have been $127,000—38 times the value in 1870 and 3.8 times the actual value in 2000 of $33,330. A per capita GDP of $127,000 is well outside the historical experience of any country and may, in fact, be infeasible (although people in 1870 probably would have thought the same about $33,330). We can say, however, that a continuation of the long-term U.S. growth rate of 1.8 percent per year implies that the United States will not attain a per capita GDP of $127,000 until 2074.

1. The long-term data on GDP come from Maddison (1991) and are discussed in chapter 12. Recent data are from Heston, Summers, and Aten (2002) and are also discussed in chapter 12.

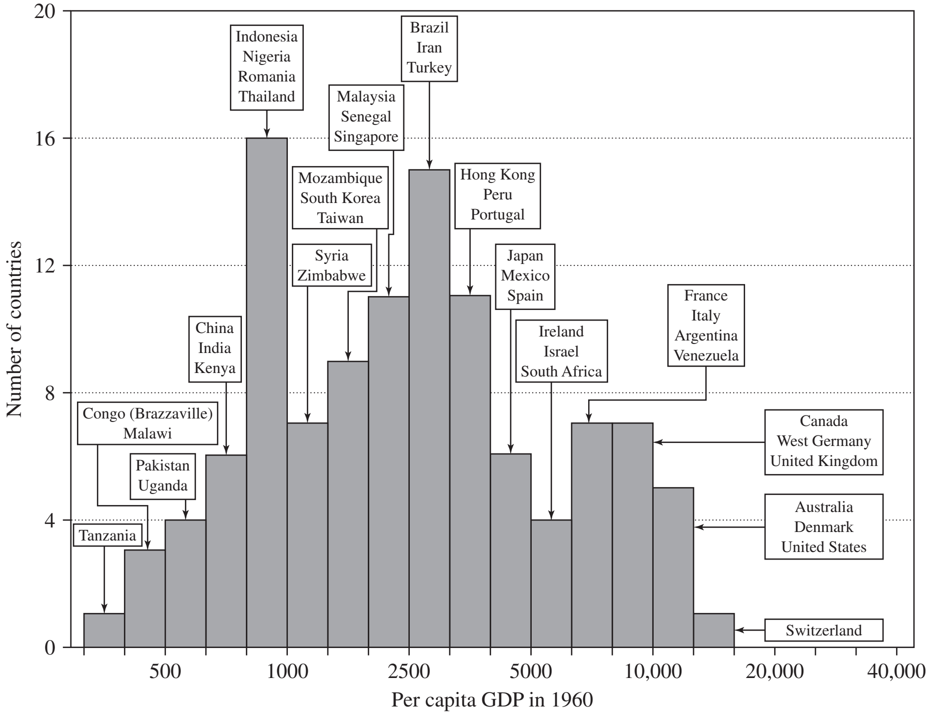

Figure I.1
Histogram for per capita GDP in 1960. The data, for 113 countries, are the purchasing-power-parity (PPP) adjusted values from Penn World Tables version 6.1, as described in Summers and Heston (1991) and Heston, Summers, and Aten (2002). Representative countries are labeled within each group.

The comparison of levels of real per capita GDP over a century involves multiples as high as 20; for example, Japan's per capita GDP in 1990 was about 20 times that in 1890. Comparisons of levels of per capita GDP across countries at a point in time exhibit even greater multiples. Figure I.1 shows a histogram for the log of real per capita GDP for 113 countries (those with the available data) in 1960. The mean value corresponds to a per capita GDP of $3390 (1996 U.S. dollars). The standard deviation of the log of real per capita GDP—a measure of the proportionate dispersion of real per capita GDP—was 0.89. This number means that a 1-standard-deviation band around the mean encompassed a range from 0.41 of the mean to 2.4 times the mean. The highest per capita GDP of $14,980 for Switzerland was 39 times the lowest value of $381 for Tanzania. The United States was second with a value of $12,270. The figure shows representative countries for each range of per capita GDP. The broad picture is that the richest countries included the OECD and

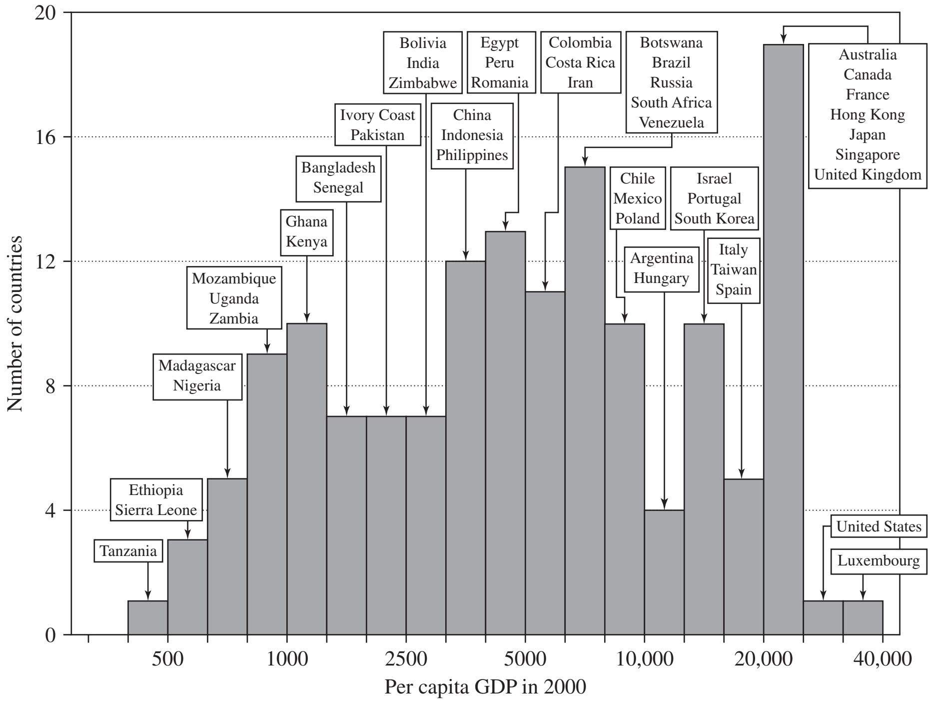

Figure I.2
Histogram for per capita GDP in 2000. The data, for 150 countries, are from the sources noted for figure I.1. Representative countries are labeled within each group.

a few places in Latin America, such as Argentina and Venezuela. Most of Latin America was in a middle range of per capita GDP. The poorer countries were a mixture of African and Asian countries, but some Asian countries were in a middle range of per capita GDP.

Figure I.2 shows a comparable histogram for 150 countries in 2000. The mean here corresponds to a per capita GDP of $8490, 2.5 times the value in 1960. The standard deviation of the log of per capita GDP in 2000 was 1.12, implying that a 1-standard-deviation band ranged from 0.33 of the mean to 3.1 times the mean. Hence, the proportionate dispersion of per capita GDP increased from 1960 to 2000. The highest value in 2000, $43,990 for Luxembourg, was 91 times the lowest value—$482 for Tanzania. (The Democratic Republic of Congo would be poorer, but the data are unavailable for 2000.) If we ignore Luxembourg because of its small size and compare Tanzania's per capita GDP with the second-highest value, $33,330 for the United States, the multiple is 69. Figure I.2 again marks out representative countries within each range of per capita GDP. The OECD countries still dominated the top group, joined by some East Asian countries. Most other Asian countries were in the middle range of per capita GDP, as were most Latin American countries. The lower range in 2000 was dominated by sub-Saharan Africa.

To appreciate the spreads in per capita GDP that prevailed in 2000, consider the situation of Tanzania, the poorest country shown in figure I.2. If Tanzania were to grow at the long-term U.S. rate of 1.8 percent per year, it would take 235 years to reach the 2000 level of U.S. per capita GDP. The required interval would still be 154 years if Tanzania were to grow at the long-term Japanese rate of 2.75 percent per year.

For 112 countries with the necessary data, the average growth rate of real per capita GDP between 1960 and 2000 was 1.8 percent per year—coincidentally the same as the long-term U.S. rate—with a standard deviation of 1.7.[2] Figure I.3 has a histogram of these growth rates; the range is from −3.2 percent per year for the Democratic Republic of Congo (the former Zaire) to 6.4 percent per year for Taiwan. (If not for missing data, the lowest-growing country would probably be Iraq.) Forty-year differences in growth rates of this magnitude have enormous consequences for standards of living. Taiwan raised its real per capita GDP by a factor of 13 from $1430 in 1960 (rank 76 out of 113 countries) to $18,730 in 2000 (rank 24 of 150), while the Democratic Republic of Congo lowered its real per capita GDP by a factor of 0.3 from $980 in 1960 (rank 93 of 113) to $320 in 1995—if not for missing data, this country would have the lowest per capita GDP in 2000.

A few other countries had growth rates from 1960 to 2000 that were nearly as high as Taiwan's; those with rates above 5 percent per year were Singapore with 6.2 percent, South Korea with 5.9 percent, Hong Kong with 5.4 percent, and Botswana with 5.1 percent. These countries increased their levels of per capita GDP by a multiple of at least 7 over 40 years. Just below came Thailand and Cyprus at 4.6 percent growth, China at 4.3 percent, Japan at 4.2 percent (with rapid growth mainly into the 1970s), and Ireland at 4.1 percent. Figure I.3 shows that a number of other OECD countries came in the next-highest growth groups, along with a few countries in Latin America (including Brazil and Chile) and more in Asia (including Indonesia, India, Pakistan, and Turkey). The United States ranked 40th in growth with a rate of 2.5 percent.

At the low end of growth, 16 countries aside from the Democratic Republic of Congo had negative growth rates of real per capita GDP from 1960 to 2000. The list (which would be substantially larger if not for missing data), starting from the bottom, is Central African Republic, Niger, Angola, Nicaragua, Mozambique, Madagascar, Nigeria, Zambia,

2. These statistics include the Democratic Republic of Congo (the former Zaire), for which the data are for 1960 to 1995.

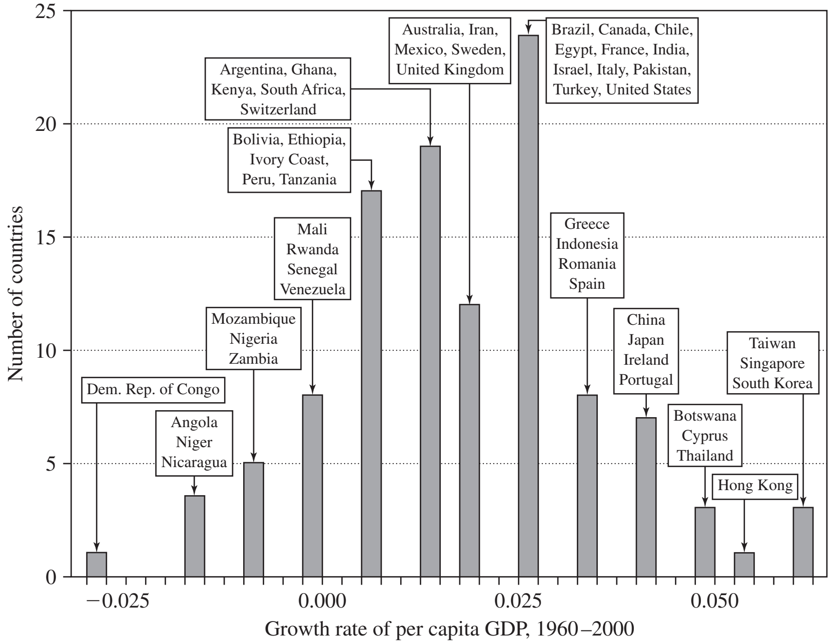

Figure I.3
Histogram for growth rate of per capita GDP from 1960 to 2000. The growth rates are computed for 112 countries from the values of per capita GDP shown for 1960 and 2000 in figures I.1 and I.2. For Democratic Republic of Congo (former Zaire), the growth rate is for 1960 to 1995. West Germany is the only country included in figure I.1 (for 1960) but excluded from figure I.3 (because of data problems caused by the reunification of Germany). Representative countries are labeled within each group.

Chad, Comoros, Venezuela, Senegal, Rwanda, Togo, Burundi, and Mali. Thus, except for Nicaragua and Venezuela, this group comprises only sub-Saharan African countries. For the 38 sub-Saharan African countries with data, the mean growth rate from 1960 to 2000 was only 0.6 percent per year. Hence, the typical country in sub-Saharan Africa increased its per capita GDP by a factor of only 1.3 over 40 years. Just above the African growth rates came a few slow-growing countries in Latin America, including Bolivia, Peru, and Argentina.

As a rough generalization for regional growth experiences, we can say that sub-Saharan Africa started relatively poor in 1960 and grew at the lowest rate, so it ended up by far the poorest area in 2000. Asia started only slightly above Africa in many cases but grew rapidly and ended up mostly in the middle. Latin America started in the mid to high range, grew somewhat below average, and therefore ended up mostly in the middle along with Asia.

Finally, the OECD countries started highest in 1960, grew in a middle range or better, and therefore ended up still the richest.

If we want to understand why countries differ dramatically in standards of living (figures I.1 and I.2), we have to understand why countries experience such sharp divergences in long-term growth rates (figure I.3). Even small differences in these growth rates, when cumulated over 40 years or more, have much greater consequences for standards of living than the kinds of short-term business fluctuations that have typically occupied most of the attention of macroeconomists. To put it another way, if we can learn about government policy options that have even small effects on long-term growth rates, we can contribute much more to improvements in standards of living than has been provided by the entire history of macroeconomic analysis of countercyclical policy and fine-tuning. Economic growth—the subject matter of this book—is the part of macroeconomics that really matters.

I.2 The World Income Distribution

Although we focus in this book on the theoretical and empirical determinants of aggregate economic growth, we should keep in mind that growth has important implications for the welfare of individuals. In fact, aggregate growth is probably the single most important factor affecting individual levels of income. Hence, understanding the determinants of aggregate economic growth is the key to understanding how to increase the standards of living of individuals in the world and, thereby, to lessen world poverty.

Figure I.4 shows the evolution of the world's per capita GDP from 1970 to 2000.[3] It is clear that the average person on the planet has been getting richer over time. But the positive average growth rate over the last three decades does not mean that the income of all citizens has increased. In particular, it does not mean that the incomes of the poorest people have grown nor that the number of people whose incomes are below a certain poverty line (say one dollar a day, as defined by the World Bank) has declined.[4] Indeed, if inequality

3. The "world" is approximated by the 126 countries (139 countries after the breakup of the Soviet Union in 1989) in Sala-i-Martin (2003a, 2003b). The individuals in these 126 countries made up about 95 percent of the world's population. World GDP per capita is estimated by adding up the data for individual countries from Heston, Summers, and Aten (2002) and then dividing by the world's population.

4. The quest for a "true" poverty line has a long tradition, but the current "one-dollar-a-day" line can be traced back to World Bank (1990). The World Bank originally defined the poverty line as one dollar a day in 1985 prices. Although the World Bank's own definition later changed to 1.08 dollars a day in 1993 dollars (notice that one 1985 dollar does not correspond to 1.08 1993 dollars), we use the original definition of one dollar a day in 1985 prices. One dollar a day (or 365 dollars a year) in 1985 prices becomes 495 dollars per year in 1996 prices, which is the base year of the Heston, Summers, and Aten (2002) data used to construct the world income distributions. Following Bhalla (2002), Sala-i-Martin (2003a) adjusts this poverty line upward by 15 percent to correct for the bias generated by the underreporting of the rich. This adjustment means that our "one-dollar-a-day" poverty line represents 570 dollars a year (or 1.5 dollars a day) in 1996 dollars.

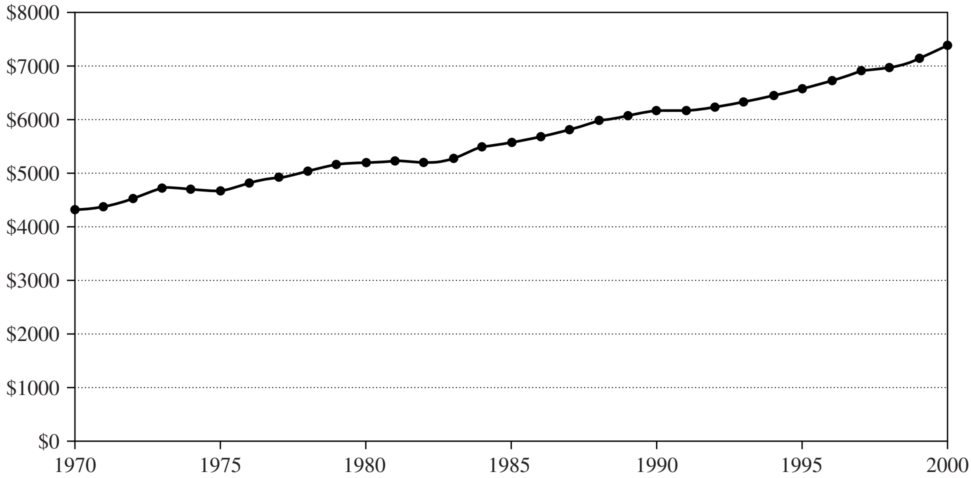

Figure I.4
World per capita GDP, 1970–2000. World per capita GDP is the sum of the GDPs for 126 countries (139 countries after the collapse of the Soviet Union) divided by population. The sample of 126 countries is the one used in Sala-i-Martin (2003a) and accounts for 95 percent of the world's population.

increased along with economic growth, it is possible for the world to have witnessed both positive per capita GDP growth and an increasing number of people below the poverty line. To assess how aggregate growth affects poverty, Sala-i-Martin (2003a) estimates the world distribution of individual income. To do so, he combines microeconomic survey and aggregate GDP data for each country, for every year between 1970 and 2000.[5] The result for 1970 is displayed in figure I.5. The horizontal axis plots the level of income (on a logarithmic scale), and the vertical axis has the number of people. The thin lines correspond to the income distributions of individual countries. Notice, for example, that China (the most populated country in the world) has a substantial fraction of the distribution below the $1/day line. The same is true for India and a large number of smaller countries. This pattern contrasts with the position of countries such as the United States, Japan, or even the USSR, which have very little of their distributions below the $1/day line. The thick line in figure I.5 is the integral of all the individual distributions. Therefore,

5. Sala-i-Martin (2003b) constructs an analogous distribution from which he estimates the number of people whose personal consumption expenditure is less than one dollar a day. The use of consumption, rather than income, accords better with the concept of "extreme poverty" used by international institutions such as the World Bank and the United Nations. However, personal consumption has the drawbacks of giving no credit to public services and saving.

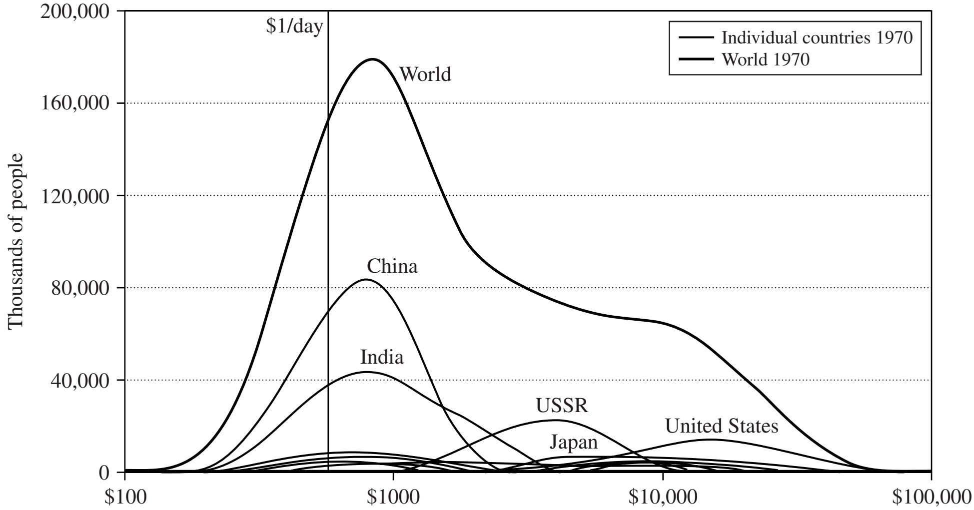

Figure I.5
The world distribution of income in 1970. The level of income is on the horizontal axis (on a logarithmic scale), and the number of people is on the vertical axis. The thin curves correspond to the income distributions of individual countries. The thick curve is the integral of individual country distributions and corresponds to the world distribution of income. The vertical line marks the poverty line (which corresponds to one dollar a day in 1985 prices). Source: Sala-i-Martin (2003a).

this line corresponds to the world distribution of income in 1970. Again, a substantial fraction of the world's citizens were poor (that is, had an income of less than $1/day) in 1970.

Figure I.6 displays the corresponding distributions for 2000. If one compares the 1970 with the 2000 distribution, one sees a number of interesting things. First, the world distribution of income has shifted to the right. This shift corresponds to the cumulated growth of per capita GDP. Second, we see that, underlying the evolution of worldwide income, there is a positive evolution of incomes in most countries in the world. Most countries increased their per capita GDP and, therefore, shifted to the right. Third, we see that the dispersion of the distributions for some countries, notably China, has increased over this period. In other words, income inequality rose within some large countries. Fourth, the increases in inequality within some countries have not been nearly enough to offset aggregate per capita growth, so that the fraction of the world's people whose incomes lie below the poverty line has declined dramatically.

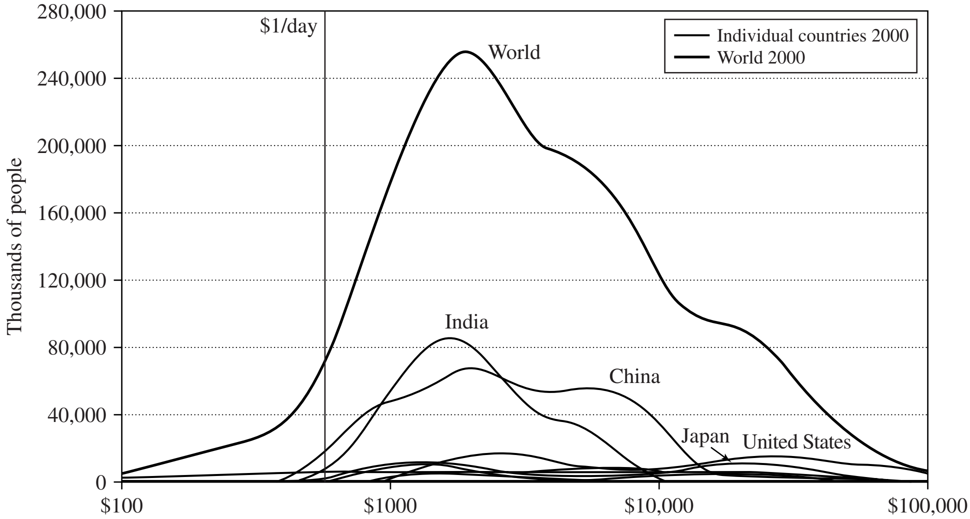

Figure I.6
The world distribution of income in 2000. The level of income is on the horizontal axis (on a logarithmic scale), and the number of people is on the vertical axis. The thin curves correspond to the income distributions of individual countries. The thick curve is the integral of individual country distributions and corresponds to the world distribution of income. The vertical line marks the poverty line (which corresponds to one dollar a day in 1985 prices). Source: Sala-i-Martin (2003a).

The exact fraction of the world's citizens that live below the poverty line can be computed from the distributions estimated by Sala-i-Martin (2003a).[6] These poverty rates, reported in figure I.7, have been cut by a factor of 3: whereas 20 percent of the world's citizens were poor in 1970, only 7 percent were poor in 2000.[7] Between 1970 and 1978, population growth more than offset the reduction in poverty rates. Indeed, Sala-i-Martin (2003a) shows that, during that period, the overall number of poor increased by 20 million people. But, since 1978, the total number of people with income below the $1/day threshold declined by more than 300 million. This achievement is all the more remarkable if we take into acount that overall population increased by more than 1.6 billion people during this period.

6. The World Bank, the United Nations, and many individual researchers define poverty in terms of consumption, rather than income. Sala-i-Martin (2003b) estimates poverty rates and head counts using consumption. The evolution of consumption poverty is similar to the one reported here for income although, obviously, the poverty rates are higher if one uses consumption instead of income and still uses the same poverty line.

7. Sala-i-Martin (2003a) reports cumulative distribution functions (CDFs) for 1970, 1980, 1990, and 2000. Using these CDFs, one can easily see that poverty rates have fallen dramatically over the last thirty years regardless of what poverty line one adopts. Thus, the conclusion that aggregate growth has reduced poverty is quite robust.

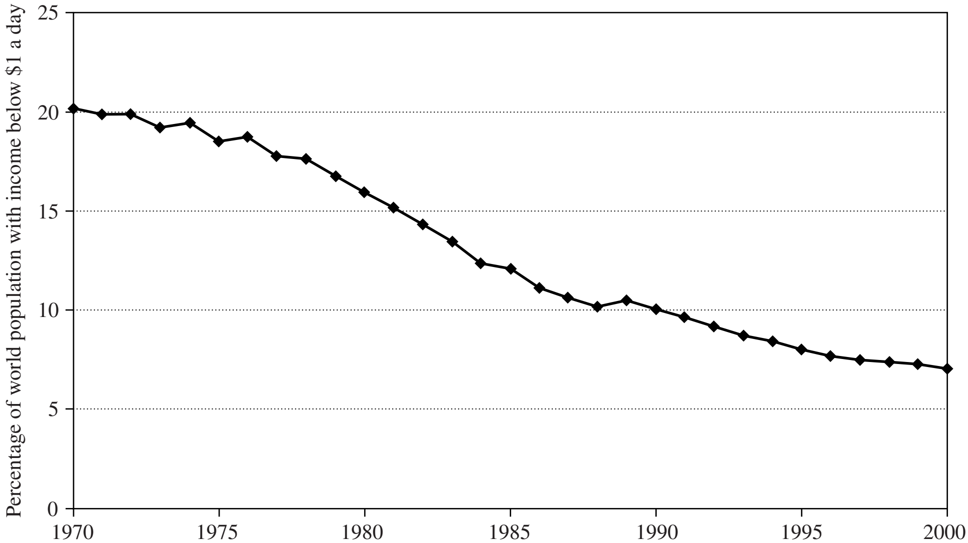

Figure I.7
World poverty rates. The graphs show the fraction of overall population with income below the poverty line. Source: Sala-i-Martin (2003a).

The clear conclusion is that economic growth led to substantial reductions in the world's poverty rates and head counts over the last thirty years. As mentioned earlier, this outcome was not inevitable: if aggregate growth had been accompanied by substantial increases in income inequality, it would have been possible for the mean of the income distribution to increase but also for the fraction of the distribution below a specified poverty threshold to also increase. Sala-i-Martin (2003a) shows that, even though this result is theoretically possible, the world did not behave this way over the last thirty years. Moreover, he also shows that world income inequality actually declined slightly between 1980 and 2000. This conclusion holds whether inequality is measured by the Gini coefficient, the Theil Index, the mean logarithmic deviation, various Atkinson indexes, the variance of log-income, or the coefficient of variation.

Sala-i-Martin (2003a) decomposes the world into regions and notes that poverty erradication has been most pronounced in the regions where growth has been the largest. Figure I.8 reports poverty rates for the poorest regions of the world: East Asia, South Asia, Latin America, Africa, the Middle East and North Africa (MENA), and Eastern Europe and Central Asia. In 1970, three of these regions had poverty rates close to or above 30 percent. Two of them (East Asia and South Asia) have experienced substantial reductions in poverty

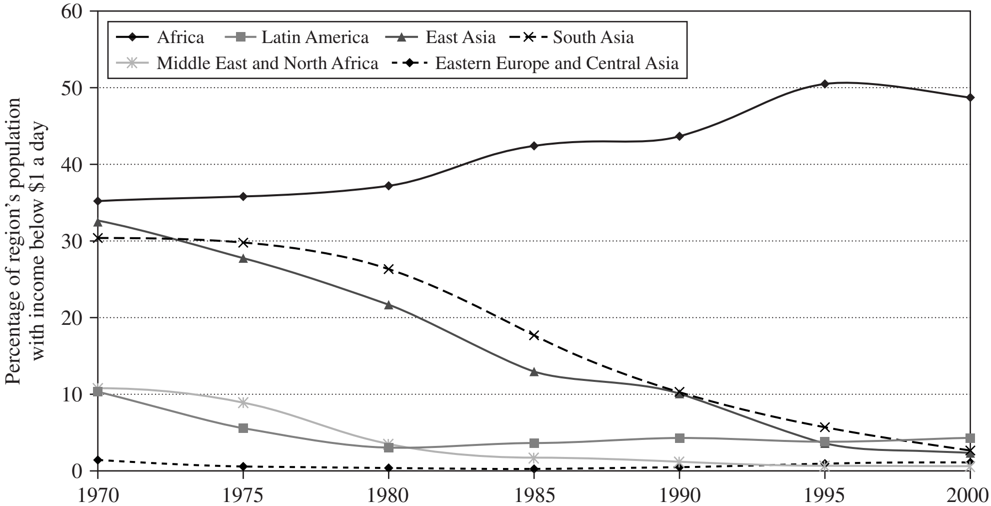

Figure I.8
Regional poverty rates. The graphs show the fraction of each region's population with income below the poverty line. The regions are the ones defined by the World Bank: East Asia, South Asia, Latin America, Africa, the Middle East and North Africa (MENA), and Eastern Europe and Central Asia. Source: Sala-i-Martin (2003a).

rates. These are the regions that also experienced large positive aggregate growth rates. The other region (Africa) has witnessed a dramatic increase in poverty rates over the last thirty years. We also know that per capita growth rates have been negative or close to zero for most countries in Africa. Figure I.8 also shows that two regions had poverty rates near 10 percent in 1970: Latin America and MENA. Both have experienced reductions in poverty rates. Latin America witnessed dramatic gains in the 1970s, when growth rates were substantial, but suffered a setback during the 1980s (the "lost decade," which featured negative growth rates). Poverty rates in Latin America stabilized during the 1990s. Poverty rates in MENA declined slightly between 1970 and 1975. The decline was very large during the high-growth decade that followed the oil shocks and then stabilized when aggregate growth stopped.

Finally, Eastern Europe and Central Asia (a region that includes the former Soviet Union) started off with very small poverty rates. The rates multiplied by a factor of 10 between 1989 and 2000. There are two reasons for the explosion of poverty rates in Eastern Europe and Central Asia. One is the huge increase in inequality that followed the collapse of the communist system. The second factor is the dismal aggregate growth performance of these countries. Notice, however, that the average levels of income for these countries remain far above the levels of Africa or even Asia. Therefore, even after the deterioration in mean income and the rise of income dispersion, poverty rates remain relatively low in Eastern Europe and Central Asia.

I.3 Empirical Regularities about Economic Growth

Kaldor (1963) listed a number of stylized facts that he thought typified the process of economic growth:

1. Per capita output grows over time, and its growth rate does not tend to diminish.
2. Physical capital per worker grows over time.
3. The rate of return to capital is nearly constant.
4. The ratio of physical capital to output is nearly constant.
5. The shares of labor and physical capital in national income are nearly constant.
6. The growth rate of output per worker differs substantially across countries.[8]

Fact 6 accords with the cross-country data that we have already discussed. Facts 1, 2, 4, and 5 seem to fit reasonably well with the long-term data for currently developed countries. For discussions of the stability of the long-run ratio of physical capital to GDP in Japan, Germany, Italy, the United Kingdom, and the United States, see Maddison (1982, chapter 3). For indications of the long-term stability of factor shares in the United States, see Denison (1974, appendix J) and Jorgenson, Gollop, and Fraumeni (1987, table 9.3). Young (1995) reports that factor shares were reasonably stable in four East Asian countries—Hong Kong, Singapore, South Korea, and Taiwan—from the early or middle 1960s through 1990. Studies of seven developed countries—Canada, France, Germany, Italy, Japan, the Netherlands, and the United Kingdom—indicate that factor shares are similar to those in the United States (Christensen, Cummings, and Jorgenson, 1980, and Dougherty, 1991). In some Latin-American countries considered by Elias (1990), the capital shares tend, however, to be higher than those in the United States.

Kaldor's claimed fact 3 on the stability of real rates of return appears to be heavily influenced by the experience of the United Kingdom; in this case, the real interest rate seems

8. Kuznets (1973, 1981) brings out other characteristics of modern economic growth. He notes the rapid rate of structural transformation, which includes shifts from agriculture to industry to services. This process involves urbanization, shifts from home work to employee status, and an increasing role for formal education. He also argues that modern growth involves an increased role for foreign commerce and that technological progress implies reduced reliance on natural resources. Finally, he discusses the growing importance of government: "The spread of modern economic growth placed greater emphasis on the importance and need for organization in national sovereign units. . . . The sovereign state unit was of critical importance as the formulator of the rules under which economic activity was to be carried on; as a referee . . . ; and as provider of infrastructure" (1981, p. 59).

to have no long-run trend (see Barro, 1987, figures 4 and 7). For the United States, however, the long-term data suggest a moderate decline of real interest rates (Barro, 1997, table 11.1). Real rates of return in some fast-growing countries, such as South Korea and Singapore, are much higher than those in the United States but have declined over time (Young, 1995). Thus it seems likely that Kaldor's hypothesis of a roughly stable real rate of return should be replaced by a tendency for returns to fall over some range as an economy develops.

We can use the data presented in chapter 12 to assess the long-run tendencies of the growth rate of real per capita GDP. Tables 12.10 and 12.11 contain figures from Angus Maddison for 31 countries over periods of roughly a century. These numbers basically exhaust the available information about growth over very long time intervals.

Table 12.10 applies to 16 currently developed countries, the major countries in Europe plus the United States, Canada, and Australia. These data show an average per capita growth rate of 1.9 percent per year over roughly a century, with a breakdown by 20-year periods as shown in table I.1. These numbers are consistent with Kaldor's proposition that the growth rate of real per capita GDP has no secular tendency to decline; in fact, the periods following World War II show growth rates well above the long-run average. The reduction in the growth rate from 3.7 percent per year in 1950–70 to 2.2 percent per year in 1970–90 corresponds to the often-discussed *productivity slowdown*. It is apparent from the table, however, that the growth rate for 1970–90 is high in relation to the long-term history.

Table 12.11 contains figures for 15 currently less-developed countries in Asia and Latin America. In this case, the average long-run growth rate from 1900 to 1987 is 1.4 percent per year, and the breakdown into four subperiods is as shown in table I.2. Again, the post–World War II period (here, 1950–87) shows growth rates well above the long-term average.

Table I.1
Long-Term Growth Rates for Currently Developed Countries

Period	Growth Rate (percent per year)	Number of Countries
1830–50	0.9	10
1850–70	1.2	11
1870–90	1.2	13
1890–10	1.5	14
1910–30	1.3	16
1930–50	1.4	16
1950–70	3.7	16
1970–90	2.2	16

Source: Table 12.10.
Note: The growth rates are simple averages for the countries with data.

Table I.2
Long-Term Growth Rates for Currently Less-Developed Countries

Period	Growth Rate (percent per year)	Number of Countries
1900–13	1.2	15
1913–50	0.4	15
1950–73	2.6	15
1973–87	2.4	15

Source: Table 12.11 in chapter 12.
Note: The growth rates are simple averages for the countries with data.

The information depicted in figures I.1–I.3 applies to the behavior of real per capita GDP for over 100 countries from 1960 to 2000. We can use these data to extend the set of stylized facts that was provided by Kaldor. One pattern in the cross-country data is that the growth rate of per capita GDP from 1960 to 2000 is essentially uncorrelated with the level of per capita GDP in 1960 (see chapter 12). In the terminology developed in chapter 1, we shall refer to a tendency for the poor to grow faster than the rich as β convergence. Thus the simple relationship between growth and the starting position for a broad cross section of countries does not reveal β convergence. This kind of convergence does appear if we limit attention to more homogeneous groups of economies, such as the U.S. states, regions of several European countries, and prefectures of Japan (see Barro and Sala-i-Martin, 1991, 1992a, and 1992b, and chapter 11). In these cases, the poorer places tend to grow faster than the richer ones. This behavior also appears in the cross-country data if we limit the sample to a relatively homogeneous collection of currently prosperous places, such as the OECD countries (see Baumol, 1986; DeLong, 1988).

We say in chapter 1 that *conditional* β convergence applies if the growth rate of per capita GDP is negatively related to the starting level of per capita GDP after holding fixed some other variables, such as initial levels of human capital, measures of government policies, the propensities to save and have children, and so on. The broad cross-country sample—that is, the data set that does not show β convergence in an absolute sense—clearly reveals β convergence in this conditional context (see Barro, 1991; Barro and Sala-i-Martin, 1992a; and Mankiw, Romer, and Weil, 1992). The rate of convergence is, however, only about 2 percent per year. Thus, it takes about 35 years for an economy to eliminate one-half of the gap between its initial per capita GDP and its long-run or target level of per capita GDP. (The target tends to grow over time.)

The results in chapter 12 show that a number of variables are significantly related to the growth rate of per capita GDP, once the starting level of per capita GDP is held constant. For example, growth depends positively on the initial quantity of human capital in the form of educational attainment and health, positively on maintenance of the rule of law and the

Table I.3
Ratios to GDP of Gross Domestic Investment and Gross National Saving (percent)

Period	Australia	Canada	France	India	Japan	Korea	United Kingdom	United States
				1. Gross Domestic Investment				
1870–89	16.5	16.0	12.8	—	—	—	9.3	19.8
1890–09	13.7	17.2	14.0	—	14.0	—	9.4	17.9
1910–29	17.4	19.8	—	6.4	16.6	5.1[a]	6.7	17.2
1930–49	13.3	13.1	—	8.4	20.5	—	8.1	12.7
1950–69	26.3	23.8	22.6	14.0	31.8	16.3[b]	17.2	18.9
1970–89	24.9	22.8	23.2	20.2	31.9	29.1	18.2	18.7
				2. Gross National Saving				
1870–89	11.2	9.1	12.8	—	—	—	13.9	19.1
1890–09	12.2	11.5	14.9	—	12.0	—	13.1	18.4
1910–29	13.6	16.0	—	6.4	17.1	2.38	9.6	18.9
1930–49	13.0	15.6	—	7.7	19.8	—	4.8	14.1
1950–69	24.0	22.3	22.8	12.2	32.1	5.9[b]	17.7	19.6
1970–89	22.9	22.1	23.4	19.4	33.7	26.2	19.4	18.5

Source: Maddison (1992).
[a] 1911–29
[b] 1951–69

ratio of investment to GDP, and negatively on fertility rates and the ratio of government consumption spending to GDP.

We can assess regularities in investment and saving ratios by using the long-term data in Maddison (1992). He provides long-term information for a few countries on the ratios of gross domestic investment to GDP and of gross national saving (the sum of domestic and net foreign investment) to GDP. Averages of the investment and saving ratios over 20-year intervals for the eight countries that have enough data for a long-period analysis are shown in table I.3. For an individual country, the table indicates that the time paths of domestic investment and national saving are usually similar. Domestic investment was, however, substantially higher than national saving (that is, borrowing from abroad was large) for Australia and Canada from 1870 to 1929, for Japan from 1890 to 1909, for the United Kingdom from 1930 to 1949, and for Korea from 1950 to 1969 (in fact, through the early 1980s). National saving was much higher than domestic investment (lending abroad was substantial) for the United Kingdom from 1870 to 1929 and for the United States from 1930 to 1949.

For the United States, the striking observation from the table is the stability over time of the ratios for domestic investment and national saving. The only exception is the relatively low values from 1930 to 1949, the period of the Great Depression and World War II. The United States is, however, an outlier with respect to the stability of its investment and saving ratios; the data for the other seven countries show a clear increase in these ratios over time. In particular, the ratios for 1950–89 are, in all cases, substantially greater than those from before World War II. The long-term data therefore suggest that the ratios to GDP of gross domestic investment and gross national saving tend to rise as an economy develops, at least over some range. The assumption of a constant gross saving ratio, which appears in chapter 1 in the Solow–Swan model, misses this regularity in the data.

The cross-country data also reveal some regularities with respect to fertility rates and, hence, rates of population growth. For most countries, the fertility rate tends to decline with increases in per capita GDP. For the poorest countries, however, the fertility rate may rise with per capita GDP, as Malthus (1798) predicted. Even stronger relations exist between educational attainment and fertility. Except for the most advanced countries, female schooling is negatively related with the fertility rate, whereas male schooling is positively related with the fertility rate. The net effect of these forces is that the fertility rate—and the rate of population growth—tend to fall over some range as an economy develops. The assumption of an exogenous, constant rate of population growth—another element of the Solow–Swan model—conflicts with this empirical pattern.

I.4 A Brief History of Modern Growth Theory

Classical economists, such as Adam Smith (1776), David Ricardo (1817), and Thomas Malthus (1798), and, much later, Frank Ramsey (1928), Allyn Young (1928), Frank Knight (1944), and Joseph Schumpeter (1934), provided many of the basic ingredients that appear in modern theories of economic growth. These ideas include the basic approaches of competitive behavior and equilibrium dynamics, the role of diminishing returns and its relation to the accumulation of physical and human capital, the interplay between per capita income and the growth rate of population, the effects of technological progress in the forms of increased specialization of labor and discoveries of new goods and methods of production, and the role of monopoly power as an incentive for technological advance.

Our main study begins with these building blocks already in place and focuses on the contributions in the neoclassical tradition since the late 1950s. We use the neoclassical methodology and language and rely on concepts such as aggregate capital stocks, aggregate production functions, and utility functions for representative consumers (who often have infinite horizons). We also use modern mathematical methods of dynamic optimization and differential equations. These tools, which are described in the appendix at the end of this book, are familiar today to most first-year graduate students in economics.

From a chronological viewpoint, the starting point for modern growth theory is the classic article of Ramsey (1928), a work that was several decades ahead of its time. Ramsey's treatment of household optimization over time goes far beyond its application to growth theory; it is hard now to discuss consumption theory, asset pricing, or even business-cycle theory without invoking the optimality conditions that Ramsey (and Fisher, 1930) introduced to economists. Ramsey's intertemporally separable utility function is as widely used today as the Cobb–Douglas production function. The economics profession did not, however, accept or widely use Ramsey's approach until the 1960s.

Between Ramsey and the late 1950s, Harrod (1939) and Domar (1946) attempted to integrate Keynesian analysis with elements of economic growth. They used production functions with little substitutability among the inputs to argue that the capitalist system is inherently unstable. Since they wrote during or immediately after the Great Depression, these arguments were received sympathetically by many economists. Although these contributions triggered a good deal of research at the time, very little of this analysis plays a role in today's thinking.

The next and more important contributions were those of Solow (1956) and Swan (1956). The key aspect of the Solow–Swan model is the neoclassical form of the production function, a specification that assumes constant returns to scale, diminishing returns to each input, and some positive and smooth elasticity of substitution between the inputs. This production function is combined with a constant-saving-rate rule to generate an extremely simple general-equilibrium model of the economy.

One prediction from these models, which has been exploited seriously as an empirical hypothesis only in recent years, is conditional convergence. The lower the starting level of per capita GDP, relative to the long-run or steady-state position, the faster the growth rate. This property derives from the assumption of diminishing returns to capital; economies that have less capital per worker (relative to their long-run capital per worker) tend to have higher rates of return and higher growth rates. The convergence is conditional because the steady-state levels of capital and output per worker depend, in the Solow–Swan model, on the saving rate, the growth rate of population, and the position of the production function—characteristics that might vary across economies. Recent empirical studies indicate that we should include additional sources of cross-country variation, especially differences in government policies and in initial stocks of human capital. The key point, however, is that the concept of conditional convergence—a basic property of the Solow–Swan model—has considerable explanatory power for economic growth across countries and regions.

Another prediction of the Solow–Swan model is that, in the absence of continuing improvements in technology, per capita growth must eventually cease. This prediction, which resembles those of Malthus and Ricardo, also comes from the assumption of diminishing returns to capital. We have already observed, however, that positive rates of per capita growth can persist over a century or more and that these growth rates have no clear tendency to decline.

The neoclassical growth theorists of the late 1950s and 1960s recognized this modeling deficiency and usually patched it up by assuming that technological progress occurred in an exogenous manner. This device can reconcile the theory with a positive, possibly constant per capita growth rate in the long run, while retaining the prediction of conditional convergence. The obvious shortcoming, however, is that the long-run per capita growth rate is determined entirely by an element—the rate of technological progress—that is outside of the model. (The long-run growth rate of the level of output also depends on the growth rate of population, another element that is exogenous in the standard theory.) Thus we end up with a model of growth that explains everything but long-run growth, an obviously unsatisfactory situation.

Cass (1965) and Koopmans (1965) brought Ramsey's analysis of consumer optimization back into the neoclassical growth model and thereby provided for an endogenous determination of the saving rate. This extension allows for richer transitional dynamics but tends to preserve the hypothesis of conditional convergence. The endogeneity of saving also does not eliminate the dependence of the long-run per capita growth rate on exogenous technological progress.

The equilibrium of the Cass–Koopmans version of the neoclassical growth model can be supported by a decentralized, competitive framework in which the productive factors, labor and capital, are paid their marginal products. Total income then exhausts the total product because of the assumption that the production function features constant returns to scale. Moreover, the decentralized outcomes are Pareto optimal.

The inclusion of a theory of technological change in the neoclassical framework is difficult, because the standard competitive assumptions cannot be maintained. Technological advance involves the creation of new ideas, which are partially nonrival and therefore have aspects of public goods. For a given technology—that is, for a given state of knowledge—it is reasonable to assume constant returns to scale in the standard, rival factors of production, such as labor, capital, and land. In other words, given the level of knowledge on how to produce, one would think that it is possible to replicate a firm with the same amount of labor, capital, and land and obtain twice as much output. But then, the returns to scale tend to be increasing if the nonrival ideas are included as factors of production. These increasing returns conflict with perfect competition. In particular, the compensation of nonrival old ideas in accordance with their current marginal cost of production—zero—will not provide the appropriate reward for the research effort that underlies the creation of new ideas.

Arrow (1962) and Sheshinski (1967) constructed models in which ideas were unintended by-products of production or investment, a mechanism described as learning by doing. In these models, each person's discoveries immediately spill over to the entire economy, an instantaneous diffusion process that might be technically feasible because knowledge is nonrival. Romer (1986) showed later that the competitive framework can be retained in this case to determine an equilibrium rate of technological advance, but the resulting growth rate would typically not be Pareto optimal. More generally, the competitive framework breaks down if discoveries depend in part on purposive R&D effort and if an individual's innovations spread only gradually to other producers. In this realistic setting, a decentralized theory of technological progress requires basic changes in the neoclassical growth model to incorporate an analysis of imperfect competition.[9] These additions to the theory did not come until Romer's (1987, 1990) research in the late 1980s.

The work of Cass (1965) and Koopmans (1965) completed the basic neoclassical growth model.[10] Thereafter, growth theory became excessively technical and steadily lost contact with empirical applications. In contrast, development economists, who are required to give advice to sick countries, retained an applied perspective and tended to use models that were technically unsophisticated but empirically useful. The fields of economic development and economic growth drifted apart, and the two areas became almost completely separated.

Probably because of its lack of empirical relevance, growth theory effectively died as an active research field by the early 1970s, on the eve of the rational-expectations revolution and the oil shocks. For about 15 years, macroeconomic research focused on short-term fluctuations. Major contributions included the incorporation of rational expectations into business-cycle models, improved approaches to policy evaluation, and the application of general-equilibrium methods to real business-cycle theory.

After the mid-1980s, research on economic growth experienced a boom, beginning with the work of Romer (1986) and Lucas (1988). The motivation for this research was the observation (or recollection) that the determinants of long-run economic growth are crucial issues, far more important than the mechanics of business cycles or the countercyclical effects of monetary and fiscal policies. But a recognition of the significance of long-run growth was only a first step; to go further, one had to escape the straitjacket of the neoclassical growth model, in which the long-term per capita growth rate was pegged by the rate of exogenous technological progress. Thus, in one way or another, the recent contributions determine the long-run growth rate within the model; hence, the designation *endogenous-growth* models.

The initial wave of the new research—Romer (1986), Lucas (1988), Rebelo (1991)—built on the work of Arrow (1962), Sheshinski (1967), and Uzawa (1965) and did not really introduce a theory of technological change. In these models, growth may go on indefinitely because the returns to investment in a broad class of capital goods—which includes human

9. Another approach is to assume that all of the nonrival research—a classic public good—is financed by the government through involuntary taxes; see Shell (1967).

10. However, recent research has shown how to extend the neoclassical growth model to allow for heterogeneity among households (Caselli and Ventura, 2000) and to incorporate time-inconsistent preferences (Barro, 1999).

capital—do not necessarily diminish as economies develop. (This idea goes back to Knight, 1944.) Spillovers of knowledge across producers and external benefits from human capital are parts of this process, but only because they help to avoid the tendency for diminishing returns to the accumulation of capital.

The incorporation of R&D theories and imperfect competition into the growth framework began with Romer (1987, 1990) and included significant contributions by Aghion and Howitt (1992) and Grossman and Helpman (1991, chapters 3 and 4). In these models, technological advance results from purposive R&D activity, and this activity is rewarded by some form of ex post monopoly power. If there is no tendency for the economy to run out of ideas, the growth rate can remain positive in the long run. The rate of growth and the underlying amount of inventive activity tend, however, not to be Pareto optimal because of distortions related to the creation of the new goods and methods of production. In these frameworks, the long-term growth rate depends on governmental actions, such as taxation, maintenance of law and order, provision of infrastructure services, protection of intellectual property rights, and regulations of international trade, financial markets, and other aspects of the economy. The government therefore has great potential for good or ill through its influence on the long-term rate of growth. This research program remained active through the 1990s and has been applied, for example, to understanding scale effects in the growth process (Jones, 1999), analyzing whether technological progress will be labor or capital augmenting (Acemoglu, 2002), and assessing the role of competition in the growth process (Aghion et al., 2001, 2002).

The new research also includes models of the diffusion of technology. Whereas the analysis of discovery relates to the rate of technological progress in leading-edge economies, the study of diffusion pertains to the manner in which follower economies share by imitation in these advances. Since imitation tends to be cheaper than innovation, the diffusion models predict a form of conditional convergence that resembles the predictions of the neoclassical growth model. Some recent empirical work has verified the importance of technological diffusion in the convergence process.

Another key exogenous parameter in the neoclassical growth model is the growth rate of population. A higher rate of population growth lowers the steady-state level of capital and output per worker and tends thereby to reduce the per capita growth rate for a given initial level of per capita output. The standard model does not, however, consider the effects of per capita income and wage rates on population growth—the kinds of effects stressed by Malthus—and also does not take account of the resources used up in the process of child rearing. Another line of recent research makes population growth endogenous by incorporating an analysis of fertility choice into the neoclassical model. The results are consistent, for example, with the empirical regularity that fertility rates tend to fall with per capita income over the main range of experience but may rise with per capita income for the poorest countries. Additional work related to the endogeneity of labor supply in a growth context concerns migration and labor/leisure choice.

The clearest distinction between the growth theory of the 1960s and that of the 1990s is that the recent research pays close attention to empirical implications and to the relation between theory and data. However, much of this applied perspective involved applications of empirical hypotheses from the older theory, notably the neoclassical growth model's prediction of conditional convergence. The cross-country regressions motivated by the neoclassical model surely became a fixture of research in the 1990s. An interesting recent development in this area, which we explore in chapter 12, involves assessment of the robustness of these kinds of estimates. Other empirical analyses apply more directly to the recent theories of endogenous growth, including the roles of increasing returns, R&D activity, human capital, and the diffusion of technology.

I.5 Some Highlights of the Second Edition

This second edition of *Economic Growth* includes changes throughout the book. We mention here a few of the highlights. In this introduction we already described new estimates of the distribution of income of individuals throughout the world from 1970 to 2000.

Chapter 1 has been made easier and more accessible. We added a section on markets in the Solow–Swan model. We also discussed the nature of the theoretical dissatisfaction with neoclassical theory that led to the emergence of endogenous growth models with imperfect competition.

Chapter 2 expands the treatment of the basic neoclassical growth model to allow for heterogeneity of households. There is an improved approach to ruling out "undersaving" paths and for deriving and using transversality conditions. We also include an analysis of models with nonconstant time-preference rates.

Chapter 3 has various extensions to the basic neoclassical growth model, including an expanded treatment of the government sector. The framework allows for various forms of tax rates and allows for a clear distinction between taxes on capital income and taxes on labor or consumption.

Chapters 6 and 7 discuss models of endogenous technological progress. The new material includes an analysis of the role and source of scale effects in these models. We refer in chapter 6 to Thomas Jefferson's mostly negative views on patents as a mechanism for motivating inventions. Chapter 7 has an improved analysis of models where technological advances take the form of quality improvements. We have particularly improved the treatment of the interplay between industry leaders and outsiders and, hence, of the role of outside competition in the growth process.

Chapter 8 has a model of technological diffusion. The basic model is improved, and the theoretical results are related to recent empirical findings.

Chapter 9 has an extended treatment of endogenous population growth. Chapter 10 has an improved analysis of growth accounting, including its relation to theories of endogenous technological progress. Chapter 11, which deals with regional data sets, extends the analysis of U.S. states through 2000.

In chapter 12 we include an updated treatment of cross-country growth regressions, using the new Summers–Heston data set, Penn World Tables version 6.1, which has data through 2000 (see Heston, Summers, and Aten, 2002). We also discuss in this chapter various issues about the reliability of estimates from cross-country regressions, including ways to assess the robustness of the results.