Schooling, Labor-Force Quality, and the Growth of Nations

Eric A. Hanushek; Dennis D. Kimko

The American Economic Review, Vol. 90, No. 5 (Dec., 2000), 1184-1208.

Stable URL:
http://links.jstor.org/sici?sici=0002-8282%28200012%2990%3A5%3C1184%3ASLQATG%3E2.0.CO%3B2-2

The American Economic Review is currently published by American Economic Association.

Schooling, Labor-Force Quality, and the Growth of Nations

By ERIC A. HANUSHEK AND DENNIS D. KIMKO*

Direct measures of labor-force quality from international mathematics and science test scores are strongly related to growth. Indirect specification tests are generally consistent with a causal link: direct spending on schools is unrelated to student performance differences; the estimated growth effects of improved labor-force quality hold when East Asian countries are excluded; and, finally, home-country quality differences of immigrants are directly related to U.S. earnings if the immigrants are educated in their own country but not in the United States. The last estimates of micro productivity effects, however, introduce uncertainty about the magnitude of the growth effects. (*JEL* O40, I20, J24)

Recent theoretical analyses of international differences in growth rates have focused attention on the role of human capital. Most cross-country empirical studies of long-run economic growth now include some proxy for human capital, and these are invariably significant. Data limitations have, however, forced severe compromises. Paralleling analyses of wage determination, empirical implementation virtually always employs some readily available measure of the quantity of formal schooling to reflect human capital, but this appears inadequate. The analysis of international differences in growth rates here suggests that math and science skill is a primary component of human capital relevant for the labor force. Such cognitive skill of a population is not well proxied by measures of school quantities or measures of resources devoted to schools. Accounting for differences in labor-force quality significantly improves our ability to explain growth rates.

Two issues arise in considering the effect of human capital on economic growth: how should any relationship be specified and how should human capital be measured? The focus of this paper is the second issue. This paper does not consider alternative formulations of the underlying growth relations but instead is a direct application of models of endogenous growth developed theoretically by a variety of people (e.g., Richard R. Nelson and Edmund Phelps, 1966; Paul Romer, 1990a; Sergio Rebelo, 1991). In the simplest formulation, growth rates are affected by ideas and invention, which in turn are related to the stock of human capital either through research and development (R&D) activities or through adoption behavior. These formulations indicate not only why the level of output is higher when a country has more human capital but also why the growth rate is higher.

Previous investigations of growth have concentrated on various measures of formal schooling activities as proxies for relevant human capital. The most frequently employed measure is either the primary- or secondary-school enrollment rate, used, for instance, in Romer (1990b), Robert J. Barro (1991), and N. Gregory Mankiw et al. (1992) and highlighted in the influential sensitivity studies of Ross Levine and David Renelt (1992) and Levine and Sara J. Zervos (1993). These schooling flow variables, however, will not accurately represent either the relevant stock of human capital of the labor force or even changes in the stock during periods of educational and demographic transition. To deal with these problems, Barro and Jong-Wha Lee (1993) pioneered the development of better schooling stock variables through the use

* Hanushek: Hoover Institution, Stanford University, Stanford, CA 94305, and National Bureau of Economic Research; Kimko: Institute for Defense Analyses, 1801 North Beauregard Street, Alexandria, VA 22311. We have benefited from comments and discussion with Mark Bils, Byron Brown, Stanley Engerman, Douglas Hodgson, Peter Klenow, Paul Romer, Sergio Rebelo, Michael Wolkoff, two very helpful referees, and participants in seminars at Boston University, Columbia University, the Rand Corporation, the University of Oregon, and the University of Rochester. Jin-Yeong Kim provided excellent research assistance.

of individual country survey and census data. A challenging problem made clear by this alternative, however, comes from the lack of adjustment for schooling quality. Few people, for example, would believe that a year of secondary schooling in the United States was equivalent to a year at the same grade level in Egypt. Indeed, Barro (1991) explored the inclusion of differences in real school resources as a crude measure of quality differences across countries in his growth regressions. While he found that the student-teacher ratio in primary schools in 1960 had a negative relationship to economic growth, the student-teacher ratio in secondary schools was statistically insignificant with a positive sign—giving little faith that these adequately captured any quality differences in schools.

There is also a conceptual issue with many human-capital formulations of growth models, since continued growth arising from human capital frequently requires continued growth in human capital. Yet, for simple investment reasons one does not expect that years of schooling will expand in an unbounded manner. When phrased in terms of cognitive skills and human-capital quality, however, continual quality growth is more natural, and the underlying growth models are much more readily interpreted.

This paper addresses the measurement problem of labor-force quality directly. Rather than concentrating on conventional measures of schooling inputs, we construct new measures of quality based on student cognitive performance on various international tests of academic achievement in mathematics and science. Labor-force quality differences measured in this way prove to have extremely strong effects on growth rates.

Direct observations of cognitive skills are available for 39 countries which participated in international assessments of student achievement at least once, although only 31 countries also have the measurement of economic performance that is needed for subsequent growth estimation. These quality measures can be extended to other countries by imputing missing values from international test score regressions. In both the subset with direct testing and in the augmented data set, the same qualitative conclusion about growth holds: Labor-force quality is significant with the correct sign even when school quantity tends to lose significance. The quality measures are also quite robust in the Levine and Renelt (1992) sense of being impervious to the precise empirical specification. Moreover, these quality measures are particularly important for explaining which countries are at the top and at the bottom of the distribution of economic growth rates.

Investigations of the empirics of growth such as this are necessarily subject to a variety of important questions and qualifications related to the stylized characterizations of different economies and to ambiguity about the underlying causal structure. While instrumental variables strategies for dealing with the causality issues are impractical, we pursue three different strategies that overall suggest we have identified important elements of the causal structure of growth. First, if stronger growth leads nations to increase investment in schools, growth could cause increased achievement. But direct estimation of international production functions does not support this as an avenue of effect. Second, if some unmeasured characteristics of countries affect both the performance of schools and the performance of other sectors in the economy, the observed relationship could be spurious. But, estimation of U.S. earnings models for immigrants that relate to country of schooling and our cognitive estimates of labor-force quality indicate that our measures of quality are directly related to labor-force skills and productivity of individuals. Third, the tests could just be pinpointing the high growth of East Asian countries that also score highly on international tests. The growth results here, however, hold for samples of countries which exclude various subsets of East Asian countries, although with somewhat lesser force.

The one caution of the tests involves the magnitude of estimated quality impacts on growth. The micro productivity estimates for immigrants, which have more modest earnings impacts of quality differences than found in the growth equations, suggest uncertainty about the potential magnitude of causal growth effects. Depending on how the level effects of productivity differences are translated into growth effects, the quality impacts could be noticeably smaller than those estimated in the growth equations—raising the possibility of other omitted factors. Nonetheless, even with uncertainty

about magnitudes, we conclude that labor-force quality is directly related to productivity and growth.

I. The Measurement of Labor-Force Quality

Qualitative descriptions of human capital, when considered, generally come from one of two sources: measures of schooling inputs (such as expenditure or teacher salaries) or direct measures of cognitive skills of individuals.[1] Employing direct cognitive skill measures has the significant advantage of permitting quality differences to arise from factors outside of formal schools, while employing measures of school inputs has potential advantages if important aspects of the relevant human capital are only partially measured by cognitive tests. While we consider both alternatives, the heart of this analysis is the development and use of a consistent set of cognitive test measures of quality.

The comparison of cognitive achievement across countries capitalizes on six voluntary international tests of student achievement in mathematics and science that were conducted over the past three decades.[2] Four were administered by the International Association for the Evaluation of Educational Achievement (IEA) and two by International Assessment of Educational Progress (IAEP).[3] The IEA, since its establishment in 1959, has a long and unique role in developing comparative education research for almost all aspects of primary and secondary education. On the other hand, the IAEP, starting in 1988, builds on the statistical techniques and procedures developed in the United States for the National Assessment of Educational Progress (NAEP), the main national testing instrument in the United States since 1969. While the IAEP is geared to the U.S. curriculum, the IEA has an international focus not associated with the curriculum in any particular country.

The concentration on mathematics and science corresponds to the theoretical emphasis on the importance of research and development activities as the source of growth (e.g., Romer, 1990a). Able students with a good understanding of mathematics and science form a pool of future engineers and scientists. At least for the United States, John Bishop (1992) provides separate confirmation of the importance of mathematics in determining individual productivity and income. Additionally, while some test information exists for other subjects, it cannot be compared readily with the mathematics and science scores and therefore is not used here.[4]

To develop a single measure of labor-force quality, we combine all of the information on international mathematics and science tests available for each country through 1991. The testing does not directly measure skills of the stock of workers (who received their schooling at varying times), although the mixture of the different tests employed here does approximate the relevant time range. While data are hard to come by, the usual presumption is that quality of schooling systems evolves slowly over time, in part related to the stationarity of the teaching technology and to the slow turnover of teachers and other personnel.[5] The approach of combining scores in order to estimate the aggregate skills of the labor force does, nonetheless, preclude any consideration of how estimated quality and growth evolve over any subperiods.

Two approaches were taken to combining the separate tests available for each country.

[1] As an alternative, Dale W. Jorgenson and Barbara M. Fraumeni (1992) calculate the human-capital stock from analysis of lifetime earnings (by education, age, and sex categories) and, going even further, adjust for the value of nonmarket activities.

[2] The Third International Mathematics and Science Study (TIMSS) was conducted in 1995 but is not used in the direct quality measurement. Its results are related to our labor-force quality measures below (U.S. Department of Education, 1996a, b).

[3] Details of participating countries, test administration, and sample sizes can be found in Hanushek and Kim (1995). Lee and Barro (1997) expand international quality measures by including reading and literacy scores along with more recent TIMSS data. We do not include reading and literacy because of concerns about valid testing across languages and doubts about putting these scores into a common one-dimensional scale with science and mathematics tests.

[4] Reading literacy assessments, for example, are available for 30 countries in 1991 (U.S. Department of Education, 1994).

[5] The test performance in Figure 1 provides some evidence about the stability of scores. The United States and United Kingdom participate in all six testing programs. Throughout the period, the United Kingdom consistently performs a little better than the United States. Further, with a few exceptions, countries that outperform either the United States or United Kingdom on one test also tend to do so when they participate in other tests and vice versa.

Twenty-six performance series—reflecting different age, subtest scores, and years—are available (for varying subsets of countries), and these series have different mean percent correct. The first summary method uses a multiplicative transformation to convert each performance series to a mean of 50. This transformation relies on a strong assumption that the intertemporal mean in world science and mathematics performance is constant and that the countries taking the tests are a random draw from the world distribution. The second method incorporates the additional information that is provided by the time series information from the National Assessment of Educational Progress (NAEP). At varying times from 1969 to the present, U.S. students aged 9, 13, and 17 have taken NAEP tests in both mathematics and science. These tests, constructed on a consistent basis to provide comparisons over time, provide an absolute benchmark of performance to which the U.S. scores on international tests can be keyed. Thus, the mean for each international test series is allowed to drift in accordance to U.S. NAEP score drift and the mean U.S. performance on each international comparison.[6] The constructed measures of schooling quality for each country are the weighted average over all available transformed test scores where weights are the (normalized) inverse of the country-specific standard error (σ).

Figure 1 gives a graphical depiction of the available test information over time. In this, scores for each age-group and subtest are combined into a country score on each of the six underlying assessments (with the world mean for each assessment set to 50). The quality measures subsequently used for each country combine these scores across years to obtain a single measure.

[6] Since the United States participated in all of the international comparison studies, the performance of U.S. students participating international tests can be transformed to mimic the performance of U.S. students in NAEP in the closest time and age-group; see U.S. Department of Education (1994). Scores of all other countries are adjusted proportionately on the basis of matched U.S. score adjustments. The first IEA mathematics test in 1963 and 1964 does not have an NAEP comparison, because it pre-dated the NAEP. While we experimented with discarding these scores in the calculation of labor-force quality, subsequent empirical results were unaffected.

The overall NAEP scores for U.S. students fell during the 1970's and rose in the 1980's, just the pattern that is seen in Figure 1 for U.S. performance on the separate international tests. Thus, our two alternative measures are highly correlated ($r = 0.92$). Table 1 shows summary statistics for the two constructed quality measures for the 31 countries that also have complete economic data for the subsequent growth analysis. The first series, *QL*1, is based on world average scores of 50 for all six tests; the second, *QL*2, is benchmarked to the U.S. performance on the NAEP.

II. Quality Effects on Growth

The underlying formulation guiding this empirical work follows from endogenous growth models where a country's growth rate is directly related to the stock of human capital. A variety of underlying formulations point to this empirical specification. In Romer (1990a), human capital influences the supply of ideas and new technologies. In the AK models of Rebelo (1991), growth is directly related to the aggregate capital stock that includes human capital with the key element being the lack of diminishing returns to human capital. Such models also have earlier antecedents in the models of technology adoption of Nelson and Phelps (1966) and Finis Welch (1970), where a country's stock of human capital influences the rate at which it puts in place new technologies and thus its growth rate. While having different policy implications, each can motivate the basic formulation employed here.

The endogenous growth formulation is, of course, not the only model of growth differences, and considerable controversy exists about the best formulation. (For a discussion of alternative formulations, see Barro and Xavier Sala-i-Martin, 1995.) Part of the debate is theoretically based, largely related to the long-run properties of the alternative models. Part is empirically based. The fundamental issue for empirical specifications is whether stocks of human capital or changes in the human-capital stock should enter into the determination of growth rates. If education is viewed as a direct input into production, then growth rates would be related to growth in the different inputs, and changes in the human-capital stock would be

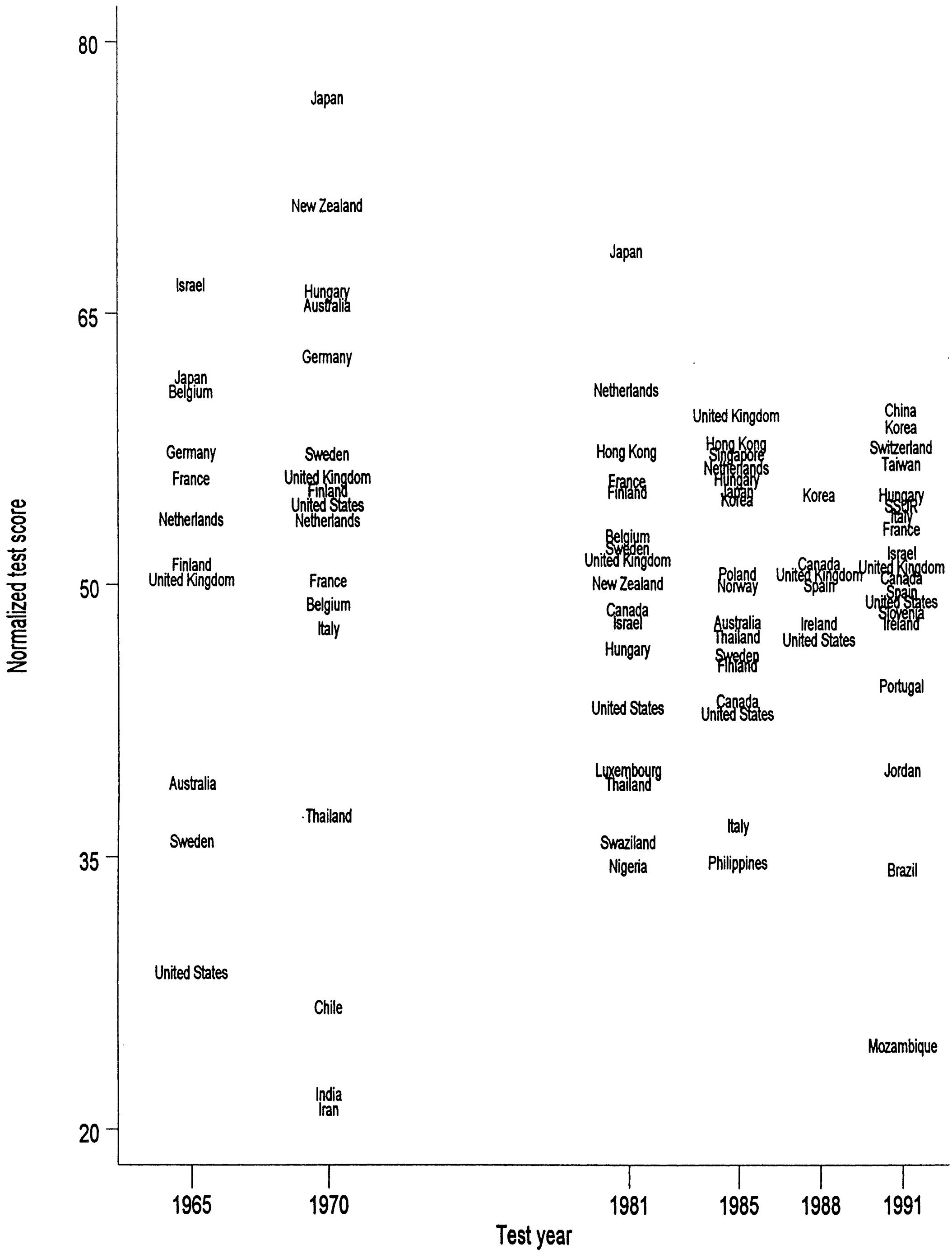

FIGURE 1. SCORES ON INTERNATIONAL MATH AND SCIENCE TESTS BY YEAR

TABLE 1—SUMMARY STATISTICS FOR CONSTRUCTED LABOR-FORCE QUALITY MEASURES (31 COUNTRIES)

Quality measure	Median	Mean	Standard deviation	Minimum	Maximum
*QL*1	48.76	46.61	10.86	20.79	60.65
*QL*2	54.52	51.28	13.48	18.26	72.13

Notes: See Hanushek and Kim (1995) for a detailed discussion of the included tests along with the construction of labor-force quality measures. *QL*1 sets the world mean on each of the six underlying tests equal to 50. *QL*2 adjusts all scores based on the U.S. international performance modified for the national time pattern of scores on the National Assessment of Educational Progress. While 39 countries participated in one or more test, these data are restricted to countries that also have the economic data required for the subsequent growth analysis.

the relevant explanatory factor in growth (e.g., Mankiw et al., 1992). Various aspects of these alternative models have been tested, but the results, which depend on specific formulations and implications, have not been conclusive.[7] Much of the past discussion has focused on the specification and interpretation of quantity measures of human capital (cf., Mark Bils and Peter J. Klenow, 2000). A central argument here is that the existing testing is likely to be misleading if quality concerns such as those raised in this paper are key to growth differences.

The emphasis in this paper is the comparison of alternative empirical versions of endogenous growth models with special attention to how the explanation of cross-country growth is affected by inclusion of quality measures. This analysis cannot provide tests of alternative growth formulations, because just cross-sectional variations in labor-force quality are observed.[8]

Table 2 reports the results of our baseline cross-country regressions that describe growth in average real per capita GDP between 1960 and 1990 (Robert Summers and Alan Heston, 1991). This begins with a small set of determinants of growth rates and investigates the magnitude and stability of the influence of labor-force quality. Subsequent estimation considers expansion to a broader set of countries and, in the tradition of Levine and Renelt (1992), the potential confounding effects of other frequently assessed factors.

The simplest models for the sample of 31 countries with complete data relate growth to initial income (*Y*60) and the Barro-Lee measure of school quantity (*S*) [column (1)]; an alternative baseline employs these plus the annual growth rate in population times 100 (*GPOP*) [column (4)]. (Variable definitions and sources of data are found in Appendix A.) These baseline estimates give rather common results and explain between 33–41 percent of the variation in economic performance for the restricted set of countries considered. The basic results are consistent with past estimation. Initial income negatively affects growth, supporting notions of conditional convergence in growth rates.[9]

[7] For example, Jess Benhabib and Mark M. Spiegel (1994) develop alternative estimation approaches within an international growth accounting framework in order to compare the factor accumulation view with various endogenous growth versions. Their empirical application of endogenous growth models involves adding Barro and Lee schooling levels (simple endogenous growth) and schooling levels interacted with the countries productivity gap (Nelson and Phelps growth). They argue from their empirical analysis that there is weak support for viewing human capital as a simple input into an aggregate production function. Simple endogenous growth models, defined in terms of just level of schooling attained, also receive limited support, but there is stronger support for a Nelson and Phelps version with evolving adoption of best technologies.

[8] The international math and science tests are given at different points in time, suggesting that it could be possible to look at changes in quality over the 30-year period considered here. This approach is, however, impractical because the emphasis here is on the quality of the labor force and not the quality of students. A change in observed test performance of current students might indicate future growth effects but not contemporaneous effects, because the achievement change would not have propagated through the labor force.

[9] William Easterly and Sergio Rebelo (1993) point out that using World Bank data reduces the possibility of the

TABLE 2—BASELINE ESTIMATES OF 1960–1990 CROSS-COUNTRY GROWTH MODELS WITH LABOR-FORCE QUALITY (Dependent variable: Average annual growth rate in real per capita GDP (×100) [31 countries])

	(1)	(2)	(3)	(4)	(5)	(6)
Initial per capita income (*Y*60)	−0.609	−0.472	−0.460	−0.745	−0.481	−0.517
[$1,000]	(0.186)	(0.096)	(0.103)	(0.181)	(0.093)	(0.112)
Quantity of schooling (*S*)	0.548	0.103	0.100	0.519	0.106	0.116
	(0.209)	(0.126)	(0.146)	(0.195)	(0.119)	(0.139)
Annual population growth				−0.713	−0.038	−0.250
(*GPOP*)				(0.224)	(0.215)	(0.211)
Labor-force quality (*QL*1)		0.134			0.133	
		(0.023)			(0.024)	
Labor-force quality (*QL*2)			0.104			0.098
			(0.015)			(0.015)
Constant	2.265	−1.900	−0.989	4.092	−1.756	−0.151
	(0.863)	(1.004)	(0.910)	(0.974)	(1.346)	(1.142)
R^2	0.33	0.73	0.68	0.41	0.73	0.69

Note: Huber-White estimated standard errors are in parentheses below coefficients.

Quantity of schooling (S) has a strong, positive impact on growth.

The corresponding estimates with the addition of our alternative measures of labor-force quality, found in the remaining columns, indicate a very strong relationship between quality and per capita growth rates. In the simplest form, adding either quality measure ($QL1$ or $QL2$) boosts the adjusted R^2 to about 0.7, a substantial increase from the simpler models. An increase of one standard deviation of labor-force quality, either measured by $QL1$ or $QL2$, enhances the real per capita growth rate by over 1.4 percentage points per year (e.g., $0.134 \cdot 10.86 = 1.46$). In contrast, a one-standard-deviation increase in the quantity of schooling is associated with only a quarter-percentage-point increase in growth ($0.10 \cdot 2.63 = 0.26$). (Note that the school quantity effect falls dramatically with inclusion of direct performance measures.)[10] The magnitude of the estimated quality effect is huge, given that the unconditional standard deviation of growth rates for these countries is 1.75 percent. (We return to issues of the magnitude of effects below.)

Replication of the basic models with the addition of population growth [columns (4)–(6)] leaves the magnitude and significance of the quality of schooling little changed. The estimates of the impact of population growth on the rate of economic growth, while always negative, are quite sensitive to specification and are not significantly different from zero when quality is considered. Also, as with other influences that have been estimated in the past, causality is a concern, since higher income may lead a country to reduce its birthrate.

The basic model concentrates just on additive effects of school quantity and quality, but their effects may be complementary. Some experimentation with a linear interaction term produced implausible results for the sample extremes for both quantity of schooling and measured quality. The substitution of logarithmic models on the other hand yielded virtually identical qualitative results as the basic additive

negative coefficient on initial income arising from potential measurement error in initial income. Nevertheless, for consistency with other studies, we rely on Summers and Heston (1991) data for initial income and growth rate calculation.

[10] This fall in returns to quantity with inclusion of quality is much larger than those typically found in micro estimates of wage determination models and those reported below. Bils and Klenow (2000) point out that the simple estimates of the effects of quantity of schooling are unbelievably large, a situation they attribute to reverse causality. The significantly smaller estimates here when quality effects are considered make the quantity effects more plausible, although, as we discuss below, there is uncertainty about the best way to compare micro productivity estimates with the macro growth equations.

models without interactions. Given the paucity of data, it remains impossible to investigate satisfactorily alternative functional forms.

The basic results provide strong support for the importance of labor-force quality difference as measured by cognitive achievement in mathematics and science. They also comprise the baseline for subsequent analyses both of the avenue of effects and of the generalizability to other countries.

III. Causality, Part A: The Determinants of Schooling Quality

Growth provides increased resources to a nation, and a portion of these resources may be ploughed back into human-capital investments, thus suggesting that the relationships previously estimated could overstate the causal impact of higher labor-force quality. Consider the following such description:

$$g_i = \mathbf{X}_i\boldsymbol{\beta} + \gamma QL_i + \varepsilon_i \tag{1}$$

$$R_i = \mathbf{W}_i\boldsymbol{\delta} + \eta g_i + \nu_i \tag{2}$$

$$QL_i = \mathbf{Z}_i\boldsymbol{\alpha} + \pi R_i + v_i. \tag{3}$$

Growth (g_i) of nation i is determined by labor-force quality (QL) plus a vector of other factors ($\mathbf{X}$) [equation (1)], while growth also contributes along with other factors $\mathbf{W}$ to determining the amount of resources devoted to schools and human-capital production (R_i) [equation (2)]. This formulation highlights the fact that governments cannot directly affect outcomes but instead must pursue various indirect policies that depend on the organization of schools and the underlying production function. If, however, resources in combination with other inputs ($\mathbf{Z}$) determine labor-force quality [equation (3)], simple estimation of equation (1) will not provide estimates of the causal effect of quality on growth (γ). Instead the estimation will also reflect the impact of growth on quality, embodied in the structural parameters η and π.

It is infeasible to estimate the complete system of equations with available data. Instead, the approach here is direct estimation of equation (3), the human-capital production function that relates measured labor-force quality to aggregate resources for schools and characteristics populations across countries ($\mathbf{Z}$). As long as ν and v are uncorrelated, estimation of equation (3) provides consistent estimates of the production parameters. The estimation follows the micro level studies of school performance (Hanushek, 1979, 1986). While past estimation of such models, both in the United States and in developing countries, has failed to find a consistent relationship between school resources and student performance (Hanushek, 1995, 1996), these results do not necessarily carry over to international comparisons. Within-country variations of resources are dwarfed by between-country variation, implying that, even if small marginal differences in resources have little effect, the large order-of-magnitude differences found across countries could have important outcome effects.

We assume that the international level of average ability of students does not vary across countries (or at least is exogenous to the other determinants considered here), leading us to concentrate on standard, readily available resource measures and aggregate population characteristics.[11] For the previous growth models, we were interested in measuring the quality of the entire labor force, leading to aggregation of different test scores into a single measure for each country. Here, however, our interest is the underlying production relationships, and we can exploit the disaggregated test data to explore

[11] The sources of variables, as explained in Appendix A, are mainly from Marlaine E. Lockheed and Adrian Verspoor (1991) and Barro and Lee's (1993) data set. For characteristics of education system, we consider: pupil/teacher ratio in primary schools (*PT-pri*) and in secondary schools (*PT-sec*), pupils/school ratios, teaching materials and teacher salary in primary schools, repeaters in primary schools, percentage of primary-school cohort reaching the last grade, public recurrent expenditure in primary school relative to GNP, ratio of recurring nominal government expenditure on education to nominal GDP (*RECUR*), current expenditure per pupil (*PPE*), and ratio of total nominal government expenditure on education to nominal GDP (*EXPEND*). The school spending variables are converted into spending per pupil by using purchasing-power parity estimates from Summers and Heston (1991) along with relevant school-enrollment data. For socioeconomic variables, we consider: general education level (schooling quantity—*S*), per capita income at 1960 (*Y60*) and average income, and demographic variables (fertility rate, growth rate of population (*GPOP*), infant mortality rate, life expectancy at birth).

TABLE 3—PRODUCTION OF MATHEMATICS AND SCIENCE ACHIEVEMENT: COUNTRY- AND COHORT-SPECIFIC SCORES
(Dependent variable: Normalized test performance in test year t)

	(1)	(2)	(3)	(4)	(5)	(6)
IEA math1				44.09	50.15	40.78
				(5.16)	(5.61)	(7.03)
IEA science1				46.44	54.30	41.98
				(5.87)	(5.48)	(7.01)
IEA math2				48.49	55.49	41.30
				(6.04)	(5.65)	(7.11)
IEA science2				45.94	51.66	37.97
				(6.35)	(6.16)	(7.62)
IAEP math and science				47.14	52.45	38.24
				(5.47)	(4.91)	(6.57)
Adult schooling (S_{t-1})	2.04	1.62	1.54	2.70	1.75	1.59
	(0.82)	(0.76)	(0.64)	(0.70)	(0.73)	(0.64)
Pupil-teacher ratio in primary schools ($PT\text{-}pri_{t-1}$)			0.066			0.09
			(0.16)			(0.15)
Current public expenditure per student (PPE_{t-1})	−0.69			−0.766		
	(0.19)			(0.21)		
Total expenditure on education/GDP ($EXPEND_{t-1}$)		−165.90			−189.78	
		(90.66)			(88.69)	
Annual population growth ($GPOP_{t-1}$)	−4.65	−4.60	−2.64	−4.86	−4.98	−2.81
	(1.68)	(1.36)	(1.96)	(1.94)	(1.42)	(1.91)
Constant	46.46	52.27	40.80			
	(5.17)	(4.94)	(6.55)			
Number of countries	69	67	70	69	67	70
R^2 (adjusted)	0.25	0.19	0.25	0.22	0.26	0.25

Notes: The sample consists of one observation per country for each of the six test years that each country participated in the international testing. Scores are normalized to 50 in each test year. Huber-White estimated standard errors are in parentheses below coefficients.

how school and family resources relate to the performance of specific cohorts. To estimate the school production relationships, we consider normalized individual country test scores on the six separate tests (i.e., the individual country data points plotted in Figure 1) and relate them to the relevant cohort-specific family characteristics and school resources (see Appendix B). This approach provides up to 70-country-test-cohort observations with necessary school input data, while ensuring that the inputs are exogenous to the achievement outcomes.[12]

Table 3 displays several variants of the production models. The overall story is that variations in school resources do not have strong effects on test performance. The estimated effects of various measures of resources are either statistically insignificant or, more frequently, statistically significant but with an unexpected sign. This finding holds regardless of the specific measure of school resources—whether pupil-teacher ratios, recurrent expenditure per student, total expenditure per student, or a variety of other measures. The education of parents, proxied by the quantity of schooling of the adult population, tends to be positive and significant at conventional levels. Also, countries with higher population growth rates tend to have lower achievement, consistent with standard arguments about the trade-off between quantity and quality of children and the impact of larger family sizes (Gary S. Becker and H. Gregg Lewis, 1973; Robert J. Willis, 1973; Hanushek, 1992). Alternative models (not shown) included region-of-the-world dummy variables, but these did not change any of the

[12] A total of 87 country-test observations are available, but limits on school input data reduce the usable sample.

estimated school resource effects. Most significantly, the perverse effects of pupil-teacher ratios are not a simple artifact of the large class sizes found in many East Asian countries.

Finally, the structure of the testing itself might influence the pattern of observed performance. Variations in school attendance could lead to variations in scores due to pure selection effects, because countries with lower school completion rates might typically test a more selective portion of the total age cohort. The models in Table 3 were reestimated adding the secondary-school enrollment rate for the five-year period covering the specific testing program (not shown). The secondary-school enrollment rate was uniformly positive and statistically insignificant (the opposite of what might be expected from important selection effects), while the pattern of the school resource effects was unaffected.

A significant concern with previous analyses of cross-country growth relationships has been the likelihood of simultaneity—fast-growing countries are likely to invest in more schooling, more plant and equipment, and the like. The effect of growth on human-capital development, as opposed to the reverse, has specifically been emphasized, for example, by Jacob Mincer (1996) and Bils and Klenow (2000). The lack of systematic income and expenditure effects on labor-force quality strengthens the causal interpretation of labor-force quality in our growth models, because $\pi = 0$ eliminates the feedback loop in equations (1)–(3).

IV. An Expanded Sample of Countries

To explore growth differences among countries further, we significantly expand the group of countries analyzed by projecting labor-force quality based on observed characteristics. The estimation strategy concentrates on equations (1) and (3) above. We observe g, $\mathbf{X}$, QL, and $\mathbf{Z}$ for a set of n_1 countries, while we just observe g, $\mathbf{X}$, and $\mathbf{Z}$ for another set to countries $n_1 + 1, \ldots, n_2$. We begin by obtaining consistent estimates of $\boldsymbol{\alpha}$, the effect of country factors on measured quality, from estimation of (3) for the first set of n_1 observations with complete data. We then use this estimate ($\hat{\boldsymbol{\alpha}}$) to estimate jointly equation (1) for the first n_1 countries and

$$(4) \quad g_i = \mathbf{X}_i\boldsymbol{\beta} + \gamma\mathbf{Z}_i\hat{\boldsymbol{\alpha}} + [\gamma v_i + \varepsilon_i + \gamma\mathbf{Z}_i(\boldsymbol{\alpha} - \hat{\boldsymbol{\alpha}})]$$

for the remaining $n_2 - n_1$ countries, where the bracketed term indicates the composite error term in equation (4).

When ε and v are uncorrelated for the countries with complete data, estimation of this system will, under quite general conditions, yield consistent estimates of γ.[13] Clearly, however, the full error term in (4) will have a larger variance than that in (1), so correction for heteroskedasticity is important both to obtain efficient parameter estimates and to obtain consistent estimates of the standard errors. If the only source of heteroskedasticity arises from estimation of QL for the expanded set of countries in equation (3), a simple two-stage estimator that employed different variance estimates for the first n_1 observations and for the next $n_1 + 1, \ldots, n_2$ observations would be appropriate. Because other sources of heteroskedasticity are likely, however, we employ more general robust estimation techniques.[14]

To project labor-force quality in equation (3), we return to the simple cross-sectional sample of aggregate country values for labor-force quality, $QL1$ and $QL2$, and relate them to average population and schooling characteristics for the entire 1960–1990 period. Our primary purpose is developing projection equations, and we employ expanded models that go beyond the previous production function estimates.

As shown in Table 4, three basic factors are

[13] The prior production function estimates plus the further analysis of U.S. earnings of immigrants (in the text) provide some evidence about the appropriateness of the error assumption. With a consistent estimate of $\boldsymbol{\alpha}$, the last term in equation (4) vanishes in the limit, yielding consistent estimates of $\boldsymbol{\beta}$ and γ. Based on the prior estimation, we ignore possible feedback in equation (2) and πR_i in equation (3).

[14] Subsequent empirical analysis indicates that this simple weighting and the use of Huber-White corrected standard errors yield very similar results (Halbert White, 1980). Note that this estimation problem is not the same as a standard generated regressor case, given the partial observability of QL. Ignoring observed values of QL would effectively increase the error variance in (1) and thus reduce the efficiency of the estimator. This approach is also not the same as errors-in-variables, because it rests on the consistency of the underlying estimator in equations (3) and (4).

TABLE 4—PREDICTION MODELS FOR LABOR-FORCE QUALITY

	Dependent variable = *QL*1			Dependent variable = *QL*2		
	(1)	(2)	(3)	(4)	(5)	(6)
Primary-school enrollment rate (*ENROLL-pri*)	81.50	85.78	86.23	62.36	73.28	75.32
	(16.54)	(23.32)	(24.81)	(24.44)	(29.54)	(31.30)
Quantity of schooling (*S*)	0.714	0.038	0.320	1.651	0.97	1.181
	(0.53)	(0.65)	(0.64)	(0.77)	(1.18)	(1.14)
Pupil-teacher ratio in primary schools (*PT-pri*)	0.142			−0.043		
	(0.06)			(0.13)		
Recurring expenditure on education/GDP (*RECUR*)			55.49			133.23
			(119.3)			(176.1)
Total expenditure on education/GDP (*EXPEND*)		121.07			170.37	
		(109.9)			(169.1)	
Annual population growth (*GPOP*)	−3.11	−3.08	−2.80	−4.105	−4.17	−3.85
	(1.76)	(1.98)	(2.06)	(2.36)	(2.65)	(2.62)
Asian (=1)	5.123	7.52	6.33	12.25	13.77	12.76
	(3.56)	(2.76)	(3.07)	(6.54)	(4.76)	(5.02)
Latin American (=1)	−4.004	−3.87	−4.31	−1.24	0.203	0.432
	(2.31)	(2.93)	(3.39)	(3.689)	(4.07)	(4.53)
African (=1)	3.170	2.94	3.30	11.97	8.71	9.16
	(2.38)	(2.32)	(2.68)	(5.73)	(3.44)	(3.72)
Constant	−35.59	−37.19	−36.19	−13.10	−28.40	−29.27
	(17.20)	(22.41)	(24.55)	(25.76)	(27.65)	(30.17)
Number of countries	31	30	30	31	30	30
R^2	0.73	0.72	0.71	0.68	0.66	0.65

Note: Huber-White estimated standard errors are in parentheses below coefficients.

generally important in determining variations in labor-force quality. Primary-school enrollment rates strongly influence performance, probably through indicating the overall importance that each country puts on education. Higher population growth rates are associated with lower labor-force quality. Finally, distinct regional differences influence performance with Asian countries doing best (given the other characteristics).[15] The average quantity of schooling is generally positively related to performance, although it is typically insignificant at conventional levels. School resources again are not strongly related to quality. The incorrect positive sign for pupil-teacher ratio appears whether or not there is a dummy variable for the Asian region, a region with traditionally high pupil-teacher ratios and high student performance. The expenditure measures, while positive, are statistically insignificant.

We utilize the estimates in the second and fifth columns to construct an expanded set of labor-force quality measures, *QL*1* and *QL*2*, which combine observed quality with predicted quality for all countries without observed test data but with data on the right-hand-side measures. A significant concern, however, is that relatively few developing countries ever participated in the testing, implying limited observations at the low-achieving end of the distribution. Because of the thinness of observations at the lower end (see Figure 1), our main growth estimates eliminate all countries with predicted achievement below 20, the lower bound on actual observations. The subsequent

[15] The limited number of countries makes the regional differences suspect. While there are six Asian countries, there are only two African (Mozambique and Swaziland) and one Latin American (Brazil) country. This implies that the estimates for Latin American countries are calculated relative to the Brazilian mean instead of the world mean. Similarly, African countries are calculated relative to the mean for Mozambique and Swaziland.

TABLE 5—1960–1990 CROSS-COUNTRY GROWTH MODELS WITH LABOR-FORCE QUALITY FOR AUGMENTED QUALITY SAMPLE
[Dependent variable: Average annual growth rate in real per capita GDP (×100)]

	(1)	(2)	(3)	(4)	(5)	(6)
Initial per capita income (*Y*60)	−0.382	−0.390	−0.453	−0.370	−0.384	−0.442
[$1,000]	(0.081)	(0.079)	(0.078)	(0.084)	(0.082)	(0.081)
Quantity of schooling (*S*)	0.127	0.117	0.112	0.120	0.103	0.112
	(0.089)	(0.093)	(0.093)	(0.096)	(0.100)	(0.104)
Annual population growth		−0.097			−0.161	
(*GPOP*)		(0.212)			(0.209)	
Labor-force quality (*QL*1*)	0.108	0.104	0.076			
	(0.021)	(0.023)	(0.027)			
Labor-force quality (*QL*2*)				0.094	0.090	0.072
				(0.016)	(0.016)	(0.021)
Assessment available (*TEST*)			−1.392			−0.628
			(1.455)			(1.436)
Observed labor-force quality			0.054			
(*TEST QL*1*)			(0.032)			
Observed labor-force quality						0.034
(*TEST QL*2*)						(0.028)
Constant	−1.606	−1.184	−0.475	−1.483	−0.869	−0.657
	(0.749)	(1.241)	(1.069)	(0.584)	(0.984)	(0.798)
Number of countries	78	78	78	80	80	80
R^2	0.41	0.42	0.49	0.41	0.41	0.48

Note: Huber-White estimated standard errors are in parentheses below coefficients.

growth estimation, however, is not very sensitive to this sampling criteria.[16] Also, there ultimately proves to be little difference between the two measures of labor-force quality, which have a simple correlation of 0.95 in the augmented country sample.

As an independent check on the projections, the estimates of labor-force quality can be related to the eighth-grade country scores from 1995 on the Third International Math and Science Study testing, or TIMSS (U.S. Department of Education, 1996a, b). The correlation of the combined TIMSS score with *QL*1* and *QL*2* is 0.67 and 0.66, respectively.[17] If South Africa, which is a full standard deviation below the score of the next lowest of the 32 matched countries, is excluded, the correlations rise to 0.76 and 0.78. Importantly, there is no significant difference in these relationship for the eight newly tested, as opposed to the 24 previously tested, countries in the TIMSS sample.[18] The high correlations also reinforce the previously assumed relative stability of schooling systems.

The growth estimates in Table 5, which replicate the basic models in Table 2, are very consistent with the baseline estimates. Considering *QL*1*, the first column indicates significant conditional convergence with higher initial income translating into lower growth rates. While quantity of schooling is positive, it is statistically insignificant, and its estimated

[16] The cutoff, which lies more than two and one-half standard deviations below the country means, eliminates nine countries which meet overall data availability criteria. The sampling criterion has slightly different sample implications for the two separate quality measures with 77 common observations. To test the sensitivity of the results to truncation, reestimation of the growth models using all countries with predicted values of *QL*1* and *QL*2* greater than zero yielded estimates that are highly statistically significant but slightly lower in magnitudes from those estimated on the more restricted samples.

[17] It was possible to construct eighth-grade scores for 32 countries that enter into our analysis, including eight countries not previously directly observed (U.S. Department of Education, 1996a, b). The relationship of TIMSS scores and growth is explored in Paul T. Decker and Larry M. Radbill (1999).

[18] If *QL*1* and *QL*2* are regressed on the TIMSS aggregate scores, a dummy variable for direct observations of prior tests or allowance for a different slope on the TIMSS variable if labor-force quality was observed yields insignificant differences.

effect on growth is close to what was observed previously in Table 2. Most important, labor-force quality—measured by $QL1^*$—is statistically significant and very similar in magnitude to the estimates for the restricted set of countries in the baseline. A one-standard-deviation change in quality translates into a slightly greater than one-percentage-point difference in annual real growth rates. The effect of quality improvements also appears much more important than increases in quantity: a one-standard-deviation change in school quantity translates into 0.32 percentage points in average growth (and even this might be overstated by the arguments of Bils and Klenow, 2000).

The remainder of the table demonstrates stability of the estimated effects of quality. In column (2), the addition of population growth rates has no impact on the estimated quality effect. Column (3) provides information on the projection of labor-force quality in growth by separating countries with direct observations from those with projections. The point estimates on marginal test score effects ($TEST \cdot QL1^*$) indicate somewhat stronger test influences in countries where there are direct observations, although the coefficient is not significant at the 5-percent level.[19] As found before, the results from the alternative quality measure ($QL2^*$) are virtually the same.

Following Levine and Renelt (1992), a variety of other common measures of economies were added (not shown), but the importance of quality differences remains both in terms of consistent strength and statistical significance.[20] Barro (1991) and others have emphasized a variety of political factors which might influence growth. When we introduce measures of political assassinations or of revolutions and coups into our models, they are uniformly insignificant, and they do not lessen the importance of variations in labor-force quality.[21] Importantly, discussions below about the estimated magnitude of quality effects raise the possibility that quality partially proxies for some set of omitted variables. This sensitivity analysis, however, rules out many commonly hypothesized candidates.

The estimates using this augmented sample confirm the appropriateness of projection to the expanded set of countries. This expanded sample provides increased precision in testing further hypotheses. It also points to the broader generalizations across world economies that are possible.

V. Causality, Part B—The Quality-Productivity Relationship[22]

Confusion about causality can also arise from a simple omitted variable problem. To the extent that other aspects of a country influence both scores and the success of the economy, the measures of labor-force quality may simply be proxying for the true influences. For example, if growth were related to more open labor markets, nations with more open labor markets may also have better allocations of workers to teaching and thus better student performance. Or, investment in health of individuals may lead both to higher productivity and growth and to better school performance. These and other examples could produce the kind of quality-growth relationships estimated here, even when cognitive skills of workers per se are not driving growth. For example, consider a situation where a common factor, θ, affects both growth of the economy and efficiency of the schools, but where measured quality, QL, does not enter into growth:

$$g_i = f(\theta_i, \mathbf{X}_i\boldsymbol{\beta}) + \varepsilon_i \tag{5}$$

$$QL_i = g(\theta_i, \mathbf{Z}_i\boldsymbol{\alpha}) + v_i. \tag{6}$$

[19] The F-test on the addition of both the intercept and slope dummies does reject the hypothesis that both parameters are jointly zero ($F[2, 72] = 5.59$), but slope equality seems most important.

[20] The additional factors included: the ratio of real government consumption expenditure net of spending on defense and education to real GDP (*GCN*), the ratio of private investment to GDP (*PRIINV*), and ratio of total trade to GDP (*TRD*). While school quantity (S) was insignificant, the consistent result was that labor-force quality remained statistically significant, although at times slightly smaller in magnitude than in Table 5. Alternative functional forms, including log specifications and interactions between quality and the other variables, likewise did not change the overall results.

[21] We average assassinations per million population and number of revolutions and coups per year across the observed periods. While such measures would imprecisely show the impact of political upheaval on short-run growth, they should still capture any overall effects.

[22] We thank a prior referee for suggesting this analysis.

This is a simplified structural form of the kind of omitted variables problem that has been hypothesized to enter in one form or another into estimation and interpretation of many previous growth models.[23] Without measures of θ, it is extraordinarily difficult to deal with such situations, because it generally requires finding instrumental variables which are correlated with true student and worker quality but which are uncorrelated with growth.

In our analysis, we employ a different strategy. We concentrate on immigrants working in the United States and relate variations in their earnings to our measures of labor-force quality. If our measure of labor-force quality does not indicate differences in productivity but is merely a proxy for other differences in each country's economy, we should see no relationship with immigrant earnings differences within the U.S. economy. Moreover, by distinguishing where each immigrant's education was received—in the home country, in the United States, or in a combination of the two—it is possible to relate the quality measures more closely to schools than to other characteristics of the immigrants, be it cultural, family behavior, or whatever.

We constructed a sample of all foreign-born male workers, aged 25–60, with at least $1,000 in earnings in 1989 from the 1990 Census of Population 5-percent public use micro data files (PUMS). The PUMS data set provides information about 1989 labor earnings, schooling attained, age, country of origin, and age of entry into the United States. Two overlapping samples were constructed: (1) those who were born in a country for which direct quality measures were available [37 countries]; and (2) those in the augmented sample that adds countries for which quality measures are estimated [87 countries].[24] The OLS estimated earnings equations begin with the standard Mincer form (Mincer, 1974) where log of annual earnings is regressed on years of school, potential experience (age-schooling-6) and its square and then add our measure of labor-force quality ($QL2^*$) based on country of origin.

The simple Mincer models in Table 6 for male immigrants (the first column for each of the two samples) provide rather standard results: earnings gradients of 9–10 percent per year of schooling and declining returns to experience over a career. The interesting portion is the estimated effects of quality. In the sample with observed tests, one point of test score approximately translates into an average of 0.19-percent higher individual earnings. The following three columns then disaggregate the sample by where the immigrant obtained schooling (based on age of entry into the United States). Those entirely educated in the home country average 0.21-percent higher earnings for each point of test score, while those educated partially or totally in the United States show a statistically insignificant relationship between home-country labor-force quality ($QL2^*$) and individual earnings. With the augmented sample of countries on the right-hand side of the table, the pattern of results is the same, and estimated quality effects on individual earnings are even stronger.

If the growth effects of labor-force quality previously estimated just reflected some omitted factor related to the home country and the efficiency of its markets, we would not expect to see productivity effects for immigrants in the U.S. economy. Further, if the growth effects simply reflected cultural or family factors and not school achievement and skill effects, we would expect the influence on productivity in the U.S. economy to hold regardless of where the schooling took place. We interpret the results in Table 6 as providing consistent evidence that our measures of labor-force quality are related to individual productivity and signal a causal influence in the growth relationship.

Nevertheless, while the qualitative evidence supports a causal relationship between country-specific quality differences and individual productivity, the question remains whether the magnitude of the estimated effects is sufficient to explain the strong relationship with growth rates. The prior investigation of possible simultaneous provision

[23] The same set of problems arise without θ entering both structural equations if θ is correlated with the error in one of the equations and enters directly into the other.

[24] Countries are included if there are at least 50 immigrants in the United States meeting the work and earnings restrictions. Note that a larger set of countries are used in both samples here, because information on the own-country growth and schooling levels is not required.

TABLE 6—EFFECTS OF SCHOOL QUALITY ON U.S. EARNINGS OF IMMIGRANTS, 1989[a]

	Only immigrants from countries with observed quality measures (n = 37)					Immigrants from countries with observed or predicted quality measures (n = 87)				
			Schooling location					Schooling location		
	All		Home only	United States and home	United States only	All		Home only	United States and home	United States only
Years of school	0.089 (0.002)	0.090 (0.002)	0.080 (0.002)	0.103 (0.004)	0.132 (0.005)	0.103 (0.001)	0.099 (0.001)	0.089 (0.001)	0.117 (0.003)	0.130 (0.004)
Potential experience	0.061 (0.002)	0.061 (0.002)	0.069 (0.003)	0.089 (0.004)	0.067 (0.008)	0.058 (0.002)	0.058 (0.002)	0.067 (0.002)	0.085 (0.003)	0.066 (0.006)
Potential experience squared	−0.0009 (0.0000)	−0.0009 (0.0000)	−0.0010 (0.0001)	−0.0016 (0.0001)	−0.0011 (0.0002)	−0.0008 (0.0000)	−0.0008 (0.0000)	−0.0009 (0.0000)	−0.0015 (0.0001)	−0.0011 (0.0002)
*QL2**		0.0019 (0.0004)	0.0021 (0.0005)	−0.00018 (0.0006)	0.00004 (0.0016)		0.0060 (0.0003)	0.0064 (0.0004)	−0.0025 (0.0006)	0.0018 (0.0011)
Constant	8.154 (0.035)	8.046 (0.041)	7.977 (0.056)	7.881 (0.078)	7.597 (0.127)	7.883 (0.022)	7.678 (0.024)	7.589 (0.031)	7.498 (0.054)	7.542 (0.093)
Observations	20,644	20,644	12,955	4,956	3,412	3,9840	3,9840	2,6643	8,639	4,558
R^2	0.15	0.15	0.13	0.22	0.18	0.23	0.23	0.23	0.26	0.19

[a] Mincer earnings models are estimated with the dependent variable being log of labor earnings. Samples include immigrants with 1989 labor-market earnings who were born in a country for which measures of labor-force quality (*QL2**) are observed or, in the right panel, estimated.

of school resources with growth concluded that this path for growth effects was inconsistent with the data. But, this and the further investigation of variations related to East Asian countries (see the next section) are qualitative conclusions and say nothing about the estimated magnitude of quality-growth effects. The individual earnings estimation, which provides direct evidence about the magnitudes of productivity differences associated with varying school-induced skills, holds the possibility that something can be said about the magnitude of the growth effects.

Unfortunately, there is no simple way to translate the individual productivity effects, which relate to levels of income, into aggregate growth effects. Much depends upon the particular model of growth that is selected. For example, one approach would be to model achievement effects as affecting the aggregate steady-state level of output for an economy with growth coming through conditional convergence based on initial income levels.[25] In this, the relative magnitude of the quality coefficient and the coefficient on initial income provide an estimate of the impact of quality on steady-state levels of income—a parameter that would be comparable to the Mincer earnings parameter if growth comes directly from varying levels of human-capital inputs. The growth estimates, however, provide much larger estimates of quality effects than found in the earnings estimation for immigrants, suggesting that more than just direct productivity effects are entering into the growth relationship.

An alternative view from an endogenous growth model perspective would relate growth rates to stocks of human capital—here, either in terms of quantity or quality of human capital. In this, strong externalities or endogenous growth in terms of the aggregate stock of labor-force quality might be able to explain the apparent disparity noted in terms of individual productivity and growth effects of quality differences.[26] Two things are important in assessing this. Be-

[25] This approach is an extension of Barro and Sala-i-Martin (1995). We thank a referee for suggesting this way of estimating the growth effects of quality differences.

[26] As noted previously, the form could be from, for example, quality influencing the rate of technological change (Romer, 1990a) or an expanding steady-state level through closing gaps in innovation (Nelson and Phelps, 1966) or a variant of an AK model (Rebelo, 1991).

cause the quality coefficient is so large in the growth equations, both absolutely and relative to school quantity, the emphasis must be on the strength of externalities to quality that differ from those for quantity.[27] Second, the magnitudes of quality effects in an endogenous growth case still appear implausibly large. The quality estimates in the baseline and in the augmented growth models suggest that one standard deviation in test performance translates into more than 1-percent higher annual growth in GDP per capita. But, by the estimates of Klenow and Andrés Rodríguez-Clare (1997 p. 94), the average growth across a broad set of countries that is attributable to technological change is slightly over 1 percent per year—roughly the same as one standard deviation of test performance. Thus, the estimated quality effect in growth, even if coming from its aggregate influence on technological change, appears too large.

While other models allowing a crosswalk between quality effects on the level of earnings and on the rate of aggregate growth might reconcile the magnitudes,[28] we conclude that there is currently uncertainty about how much of the estimated growth effects comes from direct causal influences of measured labor-force quality. A portion could come from omitted variables that are strongly correlated with measured quality. The exact nature of these omitted variables is unclear, however, because the sensitivity analyses and other considerations of the causal structure rule out many of the most plausible factors. Thus, we conclude that the evidence here suggests that the quality measures do indicate productivity differences. It also confirms the role of schooling differences across countries, as opposed to cultural, racial, or parental factors, in affecting these quality differences. But, because the micro effects would appear to translate directly into smaller growth influences, questions about the complete causal structure remain.

VI. Causality, Part C—Sensitivity to East Asian Countries

The well-known growth experience of East Asian countries over our sample period raises concern that these results have been dominated by countries in that region. As seen in Figure 1, East Asian countries tend to score highly on the international tests, so it is possible that the test scores might be merely identifying these countries even when labor-force quality is not causally related to growth. This can be thought of as a special case of the omitted variables discussion of the previous section.

Table 7 provides evidence about the impact of these countries on the estimates of growth relationships. This table compares results for the entire sample with results that come from excluding different subsets of East Asian countries. Three subsets of East Asian countries are explored: Four Tigers (Hong Kong, Korea, Singapore, and Taiwan); High Performing (Four Tigers plus Japan); and Newly Industrialized (High Performing plus

[27] Considerable discussion has surrounded the magnitude of the effect of school quantity (cf., Bils and Klenow, 2000), so the relatively larger quality effects obtained here lead to similar questions, although the underlying structure is less clear.

[28] For example, this difference in estimated productivity effects between micro and macro estimates may reflect an errors-in-variables problem related to the measurement of the quantity of schooling (see Alan B. Krueger and Mikael Lindahl, 1999). Their analysis, however, concentrates more on the estimated effect of changes in human capital, whereas the averages employed here should lessen any errors-in-variables problems. Another possibility to reconcile the magnitude of the micro and macro estimates is that the productivity effects in the Mincer equation are poorly estimated. The Mincer estimation attributes overall country average quality to individuals, instead of including the appropriate individual measure. Moreover, the estimates are sensitive to sample. The quality estimates in Table 6 are three times as large in the augmented sample as in the smaller sample of test-taking nations. Nonetheless, while other labor-market factors influence immigrants' earnings picture and have some effect on the precise magnitude of the estimates, investigation of a variety of more complicated earnings specifications confirms the overall pattern, strength, and significance of the country-specific labor-force quality measures in the individual Mincer models for the United States. Hanushek and Jin-Yeong Kim (1999) investigate a variety of specifications of the individual earnings models, including ones with measures of immigrants' language ability, race, part-time work status, and selectivity of immigration. These alternative specifications affect the magnitude of coefficient estimates but not the overall strength and significance. Moreover, this latter perspective would not change the magnitude of the quality coefficient in the growth equation, which, independent of the issues about compatibility of the micro and macro estimates, raises concerns.

TABLE 7—THE IMPORTANCE OF EAST ASIAN COUNTRIES

	Countries with observed quality measures				Countries with observed or predicted quality measures			
	Full sample	Excluding Four Tigers[a]	Excluding High Performing[b]	Excluding Newly Industrialized[c]	Full sample	Excluding Four Tigers[a]	Excluding High Performing[b]	Excluding Newly Industrialized[c]
Initial per capita income	−0.472	−0.357	−0.328	−0.270	−0.382	−0.308	−0.291	−0.256
($Y60$)	(0.090)	(0.092)	(0.104)	(0.086)	(0.078)	(0.081)	(0.084)	(0.086)
Quantity of schooling (S)	0.103	0.117	0.106	0.085	0.127	0.148	0.140	0.143
	(0.118)	(0.116)	(0.118)	(0.113)	(0.086)	(0.087)	(0.089)	(0.091)
Labor-force quality ($QL1^*$)	0.134	0.101	0.095	0.091	0.108	0.079	0.075	0.069
	(0.020)	(0.022)	(0.024)	(0.023)	(0.020)	(0.020)	(0.020)	(0.020)
Constant	−1.901	−1.111	−0.929	−0.966	−1.606	−0.848	−0.709	−0.648
	(0.936)	(0.934)	(1.010)	(0.986)	(0.730)	(0.719)	(0.739)	(0.755)
Number of countries	31	27	26	25	78	74	73	70
R^2 (adjusted)	0.70	0.49	0.39	0.40	0.39	0.26	0.22	0.21

Note: Huber-White estimated standard errors are in parentheses below coefficients.
[a] Four Tigers: Hong Kong, Korea, Singapore, and Taiwan.
[b] High Performing: Hong Kong, Korea, Singapore, Taiwan plus Japan.
[c] Newly Industrialized: Hong Kong, Korea, Singapore, Taiwan, Japan plus Indonesia, Malaysia, and Thailand. (Indonesia and Malaysia did not participate in any of the international examinations).

Indonesia, Malaysia, and Thailand).[29] These models are estimated both for samples involving only directly observed test information and for the augmented samples (i.e., including countries with predicted tests). In all cases, negative effects of initial income level are significant, demonstrating conditional convergence, although these effects are stronger when the East Asian countries are included. The estimated effect of quantity of schooling rises some when East Asian countries are excluded from the larger sample and shows no consistent change when the sample is restricted just to countries with observed test scores—but it is never statistically significant.

For the sample with observed tests, the effect of labor-force quality drops in magnitude by one-quarter to one-third, depending on the precise sample, and the R^2 drops noticeably when any of the East Asian countries are deleted. These results are consistent with quality of human capital contributing to the growth of East Asian economies. Nonetheless, labor-force quality maintains a strong and significant effect on estimated growth in all subsets of countries. In the augmented samples, similar results hold.

[29] These subsets follow those of the World Bank (1993). High Performing would include China, but China is not included in the estimation because of missing data on initial income for the basic growth models.

In short, the importance of labor-force quality is not merely an artifact of the data driven by East Asian countries but holds in areas across the world. Outside of East Asia, a one-standard-deviation change in labor-force quality is still associated with a 0.7–1.0-percentage point higher growth rate (depending upon the specific sample employed in the estimation).

VII. The Importance of Quality Measurement

Table 8 provides direct comparisons of estimated models with quality measures versus no quality measures or schooling input measures. Without quality measures [columns (1–3)], school quantity is consistently significant. Further, from column (2), primary-school pupil-teacher ratio is significant at the 10-percent level, but the magnitude is small. Cutting average pupil-teacher ratios in half (to 18 from its current world mean of 36), an extraordinarily expensive policy, would increase growth rates by 0.7 percent per year. At the secondary level, pupil-teacher ratios have a perverse but statistically insignificant effect. Considering total expenditure per pupil provides little additional information about growth.

Direct measures of labor-force quality, however, have a very different impact [columns (4) and (5)]. The proportion of variance explained

TABLE 8—COMPARISON OF ALTERNATIVE MEASURES OF SCHOOL QUALITY AND GROWTH

	Baseline	School inputs		Quality measures		Combined input and quality	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Initial per capita income	−0.407	−0.455	−0.408	−0.382	−0.370	−0.393	−0.368
(*Y60*) [$1,000]	(0.139)	(0.114)	(0.137)	(0.081)	(0.084)	(0.095)	(0.095)
Quantity of schooling (*S*)	0.529	0.481	0.486	0.127	0.120	0.070	0.065
	(0.122)	(0.124)	(0.133)	(0.089)	(0.096)	(0.104)	(0.117)
Pupil-teacher ratio in primary		−0.040				0.001	0.006
schools (*PT-pri*)		(0.024)				(0.026)	(0.024)
Pupil-teacher ratio in		0.027				−0.038	−0.038
secondary schools (*PT-sec*)		(0.042)				(0.044)	(0.045)
Total expenditure on			14.42			7.388	3.968
education/GDP (*EXPEND*)			(14.73)			(16.060)	(15.100)
Labor-force quality (*QL1**)				0.108		0.112	
				(0.021)		(0.020)	
Labor-force quality (*QL2**)					0.094		0.100
					(0.016)		(0.015)
Constant	0.933	2.140	0.558	−1.606	−1.483	−1.113	−1.042
	(0.238)	(2.060)	(0.418)	(0.749)	(0.584)	(1.091)	(0.992)
Number of countries	100	96	96	78	80	76	78
R^2	0.23	0.26	0.22	0.41	0.41	0.42	0.42

Note: Huber-White estimated standard errors are in parentheses below coefficients.

doubles, reaching almost 40 percent. Moreover, quality measures replace quantity in demonstrating the importance of human capital for national growth.[30]

The final two columns of Table 8 incorporate both direct quality measures (QL) and the input measures of the second and third columns. None of the inputs has a significant influence on growth after conditioning on either of the labor-force quality measures. These final specifications simply reinforce the higher informational content of the test measures as opposed to the school inputs.

The importance of adequate measures of quality becomes clear in Table 9, which provides comparisons of the root mean-squared error (RMSE) in explaining growth rates across a consistent set of countries. The root mean-squared error in models with quality falls by 0.2 percentage points (in annual growth) when compared with the other models from Table 8. Moreover, the quality measures are especially powerful in explaining anomalies in growth rates. The final three columns in Table 8 divide the countries into slow growers (average annual growth less than 1 percent), fast growers (average annual growth greater than 3.5 percent), and the remainder. The reduction in root mean-squared error comes at the extremes of the growth distribution, and especially from better explaining rapid growth. The RMSE for fast growers falls from about 2.2 percent to 1.7 percent when $QL1^*$ is used to explain growth differences. While the improvement at the bottom of the distribution is less in absolute terms, it is still noticeable at about 0.2 percent. The results for $QL2^*$ are similar.

The previous analyses have concentrated on variations in growth rates that can be explained by initial conditions and by quantity and quality of human capital. At the same time, these highly simplified models leave out a wide range of characteristics of country economies. The distribution of observed growth rates reflects

[30] There is an artificially high correlation of the quantity and quality of schooling ($r = 0.71$ with $QL1^*$) that comes from the projection equation for predicting QL in countries that lack direct test information. The separate influences of quantity and quality, however, are identified by the countries with test observations, and the reduction in estimated effect of school quantity seen here simply mirrors that previously seen in the base case for just countries with direct observations (Table 2).

TABLE 9—ROOT MEAN-SQUARED ERROR FOR ALTERNATIVE GROWTH MODELS

Model (from Table 8)	All countries	Average growth-rate category		
		<1 percent	1–3.5 percent	>3.5 percent
A. *QL1* quality measure*				
Baseline: *C, Y60, S*	1.50	2.16	0.87	2.16
Input: *C, Y60, S, PT-pri, PT-sec*	1.50	2.16	0.87	2.17
Input: *C, Y60, S, EXPEND*	1.50	2.14	0.86	2.18
Quality: *C, Y60, S, QL1**	1.28	1.95	0.79	1.71
Number of countries	76	13	47	16
B. *QL2* quality measure*				
Baseline: *C, Y60, S*	1.52	2.08	0.89	2.21
Input: *C, Y60, S, PT-pri, PT-sec*	1.52	2.07	0.89	2.22
Input: *C, Y60, S, EXPEND*	1.52	2.07	0.88	2.23
Quality: *C, Y60, S, QL2**	1.31	1.90	0.87	1.67
Number of countries	78	15	47	16

the combination of "growth conditions" and unmeasured factors, and, by the analysis in Table 9, it is clear that the importance of measured and unmeasured factors differs across countries. By taking out the measured effects, we can identify which countries are growing surprisingly fast or surprisingly slow given human-capital stocks. Figure 2 displays the percentile distributions of countries in terms of average annual growth rates over the period 1960–1990. Figure 3 displays the distribution after conditioning on initial income and quantity and quality of human capital; i.e., it displays countries that do significantly better or worse than would be expected. The countries in boldface are the 13 with a percentile ranking in conditional growth that is 20 points or more higher in the distribution than that in terms of unconditional growth. The countries in italics are the 13 with a percentile ranking in conditional growth that is 20 points or more lower in the distribution than that in terms of unconditional growth. For example, Egypt, with an average growth rate of 2.9 percent, moves from the twenty-seventh-ranked country in unconditional growth to the third-ranked country in growth after conditioning on its human-capital stock. At the same time, Japan, with an average growth rate of 5.3 percent, falls from seventh in unconditional rankings to twenty-fourth after its favorable growth conditions are taken into account.

The distribution of growth rates after eliminating the influence of the measured initial income and human capital provides some guidance in looking to identify a richer set of factors that affect growth. For example, the performance of the United States, identified as having significantly faster growth than expected given its human-capital characteristics (moving from below the median almost to the top 10 percent of the distribution), may partially reflect additional imperfections in the human-capital measure. None of this analysis incorporates information about higher education, where U.S. schools are arguably the world's best. Moreover, in versions of endogenous growth models which emphasize ideas and inventions (e.g., Romer, 1990a), higher education's input into the production of scientists and engineers would be especially important. Alternatively, it may simply reflect the more open and competitive markets in the United States.

It is interesting to put this in the framework of the often-noted growth performance of East Asia (see World Bank, 1993). After taking into account measures of labor-force quality, Singapore, Hong Kong, Korea, Thailand, and Taiwan are still identified as unusually fast growers. In other words, the East Asian miracle includes a significant component that is over and above any emphasis and performance in the development of human capital. At the same time, the Japanese growth pattern looks much less unusual. Thus, without minimizing the importance of human capital for development, it is important to note that a significant component of growth falls elsewhere.

		Panama	Tunisia		
		Ecuador	Norway		
		Sweden	Israel		
		Colombia	Iceland		
		Mexico	Austria		
		United States	Belgium		
		Paraguay	Canada		
	Trinidad&Tobago	Fiji	Egypt	Thailand	
	South Africa	Australia	Yugoslavia	Malaysia	
	New Zealand	Swaziland	France	Portugal	
	Bolivia	Costa Rica	Turkey	Lesotho	
Peru	Kenya	Switzerland	Congo	Greece	Korea
Nicaragua	Jamaica	Cameroon	Brazil	Indonesia	Hong Kong
Venezuela	Zimbabwe	Sri Lanka	Germany	Spain	Singapore
Iraq	Honduras	Dominican Rep.	Netherlands	Barbados	Botswana
Zaire	El Salvador	Togo	Denmark	Ireland	Taiwan
Guyana	Iran	India	Mauritius	Italy	Malta
Zambia	Uruguay	Philippines	Jordan	Finland	Japan
Mozambique	Argentina	Algeria	United Kingdom	Syria	Cyprus
0-10	**10-25**	**25-50**	**50-75**	**75-90**	**90-100**

FIGURE 2. DISTRIBUTION OF COUNTRIES BY AVERAGE ANNUAL GROWTH RATES OF REAL GDP/CAPITA

		Kenya	Spain		
		Trinidad&Tobago	Norway		
		Switzerland	Iceland		
		Ecuador	*Japan*		
		Belgium	Tunisia		
		Swaziland	Finland		
		Barbados	Sweden		
	Venezuela	Jordan	**Algeria**	Taiwan	
	Nicaragua	**Honduras**	*Lesotho*	**United States**	
	Costa Rica	United Kingdom	**Togo**	Portugal	
	Mauritius	Denmark	**Bolivia**	Indonesia	
Jamaica	Phillippines	Paraguay	Ireland	Italy	Botswana
Zaire	Iraq	*Israel*	*Greece*	Malta	Singapore
Peru	*Yugoslavia*	*Congo*	Turkey	Canada	**Egypt**
Mozambique	Sri Lanka	Domican Rep.	**El Salvador**	**Brazil**	Hong Kong
Uruguay	New Zealand	*Netherlands*	Colombia	**India**	Korea
Argentina	Zimbabwe	*Germany*	France	**Iran**	Syria
Zambia	South Africa	Cameroon	Austria	Malaysia	Cyprus
Guyana	*Fiji*	*Panama*	Australia	**Mexico**	Thailand
0-10	**10-25**	**25-50**	**50-75**	**75-90**	**90-100**

FIGURE 3. DISTRIBUTION OF COUNTRIES BY CONDITIONAL GROWTH RATES OF REAL GDP/CAPITA ALLOWING FOR INITIAL INCOME AND QUANTITY AND QUALITY OF HUMAN CAPITAL

Notes: **Bold** (Kenya, Trinidad & Tobago, Honduras, Algeria, Togo, Bolivia, El Salvador, United States, Brazil, India, Iran, Mexico, and Egypt) indicates the country is 20 percentage points or more higher in the conditional distribution than in the unconditional distribution. *ITALICS* (Mauritius, Yugoslavia, Fiji, Belgium, Barbados, Israel, Congo, Netherlands, Germany, Panama, Japan, Lesotho, and Greece) indicates the country is 20 percentage points or more lower in the conditional distribution than in the unconditional distribution.

VIII. Conclusions

Throughout consideration of varying national growth rates, one of the most robust and readily accepted conclusions involves the centrality of a nation's human capital. This conclusion has, nonetheless, come from models with very different specifications of human capital. Virtually all have ignored quality issues, implicitly assuming that any variations in quality of human capital are small relative to the importance and variation in pure quantity of human capital.

The analysis here explicitly considers the quality of the labor force as measured by comparative tests of mathematics and scientific skills. A single conclusion emerges from the various analytical specifications: Labor-force quality has a consistent, stable, and strong relationship with economic growth. A series of indirect investigations of the underlying specification are qualitatively consistent with a causal interpretation. The growth relationship does not appear to be the result of growth causing higher quality through investing resources in schools. Nor does it appear to be driven by high test performance simply being correlated with East Asian countries

that, for other reasons, also achieve high growth. Finally, by looking at how our quality measures relate to immigrants' earnings in the United States, we find clear evidence that international test performance relates to productivity differences. Moreover, these productivity differences appear related to schooling differences and not cultural factors, family support and attitudes, and the like. This direct linkage to productivity suggests a causal impact in international economic performance.

Nonetheless, the estimated impact of quality on growth, indicating that one standard deviation in mathematics and science skills translates into more than one percentage point in average annual real growth, also looks implausibly large. There is no simple and agreed upon way to translate the micro productivity differences from individual earnings into effects for economic growth, making it difficult to be precise. If, for example, quality enters through raising the steady-state level of income and growth reflects a process of conditional convergence, the growth equation results are much larger than the corresponding results for individual earnings. An alternative, which is more consistent with the underlying assumed endogenous growth approach of this paper, would highlight the role of externalities from higher levels of human capital. The growth model results, however, imply that the externalities must be significantly stronger for quality than for quantity. The estimated growth effect of one standard deviation of quality is larger than would be obtained from over nine years in average schooling. Moreover, in absolute terms, this effect is roughly equal to estimates of average rates of technological progress over this period. These arguments suggest the possibility of omitted variables in the growth equations, but we have little indication of what these might be. A number of plausible factors have been ruled out by the specification analyses and by the consideration of alternative models of causal effects.

We conclude that labor-force quality differences are important for growth; that these quality differences are related to schooling (but not necessarily the resources devoted by a country to schooling); and that quality has a causal impact on growth. At the same time, the simple estimates of cross-country growth relationships appear to overstate the causal impact of quality. The precise cause or magnitude of this overstatement is unclear.

There is also a distinct policy dilemma, because standard school resource policies do not relate to the variations in quality that have been identified. We believe that further investigation of the underlying factors guiding quality differences is important.

Appendix A: List of Variables and Sources

Sources of Data

(BL): Barro and Lee (1993 [with 1994 data update])
(ER): Easterly and Rebelo (1993)
(KL): King and Levine (1993)
(SH): Summers and Heston (1991)
(UNESCO): *UNESCO Statistical Yearbook* (various years)

Variables in Analysis

Asia, Latin America, and Africa: regional dummies for Asian, Latin American, and Sub-Saharan African countries (BL: *ASIAE, LAAM,* and *SAFRICA*).

GCN: ratio of real government consumption expenditure net of spending on defense and on education to real GDP (ER: *HSGVXDXE*).

ENROLL-pri: average of total gross enrollment rate for primary education (BL: *Pxx*).

EXPEND: average ratio of nominal government expenditure on education to nominal GDP (BL: *GEETOT*).

GPOP: average of population growth rate (BL: *GPOP*).

GR: annual average growth rate of per capita real GDP for the period of 1960–1990. If data are not available for both end points, subperiods are used for which data are available. Real GDP per capita is SH: *RGDPCH* in 1985 international prices.

PPE: per pupil current public expenditure calculated as *GDP**(*POP*)*(Current expend/GDP)/(Primary + Secondary students)

[$100] (SH: *GDP;* UNESCO: current expenditure, number of primary and secondary students).

PRIINV: average ratio of private investment relative to GDP (BL: *INVWB-INVPUB*).

PT-pri: average pupil/teacher ratio in primary schools (BL: *TEAPRI*).

PT-sec: average pupil/teacher ratio in primary schools (BL: *TEASEC*).

*QL*1, *QL*2: measures of schooling quality constructed as described in text; see below.

*QL*1*, *QL*2*: quality measures expanded to nonparticipants and participants in international comparison studies for student achievement with predicted values for nonparticipants; see text and below.

RECUR: ratio of recurring nominal government expenditure on education to nominal GDP (BL: *GEEREC*).

S: arithmetic average years of schooling for 1960, 1965, 1970, 1975, 1980, and 1985 (BL: *HUMAN*).

TEST: dummy for participants in international comparison studies of student achievement; see text.

TRD: ratio of total trade to GDP (KL).

Y60: initial income, 1960 [$1,000] (SH: *RGDPCH*).

Appendix B: Data for Production Function Estimates

The production function estimates (see Table 3) employ a sample combining different test years in which the dependent variable is the average normalized score for each country and test. Table B1 below describes the dating of exogenous variables for the production function models, based on the specific test year for the performance measure.

Table B1—Dating of Exogenous Variables for the Production Models

	IEA math 1964–1966	IEA science 1966–1973	IEA math 1980–1982	IEA science 1983–1986	IAEP 1988	IAEP 1991
S_{t-1}	1960	1960–1965	1965–1975	1970–1980	1975–1985	1975–1985
$EXPEND_{t-1}$	1960–1964	1960–1969	1965–1974	1970–1979	1975–1984	1975–1984
$PT\text{-}pri_{t-1}$	1950–1960	1955–1965	1960–1970	1965–1975	1970–1980	1970–1980
PPE_{t-1}	1960	1960–1965	1965–1975	1970–1980	1975–1985	1975–1985
$GPOP_{t-1}$	1960–1964	1960–1969	1965–1974	1970–1979	1975–1984	1975–1984

Appendix C: Labor-Force Quality Data

The following data (Table C1) are generated from the methods described in the text. *TEST* = 1 if performance is observed, at which time *QL1** and *QL2** are the observed test data. If *TEST* = 0, *QL1** and *QL2** are projected scores. If blank, the data necessary for projecting scores or for estimating growth models were missing or the projected scores were less than 20 (see text).

Table C1—Labor-Force Quality Data

COUNTRY	*TEST*	*QL1**	*QL2**
1 ALGERIA	0	28.28	28.06
2 ANGOLA	0		
3 ARGENTINA	0	42.99	48.50
4 AUSTRALIA	1	48.13	59.04
5 AUSTRIA	0	53.20	56.61
6 BAHAMAS	0		
7 BAHRAIN	0	26.03	23.19
8 BANGLADESH	0		
9 BARBADOS	0	50.41	59.80
10 BELGIUM	1	53.25	57.08
11 BELIZE	0		
12 BENIN	0		
13 BHUTAN	0		
14 BOLIVIA	0	22.10	27.47
15 BOTSWANA	0	25.05	31.71
16 BRAZIL	1	33.91	36.60
17 BULGARIA	0		
18 BURKINA FASO	0		
19 BURUNDI	0		
20 CAMEROON	0	39.00	42.36
21 CANADA	1	47.57	54.58
22 CAPE VERDE ISLAND	0		
23 CENTRAL AFRICA, REPUBLIC OF	0		24.77
24 CHAD	0		
25 CHILE	1	26.30	24.74
26 CHINA	1	59.28	64.42
27 COLOMBIA	0	34.78	37.87
28 COMOROS	0		
29 CONGO	0	46.04	50.90
30 COSTA RICA	0	42.15	46.15
31 CYPRUS	0	42.24	46.24
32 CZECHOSLOVAKIA	0		
33 DENMARK	0	53.48	61.76
34 DJIBOUTI	0		
35 DOMINICA	0		
36 DOMINICAN REPUBLIC	0	37.41	39.34
37 ECUADOR	0	35.78	38.99
38 EGYPT	0	26.01	26.43
39 EL SALVADOR	0	21.81	26.21
40 ETHIOPIA	0		
41 FIJI	0	50.02	58.10
42 FINLAND	1	48.76	59.55
43 FRANCE	1	54.15	56.00
44 GABON	0		
45 GAMBIA	0		
46 GERMANY, WEST	1	59.03	48.68
47 GHANA	0		25.58
48 GREECE	0	49.11	50.88
49 GRENADA	0		
50 GUATEMALA	0		
51 GUINEA-BISS	0		
52 GUINEA	0		
53 GUYANA	0	45.71	51.49
54 HAITI	0		
55 HONDURAS	0	26.43	28.59
56 HONG KONG	1	56.93	71.85
57 HUNGARY	1	53.85	61.23
58 ICELAND	0	48.13	51.20
59 INDIA	1	21.63	20.80
60 INDONESIA	0	37.98	42.99
61 IRAN	1	20.79	18.26
62 IRAQ	0	29.34	27.50
63 IRELAND	1	47.59	50.20
64 ISRAEL	1	51.29	54.46
65 ITALY	1	44.59	49.41
66 IVORY COAST	0		
67 JAMAICA	0	44.19	48.62
68 JAPAN	1	60.65	65.50
69 JORDAN	1	39.38	42.28
70 KENYA	0	24.43	29.73
71 KOREA, REPUBLIC OF	1	56.21	58.55
72 KUWAIT	0	28.36	22.50
73 LAOS	0		
74 LESOTHO	0	46.14	51.95
75 LIBERIA	0		
76 LUXEMBOURG	1	39.45	44.49
77 MADAGASCAR	0		
78 MALAWI	0		
79 MALAYSIA	0	47.89	54.29
80 MALI	0		
81 MALTA	0	53.16	57.14
82 MAURITANIA	0		
83 MAURITIUS	0	49.53	54.95
84 MEXICO	0	35.06	37.24
85 MONGOLIA	0		
86 MOROCCO	0		
87 MOZAMBIQUE	1	24.26	27.94
88 MYANMAR	0		
89 NAMIBIA	0		
90 NEPAL	0		
91 NETHERLANDS	1	56.84	54.52
92 NEW ZEALAND	1	52.44	67.06
93 NICARAGUA	0	24.19	27.30
94 NIGERIA	1	34.15	38.90
95 NIGER	0		
96 NORWAY	1	49.60	64.56
97 OMAN	0		
98 PAKISTAN	0		
99 PANAMA	0	42.02	46.78
100 PAPUA NEW GUINEA	0		22.58
101 PARAGUAY	0	37.99	39.96
102 PERU	0	37.83	41.18

TABLE C1—*CONTINUED*

COUNTRY	*TEST*	*QL1**	*QL2**
103 PHILIPPINES	1	34.35	33.54
104 POLAND	1	50.28	64.37
105 PORTUGAL	1	44.09	44.22
106 PUERTO RICO	0		
107 QATAR	0		
108 REUNION	0		
109 ROMANIA	0		
110 RWANDA	0		
111 SAUDI ARABIA	0		
112 SENEGAL	0		
113 SEYCHELLES	0		
114 SIERRA LEONE	0		
115 SINGAPORE	1	56.51	72.13
116 SOLOMON ISLANDS	0		
117 SOMALIA	0		
118 SOUTH AFRICA	0	45.25	51.30
119 SPAIN	1	49.40	51.92
120 SRI LANKA	0	41.54	42.57
121 ST. LUCIA	0		
122 ST. VINCENT	0		
123 SUDAN	0		
124 SURINAME	0		
125 SWAZILAND	1	35.46	40.26
126 SWEDEN	1	47.41	57.43
127 SWITZERLAND	1	57.17	61.37
128 SYRIA	0	31.66	30.23
129 TAIWAN	1	56.28	56.31
130 TANZANIA	0		
131 THAILAND	1	39.83	46.26
132 TOGO	0	28.08	32.69
133 TONGA	0		
134 TRINIDAD AND TOBAGO	0	40.57	46.43
135 TUNISIA	0	41.79	40.50
136 TURKEY	0	41.52	39.72
137 UGANDA	0		
138 URUGUAY	0	46.33	52.27
139 UNITED ARAB EMIRATES	0		
140 UNITED KINGDOM	1	53.98	62.52
141 UNITED STATES	1	43.43	46.77
142 UNION OF SOVIET SOCIALIST REPUBLICS	1	53.89	54.65
143 VAUATU	0		
144 VENEZUELA	0	36.78	39.08
145 WESTERN SAMOA	0		
146 YEMEN	0		
147 YUGOSLAVIA	0	50.91	53.97
148 ZAIRE	0	30.03	33.53
149 ZAMBIA	0	30.54	36.61
150 ZIMBABWE	0	35.97	39.64

REFERENCES

Barro, Robert J. "Economic Growth in a Cross Section of Countries." *Quarterly Journal of Economics*, May 1991, *106*(2), pp. 407–43.

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

Barro, Robert J. and Sala-i-Martin, Xavier. *Economic growth.* New York: McGraw-Hill, 1995.

Becker, Gary S. and Lewis, H. Gregg. "On the Interaction between the Quantity and Quality of Children." *Journal of Political Economy*, March–April 1973, *81*(2), pp. S279–88.

Benhabib, Jess and Spiegel, Mark M. "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data." *Journal of Monetary Economics*, October 1994, *34*(2), pp. 143–74.

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

Bishop, John. "The Impact of Academic Competencies of Wages, Unemployment, and Job Performance." *Carnegie-Rochester Conference Series on Public Policy*, December 1992, *37*, pp. 127–94.

Decker, Paul T. and Radbill, Larry M. "Learning to Grow: Educational Quality and Economic Growth in 32 Countries." Mathematica Policy Research, 1999.

Easterly, William and Rebelo, Sergio. "Fiscal Policy and Economic Growth: An Empirical Investigation." *Journal of Monetary Economics*, December 1993, *32*(3), pp. 417–58.

Hanushek, Eric A. "Conceptual and Empirical Issues in the Estimation of Educational Production Functions." *Journal of Human Resources*, Summer 1979, *14*(3), pp. 351–88.

______. "The Economics of Schooling: Production and Efficiency in Public Schools." *Journal of Economic Literature*, September 1986, *24*(3), pp. 1141–77.

______. "The Trade-Off Between Child Quantity and Quality." *Journal of Political Economy*, February 1992, *100*(1), pp. 84–117.

______. "Interpreting Recent Research on Schooling in Developing Countries." *World Bank Research Observer*, August 1995, *10*(2), pp. 227–46.

______. "School Resources and Student Performance," in Gary Burtless, ed., *Does money matter? The effect of school resources on student achievement and adult success.* Washington, DC: Brookings Institution Press, 1996, pp. 43–73.

Hanushek, Eric A. and Kim, Dongwook. "Schooling, Labor Force Quality, and Economic Growth." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 5399, December 1995.

Hanushek, Eric A. and Kim, Jin-Yeong. "Immigration and Home-Country School Quality." Mimeo, University of Rochester, 1999.

Jorgenson, Dale W. and Fraumeni, Barbara M. "Investment in Education and U.S. Economic Growth." *Scandinavian Journal of Economics*, 1992, Supp., *94*, pp. 51–70.

King, Robert G. and Levine, Ross. "Finance, Entrepreneurship, and Growth: Theory and Evidence." *Journal of Monetary Economics*, December 1993, *32*(3), pp. 513–42.

Klenow, Peter J. and Rodríguez-Clare, Andrés. "The Neoclassical Revival in Growth Economics: Has It Gone Too Far?" in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER macroeconomics annual 1997*. Cambridge, MA: MIT Press, 1997, pp. 83–103.

Krueger, Alan B. and Lindahl, Mikael. "Education for Growth in Sweden and the World." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 7190, 1999.

Lee, Jong-Wha and Barro, Robert J. "Schooling Quality in a Cross Section of Countries." National Bureau of Economic Research (Cambridge, MA) Working Paper No. 6198, September 1997.

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

Levine, Ross and Zervos, Sara J. "What We Have Learned About Policy and Growth From Cross-Country Regressions?" *American Economic Review*, May 1993 (*Papers and Proceedings*), *83*(2), pp. 426–30.

Lockheed, Marlaine E. and Verspoor, Adriaan. *Improving primary education in developing countries*. New York: Oxford University Press, 1991.

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

Mincer, Jacob. *Schooling, experience, and earnings*. New York: National Bureau of Economic Research, 1974.

______. "Economic Development, Growth of Human Capital, and the Dynamics of the Wage Structure." *Journal of Economic Growth*, March 1996, *1*(1), pp. 29–48.

Nelson, Richard R. and Phelps, Edmund. "Investment in Humans, Technology Diffusion, and Economic Growth." *American Economic Review*, May 1966 (*Papers and Proceedings*), *56*(2), pp. 69–75.

Rebelo, Sergio. "Long-Run Policy Analysis and Long-Run Growth." *Journal of Political Economy*, June 1991, *99*(3), pp. 500–21.

Romer, Paul. "Endogenous Technological Change." *Journal of Political Economy*, October 1990a, Pt. 2, *99*(5), pp. S71–S102.

______. "Human Capital and Growth: Theory and Evidence." *Carnegie-Rochester Conference Series on Public Policy*, Spring 1990b, *32*, pp. 251–86.

Summers, Robert and Heston, Alan. "The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950–1988." *Quarterly Journal of Economics*, May 1991, *106*(2), pp. 327–68.

UNESCO Statistical Yearbook. Paris: UNESCO Publishing, various years.

U.S. Department of Education. *The condition of education 1994*. Washington, DC: National Center for Education Statistics, 1994.

______. *The condition of education 1996*. Washington, DC: National Center for Education Statistics, 1996a.

______. *Pursuing excellence: A study of U.S. eighth-grade mathematics and science teaching, learning, curriculum, and achievement in international context*. Washington, DC: National Center for Education Statistics, 1996b.

Welch, Finis. "Education in Production." *Journal of Political Economy*, January–February 1970, *78*(1), pp. 35–59.

White, Halbert. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica*, May 1980, *48*(4), pp. 817–38.

Willis, Robert J. "A New Approach to the Economic Theory of Fertility Behavior." *Journal of Political Economy*, April–March 1973, Pt. 2, *81*(2), pp. S14–64.

World Bank. *The East Asian miracle*. New York: Oxford University Press, 1993.