BARRY P. BOSWORTH
Brookings Institution

SUSAN M. COLLINS
Brookings Institution and Georgetown University

The Empirics of Growth: An Update

THE PAST DECADE HAS seen an explosion of empirical research on economic growth and its determinants, yet many of the central issues of interest remain unresolved. For instance, no consensus has emerged about the relative contributions of capital accumulation and improvements in total factor productivity in accounting for differences in growth across countries and time. Nor is there agreement about the role of increased education or the importance of economic policy. Indeed, results from the many studies on a given issue frequently reach opposite conclusions. And two of the main empirical approaches—growth accounting and growth regressions—have themselves come under attack, with some researchers going so far as to label them as irrelevant to policymaking.

In this paper we argue that, properly implemented and interpreted, both growth accounts and growth regressions are valuable tools, which can improve—and have improved—our understanding of growth experiences across countries. We also show that careful attention to issues of measurement and consistency goes a long way in explaining the apparent contradictions among findings in the literature. Our analysis combines growth accounts and growth regressions with a focus on measurement and procedural consistency to address the issues raised. The growth accounts are constructed for eighty-four countries that together represent 95 percent of gross world product and 84 percent of world population, over a period of forty years from 1960 to 2000. Appendix A lists the

The research for this paper was financed in part by a grant from PRMEP-Growth of the World Bank and the Tokyo Club Foundation for Global Studies. We would like to thank participants at seminars at the World Bank and the London School of Economics for comments. We are very indebted to Kristin Wilson, who prepared the data and performed the statistical analysis.

countries in the sample by region.[1] This large data set also enables us to compare growth experiences across two twenty-year time periods: 1960–80 and 1980–2000.

Understanding the characteristics and determinants of economic growth requires an empirical framework that can be applied to large groups of countries over a relatively long period. Growth accounts and growth regressions provide such frameworks in a way that is particularly informative because the two approaches can be used in concert, enabling researchers to explore the channels (factor accumulation versus increased factor productivity) through which various determinants influence growth. Although the information thus provided is perhaps best considered descriptive, it can generate important insights that complement those gained from in-depth case studies of selected countries, or from estimation of carefully specified econometric models designed to test specific hypotheses.

Growth accounts provide a means of allocating observed output growth between the contributions of changes in factor inputs and a residual, total factor productivity (TFP), which measures a combination of changes in efficiency in the use of those inputs and changes in technology. These accounts are used extensively within the industrial countries to evaluate the sources of change in productivity growth, the role of information technology, and differences in the experience of individual countries.[2] In his recent, comprehensive assessment, Charles Hulten aptly describes the approach as "a simple and internally consistent intellectual framework for organizing data. . . . For all its flaws, real and imagined, many researchers have used it to gain valuable insights into the process of economic growth."[3]

Despite its extensive use within the industrial countries, growth accounting has done surprisingly little to resolve some of the most funda-

1. Our sample covers all world regions and includes all countries with population in excess of 1 million for which we have national accounts spanning the last forty years. The largest groups of excluded countries are those of Eastern Europe and the former Soviet Union. The share of world GDP is the 95 percent share measured between 1995 and 2000 using market exchange rates, and population data are from World Development Indicators.

2. Recent examples of growth accounting analyses for industrial countries are Oliner and Sichel (2000), Jorgensen (2001), and Organization for Economic Cooperation and Development (2003).

3. Hulten (2001, p. 63).

mental issues under debate in the development literature. For example, the major objective of growth accounting is to distinguish the contribution of increased capital per worker from that of improvements in factor productivity. Yet one can observe widely divergent views on this issue, with some researchers claiming that capital accumulation is an unimportant part of the growth process and others that it is the fundamental determinant of growth.

Criticism of growth accounting has been concentrated in three areas. The first focuses on the fact that TFP is measured as a residual. As discussed in detail by Hulten, this residual provides a measure of gains in economic efficiency (the quantity of output that can be produced with a given quantity of inputs), which can be thought of as shifts in the production function. But such shifts reflect myriad determinants, in addition to technological innovation, that influence growth but that the measured increases in factor inputs do not account for. Examples include the implications of sustained political turmoil, external shocks, changes in government policies, institutional changes, and measurement error. Therefore this residual should *not* be taken as an indicator of technical change.

A second concern focuses on whether the growth decomposition is sensitive to underlying assumptions about the nature of the production process and to the indicators chosen to measure changes in output and inputs. In principle, growth accounts can be constructed to yield estimates of TFP that are independent of the functional form and the parameters of the production process. This requires assuming both a sufficient degree of competition so that factor earnings are proportionate to factor productivities, and the availability of accurate data on factor shares of income. In practice, data limitations require the approximation of fixed factor income shares, which is consistent with a more limited set of production functions. But given that these factor shares (appropriately adjusted for self-employment) do not appear to differ systematically across countries, this approach seems quite reasonable. We address some of the key measurement concerns in detail later in the paper.

Finally, an accounting decomposition cannot (and is not intended to) determine the fundamental *causes* of growth. Consider a country that has had rapid increases in both accumulation of capital per worker and factor productivity. Growth accounting does not provide a means to identify whether the productivity growth caused the capital accumulation (for example, by increasing the expected returns to investment), or whether the capital accumulation made additional innovations possible. Growth accounting is a framework for examining the *proximate sources* of growth. And the application of a consistent and transparent procedure across a wide range of countries, combined with robustness checks, generates useful benchmarks that facilitate broad cross-country comparisons of economic performance.

Growth regressions have also come under considerable fire. A great many researchers have regressed various indicators of output growth on a vast array of potential determinants. But recent summaries of this literature have called the usefulness of this approach into question, largely because the resulting parameter estimates are claimed to be unstable.[4] We will argue that this critique has gone too far. In fact, most of the variability in the results can be explained by variation in the sample of countries, the time period, and the additional explanatory variables included in the regression. We maintain that there is a core set of explanatory variables that has been shown to be consistently related to economic growth and that the importance of other variables should be examined conditional on inclusion of this core set. Thus, in implementing our growth regressions, we restrict our attention to estimations based on a common sample of countries, a common time period, and a common set of conditioning variables.

A second concern with growth regressions is that many of the explanatory variables of interest are likely to be endogenous. We note that our conditioning variables include initial conditions that can be considered predetermined for an individual country. However, for other variables, including institutional quality, openness to trade, and especially policy measures, the concern about endogeneity is certainly valid. Recent work has identified certain variables that can be used, in instrumental variables regressions, as instruments for institutional quality and for trade share-based indicators of openness, and we use these in our analysis. However, no effective instruments are available for the key macroeconomic policy variables. In this context we interpret the regression results descriptively.

We also limit our discussion to a consideration of the variations in income *growth* over the past forty years. Although analysis of the sources of international differences in income *levels* (sometimes called develop-

4. For example, see Levine and Renelt (1992), Durlauf and Quah (1999), and Lindauer and Pritchett (2002).

ment accounting) has become increasingly popular in recent years, we believe it paints too pessimistic a picture of the opportunities that countries have to improve economic performance. In a levels formulation it is difficult to define a set of exogenous initial conditions beyond geography and perhaps colonial governance, and the income differences seem largely predetermined by events in the distant past.

Furthermore, the differences between analysis of levels and analysis of rates of growth are less than they seem. The level of income in 2000 can be viewed as a simple combination of the level of income in an earlier year (say, 1960) and the change since then. Given the importance of convergence issues (that is, whether incomes of developing countries are converging toward those of the industrial countries), nearly all empirical studies of growth include the initial level of income as a conditioning variable. Thus, at the empirical level, the two approaches differ primarily in that the growth studies treat developments up to the initial year as predetermined and do not attempt to explain the earlier history. In our data set, 30 percent of the cross-country variance in income per capita at purchasing power parity (international prices) in 2000 is attributed to events since 1960, and 70 percent to the preceding millennium.

We begin by explaining our construction of a consistent set of growth accounts covering most of the global economy. We then use growth accounts and growth regressions to examine three issues: the relative importance of capital accumulation and TFP in raising income per capita; the significance for economic growth of improvements in the quantity and quality of education (a factor emphasized by the international aid organizations, among others); and the sources of the sharp differences in growth performance before and after 1980.

Construction of the Accounts

Growth accounts have long provided a conceptual structure for analyzing growth in the industrial countries. However, it is only in the last decade that the development of multicountry data sets has made it possible to extend the analysis to a large number of developing countries. Among the more important data sets are the World Development Indicators (WDI) of the World Bank, the Penn World Tables (PWT) produced at the University of Pennsylvania, population and labor force data compiled by the United Nations, and measures of educational attainment compiled by Robert Barro and Jong-Wha Lee. We have relied primarily on data from the WDI for developing countries and from the Organization for Economic Cooperation and Development (OECD) for industrial countries. We have also been able to compare the basic information for some of the early years with the national income accounts file underlying version 6 of the PWT, allowing us to construct consistent measures of GDP and investment in national prices for eighty-four countries over the period 1960–2000.

Growth accounts are consistent with a wide range of alternative formulations of the relationship between factor inputs and output. It is necessary only to assume a degree of competition sufficient to ensure that the earnings of each factor are proportionate to its productivity. The shares of income paid to the factors can then be used to measure their importance in the production process. Unfortunately, consistent measures of factor income are unavailable for individual countries, compelling us to use fixed income-share weights to construct the indexes. In those countries where factor shares can be measured appropriately, labor shares are considerably more similar across countries (and over time) than conventional measures imply, suggesting that this simplification does not raise serious problems.[5] Although the assumption that income-share weights are fixed over time is consistent only with a more limited set of production functions, it is consistent with the data available for the OECD countries.

In this exercise we assume a constant-returns-to-scale production function of the form

$$Y = AK^{\alpha}(LH)^{1-\alpha}. \quad (1)$$

5. Most of the debate has been over the magnitude of the capital share. Cross-country variations in this share can be traced largely to differences in the importance of the self-employed, whose earnings are assigned to property income in the national accounts. After adjusting for the labor component of the earnings of the self-employed, Englander and Gurney (1994) found that income shares in OECD countries were relatively stable and largely free of trend but that there were significant cyclical variations. Gollin (2002) concludes that the adjusted measures of factor shares are roughly similar across a broad range of industrial and developing countries. He finds no systematic differences between rich and poor countries. In contrast, Harrison (2003) argues that labor shares do vary over time in most countries, but she is unable to differentiate between the capital and labor income of the self-employed.

The capital share, α, is assumed equal to 0.35 for the entire sample. *H* is a measure of educational attainment, or human capital, used to adjust the work force (that is, the labor input, *L*) for quality change. We report our results in a form that decomposes growth in output per worker $\Delta\ln(Y/L)$ into the contributions of increases in capital per worker $\Delta\ln(K/L)$, increases in education per worker $\Delta\ln(H)$, and improvements in TFP $\Delta\ln(A)$:

$$(2) \qquad \Delta\ln(Y/L) = \alpha[\Delta\ln(K/L)] + (1-\alpha)\Delta\ln H + \Delta\ln A.$$

Much of the controversy over the relative contributions to growth from increases in factor inputs and from changes in TFP results from differences in the measures of capital and labor inputs. We discuss these measures briefly here and in more detail in the following two sections.

We assume that growth in capital services is proportional to the capital stock, which we estimate with a perpetual inventory model:

$$(3) \qquad K_t = K_{t-1}(1-d) + I_t,$$

where the depreciation rate, *d,* equals 0.05, and *I* is gross fixed investment. The basic investment data are taken from a World Bank study that incorporated information extending back as far as 1950.[6]

Our measure of labor input is based on labor force data from the International Labour Organization. The use of labor force instead of population data implies that our measure reflects variations in the proportion of the population that is of working age, and in age- and sex-specific labor force participation rates. (However, for many countries participation rates are interpolated between census years.) Comprehensive measures of unemployment rates and annual hours of work are unavailable. We also allow for differences in educational attainment by relating human capital to average years of schooling, *s,* assuming a 7 percent return to each year:[7]

6. Nehru and Dhareshwar (1993). We adjusted their estimates for revisions in the investment series after 1960 and a higher rate of depreciation, and we extended the series to 2000.

7. Estimated returns to schooling average 7 percent in high-income countries but 10 percent in Latin America and Asia and 13 percent in Africa. (See the summary in Bils and Klenow, 2000.) Our earlier work also explored the implications of assuming a 12 percent rate of return.

(4) $$H = (1.07)^s.$$

Table 1 and figure 1 present the results of our growth accounting decomposition for seven major world regions.[8] Table 1 reports the results for each decade and over the entire period, distinguishing the contribution of physical from that of human capital.[9] Figure 1 shows how growth patterns have evolved over time. Each panel shows, for one region, the contribution of increased (physical and human) capital per worker to growth in output per worker, the contribution of changes in TFP, and output per worker, which is the product of the two. We note that growth accounting is not intended for analysis of short-term fluctuations but rather for analysis of economic growth over the longer term. Not surprisingly, capital's contribution exhibits a relatively smooth trend over time, with most of the year-to-year fluctuations in output per worker reflected in TFP.

Consider first the total of all the countries in our sample. As table 1 shows, over the entire period (1960–2000) world output grew, on average, by 4 percent a year, while output per worker grew by 2.3 percent a year. The table also shows that increases in physical capital per worker and improvements in TFP each contributed roughly 1 percentage point a year to growth, while increased human capital (education) added about 0.3 percentage point a year.

East Asia (excluding China) has been the fastest-growing region, with output per worker increasing by 3.9 percent a year over 1960–2000. (In this comparison China is treated separately because of its dominant size, phenomenal growth since 1980, and questions about the accuracy of reported measures of its GDP growth.[10]) But TFP among these countries grew barely more rapidly than the overall world average over that period.[11] A now-common finding in growth accounting studies is that East Asia's rapid growth does not appear to have been a story in which

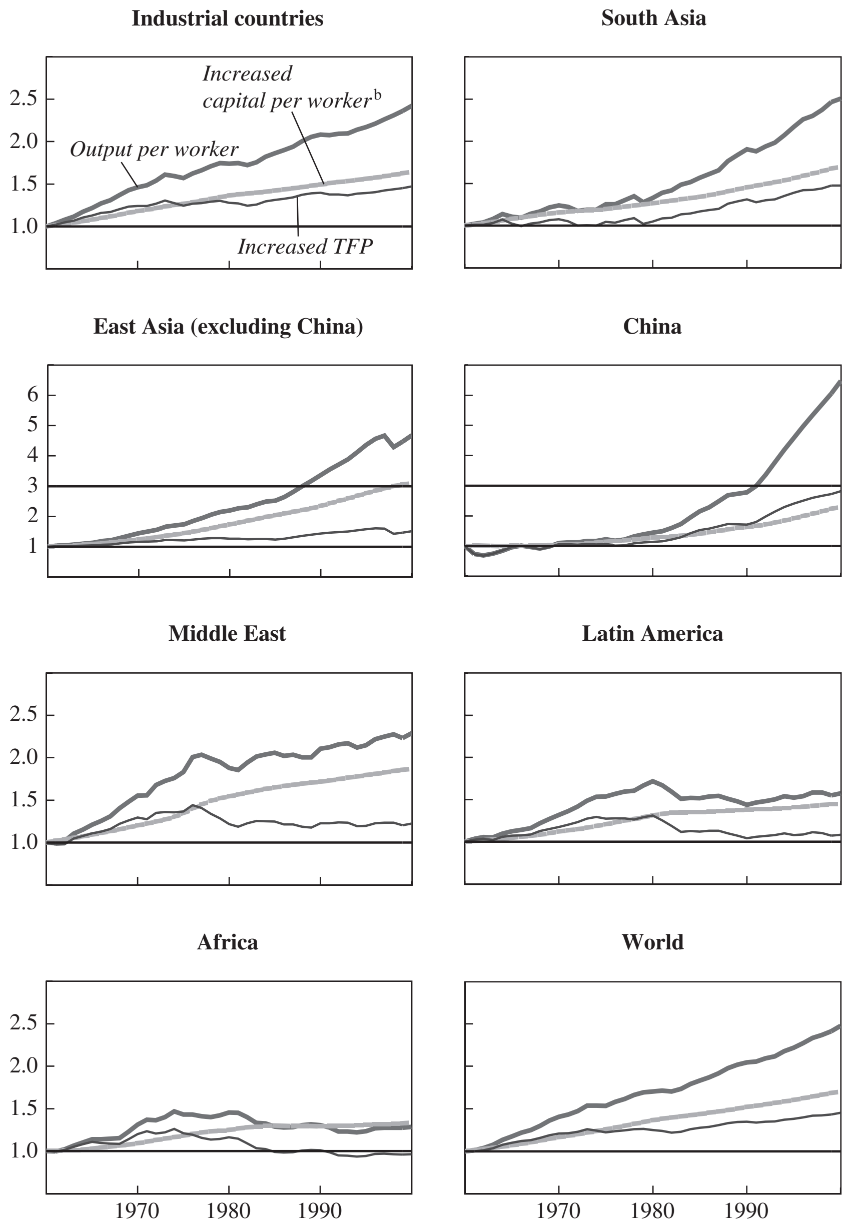
8. GDP weights have been used to construct the regional averages. The weights are averages of GDP over 1960–2000 using the 1996 purchasing-power-parity exchange rates of version 6 of the PWT. Regional weights in the "world" are as follows: industrial countries, 0.67; Latin America, 0.10; East Asia, 0.05; China, 0.06; South Asia, 0.07; Africa, 0.03; and Middle East, 0.03.

9. Results for individual countries are available from the authors.

10. Our measures are based on the WDI data, but several researchers have argued that China's growth rate is overstated in those data. See, for example, Heston (2001), Wu (2002), and Young (2000).

11. Alwyn Young's (1994) careful analysis was one of the first to document this point.

Figure 1. Output per Worker and Its Components, by Region, 1960–2000[a]



Source: Authors' calculations as detailed in the text.
a. See appendix A for country groupings.
b. Share of growth directly attributable to growth in the capital stock.

Table 1. Sources of Growth by Region and Period, 1960–2000[a]

			Contribution by component (percentage points)		
Region and period	*Growth in output (percent a year)*	*Growth in output per worker (percent a year)*	*Physical capital per worker[b]*	*Education per worker[c]*	*Total factor productivity[d]*
World (84 countries)					
1960–70	5.1	3.5	1.2	0.3	1.9
1970–80	3.9	1.9	1.1	0.5	0.3
1980–90	3.5	1.8	0.8	0.3	0.8
1990–2000	3.3	1.9	0.9	0.3	0.8
1960–2000	4.0	2.3	1.0	0.3	0.9
Industrial countries (22)					
1960–70	5.2	3.9	1.3	0.3	2.2
1970–80	3.3	1.7	0.9	0.5	0.3
1980–90	2.9	1.8	0.7	0.2	0.9
1990–2000	2.5	1.5	0.8	0.2	0.5
1960–2000	3.5	2.2	0.9	0.3	1.0
China					
1960–70	2.8	0.9	0.0	0.3	0.5
1970–80	5.3	2.8	1.6	0.4	0.7
1980–90	9.2	6.8	2.1	0.4	4.2
1990–2000	10.1	8.8	3.2	0.3	5.1
1960–2000	6.8	4.8	1.7	0.4	2.6
East Asia except China (7)					
1960–70	6.4	3.7	1.7	0.4	1.5
1970–80	7.6	4.3	2.7	0.6	0.9
1980–90	7.2	4.4	2.4	0.6	1.3
1990–2000	5.7	3.4	2.3	0.5	0.5
1960–2000	6.7	3.9	2.3	0.5	1.0

countries achieved strong gains in TFP by adopting existing technologies and catching up to the efficiency frontier. Instead, the region's rapid growth is associated in part with an above-average contribution from gains in human capital and, most important, with large and sustained increases in physical capital. The contribution of increased physical capital per worker is more than twice the global average. In contrast, the industrial countries as a group enjoyed very rapid TFP growth before 1970, consistent with their "catching up" to the United States.

Sub-Saharan Africa was the slowest-growing region over 1960–2000 as a whole, with output per worker rising just 0.6 percent a year. Increased

Table 1. Sources of Growth by Region and Period, 1960–2000[a] (*continued*)

			Contribution by component (percentage points)		
Region and period	*Growth in output (percent a year)*	*Growth in output per worker (percent a year)*	*Physical capital per worker*[b]	*Education per worker*[c]	*Total factor productivity*[d]
Latin America (22)					
1960–70	5.5	2.8	0.8	0.3	1.6
1970–80	6.0	2.7	1.2	0.3	1.1
1980–90	1.1	–1.8	0.0	0.5	–2.3
1990–2000	3.3	0.9	0.2	0.3	0.4
1960–2000	4.0	1.1	0.6	0.4	0.2
South Asia (4)					
1960–70	4.2	2.2	1.2	0.3	0.7
1970–80	3.0	0.7	0.6	0.3	–0.2
1980–90	5.8	3.7	1.0	0.4	2.2
1990–2000	5.3	2.8	1.2	0.4	1.2
1960–2000	4.6	2.3	1.0	0.3	1.0
Africa (19)					
1960–70	5.2	2.8	0.7	0.2	1.9
1970–80	3.6	1.0	1.3	0.1	–0.3
1980–90	1.7	–1.1	–0.1	0.4	–1.4
1990–2000	2.3	–0.2	–0.1	0.4	–0.5
1960–2000	3.2	0.6	0.5	0.3	–0.1
Middle East (9)					
1960–70	6.4	4.5	1.5	0.3	2.6
1970–80	4.4	1.9	2.1	0.5	–0.6
1980–90	4.0	1.1	0.6	0.5	0.1
1990–2000	3.6	0.8	0.3	0.5	0.0
1960–2000	4.6	2.1	1.1	0.4	0.5

Source: Authors' calculations.

a. Regional averages are aggregated with purchasing-power-parity GDP weights.

b. Growth rate of physical capital per worker multiplied by capital's productions share (0.35).

c. Growth rate of the labor quality index multiplied by labor's production share (0.65).

d. Difference between the growth rate of output per worker and the summed contributions of physical capital per worker and education per worker.

capital per worker contributed only 0.5 percentage point a year to growth in output per worker, half the global average. Modest increases in education before 1980 implied a smaller contribution from increased human capital than for most other nonindustrial regions. But the primary culprit in Africa's slow growth is TFP, which declined in every decade after 1970.

Capital versus TFP

The summary of the growth accounts in table 1 highlights the fact that *both* capital accumulation and TFP growth have made important contributions to growth in output per worker. At the global level we find that their contributions are roughly equal, but there have been substantial variations in their relative importance across regions and time.

The relative importance of capital accumulation and higher TFP as sources of growth has long been a subject of contention, dating back to the famous debate between Edward Denison, on one side, and Zvi Griliches and Dale Jorgenson, on the other.[12] It has reemerged with surprising vehemence, however, in the development literature.[13] The neoclassical growth model of Robert Solow highlights the importance of technological change as the primary determinant of long-run, steady-state growth. However, by assuming that everyone has access to the same technology, the model also assigns a large role to the accumulation of physical and human capital for countries that are in a transitional or catch-up phase. In contrast, endogenous growth theories often incorporate a role for physical and human capital in determining steady-state growth and argue that differences in technology contribute to variations in the speed of convergence.

Empirical studies reach surprisingly different conclusions about the role of capital accumulation versus that of TFP. Representing the neoclassical perspective, Gregory Mankiw, David Romer, and David Weil find that differences in physical and human capital account for roughly 80 percent of the observed international variation in income per capita.[14] In contrast, Peter Klenow and Andrés Rodríguez-Clare argue in favor of a more substantial role for differences in technological efficiency, claiming that TFP accounts for 90 percent of the cross-country variation in growth rates.[15] Particularly sharp rejections of the importance of capital accumulation are provided by William Easterly and Ross Levine.[16]

12. See the series of articles and replies in the May 1972 *Survey of Current Business.*

13. The dispute over the relative importance for output growth of increases in capital per worker and improvements in TFP is discussed in the survey by Temple (1999, especially pp. 134–41). For a perspective that emphasizes the role of TFP, see Easterly and Levine (2001).

14. Mankiw, Romer, and Weil (1992).

15. Klenow and Rodríguez-Clare (1997).

16. Easterly and Levine (2001); Easterly (2001).

Why are the empirical results so divergent? In large part the differences reflect three basic measurement issues. First, some researchers rely on the share of investment in GDP to proxy changes in the capital stock, whereas others construct a direct measure of the capital stock. Second, some value investment in terms of domestic prices, whereas others use an international price measure. Finally, some measure the contribution of capital by the change in the capital-output ratio, instead of by the change in the capital-labor ratio. We discuss each of these issues in turn.

The Investment Rate Versus the Capital Stock

The choice between the investment rate and the change in a constructed measure of the capital stock has surprisingly important implications for empirical analysis. Several growth accounting studies, including that of Mankiw, Romer, and Weil, use a formulation of the production relationship that replaces the growth in the capital stock in our equation 2 with an approximation based on its steady-state relationship with investment as a share of GDP. The change in the capital stock is given by

$$\Delta K = I - \mathrm{d}K. \tag{5}$$

Dividing through by K and assuming a steady-state constant value (γ) for the inverse of the capital-output ratio allows the rate of change of the capital stock (K) to be measured by the investment rate ($i = I/Y$):

$$\Delta \ln \mathrm{K} = i\gamma - d. \tag{6}$$

A production relationship such as that given by equation 2 can be rewritten to replace $\Delta\ln(K/L)$ with the steady-state approximation in equation 6, yielding the formulation used in many past cross-country growth studies:

$$\Delta \ln(Y/L) = \alpha(\gamma i - d) + (1-\alpha)\Delta \ln(H) + \Delta \ln(A). \tag{2$'$}$$

The use of the investment rate has an obvious advantage. It avoids the measurement problems introduced by the choice of an initial capital stock and an assumed rate of depreciation. However, the assumption of a constant capital-output ratio seems particularly unreasonable for studying the growth experiences of a highly diverse group of countries, many of which seem very far from steady-state conditions. It also seems unreasonable to assume the same capital-output ratio across a sample of countries at very different stages of development.

Strikingly, figure 2 shows that there is very little correlation between the change in the capital stock and the mean investment rate in our sample, even over a period as long as forty years. (The R^2 for the bivariate regression is just 0.08.) Some of the newly industrializing economies of Asia stand out with very high capital stock growth rates, but, because their output growth has been equally rapid, they do not have unusually large shares of output devoted to investment. In contrast, Guyana and especially Zambia—two countries with very slow output growth—are conspicuous for the small changes in their capital stocks despite high average investment shares.

It is also easy to show that the change in the capital stock—not the investment rate—is the better measure of the contribution of capital to output growth. The regressions reported in table 2 confirm that changes in the capital stock explain far more of the growth in output per worker over the full forty-year period: the R^2 equals 0.67 in the regression that includes the capital stock, but just 0.26 in the regression using the investment rate. The same is true for both twenty-year subperiods.[17] Indeed, this basic result is robust to all the specifications we tried, including those discussed later in the paper, which control for additional right-hand-side variables.

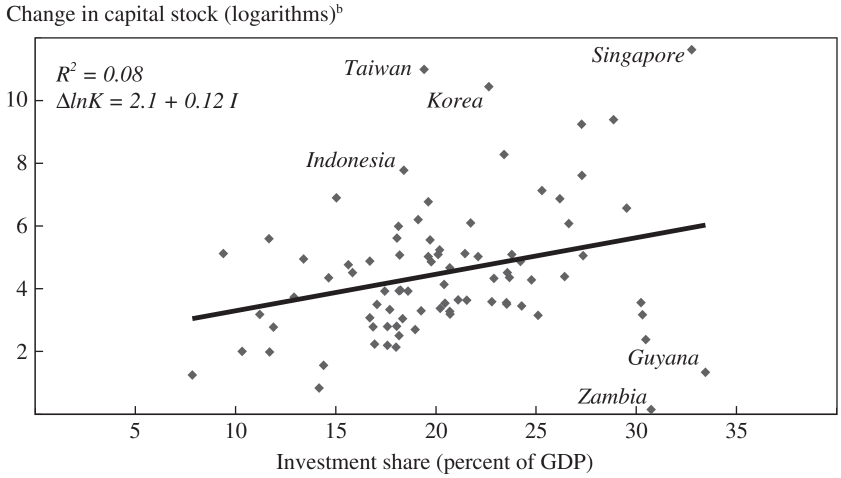
Purchasing Power Parity

The second source of variation in the empirical findings arises from the use of international price data from the PWT in some studies and data in national currency values in others.[18] International prices are strongly preferred over national prices converted at commercial exchange rates in cross-country comparisons of measures of standards of living (such as GDP per capita). However, the choice is much less clear for other comparisons, particularly those involving the composition of aggregate demand. The PWT uses three separate purchasing power parities (PPP) to

17. In these regressions all variables are scaled by the change in the labor force. The stronger correlation between output growth and the investment rate in 1980–2000 is consistent with the finding of a stronger correlation between investment and capital accumulation in the later decades.

18. We have made this point in previous work (Bosworth, Collins, and Chen, 1996). Related issues have recently been explored in Hsieh and Klenow (2003).

Figure 2. Comparison of Investment Share and Change in the Capital Stock, 1960–2000[a]



Source: Authors' calculations using data from the *OECD Statistical Compendium* and WDI.
a. Data are for the eighty-four countries listed in appendix A at national prices.
b. Capital stock is constructed as explained in the text.

construct its international price measures of investment, consumption, and government expenditure. Thus, in the process of converting to international prices, the PWT dramatically alters the expenditure shares of GDP in each country.[19]

In particular, conversion to international prices results in a new and very different measure of the investment share of GDP. Note that average international prices are dominated by the experience in the large industrial countries, where labor is relatively expensive and capital relatively cheap. Because investment heavily reflects the prices of capital goods, investment shares for low-income countries are much smaller when measured in international prices than when measured in national prices. The opposite is true for the share of GDP devoted to government expenditure, which typically has a large labor component. As a result, conversion to international prices induces a large and systematic change in investment

19. Other published measures of PPP often report a single PPP exchange rate at the level of total GDP, leaving its composition unchanged.

Table 2. Regressions Comparing Alternative Measures of Capital's Contribution[a]

	1960–2000		*1960–80*		*1980–2000*	
Independent variable	*2-1*	*2-2*	*2-3*	*2-4*	*2-5*	*2-6*
Growth in physical capital per worker[b]	0.56 (13.0)		0.38 (8.9)		0.70 (13.5)	
Investment per worker[c]		0.13 (5.3)		0.05 (2.5)		0.21 (7.7)
Summary statistics						
R^2	0.67	0.26	0.49	0.07	0.69	0.42
Standard error	0.82	1.24	1.08	1.46	1.04	1.42

Source: Authors' regressions.
a. The dependent variable is the average annual log change in output per worker times 100. A constant term is included in all regressions but not reported. The sample for all regressions consists of the eighty-four countries listed in appendix A. Numbers in parentheses are *t* statistics.
b. Measured as the average annual log change times 100. The capital stock is constructed as explained in the text.
c. Investment is measured as a percent share of GDP in constant national prices.

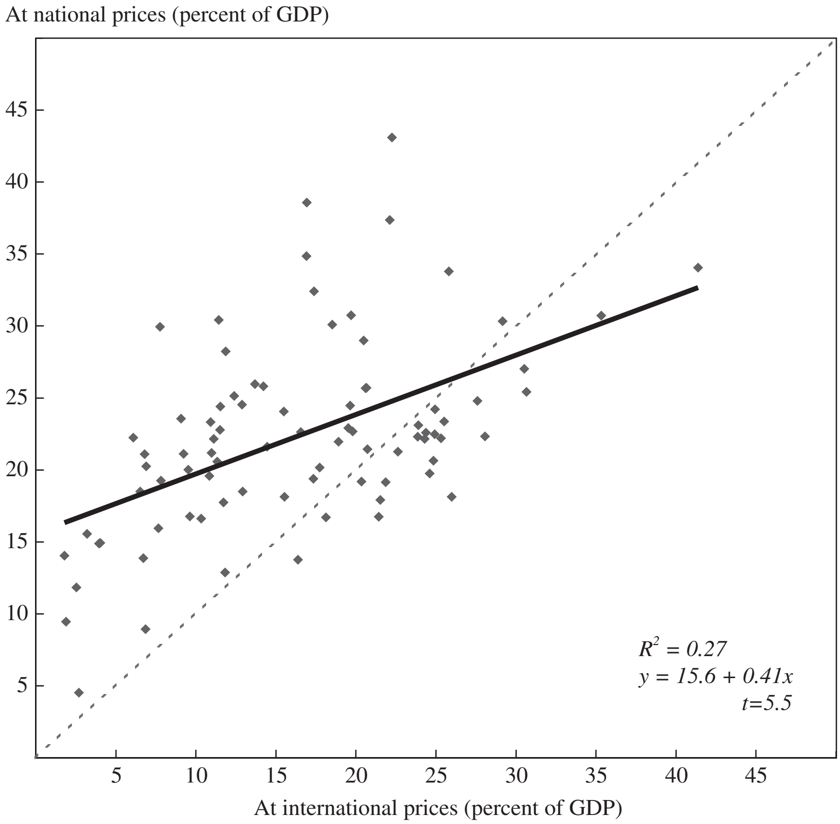
shares, reducing them in low-income countries while raising them for the most developed countries.[20]

Figure 3 provides a cross-country comparison of average investment shares based on national and international prices for our eighty-four-country sample.[21] The correlation between the two measures is surprisingly low (the R^2 from the bivariate regression of the share in international prices on that in national prices is only 0.27). From a comparison of the two panels of figure 4, it is also evident that conversion to international prices introduces a strong positive correlation between the investment rate and the level of income per capita. No such correlation exists when investment is measured in national prices. For these reasons the choice between national and international prices will play an important role in studies, such as that by Mankiw, Romer, and Weil, that rely on the investment rate to measure the capital input. We also note that nearly all of the studies that estimate the relationship between the *level* of income (as opposed to its rate of change) and the capital stock value the latter at international prices using the PPP exchange rate for investment

20. The change in the expenditure shares in the conversion to international prices also results in a somewhat different measure of the growth in aggregate output compared with the estimate derived from domestic prices. For more discussion of these issues in the context of the Gerschenkron effect, see Nuxoll (1994).

21. Figures 3 and 4 cover the shorter period 1960-98 because of the more limited availability of the PWT data.

Figure 3. Comparison of Investment Shares in National and International Prices, 1960–98[a]



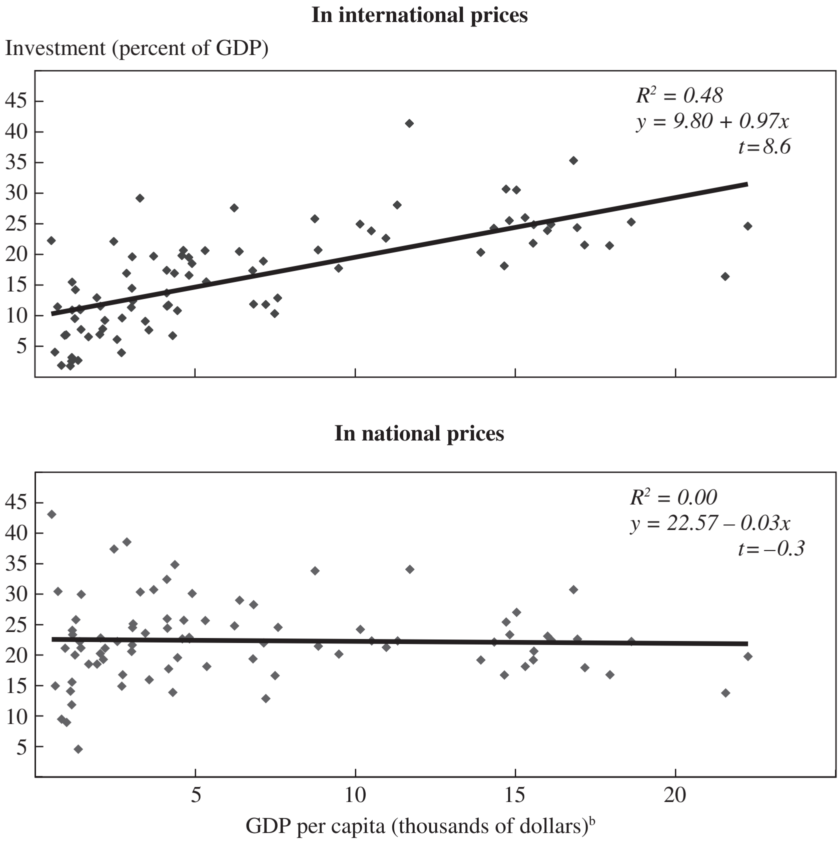
Source: Data from Penn World Tables and WDI.
a. Data are for the eighty-four countries listed in appendix A.

goods. Such a construction builds in a strong positive correlation between income and capital per worker.[22]

In a growth accounting context, we believe that the capital input should be valued in the prices of the country in which it is used. Profit-maximizing firms make production decisions based on the relative prices

22. The correlation between average investment rates and growth in output per worker is somewhat larger when the national price measure used in the regressions reported in table 2 is replaced with the international price measure. Nonetheless, changes in capital stocks continue to significantly outperform investment measured in international prices.

Figure 4. Investment Shares and GDP per Capita, 1960–98[a]



Source: Data from Penn World Tables and WDI.
a. Data are for the eighty-four countries listed in appendix A.
b. At international prices.

of capital and labor in their own domestic markets. It seems unreasonable to evaluate the production of poor countries using high international wage rates and low international capital prices. In addition, the average international price of capital does not reflect the influence of trade policy.

However, because the PWT converts national prices to international prices using the PPP exchange rate of a single year, the growth of real investment spending is the same in international and in national prices. Thus the choice between these two measures matters less for those studies that rely on changes in a constructed measure of the capital stock to measure capital services. It will alter only the level of the capital stock, and not its rate of growth.

Induced Investment

Our accounting decomposition measures the importance of capital to the growth of output per worker in terms of the change in the capital-labor ratio. However, some researchers argue that the focus on changes in the capital-labor ratio overstates the role of capital and undervalues TFP, because it ignores the fact that investment is endogenous in the sense that a portion of the change in capital is induced by an increase in TFP. Thus they maintain that "growth in physical capital induced by rising productivity should be attributed to productivity."[23]

These researchers propose an alternative benchmark that limits capital's contribution to increases in the capital-output ratio. That is, they rewrite equation 2 as

$$(2'') \qquad \Delta \ln(Y/L) = \left(\frac{\alpha}{1-\alpha}\right)[\Delta \ln(K/Y)] + \Delta \ln H + \left(\frac{1}{1-\alpha}\right)\Delta \ln A.$$

These alternative formulations are based on exactly the same measures of the changes in *A, K,* and *H*. However, they imply very different interpretations of how each contributes to growth. In particular, the last term in equation 2″ can be interpreted as the contribution of changes in TFP under the strong assumption that higher TFP induces increases in capital just sufficient to maintain the capital-output ratio. In effect, the investment rate is assumed equal to the capital-output ratio times the rate of growth of output (a simple accelerator relationship).

By restricting the contribution of capital to variations in the capital-output ratio, equation 2″ automatically expands the role of TFP. Compared with equation 2, TFP is "adjusted," or scaled upward, by $1/(1 - \alpha)$—an amount equal to 1.54 in our formulation. Equation 2″ is analogous to the formulation used by Robert Hall and Charles Jones.[24] Klenow and Rodríguez-Clare go even further in that they also assume a fully endogenous response of human capital to income growth.[25] In their version the

23. Klenow and Rodríguez-Clare (1997, p. 97). See also Barro and Sala-i-Martin (1995, p. 352).

24. Hall and Jones (1999).

25. Klenow and Rodríguez-Clare (1997).

contributions of both physical and human capital are restricted to increases in excess of the growth in output. As we illustrate below, the result, of course, is a dominant role for the TFP residual.

It is certainly true that capital investment is partly an induced response to changes in GDP associated with variations in TFP. Thus it has long been recognized that the assumption of wholly exogenous capital, as in the decomposition given by equation 2, leads to an overstatement of capital's contribution and an understatement of the contribution of TFP growth.[26] But it seems equally extreme to assume that the capital stock will simply and automatically adjust in a proportionate fashion to *all* deviations in the rate of growth of output induced by changes in TFP. Investment decisions are likely to be influenced by a large number of factors, such as the availability of finance, taxes, and other aspects of the investment environment, as well as changes in TFP.

This perspective suggests that an ideal representation would be somewhere between the two extremes of changes in the capital-labor ratio and changes in the capital-output ratio. However, the extent to which investment is in fact endogenous is a distinct issue from the preferred approach to measuring capital's contribution to growth. One can recognize that changes in the capital stock are at least partly induced by changes in TFP without concluding that measures of capital's contribution to growth should exclude this induced portion. Indeed, some growth models suggest that technological gains are embodied in new capital, creating a potential two-way interaction between capital accumulation and TFP growth.[27] The potential endogeneity of *both* the factor inputs and TFP reinforces our caution against using growth accounts to infer a causal interpretation of the growth process.

A dispute over the relative importance of capital accumulation and TFP can hardly be resolved by definitional changes. We have chosen to report capital's contribution in terms of the capital-labor ratio because, as discussed above, the steady-state assumption of a constant capital-output

26. Hulten (1975).

27. Klenow and Rodríguez-Clare (1997) assume a constant investment rate in a steady-state growth path in arguing that this concern is unfounded. However, as noted by Jones (1997, p. 110) in his comment on their paper, "If all countries were in their steady-states, then, as is well known, all growth would be attributable to TFP growth. In this sense, the [Klenow and Rodríguez-Clare] methodology is, in some ways, set up to deliver their result."

ratio seems unreasonable for a sample that is dominated by developing countries. Including the induced portion seems to us consistent with a focus on the proximate sources of growth. We also note that, despite long-standing awareness of this issue in the extensive literature that applies growth accounting to industrial countries, every study of which we are aware measures capital's contribution in terms of the capital-labor ratio.

A Variance Decomposition

Is the *variation* in growth of output per worker across countries primarily accounted for by variations in the growth of TFP, as researchers such as Klenow and Rodríguez-Clare, and Easterly and Levine, have claimed?[28] In this section we discuss alternative perspectives on the relative importance of capital accumulation, educational attainment, and TFP for our sample of countries.

The variance of $\Delta\ln(Y/L)$ is equal to the sum of the variances of each of three components (physical capital, education, and TFP) plus twice the sum of the three covariances. The existence of nonzero covariances implies that there is no unique way to allocate the variance of $\Delta\ln(Y/L)$ among the three components. Following Klenow and Rodríguez-Clare, we first divide the covariance terms equally among components, measuring the contribution of each component as its covariance with $\Delta\ln(Y/L)$ divided by the variance of $\Delta\ln(Y/L)$.[29] The top row of table 3 reports the resulting decomposition of the variance in growth of output per worker among the contributions of capital per worker, education, and the residual estimate of TFP. This is based on the decomposition in equation 2 for the entire forty-year period. We find that 43 percent of the variation in growth of output per worker is associated with variations in physical capital per worker, compared with only 3 percent with education and 54 percent with TFP. If the sample is weighted by population (second row of table 3), the importance of education is increased and that of physical capital declines.

28. Surprisingly, researchers who use the capital-output formulation continue to attribute most of the *trend* growth in GDP per capita to growth in the inputs (Klenow and Rodríguez-Clare, 1997, p. 94).

29. In other words we define the contribution of each component as its own variance plus its covariances with both other components, scaled by the variance of growth in output per worker. For example, the figure 0.43 in the top row of table 3 is equal to $0.27 + 0.5(0.27 + 0.3) + 0.02$. Baier, Dwyer, and Tamura (2002) explore alternative decompositions.

Table 3. Variance-Covariance Analysis of Output per Worker

	Contribution to variation in growth of output per worker			*Underlying decomposition*					
				Variance			*Covariance × 2*		
Equation	*Physical capital*	*Education*	*Factor productivity*	*K**	*H**	*A**	*K*, H**	*K*, A**	*H*, A**
Capital-labor ratio, unweighted[a]	0.43	0.03	0.54	0.27	0.01	0.40	0.03	0.27	0.02
Capital-labor ratio, population weighted[a]	0.37	0.09	0.54	0.14	0.01	0.30	0.06	0.39	0.09
Capital-output ratio, unweighted[b]	0.12	0.05	0.83	0.24	0.01	0.95	0.03	–0.26	0.04

Source: Authors' calculations as described in the text.

a. The contribution of each factor to growth in output per worker is defined as in equation 2; $K^* = \alpha\Delta\ln K/L$, $H^* = (1 - \alpha)\Delta\ln H$, and $A^* = \Delta\ln A$.

b. The contribution of each of factors K^*, H^*, and A^* to growth in output per worker is defined as in equation 2″; $K^* = [\alpha/(1 - \alpha)]\Delta\ln K/Y$, $H^* = \Delta\ln H$, and $A^* = [1/(1 - \alpha)]\Delta\ln A$.

The third row of table 3 reports the alternative decomposition based on equation 2″, limiting the capital contribution to variation in the capital-output ratio. As shown, the contribution of TFP rises to 83 percent of the total, whereas the contribution of physical capital falls to just 12 percent, consistent with the claim that capital accumulation is an unimportant contributor to growth. The result is clearly related to the shift in the accounting framework from the equation 2 formulation of TFP to the equation 2″ formulation.

Table 3 also shows the pieces underlying the variance decomposition. The entries can be used to construct upper and lower bounds for the contributions of capital and TFP under each of the alternatives. For example, in the $\Delta\ln(K/L)$ formulation, the weight on the contribution of physical capital ranges from 0.27 to 0.57, depending on whether the relevant covariance terms are allocated to capital.[30] Similarly, the weight on the contribution of capital could range from 0.24 to 0.01 under the $\Delta\ln(K/Y)$ formulation.

The main source of the relatively large contribution of TFP in the $\Delta\ln(K/Y)$ formulation is its much larger variance. This is a direct consequence of scaling: $0.95 = 0.40/(1 - \alpha)^2$. However, the reduced contribution of physical capital is not due to the difference between the variance of $\Delta\ln(K/L)$ and that of $\Delta\ln(K/Y)$.[31] Instead it arises from the fact that the covariance between the contribution of capital and that of TFP switches from positive, based on the $\Delta\ln(K/L)$ formulation, to negative, based on the $\Delta\ln(K/Y)$ formulation. The positive correlation between growth in $\Delta\ln(K/L)$ and TFP in the first line of table 3 could be taken as supporting the view that the capital accumulation was induced by productivity gains. However, this is just one of a number of plausible explanations, including the possibility that both productivity gains and capital accumulation were spurred by other common factors. Indeed, one would expect to observe a positive correlation between these variables. On the other hand, the negative correlation between growth in TFP and $\Delta\ln(K/Y)$ suggests to us that this formulation has indeed gone too far.[32] It is also somewhat surprising

30. The upper bound for the contribution of increases in $\Delta\ln(K/L)$ is given by the variance of the contribution of $\Delta\ln(K/L)$ plus twice the sum of the two relevant covariances, divided by the variance of $\Delta\ln(Y/L)$. This is equal to $0.27 + 0.03 + 0.27 = 0.57$.

31. The variances of $\Delta\ln(K/L)$ and $\Delta\ln(K/Y)$ are 0.55 and 0.48, respectively. In table 3, recall that all entries are scaled by the variance of $\Delta\ln(Y/L)$ (equal to 2.03).

32. Klenow and Rodríguez-Clare also report a negative correlation between growth in TFP and $\Delta\ln(K/Y)$. However, they suggest that this “could indicate an overstatement of the

that these variables show so little relation to each other under either formulation. Regressing the changes in ln(*K*/*L*) or in ln(*K*/*Y*) on changes in TFP results in R^2s of just 0.13 and 0.06, respectively.

We conclude that both capital (physical and human) accumulation and improvements in economic efficiency are central to the growth process. For most purposes the emphasis on determining which is more exogenous or more important seems misplaced. Policies that aim to promote TFP growth will also tend to promote capital formation, and vice versa. An emphasis on either of the two extremes offers few insights into the growth process. In sum, we agree strongly with Charles Jones, who states in his comment on Klenow and Rodríguez-Clare that "oftentimes readers want an all or nothing answer. . . . A better answer, I think, is that both traditional inputs and productivity play large and important roles."[33]

The Contribution of Education

A second area of dispute involves the role of education. Relying on a large body of microeconomic evidence of a strong relationship between education and earnings, several growth accounting studies, including our own, adjust the work force for improvements in educational attainment.[34] However, as discussed below, at the macroeconomic level a number of recent studies have been unable to find a correlation between economic growth and increased educational attainment. This result has been used as a basis for rejecting the microeconomic evidence and for arguing that the focus of governments and the multilateral organizations on raising levels of literacy and average educational attainment has been misplaced.[35]

As an aside, the problem may be unrealistic expectations. Given that average years of schooling change very slowly, the effects on output growth may be hard to detect in the international data. Under our assumption that the social and private returns are equal to 7 percent a year, the

---

contribution of [Δln(*K*/*L*)] . . . [implying that] . . . the role of A is even larger" (1997, p. 96). We find this view unconvincing.

33. Jones (1997, p. 110).

34. Summaries of the microeconomic studies covering a variety of countries are available in Psacharopoulos (1994) and Bils and Klenow (2000).

35. Easterly (2001).

average annual contribution of education to output growth is only 0.3 percent a year (table 1), and the standard deviation across the eighty-four countries is just 0.1 percent.

Increases in education could have an impact on economic growth through two different channels. First, more education may improve the productivity, or quality, of workers. This is the formulation we employ in multiplying the quantity of workers by an index of average educational attainment and imputing the return to education from microeconomic studies. Specifically, we assume a 7 percent return to an additional year of schooling—a value near the lower boundary of the results from the microeconomic studies. Thus, for a country such as the United States in 1990, with an average level of educational attainment near twelve years, the effective labor supply is treated as 2.25 times the number of persons.

An alternative formulation, adopted by Mankiw, Romer, and Weil and by Klenow and Rodríguez-Clare,[36] specifies human capital (education) as an independent factor in the growth process, one that can augment labor, physical capital, and TFP. The relationship with TFP reflects the view that an educated work force is better able to implement new technologies and to generate ideas for improving efficiency. Designing an empirical test to distinguish between these two channels is very difficult. Thus both suggest potential justifications for expecting a positive correlation between gains in educational attainment and growth.

Whereas microeconomic studies aimed at estimating the relationship between income and educational attainment have a long history, empirical macroeconomic studies of the issue are relatively recent. This work was greatly stimulated by Barro and Lee, who developed a comprehensive data set on schooling, covering a large number of countries over an extended time period.[37] They use national censuses and surveys taken at five-year intervals beginning in 1960 to infer the proportions of the working-age population with various levels of schooling.[38] However, large gaps in the census data require them to make extensive use of enrollment information to extrapolate and interpolate the census information.

36. Mankiw, Romer, and Weil (1992); Klenow and Rodríguez-Clare (1997).

37. Barro and Lee (1993, 2000).

38. The Barro-Lee data distinguish among three levels of schooling—primary, secondary, and tertiary—and between those who initiate a level of schooling and those who complete it.

Early studies, including those of Mankiw, Romer, and Weil and of Barro and Xavier Sala-i-Martin,[39] found a significant positive association between cross-country differences in the initial endowment of education and subsequent rates of growth. Barro has explored the link between growth and a variety of schooling level indicators.[40] However, later studies that examined the relationship between changes in years of schooling and changes in average incomes failed to find a significant association.[41]

The failure to replicate the microeconomic results at the aggregate level has three possible explanations. First, the social return to education, as reflected in the aggregate data, may be much less than the private return that underlies the microeconomic analysis. Second, there may be measurement errors in the data. Third, cross-country variations in educational attainment may fail to account for variations in the quality of education. We examine each of these issues in turn.

Social versus Private Returns

There is a long-standing debate over how to interpret the coefficient on years of schooling in the microeconomic analyses of wage differentials. Does it reflect the skill gains from education, as argued by Gary Becker?[42] Or does the educational process simply sort people by native abilities, thereby providing a convenient indicator (or "signal") of hard-to-observe characteristics, as argued by Michael Spence?[43] If the latter process dominates, aggregate gains would be limited to a somewhat better matching of workers and jobs and would be substantially overstated by estimates of the private return. On the other hand, a case can also be made that the true social or aggregate gains exceed the private returns because an educated work force accelerates innovation and its introduction into the production process.

Problems in designing a microeconomic study that can fully distinguish between the roles of signaling and of skill improvement make it dif-

39. Mankiw, Romer, and Weil (1992); Barro and Sala-i-Martin (1995).

40. Barro (2001). Even the linkage between growth and schooling level has been called into question. Bils and Klenow (2000) use a calibration model to argue that less than a third of this relationship should be interpreted as reflecting the impact of schooling on growth.

41. See, for example, Benhabib and Spiegel (1994), Bils and Klenow (2000), Pritchett (2001), Easterly and Levine (2001), and Temple (2001).

42. Becker (1964).

43. Spence (1973).

ficult to rule out the possibility that empirical estimates do reflect signaling, thereby overstating the actual return to schooling. However, efforts to use natural experiments, such as episodes of change in compulsory education requirements or other changes in schooling that are uncorrelated with ability, have found little evidence of a significant upward bias in the estimated return.[44] From this perspective, the fact that macroeconomic analysis has had such difficulty finding a positive association between increased average years of schooling and economic growth, even in those studies that control for other factors, is puzzling.

Some researchers suggest that the benefits of education are not fully realized because of a failure to integrate improvements in education with other important elements of the growth process. That is, the creation of skills offers no benefits if the technology and infrastructure do not exist to make use of them. Although this explanation sounds plausible, it is not consistent with the fact that the correlation between growth and changes in educational attainment is also weak in samples limited to OECD economies.

Data Measurement

Nearly all of the contributors to the empirical literature recognize that measurement error might account for the lack of association between economic growth and gains in educational attainment. In one of the first efforts to seriously explore this issue, Ángel de la Fuente and Rafael Doménech found large variations in the classification of educational attainment over successive censuses in many OECD countries.[45] They developed a new estimate of educational attainment that adjusts for classification changes and that appears to evolve over time in a much smoother fashion (with a reduced signal-to-noise ratio) than the Barro-Lee data.[46] The two measures of average educational attainment have similar levels, but there is almost no correlation of the changes over a thirty-year period. When they used their data to estimate a model similar to that of the earlier studies, they obtained a much stronger correlation

44. Ashenfelter, Harmon, and Oosterbeek (1999) survey some of the major studies. See also an earlier paper by Griliches (1977).

45. De la Fuente and Doménech (2000, 2001).

46. Classification issues include inconsistencies in the treatment of vocational and technical training, as well as changes in the numbers of years associated with different levels of schooling.

between the accumulation of human capital and economic growth: their estimated coefficient on the change in educational attainment was near that implied by the Mankiw, Romer, and Weil study. De la Fuente and Doménech concluded that measurement problems were responsible for most of the earlier difficulties in discerning a positive return to education. However, their analysis was limited to the OECD countries.

A second global data set, covering ninety-five countries, has been developed by Daniel Cohen and Marcelo Soto as an extension of earlier work done at the OECD.[47] They compute educational attainment at the beginning of each decade for the period 1960–2000. For some countries they had more recent census information than that used by Barro and Lee. But a more important difference arises from their use of age-specific data in the available censuses to construct estimates of educational attainment for each age cohort in other years for which direct observations were missing. That is, the educational attainment of a specific age cohort is assumed to be the same at successive ten-year periods. Thus they use enrollment data only to fill missing age cohort cells. They also report significant differences between data from national sources and the data available from the multilateral agencies used by Barro and Lee. However, for many countries their series are based on information from a single census. Like de la Fuente and Doménech, Cohen and Soto point to excessive volatility over time in the Barro-Lee data, which appears to reflect changes in national classifications.

Cohen and Soto use their data to examine the relationship between economic growth and years of schooling.[48] Using a variety of specifications and econometric techniques, they estimate annual returns to schooling in the range of 7 to 10 percent, close to the average of the microeconomic studies. They argue that earlier difficulties in finding a positive correlation were partly related to measurement problems. Alan Krueger and Mikael Lindahl also stressed the importance of measurement error in their evaluation of the micro- and macroeconomic evidence.[49] As we argue below, measurement error does seem to be a major problem for the macroeconomic studies. However, we are not convinced that any one of the available data series is clearly preferable to the alternatives.

47. Cohen and Soto (2001).
48. Cohen and Soto (2001); Soto (2002).
49. Krueger and Lindahl (2001).

We have data from the Barro-Lee and Cohen-Soto data sets for seventy-three of the countries in our sample for the period 1960–2000.[50] The top panel of figure 5 compares average years of schooling in the two data sets over the period. The two are very highly correlated, with an R^2 of 0.93. But the Cohen-Soto estimates are generally higher, and for a few countries the difference exceeds two years of schooling. Some of the variation can be traced to different methods for estimating completion rates, where the census information is particularly limited.[51]

The correspondence between the two education data sets is much poorer in terms of changes over time, however. For the forty-year changes, shown in the bottom panel of figure 5, the R^2 of the bivariate regression is 0.28. On average, the Cohen-Soto data indicate greater improvement in years of schooling, and the differences are large in a significant number of countries. The correspondence is even worse for ten-year changes, with R^2s for the four subperiods ranging between 0.1 and 0.2.[52]

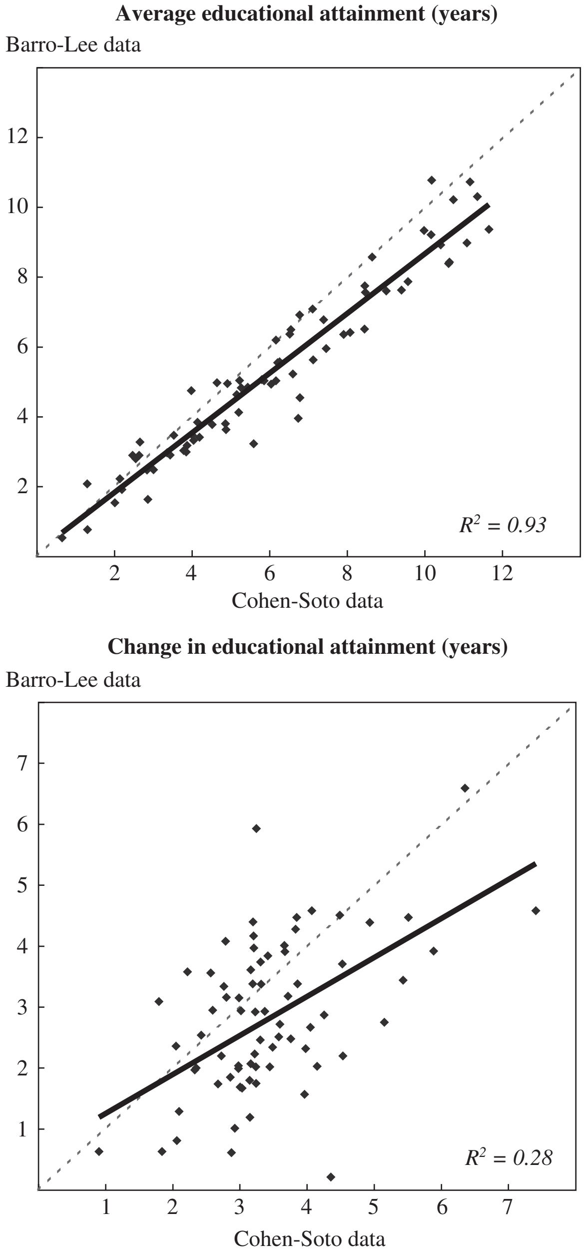
Both the Cohen-Soto and the Barro-Lee data differ substantially from those of de la Fuente and Doménech. For a common group of twenty OECD countries, the thirty-year change (1960–90) in years of schooling reported by de la Fuente and Doménech has no correlation with the corresponding changes reported by Barro and Lee and an R^2 of 0.23 with the changes reported by Cohen and Soto. For the same data, the Cohen-Soto measures are also uncorrelated with those of Barro and Lee. Such large differences among the estimates are surprising, given the expectation that information from the OECD countries would be the most reliable. Cohen and Soto report results from growth regressions in which their measure of

50. We can also make use of a third data set compiled by Nehru, Swanson, and Dubey (1995), which is attractive because it relies only on school enrollment data. The disadvantage is that this data set is limited to 1960–87. We also undertook a three-way comparison that included these alternative data. None of the conclusions reported in the text were altered, and these results are omitted from the discussion.

51. Cohen and Soto, like de la Fuente and Doménech, typically assume that everyone who starts a given level of schooling completes it. This implies large differences between alternative data series in the ratio of those completing given levels of schooling to those attending. For example, in the Cohen and Soto data, for many countries, everyone who enters higher education is assumed to complete it. In contrast, in the Barro-Lee data, the ratio of completers to attenders varies widely in some countries over short periods. Both approaches generate some implausible results.

52. The Nehru, Swanson, and Dubey (1995) measure also shows a relatively low correlation with the other two series (results not shown).

Figure 5. Comparison of Measures of Educational Attainment, 1960–2000[a]



Source: Data from Barro and Lee (2000) and Cohen and Soto (2001).
a. Data are for seventy-three countries.

changes in schooling enters significantly but the Barro-Lee measure does not. On this basis they argue that their series should be preferred over the Barro-Lee data. However, we find no such result using our sample. Instead a variety of growth regressions that incorporate the two measures of human capital provide little basis for choosing between them.

Alternatively, we could view the two measures as proxies for the true value and search for alternative ways to combine them that would yield the best measure. Unfortunately, none of the approaches we explored to combining the two proxies proved satisfactory.

First, we use instrumental variables in a regression equation relating growth in income per capita to growth in physical capital per worker and to changes in schooling. This is simply a regression version of our growth accounting formulas, equations 2 and 4:

$$\Delta \ln(Y/L) = \beta_1 \Delta \ln(K/L) + \beta_2 \Delta s + \mu. \quad (7)$$

If the private and social returns to schooling are equal, we would expect the coefficient on Δs to be about 0.045, or $0.07 \times (1 - \alpha)$. Under the assumption that the measurement errors are uncorrelated, each of the proxies is a valid instrumental variable for the other. However, all variations of this regression resulted in estimates of the coefficient on the change in schooling that were small or negative and always statistically insignificant.

Second, following Krueger and Lindahl,[53] we construct a reliability measure for each proxy, based on its covariance with the alternative measure divided by its variance. We obtain results of 0.63 for the Cohen-Soto series and 0.43 for Barro-Lee. These reliability measures suggest that the larger weight should be assigned to the Cohen-Soto series. However, we are doubtful about the value of this criterion. Since they share a common covariance, the reliability measures will differ only because the two proxies have different variances. But there is no particular reason here to believe that the variable with lower cross-country variance has less measurement error.

Finally, we implemented an approach suggested by Darren Lubotsky and Martin Wittenberg.[54] Here both proxy measures are included in the estimate of equation 7, and the regression weights are used to form a new

53. Krueger and Lindahl (2001).
54. Lubotsky and Wittenberg (2001).

composite variable. Lubotsky and Wittenberg argue that this "post hoc" estimate is superior to any "a priori" set of weights. In an equation based on forty-year changes that included physical capital per worker, the coefficient on the Barro-Lee measure was positive, but the coefficient on the Cohen-Soto measure was negative, and neither approached statistical significance.

We conclude that there is substantial evidence of measurement error. However, none of the alternative approaches yields a convincing way to choose between or to combine information from the available schooling proxies. Thus we opted to adopt a measure of educational attainment based on the simple average of the Cohen-Soto and Barro-Lee estimates of years of schooling.

Educational Quality

The use of years of schooling as the measure of educational attainment does not incorporate any adjustment for variations in the quality of education. This is likely to be a far more serious problem for international comparisons of the correlation between incomes and education than for microeconomic studies, since the quality of education within a country might be relatively homogeneous. Despite general agreement that the quality of education varies substantially across countries, obtaining quality measures for a large number of countries is difficult.

The most extensive empirical analysis is that of Eric Hanushek and Dennis Kimko, who developed indexes of educational quality for thirty-eight countries based on international tests of academic performance in mathematics and science over 1965–91.[55] To infer that differences in academic performance are reflected in the quality of the work force, we must assume that country differences in test performance persist over long periods. Hanushek and Kimko also sought to associate their quality indexes with other correlates of educational performance in order to extend the quality measure to a larger number of countries. Thus, for thirty countries where the variables were measured directly, they estimated a statistical relationship between the educational quality index and indicators such as primary school enrollment rate, average years of schooling, expenditure

55. Hanushek and Kimko (2000).

per student, population growth, and regional dummies.[56] The resulting estimates were used to generate predicted values for an additional forty-nine countries, thirty-six of which are in our sample. Using this expanded data set, Hanushek and Kimko found a strong correlation between their measure of educational quality and increases in GDP per capita. At the same time, the quality variable had the effect of eliminating any significant correlation between the quantity of schooling and economic growth.

We used an updated set of correlates from the 2002 WDI to reestimate and extend the Hanushek-Kimko measure of educational quality to our full set of eighty-four countries. In addition, we added Chile to the analysis of directly measured countries in order to have at least two countries, Chile and Brazil, on which to base a regional adjustment for Latin America.[57] Appendix B provides the details.

Empirical Estimates

Table 4 presents regression estimates of the relationship between education and economic growth. The dependent variable is the average annual change (in logarithms) in real GDP per worker over 1960–2000 for our eighty-four-country sample. Column 4-1 shows the results of a regression in which growth in physical and human capital per worker and the initial 1960 level of years of schooling are the explanatory variables. The estimated coefficient on physical capital is larger than the 0.35 that we assumed for construction of the growth accounts, but such a result would be expected because of the bias that results from the endogeneity of both capital accumulation and total GDP. The coefficient on the change in human capital is closer to the value of 0.65 used in the growth accounts, but it is statistically insignificant. We find a stronger correlation between growth and the initial level of educational attainment than between growth and the change in educational attainment (growth in human capital per worker).

In column 4-2 the coefficient on physical capital is constrained to equal the hypothesized value of 0.35. In this case the coefficients on the level

56. One problem with the estimated relationship is that most of the correlation of quality is with the enrollment rate (a quantity measure) and the regional dummies. Thus there is no right-hand-side variable that is strongly identifiable as a measure of quality.

57. The Hanushek-Kimko estimates used only Brazil, and data for all other Latin American countries were scaled relative to Brazil.

Table 4. Regressions Explaining Growth in Output per Worker with Measures of Educational Attainment and Quality[a]

Independent variable	*4-1*	*4-2*	*4-3*	*4-4*	*4-5*
Growth in physical capital per worker[b]	0.51	0.35	0.27	0.48	0.27
	(11.5)		(6.2)	(10.5)	(6.0)
Growth in human capital per worker[b]	0.74	1.55	0.55	0.82	0.53
	(1.4)	(3.0)	(1.3)	(1.6)	(1.3)
Initial level of average years of schooling	0.11	0.13	0.08	0.07	0.07
	(3.5)	(3.7)	(1.4)	(1.8)	(1.0)
Educational quality[c]				0.02	0.01
				(2.2)	(0.7)
Initial conditions included[d]	No	No	Yes	No	Yes
Constant	–0.41	–0.51	–4.25	–0.90	–4.53
	(–1.4)	(–1.6)	(–3.9)	(–2.5)	(–3.9)
Adjusted R^2	0.71	0.70	0.84	0.72	0.84

Source: Authors' regressions.

a. The dependent variable is the average annual log change in real output per worker times 100, over 1960–2000. The number of observations in all regressions is eighty-four. Numbers in parentheses are *t* statistics.

b. Calculated as the average annual log change times 100.

c. Measure is expanded to eighty-four countries using data from the 2002 World Development Indicators as shown in equation B3 in appendix B. See appendix B for details on the sources and construction of the variable.

d. Variables for initial conditions include GDP per capita in 1960, life expectancy in 1960, log of population in 1960, trade instrument, geography, and institutional quality and are described in table C1 in appendix C.

and the rate of change of educational attainment become statistically very significant. Thus we find support for the argument of Krueger and Lindahl that it is critical to control for the role of physical capital. However, the coefficient on the change in human capital is now too large to represent the augmenting effect of education on the work force. Column 4-3 shows the effects of adding a set of additional conditioning variables that have been found to be consistently correlated with growth.[58] These variables are discussed more fully in the next section, but they have the effect of reducing the coefficients on the changes in physical and human capital to values close to our hypothesized values. However, neither of the coefficients on the education variables is statistically significant.

Finally, columns 4-4 and 4-5 show the effect of including the measure of educational quality.[59] Our results in column 4-4 are very similar to

58. These variables include initial income and life expectancy, the standard deviation of the terms of trade, a measure of geographical distance from the equator, and the quality of government institutions.

59. These results are based on the quality measure that we derived from the WDI data, but the measure of Hanushek and Kimko, as augmented by Woessman (2000), performed nearly as well.

those of Hanushek and Kimko in that the quality variable is statistically significant, but this result comes at the cost of reducing the role of educational attainment. However, the finding of a significant coefficient on educational quality is not robust to inclusion of the set of conditioning variables, as shown in column 4-5. In particular, the loss of significance is associated specifically with a measure of the quality of government institutions (we discuss this measure more fully in the next section). Although the notion that the quality of education matters for growth is eminently sensible, we cannot distinguish it from more general concepts of the quality of government institutions. As we demonstrate in appendix B, the quality of governing institutions is the single best correlate with which to identify cross-country differences in measures of educational quality.

Macroeconomic evidence of the contribution of education to growth is clearly much weaker than that derived from microeconomic studies. Resolution of this issue is complicated by the substantial measurement error implied by differences in the magnitude of cross-country improvements in educational attainment among alternative data sets. We were not able to distinguish among the various measures and have instead relied on a simple average of the results from two large independent studies. Finally, like Hanushek and Kimko, we find quality of education to be significantly correlated with growth. But we lack any effective means of measuring its change over time. Furthermore, educational quality is highly correlated with measures of the quality of governing institutions and may simply be a proxy for this broader concept.

Economic Growth: 1960–80 versus 1980–2000

The last two decades witnessed a remarkable collapse of growth in much of the global economy. Table 5 provides a region-by-region summary of the change in growth for our sample of eighty-four countries over the twenty years before and after 1980. On average, growth in output per worker slowed from an annual rate of 2.5 percent in 1960–80 to only 0.8 percent in 1980–2000. The slowdown was of almost equal magnitude in the industrial and the developing countries, and it is apparent in all regions except South Asia. Furthermore, lower rates of both physical capital accumulation and TFP growth were important contributors to the slowdown. On average across countries, a lower rate of physical capital

Table 5. Decomposition of the Change in Output Growth between 1960–80 and 1980–2000

	Average annual growth of output per worker (percent)		*Change in growth rate of output per worker (percentage points)*	*Contribution of component (percentage points)*		
Region	*1960–80*	*1980–2000*		*Physical capital*	*Education*	*Total factor productivity*
World (84 countries)						
Mean	2.5	0.8	–1.7	–0.7	0.0	–1.0
Share of variation[a]				0.28	–0.01	0.73
Developing countries (62)						
Mean	2.3	0.6	–1.7	–0.7	0.0	–1.1
Share of variation				0.26	–0.01	0.75
Africa (19)						
Mean	1.4	–0.3	–1.7	–0.9	0.1	–1.0
Share of variation				0.25	–0.02	0.78
East Asia with China (8)						
Mean	4.1	3.9	–0.2	–0.3	0.0	0.0
Share of variation				0.40	0.01	0.59
East Asia without China						
Mean	4.3	3.4	–1.0	–0.5	0.0	–0.4
Share of variation				0.44	0.06	0.51
Latin America (22)						
Mean	2.0	–0.5	–2.4	–0.6	0.0	–1.8
Share of variation				0.14	0.00	0.86

Middle East (9)						
Mean	3.3	1.0	–2.3	–1.1	0.1	–1.3
Share of variation				0.35	0.02	0.63
South Asia (4)						
Mean	2.0	2.6	0.6	–0.1	0.1	0.6
Share of variation				0.51	–0.02	0.51
Industrial countries (22)						
Mean	3.1	1.5	–1.6	–0.7	0.0	–0.8
Share of variation				0.43	–0.02	0.59
25 countries with greatest increase in growth						
Mean	2.0	2.4	0.4	0.0	0.0	0.4
Share of variation				0.24	–0.01	0.76
China	2.2	7.1	4.9	1.8	0.0	3.2
India	1.3	3.5	2.2	0.4	0.1	1.7
Uganda	–1.1	2.3	3.4	–0.5	0.2	3.7
25 countries with greatest decrease in growth						
Mean	3.0	–0.5	–3.5	–1.1	0.0	–2.5
Share of variation				0.31	–0.02	0.71

Source: Authors' calculations as explained in the text.

a. Measured as the covariance of the average annual log change in output per worker with the change in the factor contribution, divided by the total variance of the average annual log change in output per worker.

accumulation accounted for about 40 percent (–0.7/–1.7) of the slowdown, but it was less important as a source of the cross-country variation (28 percent). (As in the analysis above, this variance decomposition splits the covariance terms equally among pairs of components.) Changes in educational attainment were of minor importance in all cases.

However, the averages also mask some important exceptions. The three countries with the largest acceleration of growth were China, India, and Uganda. The situation of Uganda differs from that of the other two because the 1970s and early 1980s were a time of internal strife and chaos in that country. Thus, although growth was relatively strong in 1980–2000, the turnaround mainly reflects the rebound from negative growth in the earlier period. On the other hand, the performance of China and India has been extraordinarily important both because they achieved a significant acceleration of growth after 1980, and because they account for a large proportion of the world's population and an even larger proportion of the world's poor. In much of our empirical analysis, they are only two out of eighty-four countries, but they represent 45 percent of the population of our global sample and 56 percent of the population of developing countries in our sample.

The disproportionate impacts of China and India are highlighted in table 6, where the post-1980s slowing of global growth is transformed into an acceleration when the country experiences are weighted by population. Measured as a simple average of the eighty-four countries, global growth in income per capita slowed by 1.7 percentage points; based on population weighting, however, it accelerated by 0.7 percentage point. The fact that the two largest (and two of the poorest) countries in the sample grew far more than the average translates into a major reduction in global poverty and suggests a much different perspective on the post-1980 experience.[60]

Alternatively, the use of GDP weights has the effect of translating the sample into one that emphasizes the growth experience of the industrial countries, which account for two-thirds of aggregate output but only one-fifth of total population of the sample. The implication for the magnitude of the growth slowdown, shown in the third column of table 6, is intermediate between that for the simple average and that using population

60. See Deaton (2003) for a detailed discussion of the use of national accounts versus survey data to study trends in global poverty.

Table 6. Alternative Measures of the Change in Annual Growth in Output per Worker between 1960–80 and 1980–2000
Percentage points

Region	*Unweighted average*	*Countries weighted by population*	*Countries weighted by GDP*
World	–1.7	0.7	–0.9
Developing countries	–1.7	1.3	–0.4
Africa	–1.7	–2.1	–2.6
East Asia (excluding China)	–1.0	–0.9	–0.9
East Asia (including China)	–0.2	3.5	2.2
Latin America	–2.4	–3.2	–3.1
Middle East	–2.3	–2.2	–2.3
South Asia	0.6	1.7	1.8
Industrial countries	–1.6	–1.5	–1.2
25 countries with greatest increase in growth	0.4	2.5	0.8
25 countries with greatest decrease in growth	–3.5	–3.6	–3.7

Source: Authors' calculations as explained in the text; see also appendix C.

weights: the change in the GDP-weighted average growth rate is –0.9 percentage point.

In this section we use a combination of the growth accounts and regression analysis to explore the sources of the change in growth before and after 1980. Although the last decade has witnessed a large number of empirical studies of the growth process, nearly all of that research has focused on identifying the sources of variation across countries for a single period stretching from 1960 to the present. Much less effort has been devoted to exploring the sources of change between subperiods.[61] In part the emphasis on a single period has been dictated by the need to measure changes over relatively long periods in order to exclude cyclical complications. However, two twenty-year periods should be of sufficient length to minimize the role of cyclical factors, while greatly expanding the range of observed variation in growth rates. In addition, following the approach in our 1996 Brookings Paper,[62] we can combine the growth accounting decomposition with regression analysis to explore the channels through

61. A recent exception is Dollar and Kraay (2002), who included an analysis of the role of trade using decadal data. Islam (1995) used panel data based on five-year averages.

62. Collins and Bosworth (1996).

which various factors influence growth in income per worker. Does that growth occur principally through the effects of these factors on capital accumulation or through stimulating improvements in the efficiency of resource use, that is, TFP?

As part of our effort to compile a standard set of growth accounts over a forty-year period, we have also culled a set of principal determinants of growth from the empirical literature and expanded those data where necessary to cover our eighty-four countries. We use regression analysis to relate economic growth over the full forty-year period to some basic measures of initial conditions, external shocks, and policy. This specification is largely drawn from the existing empirical literature. We then use that specification to explore the extent to which the various determinants operate through the channel of capital accumulation and to what extent through improvement in TFP. Finally, the specification is applied to the two twenty-year subperiods to determine to what extent we can account for the sharp changes in growth performance.

In recent years the use of regression analysis to explore the determinants of growth has encountered significant criticism. Some surveys of that literature, for example, conclude that it has all been for naught, and that the regression analysis has been a disservice to policymakers because the research has failed to adequately communicate the extent of parameter instability.[63] Levine and David Renelt argue that few of the empirical relationships in the growth literature display a "robust" correlation with economic growth.[64] Many of the concerns arise out of the extreme heterogeneity of the sampled population of economies.[65] On the other hand, although cross-country regressions should be only one of several meth-

63. Lindauer and Pritchett (2002).

64. Levine and Renelt (1992). Sala-i-Martin (1997) argues that the extreme-bounds test used by Levine and Renelt is too strict. Using a cumulative density function, he finds that nearly half of the fifty-nine variables that he tested should be viewed as potentially important regressors. Fernandez, Ley, and Steel (2001) arrived at similar conclusions using a Bayesian framework. However, one disappointing aspect is that most of the variables identified in these studies are religious or geographic measures that are largely beyond the control of policymakers.

65. Kenny and Williams (2001); Brock, Durlauf, and West (2003). The latter (p. 296) apply econometric techniques that account for model uncertainty to growth analysis. Somewhat surprisingly, they conclude that there are "important respects in which our new approach did not provide particularly different insights from what one obtains from [ordinary least squares] exercises."

ods, evaluation of hypotheses in terms of their consistency with a wide range of empirical experiences needs to be a central component of any effort to build a coherent explanation of the growth process.[66]

We try to improve on this evaluation process through standardization, thereby removing an important reason for the varying results in the literature. It is important to compare regression results using a standard set of countries, standard time periods, and a standard set of conditioning variables.[67] By conditioning variables we mean a set of determinants that have been found to be important in a large number of studies. In this process some experimentation with alternative, but equally plausible, measures of a given determinant is unavoidable. But the emphasis ought to be on the consistency of the results for a general determinant, not the specific measure. In addition, we examine the robustness of the regression results across subsets of the data set (rich countries and poor countries) and subperiods (pre- and post-1980). Finally, we note that restricting the data to exclude transition economies as well as the smallest countries and city-states omits country groups that may be particularly unusual.

Results from the Forty-Year Sample

Table 7 summarizes our basic measures of the determinants of growth. In developing the indicators of initial conditions, we have relied heavily on prior work by Barro and Lee and by Hall and Jones.[68] The 1960 level of income per capita (in international prices) is measured as a ratio to the U.S. level and serves to capture the convergence process. Life expectancy in 1960 is included as a measure of initial health conditions.[69] At the suggestion of one of our discussants, we have included the logarithm of population in 1960 and a trade instrument among our measures of initial

66. One might think that these concerns could be addressed using panel data to increase the number of observations, by defining the dependent variable as growth in each country over a shorter period. We do consider twenty-year periods later in this section. However, we are not convinced that analysis of even shorter periods provides a means to address the same issues.

67. In contrast, studies such as Baier, Dwyer, and Tamura (2002) combine observations in which growth is measured over different time periods of varying length across countries.

68. Barro and Lee (1993, 1994); Hall and Jones (1999).

69. Other studies have included the initial level of income per capita and life expectancy in logarithmic form. We use the levels versions only because they fit slightly better and are less collinear with other included variables.

Table 7. Means and Standard Deviations of Conditioning and Policy Variables Used in the Growth Regressions, by Region, 1960–2000[a]

	Variables representing initial conditions						*Policy variables*		
Region	*Initial income per capita*	*Life expectancy*	*Log of population*	*Frankel-Romer-Rose trade instrument*	*Geography*	*Institutional quality*	*Budget balance*	*Inflation*	*Average Sachs-Warner openness*
Developing countries (62 countries)	0.17	49.90	15.76	0.08	–0.55	0.48	–3.53	16.43	0.27
	(0.11)	(9.67)	(1.48)	(0.07)	(0.73)	(0.14)	(3.03)	(16.35)	(0.33)
Africa (19)	0.11	41.59	15.52	0.06	–0.90	0.47	–4.98	13.30	0.08
	(0.09)	(5.93)	(0.91)	(0.04)	(0.37)	(0.12)	(2.15)	(7.65)	(0.24)
East Asia incl. China (8)	0.12	52.54	17.04	0.14	–0.34	0.61	–0.92	8.69	0.66
	(0.05)	(9.67)	(1.77)	(0.15)	(1.11)	(0.20)	(2.39)	(9.01)	(0.38)
China	0.04	36.32	20.32	0.04	1.32	0.57	–1.03	4.03	0.00
Latin America (22)	0.25	55.44	15.19	0.08	–0.71	0.43	–3.16	25.07	0.28
	(0.11)	(7.64)	(1.24)	(0.04)	(0.56)	(0.13)	(3.37)	(23.02)	(0.25)
Middle East (9)	0.21	52.79	15.59	0.11	0.26	0.49	–3.08	12.51	0.36
	(0.08)	(9.98)	(1.48)	(0.07)	(0.66)	(0.10)	(2.58)	(10.23)	(0.40)
South Asia (4)	0.08	47.04	17.85	0.10	–0.23	0.42	–4.98	8.03	0.06
	(0.02)	(8.97)	(1.55)	(0.03)	(0.67)	(0.12)	(3.53)	(0.56)	(0.13)
India	0.07	44.33	19.89	0.06	–0.36	0.58	–5.08	7.67	0.00
Industrial countries (22)	0.62	70.26	16.24	0.13	1.08	0.91	–1.96	6.35	0.91
	(0.22)	(2.24)	(1.53)	(0.09)	(0.45)	(0.11)	(2.56)	(3.25)	(0.20)

Source: Authors' calculations using sources listed in table C1 in appendix C.
a. Numbers in parentheses are regional standard deviations.

conditions.[70] Population is a dimension of country size, and the trade instrument can be viewed as a measure of a country's predisposition to trade. We examined several measures of geographical factors and obtained the most significant results with a composite average of the number of frost days and area within the tropics.[71] Table C2 in appendix C lists the other geographic indicators we considered. All five of the variables discussed above are considered exogenous in our regressions.

We also explored a number of alternative indicators of institutional quality (listed in table C2) and obtained the most significant results with a composite variable constructed by Stephen Knack and Philip Keefer from information in the International Country Risk Guide.[72] This index proved superior to alternatives obtained from Daniel Kaufman, Aart Kraay, and Pablo Zoido-Lobatón and substituted for a large number of cultural measures, such as the proportion of the population identified with specific religions (used by Sala-i-Martin).[73] It also largely eliminated any independent role for the constructed measure of educational quality reported in the preceding section. We included the institutional quality measure with the initial conditions, even though it is likely to be somewhat endogenous and determined by policy.[74] Unfortunately, we have no effective measure of the change in institutional quality between our two twenty-year subperiods, because our indicator is drawn from survey data for 1982.

Results of a regression that relates these six conditioning variables to growth in output per worker are reported in column 8-1 of table 8. Those

70. The significance of the population measure is not sensitive to the inclusion of China and India. The trade instrument is taken from Frankel and Romer (1999) and Frankel and Rose (2002). It is created by regressing the bilateral trade of countries *i* and *j* on distance between their principal cities, the extent of common borders, the presence or absence of a common language, land area, the population of the trading partner, and whether or not either country is landlocked. The predicted values are aggregated over all trading partners.

71. The percentage of tropical land area is from Gallup and Sachs (1998). The average number of frost days is from Masters and McMillan (2001). The two measures were converted to standard deviates and assigned equal weights. Noting that tropical land area and frost days are negatively and positively correlated with growth, respectively, our weights are –0.5 and +0.5.

72. Knack and Keefer (1995). Their variable is an equally weighted average of 1982 values for law and order, bureaucratic quality, corruption, risk of expropriation, and government repudiation of contracts. It is scaled from 0 to 1, with higher values representing better institutions.

73. Kaufman, Kraay, and Zoido-Lobatón (2002); Sala-i-Martin (1997).

74. All of our results were robust to the use of colonial mortality rates as an instrument for institutional quality, as suggested by Acemoglu, Johnson, and Robinson (2001). However, this instrument is available for only fifty-two of the countries in our sample.

Table 8. Regressions Explaining Growth and Its Components: Conditioning and Policy Variables, 1960–2000[a]

	Dependent variable								
	Growth in output per worker					*Contribution of capital per worker*		*Contribution of TFP*	
Independent variable	*8-1*	*8-2*	*8-3*	*8-4*	*8-5*	*8-6*	*8-7*	*8-8*	*8-9*
Constant	–2.27	–1.49	–2.94	–1.64	–0.57	–0.52	0.64	–1.12	–1.22
	(–2.2)	(–1.8)	(–2.7)	(–1.6)	(–0.5)	(–0.6)	(0.8)	(–1.3)	(–1.1)
Contribution of capital per worker		0.78							
		(6.9)							
Investment share			2.76						
			(1.6)						
Initial income per capita	–6.29	–4.02	–5.89	–6.24	–5.18	–2.89	–2.18	–3.35	–3.00
	(–10.4)	(–7.0)	(–9.2)	(–10.7)	(–8.9)	(–6.2)	(–4.8)	(–6.7)	(–5.3)
Life expectancy	0.07	0.04	0.06	0.06	0.06	0.02	0.02	0.04	0.04
	(5.6)	(4.4)	(4.9)	(5.0)	(4.1)	(2.4)	(1.6)	(3.6)	(2.9)
Log of population	0.29	0.16	0.29	0.28	0.19	0.16	0.08	0.12	0.11
	(4.8)	(3.0)	(4.8)	(4.7)	(3.0)	(3.4)	(1.6)	(2.4)	(1.8)
Trade instrument	4.77	2.26	4.53	3.55	2.51	2.24	1.66	1.31	0.85
	(4.1)	(2.3)	(3.9)	(3.0)	(2.2)	(2.4)	(1.9)	(1.3)	(0.8)
Geography	0.53	0.33	0.51	0.48	0.40	0.22	0.17	0.27	0.23
	(4.1)	(3.1)	(3.9)	(3.9)	(3.0)	(2.2)	(1.6)	(2.5)	(1.8)
Institutional quality	2.84	2.29	2.72	2.34	2.66	0.34	0.35	2.00	2.31
	(4.5)	(4.5)	(4.3)	(3.6)	(3.7)	(0.7)	(0.6)	(3.6)	(3.3)
Inflation				–0.01	0.00	0.00	0.00	0.00	0.00
				(–1.1)	(–0.9)	(–0.7)	(–0.4)	(–0.6)	(–0.6)
Budget balance				0.06	0.03	0.06	0.03	0.01	0.00
				(2.3)	(1.1)	(2.6)	(1.2)	(0.3)	(0.1)
Sachs-Warner openness				0.48	0.18	0.44	0.04	0.04	0.13
				(1.7)	(0.6)	(1.9)	(0.2)	(0.2)	(0.5)
Regional dummies included	No	No	No	No	Yes	No	Yes	No	Yes
Summary statistics									
Adjusted R^2	0.75	0.84	0.75	0.77	0.82	0.53	0.64	0.57	0.56
Standard error	0.72	0.57	0.71	0.69	0.61	0.55	0.48	0.59	0.60

Source: Authors' regressions using sources listed in table C1 in appendix C.

a. The number of observations in all regressions is eighty-four. Numbers in parentheses are *t* statistics.

variables account for three-fourths of the cross-country variation in growth over 1960–2000. All six of our conditioning variables are highly significant, the convergence variable especially so. Column 8-2 shows the effect of adding the contribution of capital as a regressor. Although this variable is obviously highly endogenous, it shows that the growth account measure of the capital contribution greatly improves the R^2. In contrast, inclusion of the investment rate (column 8-3) results in a statistically insignificant coefficient, supporting the conclusion above that it is a very poor proxy for the capital contribution.

In column 8-4 the model is expanded to include three policy indicators: the average rate of inflation, the government budget balance, and a measure of trade openness computed by Jeffrey Sachs and Andrew Warner.[75] All three of these measures have the expected sign, but only the budget balance is statistically significant at conventional levels. We note that the coefficient estimate and significance of the Sachs-Warner index are unaffected by exclusion of the trade instrument (results not shown).

The weak, negative role of inflation is particularly noteworthy in view of the emphasis frequently placed on it in policy discussions. However, our analysis examines the long-run association between inflation and growth, not the obvious short-run inverse relationship between inflation crises and growth. Following Michael Bruno and Easterly,[76] in a separate regression we allowed the set of ten countries with inflation rates more than 1 standard deviation above the mean to enter with a separate coefficient, but the coefficient was near zero and statistically insignificant.

As noted above, these policy indicators should be considered endogenous. However, we have no plausible instruments for inflation or the budget balance. Furthermore, the trade instrument is only weakly correlated with the Sachs-Warner openness indicator. When we removed it from the regression and attempted to use it as an instrumental variable for the Sachs-Warner indicator, it performed poorly in the first-stage regression. Therefore we present ordinary least squares results, which should be interpreted descriptively. Like growth accounts, these regressions cannot be used to infer the underlying causes of growth.

75. Sachs and Warner (1995). Although we recognize that researchers disagree on the best interpretation of the Sachs-Warner indicator, we considered it because so many other studies have used it as a trade policy measure.

76. Bruno and Easterly (1998).

We explored the significance of a large number of other potential explanatory variables, but we omitted them from the reported regressions because they did not play a role when our set of conditioning variables was also included. Table C2 in appendix C provides the complete list. In particular, we tested several measures of financial development from Levine, Norman Loayza, and Thorsten Beck.[77] Their preferred measure, the ratio to GDP of private credit extended by financial intermediaries, was available for just sixty-one of our countries. We computed a comparable series, covering only deposit banks, for eighty-one countries. The forty-year average of this variable is statistically significant, with a p value of 0.035, in a regression with the set of conditional variables used in column 8-1 of table 8, and marginally significant in the presence of the policy variables. However, we were concerned about the obvious endogeneity of this variable. In regressions that restricted the measure to its average value in the first ten years of the sample, it was very insignificant.[78] We were unable to identify other instruments that could be used to control more explicitly for the endogeneity problems. We also found no role for variations in the real exchange rate, and the standard deviation of the terms of trade (a measure of external shocks) was not consistently significant.

We also experimented with a number of alternative measures of trade openness, reflecting the extensive literature that has developed around the issue. In addition to the indicators discussed above, we tried various measures of the share of trade in GDP as well as an openness index from Dennis Quinn and Carla Inclán.[79] We found these measures to be positively correlated with growth when the number of other conditioning variables was limited, but inclusion of the full set of conditioning variables reduced the coefficient on the trade variable to near zero, sometimes turning it negative. This was particularly evident for regressions that included the measure of institutional quality. In this respect, our results are very similar to those of Dani Rodrik, Arvind Subramanian, and Francesco Trebbi (2002).[80]

77. Levine, Loayza, and Beck (2000).

78. We encountered a similar problem in the twenty-year samples. The financial depth variable was very insignificant when limited to the average of the first five years.

79. Dollar and Kraay (2003); Alcalá and Ciccone (2001); Quinn and Inclán (1997, 2001).

80. Rodrik, Subramanian, and Trebbi (2002). We also experimented with various instrumental variables estimates for the trade and institutions variables in ways that parallel

The implications for the channels through which the variables influence growth are shown in columns 8-6 and 8-8. By construction, the coefficient for each variable in column 8-4 is identically equal to the sum of its coefficients in columns 8-6 and 8-8. In most respects the results are in accord with our expectations. The convergence process is evident both through capital accumulation and through the efficiency of resource use. Similarly, the influences of geographical factors and population are equally evident through both channels. Life expectancy and, especially, the quality of institutions have relatively greater effects through the channel of TFP improvements. Variations in the budget balance have their primary impact on capital accumulation, presumably because budget deficits are a competing use of national saving. One surprise is that the correlation of both the trade instrument and trade openness with growth appears to operate through capital accumulation rather than through TFP. Much of the theoretical literature has emphasized the efficiency gains from trade.

Finally, the implications of including regional effects are shown in columns 8-5, 8-7, and 8-9 (individual coefficients and their significance not reported). These further reduce the significance of the policy variables. They have the largest impact on capital accumulation, with a significant positive effect for East Asia and negative effects for Latin America and Africa. However, none of the regional variables is statistically significant in the regression for TFP.

A striking aspect of these regressions is the relatively minor evidence of a direct role for conventional government policies. Instead the most important determinants of growth appear to be factors that cannot be changed substantially in the short run. We also stress that combining the growth account decomposition with regression analysis affords a focus on the determinants of TFP in a fashion that cannot be duplicated by the simple inclusion of the investment share as a regressor.

Results from the Twenty-Year Samples

Table 9 reports the regional means for six of our variables for each of the two twenty-year subperiods. The other three measures—institutional quality, the trade instrument, and geography—do not change across the

---

the work by Rodrik, Subramanian, and Trebbi and by Dollar and Kraay. These did not materially alter our results and are not reported.

Table 9. Means and Standard Deviations of Variables Used in the Growth Regressions, by Region and Subperiod, 1960–2000[a]

	Income per capita		*Life expectancy*		*Log of population*		*Budget balance*		*Inflation*		*Average Sachs-Warner openness*	
Region	*1960–80*	*1980–2000*	*1960–80*	*1980–2000*	*1960–80*	*1980–2000*	*1960–80*	*1980–2000*	*1960–80*	*1980–2000*	*1960–80*	*1980–2000*
Developing countries (62 countries)	0.17	0.18	49.90	58.84	15.76	16.27	–3.18	–3.96	10.90	22.68	0.19	0.37
	(0.11)	(0.13)	(9.67)	(9.32)	(1.48)	(1.49)	(2.77)	(4.06)	(10.94)	(28.45)	(0.35)	(0.37)
Africa (19)	0.11	0.09	41.59	48.92	15.52	16.06	–4.84	–5.26	8.33	18.27	0.07	0.15
	(0.09)	(0.09)	(5.93)	(6.71)	(0.91)	(0.94)	(2.95)	(2.46)	(4.23)	(13.00)	(0.24)	(0.30)
East Asia incl. China (8)	0.12	0.21	52.54	65.48	17.04	17.52	–1.53	–0.56	11.52	6.10	0.60	0.82
	(0.05)	(0.15)	(9.67)	(5.60)	(1.77)	(1.76)	(1.79)	(3.44)	(15.82)	(3.16)	(0.43)	(0.37)
China	0.04	0.04	36.32	66.84	20.32	20.70	n.a.	–1.04	0.97	6.84	0.00	0.00
Latin America (22)	0.25	0.24	55.44	63.83	15.19	15.67	–2.28	–4.07	14.50	37.40	0.16	0.39
	(0.11)	(0.11)	(7.64)	(6.11)	(1.24)	(1.28)	(2.15)	(5.23)	(14.42)	(41.32)	(0.31)	(0.27)
Middle East (9)	0.21	0.26	52.79	62.77	15.59	16.10	–2.92	–3.21	8.25	16.92	0.25	0.50
	(0.08)	(0.13)	(9.98)	(6.64)	(1.48)	(1.51)	(3.14)	(2.51)	(5.22)	(15.79)	(0.43)	(0.41)
South Asia (4)	0.08	0.07	47.04	56.38	17.85	18.34	–5.94	–5.63	7.39	8.75	0.03	0.09
	(0.02)	(0.02)	(8.97)	(8.04)	(1.55)	(1.57)	(1.77)	(3.96)	(2.02)	(1.88)	(0.05)	(0.19)
India	0.07	0.06	44.33	54.18	19.89	20.35	–3.93	–6.28	6.68	8.66	0.00	0.00
Industrial countries (22)	0.62	0.74	70.26	74.26	16.24	16.41	–0.84	–3.06	7.16	5.73	0.88	0.98
	(0.22)	(0.16)	(2.24)	(1.38)	(1.53)	(1.52)	(2.63)	(2.95)	(3.14)	(3.75)	(0.30)	(0.07)

Source: Authors' calculations using sources listed in table C1 in appendix C.

a. Variables in table 7 that are not listed here do not change from one period to the next. Numbers in parentheses are regional standard deviations.

two subperiods. Initial income, life expectancy, and population are measured at the beginning of each subperiod. All the other variables are averages over the subperiod.

Table 10 reports regression results for the two subperiods. The basic results for growth in output per worker are shown in columns 10-1 and 10-2. Overall they are quite similar to those reported for the full forty-year sample, although there is some decline in statistical significance.[81] Budget policy plays a less important role in the second period, whereas geography, institutional quality, the trade instrument, and the Sachs-Warner index all become more significant, with larger coefficients as well. These results are consistent with the view that trade and openness to trade became more important contributors to growth after 1980. The reduced role for the institutional quality variable in the first period may reflect the fact that all of the observations on that indicator are drawn from survey information for 1982; however, its statistical significance in the second period supports the argument that causation runs primarily from institutional quality to growth rather than the converse. The largest change between the two subperiods is in the size of the constant term, which shows a decline of 5 percentage points of growth between the first and the second subperiod.[82] Finally, the inclusion of the regional effects had little substantive impact, and they are not reported.

The corresponding channel regressions are reported in columns 10-5, 10-6, 10-8, and 10-9. Again, in most respects the results are consistent with those from the forty-year sample. The convergence process is evident in both capital accumulation and TFP, as are the effects of both life expectancy and population in the second period. However, the previously noted differences between the two subperiods in the relative roles of geography, institutional quality, and the trade instrument are all concentrated in the TFP component. Indeed, both geography and the trade instrument are statistically insignificant in the first period in the TFP regressions. The Sachs-Warner trade measure continues to be significant only in the regressions for capital accumulation.

81. The number of observations declines for the earlier subperiod because of missing values for the measures of fiscal balance and inflation. We imposed a requirement that values had to exist for at least half of the period for a specific observation to be included in the cross-country regression.

82. The net change is much less because of offsetting changes in the coefficients on other variables, such as the trade instrument and geography.

Table 10. Regressions Explaining Growth and Its Components: Conditioning and Policy Variables, 1960–80 and 1980–2000[a]

	Dependent variable									
	Growth in output per worker				*Contribution of capital per worker*			*Contribution of TFP*		
	1960–1980	*1980–2000*	*Pooled 1960–2000*	*Population-weighted 1960–2000*	*1960–1980*	*1980–2000*	*Pooled 1960–2000*	*1960–1980*	*1980–2000*	*Pooled 1960–2000*
Independent variable	*10-1*	*10-2*	*10-3*	*10-4*	*10-5*	*10-6*	*10-7*	*10-8*	*10-9*	*10-10*
Constant	0.28	–4.79	–0.90	1.41	–0.04	–1.81	–0.43	0.32	–2.99	–0.48
	(0.2)	(–2.8)	(–0.8)	(0.9)	(0.0)	(–2.0)	(–0.6)	(0.2)	(–2.1)	(–0.5)
Shift in constant			–2.14	–0.96			–0.73			–1.41
			(–7.5)	(–2.6)			(–3.8)			(–5.7)
Income per capita	–6.51	–7.42	–6.71	–7.98	–3.20	–2.69	–2.84	–3.32	–4.73	–3.87
	(–7.9)	(–6.8)	(–10.1)	(–11.8)	(–4.5)	(–4.8)	(–6.3)	(–4.3)	(–5.3)	(–6.7)
Life expectancy	0.07	0.07	0.07	0.14	0.02	0.03	0.02	0.05	0.04	0.05
	(3.7)	(3.2)	(5.2)	(6.6)	(1.1)	(2.8)	(2.6)	(2.9)	(2.1)	(4.0)
Log of population	0.25	0.35	0.29	0.29	0.15	0.20	0.17	0.09	0.16	0.13
	(2.8)	(3.8)	(4.6)	(3.6)	(2.1)	(4.0)	(3.9)	(1.1)	(2.1)	(2.3)
Trade instrument	3.46	6.54	2.29	–1.51	4.12	2.08	3.10	–0.66	4.46	–0.80
	(2.0)	(3.5)	(1.4)	(–0.5)	(2.8)	(2.2)	(2.7)	(–0.4)	(3.0)	(–0.6)

Shift in trade instrument			4.61	–0.35			–0.16			4.78
			(2.18)	(–0.09)			(–0.11)			(2.60)
Geography	0.37	0.71	0.24	0.49	0.41	0.15	0.21	–0.05	0.56	0.03
	(1.9)	(3.3)	(1.4)	(2.3)	(2.5)	(1.4)	(1.8)	(–0.3)	(3.2)	(0.2)
Shift in geography			0.59	0.55			0.13			0.46
			(3.36)	(2.71)			(1.11)			(3.02)
Institutional quality	2.09	3.78	2.74	0.86	0.22	0.69	0.42	1.88	3.09	2.31
	(2.2)	(3.4)	(3.8)	(0.8)	(0.3)	(1.2)	(0.9)	(2.1)	(3.4)	(3.7)
Inflation	–0.01	–0.01	–0.01	–0.01	–0.01	–0.01	–0.01	0.00	0.00	0.00
	(–0.5)	(–1.8)	(–2.2)	(–3.2)	(–0.6)	(–1.9)	(–1.8)	(0.0)	(–1.0)	(–1.2)
Budget balance	0.14	0.05	0.08	0.06	0.04	0.04	0.04	0.10	0.01	0.04
	(3.2)	(1.4)	(2.9)	(1.3)	(1.1)	(2.4)	(2.3)	(2.4)	(0.2)	(1.6)
Sachs-Warner openness	0.32	1.19	0.66	0.76	0.19	0.68	0.45	0.13	0.51	0.21
	(0.9)	(2.6)	(2.3)	(2.1)	(0.6)	(2.8)	(2.4)	(0.4)	(1.4)	(0.9)
Summary statistics										
Adjusted R^2	0.60	0.64	0.70	0.77	0.34	0.52	0.49	0.30	0.50	0.50
Standard error	0.96	1.12	1.05	1.10	0.83	0.58	0.71	0.91	0.91	0.91
No. of observations	77	84	161	161	77	84	161	77	84	161

Source: Authors' regressions using sources listed in table C1 in appendix C.

a. Numbers in parentheses are *t* statistics.

The regressions for the two subperiods seem quite similar in their basic conclusions, yet a test statistic for a structural change in the relationship between the two periods is highly significant. This significance comes largely from the constant term, geography, and the trade instrument. Allowing these three to vary between the two subperiods raises the p value in the test for structural change from 0.000 to 0.23. Similarly, for TFP we can strongly reject the null hypothesis of no structural change if only the constant term is allowed to shift, but we cannot reject this null if the coefficients on both geography and the trade instrument are allowed to shift also. (The corresponding p values are 0.00 and 0.27, respectively.) In contrast, when only the constant is allowed to vary, the test for structural change in all other parameters of the channel regression for capital's contribution yields a p value of 0.31.[83]

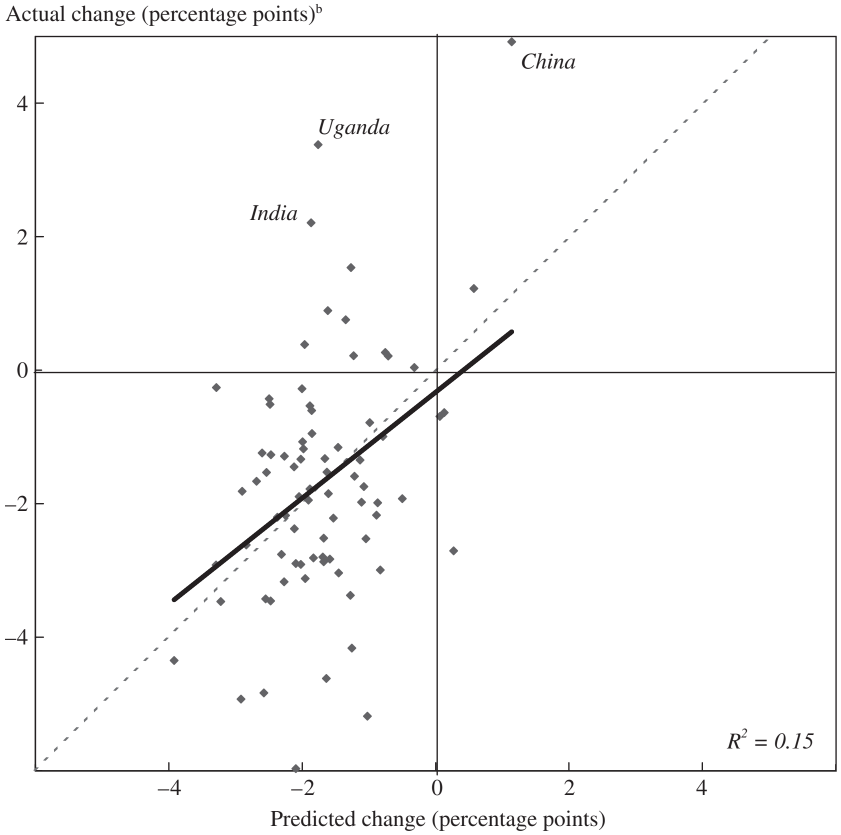
These points are even more evident in the pooled regressions shown in columns 10-3, 10-7, and 10-10, where we allowed for shifts in the constant and in the coefficients on the trade instrument and geography between the two subperiods. The exogenous decline in the growth rate is estimated at –2.1 percentage points, compared with the simple sample average of –1.7 percentage points reported in table 5. The shift in the growth rate is equally evident for both the contribution of capital per worker and TFP, but the changing roles of the trade instrument and geography are evident only in the TFP equation.

A final exploration of the stability of the statistical relationship is provided by the population-weighted regression reported in column 10-4. The weights will, of course, give a dominant role to the experiences of China and India, but a weighted regression provides a useful test of the stability of the specification. In this case there is still evidence of a large shift in the constant term and geography, but no role for the trade instrument or institutional quality. It also strengthens the negative role for inflation, while weakening the association with the budget.

Does this model account for the sharp changes in growth rates after 1980? We explore that issue by using the coefficients from the pooled regression in column 10-3 to calculate the expected change in growth between the pre- and the post-1980 period. For the eighty-four-country sample, the R^2 of the regression of predicted on actual changes, shown in

83. The instability of the coefficient estimates for the trade instrument and population raises questions about the interpretation of the role of these variables.

Figure 6. Actual and Predicted Change in Annual Growth of Output per Worker, 1960–80 and 1980–2000[a]



Source: Authors' calculations.

a. Data are for seventy-eight countries listed in appendix A for which data are available in the first period. Predicted values are generated from the pooled regressions in column 10-3 of table 10.

b. In logarithms.

figure 6, is only 0.15, and the variance of the predicted changes is only one-quarter that of the actual changes. Table 11 provides a regional summary in which the total predicted change in growth is divided into two pieces. The column labeled "Shift terms" separates out the combined effects of shifts in the constant and in the coefficients on geography and the trade instrument. (Recall that these variables are identical across subperiods.) The contribution of those variables that actually change between

Table 11. Actual and Predicted Changes in Growth of Output per Worker between 1960–80 and 1980–2000[a]

Region	Actual change	Predicted change: Total	Predicted change: Variables[b]	Predicted change: Shift terms[c]	Residual
Developing countries	–1.7	–1.7	0.4	–2.1	–0.1
Africa	–1.7	–2.0	0.4	–2.4	0.2
East Asia (excluding China)	–1.0	–1.5	0.2	–1.8	0.6
East Asia (including China)[b]	–0.2	–1.2	0.5	–1.7	1.0
Latin America	–2.4	–1.9	0.3	–2.2	–0.5
Middle East	–2.3	–1.3	0.3	–1.5	–1.1
South Asia	0.6	–1.5	0.5	–1.9	1.9
Industrial countries	–1.6	–1.7	–0.8	–0.9	0.2
25 countries with greatest increase in growth	0.4	–1.3	0.3	–1.6	1.6
China[b]	4.9	0.8	2.0	–1.2	4.1
India	2.2	–1.6	0.5	–2.1	3.8
Uganda	3.4	–1.8	0.8	–2.6	5.1
25 countries with greatest decrease in growth	–3.5	–2.0	0.0	–2.0	–1.6

Source: Authors' calculations.
a. Shift terms and coefficients are based on results of the regression reported in column 10-3 of table 10. Because China has no budget data for 1960–80, it is assumed that the change in the budget balance variable between the two periods was zero.
b. Contribution to predicted change from the change in value of variables that appear in both regressions.
c. Contribution to predicted change from the shift variables included in column 10-3 of table 10.

the two periods is reported in the column labeled "Variables." This makes it clear that most of the predicted change in the growth rates is coming from the negative shifts in the constant term and the coefficients on geography and the trade instrument.

The basic problem is that some of the most significant variables are those that do not change between the two periods, whereas the measures of policy, which do change over time, have small coefficients with limited statistical significance. The regression analysis has focused on identifying factors that are correlated with the cross-country variation in growth rates, but those same factors appear to do little to account for the variation in growth over time. The exceptions are China and the industrial countries. For China the predicted acceleration is coming from a large improvement in life expectancy between 1960 and 1980. The predicted slowing of growth in the industrial countries can be traced to a sharp deterioration in

Table 12. Growth Regressions by Level of Income per Capita, 1960–2000[a]

Independent variable	*Full sample*	*Developing countries*	*Higher-income countries*[b]	*Lower-income countries*[c]
Constant	–1.64	–2.19	–0.43	–1.68
	(–1.6)	(–1.4)	(–0.3)	(–0.7)
Income per capita	–6.24	–7.26	–5.08	–8.19
	(–10.7)	(–5.6)	(–7.7)	(–2.4)
Life expectancy	0.06	0.07	0.06	0.07
	(5.0)	(5.0)	(3.2)	(3.6)
Log of population	0.28	0.33	0.18	0.31
	(4.7)	(3.9)	(2.4)	(2.8)
Trade instrument	3.55	4.03	3.47	3.14
	(3.0)	(2.3)	(2.2)	(1.5)
Geography	0.48	0.61	0.55	0.62
	(3.9)	(4.2)	(3.2)	(2.9)
Institutional quality	2.34	2.60	2.13	2.43
	(3.6)	(3.5)	(1.8)	(2.6)
Inflation	–0.01	–0.01	0.00	–0.01
	(–1.1)	(–1.0)	(–0.4)	(–1.2)
Budget balance	0.06	0.06	0.04	0.04
	(2.3)	(1.6)	(0.8)	(1.1)
Sachs-Warner openness	0.48	0.69	0.07	1.21
	(1.7)	(2.1)	(0.2)	(2.8)
Summary statistics				
Adjusted R^2	0.77	0.80	0.77	0.80
Standard error	0.69	0.68	0.61	0.71
No. of observations	84	62	42	42

Source: Authors' regressions using sources listed in table C1 in appendix C.

a. The dependent variable is the average annual log change in output per worker times 100. Numbers in parentheses are *t* statistics.

b. Countries with income per capita above the sample median in 1960.

c. Countries with income per capita below the sample median in 1960.

public budgets and a reduced role for convergence after 1980. Finally, both East and South Asia have performed better than expected since 1980, while Latin America and the Middle East have fallen short.

Sensitivity to Country Groupings

Our sample includes a very heterogeneous group of countries. Do our averaged results apply to specific country groups? To some extent we explored this issue earlier in terms of the sensitivity of some of the results to population weighting. Here we extend the sensitivity analysis by estimating the relationships reported in tables 8 and 10 for various country subgroups. We report in table 12 the results for the sixty-two developing economies in our sample, the forty-two countries with income per capita in 1960 above the median ("higher-income countries" in the table), and the forty-two countries with 1960 income per capita below the median ("lower-income countries"). The full-sample results from table 8 are reproduced for comparison in the first column of the table. The regression results are strikingly similar across these groups. Convergence is somewhat more evident in the lower-income countries (as expected), and the Sachs-Warner openness measure seems least important for the richer countries (which was unexpected). Any effort to explore smaller, more specific samples resulted in serious problems of multicollinearity. However, for the nineteen African countries, the problems extended beyond multicollinearity: the adjusted R^2 in those regressions was only 0.25.

We can also combine the analysis of the subgroups with the examination of the twenty-year periods before and after 1980. As shown in table 13, the shifts in the intercept and in the coefficients on the trade instrument and geography variables between the two periods in the aggregate sample are all primarily due to shifts for the lower-income group. In addition, the convergence term for the lower-income countries becomes much more important in the second period. Somewhat surprisingly, the increased importance of institutional quality after 1980 is entirely due to a change for the higher-income countries.

Thus we find surprisingly small differences between determinants of growth between higher- and lower-income countries over our entire time period. However, we do find evidence that shifts in parameter estimates across time periods are sensitive to country groupings, with more substantive shifts for the lower-income half of the sample.

Conclusion

We conclude that, contrary to much of the recent literature, both growth accounting and growth regressions—the main tools for empirical analysis of cross-country differences in economic growth—can yield consistent and useful results. In addition, we have argued that much of the apparent variability in the conclusions from earlier studies can be traced to measurement problems, differences in data or definitions, and, in the regression analyses, failure to include other conditioning variables. To address some of these problems, we developed a set of growth accounts

Table 13. Growth Regressions by Level of Income per Capita, 1960–80 and 1980–2000[a]

	Higher-income countries		*Lower-income countries*	
Independent variable	*1960–80*	*1980–2000*	*1960–80*	*1980–2000*
Constant	–0.05 (0.0)	–1.41 (–0.7)	1.64 (0.4)	–6.13 (–1.7)
Income per capita	–6.28 (–7.1)	–5.24 (–3.9)	–4.35 (–0.8)	–11.67 (–2.5)
Life expectancy	0.09 (3.0)	0.08 (1.7)	0.06 (1.8)	0.09 (2.7)
Log of population	0.35 (3.4)	0.06 (0.5)	0.09 (0.5)	0.50 (2.7)
Trade instrument	4.46 (2.1)	4.47 (2.0)	2.13 (0.7)	12.17 (3.1)
Geography	0.72 (2.8)	0.60 (2.3)	0.36 (0.9)	0.78 (2.3)
Institutional quality	0.45 (0.3)	4.10 (2.3)	3.38 (1.9)	2.59 (1.7)
Inflation	–0.02 (–1.5)	0.00 (0.0)	0.01 (0.3)	–0.01 (–0.6)
Budget balance	0.11 (1.7)	0.05 (0.8)	0.13 (1.8)	0.02 (0.5)
Sachs-Warner openness	–0.31 (–0.7)	1.04 (1.3)	0.78 (1.2)	1.56 (2.6)
Summary statistics				
Adjusted R^2	0.65	0.64	0.53	0.72
Standard error	0.84	0.91	1.14	1.15
No. of observations	42	42	35	42

Source: Authors' regressions using sources listed in table C1 in appendix C.

a. The dependent variable is the average annual log change in output per worker times 100. Numbers in parentheses are *t* statistics.

for the period from 1960 to 2000 covering eighty-four countries, which together represent a preponderance of the world economy and of world population. Combining these data with additional variables allows us to examine a wide range of competing hypotheses over a common group of countries and common time periods.

Much of the debate and dissatisfaction with past empirical analyses centers around disputes over the relative contributions of capital accumulation and improvements in TFP and the importance of education. In both of these questions, measurement issues play a central role. With respect to the debate over capital accumulation versus TFP, we emphasize that both are important and that some of the earlier research understates the role of capital because of inadequate measurement of the capital input. In particular, we caution against the inappropriate use of an average of the investment rate as a proxy for the change in capital. Despite concerns about the assumed rate of capital depreciation, relatively simple measures of the change in the capital stock display a much stronger correlation with output growth than does the investment rate, yielding estimates of capital's contribution that are close to hypothesized values. We have identified two additional issues. First, some studies fail to recognize that measuring investment in international prices induces a positive correlation between investment and income, further compounding the problem. Second, some studies formulate the decomposition so as to focus on variations in the ratio of capital to output instead of capital per worker. We show that this definitional presentation change underplays the role of capital relative to changes in TFP.

We agree with the critics in finding only a weak correlation between economic growth and aggregate measures of improvements in educational attainment. However, rather than conclude that education does not matter, we stress the problems introduced by difficulties in accurately measuring cross-country variations in educational attainment and adjusting for differences in educational quality. We find a surprisingly low correlation among the alternative measures of changes in educational attainment. And although we find strong evidence that available indicators of educational quality are highly correlated with growth, this finding is not robust to the inclusion of broader indicators of institutional quality. We also note that even an optimistic valuation of the return to education would lead to only small differences in economic growth rates.

Within our framework, a very large portion of the cross-country variation in economic growth experiences over the past forty years can be related to differences in initial conditions and government institutions. In particular, the finding of a strong negative association between initial income and subsequent growth provides very robust support for a process of conditional convergence. Similarly, life expectancy in the initial year (as a measure of health) and population are positively associated with growth. There is also a strong correlation between growth and the quality of governing institutions (such as law and order, absence of corruption, and protection of property rights). Other variables that are consistently significant are geographical location (temperate versus tropical climate) and an indicator of a country's predisposition to trade.

In contrast, we found only limited evidence associating macroeconomic policies and the Sachs-Warner indicator of openness with growth. Equally notable, some factors often cited in the literature as important for growth did not display a consistent correlation with growth. For example, we experimented with a wide variety of alternative measures of trade openness and found their role to be insignificant in the presence of the other variables mentioned above. Much of the variation in growth experiences appears to be more closely tied to differences in initial conditions, rather than in the short-term policies of governments. In addition, although research has identified some of the factors responsible for cross-country variations in rates of economic growth, it has been far less effective in identifying the sources of change over time. Overall, we find that the variables important in accounting for differences among countries provide little insight into the change in growth rates from the twenty-year period before 1980 to that after 1980.

By combining growth regressions with growth accounting, we are able to explore the channels through which various determinants are related to economic growth. In particular, the accounting decomposition provides a much more informative way to focus on determinants of changes in TFP than the frequently adopted alternative of including investment rates as a regressor. We find that geography and, especially, initial income are related to growth through both channels. Thus both capital accumulation and TFP exhibit convergence. Changes in TFP are positively related to institutional quality and life expectancy. Capital accumulation is more closely associated with budget balance and, somewhat surprisingly, with measures of trade predisposition and openness.

Furthermore, some of the parameter estimates exhibit sensitivity to variations in the sample, especially when parameter shifts over time are compared for different country groups. Of particular interest is that indicators of geography and predisposition to trade appear to have become more important (especially for lower-income countries) since 1980. There is also considerably more evidence of catch-up for the poorer countries in that period.

In conclusion, we believe that the cross-country analysis provides some confirming evidence of the role of various contributors to growth. But it cannot stand alone, and it requires careful attention to measurement concerns. The disappointments are that the analysis yields surprisingly little insight into the sources of the widespread (except in China and India) slowing of growth observed after 1980, and that we find a relatively minor role for macroeconomic policies.

APPENDIX A

Country Sample

East Asia (8 countries)	*Industrial countries (22 countries)*	*Middle East and N. Africa (9 countries)*
China	Australia	Algeria
Indonesia	Austria	Cyprus
Korea	Belgium	Egypt
Malaysia	Canada	Iran
Philippines	Denmark	Israel
Singapore	Finland	Jordan
Taiwan	France	Morocco
Thailand	Germany	Tunisia
	Greece	Turkey
Latin America (22 countries)	Iceland	
Argentina	Ireland	*Sub-Saharan Africa (19 countries)*
Bolivia	Italy	
Brazil	Japan	Cameroon
Chile	Netherlands	Côte d'Ivoire
Colombia	New Zealand	Ethiopia
Costa Rica	Norway	Ghana
Dominican Rep.	Portugal	Kenya
Ecuador	Spain	Madagascar
El Salvador	Sweden	Malawi
Guatemala	Switzerland	Mali
Guyana	United Kingdom	Mauritius
Haiti	United States	Mozambique
Honduras		Nigeria
Jamaica	*South Asia (4 countries)*	Rwanda
Mexico	Bangladesh	Senegal
Nicaragua	India	Sierra Leone
Panama	Pakistan	South Africa
Paraguay	Sri Lanka	Tanzania
Peru		Uganda
Trinidad and Tobago		Zambia
Uruguay		Zimbabwe
Venezuela		

APPENDIX B

Measures of Educational Quality

THE ORIGINAL ANALYSIS by Hanushek and Kimko (2000) estimated a statistical relationship between their index of educational quality and a set of indicators from the Barro-Lee (1993) data set for thirty countries that participated in the testing. This relationship was then used to predict educational quality for an additional forty-nine countries, thirty-six of which are in our sample. That relationship is reported in the first column of table B1. We expanded the thirty-country sample to include Chile because we wanted to have at least two countries, Chile and Brazil, on which to base the placement of Latin American countries. In the Hanushek-Kimko series the Latin America measures are all calculated relative to Brazil and appeared to be implausibly high. The result of that addition is shown in the second column of table B1.[84] The right-hand-side variables, except population growth and educational attainment, were updated from the 2002 World Development Indicators and are average values over the period 1970–2000. Population growth and average years of schooling are both measured over the period 1960–2000. We were also able to add China, Mozambique, and Nigeria, for which data were not reported in the Barro-Lee data set. The resulting equation, which we used to construct the revised index of educational quality, is reported in the third column of the table. Finally, because of the correlation reported in the text between the measures of educational quality and the quality of government institutions, we show in the fourth column a regression for the thirty-four-country sample that includes the measure of institutional quality. It is highly significant, but it alters and reduces the role of several other variables.

The index of educational quality was extended to the remaining countries in our sample using the equation reported in the third column of table B1 and data drawn from the WDI. Two of the countries in the thirty-four-country sample, Swaziland and Hong Kong, are not in our sample. In the first column of table B2 we show the original Hanushek-Kimko index. For those countries that were not in their sample, we show estimates provided by Woessman (2000). His estimates are based on

84. The Hanushek and Kimko study excluded Chile because the test score data came from an earlier decade.

countries that are similar in regional distribution and income level. The second column reports our estimates based on the equation in the third column of table B1. Finally, the estimate of school quality using the quality of government institutions is presented in the third column.

Table B1. Regressions Explaining Educational Quality[a]

	Barro-Lee data		*WDI data*	
Independent variable	*B1-1*[b]	*B1-2*[c]	*B1-3*[d]	*B1-4*[e]
Constant	–28.40	–23.01	9.97	–47.98
	(–1.0)	(–0.6)	(0.5)	(–2.3)
Primary enrollment	73.28	68.71	15.78	24.87
	(2.5)	(1.7)	(0.9)	(1.7)
Education expenditure	170.37	155.52	96.68	–198.02
	(1.0)	(0.9)	(0.5)	(–1.2)
Population growth	–417.00	–398.13	–265.28	483.20
	(–1.6)	(–1.4)	(–1.2)	(2.0)
Education years	0.97	0.90	2.99	0.58
	(0.8)	(0.8)	(3.1)	(0.6)
Institutional quality				85.65
				(4.4)
East Asia dummy	13.77	13.15	15.22	10.85
	(2.9)	(2.0)	(2.6)	(2.30)
Latin America dummy	0.20	–12.65	–11.09	–9.67
	(0.0)	(–1.6)	(–1.5)	(–1.7)
Africa dummy	8.71	7.50	9.47	11.46
	(2.5)	(0.7)	(1.2)	(1.9)
Summary statistics				
Adjusted R^2	0.56	0.55	0.55	0.74
No. of observations	30	31	34	34
Mean predicted value of educational quality	40.6	35.2	33.1	25.4

Sources: Hanushek and Kimko (2000); authors' regressions using sources listed in table C1 in appendix C.

a. The dependent variable is Hanushek and Kimko's index of institutional quality.

b. Results as reported by Hanushek and Kimko (2000) for a sample of thirty countries.

c. Results from the same equation and data as column B1-1 but with the addition of Chile.

d. Results from the same equation as column B1-2 but using updated data from the 2002 World Development Indicators and adding China, Mozambique, and Nigeria.

e. Results from the same equation and data as column B1-3 but adding the institutional quality variable.

Table B2. Measures of Educational Quality

Country	*Original Hanushek-Kimko index*	*Without institutional quality variable*[a]	*With institutional quality variable*[a]	*Test countries*[b]	*Woessman extension*[c]
Algeria	28.1	24.9	28.5	No	No
Argentina	48.5	26.0	22.3	No	No
Australia	59.0	59.0	59.0	Yes	No
Austria	56.6	55.6	55.4	No	No
Bangladesh	43.0	5.3	2.8	No	Yes
Belgium	57.1	57.1	57.1	Yes	No
Bolivia	27.5	4.0	5.4	No	No
Brazil	36.6	36.6	36.6	Yes	No
Cameroon	42.4	0.2	42.5	No	No
Canada	54.6	54.6	54.6	Yes	No
Chile	24.7	24.7	24.7	Yes	No
China	64.4	64.4	64.4	Yes	No
Colombia	37.9	25.9	24.2	No	No
Costa Rica	46.2	32.1	35.2	No	No
Côte d'Ivoire	39.1	40.4	44.7	No	Yes
Cyprus	46.2	34.2	33.2	No	No
Denmark	61.8	54.5	57.9	No	No
Dominican Rep.	39.3	22.8	20.3	No	No
Ecuador	39.0	28.3	28.3	No	No
Egypt	26.4	23.6	25.8	No	No
El Salvador	26.2	–0.6	–1.6	No	No
Ethiopia	37.6	12.7	12.5	No	Yes
Finland	59.6	59.6	59.6	Yes	No
France	56.0	56.0	56.0	Yes	No
Germany	48.7	48.7	48.7	Yes	No
Ghana	25.9	35.8	34.9	No	Yes
Greece	50.9	41.4	37.1	No	No
Guatemala	40.1	3.4	1.5	No	Yes
Guyana	51.5	–3.9	0.8	No	No
Haiti	38.4	–11.1	–12.1	No	Yes
Honduras	28.6	13.4	14.5	No	No
Iceland	51.2	64.0	61.4	No	No
India	20.8	20.8	20.8	Yes	No
Indonesia	43.0	39.7	38.9	No	No
Iran	18.3	18.3	18.3	Yes	No
Ireland	50.2	50.2	50.2	Yes	No
Israel	54.5	54.5	54.5	Yes	No
Italy	49.4	49.4	49.4	Yes	No
Jamaica	48.6	13.1	17.6	No	No
Japan	65.5	65.5	65.5	Yes	No
Jordan	42.3	42.3	42.3	Yes	No
Kenya	29.7	42.3	46.1	No	No
Korea	58.6	58.6	58.6	Yes	No

(*continued*)

Table B2. Measures of Educational Quality (*continued*)

Country	*Original Hanushek-Kimko index*	*Without institutional quality variable*[a]	*With institutional quality variable*[a]	*Test countries*[b]	*Woessman extension*[c]
Madagascar	37.6	34.3	32.4	No	Yes
Malawi	37.1	32.4	31.2	No	Yes
Malaysia	54.3	51.8	56.2	No	No
Mali	37.9	2.4	3.3	No	Yes
Mauritius	55.0	52.3	51.2	No	No
Mexico	37.2	29.8	29.2	No	No
Morocco	35.8	32.1	34.8	No	Yes
Mozambique	27.9	27.9	27.9	Yes	No
Netherlands	54.5	54.5	54.5	Yes	No
New Zealand	67.1	67.1	67.1	Yes	No
Nicaragua	27.3	19.3	19.9	No	No
Nigeria	38.9	38.9	38.9	Yes	No
Norway	64.6	64.6	64.6	Yes	No
Pakistan	42.8	13.9	11.3	No	Yes
Panama	46.8	9.7	12.0	No	No
Paraguay	40.0	23.0	19.8	No	No
Peru	41.2	17.6	16.8	No	No
Philippines	33.5	33.5	33.5	Yes	No
Portugal	44.2	44.2	44.2	Yes	No
Rwanda	37.2	22.7	22.3	No	Yes
Senegal	39.1	24.2	25.5	No	Yes
Sierra Leone	37.6	16.6	14.7	No	Yes
Singapore	72.1	72.1	72.1	Yes	No
South Africa	51.3	52.0	54.2	No	No
Spain	51.9	51.9	51.9	Yes	No
Sri Lanka	42.6	22.8	20.8	No	No
Sweden	57.4	57.4	57.4	Yes	No
Switzerland	61.4	61.4	61.4	Yes	No
Taiwan	56.3	56.3	56.3	Yes	No
Tanzania	37.5	34.7	35.4	No	Yes
Thailand	46.3	46.3	46.3	Yes	No
Trinidad and Tobago	46.4	22.7	22.5	No	No
Tunisia	40.5	25.8	30.6	No	No
Turkey	39.7	38.1	35.4	No	No
Uganda	37.4	20.8	19.0	No	Yes
United Kingdom	62.5	62.5	62.5	Yes	No
United States	46.8	46.8	46.8	Yes	No
Uruguay	52.3	19.5	18.1	No	No
Venezuela	39.1	26.0	28.0	No	No
Zambia	36.6	33.0	32.2	No	No
Zimbabwe	39.6	38.6	43.7	No	No

Sources: Hanushek and Kimko (2000); Woessman (2000); authors' calculations.
a. Using Hanushek and Kimko data extended with updated WDI data.
b. "Yes" if Hanushek and Kimko data included test scores for the indicated country, "No" if test scores are predicted.
c. "Yes" if data in the first column came from Woessman's estimation, "No" otherwise.

APPENDIX C

Variables Used in the Analysis

Table C1. Variable Sources and Definitions

Variable	*Source and definition*
Investment	Domestic fixed investment in national prices is taken from the *OECD Statistical Compendium* for industrial countries and World Bank, World Development Indicators (WDI), for developing countries. Investment in international prices is taken from PWT 6.0.
GDP	Gross domestic product in real national prices, used for constructing the growth accounts, is taken from WDI and filled in with data from the *OECD Statistical Compendium*. Gross domestic product in real international prices, used for computing GDP weights and calculating the investment share, is taken from PWT 6.0.
Labor force	Economically active population, taken from WDI.
Educational attainment	Average educational attainment, in years, of the population aged 15 and over; data are averages of series from Barro and Lee (2000) and Cohen and Soto (2001). The annual average is used to construct the human capital index.
Population	Total population, used in constructing population weights, is taken from WDI.
Initial average years of schooling	Average educational attainment, in years, of the population aged 15 and over; data are averages of series from Barro and Lee (2000) and Cohen and Soto (2001). Initial year is 1960 or 1980.
Initial income per capita	Income per capita in 1960 or 1980 relative to the United States, from PWT 6.0 and WDI.
Life expectancy	In years, expressed as the difference in 1960 or 1980 from the U.S. level, from WDI.
Log of population	Natural logarithm of the total population, from WDI. Data are period averages.
Frankel-Romer-Rose trade instrument	Computed as the predicted values from a regression in which the bilateral trade share is related to a set of fixed characteristics and averaged over the trading partners, as explained in Frankel and Rose (2002). Higher values represent a greater predisposition to trade openness.
Geography	Average of frost days and tropical land area, from Rodrik, Subramanian, and Trebbi (2002). Measures are scaled by standard deviates. Higher values represent "better" geography.
Institutional quality	International Country Risk Guide assessment of institutional quality as of 1982, data from Knack and Keefer (1995). Higher values represent better institutional quality.
Budget balance	Average annual general government budget surplus or deficit as a percentage of GDP, from *OECD Statistical Compendium* for industrial countries, African Development Bank since 1980 for African countries, and WDI and International Monetary Fund data for all other countries.

(*continued*)

Table C1. Variable Sources and Definitions (*continued*)

Variable	*Source and definition*
Inflation	Average annual log change in the national consumer price index from International Monetary Fund, *International Financial Statistics*.
Sachs-Warner openness	Average years during the period in which the economy is "open," as determined by the openness dummy variable constructed by Sachs and Warner (1995).

Table C2. Additional Variables Used in Unreported Regressions

Variables tried in growth regressions

Financial depth
Private credit as share of GDP (Levine, Loayza, and Beck, 2000; WDI)

International integration
Current and capital account openness, levels and change (Dennis Quinn)
Capital account openness indicator (International Monetary Fund)
Trade as share of GDP, in real and nominal terms (WDI, PWT)

Exchange rate indicators
Real exchange rate, change and standard deviation (three measures constructed by the authors)
Average black market premium (Levine, Loayza, and Beck, 2000; Barro and Lee, 1994)

Educational indicators
Average years of education, levels and change (three measures: Cohen and Soto, 2001; Barro and Lee, 2000; Nehru and Dhareshwar, 1993)
Educational quality (three measures: Hanushek and Kimko, 2000; authors' calculations)

Social and political indicators
Index of ethnolinguistic fractionalization (Easterly and Levine, 1997)
Index of civil and political freedoms (Freedom House)
Population growth (WDI)
Revolutions (Rodrik, Subramanian, and Trebbi, 2002)
War casualties (Rodrik, Subramanian, and Trebbi, 2002)

Institutions
Government antidiversion policies (Hall and Jones, 1999)
Institutional quality measures (Kaufmann, Kraay, and Zoido-Lobatón, 2002)
- Corruption
- Government effectiveness
- Regulatory quality
- Rule of law
- Political stability
- Voice and accountability

Institutional Quality Composite Index components (International Country Risk Guide)
- Political risk
- Economic risk
- Financial risk

Constraint on the executive (Rodrik, Subramanian, and Trebbi, 2002)
Economic Organization Indicator (Hall and Jones, 1999)

(*continued*)

Table C2. Additional Variables Used in Unreported Regressions (*continued*)

Geography
Frost area (Masters and McMillan, 2001)
Days of frost a year (Masters and McMillan, 2001)
Latitude (Rodrik, Subramanian, and Trebbi, 2002)
Average temperature (Rodrik, Subramanian, and Trebbi, 2002)
Percentage of land in tropics (Gallup and Sachs, 1998)
Total land area (Gallup and Sachs, 1998)
Landlocked dummy (Gallup and Sachs, 1998)
Malaria index (Gallup and Sachs, 1998)

Variables tried as instruments

Share of population speaking European languages (Hall and Jones, 1999)
Predicted trade from Frankel-Romer gravity model (Frankel and Romer, 1999; Frankel and Rose, 2002)
Settler mortality (Acemoglu, Johnson, and Robinson, 2001)

Comments and Discussion

Steven N. Durlauf: In attempting to provide an "update" of the immense empirical literature on cross-country growth, Barry Bosworth and Susan Collins have assigned themselves an exceptionally ambitious task. As they note, the empirical growth literature is filled with conflicting claims and strong disagreements, whether one considers questions of econometric methodology, substantive conclusions on the predictors or determinants of cross-country growth differences, or even the appropriate ways to measure possible growth determinants. Bosworth and Collins argue that through careful attention to issues of variable selection and measurement, it is possible to develop a coherent perspective on cross-country growth determinants and thereby bring some clarity to the morass of empirical growth studies.

In this effort the authors are partially successful. They have provided a particular set of specifications of growth regressions which reflect many good judgments. On issues such as the appropriate measurement of capital accumulation, for example, Bosworth and Collins make a persuasive case that previous growth studies have employed a poor proxy for physical capital accumulation. I am also sympathetic with their conclusion that the worldwide slowdown in economic growth since 1980 cannot be well explained by the sorts of long-run factors typically employed in growth studies. However, as in so many studies of this type, a number of the authors' claims are overstated. More generally, despite suggestions the authors make to the contrary, the paper fails to engage, let alone address, the major criticisms that have been leveled against growth regressions.

My first example of how the authors overstate their claims concerns the use of growth accounting to understand cross-country growth differences.

The authors argue that growth accounting can provide an important complement to cross-country growth regressions. Although I am sympathetic with this claim, the paper does a relatively poor job of acknowledging the measurement problems inherent in such exercises. Specifically, the growth accounting exercises in the paper require strong empirical assumptions in order to compute the relative contributions of human capital accumulation, physical capital accumulation, and total factor productivity growth to cross-country growth variation, assumptions whose validity is questionable. And no evidence is provided that the results in the paper are robust if these assumptions do not hold.

One such assumption is that the labor share in income is constant across countries and is equal to 0.65. The only evidence provided for making this assumption for the full set of countries under study is a reference to a study by Douglas Gollin.[1] However, Gollin's calculations do not (in my view) show this; rather they show that there is no systematic relationship between labor shares and income per capita. This is true at two levels. First, Gollin provides two separate calculations of the labor share, one of which would imply a constant labor share (if that is what the data in fact indicate) of 0.65, whereas the other would imply a share of 0.75. Second, and more important, I do not think Gollin's results justify the authors' assumption. For the African countries in Gollin's sample, the two main ways to compute shares yield results as follows: Botswana, 0.37 and 0.34; Burundi, 0.91 and 0.73; Congo, 0.69 and 0.58; Côte d'Ivoire, 0.81 and 0.69; and Mauritius, 0.77 and 0.67. Showing that there does not appear to be any systematic relationship between income per capita and the labor share is not equivalent to showing that the labor share is constant. Further, there is a very large dispersion of labor shares for lower-income economies. Hence I see little justification for the assumption of a constant factor share, let alone for assuming a particular value.

The growth accounting exercises also require that the stock of human capital be measured. The authors measure a country's human capital H_i by the equation $H_i = (1.07)^{s_i}$, where s_i is the average number of years of schooling and 0.07 is the assumed return to a year of schooling. As indicated in a paper by Mark Bils and Peter Klenow,[2] which the authors use to justify their return-to-schooling assumption, it is far from clear that the

1. Gollin (2002).
2. Bils and Klenow (2000).

return to schooling is constant across countries at different stages of development.

A second example of overclaiming occurs with respect to the interpretation of results. A narrow example is found in the variance decompositions undertaken to show that both physical capital accumulation and TFP are important in understanding cross-country growth differences. I certainly agree with the conclusion, but it is unclear that the data the authors present add to the evidence for it found in other studies. One reason is that, in their variance decompositions, the authors understate the uncertainty that exists in measuring the contributions of particular components when the components are not orthogonal. As is well understood from the literature on vector autoregressions, these bounds are determined by the way in which one component is treated as causally determining another; so, if one wants to identify bounds on the variance contributions to z of x and y (where $z = x + y$ by definition), these bounds are determined by whether one attributes the covariance between the two components to x or to y. The authors' bounds, as described in a note to their paper, do not do this and in fact understate the bounds.[3] Further, the point estimates used in the variance decomposition exercises are not associated with standard errors or any measure of uncertainty. Hence, although it may be the case that the authors' data support an important role for both physical capital accumulation and TFP growth, the case is not made.

A more serious example concerns the way in which policies are evaluated. I will focus on the authors' discussion of education and growth. The paper makes an extended argument that the education data employed in cross-country growth regressions are very badly measured. They nevertheless report a set of growth regressions that include education variables, and for a couple of these regressions they report statistically significant coefficients for these variables. This analysis leads the authors to conclude as follows:

> We agree with the critics in finding only a weak correlation between economic growth and aggregate measures of improvements in educational attainment. However, rather than conclude that education does not matter, we stress the problems introduced by difficulties of accurately measuring cross-country vari-

3. To see this, suppose that $z = x + y$ and $x = y$; then Bosworth and Collins would bound the contributions of x and y between 25 percent and 75 percent, whereas the correct bounds are zero and 100 percent, depending on whether all covariance is attributed to the one variable or the other.

> ations in educational attainment and adjusting for differences in educational quality.

It is useful to ask to what extent this conclusion is better justified than one of the following form:

> Overall, we find no systematic evidence that education, measured in terms of either quantity or quality, affects aggregate growth. In light of the fact that microeconomic evidence on returns to schooling across countries does not incorporate general-equilibrium effects (let alone full treatment of issues of self-selection), we find no justification, on the basis of existing econometric evidence, to recommend educational improvements as a path to more rapid growth.

Can these two statements be distinguished by findings in the paper? I do not see any justification for preferring one to the other. Certainly the discussion of measurement error does not distinguish the two. (One should also note that the claims about measurement error are themselves overstated, in that, at best, all the authors show is that at least one of the two series they study is mismeasured.) In addition, taken on their own terms, the authors' conclusions are inconsistent with the way empirical evidence is assessed elsewhere in the paper. For example, statistical insignificance is used as the criterion for variable exclusion in most of the growth regressions, yet here the general absence of statistical significance of the education variables does not justify the conclusion that they are not important. And, of course, if the measurement error issues are as serious as the authors claim, this undermines the utility of the growth accounting supplement to growth regressions that the authors advocate.

Further, the discussion of policy here and elsewhere is flawed by a lack of integration between the statistical work and formal policy evaluation. Put differently, the authors fail to adopt a decision-theoretic approach to policy analysis; hence it is very difficult to know how to interpret their statistical work. Throughout, the authors use the statistical significance of coefficients to evaluate whether a policy matters or not, with no attention to whether this makes sense if one is actually solving a decision problem. To be fair, this is a limitation of virtually all empirical growth work.[4]

My broader reservations about this paper revolve around its methodological assertions. Bosworth and Collins argue that the approach taken in

4. Brock, Durlauf, and West (2003) discuss ways to achieve integration between econometric analysis and policy evaluation.

this paper addresses some of the many criticisms that have been levied against cross-country growth regressions. The empirical growth literature is filled with virtually as many specifications of growth regressions as there are studies, and it is well known that many claims about the relevance of particular growth determinants differ dramatically across studies. Bosworth and Collins argue that they are able to substantially overcome concerns over the instability of findings across studies:

> . . . this critique has gone too far. In fact, most of the variability in the results can be explained by variation in the sample of countries, the time period, and the additional explanatory variables included in the regression. We maintain that there is a core set of explanatory variables that has been shown to be consistently related to economic growth and that the importance of other variables should be examined conditional on inclusion of this core set.

However, there is *no* sense in which Bosworth and Collins directly answer the criticisms that have been made of growth regressions. Nor do they even establish the reasons for differences in findings between their paper and others: they do not reestimate previous models under alternative assumptions about data and model specifications, and thereby do not establish the reasons why their results differ from those of other papers. What they do, rather, is assert that they have identified a set of countries, a set of growth determinants, a set of measurement rules, and a time horizon, which are the appropriate baseline from which to analyze growth regressions.

One problem with the authors' approach is that they simply assert that certain features of the data they analyze are superior to other data sets that have been employed. For example, relative to the standard study by Gregory Mankiw, David Romer, and David Weil,[5] Bosworth and Collins omit the following countries from their data set: Afghanistan, Angola, Bahrain, Barbados, Benin, Botswana, Burkina Faso, Burma, Burundi, Central African Republic, Chad, Congo, Fiji, Gabon, Gambia, Guinea, Hong Kong, Iraq, Kuwait, Lesotho, Liberia, Luxembourg, Malta, Mauritania, Nepal, Niger, Oman, Papua New Guinea, Saudi Arabia, Somalia, Sudan, Surinam, Swaziland, Syria, Togo, United Arab Emirates, Yemen, and Zaire. On the other hand, unlike Mankiw, Romer, and Weil, they do include China. Perhaps the Bosworth and Collins country choices are "better" than those of Mankiw, Romer, and Weil or those made in other

5. Mankiw, Romer, and Weil (1992).

studies, but there is no systematic argument as to why this is so. There is some suggestion that the countries in Bosworth and Collins are more homogeneous; the authors also say they omitted very small countries. But this does not constitute, in my view, an adequate defense, nor does it elucidate differences across studies. And I fail to see why the introduction of an economy making a transition from communism (China) makes the sample more homogeneous than otherwise.

At a conceptual level, criticisms of cross-country growth regressions fall into two main categories.[6] First, endogeneity of growth determinants renders any causal inferences from a given growth regression problematic. Second, growth regressions suffer from model uncertainty, so that there are many possible specifications of a growth regression that are consistent with the body of modern growth theory. This model uncertainty takes three general forms. First, there is theory uncertainty: different growth regressions almost universally incorporate different combinations of growth theories. Why does this happen? Growth models are open ended in the sense that one growth mechanism typically has no necessary link to another;[7] hence the theory that civil liberties affect growth does not speak to the validity of the theory that trade openness affects growth. Second, there is specification uncertainty. The construction of empirical analogues to growth theories is often difficult; this is very obviously the case when the theories relate to cultural or political factors. In addition, many growth theories are nonlinear, leading to nontrivial specification questions about how to capture potential nonlinearities. Finally, there is heterogeneity uncertainty: different analyses take different stances on which countries are assumed to obey a common growth model. I will now examine how the Bosworth and Collins analysis addresses these criticisms.

Bosworth and Collins address the endogeneity of growth determinants in two ways. First, they distinguish between proximate and ultimate growth determinants. This means that they try to partially restrict themselves to variables for which endogeneity seems a second-order concern. Second, some limited work is done with instrumental variables.

However, both of these solutions are unsatisfying. The distinction between ultimate and proximate employed in variable selection suffers from

6. See Brock and Durlauf (2001), Durlauf (2000), and Brock, Durlauf, and West (2003) for elaborations of this argument.

7. This argument follows Brock and Durlauf (2001) and Durlauf (2000).

several problems. First, the division between ultimate and proximate causes seems arbitrary in two respects. Some variables, such as life expectancy, are not obvious candidates for ultimate causes; life expectancy, for example, presumably has something to do with equality and government quality, neither of which is an ultimate cause in the sense that climate is. Further, the authors' choice of ultimate causes is ad hoc; they exclude many ultimate causes that previous authors have identified, even though other studies have found them to be robust predictors of growth. For example, Carmen Fernandez, Eduardo Ley, and Mark Steel,[8] who have performed the most comprehensive assessment, assign a posterior model inclusion probability of 0.995 to a variable measuring the percentage of the population that is Confucian and a posterior model inclusion probability of 0.76 to a sub-Saharan Africa dummy variable, yet neither appears here, and each is as much an ultimate cause as the variables that are included.

The instrumental variables exercises are similarly unhelpful, as it is unclear why the instruments are valid. Bosworth and Collins employ three instruments, two of which (settler mortality and the Frankel-Romer measure of predicted trade based on a gravity model) are in essence geography-based instruments, and the third measures the percentage of a country's population speaking a European language. At first glance each would seem to be a valid instrument, since it is implausible to argue that either is "caused" by growth. However, that is not sufficient for an instrument to be valid. Validity also requires that the instrument be orthogonal to the residual in the growth model that is being analyzed. But this residual contains all growth determinants not included in the specification. Validity therefore requires that an instrument be orthogonal to these additional omitted growth determinants. Nothing in Bosworth and Collins's paper (or, to be fair, in other papers that have employed instruments of this type) gives any reason to believe this additional orthogonality condition is fulfilled. This problem has been misunderstood as asserting that instruments such as geographic measures are endogenous, but that is not what the argument says. The problem in fact derives from theory open-endedness: since the inclusion of certain theories does not preclude the causal role of others, the use of instrumental variables in growth regressions needs to account for their correlation with omitted growth determinants.

8. Fernandez, Ley, and Steel (2001).

Bosworth and Collins address theory uncertainty and specification uncertainty through their choice of variables to include in their reported regressions. Their criteria by which some variables are included and others are excluded are generally unclear or informal, or both. The choices of variables to include apparently derive from the prior conclusions of Fernandez, Ley, and Steel and of Xavier Sala-i-Martin that many growth variables are ex ante plausible,[9] and from the statistical significance of certain variable coefficients in this paper and in other studies.

Does this constitute an adequate treatment of theory-based and specification-based model uncertainty? At one level the authors' own empirical criteria suggest that it is not. The results obtained by Fernandez, Ley, and Steel and by Sala-i-Martin would imply that the current regressions are far too parsimonious; Sala-i-Martin concluded that twenty-five regressors should be included; Fernandez, Ley, and Steel concluded that there are twenty-four variables associated with ex post inclusion probabilities greater than 0.05, which is apparently the standard Bosworth and Collins use in evaluating that paper's findings to conclude that many variables are plausibly included in a growth regression. Hence the first evidentiary criterion used by Bosworth and Collins implies that they are working with misspecified models.

But leaving aside the question of internal inconsistencies in how Bosworth and Collins employ evidence in other studies, it is difficult to regard the model selection analyses conducted in this study as anything but ad hoc. In some cases the preferred set of included variables is determined by the standard that, given these variables, other variables do not seem statistically significant. Bosworth and Collins state, for example, that

> We also explored a number of alternative indicators of institutional quality . . . and obtained the most significant results with a composite variable. . . . [It] substituted for a large number of cultural measures, such as the proportion of the population identified with specific religions.

This sort of procedure does not correspond to a coherent way of engaging in model selection. Results based on an unsystematic search across alternative specifications, with a specification surviving when other variables do not augment it in the sense of producing statistically significant coefficients, will produce results that are path dependent (that is, that depend on

9. Fernandez, Ley, and Steel (2001); Sala-i-Martin (1997).

the order in which specifications are compared) and will be subject to pretest bias. Further, it is not even clear why the authors want to engage in model selection. If their objective is to identify how certain factors affect growth, then model uncertainty constitutes an important part of the overall uncertainty in the exercise. Put differently, employing a single model specification (or a small set of similar specifications) produces a systematic understatement of parameter uncertainty by failing to acknowledge that the specification itself is uncertain.[10] Therefore I see no reason to conclude that the choice of variables by Bosworth and Collins makes any progress in addressing the problems of theory uncertainty and specification uncertainty; rather, theirs simply represents one more cross-country growth regression specification to add to the many hundreds that have appeared.

Bosworth and Collins address heterogeneity uncertainty by comparing the parameters of growth regressions that are estimated over subsets of countries, typically comparing groupings of countries with high and low initial incomes. They conclude that there is relatively little evidence of parameter differences for the different groupings. This result is interesting, and I commend the authors for the exercise. However, it should be recognized that this approach to identifying heterogeneity in growth models is quite narrow. There is evidence from a variety of studies that the single linear model assumption is not appropriate in understanding cross-country growth differences.[11] The critical difference in these studies is that the search for multiple growth models is treated as a classification problem, not as an ex post check on a model after it has been subjected to various model selection procedures. So I suspect there may be more here than is reported.

In conclusion, this paper makes useful contributions toward clarifying issues of measurement of capital accumulation and on aspects of education. The paper also makes a good case that growth accounting exercises can supplement growth regressions in empirical work. However, many of the empirical claims should have been stated with more circumspection and with far greater analysis of the sensitivity of the results to particular modeling assumptions. At a methodological level, there is no reason to

10. Brock, Durlauf, and West (2003) argue this point, following ideas developed in Leamer (1978).

11. See, for example, Durlauf and Johnson (1995), Desdoigts (1999), and Canova (1999).

believe that any of the main problems associated with interpreting growth regressions have been addressed systematically, let alone solved. Nevertheless, the methodological weaknesses of the analysis and the imprecision in the discussion about the strength of some of the evidence should not obscure the range of valuable insights that the paper does provide.

Jeffrey A. Frankel: I am very sympathetic with the general approach of this paper and, for that matter, with the specifics as well, which in many ways are a statement of the state of the art in empirical growth analysis. I find little to criticize, although perhaps if I were a drama critic, I might wish for a more exciting conclusion.

I agree with the authors that much of the growth literature has set up some artificial all-or-nothing choices: convergence versus divergence, changes versus levels, factor accumulation versus total factor productivity growth, history versus policies, trade versus institutions, and so on. In each case the right answer is not all or nothing, but a balance of both. Let's call it "balanced growth theory."

The authors have it right on the convergence debate. I agree that, for most purposes, it is better to include initial income on the right-hand side of the regression equation along with the other variables—the conditional convergence specification. Given this specification, it does not matter if our left-hand-side variable is the end-of-sample level of income or, as in this study, the change in income. If the data have a strong opinion that the coefficient on initial income should be close to 0 (no convergence) or close to –1 (complete convergence), they will tell us. Usually the truth is in between, with a coefficient of, say, 0.7 on initial income, implying a 30 percent rate of conditional convergence over about thirty years).

The calculation I just mentioned assumed that initial income was expressed in logarithmic form. The authors have not done this. Instead initial income is expressed in level form, relative to initial income in the United States. They tell us at one point that their numbers imply 30 percent convergence since 1960, which is in line with others' estimates, but I would have preferred to be able to read this estimate directly from the reported coefficient.

I also agree that the whole debate about whether growth is determined by capital accumulation (as in the neoclassical growth model) as opposed to TFP growth has been a bit overdone. New growth theory was proclaimed to constitute an overthrowing of Robert Solow's theory; it was followed by a supposed neoclassical revival,[1] followed by an allegation that the neoclassical revival had gone too far. This is a little too fashion-conscious for me. I recall that the main conclusion of Solow's famous 1957 paper, "Technical Change and the Aggregate Production Function," was precisely that all the action was *not* in capital accumulation but rather (seven-eighths) in the residual.[2] I assume that the aficionados all recognize that the Solow residual "school of thought" is the opposite of the Solow growth model school of thought, but I am guessing that this confuses many of our students.

I do think that the decomposition into the role of factor accumulation and the role of TFP growth is useful. I just don't think we should be surprised or disappointed when Bosworth and Collins find that the shares are roughly half and half. This is what they find in the simple decomposition, as well as in the regression exercise where they ask to what extent the channel for the effects of various policies and other growth determinants runs through factor accumulation or through TFP.

I am particularly pleased to see that the authors find comparably large roles for capital formation and TFP as channels for conditional convergence. It stands to reason that both are important: On the one hand, differences in capital-labor ratios create differences in rates of return, which in turn promote equalization through such mechanisms as international capital flows, for the right countries, namely, those that are open and stable. Meanwhile, on the other hand, such countries can also be expected to catch up to the global productivity frontier through technology transfer and emulation of state-of-the-art techniques and management practices. Many authors mention only one channel of catch-up, to the exclusion of the other, but it seems obvious that both should be important.

Nevertheless, the authors support the finding of Alwyn Young that, in the case of the East Asian newly industrializing countries, the growth miracle was more a matter of factor accumulation and less one of TFP.[3] (Interestingly, however, the authors' table 1 suggests that in China it is the other way around.) This is a proposition that does matter and probably deserves the attention it has gotten. Paul Krugman popularized the finding in his famous or infamous 1994 article "The Myth of the Asian Economic

1. Mankiw, Romer, and Weil (1992).
2. Solow (1957).
3. Young (1995).

Miracle."[4] People outside of the economics profession were shocked at what seemed to be Krugman's claim that Asia's miraculous rise from poverty to prosperity, in the span of a few decades, had been an illusion. At the time, I was slightly amused by the reaction that his article created in the world of international affairs. The article could just as easily have been titled, "The Asian Economic Miracle Is Mostly Due to Saving, Education, and Urban Migration," in which case nobody would have taken much note of it.

The authors consider their most striking finding to be that there is only "minor evidence of a direct role for conventional government policies. Instead the most important determinants of growth appear to be factors that cannot be changed substantially in the short run." They are referring to the tendency of inflation, budget balance, and trade distortions to lose most of whatever statistical significance they might have had when one controls for such deeper determinants as life expectancy, geography, and institutions. I will concentrate the remainder of my comments in this area.

Their finding ties in well with some other important recent research, as well as with some current trends in the practice of aid and development policy in Washington. The current trend is to say, not that such policies as macroeconomic discipline and openness are not important, but that countries cannot be artificially forced from the outside to agree to such policies, as they are under typical International Monetary Fund (IMF) or World Bank programs. Instead the country needs to "take ownership" of the reforms. If the political economy dictates transfers from rural farmers to urban workers, or if a federalist constitution gives provinces a claim on income tax revenue, an agreement on paper with the IMF or World Bank to devalue the currency or reduce the budget deficit may be doomed to fail. This is the argument of a recent paper by Daron Acemoglu and coauthors.[5] They find econometrically that institutions are more powerful than policies in explaining growth, which is consistent with the finding of the present paper. They also use a case study of Ghana to illustrate how the impact of an IMF-encouraged devaluation, aimed at raising the real price of traded goods such as cocoa, can quickly be offset by the governing elite, because the cocoa marketing board controls the price paid to the small inland farmers for cocoa.

4. Krugman (1994).
5. Acemoglu and others (2003).

But institutions are not the only candidates for deeper determinants in growth equations. The question is well framed in a recent paper by Dani Rodrik, Arvind Subramanian, and Francesco Trebbi.[6] The rendition that follows is similar to theirs. Three big theories of deep determinants seem to have emerged, based on tropical conditions, openness, and institutions. Each has been captured by some now-standard measures. Although each of these three factors may be more exogenous than macroeconomic policies, each also has serious endogeneity problems of its own that must be addressed (table 1).

I prefer to use the phrase "tropical conditions" for what some, including the authors, have taken to calling geography. By now, geography has (belatedly) made its way deep into the literatures on trade and growth in many different ways. So it is important to clarify here what sort of geography one means. We are talking about the natural climate, biology, and geology, and especially differences between the tropics and the temperate zones, such as the presence or absence of malaria and other debilitating tropical diseases, the presence or absence of agricultural pests, the length of the growing season, and other climate effects.[7]

By "openness" we mean international integration along several dimensions, but trade is the most important. A common measure is the simple ratio of trade to GDP.

Finally, measures of "institutions," or institutional quality, are usually indicators of the rule of law, protection of property rights, and the extent of constraints on the executive power. The theory is that weak institutions lead to intermittent dictatorship, a lack of constraints on elites and politicians seeking to plunder the country, and hence low incentives for investment. Bosworth and Collins use an average of indicators from the International Country Risk Guide covering law and order, bureaucratic quality, corruption, risk of expropriation, and government repudiation of contracts.

I mentioned that each of the three has serious endogeneity problems. Fortunately, reasonable instruments have been proposed and implemented for each. The presence of malaria can be partly endogenous: it was stamped out in Panama and Singapore, despite their tropical locations, by superior technology and social organization. The instrumental

6. Rodrik, Subramanian, and Trebbi (2002).

7. Diamond (1997); Gallup, Sachs, and Messenger (1998); Hall and Jones (1999); and Sachs (2001).

Table 1. Deep Determinants of Growth

Determinant	*Measures*	*Sample endogeneity problems*	*Exogenous instrumental variables*
Tropical conditions	Malaria and other diseases, crop pests, length of growing season	Suppression by humans of malaria or pests	Distance from equator, tropical area, temperature, rainfall, frost days
Openness	Trade as share of GDP, tariffs, foreign direct investment	Imported investment or luxury goods, endogenous tariffs	Gravity model including remoteness, landlockedness, and linguistic and historical links
Institutions	Property rights, rule of law	Regulatory systems develop with income; ratings may be subjective	European settler mortality rates, extractive industries (plantation crops and mining)

Sources: Acemoglu, Johnson, and Robinson (2001); Acemoglu and others (2003); Easterly and Levine (2003); Engerman and Sokoloff (1997, 2002); Hall and Jones (1999); Rodrik, Subramanian, and Trebbi (2002); Sachs (2003).

variables adopted to capture the exogenous component of the tropical geography theory started out fairly crude but have been getting progressively better: moving from continental dummy variables to latitude, and from there to the percentage of land area in the tropics, to average temperature or number of frost days. The state of the art must be Jeffrey Sachs' latest measure of ecological predisposition to malaria, since, as director of the Earth Institute at Columbia University, he now has an army of biologists to figure it out for him.[8] But the measure that Bosworth and Collins use, a composite of tropical area and frost days, should be adequate.

Trade and trade policies are both endogenous. This is why David Romer and I proposed an instrumental variable: geographical suitability for trade as predicted by the gravity model.[9] This variable includes such exogenous determinants of trade as remoteness from large potential trading partners and landlockedness. This instrumental variable idea has been widely accepted.

8. Sachs (2003).
9. Frankel and Romer (1999).

Institutions can also be endogenous. Many institutions—such as the structure of financial markets, mechanisms of income redistribution, social safety nets, and regulatory and tax systems—tend to evolve in response to the level of income. But here the problem is thornier. Not only are institutions themselves likely to be endogenous, but the measures available are subjective evaluations of institutions. I submit that if you asked international businesspeople to rate the quality of Switzerland's fire departments compared with Colombia's fire departments, the Swiss would come out on top, even if they do not deserve to, because of the halo effect of Switzerland's general reputation. My point is that reported evaluations of institutional quality are likely to be endogenous with respect to national economic performance.

What is a good instrumental variable for institutional quality? Acemoglu, Simon Johnson, and James Robinson proposed the mortality rate among European settlers (more precisely, among soldiers and clergymen) during the period of initial colonization.[10] This is a better instrument than it sounds. In fact, it is probably the best we have. The theory is that, out of all the lands that Europeans colonized, only those where Europeans actually settled were given good European institutions. This theory is related to the idea of Stanley Engerman and Kenneth Sokoloff that lands endowed with extractive industries and plantation crops (mining, sugar, cotton) developed institutions of slavery, inequality, dictatorship, and state control, whereas those climates suited to fishing and small farms developed institutions based on individualism, democracy, egalitarianism, and capitalism.[11] Acemoglu, Johnson, and Robinson chose their instrument on the reasoning that initial settler mortality rates determined whether Europeans subsequently settled in large numbers.

Bosworth and Collins find that including the institutions variable tends to reduce the significance of policy variables, even when the Acemoglu instrument is used for institutions. A string of earlier authors have reached essentially the same finding: institutions drive out the effect of policies, and geography matters primarily as a determinant of institutions.[12]

10. Acemoglu, Johnson, and Robinson (2001).

11. Engerman and Sokoloff (1997, 2002).

12. These include Acemoglu and others (2003), Easterly and Levine (2003), and Hall and Jones (1999). Easterly and Levine simply group openness together with other policies. Hall and Jones consider latitude a proxy for European institutions and thus do not distinguish the independent effect of tropical conditions.

Nobody denies the important role of, for example, macroeconomic stability; rather the claim is that macroeconomic policies are merely the outcome of institutions. The conclusion has been phrased most aggressively by Rodrik, Subramanian, and Trebbi in the title of their paper: "Institutions Rule." Institutions trump everything else—the effects of both tropical geography and trade, in this view, pale in the blinding light of institutions. Bosworth and Collins, however, find that tropical geography remains significant against this onslaught, thus aligning themselves with Sachs, whose title retorts that "Institutions Don't Rule."[13]

I was also pleased to see that the authors found a significant role for predisposition to trade, as determined by the gravity model. Controlling for size (that is, population) is particularly important, indeed in my view essential, if one is to capture the effect of trade. Small countries are poor countries, all else equal. One of the reasons for the success of the U.S. economy is that our fifty states stretch from sea to sea and enjoy free trade with each other, creating an internal market large enough to achieve scale economies and exploit diverse endowments of natural resources and other factors. When smaller countries rely on international trade, it is as a second-best alternative to the preferred strategy, which is to be large. If a growth equation included the fifty American states as independent observations, their ratios of trade to GDP would be much higher than the ratio for the federal union; if the equation neglected to include size as another variable, it thus would erroneously predict higher levels of state income per capita than the national average. The opportunity to trade with one's fellow citizens is even more important than the opportunity to trade with people on the other side of the border. (Note that Bosworth and Collins report an even stronger relationship for size than for predisposition to trade.) Indeed, some authors who neglect to include size estimate a significant negative coefficient on openness, because the inverse correlation between size and openness is so strong.

I am not overly concerned that Bosworth and Collins did not find a big role for the Sachs-Warner measure of trade policy. I have always accepted the argument that the strategy of using trade barriers such as tariffs in a growth regression is not necessarily a solution to the problem of the endogeneity of trade.[14] One problem with it is that countries tend to move

13. Sachs (2003).
14. Sala-i-Martin (1991).

away from tariff revenue as a source of public finance as they industrialize. Further, the Sachs-Warner measure of trade distortions is a bit idiosyncratic. Francisco Rodríguez and Rodrik find that this measure is driven overwhelmingly, not by tariffs or quotas, but by the black market premium for foreign exchange and a measure of state export monopoly, and they argue that these largely reflect policies not related to trade.[15]

Any of these instrumental variables—tropical geography, gravity, or settler mortality—could in theory also be endogenous, notwithstanding that we have already pushed back some distance in the direction of exogeneity, from policies to social structure, and then to history and geography. The quality of instrumental variables is largely in the eye of the beholder, and I am repeatedly surprised at how some beholders see some things. The proper test of the ex ante plausibility of one's claim that a variable is a good candidate for an instrumental variable, in the sense that it is econometrically exogenous, is not whether one can think of a story whereby it *might* be correlated with other independent variables, but rather how convoluted and implausible the story has to be.

At a conference seven or eight years ago, a discussant took issue with the claim by Romer and me that the geographical variables specified in the gravity model made a good instrument for a country's propensity to trade. We said Botswana predictably has a relatively low ratio of trade to GDP because of its remote location, landlockedness, and high ratio of land to population. I think the discussant's point was that such geographic variables, although predetermined, might be correlated with the error term. He told a story roughly along the lines that a country with a large land area was more likely to have higher military spending, which in turn would result in slower economic growth. Over the years I have often used this as a methodological example: I tell my students that you know you have a relatively good instrument according to how convoluted is the story that the discussant has to tell about its potential endogeneity.

When I have related this incident, I have omitted the name of the discussant so as not to embarrass an accomplished macroeconomist because of what I assumed was a day when he had not had sufficient time to prepare comments that scored more effectively. Recently, however, I discovered that he had been telling a similarly one-sided story, about how foolish Frankel and Romer were for thinking that national borders were

15. Rodriguez and Rodrik (2001).

necessarily exogenous just because they were predetermined, as a way of illustrating the pitfalls of choosing instrumental variables.[16]

The point is not that the story about military spending, or many other possible stories like it, could not be true. One can of course tell such a story about any growth equation, or almost any regression equation at all for that matter. The question is how plausible or likely the story sounds to the paper's readers. And the lesson is that these judgments can be in the eye of the beholder.

With that caveat I will state my subjective ratings for the three sets of instruments in use for the three big categories of fundamental growth factors: tropical diseases and pests, trade, and institutions. I still believe that the trade predictions of the gravity model are a relatively good instrument for a country's openness to trade. The instruments available for tropical diseases and pests are even better.

The big challenge is institutions. I do not wish by any means to denigrate the importance of institutions. And, as I said, the settler mortality instrument is probably the best we have. I think everyone should use it until something better comes along, and I regret that I have never used it. But I view it as not as good as the instruments for trade and tropical conditions. For one thing, it is available only for former colonies. And there is another problem that I regard as more important. What are the big questions we are trying to answer? We already knew, long ago, that Australia, Canada, New Zealand, and the United States belonged with Europe in the list of countries that had industrialized. The big question is why they did and other countries outside Europe did not.[17] Is it policies, institutions, culture, or something else? In that light, to be told that the areas where Europeans settled did well is not exactly news. It simply repeats the big data point we already had. It does not help us all that much in choosing among policies, institutions, and cultures.

For my final point, I will move on from the topic of econometric exogeneity in historical data ("What makes a good instrument?") to the different question of conceptual exogeneity in the analysis of alternative future policies ("What would be the effect of a hypothetical reform?")

16. The discussant was Steven Durlauf; see Brock and Durlauf (2001).

17. There were exceptions to the rule: the failure of Argentina during the twentieth century and the success of Japan, the failure of Eastern Europe during the last third of the century and the success of the East Asian countries. But not everyone agrees over what lessons to draw from these cases.

What will economists say when called on to answer policy questions—such as what are the best currency arrangements—in a circumstance such as postwar Iraq? Are fundamental changes in policies, structure, and institutions possible, politically and socially? Or is all predetermined by history and geography? Are statistics from past history a guide to the consequences of future policy changes?

Needless to say, one can never be sure that a statistical or econometric pattern that characterizes the data in the past will continue to hold in the future. Particularly if one is talking about the future consequences of a deliberate change in policy, there are cases where one can even predict that the pattern of behavior will shift, after one has thought about it carefully. The Lucas critique of changes in the monetary policy regime is one famous example, but there are many more. Nevertheless, if we can *never* use past experience to predict the consequences of some innovation in policy, we might as well give up and go home. Much as with the choice of exogenous instruments within the sample period of historical data, there is little substitute for thoughtful deliberation and judicious choice in extrapolating to lessons for future policy changes. Of Rodriguez and Rodrik's critiques of Frankel-Romer, one was hard to prove wrong: even if the gravity variable is exogenous, there is no guarantee that changes in openness due to deliberate tariff policy will have identically the same effects on growth as does variation in openness due to geographical determinants of transport costs. The point is potentially valid in theory. When I think of all the arguments arrayed on both sides of the debate over trade, however, I do not believe that a debate between globalizers and antiglobalizers over the benefits of increases in trade due to reduced transport costs would be very different from a debate over increases in trade due to reduced government barriers. Evidence on one question is relevant for the other. The same is true of policy reforms in areas other than trade.

The example of settler mortality rates highlights how deeply rooted institutions can be and how infrequently and slowly they tend to change. But notwithstanding historical influences, institutions can change, and sometimes quickly. Most institutional change happens at a time of national upheaval, such as the end of a war or the birth of an independent country. We have all been reflecting recently on how successfully Japan and Germany were remade after the end of World War II. The breakup of the colonial empires in the 1950s through the 1980s offered another opportunity, which some countries seized much better than others. In the early 1990s the ruins of the Soviet Union left an opportunity for building new institutions in many transition economies; although that process appeared frustratingly slow and erratic at the time, ten years later it has begun to look better. Finally, today such new countries as East Timor and Macedonia, or countries in upheaval such as Afghanistan and Iraq, are open to guidance on institutional design from the international community, more than were the countries that became independent with the original breakup of the big colonial empires several decades ago.

The point is that even if macroeconomic or trade policies have on average been prisoner to slowly changing institutions and their historical or geographical determinants over a particular sample period, that is not necessarily a reason to despair of the possibility of genuine policy changes in the future, or of seeking to guide such changes by the light of our discipline. True, it is useful to keep one's eyes open, and to realize that well-intentioned policies may turn out instead to be the slaves of defunct nineteenth-century colonizers. But we academic scribblers must do our part as well.

General discussion: Benjamin Friedman noted that although the authors recognized that the residual in their growth equations may reflect a broad array of unmeasured factors affecting output, and thus may be a poor measure of technical change, in many places the paper followed the conventional practice of labeling the residual as a measure of total factor productivity (TFP). He suggested instead using a label that would remind readers that what is being discussed is, after all, a residual—for example, INI, for "influences not identified."

Richard Cooper aligned himself with Steven Durlauf's concerns about heterogeneity, expressing deep skepticism about the value of cross-country regressions when the countries in the sample are so diverse. As far as Cooper could tell, the only things these countries have in common are a seat in the United Nations and the fact that each is represented by a single color on a map of the world. He did not find satisfying the authors' attempt to ameliorate this problem by including a variable that putatively measures the quality of institutions, entered into the regression as if it were just another linear input. In his view institutional quality affects the importance of every other variable included in such regressions—for example, the rate of return on capital and the quality of education. That being the case, Cooper found it hard to know how to interpret the estimated coefficients on capital and education. Friedman suggested that the "institutions" variable may be a proxy for a much broader range of differences across countries, differences that noneconomists call "culture." If that is so, it would be a mistake to think that, by simply improving the institutions captured in the measure, a country would reap the estimated benefits to growth. Jeffrey Frankel mentioned that some economists are studying the effect of culture on economic performance: for example, Robert Barro and Rachel McCleary have recently examined how differences in countries' religious heritage affect growth.

Several panelists raised questions about the role of the "geography" variable in the growth equations. Cooper viewed economists' treatment of geography in recent years as very naïve. For example, countries adjacent to rich countries tend to be richer than countries adjacent to poor countries. Is it geography or neighbors that matter? He also suggested that the inclusion of geography, like the inclusion of institutions, made it difficult to interpret the coefficients on the other variables in the equations. Friedman did not think it obvious that variables that do not change over time, in this case geography, should be in a regression attempting to explain changes over time in another variable. Frankel reiterated his belief that although many different factors are packed together under the name "geography," careful use of instrumental variables can distinguish among them. Carmen Reinhart demurred, arguing that Frankel's proposed instruments for geography could be correlated with institutions, and vice versa.

Some panelists thought the paper understated the importance of inflation for growth. Miguel Savastano noted that although the estimated coefficient on inflation was relatively small, it was statistically significant in the pooled regressions and nearly so for the 1980–2000 subperiod. He thought even these results might understate inflation's importance because of its collinearity with the Sachs-Warner measure of openness or with other variables in the regression. Reinhart argued that inflation has an important negative effect on growth only when it is above a threshold, so that the authors' linear specification gives a misleading impression of the costs of hyperinflation.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91(5): 1369–1401.

Acemoglu, Daron, and others. 2003. "Institutional Causes, Macroeconomic Symptoms: Volatility, Crises and Growth." *Journal of Monetary Economics* 50(1): 49–123.

Alcalá, Francisco, and Antonio Ciccone. 2001. "Trade and Productivity." CEPR Discussion Paper 3095. London: Centre for Economic Policy Research (December).

Ashenfelter, Orley, Colm Harmon, and Hessel Oosterbeek. 1999. "A Review of Estimates of the Schooling/Earnings Relationship, with Tests for Publication Bias." *Labour Economics* 6(4): 453–70.

Baier, Scott, Gerald P. Dwyer, Jr., and Robert Tamura. 2002. "How Important Are Capital and Total Factor Productivity for Economic Growth?" Working Paper 2002-2a. Federal Reserve Bank of Atlanta (April).

Barro, Robert J. 1991. "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics* 106(2): 407–33.

———. 2001. "Education and Economic Growth." In *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being,* edited by John Helliwell. Paris: Organization for Economic Cooperation and Development.

Barro, Robert J., and Jong-Wha Lee. 1993. "International Comparisons of Educational Attainment." *Journal of Monetary Economics* 32(3): 363–94.

———. 1994. "Sources of Economic Growth." *Carnegie-Rochester Conference Series on Public Policy* 40: 1–46.

———. 2000. "International Data on Educational Attainment, Updates and Implications." Working Paper 7911. Cambridge, Mass.: National Bureau of Economic Research (September).

Barro, Robert J., and Xavier Sala-i-Martin. 1995. *Economic Growth*. McGraw-Hill.

Becker, Gary S. 1964. *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*. New York: National Bureau of Economic Research and Columbia University Press.

Benhabib, Jess, and Mark M. Spiegel. 1994. "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data." *Journal of Monetary Economics* 34(2): 143–73.

Bils, Mark, and Peter J. Klenow. 2000. "Does Schooling Cause Growth?" *American Economic Review* 90(5): 1160–83.

Bosworth, Barry P., Susan M. Collins, and Yu-chin Chen. 1996. "Accounting for Differences in Economic Growth." In *Structural Adjustment and Economic*

Reform: East Asia, Latin America, and Central and Eastern Europe, edited by Akira Kohsaka and Koichi Ohno. Tokyo: Institute of Developing Economies.

Brock, William A., and Steven N. Durlauf. 2001. "Growth Empirics and Reality." *World Bank Economic Review* 15(2): 229–72.

Brock, William A., Steven N. Durlauf, and Kenneth D. West. 2003. "Policy Evaluation in Uncertain Economic Environments." *BPEA 1:2003,* 235–301.

Canova, Fabio. 1999. "Testing for Convergence Clubs in Cross-Country Growth Data: A Predictive Density Approach." CEPR Discussion Paper 2201. London: Centre for Economic Policy Research (August).

Center for International Comparisons. 2001. *Penn World Tables Version 6.* University of Pennsylvania.

Cohen, Daniel, and Marcelo Soto. 2001. "Growth and Human Capital: Good Data, Good Results." Technical Paper 179. Paris: Organization for Economic Cooperation and Development (September).

Collins, Susan M., and Barry P. Bosworth. 1996. "Economic Growth in East Asia: Accumulation versus Assimilation." *BPEA 2:1996,* 135–203.

De la Fuente, Ángel, and Rafael Doménech. 2000. "Human Capital in Growth Regressions: How Much Difference Does Data Quality Make?" Working Paper 262. Paris: Organization for Economic Cooperation and Development (October).

———. 2001. "Educational Attainment in the OECD, 1960–1995." Institúto de Análisis Económico and Universidad de Valencia (January).

Deaton, Angus. 2003. "Measuring Poverty in a Growing World (or Measuring Growth in a Poor World)." Working Paper 9822. Cambridge, Mass.: National Bureau of Economic Research (July).

Desdoigts, Alain. 1999. "Patterns of Economics Development and the Formation of Clubs." *Journal of Economic Growth* 4(3): 305–30.

Diamond, Jared. 1997. *Guns, Germs and Steel: The Fates of Human Societies.* New York: W. W. Norton.

Dollar, David, and Aart Kraay. 2003. "Institutions, Trade, and Growth." *Journal of Monetary Economics* 50(1): 133–62.

Durlauf, Steven N. 2000. "Econometric Analysis and the Study of Economic Growth: A Skeptical Perspective." In *Macroeconomics and the Real World,* edited by Roger E. Blackhouse and Andrea Salanti. Oxford University Press.

Durlauf, Steven N., and Paul A. Johnson. 1995. "Multiple Regimes and Cross-Country Growth Behavior." *Journal of Applied Econometrics* 10(4): 365–84.

Durlauf, Steven N., and Danny T. Quah. 1999. "The New Empirics of Economic Growth." In *Handbook of Macroeconomics*, edited by John B. Taylor and Michael Woodford. Amsterdam: North-Holland.

Easterly, William. 2001. *The Elusive Quest for Growth: Economists' Adventures and Misadventures in the Tropics.* MIT Press.

Easterly, William, and Ross Levine. 1997. "Africa's Growth Tragedy: Policies and Ethnic Divisions." *Quarterly Journal of Economics* 112(3): 1203–50.

———. 2001. "It's Not Factor Accumulation: Stylized Facts and Growth Models." *World Bank Economic Review* 15(2): 177–219.

———. 2003. "Tropics, Germs, and Crops: How Endowments Influence Economic Development." *Journal of Monetary Economics* 50(1): 3–39.

Engerman, Stanley L., and Kenneth L. Sokoloff. 1997. "Factor Endowments, Institutions, and Differential Paths of Growth among New World Economies: A View from Economic Historians of the United States." In *How Latin America Fell Behind: Essays on the Economic Histories of Brazil and Mexico, 1800-1914,* edited by Stephen Haber. Stanford University Press.

———. 2002. "Factor Endowments, Inequality, and Paths of Development among New World Economies." Working Paper 9259. Cambridge, Mass.: National Bureau of Economic Research.

Englander, A. Steven, and Andrew Gurney. 1994. "OECD Productivity Growth: Medium-Term Trends." *OECD Economic Studies* 22: 111–54.

Fernandez, Carmen, Eduardo Ley, and Mark F. J. Steel. 2001. "Model Uncertainty in Cross-Country Growth Regressions." *Journal of Applied Econometrics* 16(5): 563–76.

Frankel, Jeffrey A., and David Romer. 1999. "Does Trade Cause Growth?" *American Economic Review* 89(3): 379–99.

Frankel, Jeffrey, and Andrew Rose. 2002. "An Estimate of the Effect of Common Currencies on Trade and Income." *Quarterly Journal of Economics* 117(2): 437–66.

Gallup, John Luke, and Jeffrey D. Sachs with Andrew D. Mellinger. 1998. "The Economic Burden of Malaria." Harvard Institute for International Development (October).

Gollin, Douglas. 2002. "Getting Income Shares Right." *Journal of Political Economy* 110(2): 458–74.

Griliches, Zvi. 1977. "Estimating the Returns to Schooling: Some Econometric Problems." *Econometrica* 45(1): 1–22.

Hall, Robert E., and Charles I. Jones. 1999. "Why Do Some Countries Produce So Much More Output Per Worker Than Others?" *Quarterly Journal of Economics* 114(1): 83–116.

Hanushek, Eric A., and Dennis D. Kimko. 2000. "Schooling, Labor-Force Quality, and the Growth of Nations." *American Economic Review* 90(5): 1184–1208.

Harrison, Ann. 2003. "Has Globalization Eroded Labor's Share? Some Cross-Country Evidence." University of California, Berkeley (May).

Heston, Alan. 2001. "Treatment of China in PWT 6." University of Pennsylvania (December).

Hsieh, Chang-Tai, and Peter J. Klenow. 2003. "Relative Prices and Relative Prosperity." Working Paper 9701. Cambridge, Mass.: National Bureau of Economic Research (May).

Hulten, Charles R. 1975. "Technical Change and the Reproducibility of Capital." *American Economic Review* 65(5): 956–65.

———. 2001. "Total Factor Productivity: A Short Biography." *New Developments in Productivity Analysis* (Studies in Income and Wealth, vol. 63), edited by Charles R. Hulten, Edward R. Dean, and Michael J. Harper. University of Chicago Press.

Islam, Nazrul. 1995. "Growth Empirics: A Panel Data Approach." *Quarterly Journal of Economics* 110(4): 1127–70.

Jones, Charles. 1997. "Comment on 'The Neoclassical Revival in Growth Economics: Has It Gone Too Far?'" In *NBER Macroeconomics Annual 1997,* edited by Ben S. Bernanke and Julio Rotemberg. MIT Press.

Jorgenson, Dale W. 2001. "Information Technology and the U.S. Economy." *American Economic Review* 91(1): 1–32.

Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobatón. 2002. "Governance Matters II: Updated Indicators for 2000–01." Working Paper 2772. Washington: World Bank, Policy Research Department (February).

Kenny, Charles, and David Williams. 2001. "What Do We Know about Economic Growth? Or, Why We Don't Know Very Much." *World Development* 29(1): 1–22.

Klenow, Peter J., and Andrés Rodríguez-Clare. 1997. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" In *NBER Macroeconomics Annual 1997,* edited by Ben S. Bernanke and Julio J. Rotemberg. MIT Press.

Knack, Stephen, and Philip Keefer. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Institutional Measures. " *Economics and Politics* 7(3): 207–27.

Krueger, Alan B., and Mikael Lindahl. 2001. "Education for Growth: Why and For Whom?" *Journal of Economic Literature* 39(4): 1101–36.

Krugman, Paul. 1994. "The Myth of Asia's Miracle." *Foreign Affairs* 73(6): 62–78.

Leamer, Edward E. 1978. *Specification Searches*. New York: John Wiley and Sons.

Levine, Ross, and David Renelt. 1992. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82(4): 942–63.

Levine, Ross, Norman Loayza, and Thorsten Beck. 2000. "Financial Intermediation and Growth: Causality and Causes." *Journal of Monetary Economics* 46(1): 31–77.

Lindauer, David L., and Lant Pritchett. 2002. "What's the Big Idea? The Third Generation of Policies for Economic Growth." *Economía* 3(1): 1–29.

Lubotsky, Darren, and Martin Wittenberg. 2001. "Interpretation of Regressions with Multiple Proxies." Industrial Relations Section Paper 457. Princeton University, Department of Economics (August).

Mankiw, N. Gregory, David Romer, and David N. Weil. 1992. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107(2): 407–37.

Masters, William A., and Margaret S. McMillan. 2001. "Climate and Scale in Economic Growth." *Journal of Economic Growth* 6(3): 167–86.

Nehru, Vikram, and Ashok Dhareshwar. 1993. "A New Database on Physical Capital Stock: Sources, Methodology and Results." *Revista de Análisis Económico* 8(1): 37–59.

Nehru, Vikram, Eric Swanson, and Ashutosh Dubey. 1995. "A New Database on Human Capital Stock in Developing and Industrial Countries: Sources, Methodology, and Results." *Journal of Development Economics* 46(2): 379–401.

Nuxoll, Daniel A. 1994. "Differences in Relative Prices and International Differences in Growth Rates." *American Economic Review* 84(5): 1423–36.

Oliner, Stephen D., and Daniel E. Sichel. 2000. "The Resurgence of Growth in the Late 1990s: Is Information Technology the Story?" *Journal of Economic Perspectives* 14(4): 3–22.

Organization for Economic Cooperation and Development. 2003. *Sources of Growth in OECD Countries*. Paris.

Pritchett, Lant. 2001. "Where Has All the Education Gone?" *World Bank Economic Review* 15(3): 367–91.

Psacharopoulos, George. 1994. "Returns to Investment in Education: A Global Update." *World Development* 22(9): 1325–43.

Quinn, Dennis P., and Carla Inclán. 1997. "The Origins of Financial Openness: A Study of Current and Capital Account Liberalization." *American Journal of Political Science* 41(3): 771–813.

Quinn, Dennis P., Carla Inclán, and A. Maria Toyoda. 2001. "How and Where Capital Account Liberalization Leads to Economic Growth." Paper prepared for the 2001 Annual American Political Science Association Convention, August.

Rodriguez, Francisco, and Dani Rodrik. 2001. "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence." In *NBER Macroeconomics Annual 2000,* edited by Ben S. Bernanke and Kenneth S. Rogoff. Cambridge, Mass.: National Bureau of Economic Research.

Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. 2002. "Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development." Working Paper 9305. Cambridge, Mass.: National Bureau of Economic Research (October).

Sachs, Jeffrey D. 2001. "Tropical Underdevelopment." Working Paper 8119. Cambridge, Mass.: National Bureau of Economic Research (February).

———. 2003. "Institutions Don't Rule: Direct Effects of Geography on Per Capita Income." Working Paper 9490. Cambridge, Mass.: National Bureau of Economic Research (February).

Sachs, Jeffrey D., and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *BPEA 1:1995,* 1–96.

Sala-i-Martin, Xavier. 1991. "Comments on: Growth, Macroeconomics and Development." In *NBER Macroeconomics Annual 1991,* edited by Olivier J. Blanchard and Stanley Fischer. MIT Press.

———. 1997. "I Just Ran Two Million Regressions." *American Economic Review* 87(2): 178–83.

Solow, Robert. 1957. "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics* 39(3): 312–20.

Soto, Marcelo. 2002. "Rediscovering Education in Growth Regressions." Technical Paper 202. Paris: OECD Development Centre (November).

Spence, A. Michael. 1973. "Job Market Signaling." *Quarterly Journal of Economics* 87(3): 355–74.

Temple, Jonathan R. W. 1999. "The New Growth Evidence." *Journal of Economic Literature* 37(1): 112–56.

———. 2001. "Generalizations That Aren't? Evidence on Education and Growth." *European Economic Review* 45(4–6): 905–18.

Woessmann, Ludger. 2000. "Specifying Human Capital: A Review, Some Extensions, and Development Effects." Working Paper 1007. Kiel, Germany: Kiel Institute of World Economics (October).

Wu, Harry X. 2002. "How Fast Has Chinese Industry Grown? Measuring the Real Output of Chinese Industry, 1949–97." *Review of Income and Wealth* 48(2): 179–204.

Young, Alwyn. 1994. "Lessons from the East Asian NICs: A Contrarian View." *European Economic Review* 38(3-4): 964–73.

———. 1995. "The Tyranny of Numbers: Confronting the Statistical Reality of the East Asian Growth Experience." *Quarterly Journal of Economics* 110(3): 641–80.

———. 2000. "Gold into Base Metals: Productivity Growth in the People's Republic of China During the Reform Period." Working Paper 7856. Cambridge, Mass.: National Bureau of Economic Research (August).