

The Role of Education Quality in Economic Growth*

Eric A. Hanushek

Hoover Institution
Stanford University
CESifo and NBER
Stanford, CA 94305-6010, United States
Phone: (+1) 650 / 736-0942
E-mail: hanushek@stanford.edu
Internet: www.hanushek.net

Ludger Wößmann

University of Munich,
Ifo Institute for Economic Research and CESifo
Poschingerstr. 5
81679 Munich, Germany
Phone: (+49) 89 / 9224-1699
E-mail: woessmann@ifo.de
Internet: www.cesifo.de/woessmann

Abstract

The role of improved schooling, a central part of most development strategies, has become controversial because expansion of school attainment has not guaranteed improved economic conditions. This paper reviews the role of education in promoting economic well-being, with a particular focus on the role of educational quality. It concludes that there is strong evidence that the cognitive skills of the population – rather than mere school attainment – are powerfully related to individual earnings, to the distribution of income, and to economic growth. New empirical results show the importance of both minimal and high level skills, the complementarity of skills and the quality of economic institutions, and the robustness of the relationship between skills and growth. International comparisons incorporating expanded data on cognitive skills reveal much larger skill deficits in developing countries than generally derived from just school enrollment and attainment. The magnitude of change needed makes clear that closing the economic gap with developed countries will require major structural changes in schooling institutions.

World Bank Policy Research Working Paper 4122, February 2007

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the view of the World Bank, its Executive Directors, or the countries they represent. Policy Research Working Papers are available online at <http://econ.worldbank.org>.

* This project developed through conversations with Harry Patrinos, who provided useful comments and suggestions along the way. We have also benefited from comments by Martha Ainsworth, Luis Benveniste, François Bourguignon, Deon Filmer, Paul Gertler, Manny Jimenez, Ruth Kagia, Beth King, Lant Pritchett, and Emiliana Vegas. Support has come from the World Bank, CESifo, the Program on Education Policy and Governance of Harvard University, and the Packard Humanities Institute.

The Role of School Improvement in Economic Development

By Eric A. Hanushek and Ludger Wößmann

| | |
|---|-----------|
| 1. Introduction | 1 |
| 2. Individual Returns to Education and Economic Inequality | 5 |
| 2.1 Impacts of School Attainment on Individual Incomes..... | 5 |
| 2.2 Impacts of Educational Quality on Individual Incomes—Developed Countries..... | 6 |
| 2.3 Impacts of Educational Quality on Individual Incomes—Developing Countries..... | 11 |
| 2.4 Evidence from the International Adult Literacy Survey | 14 |
| 2.5 Causality..... | 16 |
| 2.6 Income Distribution | 17 |
| 3. Quantity of Schooling and Economic Growth | 20 |
| 3.1 Results of Initial Cross-country Growth Regressions | 20 |
| 3.2 More Recent Evidence on the Effects of Levels of and Growth in Years of Schooling..... | 22 |
| 4. Quality of Education and Economic Growth | 25 |
| 4.1 A Review of the Basic Results..... | 25 |
| 4.2 Issues of Endogeneity | 29 |
| 4.3 Some New Evidence | 31 |
| 4.4 Distribution of Educational Quality and Economic Growth..... | 38 |
| 4.5 Institutions, Education and Growth..... | 40 |
| 4.6 The Implications of Improved Quality..... | 43 |
| Appendix: Data on Quality of Education..... | 47 |
| 5. Where Does the Developing World Stand? | 51 |
| 5.1 Lack of Quantity of Schooling..... | 51 |
| 5.2 Lack of Quality of Education..... | 52 |
| 5.3 The Size of the Task at Hand: Schooling Quantity and Educational Quality Combined | 55 |
| 6. Educational Spending and Student Outcomes | 59 |
| 6.1 Cross-country Evidence on Resources..... | 60 |
| 6.2 Within-country Evidence – Developed Countries | 63 |
| 6.3 Within-country Evidence – Developing Countries | 66 |
| 6.4 Is There a Minimum Resource Requirement? | 67 |
| 7. Schooling Institutions and Educational Quality | 68 |
| 7.1 Choice and Competition in Developing Countries | 68 |
| 7.2 Evidence on Autonomy of Schools..... | 70 |
| 7.3 School Accountability..... | 71 |
| 7.4 Summary of How to Improve the Quality of Education | 74 |
| 8. Conclusion | 76 |
| References | 80 |

1. Introduction

It takes little analysis to see that education levels differ dramatically between developing and developed countries. Building upon several decades of thought about human capital – and centuries of general attention to education in the more advanced countries – it is natural to believe that a productive development strategy would be to raise the schooling levels of the population. And, indeed, this is exactly the approach of the Education for All initiative and a central element of the Millennium Development Goals.

But there are also some nagging uncertainties that exist with this strategy. First, developed and developing countries differ in a myriad of ways other than schooling levels. Second, a number of countries – both on their own and with the assistance of others – have expanded schooling opportunities without seeing any dramatic catch-up with developed countries in terms of economic well-being. Third, countries that do not function well in general might not be more able to mount effective education programs than they are to pursue other societal goals. Fourth, even when schooling policy is made a focal point, many of the approaches undertaken do not seem very effective and do not lead to the anticipated student outcomes. In sum, is it obvious that education is the driving force, or merely one of several factors that are correlated with more fundamental development forces?

The objective of this study is to review what is known about the role of education in promoting economic well-being. We are interested in assessing what research says about these issues. More than that, we pay particular attention to the credibility of this research in establishing a causal relationship between education and economic outcomes and between policy initiatives and educational outcomes.

The discussion also has one distinctive element. We have come to conclude that educational quality – particularly in assessing policies related to developing countries – is THE key issue. It is both conventional and convenient in policy discussions to concentrate on such things as years of school attainment or enrollment rates in schools. These things are readily observed and measured. They appear in administrative data, and they are published on a consistent basis in virtually all countries of the world. And, they are very misleading in the policy debates.

We will show in graphic terms the differences in educational quality that exist. Most people would, in casual conversation, acknowledge that a year of schooling in a school in a Brazilian Amazon village was not the same as a year of schooling in a school in Belgium. They would also agree that families, peers, and others contribute to education. Yet, research on the economic impact of schools – largely due to expedience – almost uniformly ignores these. The data suggest that the casual conversation may actually tend to understate the magnitude of differences.

We will also provide strong evidence that ignoring quality differences significantly distorts the picture about the relationship between education and economic outcomes. This distortion occurs at three levels. It misses important differences between education and skills on the one hand and individual earnings on the other. It misses an important underlying factor determining the interpersonal distribution of incomes across societies. And, it very significantly misses the important element of education in economic growth.

The plan of this study is straightforward. We begin by documenting the importance of cognitive skills – the measure of educational quality we use – in determining individual earnings, and by implication important aspects of the income distribution. We then turn to the relationship of education and economic growth. Research into the economics of growth has itself been a growth area, but much of the research focuses just on school attainment with no consideration of quality differences or of other sources of learning. We show, in part with new evidence, that the evidence is highly biased by its concentration on just quantity of schooling.

In both of these areas, attention has been given to causality; i.e., is it reasonable to believe that changing education would directly lead to a change in economic outcomes? Again, the concentration on quantity of schooling has distorted these discussions of causality, and consideration of quality considerably alters the issues and implications.

The simple answers in the discussion of economic implications of education are that educational quality, measured by cognitive skills, has a strong impact on individual earnings. More than that, however, educational quality has a strong and robust influence on economic growth. In both areas, there is credible evidence that these are truly causal relationships.

To be sure, none of this says that schools per se are the answer. Even though it is common to treat education and schooling synonymously, it is important to distinguish between knowledge and skills on the one hand (educational quality in our terminology) and schooling. This semantic distinction has important substantive underpinnings. Cognitive skills may be developed in formal schooling, but they may also come from the family, the peers, the culture, and so forth. Moreover, other factors obviously have an important impact on earnings and growth. For example, overall economic institutions – a well-defined system of property rights, the openness of the economy, the security of the nation – can be viewed almost as preconditions to economic development. And, without them, education and skills may not have the desired impact on economic outcomes.

Yet, while recognizing the impact of these overall institutions, we find that schools can play an important role. Quality schools can lead to improved educational outcomes. Moreover, from a public

policy perspective, interventions in the schools are generally viewed as both more acceptable and more likely to succeed than, say, direct interventions in the family.

Given the evidence on the importance of educational quality for economic outcomes, the study turns to important policy issues. To begin with, what can be said about the educational quality and cognitive skills in developing countries? Although information on enrollment and attainment has been fairly widely available, quality information has not. We use newly developed data on international comparisons of cognitive skills (also employed in the analysis of growth) to show that the education deficits in developing countries are larger than previously appreciated.

Discussions of quality inevitably lead to questions about whether it can be affected by policy. An extensive literature, albeit one biased toward developed countries, now exists on a number of policy issues. Perhaps most well known is that simply putting more resources into schools – pure spending, reduced class sizes, increased teacher training, and the like – will not reliably lead to improvements in student outcomes.

These findings are, however, often misinterpreted. First, they do not imply that schools have no effect. They say simply that common measures of school quality are in reality not closely related to student outcomes, but this is not the same as finding that school quality differences do not exist. Second, the findings do not say that spending and resources never matter. Indeed, there is some indication, particularly in developing countries, that a range of resources are important – textbooks, rudimentary facilities, and the like. The potential impacts of these are nonetheless too small to be instruments for radical changes in outcomes, something that the prior evidence indicates is needed in many developing countries. Third, the findings do not say that resources cannot matter. They indicate that resources may not have any consistent effects within the current structure and institutions of schools, but the findings do not put resource discussions into the context of alternative structures.

One consistent finding that is emerging from research, albeit largely from developed country experiences, is that teacher quality has powerful impacts on student outcomes. The problem from a policy aspect is, however, that quality differences are not closely related to the common measures of quality and to the common policy instruments that are employed. Within countries where the data exist, there is little indication that quality is closely related to teacher education and training, teacher experience, teacher certification, or teacher salaries. These facts disrupt the policy discussions. They also make it clear that different sets of policies must be contemplated if schools are to improve.

A different view of schools, however, concentrates on larger institutional issues. There is growing evidence that a number of devices – things that effectively change the existing incentives in schools – have an impact. Accountability systems based upon tests of student cognitive achievement can change

the incentives for both school personnel and for students. By focusing attention on the true policy goal – instead of imperfect proxies based on inputs to schools – performance can be improved. These systems align rewards with outcomes. Moreover, increased local decision making or local autonomy, coupled with accountability, can facilitate these improvements.

The evidence on a set of larger, and potentially more powerful, policy changes is relatively limited at the current time. There is suggestive evidence that greater school choice promotes better performance. Further, direct incentives to teachers and school personnel in the form of performance pay have promise. Unfortunately, however, these policies can lead to substantial changes in the incentives within schools, and such substantial changes are frequently resisted by current school personnel. Current employees, often through their unions, generally tend to resist and to stop even experimentation with such changes. Thus, direct evidence on them is more limited, and may require more inferences. Nonetheless, there remains reason to believe that pursuing these larger changes could lead to the substantial improvements in outcomes that are desired or hoped for in the policy process.

2. Individual Returns to Education and Economic Inequality

2.1 *Impacts of School Attainment on Individual Incomes*

Most attention to the value of schooling focuses on the economic returns to differing levels of school attainment for individuals. This work, following the innovative analyses of human capital by Jacob Mincer (1970, 1974), considers how investing in differing amounts of schooling affects individual earnings. Over the past thirty years, literally hundreds of such studies have been conducted around the world.¹

These studies have uniformly shown that more schooling is associated with higher individual earnings. The rate of return to schooling across countries is centered at about 10 percent with variations in expected ways based largely on scarcity: returns appear higher for low income countries, for lower levels of schooling, and, frequently, for women (Psacharopoulos and Patrinos (2004)).

Much of the academic debate has focused on whether these simple estimates provide credible measures of the causal effect of schooling. In particular, if more able people tend also to obtain additional schooling, the estimated schooling effect could include both the impacts of schooling and the fact that those continuing in school could earn more in the absence of schooling.² For the most part, employing alternative estimation approaches dealing with the problems of endogeneity of schooling do not lead to large changes in the estimates, and many times they suggest that the returns are actually larger with the alternative estimation schemes than with the simpler modeling strategies.

The basic estimates of Mincer earnings models are typically interpreted as the private returns to schooling. As is well known, the social returns could differ from the private returns – and could be either above or below the private returns. The most common argument is that the social returns will exceed the private returns because of the positive effects of education on crime, health, fertility,

¹ A variety of studies review and interpret the basic estimation of rates of return. See Psacharopoulos (1994), Card (1999), Harmon, Oosterbeek, and Walker (2003), Psacharopoulos and Patrinos (2004), and Heckman, Lochner, and Todd (2006).

² Harmon, Oosterbeek, and Walker (2003) systematically review the various issues and analytical approaches dealing with them along with providing a set of consistent estimates of returns (largely for OECD countries). They conclude that, while the estimation approaches can have an impact on the precise value of the rate of return, it is clear that there is a strong causal impact of school attainment on earnings.

improved citizen participation,³ and (as we discuss below) on growth and productivity of the economy as a whole.⁴ If on the other hand schooling was more of a selection device than of a means of boosting knowledge and skills of individuals, the social return could be below the private return.⁵ Although there are many uncertainties about precisely how social returns might differ from private returns, there is overall little reason to believe that the social returns are less than the private returns, and there are a variety of reasons to believe that they could be noticeably higher.

2.2 Impacts of Educational Quality on Individual Incomes—Developed Countries

The concentration on school attainment in the academic literature, however, contrasts with much of the policy discussion that, even in the poorest areas, involves elements of “quality” of schooling. Most countries are involved in policy debates about the improvement of their schools. These debates, often phrased in terms of such things as teacher salaries or class sizes, rest on a presumption that there is a high rate of return to schools in general and to quality in particular.

But it is not appropriate simply to presume that any spending on schools is a productive investment that will see the returns estimated for attainment. It is instead necessary to ascertain two things: how various investments translate into quality and how that quality relates to economic returns. This section provides a summary of what is known about the individual returns to educational quality in both developed and developing countries.

One of the challenges to understanding the impact of quality differences in human capital has been simply knowing how to measure quality. Much of the discussion of quality—in part related to new efforts to provide better accountability—has identified cognitive skills as the important dimension.

³ Recent studies indeed find evidence of externalities of education in such areas as reduced crime (Lochner and Moretti (2004)), improved health of children (Currie and Moretti (2003)), and improved civic participation (Dee (2004); Milligan, Moretti, and Oreopoulos (2004)). The evidence on direct production spillovers of education among workers is more mixed, with Moretti (2004) and the studies cited therein finding favorable evidence and Acemoglu and Angrist (2000) and Ciccone and Peri (2006) finding no evidence for this kind of spillovers.

⁴ In the Mincer earnings work, social rates of return are frequently calculated. These calculations are not based on the positive externalities cited but instead on the fact that the social cost of subsidized education exceeds the private costs – thus lowering the social rate of return relative to the private rate of return (see Psacharopoulos and Patrinos (2004)).

⁵ The empirical analysis of these issues has been very difficult because the labor market outcomes of the screening/selection model and the productivity/human capital model are very similar if not identical. Lange and Topel (2006) review the theory and empirical work and conclude that there is little evidence that the social rate of return to schooling is below the private rate of return.

And, while there is ongoing debate about the testing and measurement of these skills, most parents and policy makers alike accept the notion that cognitive skills are a key dimension of schooling outcomes. The question is whether this proxy for school quality—students’ performance on standardized tests—is correlated with individuals’ performance in the labor market and the economy’s ability to grow.

Until fairly recently, little comprehensive data have been available to show any relationship between differences in cognitive skills and any related economic outcomes. The many analyses of school attainment and Mincer earnings functions rely upon readily available data from censuses and other surveys, which find it easy to collect information on earnings, school attainment, age, and other demographic information. On the other hand, it is difficult to obtain data on cognitive skills along with earnings and the other determinants of wages. Although cognitive test and school resource data are increasingly available at the time of schooling, these are seldom linked to subsequent labor market information. Such analyses generally require tracking individuals over time, a much more difficult data collection scheme. Such data are, however, now becoming available.

A variety of researchers are now able to document that the earnings advantages to higher achievement on standardized tests are quite substantial.⁶ While these analyses emphasize different aspects of individual earnings, they typically find that measured achievement has a clear impact on earnings after allowing for differences in the quantity of schooling, the experiences of workers, and other factors that might also influence earnings. In other words, higher quality as measured by tests similar to those currently being used in accountability systems around the world is closely related to individual productivity and earnings.

Three recent U.S. studies provide direct and quite consistent estimates of the impact of test performance on earnings (Mulligan (1999); Murnane, Willett, Duhaldeborde, and Tyler (2000); Lazear (2003)). These studies employ different nationally representative data sets that follow students after they leave school and enter the labor force. When scores are standardized, they suggest that one

⁶ These results are derived from different specific approaches, but the basic underlying analysis involves estimating a standard “Mincer” earnings function and adding a measure of individual cognitive skills. This approach relates the logarithm of earnings to years of schooling, experience, and other factors that might yield individual earnings differences. The clearest analyses are found in the following references for the U.S. (which are analyzed in Hanushek (2002b)). See Bishop (1989, 1991); O’Neill (1990); Grogger and Eide (1993); Blackburn and Neumark (1993, 1995); Murnane, Willett, and Levy (1995); Neal and Johnson (1996); Mulligan (1999); Murnane, Willett, Duhaldeborde, and Tyler (2000); Altonji and Pierret (2001); Murnane, Willett, Braatz, and Duhaldeborde (2001); and Lazear (2003).

standard deviation increase in mathematics performance at the end of high schools translates into 12 percent higher annual earnings.⁷

Murnane, Willett, Duhaldeborde, and Tyler (2000) provide evidence from the High School and Beyond and the National Longitudinal Survey of the High School Class of 1972. Their estimates suggest some variation with males obtaining a 15 percent increase and females a 10 percent increase per standard deviation of test performance. Lazear (2003), relying on a somewhat younger sample from NELS88, provides a single estimate of 12 percent. These estimates are also very close to those in Mulligan (1999), who finds 11 percent for the normalized AFQT score in the NLSY data.⁸ Note that these returns can be thought of as how much earnings would increase with higher quality each and every year throughout the persons' working career. Thus, the present value of the returns to higher quality is large.

These estimates are obtained fairly early in the work career (mid-20s to early 30s), and analyses of the impact of cognitive skills across the entire work life are more limited. Altonji and Pierret (2001) find that the impact of achievement on earnings grows with experience, because the employer has a chance to observe the performance of workers. The pattern of how returns change with age from their analysis is shown in Figure 2.1, where the power of school attainment differences to predict differences in earnings is replaced by cognitive skills as workers are in the labor force longer. The evidence is consistent with employers relying on readily available information on school attainment when they do not have other information and switching to observations of skills and performance as that information becomes available through job performance.⁹ On the other hand, Hanushek and Zhang (2006) do not find that this pattern holds for a wider set of countries (although it continues to hold for the United

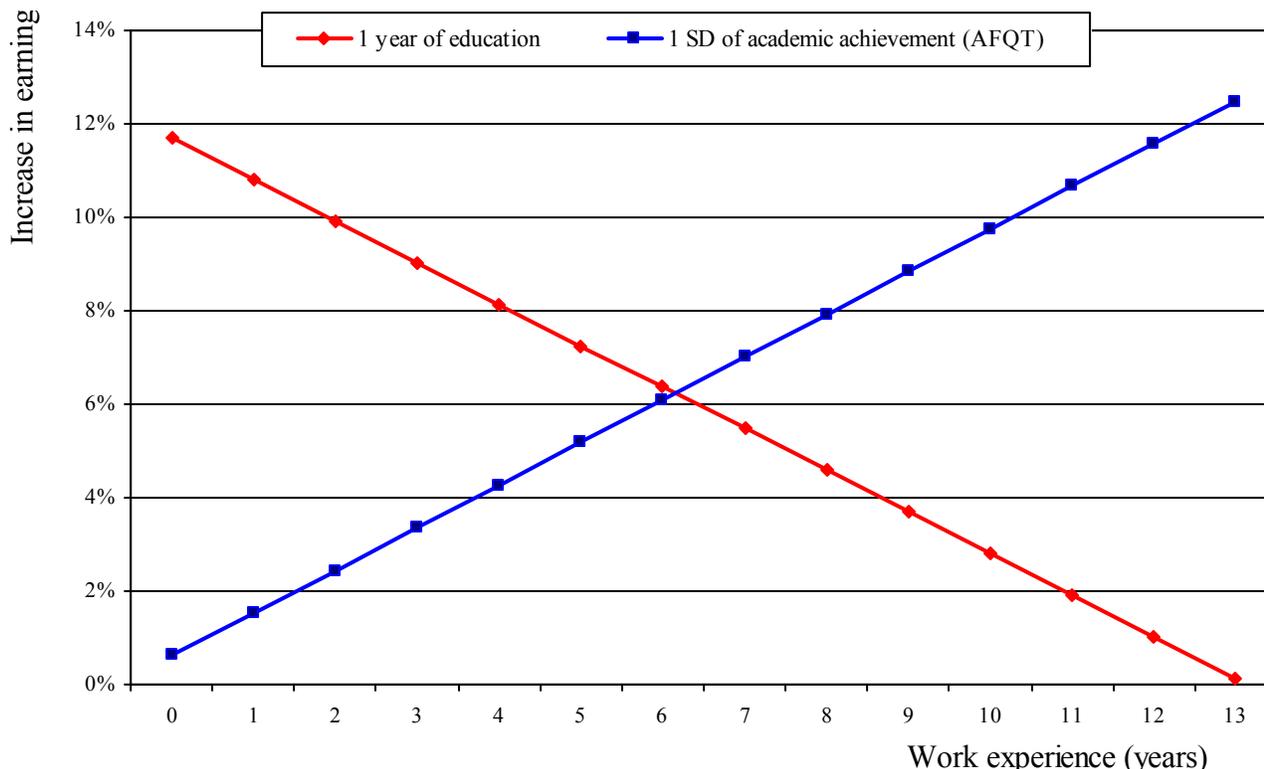
⁷ Because the units of measurement differ across tests, it is convenient to convert test scores into measures of the distribution of achievement across the population. A one-half standard deviation change would move somebody from the middle of the distribution (the 50th percentile) to the 69th percentile; a one standard deviation change would move this person to the 84th percentile. Because tests tend to follow a bell-shaped distribution, the percentile movements are largest at the center of the distribution.

⁸ By way of comparison, we noted that estimates of the value of an additional year of school attainment are typically about 10 percent. Of course, any investment decisions must recognize that quality and quantity are generally produced together and that costs of changing each must be taken into account.

⁹ Note that Altonji and Pierret (2001) observe a limited age range, so that these changing returns may well be thought of as leveling off after some amount of labor market experience.

States in their data). Thus, there is some uncertainty currently about whether cognitive skills have differential effects on economic outcomes over the work-experience profile.¹⁰

Figure 2.1: Returns to Observed Educational Quantity and Unobserved Educational Quality over the Work Life



Notes: Based on data from National Longitudinal Survey of Youth (NLSY) and Armed Forces Qualification Test (AFQT). SD = standard deviation.

Source: Based on Altonji and Pierret (2001).

There are other reasons to believe that these estimates provide a lower bound on the impact of higher achievement. First, the labor market experiences that are observed begin in the mid-1980s and extend into the mid-1990s, but other evidence suggests that the value of skills and of schooling has grown throughout and past that period. Second, extrapolating from recent patterns, future general improvements in productivity are likely to lead to larger returns to skill.¹¹ Finally, cognitive skills

¹⁰Hanushek and Zhang (2006), discussed below, do find that the importance of cognitive skills is not restricted just to younger workers (represented in most prior research) but holds across the experience spectrum.

¹¹ The considered analyses typically compare workers of different ages at one point in time to obtain an estimate of how earnings will change for any individual. If, however, productivity improvements occur in the economy, these will tend to raise the earnings of individuals over time. If recent patterns of skill bias in productivity improvement continue, the

measured by test scores are prone to considerable measurement error. Even if the tests were measuring exactly the relevant skill concept, we know that there are substantial errors in each test.¹² These errors will in general lead to downward biases in the estimated coefficients.

A limited number of additional studies are available for developed countries outside of the United States. McIntosh and Vignoles (2001) study wages in the United Kingdom and find strong returns to both numeracy and literacy.¹³ Finnie and Meng (2002) and Green and Riddell (2003) investigate returns to cognitive skills in Canada. Both suggest that literacy has a significant return, but Finnie and Meng (2002) find an insignificant return to numeracy. This latter finding stands at odds with most other analyses that have emphasized numeracy or math skills.

Another part of the return to school quality comes through continuation in school.¹⁴ There is substantial U.S. evidence that students who do better in school, either through grades or scores on standardized achievement tests, tend to go farther in school.¹⁵ Murnane, Willett, Duhaldeborde, and

impact of improvements in student skills are likely to rise over the work life instead of being constant as portrayed here (cf. Katz and Murphy (1992)). On the other hand, such skill-biased change has not always been the case, and technology could push returns in the opposite direction.

¹² In most testing situations, both the reliability of the specific test and the validity of the test are considered. Reliability relates to how well the test measures the specific material – and would include elements of specific question development and choice along with individual variations that would occur if an individual took the same test at different points in time. Validity refers to the correspondence between the desired concept (skills related to productivity or earnings differences) and the specific choice of test domains (such as mathematical concepts at some specific level).

¹³ Because they look at discrete levels of skills, it is difficult to compare the quantitative magnitudes directly to the U.S. work.

¹⁴ As noted above, much of the work by economists on differences in worker skills has actually been directed at the issue of determining the average labor market returns to additional schooling. The argument has been that higher-ability students are more likely to continue in schooling. Therefore, part of the higher earnings observed for those with additional schooling really reflects pay for added ability and not for the additional schooling. Economists have pursued a variety of analytical approaches for dealing with this, including adjusting for measured cognitive test scores, but this work generally ignores issues of variation in school quality. The various adjustments for ability differences typically make small differences on the estimates of the value of schooling, and Heckman and Vytalacil (2001) argue that it is not possible to separate the effects of ability and schooling. The only explicit consideration of school quality typically investigates expenditure and resource differences across schools, but these are known to be poor measures of school quality differences (Hanushek (2002b)).

¹⁵ See, for example, Dugan (1976); Manski and Wise (1983). Rivkin (1995) finds that variations in test scores capture a considerable proportion of the systematic variation in high school completion and in college continuation, so that test score differences can fully explain black-white differences in schooling. Bishop (1991) and Hanushek, Rivkin, and Taylor

Tyler (2000) separate the direct returns to measured skill from the indirect returns of more schooling and suggest that perhaps one-third to one-half of the full return to higher achievement comes from further schooling. Note also that the effect of quality improvements on school attainment incorporates concerns about drop out rates. Specifically, higher student achievement keeps students in school longer, which will lead among other things to higher graduation rates at all levels of schooling.¹⁶

Knighton and Bussière (2006) find that higher scores at age 15 lead to significantly higher rates of post-secondary schooling of Canadian 19-year-olds. This finding is particularly interesting for the international comparisons that we consider below, because the analysis follows up on precisely the international testing that is used in our analysis of economic growth.¹⁷

2.3 Impacts of Educational Quality on Individual Incomes—Developing Countries

Questions remain about whether the clear impacts of quality in the U.S. generalize to other countries, particularly developing countries. The literature on returns to cognitive skills in developing countries is restricted to a relatively limited number of countries: Ghana, Kenya, Morocco, Pakistan, South Africa, and Tanzania. Moreover, a number of studies actually employ the same basic data, albeit with different analytical approaches, but come up with somewhat different results. Table 2.1 provides a simple summary of the quantitative estimates available for developing countries.

(1996), in considering the factors that influence school attainment, find that individual achievement scores are highly correlated with continued school attendance. Neal and Johnson (1996) in part use the impact of achievement differences of blacks and whites on school attainment to explain racial differences in incomes. Their point estimates of the impact of cognitive skills (AFQT) on earnings and school attendance appear to be roughly comparable to that found in Murnane, Willett, Duhaldeborde, and Tyler (2000). Behrman, Kletzer, McPherson, and Schapiro (1998) find strong achievement effects on both continuation into college and quality of college; moreover, the effects are larger when proper account is taken of the various determinants of achievement. Hanushek and Pace (1995) find that college completion is significantly related to higher test scores at the end of high school.

¹⁶ This work has not, however, investigated how achievement affects the ultimate outcomes of additional schooling. For example, if over time lower-achieving students tend increasingly to attend further schooling, these schools may be forced to offer more remedial courses, and the variation of what students know and can do at the end of school may expand commensurately.

¹⁷ The OECD tested random samples of 15-year-old students across participating countries under the PISA program in 2000. Students taking these tests in Canada were then followed and surveyed in 2002 and 2004. See Section 4, below.

Table 2.1: Summary of Estimated Returns to a Standard Deviation Increase in Cognitive Skills

| Country | Study | Estimated effect ^a | Notes |
|--------------|--|--|---|
| Ghana | Glewwe (1996) | 0.21**-0.3** (government) 0.14-0.17 (private) | Alternative estimation approaches yield some differences; math effects shown generally more important than reading effects, and all hold even with Raven's test for ability. |
| Ghana | Jolliffe (1998) | 0.05-0.07* | Household income related to average math score with relatively small variation by estimation approach; effect is only observed with off-farm income, and on-farm income is not significantly related to cognitive skills. |
| Ghana | Vijverberg (1999) | ? | Income estimates for math and reading with nonfarm self-employment; highly variable estimates (including both positive and negative effects) but effects not generally statistically significant. |
| Kenya | Boissiere, Knight, and Sabot (1985); Knight and Sabot (1990) | 0.19**-0.22** | Total sample estimates: small variation by primary and secondary school leavers. |
| Morocco | Angrist and Lavy (1997) | ? | Cannot convert to standardized scores because use indexes of performance; French writing skills appear most important for earnings, but results depend on estimation approach. |
| Pakistan | Alderman, Behrman, Ross, and Sabot (1996) | 0.12-0.28* | Variation by alternative approaches and by controls for ability and health; larger and more significant without ability and health controls. |
| Pakistan | Behrman, Ross, and Sabot (forthcoming) | 0.25 | Estimates of structural model with combined scores for cognitive skill; significant effects of combined math and reading scores which are instrumented by school inputs |
| South Africa | Moll (1998) | 0.34**-0.48** | Depending on estimation method, varying impact of computation; comprehension (not shown) generally insignificant. |
| Tanzania | Boissiere, Knight, and Sabot (1985); Knight and Sabot (1990) | 0.07-0.13* | Total sample estimates: smaller for primary than secondary school leavers. |

*significant at 0.05 level; **significant at 0.01 level.

a. Estimates indicate proportional increase in wages from a one standard deviation increase in measured test scores.

The summary of the evidence in Table 2.1 permits a tentative conclusion that the returns to quality may be even larger in developing countries than in developed countries. This of course would be consistent with the range of estimates for returns to quantity of schooling (e.g., Psacharopoulos (1994) and Psacharopoulos and Patrinos (2004)), which are frequently interpreted as indicating diminishing marginal returns to schooling.

There are some reasons for caution in interpreting the precise magnitude of estimates. First, the estimates appear to be quite sensitive to the estimation methodology itself. Both within individual studies and across studies using the same basic data, the results are quite sensitive to the techniques employed in uncovering the fundamental parameter for cognitive skills.¹⁸ Second, the evidence on variations within developing countries is not entirely clear. For example, Jolliffe (1998) finds little impact of skills on farm income, while Behrman, Ross, and Sabot (forthcoming) suggest an equivalence across sectors at least on theoretical grounds.

Nonetheless, the overall summary is that the available estimates of the impact of cognitive skills on outcomes suggest strong economic returns within developing countries. The substantial magnitude of the typical estimates indicates that educational quality concerns are very real for developing countries and that this aspect of schools simply cannot be ignored – a topic that comes up below.

Evidence also suggests that educational quality is directly related to school attainment. In Brazil, a country plagued by high rates of grade repetition and ultimate school dropouts, Harbison and Hanushek (1992) show that higher cognitive skills in primary school lead to lower repetition rates. Further, Hanushek, Lavy, and Hitomi (2006) find that lower quality schools, measured by lower value-added to cognitive achievement, lead to higher dropout rates in Egyptian primary schools. Thus, as found for developed countries, the full economic impact of higher educational quality comes in part through greater school attainment.

This complementarity of school quality and attainment also means that actions that actually improve quality of schools will yield a bonus in terms of meeting goals for attainment. Conversely, simply attempting to expand access and attainment, say through starting a large number of low quality schools, will be self-defeating to the extent that there is a direct reaction to the low quality that affects the actual attainment results.

¹⁸ The sensitivity to estimation approach is not always the case; see, for example, Jolliffe (1998). A critique and interpretation of the alternative approaches within a number of these studies can be found in Glewwe (2002).

2.4 Evidence from the International Adult Literacy Survey

The preceding analyses for both developed and developing countries rely largely on panel data that follow individuals from school into the labor market. The alternative approach as found in the International Adult Literacy Survey (IALS) is to test a sample of adults and then to relate these measures to labor market experiences.¹⁹ An advantage of this data collection approach is that it provides information about the labor market experiences across a broader range of age and labor market experience.²⁰

Consistent data on basic skills of literacy and numeracy for a representative sample of the population aged 15-65 were collected for a sample of countries between 1994 and 1998.²¹ These data permit direct comparisons of the relative importance of quantity and quality of schooling across countries, although the bias toward developed economies remains. Hanushek and Zhang (2006) estimate returns to school attainment and to literacy scores for the 13 countries where continuous measures of individual earnings are available.²²

Figures 2.2 and 2.3 provide the relevant summary information on the returns to skills, estimated in a model that jointly includes school attainment and literacy scores. As in the prior analyses, both school attainment and cognitive skills enter into the determination of individual incomes. With the exception of Poland, literacy scores have a consistent positive impact on earnings (Figure 2.2). The (unweighted) average of the impact of literacy scores is 0.093, only slightly less than found previously for the U.S.

¹⁹ This design was subsequently repeated in 2003 with the Adult Literacy and Lifeskills Survey (ALL), but only six countries participated, and the data were unavailable for this study.

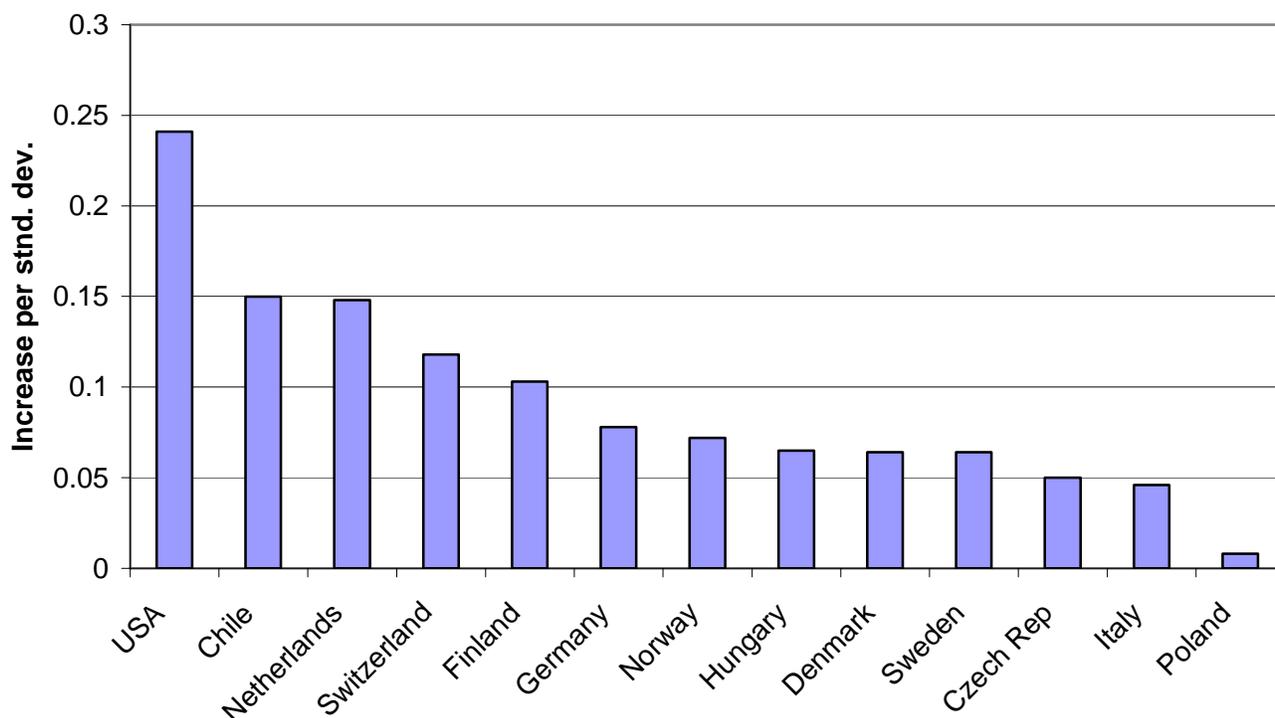
²⁰ This approach does also present some complications, because the individuals of different ages have both different adult learning experiences and different times of attending school of possibly different quality. Hanushek and Zhang (2006) consider these alternatives, but they do not change the qualitative results about the impact of cognitive skills that are presented here.

²¹ Twenty-three countries and regions participated in one of three different waves of surveys conducted in 1994, 1996 and 1998. Country participation included some sub-country divisions (e.g., Flemish speaking Belgium and Italian speaking Switzerland). Not all countries provided background information and wage information. IALS provides information on the skills and attitudes of adults aged 16-65 in a number of different areas, and the analysis here combines the scores in a single measure of literacy.

²² Samples include full-time workers between 26 and 65 years of age. In all the models, the dependent variable is the logarithm of annual earnings from employment; control variables are gender, potential experience and its square, and an indicator for living in rural area.

studies.²³ On the other hand, the estimated impact of school attainment across the 13 countries is just 0.049 after adjusting the Mincer returns for the literacy scores. This low estimate partly reflects the joint consideration of literacy scores and school attainment. The estimated return to years of schooling without considering literacy scores is 0.06, which is still below the more common estimates of about 0.1. As Figure 2.3 shows, the estimates that adjust for literacy scores are below the unadjusted returns in every country except Poland.

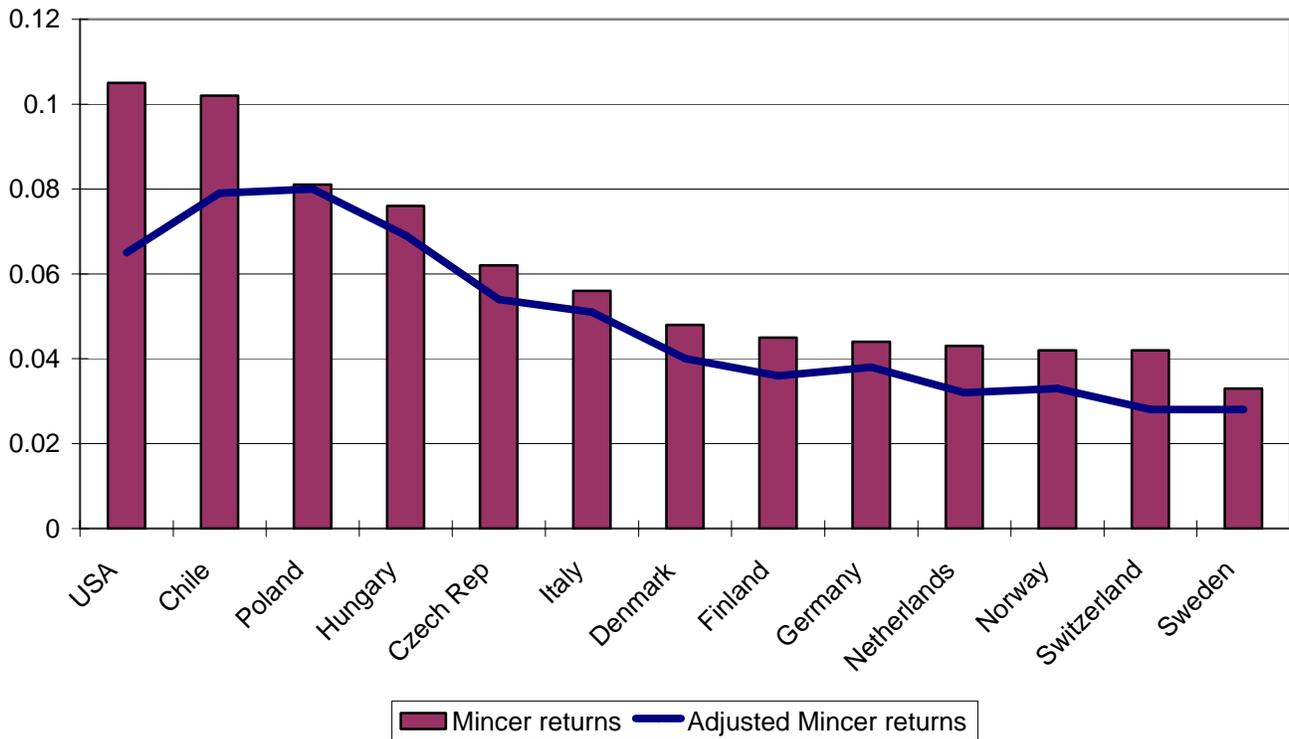
Figure 2.2: Returns to Cognitive Skills, International Adult Literacy Survey



Source: Hanushek and Zhang (2006).

²³ The U.S. is noticeably higher than other countries and the previous U.S. studies, perhaps reflecting that these earnings are obtained across the entire work life. The average excluding the U.S. is still 0.08.

Figure 2.3: Returns to School Attainment, International Adult Literacy Survey



Source: Hanushek and Zhang (2006).

The literacy tests in IALS are designed to measure quite basic skills only, and yet the differences are strongly associated with higher earnings. These results, from a broad age spectrum across a number of countries, reinforce the importance of quality.

The sample of countries for the IALS unfortunately contains just one developing country – Chile. Nonetheless, it is suggestive that the returns both to quantity of schooling and quality exceed those found across the countries with the exception of the United States. In the three transition economies (Czech Republic, Hungary, and Poland), the returns to school attainment are also near the top of the sample, but the returns to educational quality are noticeably lower – perhaps reflecting institutional aspects of their labor markets.

2.5 Causality

The focus on causality of estimated returns to school attainment has not been duplicated for the estimated impacts of cognitive skills. The potential problems are, however, considerably different. The tests, with the exception of the IALS sampling, are uniformly given at a date before the labor market

experiences – eliminating the possibility that higher income leads individuals to do things that raise their test scores.²⁴ Nevertheless, this fact alone does not solve all of the interpretative issues.

For our analysis, the most important issue is the source of any test score differences. It is natural to believe that schools have an influence on tests, but clearly other factors also enter. The extensive investigations of the determinants of achievement differences indicate that parents, peers, neighborhoods, and other factors enter along with school factors in determining achievement (see Hanushek (2002b)). Thus, most importantly, it is inappropriate to interpret test scores as simply reflecting school quality or school policy.

We return to a discussion of possible school policies below, but at this point it is sufficient to note that skills – however formed – have a systematic impact on earnings. Thus, if we can find approaches that increase skills reliably, the available evidence strongly indicates that individual earnings and productivity will also increase.²⁵

The other side of this issue is also important, however. Using just quantity of schooling in the earnings analyses assumes that formal schooling is the only source of skill development. But, if a variety of other inputs such as families or peers is also important in the formation of human capital, simple years of schooling is subject to this additional source of systematic measurement error.

2.6 Income Distribution

One implication of the impact of cognitive skills on individual earnings is that the distribution of those skills in the economy will have a direct effect on the distribution of income. Cognitive skills by themselves do not of course determine the full distribution, because other factors such as labor market institutions, taxes, and the like enter. But the importance of skills is becoming increasingly evident.

Very suggestive evidence on the impact of skills on the income distribution comes from Nickell (2004). Nickell, using the IALS data, considers how differences in the distribution of incomes across countries are affected by the distribution of skills and by institutional factors including unionization and

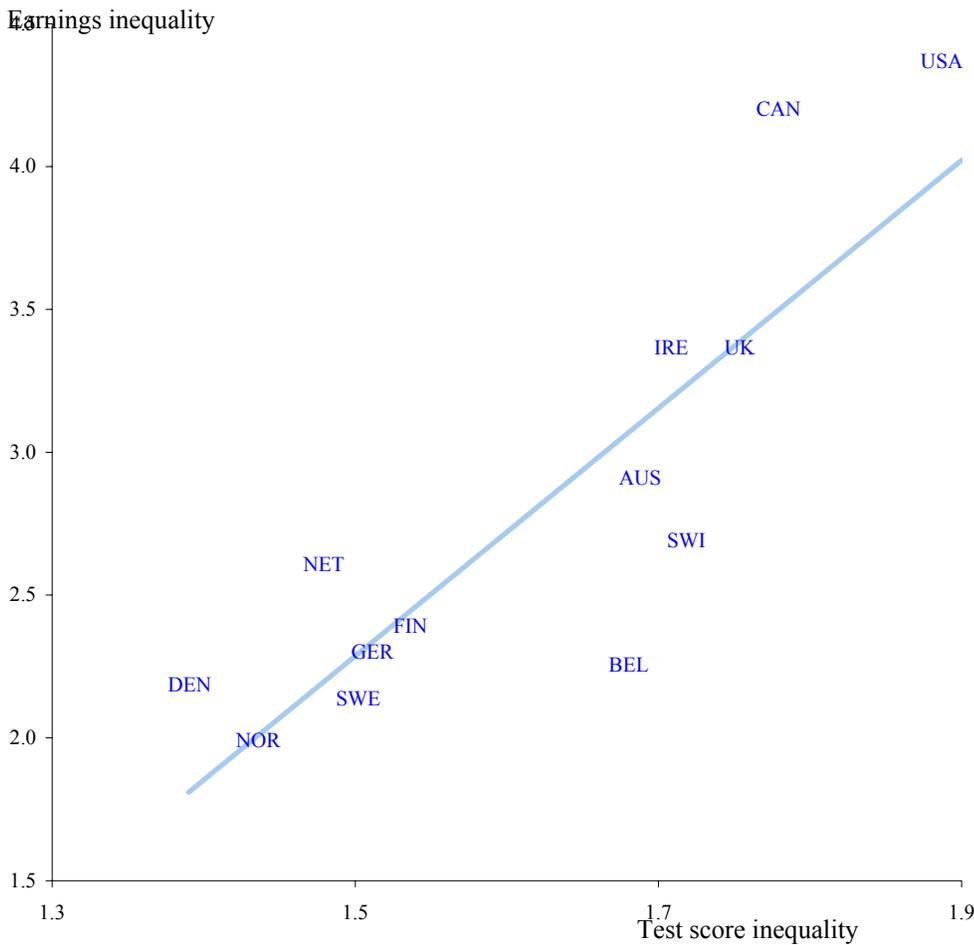
²⁴ The IALS sampling raises the concern that tests change with age, because of continual learning or simple age depreciation of skills and knowledge. An investigation of this by Hanushek and Zhang (2006) suggests that these are not large concerns in the estimation of the earnings functions.

²⁵ There could be other things that simultaneously affect both scores and earnings, such that scores are simply a proxy for some other important factor. Available evidence gives no reason to suspect that such factors are important.

minimum wages. While union coverage is statistically significant, he concludes that “the bulk of the variation in earnings dispersion is generated by skill dispersion” (page C11).²⁶

The impact of the skill distribution across countries is shown dramatically in Figure 2.4, which is derived from a comparison of the dispersion of wages and the dispersion of prose literacy scores (each measured as the ratio of the 90th to the 10th percentile). The tight pattern around the regression line reflects a simple correlation of 0.85 (which is not affected by including the other institutional factors).

Figure 2.4: Inequality of Educational Quality and Earnings



Note: Measure of inequality is the ratio of ninth decile to first decile in both cases; test performance refers to prose literacy in the International Adult Literacy Survey.

Source: Nickell (2004).

²⁶ De Gregorio and Lee (2002) find a (somewhat weaker) positive association between inequality in years of schooling and income inequality.

Other studies have also concluded that skills have an increasing impact on the distribution of income (e.g., Juhn, Murphy, and Pierce (1993)). In the U.S., the distribution of incomes within schooling groups has been rising (Levy and Murnane (1992)), i.e., holding constant schooling attainment, the income distribution has become more dispersed in reflection of growing rewards to individual skills.

Again, these studies do not attempt to describe the causal structure, and it would be inappropriate to attribute the variance in earnings simply to differences in the quantity or quality of schooling. Nonetheless, to the extent that these contribute to variations in cognitive skills, it is fair to conclude that policies aimed at improving school quality (and educational outcomes) will have direct impacts on the income distribution.

3. Quantity of Schooling and Economic Growth

Given the microeconomic evidence of the productivity-enhancing effects of education, it seems natural to extend the view to the macroeconomic perspective of long-run economic growth of countries. From a theoretical viewpoint, there are at least three mechanisms through which education may affect economic growth. First, just as in the micro perspective, education increases the human capital inherent in the labor force, which increases labor productivity and thus transitional growth towards a higher equilibrium level of output (as in augmented neoclassical growth theories, cf. Mankiw, Romer, and Weil (1992)). Second, education may increase the innovative capacity of the economy, and the new knowledge on new technologies, products and processes promotes growth (as in theories of endogenous growth, cf., e.g., Lucas (1988); Romer (1990a); Aghion and Howitt (1998)). Third, education may facilitate the diffusion and transmission of knowledge needed to understand and process new information and to successfully implement new technologies devised by others, which again promotes economic growth (cf., e.g., Nelson and Phelps (1966); Benhabib and Spiegel (2005)).

3.1 Results of Initial Cross-country Growth Regressions

Just as in the literature on microeconomic returns to education, the majority of the macroeconomic literature on economic growth that tries to test these predictions employs the quantitative measure of years of schooling, now averaged across the labor force.²⁷ Early studies used adult literacy rates (e.g., Azariadis and Drazen (1990); Romer (1990b)) or school enrollment ratios (e.g., Barro (1991); Mankiw, Romer, and Weil (1992); Levine and Renelt (1992)) as proxies for the human capital of an economy. These were followed by attempts to measure average years of schooling based on perpetual inventory methods (cf. Lau, Jamison, and Louat (1991); Nehru, Swanson, and Dubey (1995)).²⁸ Then, it was particularly the work by Barro and Lee (1993, 2001) which provided internationally comparable data on average years of schooling for a large sample of countries and years, based on a combination of census or survey data on educational attainment wherever possible, using literacy and enrollment data to fill gaps in the census data.

²⁷ More precisely, the most commonly used measure is average years of schooling in the working-age population, usually defined as the population aged 15 years and over, instead of the actual labor force.

²⁸ For a survey of measurement and specification issues from early growth accounting to current cross-country growth regressions, see Wößmann (2003b).

Using average years of schooling as the education measure implicitly assumes that a year of schooling delivers the same increase in knowledge and skills regardless of the education system. For example, a year of schooling in Papua New Guinea is assumed to create the same increase in productive human capital as a year of schooling in Japan. Additionally, this measure assumes that formal schooling is the primary (sole) source of education and, again, that variations in the quality of nonschool factors have a negligible effect on education outcomes. This neglect of cross-country differences in the quality of education is probably the major drawback of such a quantitative measure of schooling, and we will come back to this issue in great detail below.

The standard method to estimate the effect of education on economic growth is to estimate cross-country growth regressions where countries' average annual growth in gross domestic product (GDP) per capita over several decades is expressed as a function of measures of schooling and a set of other variables deemed to be important for economic growth. Following the classical contributions by Barro (1991, 1997) and Mankiw, Romer, and Weil (1992),²⁹ a vast early literature of cross-country growth regressions has tended to find a significant positive association between quantitative measures of schooling and economic growth (for extensive reviews of the literature, see, e.g., Topel (1999); Temple (2001); Krueger and Lindahl (2001); Sianesi and Van Reenen (2003)).³⁰ To give an idea of the robustness of this association, in the recent extensive robustness analysis by Sala-i-Martin, Doppelhofer, and Miller (2004) of 67 explanatory variables in growth regressions on a sample of 88 countries, primary schooling turns out to be the most robust influence factor (after an East Asian dummy) on growth in GDP per capita in 1960-1996.

Because this literature has been extensively reviewed elsewhere (see references above), rather than going into the details of the numerous important contributions, we will focus on a few important issues that emerge in the literature.

²⁹ Temple and Wößmann (2006) show that the significantly positive effect of education that Mankiw, Romer, and Weil (1992) find does not depend on their often criticized use of an education flow measure based on enrollment rates, but can be replicated when using years of schooling as a measure of the level of human capital in their model.

³⁰ A closely related literature weights years of schooling by parameters from microeconomic Mincer earnings equations (see Section 2.1 above) to construct a measure of the stock of human capital, which is then used in growth accounting and development accounting exercises. These use a given macroeconomic production function together with parameter estimates from other research to calculate the share of cross-country differences in growth or levels of income which can be accounted for by cross-country differences in education (see Klenow and Rodriguez-Clare (1997) and Hall and Jones (1999) for examples).

3.2 More Recent Evidence on the Effects of Levels of and Growth in Years of Schooling

To provide a basic representation of the association between years of schooling and economic growth on the most recent version of available data, we use a slightly extended version³¹ of the education data by Cohen and Soto (2001), representing the average years of schooling of the population aged 15 to 64, together with data on real GDP per capita in 1960-2000 from the latest update (version 6.1) of the Penn World Tables by Heston, Summers, and Aten (2002).³² Figure 3.1 plots the average annual rate of growth in GDP per capita over the 40-year period against years of schooling at the beginning of the period for a sample of 92 countries. Both growth and education are expressed conditional on the initial level of output, to account for the significant conditional convergence effect.³³

The regression results depicted by Figure 3.1 imply that each year of schooling is statistically significantly associated with a long-run growth rate that is 0.58 percentage points higher.³⁴ The positive association is substantially larger in the sample of non-OECD countries (at 0.56) than in the sample of OECD countries (at 0.26), which is in line with the pattern of larger returns to education in developing countries discussed above.³⁵ However, after controlling for the influence of openness and the security of property rights, the association becomes substantially smaller and turns insignificant, and it is close to zero when the total fertility rate is controlled for. Thus, while there is a clear positive association between years of schooling and growth in the latest available data, it is also somewhat sensitive to model specifications.

³¹ See Jamison, Jamison, and Hanushek (forthcoming) for details of the extension.

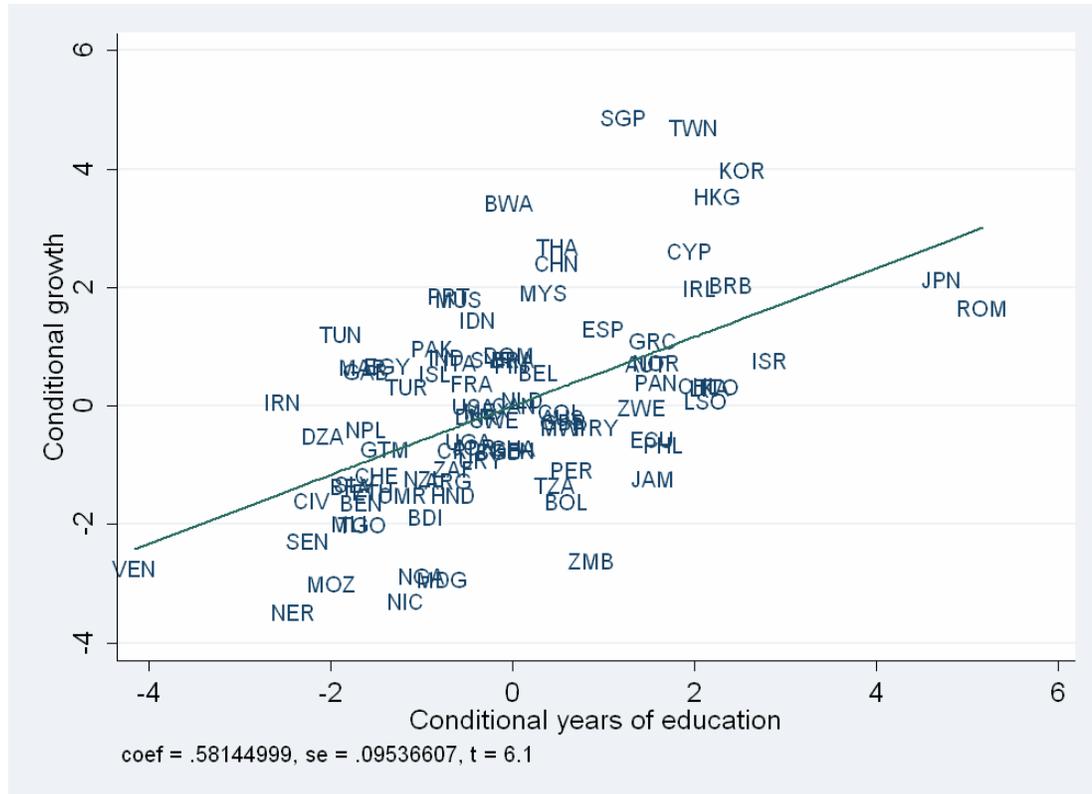
³² As discussed below, one line of investigation has been the impact of mismeasurement of the quantity of education on growth. The Cohen and Soto (2001) data improve upon the original quantity data by Barro and Lee.

³³ Added-variable plots show the association between two variables after the influences of other control variables are taken out. Thus, both of the two variables are first regressed on the other controls (in this case, initial GDP). Only the residual of these regressions, which is the part of the variation in the two variables which cannot be accounted for by the controls, is used in the graph. In so doing, the graph makes sure that the depicted association between the two variables is not driven by the control variables. The procedure is numerically equivalent to including the other controls in a multivariate regression of the dependent variable (growth) on the independent variable under consideration in the graph.

³⁴ The association is somewhat lower (at 0.32), but still significant when regional dummies are added to the regression.

³⁵ Similarly, the association is larger in the sample of countries below the median of initial output than above the median.

Figure 3.1: Added-variable Plot of Growth and Years of Schooling without Test-score Controls



Notes: Added-variable plot of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960-2000 on average years of schooling in 1960 and the initial level of real GDP per capita in 1960. Own calculations.

A considerable controversy has emerged in the literature about whether it is the *level* of years of schooling (as would be predicted by several models of endogenous growth) or the *change* in years of schooling (as would be predicted in basic neoclassical frameworks) which is the more important driver of economic growth. The early evidence, such as in Benhabib and Spiegel (1994) and discussed in Barro and Sala-i-Martin (2004), found a positive effect of educational levels, but not of changes in education. However, Temple (1999) shows that in the latter case, an existing positive association was hidden by a few unrepresentative outlying observations. Considerable evidence has also emerged that there was substantial measurement error in the education data (cf. Krueger and Lindahl (2001)), and it is well known that measurement error affects results based on changes in variables even more than results based on their levels. Subsequent evidence suggests that both levels of and changes in years of schooling may show a positive association with growth (cf. Gemmell (1996); Topel (1999); Krueger and Lindahl (2001)).

More recently, two studies have tried to overcome some problems of mismeasurement in the early Barro-Lee data on years of schooling: de la Fuente and Doménech (2001, 2006) for the sample of OECD countries and Cohen and Soto (2001) for a broader sample of countries (and the data that we use

here). Both studies find robust evidence of a positive association between changes in education and economic growth. Even more recently, Ciccone and Papaioannou (2005) find strong support for both the human capital level and the human capital accumulation effects using cross-country industry-level data that allow them to control for country-specific and industry-specific effects.

When we add the change in years of schooling over 1960-2000 to the specification depicted in Figure 3.1 and similar specifications, it never turns significant with the sole exception of the sample of 23 OECD countries.³⁶ (Because of the possibly substantial amount of mismeasurement in the education data, though, the results on the first-differenced data on changes in education may well suffer).

Thus, while recent research tends to find a positive effect of schooling quantity on economic growth, it seems beyond the scope of current data to draw strong conclusions about the relative importance of different mechanisms by which schooling quantity may affect economic growth. However, several recent studies suggest that education is important both as an investment in human capital and in facilitating research and development and the diffusion of technologies. With respect to the relative importance of the latter two mechanisms, Vandebussche, Aghion, and Meghir (2006) suggest that innovation is more important relative to imitation for countries close to the technological frontier. As a consequence, the composition of human capital between basic and higher education may be important, with initial phases of education being more important for imitation and higher education being more important for innovation. Vandebussche, Aghion, and Meghir (2006) provide evidence from a panel of OECD countries in line with this argument, which – when applied to developing countries – might suggest that a focus on basic skills seems warranted for developing countries. We will come back to the issue of the relative importance of basic versus advanced skills in more detail in Section 4.4 below.

Two more skeptical studies raise noteworthy caveats, and we will return to each of them below. First, Bils and Klenow (2000) raise the issue of causality, suggesting that reverse causation running from higher economic growth to additional education may be at least as important as the causal effect of education on growth in the cross-country association. We will address the issue of causality in cross-country regressions in Section 4.2 below. Second, one of the conclusions that Pritchett (2001, 2006) draws from the fragility of the evidence linking changes in education to economic growth is that it is important for economic growth to get other things right as well, in particular the institutional framework of the economy. We will address this issue in Section 4.4 below.

³⁶ The positive association between growth in education and economic growth in the OECD sample is sensitive to the inclusion of Korea, though.

4. Quality of Education and Economic Growth

4.1 A Review of the Basic Results

The most important caveat with the literature on education and growth reviewed in the preceding section, though, is that it sticks to years of schooling as its measure of education at the neglect of qualitative differences in ensuing knowledge. As discussed in the section on individual returns to education, this neglect probably misses the core of what education is all about. And this neglect seems even more severe in cross-country comparisons than in analyses within countries: Who would sensibly assume that the average student in a school in Ghana or Peru would gain the same amount of knowledge in any year of schooling as the average student in a school in Finland or Korea? Still, using the quantitative measure of years of schooling does exactly that. It seems beyond doubt that one year of schooling does not create the same amount of acquired knowledge regardless of the quality of the education system in which it takes place, but delivers different increases in skills depending on the efficiency of the education system, the quality of teaching, the educational infrastructure, or the curriculum. Thus, rather than counting how long students have sat in school, it seems crucial to focus on how much students have learned while in school when estimating the effect of education on economic growth.

Years of schooling has a second major shortcoming. Its use implicitly assumes that all skills and human capital formation come from formal schooling. Yet extensive evidence on knowledge development and cognitive skills indicates that a variety of factors outside of school – family, peers, and others – have a direct and powerful influence. Ignoring these nonschool factors introduces another element of measurement error into the growth analyses in the same way as it did in the analysis of individual earnings.

From the mid-1960s to today, international agencies have conducted many international tests of students' performance in cognitive skills such as mathematics and science (see the appendix to this section for details). Employing a re-scaling method that makes performance at different international tests comparable (again, see appendix), we can use performance on these standardized tests as a proxy for the quality of education. Figure 4.1 presents average student performance on twelve testing occasions³⁷ on the transformed scale which maps performance on each test to the scale of the recent

³⁷ Note that each testing occasion may combine results from several tests at different age levels and in different subjects, so that there are 36 separate test observations altogether; see appendix for details.

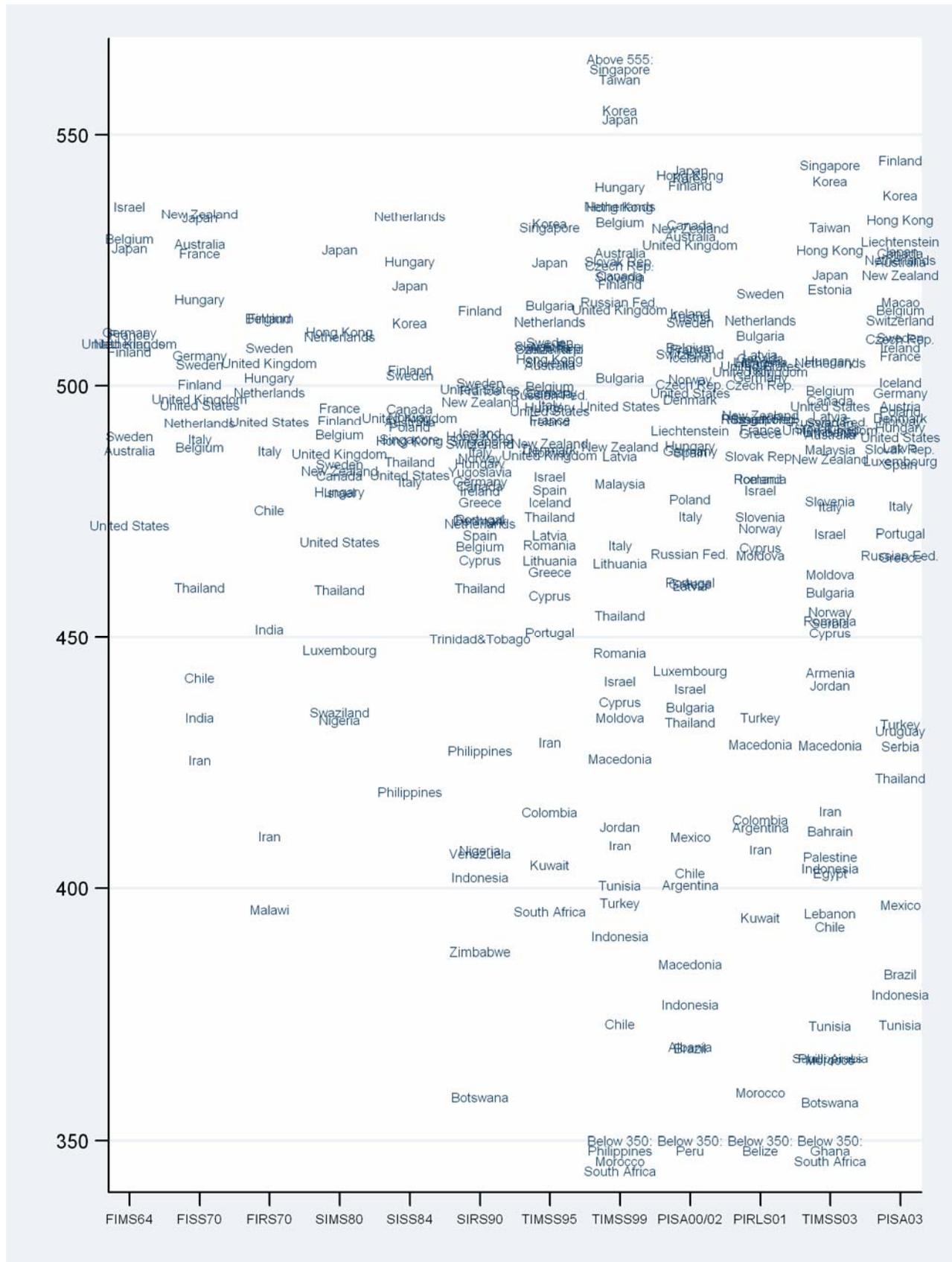
PISA international test. This scale has a mean of 500 and a standard deviation of 100 among the OECD countries in PISA.³⁸ As is obvious from the figure, the developing countries that ever participated in one of the tests perform dramatically lower than any country in the group of OECD countries. The variation in the quality of education that exists among OECD countries is already substantial, but the magnitude of the difference to developing countries in the average amount of learning that has taken place after a given number of years of schooling dwarfs any within-OECD difference.³⁹

Over the past ten years, empirical growth research demonstrates that consideration of the quality of education, measured by the cognitive skills learned, alters the assessment of the role of education in the process of economic development dramatically. When using the data from the international student achievement tests through 1991 to build a measure of educational quality, Hanushek and Kimko (2000) – first released as Hanushek and Kim (1995) – find a statistically and economically significant positive effect of the quality of education on economic growth in 1960-1990 that dwarfs the association between quantity of education and growth. Thus, even more than in the case of education and individual earnings, ignoring quality differences very significantly misses the true importance of education for economic growth. Their estimates suggest that one country-level standard deviation (equivalent to 47 test-score points in PISA 2000 mathematics, on whose scale Figure 4.1 is based) higher test performance would yield around one percentage point higher annual growth rates.

³⁸ To facilitate comparisons, the within-country standard deviation of PISA test scores ranges between 80 and 108 in mathematics in the OECD countries; the U.S. value is 98.

³⁹ The precise scaling on the transformed metric is of course subject to considerable noise, in particular for the early tests and for countries performing far below the international mean. The tests are usually not developed to provide reliable estimates of performance in the tails of the achievement distribution, which would be relevant for very poorly performing countries. However, the rough pattern of overall performance should not be severely affected by the re-scaling. For example, the average performance level of Peruvian students on PISA 2002 where no re-scaling is involved is 292 in math, 333 in science and 327 in reading.

Figure 4.1: Adjusted Performance on International Student Achievement Tests



Source: Hanushek and Wößmann (in process), based on the different tests; see appendix to this section for details.

This estimate stems from a statistical model that relates annual growth rates of real GDP per capita to the measure of educational quality, years of schooling, the initial level of income, and several other control variables (including in different specifications the population growth rates, political measures, openness of the economies, and the like). Hanushek and Kimko (2000) find that adding educational quality to a base specification including only initial income and educational quantity boosts the variance in GDP per capita among the 31 countries in their sample that can be explained by the model from 33 to 73 percent. The effect of years of schooling is greatly reduced by including quality, leaving it mostly insignificant. At the same time, adding the other factors leaves the effects of quality basically unchanged.

Several studies have since found very similar results. Another early contribution, by Lee and Lee (1995), found an effect size similar to Hanushek and Kimko (2000) using data from the 1970-71 First International Science Study on the participating 17 countries, also leaving quantitative measures of education with no significant effect on growth. Using a more encompassing set of international tests, Barro (2001) also finds that, while both the quantity and the quality of education matter for economic growth, quality is much more important. Employing the quality measure developed by Hanushek and Kimko (2000) in a development accounting framework, Wößmann (2002, 2003b) finds that the share of cross-country variation in levels of economic development attributable to international differences in human capital rises dramatically when quality of education is taken into account.⁴⁰

Extensions of the measure of Hanushek and Kimko (2000) and its imputation in Wößmann (2003b) are also used in the cross-country growth regressions by Bosworth and Collins (2003) and in the cross-country industry-level analysis by Ciccone and Papaioannou (2005). Both also find that educational quality strongly dominates any effect of educational quantity on growth.⁴¹ Coulombe, Tremblay, and Marchand (2004) and Coulombe and Tremblay (2006) use test-score data from the International Adult Literacy Survey (see Section 2.4 above) in a panel of 14 OECD countries, confirming the result that the quality measure outperforms quantitative measures of education. Finally, Jamison, Jamison, and Hanushek (forthcoming) extend the Hanushek and Kimko (2000) analysis by

⁴⁰ This work builds on Gundlach, Rudman, and Wößmann (2002) and analyzes output per worker in 132 countries in 1990. The variation that can be attributed to international differences in human capital rises from 21 percent to 45 percent once the quality of education is taken into account, and to over 60 percent in samples with reasonable data quality.

⁴¹ Bosworth and Collins (2003) cannot distinguish the effect of quality of education from the effect of quality of government institutions. The analysis in Section 4.5 below shows, however, that they can be separated when we use our new quality measure that also extends the country sample by several additional data points on international tests scores.

using the mathematics component of the transformed and extended tests shown in Figure 4.1, replicating and strengthening the previous results by using test data from a larger number of countries, controlling for a larger number of potentially confounding variables and extending the time period of the analysis. Using the panel structure of their growth data, they suggest that education seems to improve income levels mainly through speeding up technological progress, rather than shifting the level of the production function or increasing the impact of an additional year of schooling.

In sum, the existing evidence suggests that the quality of education, measured by the knowledge that students gain as depicted in tests of cognitive skills, is substantially more important for economic growth than the mere quantity of education.

4.2 Issues of Endogeneity

As noted, growth modeling is naturally subject to a common concern: Do the identified factors represent truly causal influences or mere associations that will not affect growth if altered by policy?

The concerns about causality in the relationship of cognitive skills and growth have been addressed in detail by Hanushek and Kimko (2000). They conclude that causation concerns are very different in the case of quality than quantity, being much less of an issue in interpreting the results. Because causality issues are important for assessing our results below, we describe the Hanushek and Kimko direct investigations.

One common concern in analyses such as this is that schooling might not be the actual cause of growth but, in fact, may just reflect other attributes of the economy that are beneficial to growth. For example, the East Asian countries consistently score very highly on the international tests (see Figure 4.1), and they also had extraordinarily high growth over the 1960–1990 period. It may be that other aspects of these East Asian economies have driven their growth and that the statistical analysis of labor force quality simply is picking out these countries. But in fact, even if the East Asian countries are excluded from the analysis, a strong—albeit slightly smaller—relationship is still observed with test performance. This test of sensitivity of the results seems to reflect a basic importance of school quality, a factor that contributes also to the observed growth of East Asian countries.

Another concern is that other factors that affect growth, such as efficient market organizations, are also associated with efficient and productive schools—so that, again, the test measures are really a proxy for other attributes of the country. To investigate this, Hanushek and Kimko concentrate on immigrants to the United States who received their education in their home countries. They find that immigrants who were schooled in countries that have higher scores on the international math and science examinations earn more in the United States. On the other hand immigrants receiving part or all

of their schooling in the United States do not see any earnings advantage linked to the cognitive skills of their home country. This analysis makes allowance for any differences in school attainment, labor market experience, or being native English-language speakers. In other words, skill differences as measured by the international tests are clearly rewarded in the United States labor market, reinforcing the validity of the tests as a measure of individual skills and productivity.

Finally, the observed relationships could simply reflect reverse causality, that is, that countries that are growing rapidly have the added resources necessary to improve their schools and that better student performance is the result of growth, not the cause of growth. This issue is potentially important when policy alternatives are considered below, but, as a simple test of this, Hanushek and Kimko investigated whether the international math and science test scores were systematically related to the resources devoted to the schools in the years prior to the tests. They were not. If anything, their results suggested relatively better performance in those countries spending less on their schools.

One final issue warrants consideration: The United States has never done well on these international assessments, yet its growth rate has been very high for a long period of time. The reconciliation is that quality of the labor force is just one aspect of the economy that enters into the determination of growth. A variety of factors clearly contribute, and these factors work to overcome any deficits in quality. These other factors may also be necessary for growth. In other words, simply providing more or higher-quality schooling may yield little in the way of economic growth in the absence of other elements, such as the appropriate market, legal, and governmental institutions to support a functioning modern economy. Past experiences investing in less developed countries that lack these institutional features suggest that schooling is not necessarily itself a sufficient engine of growth (e.g., Easterly (2002); Pritchett (2001)).

We take up these issues about quality of societal institutions in detail below, because the evidence suggests that they are very important and could potentially distort the analysis of the impacts of human capital. Here, however, we briefly consider the U.S. situation, both because it sets the stage for more recent analyses and because it helps to provide some balance to the overall picture of growth.

Three other factors immediately come to mind as being important in U.S. growth and as potentially masking to detrimental effects of low school quality. First, almost certainly the most important factor sustaining the growth of the U.S. economy is the openness and fluidity of its markets. The United States maintains generally freer labor and product markets than most countries in the world. The government generally has less regulation on firms, and trade unions are less extensive than those in many other countries. Even broader, the United States has less intrusion of government in the operation of the economy, including lower tax rates and minimal government production through nationalized

industries. These factors encourage investment, permit the rapid development of new products and activities by firms, and allow U.S. workers to adjust to new opportunities. While identifying the precise importance of these factors is difficult, a variety of analyses suggest that such market differences could be very important explanations for differences in growth rates.⁴²

Second, over the twentieth century, the expansion of the education system in the United States outpaced that around the world. The United States pushed to open secondary schools to all citizens. With this came also a move to expand higher education with the development of land grant universities, the G.I. bill, and direct grants and loans to students. More schooling with less learning each year still yielded more human capital than found in other nations that have less schooling but more learning in each of those years.⁴³

Third, the analysis of growth rates across countries emphasizes quality of the elementary and secondary schools of the United States. It does not include any measures of the quality of U.S. colleges. By most evaluations, U.S. colleges and universities rank at the very top in the world. A number of models of economic growth in fact emphasize the importance of scientists and engineers as a key ingredient to growth. By these views, the technically trained college students who contribute to invention and to development of new products provide a special element to the growth equation. Here, again, the United States appears to have the best programs.

Finally, in terms of endogeneity, it is important to note that our current work (reported below) replicates two of the tests in Hanushek and Kimko (2000). Specifically, in the context of extended growth analyses, we consider potential biases from including the East Asian countries, and we consider the resources-skills relationship.

4.3 Some New Evidence

To provide the most up-to-date evidence, we extend the existing evidence in several ways. The new evidence adds additional international student achievement tests not previously available and uses the most recent data on economic growth which allow an analysis for an even longer time period (1960-2000). Furthermore, the new data extend the sample of countries with available test-score and growth information from 31 countries in Hanushek and Kimko (2000) to 50 countries (see the appendix of this section for details on the data). In the subsequent sections, we will also use these data to analyze

⁴² See, for example, Krueger (1974); World Bank (1993); Parente and Prescott (1994, 1999).

⁴³ This advantage has clearly ended as many OECD countries have expanded schools to exceed the quantity of schooling found in the United States (see Organisation for Economic Co-operation and Development (2003)).

effects of the distribution of educational quality at the bottom and at the top on economic growth, as well as interactions between educational quality and the institutional infrastructure of an economy.

Our measure of the quality of education is a simple average of the mathematics and science scores over all the international tests depicted in Figure 4.1. We interpret this as a proxy for the average educational performance of the whole labor force. This measure encompasses overall cognitive skills, not just those developed in schools. Thus, whether skills are developed at home, in schools, or elsewhere, they are included in the growth analyses. As in the analyses underlying Figure 3.1, the source of the income data is version 6.1 of the Penn World Tables (cf. Heston, Summers, and Aten (2002)), and the data on years of schooling is an extended version of the Cohen and Soto (2001) data.

The basic result is reported in column (2) of Table 4.1 and depicted graphically in Figure 4.2.⁴⁴ After controlling for the initial level of GDP per capita and for years of schooling, the test-score measure features a statistically significant effect on the growth in real GDP per capita in 1960-2000. According to this specification, test scores that are larger by one standard deviation (measured at the student level across all OECD countries in PISA) are associated with an average annual growth rate in GDP per capita that is two percentage points higher over the whole 40-year period. Below we relate these numbers to outcomes of school reform policies.

When educational quality is added to a model that just includes initial income and years of schooling (column (1) of Table 4.1), the share of variation in economic growth explained by the model (the adjusted R^2) jumps from 0.25 to 0.73. As in Figure 3.1, quantity of education is statistically significantly related to economic growth in a specification that neglects educational quality, but the association between years of schooling and growth turns insignificant and is reduced to close to zero once the quality of education is included in the model (see the bottom of Figure 4.2).⁴⁵ In either specification, there is evidence for conditional convergence in the sense that countries with higher initial income tend to grow more slowly over the subsequent period.

⁴⁴ For a brief technical description of added-variable plots see Section 3.2 above.

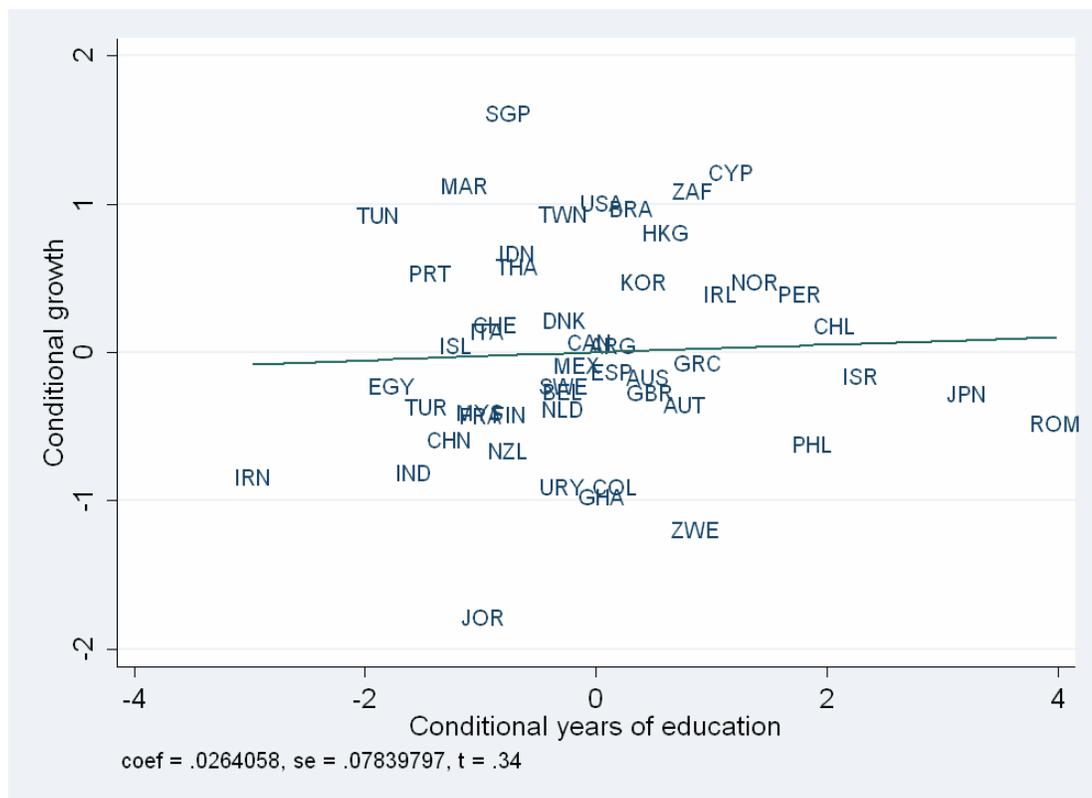
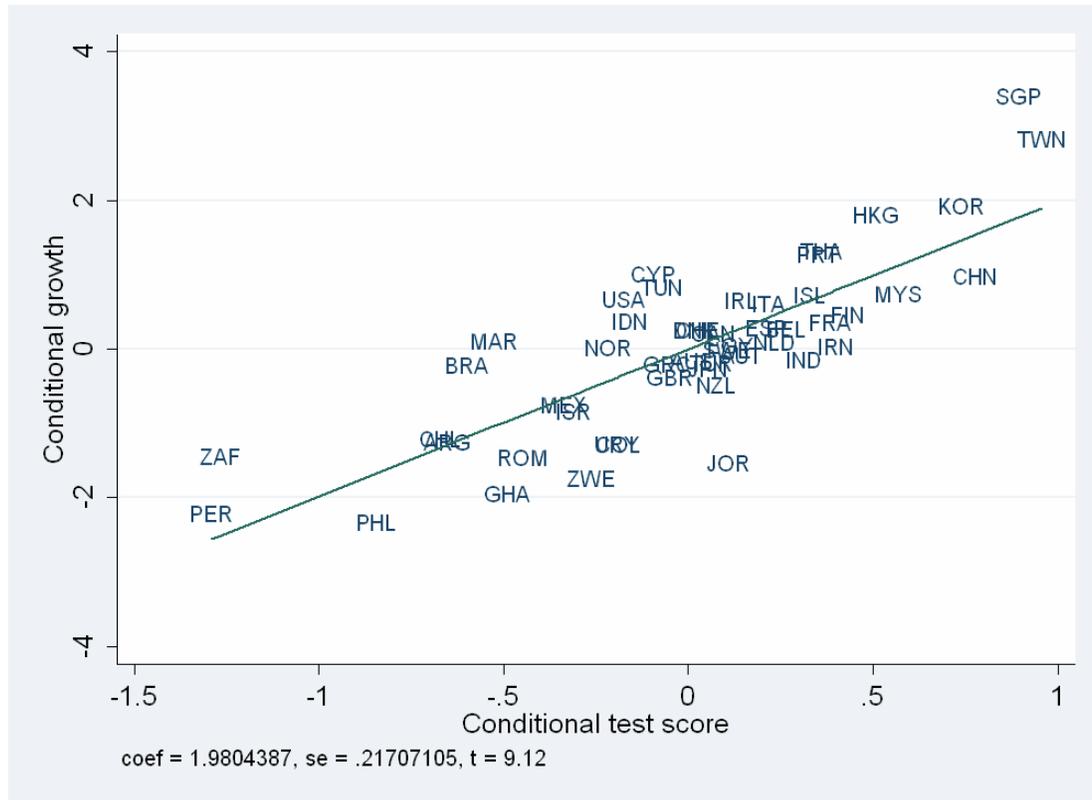
⁴⁵ The result remains the same when the measure of years of schooling refers to the average between 1960 and 2000, rather than the initial 1960 value. The coefficient estimate on years of schooling in the 50-country sample of column (1) of Table 4.1 is somewhat smaller than the one reported for the 92-country sample reported in Figure 3.1, corresponding to the fact discussed above that the estimate tends to be smaller in high-income countries.

Table 4.1: Education as Determinant of Growth of Income per Capita, 1960-2000

| | (1) | (2) | (3) ^a | (4) |
|----------------------------------|------------------|------------------|------------------|------------------|
| GDP per capita 1960 | -0.379 (4.24) | -0.302 (5.54) | -0.277 (4.43) | -0.351 (6.01) |
| Years of schooling 1960 | 0.369 (3.23) | 0.026 (0.34) | 0.052 (0.64) | 0.004 (0.05) |
| Test score (mean) | | 1.980 (9.12) | 1.548 (4.96) | 1.265 (4.06) |
| Openness | | | | 0.508 (1.39) |
| Protection against expropriation | | | | 0.388 (2.29) |
| Constant | 2.785 (7.41) | -4.737 (5.54) | -3.701 (3.32) | -4.695 (5.09) |
| N | 50 | 50 | 50 | 47 |
| R ² (adj.) | 0.252 | 0.728 | 0.741 | 0.784 |

Dependent variable: average annual growth rate in GDP per capita, 1960-2000. *t*-statistics in parentheses. ^a Regression includes five regional dummies.

Figure 4.2: Added-variable Plots of Growth and Education



Notes: Added-variable plots of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960-2000 on the initial level of real GDP per capita in 1960, average test scores on international student achievement tests, and average years of schooling in 1960. Author calculations; see Table 4.1, column (2).

The same pattern of results is preserved when we ignore any variation between world regions by including five regional dummies (column (3) of Table 4.1).⁴⁶ That is, even when considering just the variation that exists within each region, educational quality is significantly related to economic growth, with a coefficient estimate reduced from 1.98 to 1.55.

One of the most important fundamental determinants of economic growth discussed in the recent literature is the institutional framework of the economy (see Section 4.5 below for details). The most common and powerful measures of the institutional framework used in empirical work are the openness of the economy to international trade and the security of property rights.⁴⁷ When we add these two institutional variables to our model, they are jointly highly significant (column (4) of Table 4.1). But the positive effect of educational quality on economic growth is very robust to the inclusion of these controls, albeit slightly reduced in magnitude to 1.26.

Other potential determinants of economic growth often discussed in the literature are fertility and geography. However, when we add the total fertility rate and common geographical proxies, such as latitude or the fraction of the land area of a country that is located within the geographic tropics, to the specification reported in column (4) of Table 4.1, neither of these additional variables is statistically significantly associated with economic growth. Furthermore, none of the other results, in particular for educational quality, are qualitatively affected.

An important issue is whether the role of educational quality in the process of economic development differs between developing and developed countries. In Table 4.2, we divide the sample of countries into two groups in two different ways. First, in columns (5) and (6), we subdivide the sample into OECD countries and non-OECD countries. Results are remarkably similar, with the point estimate of the effect of educational quality slightly larger in non-OECD countries. However, the effect of educational quality on economic growth does not differ significantly between the two groups of

⁴⁶ The regions are East Asia, South Asia, Latin America, Middle East and North Africa, Sub-Saharan Africa, and the industrial countries.

⁴⁷ The proxy for openness used here is the fraction of years between 1960 and 1998 that a country was classified as having an economy open to international trade, based on five factors including tariffs, quotas, exchange rate controls, export controls, and whether or not a socialist economy (cf. Sachs and Warner (1995)). The proxy for security of property rights is an index of the protection against expropriation risk, averaged over 1985-1995, from Political Risk Services (a private company which assesses the risk that investments will be expropriated in different countries), ranging from 0 to 10 (high figures corresponding to low risk), as used by Acemoglu, Johnson, and Robinson (2001) and provided in McArthur and Sachs (2001). This measure of the risk of confiscation and forced nationalization of property is often used as a summary variable for institutional quality, and similar data were first used in this framework by Knack and Keefer (1995).

countries. The results remain qualitatively the same when openness and quality of institutions are again added as control variables.

Table 4.2: Education as Determinant of Growth of Income per Capita, 1960-2000: Sub-samples

| | (5) Developing countries ^a | (6) OECD sample | (7) Low-income countries ^b | (8) High-income countries ^c |
|-------------------------|---|--------------------|---|--|
| GDP per capita 1960 | -0.262 (1.77) | -0.301 (5.81) | -0.063 (0.28) | -0.294 (6.38) |
| Years of schooling 1960 | 0.025 (0.20) | 0.025 (0.26) | 0.006 (0.05) | 0.152 (1.70) |
| Test score (mean) | 2.056 (6.10) | 1.736 (4.17) | 2.286 (6.98) | 1.287 (5.37) |
| Constant | -5.139 (3.63) | -3.539 (1.96) | -6.412 (4.52) | -2.489 (2.86) |
| N | 27 | 23 | 25 | 25 |
| R ² (adj.) | 0.676 | 0.830 | 0.707 | 0.783 |

Dependent variable: average annual growth rate in GDP per capita, 1960-2000. ^a Non-OECD countries. ^b Countries below sample median of GDP per capital 1960. ^c Countries above sample median of GDP per capital 1960.

Second, in columns (7) and (8), we subdivide the sample into countries above and below the sample median of initial GDP per capita. Again, the significant effect of educational quality remains robust in both sub-samples.⁴⁸ Here, the effect of quality is considerably larger in the low-income countries, and the difference in the coefficients on test scores is statistically significant at the five percent level. Thus, if anything, the effect of educational quality is larger in developing countries than in developed countries. Furthermore, the robustness of the effect in the sub-samples is remarkable, in particular considering the small sample sizes of the specifications reported in Table 4.2.

When adding an interaction term between years of schooling and test scores to the specification, the interaction term is insignificant in the full sample of countries. But in the sub-sample of low-income countries (with a GDP per capita below the sample median in 1960), the interaction term is significantly positive. At least in this sample, the returns to increased years of schooling increase with the quality of the education. Once there is a high-quality school system, it pays to keep children longer in school – but it does not if the school system does not produce skills.

⁴⁸ Similarly, the effect of educational quality remains robust in the two sub-samples of countries above and below the median of economic growth over the considered period, with the point estimate larger in the high-growth sample.

Importantly, the results do not appear to be an artifact of the specific time period, set of countries, or achievement measurement decisions. Extensive robustness tests confirm the strong and significant relationship between educational quality and economic growth.

The results reported in Tables 4.1 and 4.2 are robust to several alternative approaches to measuring the quality of education. The results remain qualitatively the same when using only the tests performed at the level of lower secondary education, excluding any test in primary schooling or in the final year of secondary education. Arguably, test scores at the end of the secondary level, which combine the knowledge accumulated over primary and secondary schooling, may be more relevant for the human capital of the labor force than test scores that capture only the knowledge at the end of either primary or lower secondary school. At the same time, the duration of secondary education differs across countries, so that tests performed in the final year of secondary schooling in each country may not be as readily comparable across countries as the grade-based target populations. Further, given differing school completion rates, the test for the final year of secondary schooling may imply cross-country samples with differential selectivity of test takers. Neither the primary-school tests nor the tests in the final secondary year are crucial for the results.

Furthermore, results are qualitatively the same when using only scores on tests performed since 1995. These recent tests have not been used in the previous analyses and are generally viewed as having the highest standard of sampling and quality control. Likewise, results are robust to using tests scores since 1995 for just lower secondary grades.

A drawback of using only the more recent tests is that such an approach requires a strong version of the assumption that test performance is reasonably constant over time, because it relates test performance measured since 1995 to the economic data for 1960-2000. To make sure that higher previous economic growth is not driving the measured test performance, we also used a test score measure that disregards all tests since the late 1990s. Our results turn out to be robust, with a point estimate on the test score variable that is significantly higher, when we restrict the tests to only those conducted until 1995 (sample reduced to 34 countries) and until 1984 (22 countries).⁴⁹ In sum, the results are not driven by either early or late test scores alone.

The results are also robust to performing the analyses in two sub-periods, 1960-1980 and 1980-2000. The most recent period includes the Asian currency crisis and other major economic disruptions

⁴⁹ We can also use the average early test scores (either until 1984 or until 1995) as an instrument for the average of all test scores in a two-stage least-squares regression, in order to utilize only that part of the total test-score measure that can be traced back to the early test scores. Again, our results are robust to such an instrumental-variable approach.

which could affect the apparent impact of educational quality on growth – but they do not. Test scores exert a positive effect on growth in both sub-periods, while years of schooling remain insignificant in both sub-periods. In addition, all results remain qualitatively the same when we use test scores in math and science separately.

Finally, as discussed in the previous section, one may worry about the extent to which East Asian countries are driving the association between educational quality and economic growth. As is obvious from Figure 4.2, several East Asian countries feature both high educational quality and high economic growth, and the top right corner of the figure is noticeably inhabited by East Asian countries. Still, the regression reported above that controls for regional dummies already showed that the association between educational quality and growth is not solely due to a difference between the group of East Asian countries and the rest of the countries. Furthermore, similar to the results reported in Hanushek and Kimko (2000), when we drop all ten East Asian countries from our sample of 50 countries, the estimate on educational quality remains statistically highly significant at a point estimate of 1.3. The significant effect in the sample without East Asian countries is also evident in the two separate sub-periods, with the point estimates being larger in the separate regressions.

4.4 Distribution of Educational Quality and Economic Growth

The results so far consider only the mean of the quality of education in a country. From a policy viewpoint, though, it is important to know whether different parts of the distribution of education affect growth differently. Loosely speaking, is it a few “rocket scientists” at the very top of the distribution who are needed to spur economic growth, or is it “education for all” that is needed to lay a broad base at the lower parts of the educational distribution? Does educational performance at different points in the distribution of the population have separate effects on economic growth?

To estimate such effects, we measure the share of students in each country that reaches a certain threshold of basic literacy at the international scale, as well as the share of students that surpasses an international threshold of top performance. As the two thresholds, we use 400 and 600 test-score points, respectively, on the transformed international scale as shown in Figure 4.1.

The threshold of 400 points is meant to capture a notion of basic literacy. On the PISA 2003 math test, for example, this would correspond to the middle of the level 1 range which denotes that students can answer questions involving familiar contexts where all relevant information is present and the questions are clearly defined.⁵⁰ While the PISA 2003 science test does not define a full set of

⁵⁰ Level 1 in PISA 2003 math was defined by students scoring from 358 to 420.

proficiency levels, the threshold of 400 points is used as the lowest bound for a basic level of science literacy (Organisation for Economic Co-operation and Development (2004), p. 292). In general, a level of 400 points means performance at one standard deviation below the OECD mean. The share of students achieving this level ranges from 18% in Peru to 97% in the Netherlands and Japan, with an international median of 86% in our sample. The threshold of 600 points captures the notion of very high performers, performing at more than one standard deviation above the OECD mean. The share of students achieving this level ranges from below 0.1% in Colombia and Morocco to 18% in Singapore and Korea and 22% in Taiwan, with an international median of 5% in our sample.

When we enter the share of students above the two thresholds jointly in our growth model (see column (9) of Table 4.3), both turn out to be separately significantly related to economic growth. That is, both education for all and the share of absolutely top performers seem to exert separately identifiable effects on economic growth. These initial results should be viewed as suggestive rather than definite, not the least because of issues of multicollinearity between the two measures of educational quality which have a bivariate correlation of 0.73. Importantly, the relative size of the effects of performance at the bottom and at the top of the distribution depends on the specific specification used, and further research is needed to yield more detailed predictions. Nonetheless, the evidence strongly suggests that both dimensions of educational performance count for the growth potential of an economy (see Hanushek and Wößmann (in process) for greater detail).⁵¹

Additional specifications using different points of the distribution of test scores support this general view. The positive effect of educational quality is highly robust to measuring test-score performance at different percentiles of the distribution, such as the 25th, 75th, and 95th percentile, rather than at the mean. When including two measures at a time, they are always jointly, and often separately, significant. When including two percentile measures, it is not always clear whether the higher ones tend to dominate the lower ones or vice versa in different specifications, which are again hampered by strong multicollinearity of the test-score measures. The highest bivariate correlation among the different test-score measures is between mean performance and the share of students in a country reaching the threshold of 400 test-score points, at 0.99. Thus, mean performance seems to capture mainly the pattern of basic literacy across the countries in our sample. The specifications including two

⁵¹ In contrast to the result of Castelló and Doménech (2002) based on years of schooling, we do not find a significant effect of inequality of educational quality per se on economic growth, but rather significant effects at different points of the distribution.

test-score variables remain qualitatively the same when including five regional dummies, and the pattern of significance also holds when the other control variables are included.

Table 4.3: Educational Quality and Growth: Distribution and Institutional Interaction

| | (9) | (10) |
|--|------------------|------------------|
| GDP per capita 1960 | -0.287 (5.12) | -0.297 (5.64) |
| Years of schooling 1960 | 0.022 (0.28) | -0.031 (0.41) |
| Share of students above threshold of 400 | 2.732 (3.61) | |
| Share of students above threshold of 600 | 12.880 (4.35) | |
| Test score (mean) | | 0.942 (2.30) |
| Openness | | 0.732 (2.13) |
| Test score * openness | | 1.609 (2.34) |
| Constant | 1.335 (2.97) | 3.814 (11.24) |
| N | 50 | 47 |
| R ² (adj.) | 0.719 | 0.785 |

Dependent variable: average annual growth rate in GDP per capita, 1960-2000. *t*-statistics in parentheses.

We can also divide our combined test-score measure into one using only the math tests and one using only the science tests. Both subject-specific test scores are significantly associated with growth when entered separately or jointly. There is some tendency for math performance to dominate science performance in different robustness specifications but overall performance in math and science carry separate weight for economic growth.

In sum, different dimensions of the quality of education seem to have independent positive effects on economic growth. This is true both for basic and top dimensions of educational performance and for the math and science dimensions of performance. Because of the thin country samples, however, we trust the pattern of results more than the specific estimates.

4.5 Institutions, Education and Growth

In recent years, there has been an increasing emphasis on the role of economic institutions as the fundamental cause of differences in economic development (cf. Acemoglu, Johnson, and Robinson (2001, 2002, 2005)). As we saw in column (4) of Table 4.1, the quality of institutions as measured by

the protection against expropriation is indeed significantly related to economic growth in our model. A second measure of institutional quality, openness to international trade, also tends to be significantly related to economic growth, at least jointly with protection against expropriation. (Note that protection against expropriation and openness are strongly correlated, with a correlation coefficient of 0.71.) At the same time, though, the results show that, on average, the quality of education exerts a positive effect on economic growth independent of these measures of the quality of institutions.⁵²

Still, the effect of educational quality on economic growth may differ depending on the economic institutions of a country. North (1990), for example, emphasizes that the institutional framework plays an important role in shaping the relative profitability of piracy versus productive activity. If the available knowledge and skills are used in the former rather than the latter activity, one may certainly expect the effect on economic growth to be substantially different, and maybe even to turn negative. Similarly, Murphy, Shleifer, and Vishny (1991) show that the allocation of talent between rent-seeking and entrepreneurship matters for economic growth: countries with relatively more engineering college majors grow faster and countries with relatively more law concentrators grow more slowly. Easterly (2002) argues that education may not have much impact in less developed countries that lack other facilitating factors such as functioning institutions for markets and legal systems. In a similar way, Pritchett (2001, 2006) suggests that due to deficiencies in the institutional environment, cognitive skills might have been applied to socially unproductive activities in many developing countries, rendering the average effect of education on growth across all countries negligible. While his result of negligible average growth effects of increases in educational quantity may have more to do with issues of measuring education than with perverse economic institutions, his point that the social returns to education may be low in countries with perverse institutional environments is certainly worth pursuing.

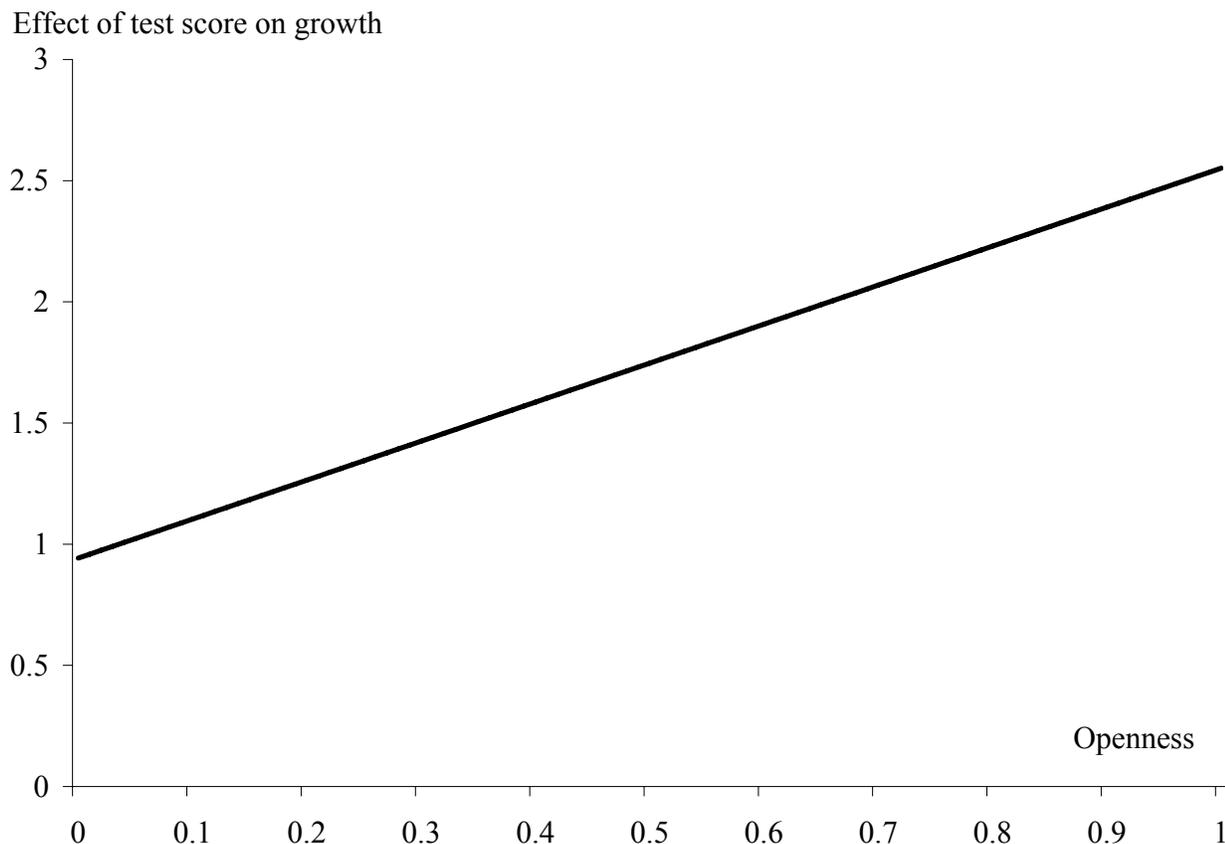
Thus, in column (10) of Table 4.3, we add an interaction term between one of our institutional measures, openness to international trade, to our growth specification.⁵³ The results suggest that openness and educational quality not only have significant individual effects on economic growth but also a significant positive interaction. This result is depicted in Figure 4.3. The effect of educational quality on economic growth is indeed significantly higher in countries that have been fully open to

⁵² While Acemoglu, Johnson, and Robinson (2001, 2002, 2005) find that geographical features do not exert an effect on economic growth independent of institutions, this remains a disputed subject (cf., e.g., McArthur and Sachs (2001)). When we add proxies for geography such as location in the geographic tropics or latitude to our basic specification, these do not enter significantly and do not change the findings on institutions and on educational quality.

⁵³ To facilitate interpretation, the test-score variable is centered in this specification.

international trade than in countries that have been fully closed.⁵⁴ The effect of educational quality on economic growth is significantly positive, albeit relatively low at 0.9, in closed economies, and it increases to a size of 2.5 in open economies.⁵⁵ The reported result is robust to including the measure of protection against expropriation. When using protection against expropriation rather than openness to trade as the measure of quality of institutions, there is similarly a positive interaction term with educational quality, although it lacks statistical significance.

Figure 4.3: The Effect of Educational Quality on Growth Depending on Openness



Notes: Estimated effect of average achievement test scores on the average annual rate of growth of real GDP per capita in 1960-2000, depending on the degree of openness to international trade of a country. Author calculations; see Table 4.3, column (10).

⁵⁴ Jamison, Jamison, and Hanushek (forthcoming) find a similar result, in that the impact of cognitive skills on technical progress is strong in countries with open trade regimes and essentially zero in closed economies.

⁵⁵ There is a similar positive interaction effect between test scores and openness when openness is specified as a dummy for countries that have been closed for the majority of the time (openness below 0.3).

In sum, both the quality of the institutional environment and the quality of education seem important for economic development. Furthermore, the effect of educational quality on economic growth seems to be significantly larger in countries with a productive institutional framework, so that good institutional quality and good educational quality can reinforce each other in advancing economic development. Thus, the macroeconomic effect of education depends on other complementary growth-enhancing policies and institutions. Nonetheless, we still also find a significant positive growth effect of cognitive skills even in countries with a poor institutional environment.

4.6 The Implications of Improved Quality

It is important to understand the implications of policies designed to improve educational outcomes. The previous estimation provides information about the long run economic implications of improvements in educational quality. To understand better the impact of improved achievement, it is useful to relate policy reforms directly to the pattern of economic outcomes that are consistent with feasible improvements.

Two aspects of any educational reform plan are important. First, what is the magnitude of the reform that is accomplished? Second, how fast does any reform achieve its results? While we return later to discuss the efficacy of a variety of potential schooling reforms, here we simply investigate the economic results that might be anticipated on the assumption that some set of schooling reforms actually leads to substantial achievement gains over an identified time period.

Our analysis puts the economic implications in terms of standard deviations of test scores. To provide a benchmark, we consider a reform that yields a 0.5 standard deviation improvement in average achievement of school completers. This metric is hard to understand intuitively, in part because most people have experiences within a single country. It is possible, however, to put this in the context of the previous estimation. Consider a developing country with average performance at roughly 400 test-score points, what we described as approximately minimal literacy on these tests. For example, on the PISA 2003 examinations, average achievement in Brazil, Indonesia, Mexico, and Thailand fell close to this level. An aggressive reform plan would be to close half the gap with the average OECD student, which would relate to a half standard deviation improvement.⁵⁶

⁵⁶ As an alternative policy change, consider what it would mean if a country currently performing near the mean of OECD countries in PISA at 500 test-score points (for example, Norway or the United States in PISA 2000 or Germany in PISA 2003, see Figure 4.1) would manage to increase its educational quality to the level of top performers in PISA at roughly 540 test-score points (for example, Finland or Korea on either PISA test). Such an increase of 40 test-score points

The timing of this reform is also important. Two aspects of timing enter. First, such movement of student performance cannot be achieved instantaneously but requires changes in schools that will be accomplished over some time (say, through systematic replacement of teachers through retirements and subsequent hiring). The time path of any reform is difficult to specify, but achieving the change of 0.5 standard deviations described above for an entire nation may realistically take 20 to 30 years.⁵⁷

Second, if the reforms succeed, their impact on the economy will not be immediate. In particular, the new graduates from school will initially be a small to negligible portion of the labor force. It will be sometime after the reform of the schools before the impact on the economy is realized. In other words, the prior estimates are best thought of as the long run, or equilibrium, outcomes of a labor force with a given educational quality.

To show the impact of these elements of reform, Figure 4.4 simulates the impact on the economy of reform policies taking 10, 20, or 30 years for a 0.5 standard deviation improvement in student outcomes at the end of upper secondary schooling – what we label as a “moderately strong knowledge improvement.” For the calibration, policies are assumed to begin in 2005 – so that a 20-year reform would be complete in 2025. The actual reform policy is presumed to operate linearly such that, for example, a 20-year reform that ultimately yielded $\frac{1}{2}$ standard deviation higher achievement would see the performance of graduates increasing by 0.025 standard deviations each year over the period. It is also necessary to characterize the impact on the economy, which we assume is proportional to the average achievement levels of prime age workers.⁵⁸ Finally, for this exercise we project the growth

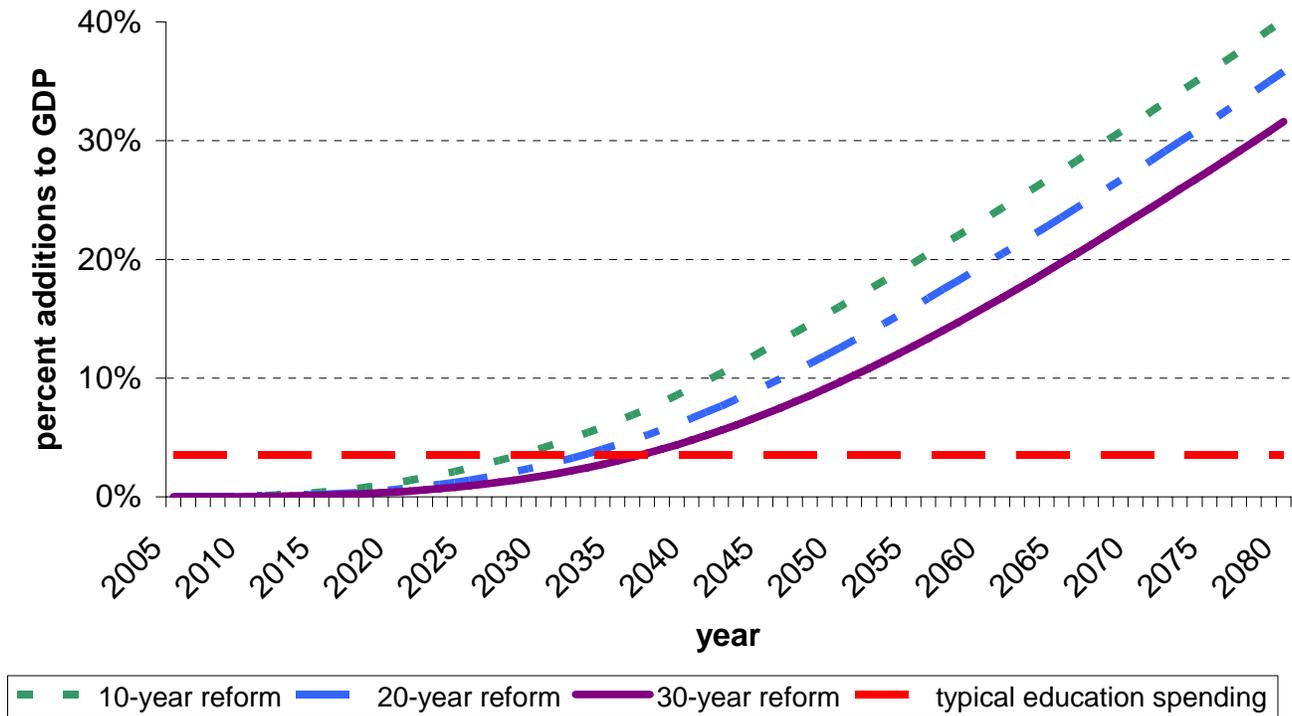
amounts to 0.4 standard deviation. These calculations also point out the asymmetry of the test score distribution with a considerably longer left tail of the country distribution.

⁵⁷ Rough calculations of the relationship between changes in teacher quality and results are found in Hanushek (2005). These are based on assumptions of teacher turnover that mirror the U.S. experience.

⁵⁸ We specifically assume that the relevant achievement levels depend on workers in the first 35 years of their work life and calculate the average achievement levels based upon the progressive improvements of school completers under the different reform periods. The full impact of an educational reform is not felt until 35 years past completion of the reform, e.g., for 45 years with a 10-year reform period. To our knowledge, no past work has indicated how achievement impacts would be felt on the economy. Our linear impact that depends on the workforce average is probably conservative, because new technologies that rely on better trained workers could probably be introduced once there is a substantial number of more skilled workers available instead of relying on improvement throughout the workforce. Alternatively, improvements that depend on innovation may depend most on the stock of young, very highly skilled workers, and the impact of educational improvements might be quite quick.

impact according to the basic achievement model that also includes the independent impact of economic institutions – column 4 of Table 4.1.

Figure 4.4: Improved GDP with Moderately Strong Knowledge Improvement (0.5 s.d.)



The figure indicates how much larger the level of GDP is at any point after the reform policy is begun as compared to that with no reform. In other words, the estimates suggest the increase in GDP expected over and above any growth resulting from other factors.

Obviously, for any magnitude of achievement improvement, a faster reform will have larger impacts on the economy, simply because the better workers become a dominant part of the workforce sooner. But, the figure shows that even a 20- or 30-year reform plan has a powerful impact on GDP. For example, a 20-year plan would yield a GDP that was five percent greater in 2037 (compared with an economy with no increase in educational quality). The figure also plots 3.5 percent of GDP, an aggressive spending level for education in many countries of the world. Five percent of GDP is significantly greater than the typical country’s spending on all primary and secondary schooling, so that it is truly a significant change that would permit the growth dividend to more than cover *all* of primary

and secondary school spending. But even a 30-year reform program (that would not be fully accomplished until 2035) would yield more than five percent higher real GDP by 2041.

Projecting these net gains from school quality further past the reform period shows vividly the long run impacts of reform. Over a 75 year horizon, a 20-year reform yields a real GDP that is 36 percent higher than would be with no change in educational quality.

It must nonetheless be clear that these effects represent the result from *actual gains* in educational outcomes. We return to this below, but there have been many attempts around the world to improve student outcomes, and many of these have failed to yield actual gains in student performance. Bad reforms, those without impacts on students, will not have these impacts.

This simulation shows that the previous estimates of impacts of educational quality on growth have really large impacts on national economies. At the same time, while the rewards are large, they also imply that policies must be considered across lengthy time periods and require patience – a patience that is not always clear in national policy making. We return below to some indications of what kind of speed of educational improvement may be feasible.

These reforms must also be put in a broader perspective, because other kinds of institutional changes and investments will also take substantial time. Changing basic economic institutions, for example, can seldom be accomplished over night, and their impacts will take time before the economy adjusts. Thus, in large part, our calculations about the time path of benefits differ more in the fact that we make it explicit than in its pattern compared to a variety of other economic policies.

Appendix: Data on Quality of Education

The analysis in this paper relies upon a variety of cognitive achievement tests. The central analysis of growth and macroeconomic performance uses information from the international tests of a set of nations voluntarily participating in a cooperative venture under the International Association for the Evaluation of Educational Achievement (IEA) and from the Organisation for Economic Co-operation and Development (OECD). The most recent of the IEA tests is the Trends in International Mathematics and Science Study (TIMSS) for 2003, although there is a much longer history identified below. The OECD tests, called the Programme for International Student Assessment (PISA), began in 2000 and cover all OECD countries plus others.

These tests have different groups of countries, sampling of students, and perspectives on what should be tests (see Neidorf, Binkley, Gattis, and Nohara (2006)). Our approach is to aggregate across the variety of tests for each country in order to develop a composite measure of performance. Perhaps the most important issue is whether or not the tests are measuring a common dimension of cognitive skills. The TIMSS tests of math and science are developed by an international panel but are related to common elements of primary and secondary school curriculum, while the PISA tests are designed to be assessments of more applied ideas. In our development of a common metric we also employ data from the U.S. National Assessment of Educational Progress (NAEP). That test, which is conceptually closest to the TIMSS tests except that it relates more directly to U.S. curriculum, provides information over time on a consistent basis.

Part of the analysis on individual returns relied on the International Adult Literacy Survey (IALS), a set of tests given to 20 countries between 1994-1998. These tests cover several functional areas including: Prose Literacy – the knowledge and skills needed to understand and use information from texts including editorials, news stories, poems, and fiction; Document Literacy – the knowledge and skills required to locate and use information contained in various formats, including job applications, payroll forms, transportation schedules, maps, tables, and graphics; and Quantitative Literacy – the knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials, such as balancing a checkbook, calculating a tip, completing an order form, or determining the amount of interest on a loan from an advertisement. They were designed to be very practical.

Interestingly, the TIMSS tests with their curricular focus and the PISA tests with their real-world application focus are highly correlated at the country level. For example, the correlation coefficients between the TIMSS 2003 tests of 8th graders and PISA 2003 tests of 15-year-olds across the 19

countries participating in both are 0.87 in math and 0.97 in science, and they are 0.86 in both math and science across the 21 countries participating both in the TIMSS 1999 tests and the PISA 2000/02 tests. Similarly, there is a high correlation at the country level between the curriculum based student tests of TIMSS and the practical literacy adult examinations of IALS (Hanushek and Zhang (2006)). Tests with very different foci and perspectives tend, nonetheless, to be highly related, lending support to our approach of aggregating different tests for each country.

The general idea behind our approach to aggregation is that of empirical calibration. We rely upon information about the overall distribution of scores on each test to compare national responses. This contrasts with the psychometric approach to scaling that calls for calibrating tests through use of common elements on each test. In reality, each of the testing situations described below is a separate activity with no general attempt to provide common scaling.

The strength of our approach is that different tests across a common subject matter tend to be highly correlated at both the individual and aggregate level. Thus, the distributional information that we use is closely related to variations in individual performance levels.

As shown in Table A4.1, there are data from international student achievement tests on twelve international testing occasions. Containing separate tests in different subjects and at different age groups, these testing occasions yield 36 separate test observations altogether, each with between 11 and 45 participating countries with internationally comparable performance data. Most of the tests were conducted by the IEA, with the exception of the OECD-conducted PISA tests.⁵⁹

In order to make performance on the different international tests comparable, Hanushek and Wößmann (in process) develop a common metric to adjust both the level of test performance and the variation of test performance through two data transformations. First, because the United States has both participated in all of the international tests and has maintained its own longitudinal testing (the NAEP), Hanushek and Wößmann (in process) calibrate the U.S. international performance over time to the external standard – thus benchmarking each of the separate international tests to a comparable level. Second, while this provides a relative comparison of countries taking each test over time, it is also necessary to establish the variance on the tests so that direct compatibility of countries taking different

⁵⁹ In this study, we do not include the two tests conducted by the International Assessment of Educational Progress (IAEP) in 1988 and 1991, because they used the U.S. NAEP test as their testing instrument, which is geared to the U.S. curriculum and may thus introduce bias to the international testing. By contrast, the tests included here are not associated with the curriculum in any particular country, but have been devised in an international cooperative process between all participating countries.

tests can be established. The calibration of the dispersion of the tests relies on holding the score variance constant within a group of countries with stable education systems (defined in terms of secondary school attendance rates) over time. For this, Hanushek and Wößmann (in process) use the thirteen OECD countries who had half or more students completing upper secondary education around the beginning of international testing in the 1970s as the “stable” country group, and standardize variances to their group performance on the 2000 PISA tests. The details of the transformation are found in Hanushek and Wößmann (in process).

Table A4.1: The International Student Achievement Tests

| Abbr. | Study | Year | Subject | Age ^{a,b} | Countries ^c |
|----------------|---|---------|--------------------------|-----------------------|------------------------|
| 1 FIMS | First International Mathematics Study | 1964 | Math | 13,FS | 11 |
| 2 FISS | First International Science Study | 1970-71 | Science | 10,14,FS | 14,16,16 |
| 3 FIRS | First International Reading Study | 1970-72 | Reading | 13 | 12 |
| 4 SIMS | Second International Mathematics Study | 1980-82 | Math | 13,FS | 17,12 |
| 5 SISS | Second International Science Study | 1983-84 | Science | 10,13,FS | 15,17,13 |
| 6 SIRS | Second International Reading Study | 1990-91 | Reading | 9,13 | 26,30 |
| 7 TIMSS | Third International Mathematics and Science Study | 1994-95 | Math/Science | 9(3+4), 13(7+8),FS | 25,39,21 |
| 8 TIMSS-Repeat | TIMSS-Repeat | 1999 | Math/Science | 13(8) | 38 |
| 9 PISA 2000/02 | Programme for International Student Assessment | 2000+02 | Reading/ Math/Science | 15 | 31+10 |
| 10 PIRLS | Progress in International Reading Literacy Study | 2001 | Reading | 9(4) | 34 |
| 11 TIMSS 2003 | Trends in International Mathematics and Science Study | 2003 | Math/Science | 9(4),13(8) | 24,45 |
| 12 PISA 2003 | Programme for International Student Assessment | 2003 | Reading/ Math/Science | 15 | 40 |

Notes: ^a Grade in parentheses where grade level was target population. ^b FS = final year of secondary education. ^c Number of participating countries that yielded internationally comparable performance data.

To create our measure of quality of education employed in this study, we use a simple average of the transformed mathematics and science scores over all the available international tests in which a country participated, combining data from up to nine international testing occasions and thirty

individual test point observations.⁶⁰ This procedure of averaging performance over a forty year period is meant to proxy the educational performance of the whole labor force, because the basic objective is not to measure the quality of students but to obtain an index of the quality of the workers in a country. If the quality of schools and skills of graduates are constant over time, this averaging is appropriate and uses the available information to obtain the most reliable estimate of quality. If on the other hand there is changing performance, this averaging will introduce measurement error of varying degrees over the sample of economic data (1960-2000). The analysis in Hanushek and Wößmann (in process) shows some variation over time, but there is no clear way to deal with this here.

When looking at effects of the distribution of educational quality, we go beyond the mean performance of a country's students and calculate the share of students above a certain test-score threshold, as well as the performance of students at different percentiles of a country, both based on our transformed metric of scores. This requires going into the details of the student-level micro data for each international test, which we can do for each test in mathematics and science with the exception of FIMS, where the micro data seem no longer accessible. See Hanushek and Wößmann (in process) for details.

By now, there are a total of 77 countries that have ever participated in any of the nine international student achievement tests in mathematics and science. 50 of these are included in the analyses of economic growth. 25 countries are not included in the growth database due to lack of data on economic output or because they drop out of the sample for a standard exclusion criterion in growth analyses (15 former communist countries, 3 countries for which oil production is the dominant industry, 2 small countries, 3 newly created countries, 2 further countries lacking early output data). Two countries (Nigeria and Botswana) turn out to be strong outliers in the growth regressions and are therefore dropped from the sample.⁶¹

⁶⁰ Note that the descriptive representation in Figure 4.1 also reports performance on the reading tests, which are not used for the growth analyses here because they may be affected by testing in different languages and may not be easily combined into a common one-dimensional scale with mathematics and science tests.

⁶¹ There are four countries with test-score data which have a few years missing at the beginning or end of the 1960-2000 period on the income data in the Penn World Tables. In particular, data for Tunisia starts in 1961 rather than 1960, and data for Cyprus (1996), Singapore (1996), and Taiwan (1998) end slightly earlier than in 2000. These countries were included in the growth regressions by estimating average annual growth over the available 36-39 year period rather than the whole 40 year period.

5. Where Does the Developing World Stand?

Given the crucial importance of cognitive skills for economic development, it is telling to document how the developing countries fare in this regard. In the following, we document both the quantity of schooling and quality of education achieved by developing countries from an international perspective, and this vividly illustrates the magnitude of the task at hand.

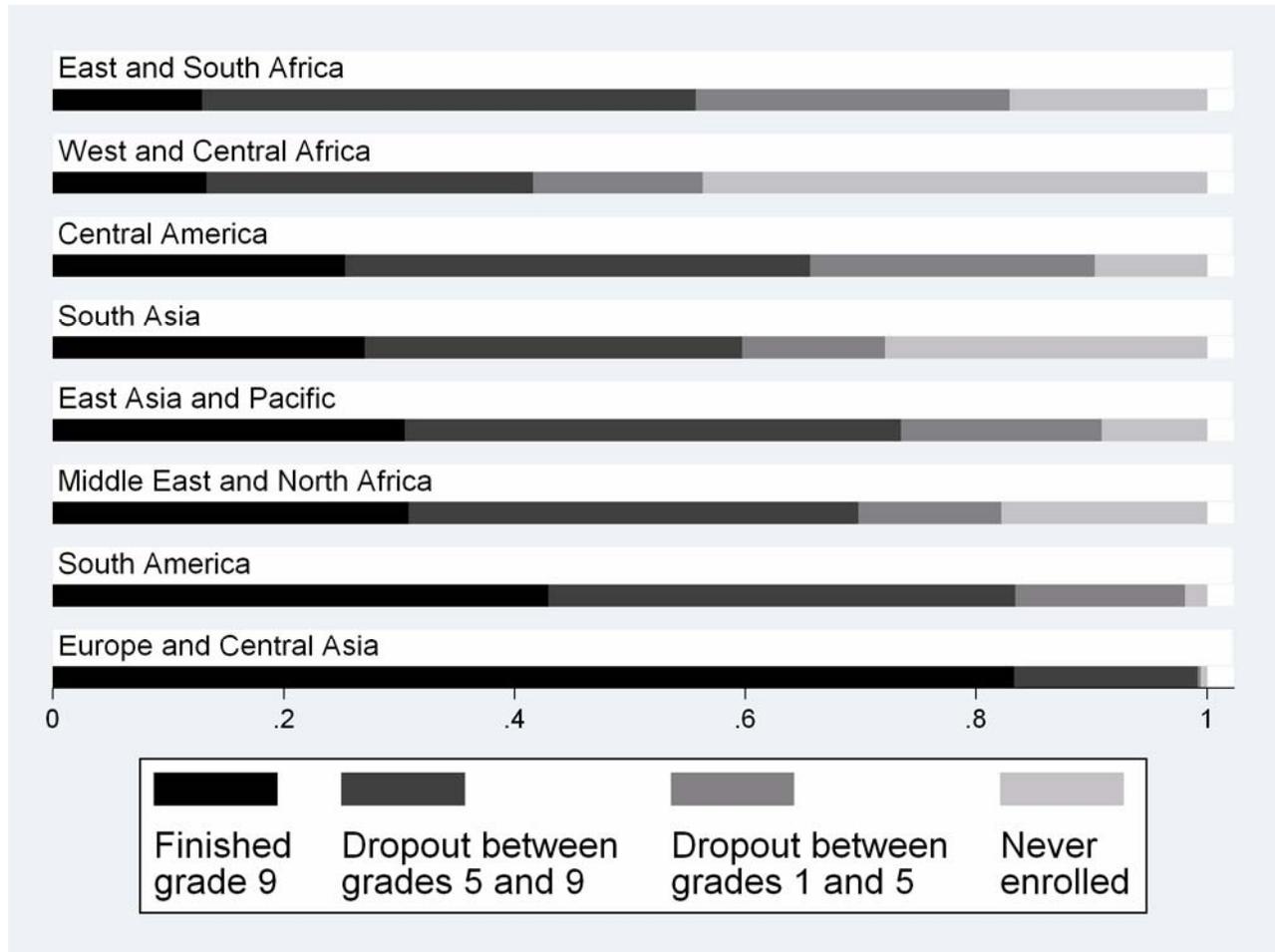
5.1 *Lack of Quantity of Schooling*

The disadvantages of less developed countries in terms of educational enrollment and attainment have been well documented and are well known. To provide an aggregate picture, Figure 5.1 presents the share of children aged 15-19 who have completed grade 9, dropped out between grades 5 and 9 or between grades 1 and 5, or never enrolled, for eight regions of developing countries.⁶² The shares are unweighted averages of the available countries in the Demographic and Health Surveys (DHS) and similar surveys (cf. Filmer (2006)), conducted in a large number of developing countries. These household survey data are much more reliable than the administrative data on school enrollment rates traditionally used, because the administrative data available in many developing countries often overstate actual enrollment rates. Further, the self-reported data on the highest grade completed is more relevant for actual education than enrollment rates, which are affected by grade repetition and the like.

While almost all OECD countries have universal school attainment to grade 9, essentially all developing regions are far from that. In the average African country in the data, only 13% of each cohort finishes grade 9, and less than 30% in Central America and South and East Asia. Even in South America, this figure is only 43%, although on the other hand only 17% of a cohort do not complete grade 5 (which often serves as an initial indication of basic literacy and numeracy rates). In West and Central Africa, 59% of each cohort do not even complete grade 5, and 44% never enroll in school in the first place. It is notable across countries, however, that the lack of grade 5 attainment is at least as often due to dropping out of school than due to never enrolling. In any event, while the pattern of educational attainment varies greatly across countries and regions, the lack of quantitative educational attainment from universal completion of basic education – be it grade 5 or grade 9 – is immense in the majority of developing countries.

⁶² Grades 5 and 9 are chosen as two possible definitions of basic schooling. The age range 15-19 is chosen to balance the relevance of the data for most recent conditions against the censoring problem of some children still being in school. The basic patterns observed are the same when using estimates adjusted for the censoring (Pritchett (2004)).

Figure 5.1: Lack of Educational Attainment in Developing Countries



Note: Based on Pritchett (2004).

Focusing on this dimension of schooling quantity, many policy initiatives of national governments and international development agencies have tried to increase the educational attainment of the population. The data in Figure 5.1 show that there remains a long way to go. But even this dire picture may understate the challenge that becomes apparent when quality is also considered.

5.2 Lack of Quality of Education

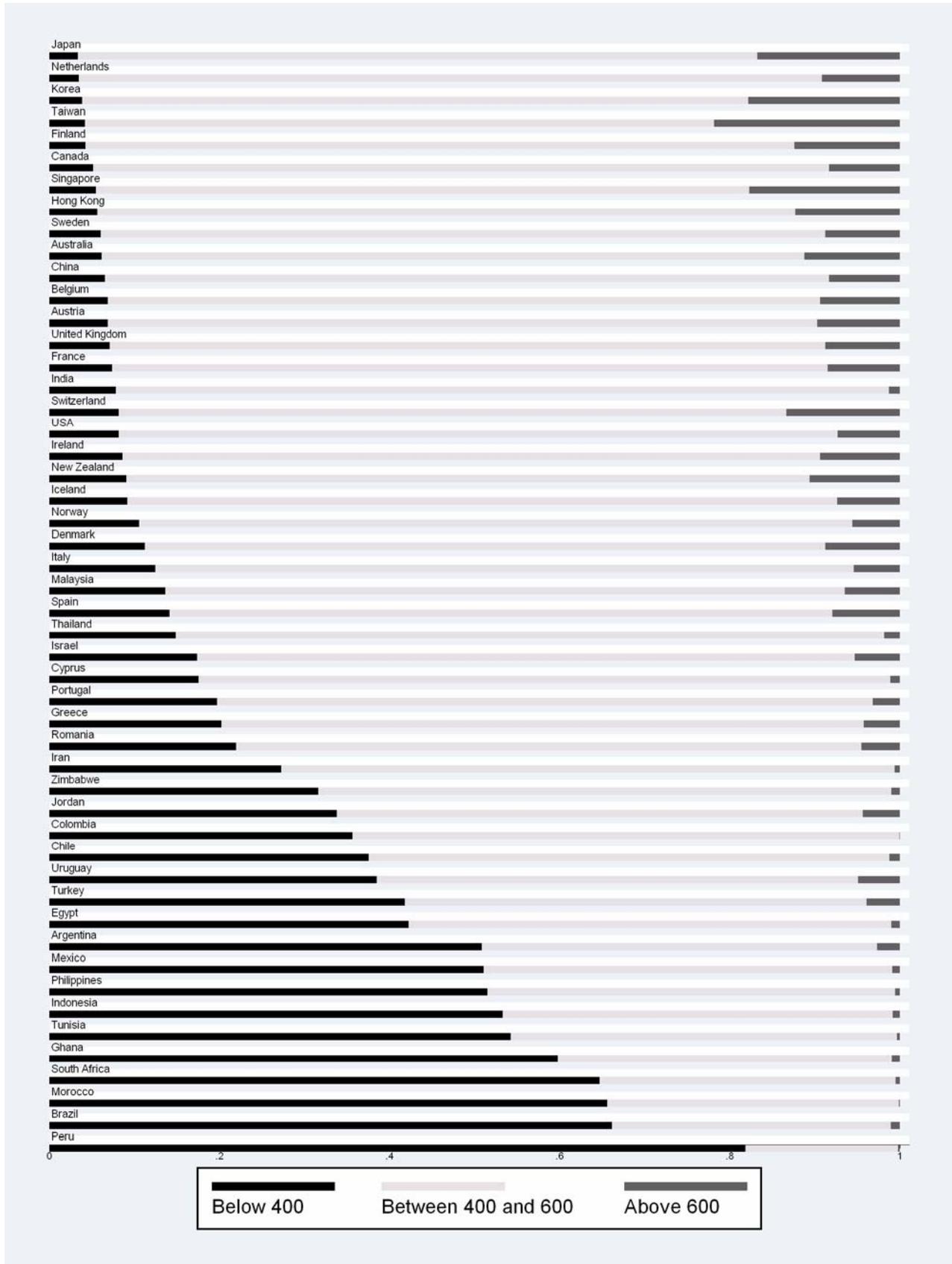
The description of school completion unfortunately ignores the level of cognitive skills actually acquired. Completing 5 or even 9 years of schooling in the average developing country does not necessarily mean that the students have become functionally literate in terms of basic cognitive skills. As a recent report by the World Bank Independent Evaluation Group (2006) documents, high priority was accorded to increasing enrollment in primary schools in developing countries over the past 15 years, but much less attention was directed to whether children are learning adequately. In Figure 4.1

above, we have already documented the particularly low levels of mean performance of students attending school in basically all the developing countries that have participated in at least one of the international student achievement tests. But of course, mean performance can hide a lot of dispersion that exists within countries, and the prior analyses of growth (Section 4.4) show that there is separate information at different percentiles of the test-score data.

Figures 5.2 and 5.3 depict the share of students in a country that surpasses the thresholds of 400 and 600 test-score points on the transformed scale of the combined international tests – the same measure and thresholds that we used in the growth analyses of Section 4.4. Figure 5.2 shows the sample of 50 countries which forms the basis of our growth analyses, and Figure 5.3 shows the remaining 27 countries that participated in one of the international student achievement tests but lack internationally comparable GDP data for the 1960-2000 period that would allow them to be included in the growth analyses.

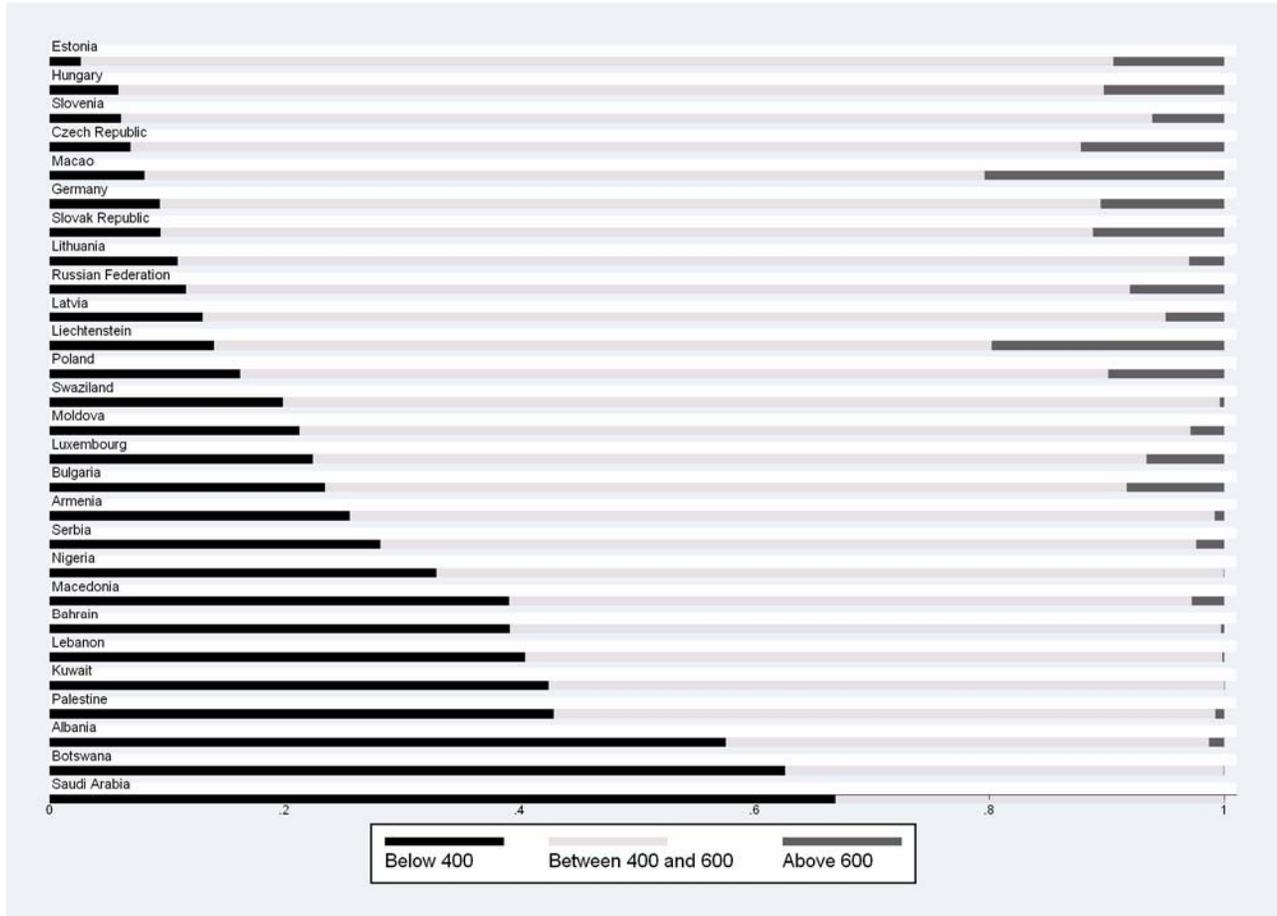
When considering the basic educational achievement of students, we are interested in the share of students who surpass the threshold of 400 test-score points, which may be viewed as a rough threshold of basic literacy in mathematics and science. As is evident from the figures, this share varies immensely across countries. In countries such as Japan, the Netherlands, Korea, Taiwan, and Finland, less than 5% of tested students fall below this literacy threshold. By contrast, in many of the developing countries participating in the international achievement tests, more than half of the tested students do not reach this threshold of literacy. The countries with the largest shares of students who are functionally illiterate are Peru (82%), Saudi Arabia (67%), Brazil (66%), Morocco (66%), South Africa (65%), Botswana (63%), and Ghana (60%). In these countries, more than 60% of those in school do not reach a level of basic literacy in cognitive skills. It should be noted that the group of developing countries participating in the international tests is probably already a select sample from all developing countries, and, furthermore, the children enrolled in school at the different testing grades are probably only a select group of all children of the respective age in these countries.

Figure 5.2: Share of Students below 400 (“Illiterate”), between 400 and 600 and above 600



Source: Hanushek and Wößmann (in process), based on several international tests; see text for details.

Figure 5.3: Share of Students below 400 (“Illiterate”), between 400 and 600 and above 600



Source: Hanushek and Wößmann (in process), based on several international tests; see text for details.

5.3 The Size of the Task at Hand: Schooling Quantity and Educational Quality Combined

We have seen that developing countries are severely lacking in terms of both schooling quantity and quality of education. Figure 5.4 shows the combination of the two. For the 14 countries that both have reliable attainment data from the household surveys and have participated in the international student achievement tests, we combine educational attainment of 15-to19-year-olds from the latest available year with test scores at the end of lower secondary education (eighth grade or 15-year-olds) from a year close by.⁶³ This allows us to calculate rough shares among recent cohorts of school-leaving

⁶³ Specifically, the years of the household survey data and the associated tests (where TIMSS always refers to the respective 8th grade sub-tests) are as follows: Albania and Peru: attainment data for 2000, combined with test scores from PISA 2002; Armenia: 2000 and TIMSS 2003; Brazil: 1996 and PISA 2000; Colombia: 2000 and TIMSS 1995; Egypt,

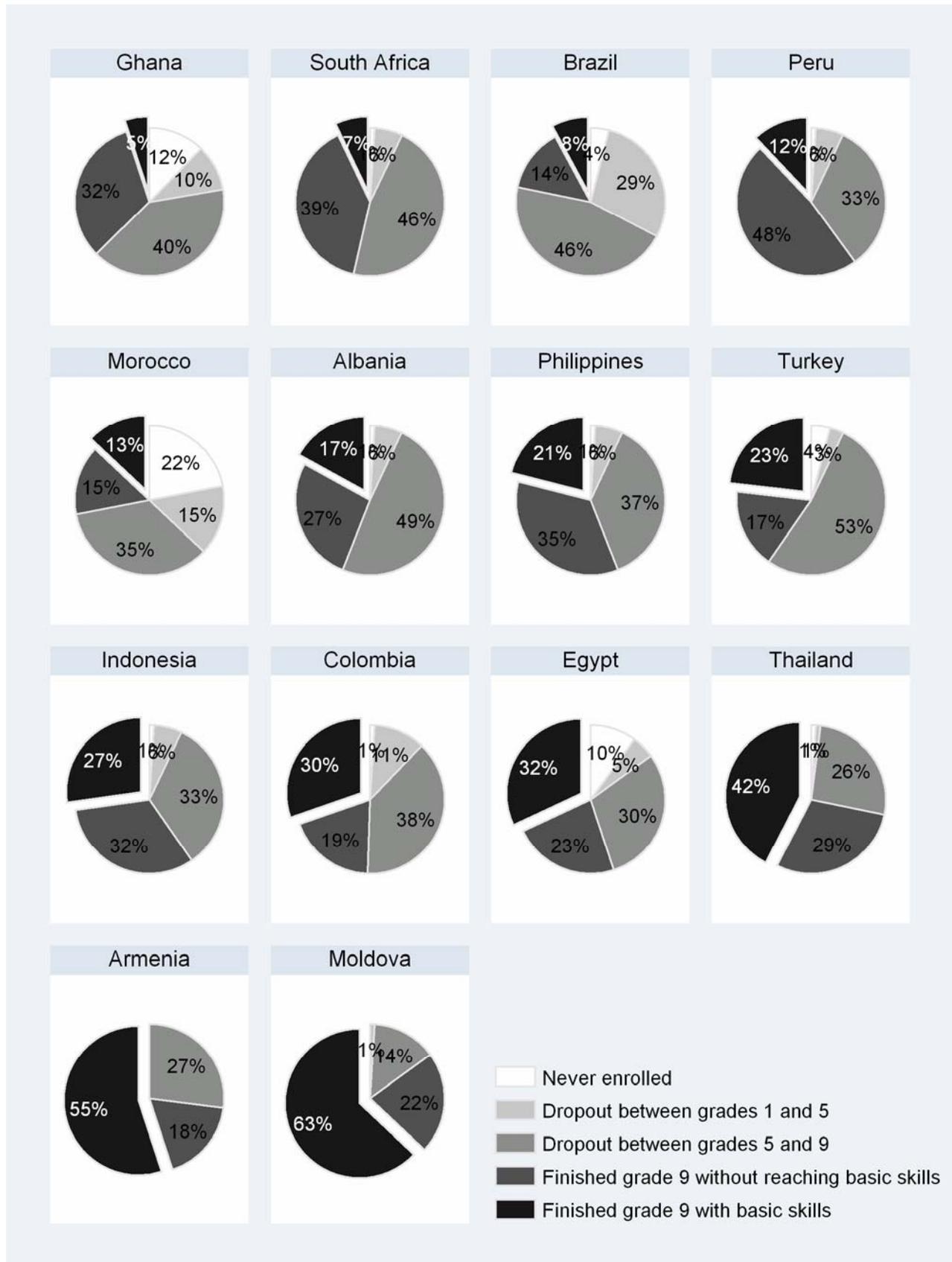
age of how many were never enrolled in school, how many dropped out of school by grade 5 and by grade 9, how many finished grade 9 with a test-score performance below 400 which signals functional illiteracy, and finally how many finished grade 9 with a test-score performance above 400. Only the last group can be viewed as having reached basic literacy in cognitive skills (cf. Pritchett (2004); Wößmann (2004)).

As the results in Figure 5.4 reveal, in 11 of the 14 countries, the share of fully literate students in recent cohorts is less than one third. In Ghana, South Africa, and Brazil, only 5%, 7%, and 8% of a cohort, respectively, reach literacy. The remaining more than 90% of the population are illiterate: because they never enrolled in school; because they dropped out of school at the primary or early secondary level; or because even after completing lower secondary education, their grasp of basic cognitive skills was so low that they have to be viewed as fundamentally illiterate. In contrast, 55% of a cohort in Armenia and 63% in Moldova can be viewed as literate at the end of lower secondary schooling.

An example of a basic test question from one of the international achievement tests can, perhaps better than anything, illustrate the scope of the problem in developing countries. One question asked to 8th graders in TIMSS 2003 was: “Alice ran a race in 49.86 seconds. Betty ran the same race in 52.30 seconds. How much longer did it take Betty to run the race than Alice? (a) 2.44 seconds (b) 2.54 seconds (c) 3.56 seconds (d) 3.76 seconds.” While 88% of 8th-grade students in Singapore, 80% in Hungary, and 74% in the United States got the correct answer (a), only 19% of students in 8th grade in Saudi Arabia, 29% in South Africa, and 32% in Ghana got the correct answer (cf. Pritchett (2004) for a similar example). Random guessing would have yielded 25% correct on average.

Ghana, and Morocco: 2003 and TIMSS 2003; Indonesia: 2002 and average of TIMSS 2003 and PISA 2003; Moldova: 2000 and average of TIMSS 1999 and TIMSS 2003; Philippines: 2003 and average of TIMSS 1999 and TIMSS 2003; South Africa: 1999 and TIMSS 1999; Thailand: 2002 and PISA 2003; Turkey: 1998 and TIMSS 1999.

Figure 5.4: Types of Lack of Education in Selected Countries (15-19-year-olds)



Note: Own calculations based on Filmer (2006) and micro data from different international student achievement tests.

When we combine data on quantitative educational attainment and qualitative achievement of cognitive skills for the countries with reliable data on both dimensions, it becomes apparent that the task at hand is truly staggering in most developing countries. In many developing countries, the share of any cohort that completes lower secondary education and passes at least a low benchmark of basic literacy in cognitive skills is *below one person in ten*. Thus, the education deficits in developing countries seem even larger than generally appreciated. Several additional references for examples of extremely low educational performance of children even after years of schooling from different developing countries are provided in Pritchett (2004). With this dismal state of the quantity and quality of education in most developing countries, the obvious remaining question is, what can be done?

6. Educational Spending and Student Outcomes

The previous discussion has stressed the importance of cognitive skills for economic outcomes – individual earnings, income distribution, and economic growth. By virtually any standard, the evidence indicates a strong causal impact of skills on these outcomes. Yet simply knowing that skill differences are important does not provide a guide to policies that might promote more skills. Indeed, a wide variety of policies have been implemented within various countries without much evidence of success in either achievement or economic terms. This history has led some to believe that education is not a useful policy lever to achieve economic development goals.

We believe that the disappointing results of the past generally reflect pursuing policies for which there is little empirical support. But the past outcomes should not be generalized to statements about all policies.

Substantial research has gone into understanding why achievement differs across students and across countries. While controversies have existed about the interpretation of various evidence, some considerable agreement now exists about what kinds of approaches are unreliable avenues for change. There is perhaps less agreement on the precise approaches that might be followed, but even here there is a growing consensus about the general sets of policies that have shown promise based on more credible research approaches.

The work on achievement determinants is generally labeled “education production functions.” The extensive work has taken a variety of perspectives and approaches. The general objective is to sort out the causal impacts of school factors (things that can potentially be manipulated through policy) from other influences on achievement including family background, peers, neighborhood influences, and the like.

The most important generally available evidence relates to the effects of resources. Many policies undertaken involve substantial flows of resources – direct spending, changes in teacher salaries, reductions in class size, and the like – made within the context of current school organization. The empirical evidence clearly documents the difficulties with such policies. Simply providing more resources gives, according to the available evidence, little assurance that student performance will improve significantly.

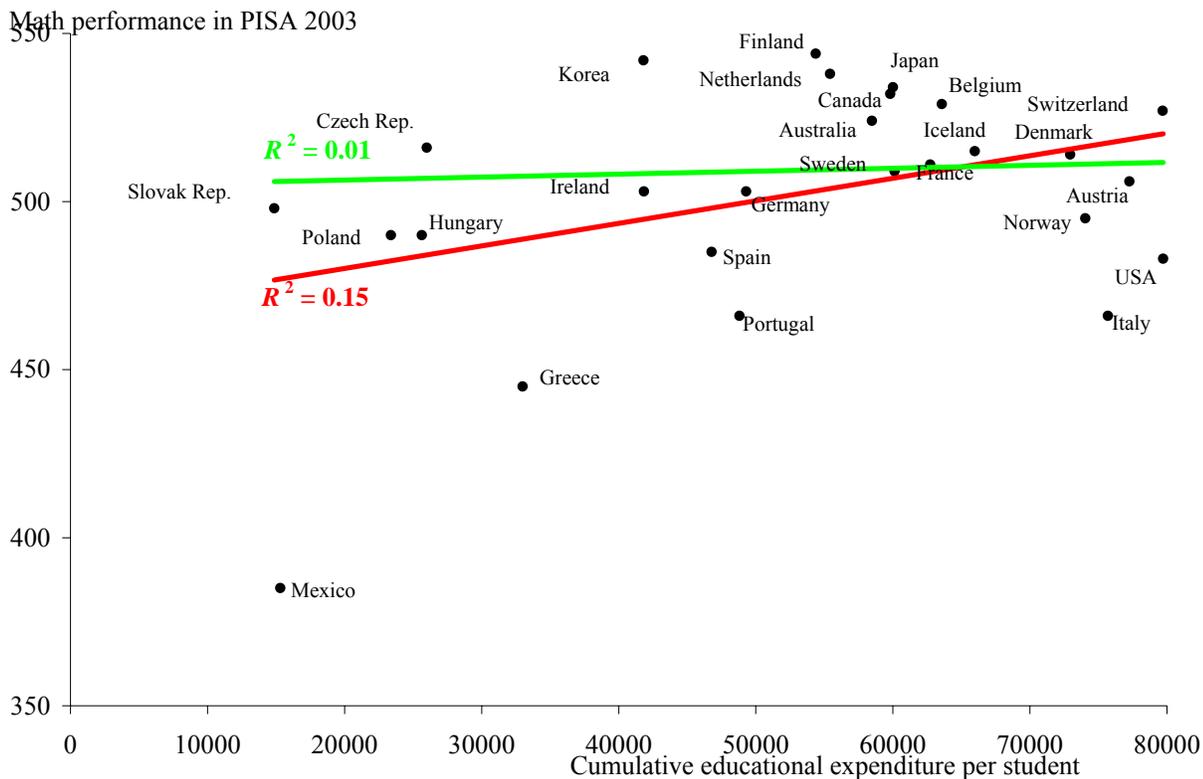
This section reviews the international evidence on resources, spending, and student outcomes. The next section discusses larger questions of institutional design.

6.1 Cross-country Evidence on Resources

The easiest starting point is a simple picture of outcomes across countries. Figure 6.1 presents the international association between spending levels and math performance in the latest international test, the 2003 cycle of PISA conducted by the OECD.

Figure 6.1: Expenditure per Student and Student Performance across Countries

Association between average math performance in PISA 2003 and cumulative expenditure on educational institutions per student between age 6 and 15, in US dollars, converted by purchasing power parities.



Source: Organisation for Economic Co-operation and Development (2004), pp. 102 and 358; Wößmann (forthcoming-a).

The dark line in Figure 6.1 represents the regression line for PISA scores on cumulative expenditure (age 6-15).⁶⁴ Taken literally, this line indicates that \$60,000 per student in additional expenditure (a quadrupling of spending in the low spending countries) is associated with about a half standard deviation improvement in scores. But, this relationship is almost entirely due to the two

⁶⁴Organisation for Economic Co-operation and Development (2004), p. 101 uses this figure to conclude that there is a positive relationship between expenditure and student performance, because the correlation is weakly significant (in statistical terms) when considering all countries and ignoring other effects (corresponding to the weakly upward sloping regression line).

spending and performance outliers – Mexico and Greece. It is impossible to believe that the only difference between these two countries and the remainder is just their spending on schools. For example, there are four countries spending less than Greece but with their students significantly outperforming Greek students. Omitting Mexico and Greece, there is no relationship whatsoever between expenditure and performance (see the lighter regression line in Figure 6.1). On average, countries with high educational expenditure perform at the same level as countries with low educational expenditure.

This picture is only the most recent demonstration that spending alone is not associated with student performance. This picture can be found with previous international student tests like TIMSS.⁶⁵ As discussed previously, Hanushek and Kimko (2000) take into account other factors including parental education in their investigation of earlier test score differences but find no relationship with expenditure per pupil, expenditure as a fraction of GDP, or pupil-teacher ratios. Similarly, even when numerous family-background and school features are considered in cross-country student-level microeconomic regressions, these results hold.⁶⁶

Nor does the picture change when we look at changes in expenditure over time within individual countries. As depicted in Figure 6.2, an in-depth study of changes over time in educational expenditure and student performance has shown that educational expenditure per student has increased substantially in real terms in all considered OECD countries between the early 1970s and the mid-1990s.⁶⁷ Yet, comparing scores in 1970 and 1994/95 suggests that no substantial average performance improvements for students have occurred in these countries.⁶⁸

⁶⁵ See, for example, Wößmann (2002), Section 3.2.

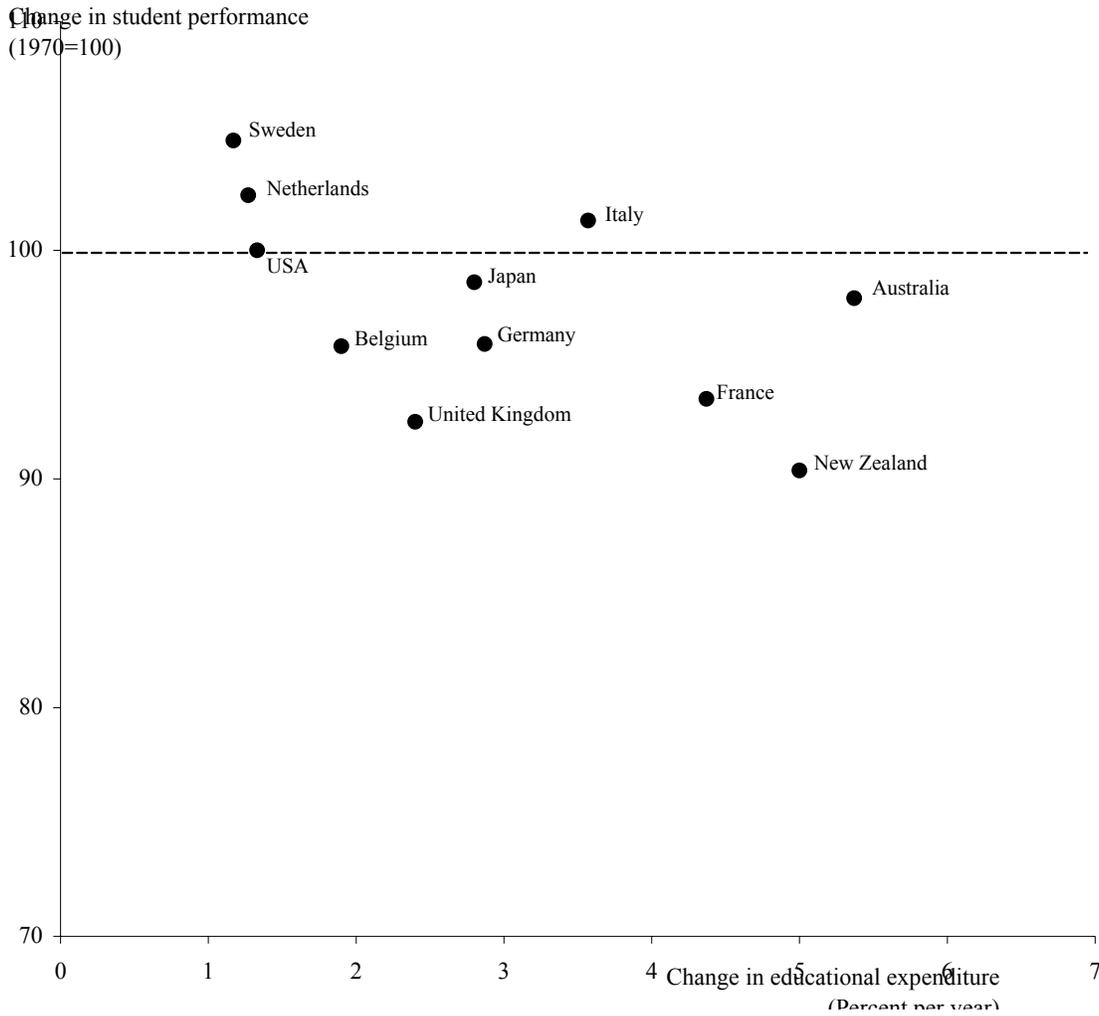
⁶⁶ See Wößmann (2001); Wößmann (2003a), pp. 117-170; Fuchs and Wößmann (forthcoming).

⁶⁷ For details of the analysis, see Gundlach, Wößmann, and Gmelin (2001), pp. C135-C147; and Wößmann (2002), Section 3.3. Tests in the two time periods are linked using U.S. performance data. Increases in educational expenditure are adjusted not only for average inflation, but also for the so-called “Baumol effect” of increasing costs in service sectors without productivity increases. Three different approaches of calculating price deflator for the schooling sector that account for this effect are averaged in the depiction of Figure 6.2. Not adjusting for this and using U.S. national data, real current expenditure per student in the United States increased by 2.4 percent each year on average over 1970-2003.

⁶⁸ One potential explanation for this bivariate pattern might of course be that students’ family backgrounds might have deteriorated on average. Today’s students may lack many of the basic capabilities required for a successful education and may thus be increasingly expensive to educate. Such effects may play a significant role in countries with a large inflow of immigrant students or with rising levels of poverty. But on average, parents in the considered countries enjoy higher incomes and are better educated than 25 years ago, and the number of children per family has declined. Hence children may

Figure 6.2: Expenditure per Student and Student Performance over Time

Association between change in student performance (1994/95 relative to 1970) and change in educational expenditure (average annual rate of change of real educational expenditure per student between 1970 and 1994/95 in percent).



Source: Wößmann (2002), p. 106.

This evidence, while covering a wide range of countries, is again lightly represented in developing countries, because developing countries have not participated very frequently in the various international tests. This nonparticipation is itself an important policy issue. It is difficult to know what improvements are needed or whether any policy changes are having an impact without accurately measuring student performance.

Comparisons across countries are nonetheless difficult in this area for a number of reasons. First, to compare spending, it is necessary to deal with exchange rate differences and with different earnings

actually start schooling with better basic capabilities than ever before. Nonetheless, these issues are better dealt with directly through econometric analyses.

and spending levels. Second, because the schooling institutions can vary in a range of ways, it is difficult to ensure that other influences on achievement are properly accounted for in the analysis. For these reasons, we turn to data generated within separate countries.

6.2 *Within-country Evidence – Developed Countries*

For policy deliberations, information on resources from within individual countries is perhaps more appropriate than cross-country information. Most importantly, it generally allows for much more detailed analyses that can more reliably isolate factors subject to policy decisions. The open question, to which we return below, is whether information gleaned from the schools in one country translates into reliable advice in other countries. We begin by reviewing evidence from developed countries and move to developing countries in the next section.

Researchers have studied the determinants of student achievement for more than 40 years. The work was begun in the United States in the “Coleman Report” (Coleman et al. (1966)), which introduced the idea of using statistical analysis to relate various inputs of schools to student outcomes. This work also underscored the importance of including nonschool factors by demonstrating that family differences were very important in explaining variations in achievement across students. While this original study has been subjected to considerable criticism, it led to an extensive line of research.

The general picture about school resources in developed countries is now well known and has been reviewed in a variety of places (see Hanushek (2002b), Hanushek (2003), Wößmann (2005b)). Here we simply review the highlights and point to the key open issues.

The available studies concentrate on various common inputs to schools such as teacher experience or class size. These factors are both readily available in both administrative and survey data sets and frequently identified as the focus of policy. The available econometric evidence now includes literally hundreds of separate estimates within the U.S. and other developed countries. Quite uniformly, however, there is little strong evidence that *any* of the following factors has a consistent impact on achievement: the level of teacher education, the pupil-teacher ratio, the characteristics of administration, or the facilities of the school. Specifically, aggregating results across studies, a minority of estimates are statistically different from zero (at the 5 percent level or better), and the studies do not even uniformly indicate improvements in performance with increased resources.

The two factors that more frequently appear to positively influence student outcomes are teacher experience and measures of teacher achievement tests. We discuss these below under the assessment of the quality of studies.

A second line of studies focuses on financial inputs. A number of studies simply relate spending per student to achievement or capture teacher differences by teacher salaries. These studies also fail to show a consistent relationship between financial resources and achievement.

These results have been controversial. A variety of debates have taken place around the correct interpretation of prior work (see, for example, Burtless (1996)). The most important line of debate has involved study quality and whether or not these works adequately control for various inputs that might complicate the interpretation of resources. For example, the statistical models may not adequately account for other inputs that affect achievement such as the quality of family inputs; the estimates might then erroneously attribute the higher achievement due to better family factors to some of the characteristics of schools.

It is easiest to think of these problems as the same as those discussed under causality when looking at the aggregate economic outcomes. Take the simple example of the kinds of complications that could come from compensatory inputs to schooling. Specifically, consider a school in which some students are performing poorly and are falling behind the rest of the students in the school. A common policy might provide extra resources to these poor performers in an effort to help them catch up. For example, the school might provide smaller classes for the poor performers so that these students could get extra attention. In such a case, if the study did not adequately account for their prior performance, the statistical analysis might indicate that smaller classes were associated with lower performance – all because the performance of the students “caused” the class size to be smaller.

A variety of approaches have been developed to circumvent problems such as these and to obtain estimates of the causal effects. One direct way has been to estimate statistical achievement models that consider “value added” of schools by introducing prior achievement levels of students into the analysis.⁶⁹ When just the results of resource impacts from the more refined approaches are considered, there is little change in the overall picture. The common measures of school resources still do not have a systematic impact on performance.

Alternative approaches beyond these statistical procedures have utilized specialized models of achievement, have introduced random assignment to different programs, or have capitalized on particular characteristics of schools in order to obtain better estimates of the impact of differing resources. A large part of this effort has been directed at uncovering the impacts of class size on student achievement. This effort reflects two aspects of class size: First, class size reduction is a popular and

⁶⁹ The details of these approaches can be seen in Hanushek (2002b).

frequently considered policy change; second, it is a very well-defined and discrete treatment that is amenable to various analytical approaches (as contrasted with vaguer concepts such as teacher quality).

The extensive focused analysis on the causal effects of class size reduction unfortunately does not provide any clear indication that such policies consistently improve student outcomes. This work, again well-summarized elsewhere (Hanushek (2003), Wößmann (forthcoming-a)), sometimes indicates benefits of reductions and equally as frequently does not.⁷⁰ From these analyses, general class size reductions do not appear to be worth their expense in terms of student outcomes.

The interpretation of this work is important, because it also carries over to discussions of developing country policies. A simplistic view of this argument – convenient as a straw man in public debates – is that ‘money never matters.’⁷¹ The research of course does not say that. Nor does it say that ‘money cannot matter.’ It simply underscores the fact that there has historically been a set of decisions and incentives in schools that have blunted any impacts of added funds, leading to inconsistent outcomes. That is, more spending on schools has not led reliably to substantially better results. We return to these issues below.

The most current research on school inputs and achievement has also led to another set of conclusions – that teacher quality is enormously important in determining student achievement. This work has concentrated on whether some teachers consistently produce more gains in student achievement than other teachers.⁷² Working with extensive panel data on individual students, these studies have confirmed large differences among teachers in terms of outcomes in the classroom.

But, they have also shown that the observed differences are not closely related to common characteristics of teachers (such as amount of teacher education). Some attributes of teachers – such as having one or two years of experience – have explained part of the differences in teacher quality, but these factors are a small part of the overall variance in teacher results. This inability to identify specific teacher qualities makes it difficult to regulate or legislate having high quality teachers in classrooms. It

⁷⁰ More details on the class size analyses and debates can be found in Angrist and Lavy (1999), Hoxby (2000b), Krueger (2000), Hanushek (2002a), and Wößmann and West (2006).

⁷¹ For the historical framing of the question, see the exchange between Greenwald, Hedges, and Laine (1996) and Hanushek (1996).

⁷² See, for example, Hanushek (1971, 1992), Murnane (1975), Aaronson, Barrow, and Sander (2003), Rockoff (2004), Rivkin, Hanushek, and Kain (2005), Hanushek, Kain, O'Brien, and Rivkin (2005), Boyd et al. (2005), and Kane, Rockoff, and Staiger (2006).

also contributes to our conclusions below that changes in the institutional structure and incentives of schools are fundamental to improving school outcomes.⁷³

6.3 Within-country Evidence – Developing Countries

The evidence on developed countries, while instructive to them, may or may not be similar in developing countries. Indeed, one concern about the within-country analyses is that they might not provide reliable guidance to other countries. For example, Heyneman and Loxley (1983) argued early in the analysis of school effects that schools may be considerably more important for achievement in developing countries, where family inputs to education may be less.⁷⁴ Further, the levels of spending differ by a factor of 100 in comparing some developed and developing countries so that simple translations of results may be inappropriate.

The research on schools in developing countries has been less extensive than that in developed countries. Moreover, the evidence is frequently weaker because of data or analytical problems with the underlying studies. Nonetheless, as Pritchett (2004) convincingly argues on the basis of ample evidence, just increasing spending within current education systems in developing countries is unlikely to improve students' performance substantially.⁷⁵ Overwhelming evidence shows that expansions on the input side, such as simple physical expansion of the educational facilities and increased spending per student, generally do not seem to lead to substantial increases in children's competencies and learning achievement. The lack of substantial resource effects in general, and class-size effects in particular, has been found across the developing world, including Africa (see, e.g., Kremer (2003); Michaelowa (2001), Latin America (see, e.g., Mizala and Romaguera (2002); Wößmann and Fuchs (2005) and East Asia (see, e.g., Gundlach and Wößmann (2001); Wößmann (2005a)).

Again, it is necessary to understand the character of the results. In particular, the evidence refers most specifically to overall infusions of resources. They do not deny that some investments are productive. A number of studies provide convincing evidence that some minimal levels of key resources are frequently valuable in promoting student learning. For example, the simple availability of

⁷³ These conclusions are also underscored by the evidence on extensive teacher absenteeism in developing countries (see Chaudhury et al. (2006); Banerjee and Duflo (2006)).

⁷⁴ Based on recent international data from TIMSS, however, Hanushek and Luque (2003) question whether this overall conclusion holds.

⁷⁵ See Hanushek (1995), Glewwe (2002), Pritchett (2004), and Glewwe and Kremer (2006) for reviews of research on the determinants of educational quality in developing countries.

textbooks has been shown quite consistently to be important (Heyneman, Jamison, and Montenegro (1984), Harbison and Hanushek (1992)). Similarly, Duflo (2001) shows that school facilities can yield high returns. Nonetheless, while there are suggestive findings of positive resource effects scattered across the literature (see, for example, Lockheed and Hanushek (1988), the main message is still not one of broad, resource-based policy initiatives. The impact of these policies and programs, even if we presume that they could be replicated elsewhere, is limited.

Glewwe and Kremer (2006) provide a review of the various problems associated with this work. Much of the evidence has come from observational evidence where the causal structure is not fully considered. Nonetheless, even among the relatively few studies passing their review criteria, the evidence on either specific or general resources remains mixed – and does not suggest any consistent resource strategies that might be applied.

The general story that emerges from work in developing countries remains very similar to that for developed countries. Simply providing generally increased resources, or resources along the lines commonly suggested such as reducing class sizes or across-the-board increases in teacher salaries, is unlikely to lead to substantial changes in student performance. As in developed countries, it appears extraordinarily important to get the institutional structure right.

6.4 Is There a Minimum Resource Requirement?

The general lack of any systematic relationship between student achievement and resources raises the question of whether or not there is some minimum required level of resources even if impacts are not seen at higher levels of resources. This almost certainly is the case. It is consistent with the few “resource findings” previously indicated about the availability of textbooks, the importance of basic facilities, the impact of having teachers actually show up for class, and similar minimal aspects of a school.

At the same time, there is little evidence upon which to build any serious notion of resource needs. One might suspect that resource effects would appear to be nonlinearly related to achievement such that, below some level, there is a strong relationship between added resources and student outcomes. There is little evidence supporting this kind of relationship.

Part of the difficulty arises from the inference that teacher quality is the most important school element but that teacher quality is unrelated to common measures of salary, education, experience, certification, and the like. This implies that resources, at least as currently spent, are not effective at leading to generally improved teachers. Again, this is not a necessity, but instead merely an observation about how the current institutions translate resources into student results.

7. Schooling Institutions and Educational Quality

The earlier evidence suggested that national economies were very dependent upon the quality of their economic institutions. It is hard to have a strongly growing economy without complementary institutions in the labor and product markets, without openness to trade and investments from the outside, and without effective systems of laws and property rights.

Similarly, we suggest, it is difficult to have a highly functioning education system without a supportive institutional structure. On this matter, however, there are more different opinions and perhaps a wider divergence in outcomes. Part of the reason for the divergent opinions is simply a lack of sufficient experience, analysis, and evidence.

This having been said, we see some clear general policies that are important. Foremost among these, the performance of a system is affected by the incentives that actors face. That is, if the actors in the education process are rewarded (extrinsically or intrinsically) for producing better student performance, and if they are penalized for not producing high performance, this will improve performance. The incentives to produce high-quality education, in turn, are created by the institutions of the education system – all the rules and regulations that (explicitly or implicitly) set rewards and penalties for the people involved in the education process. Therefore, one might expect that institutional features have important impacts on student learning.

We highlight three institutional features that may be part of a successful system for providing students with cognitive skills: choice and competition; decentralization and autonomy of schools; and accountability. Given the nature of this study and the limited evidence, we are only able to present main ideas here; deeper analyses, in particular of issues of design and implementation in specific contexts, have to be left for more encompassing surveys and collections on the matter (e.g., Hanushek (1994); Peterson and West (2003); Betts and Loveless (2005); and, for developing-country contexts, World Bank (2004), ch. 7; Vegas (2005)), as well as for future research.

7.1 Choice and Competition in Developing Countries

Choice and competition in schools were proposed a half century ago by Friedman (1962). The simple idea is that parents, interested in the schooling outcomes of their children, will seek out productive schools. This demand side pressure will result in incentives for each school to produce an effective education system. These incentives will also put pressure on schools to ensure high quality staff in addition to a good curriculum.

In developed countries, a number of privately managed schools particularly in Europe provide alternatives for students. These schools, which also often have a religious affiliation, are part of the natural institutional framework. Unfortunately, little thorough evaluation has been done of the choice possibilities, in large part because there is no obvious comparison group (i.e., choice is instituted for an entire country and there is no example of the no-choice alternative). In a cross-country comparison, students in countries with a larger share of privately managed schools tend to perform better on average (cf. Wößmann (2005d, forthcoming-b)).⁷⁶

In the U.S., there are limited examples of private school choice, ranging from the publicly funded school vouchers in Milwaukee, Cleveland, and Washington, DC, to privately financed voucher alternatives.⁷⁷ The evaluations of these generally show that the choice schools do at least as well as the regular public schools, if not better (see Rouse (1998), Howell and Peterson (2002)).

Outside the US, Bradley and Taylor (2002) and Levačić (2004) find similar positive effects of school competition on the performance of English schools. Sandström and Bergström (2005) and Björklund, Edin, Freriksson, and Krueger (2004) provide evidence on significant positive effects of competition from privately operated schools on the performance of public schools in Sweden. Filer and Munich (2003) show that the introduction of a voucher-type system in the Czech Republic led to the creation of private schools in areas where public schools are doing badly and that the public schools facing private competition improved their performance in obtaining university admission for their graduates.

The evidence from evaluations in developing countries is also generally consistent with the evidence from developed countries in terms of positive effects due to private competition. For example, Colombia ran a program that provided vouchers for the attendance of private schools. The benefits of this program have been found to clearly exceed its cost, which was similar to providing a place in public schools (Angrist et al. (2002); also King, Orazem, and Wohlgemuth (1999)). While evidence on the extensive voucher system in Chile is less uniform, the most elaborate studies tend to suggest that it

⁷⁶ Note that private school management does not mean private school funding; the international evidence suggests that both private school management and *public* school funding are associated with better performance across countries (Wößmann (2005d)).

⁷⁷ The largest U.S. voucher program in the State of Florida provides vouchers for special needs students (Hoxby and Murarka (2006)). While there is considerable satisfaction with this program, there is no evaluation available that is based on explicit outcome measures.

had positive effects on students' performance.⁷⁸ Mizala, Romaguera, and Farren (2002) find that private fee-paying schools are the most technically efficient ones in Chile, followed by private subsidized and public schools. Similarly, Cox and Jimenez (1991) show that private-school students perform better both in Colombia and Tanzania. James, King, and Suryadi (1996) and Bedi and Garg (2000) find that privately managed schools in Indonesia are more efficient and effective than public schools.

The major issue on choice and competition is still the limited experience. The teachers unions and administrator groups invariably dislike the idea of competition – because it puts pressure on them. Thus, not many examples of operational, large-scale attempts at competition have been evaluated. Nonetheless, the benefits of competition are so well documented in other spheres of activity that it is quite inconceivable that more competition would not be beneficial for schools.

Choice and competition are very broad terms that can encompass many different programs. It seems clear that the specific design of any choice program will be important. The importance seems most clear in terms of distributional outcomes, because programs have the potential for leading to segregation of the school population in various ways that might not be desirable (cf. McEwan (2000)). On this, there is simply a need for greater experimentation and experience, given the current levels of uncertainty.

7.2 Evidence on Autonomy of Schools

Several institutional features of a school system can be grouped under the heading of autonomy or decentralization, including local decision making on different matters, fiscal decentralization, and parental involvement. Almost any system of improved incentives for schools depends upon having school personnel in individual schools and districts heavily involved in decision making. It is difficult to compile evidence on the impact of autonomy, because the degree of local decision making is most generally a decision for a country (or state) as a whole, leaving no comparison group within countries. Across countries, students tend to perform better in schools that have autonomy in personnel and day-to-day decisions (Wößmann (2003a, forthcoming-b)), in particular when there is accountability (see below).

⁷⁸ Cf., e.g., Mizala and Romaguera (2000), Sapelli and Vial (2002), Vegas (1999) and Hsieh and Urquiola (2006). A less positive interpretation of vouchers in Chile is found in McEwan and Carnoy (2000) and of vouchers in general in McEwan (2000).

The U.S. states have varying amounts of local autonomy. One systematic form of school autonomy is “charter schools,” which are public schools that are allowed to perform quite autonomously. (Note that these are actually hybrids of choice schools and public school autonomy, because they survive only if sufficient numbers of students are attracted to them and continue to attend them). These schools are relatively new, a fact that complicates evaluation since many are still in the start up phase. Nonetheless, while controversial, the evidence on them suggests that they perform at least as well if not better than the regular public schools after the initial start up phase (Hoxby and Rockoff (2004), Sass (2006), Bifulco and Ladd (2006), Hanushek, Kain, Rivkin, and Branch (forthcoming)). Also, given U.S. residential mobility, individual public school districts compete with each other, and more competition appears to produce better results (Hoxby (2000a)).⁷⁹

There is evidence from within a few developing countries that supports positive effects of decentralization, school autonomy and community involvement. Focusing on parental involvement, Jimenez and Paqueo (1996) find that local financial contributions increased the productivity of public schools in the Philippines relative to central financing. Jimenez and Sawada (2001) show that enhanced community and local involvement improved student learning in El Salvador. Galiani and Schargrotsky (2002) show that decentralization in the Argentine secondary school system improved educational test performance. Vegas (1999) finds that teacher autonomy positively affects student outcomes in Chile when decision making authority is decentralized. Both Gertler, Rubio-Codina, and Patrinos (2006) and Álvarez, Moreno, and Patrinos (2006) find that decentralization of decision making to the local level in Mexico positively affects student outcomes, while the latter also find accountability to be very important in enhancing local decision making. Finally, Skoufias and Shapiro (2006) suggest for Mexico that the combination of increased school resources and local school management can produce small but statistically significant improvements in learning.

Nonetheless, given the available evidence, support for autonomy also strongly rests on a conceptual basis. It is simply hard to imagine a system with strong incentives that does not capitalize on local decision making.

7.3 School Accountability

Many countries around the world have been moving toward increased accountability of local schools for student performance. The United Kingdom has developed an elaborate system of “league tables” designed to give parents full information about the performance of local schools. The United

⁷⁹ See Rothstein (2005) and Hoxby (2005) for further discussion of this evidence.

States has legislated a federal law that all states develop an accountability system that meets certain general guidelines. It also sets into law a series of actions required when a school fails to bring sufficient numbers of students up to proficiency in core subjects.

Evidence on the impacts of these systems has begun to accumulate. While there is some uncertainty given the newness of the overall accountability system (introduced in 2002), the best U.S. evidence indicates that strong accountability systems in fact lead to better student performance (Carnoy and Loeb (2002); Hanushek and Raymond (2005); Jacob (2005)).

One institutional set-up that combines accountability with parental choice are systems which give students in schools that repeatedly do badly on the accountability test a voucher to attend private schools.⁸⁰ In Florida, the threat of becoming subject to private-school choice if failing on the test has been shown to increase school performance particularly for disadvantaged students (West and Peterson (2006); Figlio and Rouse (2006)).

Curriculum-based external exit exams are another means to introduce some form of accountability into the schooling system. They provide performance information which can hold both students and schools accountable. Students in countries with external exit exam systems tend to systematically outperform students in countries without such systems (Bishop (1997); Bishop (2006); Wößmann (2001, 2003a, forthcoming-b)). In the two national education systems where the existence of external exams varies within the country because some regions feature them and others not, Canada and Germany, it has similarly been shown that students perform better in regions with external exams (Bishop (1997); Jürges, Schneider, and Büchel (2005)).

Little evidence is currently available about accountability systems in developing countries. This reflects the generally weak accountability in these countries along with a general lack of systematic measurement and reporting of student achievement. However, Álvarez, Moreno, and Patrinos (2006) find a strong impact of accountability and local decision making in Mexico, and this remains even with consideration of the relative impact of teachers unions.

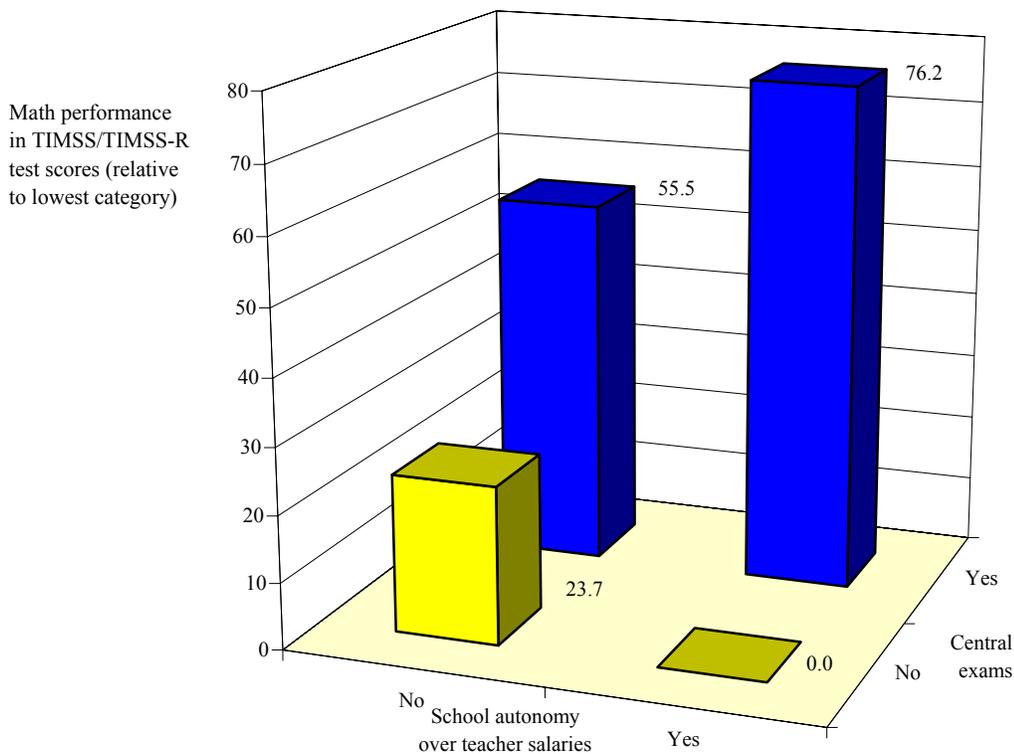
It is difficult to imagine choice or autonomy working well without a good system of student testing and accountability. Thus, the ideas about institutional structure are closely linked together (cf. World Bank (2004)). The international evidence clearly suggests that school autonomy, in particular local autonomy over teacher salaries and course content, is only effective in school systems that have external exams in place (Wößmann (2005c, forthcoming-b); Fuchs and Wößmann (forthcoming)). One

⁸⁰ The legality of this system has been challenged in the Florida courts, so that the future of the program in Florida is in doubt.

example of this evidence is depicted in Figure 7.1, which plots relative student performance under the four conditions resulting from the presence and absence of central exams and school autonomy over teacher salaries, after controlling for dozens of student, family, and school background factors. The figure shows that school autonomy regarding teacher salaries is negatively associated with student performance in systems without central exams. In systems with central exams, student performance is generally higher than in systems without central exams, reflecting the increased accountability. In addition, the effect of school autonomy is turned completely around in systems with central exams: Salary autonomy of schools has positive effects on student performance in central-exam systems. This pattern of results has been found in TIMSS, TIMSS-Repeat, and PISA, and similar cases where external exams turn a negative autonomy effect around into a positive effect have been found for other decision making areas such as school autonomy in determining course content and teacher influence on resource funding.

Figure 7.1: Accountability, Autonomy, and Student Performance across Countries

Performance difference between the four categories relative to the lowest category, after controlling for numerous additional impact factors.



Source: Wößmann (2005c).

Given the importance of high teacher quality, a promising candidate for improvement is the specific form of accountability that aims incentives directly at teachers. While convincing evidence on the effects of performance-related teacher pay is scarce, the more rigorous studies in terms of empirical identification tend to find a positive relationship between financial teacher incentives and student outcomes (cf. the survey in Atkinson et al. (2004) along with the discussions in Vegas (2005) and Vegas and Umansky (2005)). Thus, monetary incentives for teachers based on their students' performance have been shown to improve student learning in Israel immensely (Lavy (2002, 2004)). Similarly, Atkinson et al. (2004) find that the introduction of performance-related pay had a substantial positive impact on student achievement in England.

7.4 Summary of How to Improve the Quality of Education

From the perspective of policies to obtain better outcomes, a simple picture emerges. First, the evidence from extensive research shows clearly that “pure resource policies” – i.e., policies that simply provide more resources within the current incentive structure of schools within most countries – are unlikely to produce substantial systematic gains in student outcomes. There are of course qualifications and caveats. Some schools will use added resources effectively. And, some systems that fail to provide minimal resources such as basic textbooks could, if resources were applied to the key shortage areas, get improvement. Yet, both developed and developing countries demonstrate an improbable but tangible disconnect between simple resource solutions and student achievement.

Second, the evidence about what specifically counts in terms of schools is rather limited. But the central candidate is high teacher quality. Unfortunately, the characteristics of good teachers are not well understood, making regulatory solutions very difficult and supporting the development of rewards for high quality teaching. On this score, most countries – both developed and developing – have incentive systems built into teacher salaries and teacher contractual arrangements that resist significant change and that are not aligned well with teacher quality.

Third, the key to improvement appears to lie in better incentives – incentives that will lead to management keyed to student performance and that will promote strong schools with high quality teachers. Here, three interrelated policies come to the forefront: promoting more competition, so that parental demand will create strong incentives to individual schools; autonomy in local decision making, so that individual schools and their leaders will take actions to promote student achievement; and, an accountability system that identifies good school performance and leads to rewards based on this.

We have sketched three separate components of improved incentives, but it is also important to point out that they form a package. Local autonomy without strong accountability may be worse than

doing nothing. Accountability without choice is likely to be watered down and made impotent by schools that would prefer no accountability. Choice without good information about performance has uncertain outcomes attached to it. In other words, these should not be thought of as isolated policies that can be independently introduced while retaining their advantages.

8. Conclusion

This study was motivated by doubts that have been raised about the role of education and human capital in economic development. These doubts come from a variety of vantage points ranging from whether the research has correctly identified the impact of education to whether other institutional aspects of countries might be more important. They also encompass concerns about whether or not we really know how to change educational outcomes, particularly in developing countries.

Our analysis has produced some remarkably simple but clear conclusions.

1. Educational quality – measured by what people know – has powerful effects on individual earnings, on the distribution of income, and on economic growth.

The accumulated evidence from analyses of economic outcomes is that the quality of education – measured on an outcome basis of cognitive skills – has powerful effects. Much of the earlier discussion has concentrated solely on school attainment, or the quantity of schooling. This focus is unfortunate, because it distorts analysis and the policy discussions.

Individual earnings are systematically related to cognitive skills. The distribution of skills in society appears closely related to the distribution of income. And, perhaps most importantly, economic growth is strongly affected by the skills of workers.

Other factors obviously also enter into growth and may well have stronger effects. For example, having well-functioning economic institutions such as established property rights, open labor and product markets, and participation in international markets have clear importance for economic development and may also magnify the benefits of quality education. Nonetheless, existing evidence suggests that quality of education independently affects economic outcomes even after allowing for these other factors.

Moreover, the existing research provides strong reasons to believe that quality of education is causally related to economic outcomes. To be sure, quality may come from formal schools, from parents, or from other influences on students. But, a more skilled population – almost certainly including both a broadly educated population and a cadre of top performers – results in stronger economic performance for nations.

2. The current situation in developing countries is much worse than generally pictured on the basis just of school enrollment and attainment.

Available measures of school attainment uniformly indicate that developing countries lag dramatically behind developed countries. This fact has driven a variety of efforts to expand schooling in developing countries, including the Education for All initiative. Yet, much of the discussion and much of the policy making has tended to downplay the issues of quality.

International testing indicates that, even among those attaining lower secondary schooling, literacy rates (by international standards) are very low in many developing countries. By reasonable calculations, a range of countries has fewer than 10 percent of its youth currently reaching minimal literacy and numeracy levels, even when school attainment data look considerably better.

Because of the previous findings – that knowledge rather than just time in school is what counts – policies must pay more attention to quality of schools.

For developing countries, the sporadic or nonexistent assessment of student knowledge is an especially important issue – correcting this shortcoming should have the highest priority. It is impossible to develop effective policies without having a good understanding of which work and which do not. Currently available measures of program “quality” frequently rely upon various input measures that unfortunately are not systematically related to student learning. Moreover, the existing international tests – such as the PISA tests of the OECD – may not be best suited to provide accurate assessments of student performance in developing countries. The evolving capacity for adaptive testing that can adjust test content to the student’s ability level seems particularly important in the developing country context. Adaptive testing offers the possibility of meaningful within-country variation in scores along with the ability to link overall performance with global standards.

3. Just providing added resources to schools is unlikely to be successful; improving the quality of schools will take structural changes in institutions.

Even though attempts to improve quality have frustrated many policy makers around the world, we now have extensive research that reaches some clear conclusions. Research has delved deeply into the impact of adding resources within the current institutional structure (of both developed and developing countries). The overall finding is that simple resource policies – reducing class sizes, increasing teacher salaries, spending more on schools, and so forth – have little consistent impact on student performance when the overall institutional structure is not changed.

This does not say that spending never has an impact. In fact, there is reason to believe that some basic resources in the least developed schools, such as providing textbooks for all students, have a reliable impact. But, these situations have been documented just in the poorest schools and, even there, just in limited areas. There is also evidence that some schools use added resources better than others, although research does not characterize these different situations well enough to build them into overall resource policies.

There is mounting evidence that the quality of teachers is the key ingredient to student performance. Unfortunately, the characteristics of good teachers are not described well, making it impossible to legislate or regulate good teachers. This motivates our institutional discussions below.

The evidence also does not say that resources can never have an impact. In fact, the kinds of institutional changes that we identify are designed precisely to ensure that added resources are effective.

The largest problem that we see in current school policy is the lack of incentives for improved student performance. Neither students nor school personnel in most countries of the world tend to be significantly rewarded for high performance. Without such incentives, it is no surprise to find that added resources do not consistently go toward improvement of student outcomes.

Three sets of policies head our list for items needed to improve the overall incentives in schools: strong accountability systems that accurately measure student performance; local autonomy that allows schools to make appropriate educational choices; and choice and competition in schools so that parents can enter into determining the incentives that schools face.

As noted, many if not most developing countries currently lack performance measurement that would allow them to know which policies were working and which were not or where performance was most in need of change. Lack of measurement of student outcomes clearly makes any system of direct rewards for success difficult if not impossible. An early step in any reform program should be instituting reliable school accountability systems.

If schools are held responsible for results, they must have the ability to make decisions that will lead to better outcomes. Highly centralized regulatory systems simply cannot work effectively without broad knowledge of what programs are effective in different situations. Such knowledge is currently lacking, leading to centralized decision making that does not produce strong results. (The lack of results from centralized decision making can be readily seen in the data provided, particularly for developing countries.)

Finally, in terms of overall changes, more choice and competition among schools leads parents to be directly involved in evaluating the performance of schools. While experiments in various choice

plans have been limited, the existing evidence suggests that they do tend to lead to better student outcomes. The best way to introduce choice in rural systems within resource-constrained developing countries is not fully known now, but it appears clear that the existing system is not working.

Uncertainty about the best design of incentive programs for schools is most acute in developing countries, largely due to lack of relevant experiences. For this reason, it is especially important to implement a program of experimentation and evaluation – a key missing aspect of policy making in most developing countries. Education policy must be viewed from an evolutionary perspective, where on-going evaluation permits discarding policies that are ineffective while expanding those that are productive.

When asking how education policies in developing countries can create the competencies and learning achievements required for their citizens to prosper in the future, the binding constraint seems to be institutional reforms, not resource expansions within the current institutional systems. For educational investments to translate into student learning, all the people involved in the education process have to face the right incentives that make them act in ways that advance student performance.

References

- Aaronson, Daniel, Lisa Barrow, and William Sander. 2003. "Teachers and Student Achievement in the Chicago Public High Schools." WP 2002-28, Federal Reserve Bank of Chicago June.
- Acemoglu, Daron, and Joshua D. Angrist. 2000. "How large are the social returns to education? Evidence from compulsory schooling laws." In *NBER Macroeconomics Annual 2000*, edited by Ben S. Bernanke and Kenneth Rogoff. Cambridge, MA: MIT Press:9-59.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The Colonial Origins of Comparative Development: An Empirical Investigation." *American Economic Review* 91, no. 5 (December):1369-1401.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2002. "Reversal of fortune: Geography and institutions in the making of the modern world income distribution." *Quarterly Journal of Economics* 117, no. 4 (November):1231-1294.
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2005. "Institutions as a Fundamental Cause of Long-run Growth." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf. Amsterdam: North Holland:385-472.
- Aghion, Philippe, and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- Alderman, Harold, Jere R. Behrman, David R. Ross, and Richard Sabot. 1996. "The returns to endogenous human capital in Pakistan's rural wage labor market." *Oxford Bulletin of Economics and Statistics* 58, no.:29-55.
- Altonji, Joseph G., and Charles R. Pierret. 2001. "Employer learning and statistical discrimination." *Quarterly Journal of Economics* 116, no. 1 (February):313-350.
- Álvarez, Jesús, Vicente García Moreno, and Harry A. Patrinos. 2006. "Institutional effects as determinants of learning outcomes: Exploring state variations in Mexico." (mimeo), Washington, World Bank March.
- Angrist, Joshua D., Eric Bettinger, Erik Bloom, Elizabeth N. King, and Michael Kremer. 2002. "Vouchers for private schooling in Columbia: Evidence from a randomized natural experiment." *American Economic Review* 92, no. 5 (December):1535-1558.
- Angrist, Joshua D., and Victor Lavy. 1997. "The effect of a change in language of instruction on the returns to schooling in Morocco." *Journal of Labor Economics* 15, no.:S48-S76.
- Angrist, Joshua D., and Victor Lavy. 1999. "Using Maimonides' rule to estimate the effect of class size on scholastic achievement." *Quarterly Journal of Economics* 114, no. 2 (May):533-575.
- Atkinson, Adele, Simon Burgess, Bronwyn Croxson, Paul Gregg, Carol Propper, Helen Slater, and Deborah Wilson. 2004. "Evaluating the Impact of Performance-related Pay for Teachers in England." CMPO Working Paper 04/113, University of Bristol December 2004.

- Azariadis, Costas, and Allan Drazen. 1990. "Threshold Externalities in Economic Development." *Quarterly Journal of Economics* 105, no. 2:501-526.
- Banerjee, Abhijit, and Esther Duflo. 2006. "Addressing Absence." *Journal of Economic Perspective* 20, no. 1:117-132.
- Barro, Robert J. 1991. "Economic growth in a cross section of countries." *Quarterly Journal of Economics* 106, no. 2 (May):407-443.
- Barro, Robert J. 1997. *Determinants of Economic Growth: A Cross-Country Empirical Study*. Cambridge, MA: MIT Press.
- Barro, Robert J. 2001. "Human capital and growth." *American Economic Review* 91, no. 2 (May):12-17.
- Barro, Robert J., and Jong-wha Lee. 1993. "International comparisons of educational attainment." *Journal of Monetary Economics* 32, no. 3 (December):363-394.
- Barro, Robert J., and Jong-wha Lee. 2001. "International data on educational attainment: Updates and implications." *Oxford Economic Papers* 53, no. 3 (July):541-563.
- Barro, Robert J., and Xavier Sala-i-Martin. 2004. *Economic Growth*. Second ed. Cambridge, MA: The MIT Press.
- Bedi, Arjun S., and Ashish Garg. 2000. "The effectiveness of private versus public schools: The case of Indonesia." *Journal of Development Economics* 61, no. 2:463-494.
- Behrman, Jere R., Lori G. Kletzer, Michael S. McPherson, and Morton Owen Schapiro. 1998. "The microeconomics of college choice, careers, and wages: Measuring the impact of higher education." *Annals of the American Academy of Political and Social Science* 559, no. (September):12-23.
- Behrman, Jere R., David Ross, and Richard Sabot. forthcoming. "Improving the quality versus increasing the quantity of schooling: Estimates of rates of return from rural Pakistan." *Journal of Development Economics*.
- Benhabib, Jess, and Mark M. Spiegel. 1994. "The role of human capital in economic development: Evidence from aggregate cross-country data." *Journal of Monetary Economics* 34, no. 2:143-174.
- Benhabib, Jess, and Mark M. Spiegel. 2005. "Human capital and technology diffusion." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf. Amsterdam: North Holland:935-966.
- Betts, Julian R., and Tom Loveless, eds. 2005. *Getting choice right: Ensuring equity and efficiency in education policy*. Washington: Brookings.
- Bifulco, Robert, and Helen F. Ladd. 2006. "The impacts of charter schools on student achievement: Evidence from North Carolina." *Education Finance and Policy* 1, no. 1 (winter):50-90.

- Bils, Mark, and Peter J. Klenow. 2000. "Does schooling cause growth?" *American Economic Review* 90, no. 5 (December):1160-1183.
- Bishop, John H. 1989. "Is the test score decline responsible for the productivity growth decline?" *American Economic Review* 79, no. 1:178-197.
- Bishop, John H. 1991. "Achievement, test scores, and relative wages." In *Workers and their wages*, edited by Marvin H. Koster. Washington, DC: The AEI Press:146-186.
- Bishop, John H. 1997. "The Effect of National Standards and Curriculum-Based Examinations on Achievement." *American Economic Review* 87, no. 2:260-264.
- Bishop, John H. 2006. "Drinking from the fountain of knowledge: Student incentive to study and learn -- Externalities, information problems, and peer pressure." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:909-944.
- Björklund, Anders, Per-Anders Edin, Peter Freriksson, and Alan B. Krueger. 2004. "Education, Equality and Efficiency: An Analysis of Swedish School Reforms during the 1990s." IFAU Report 2004:1, Uppsala, Institute for Labour Market Policy Evaluation
- Blackburn, McKinley L., and David Neumark. 1993. "Omitted-ability bias and the increase in the return to schooling." *Journal of Labor Economics* 11, no. 3 (July):521-544.
- Blackburn, McKinley L., and David Neumark. 1995. "Are OLS estimates of the return to schooling biased downward? Another look." *Review of Economics and Statistics* 77, no. 2 (May):217-230.
- Boissiere, Maurice X., John B. Knight, and Richard H. Sabot. 1985. "Earnings, schooling, ability, and cognitive skills." *American Economic Review* 75, no. 5:1016-1030.
- Bosworth, Barry P., and Susan M. Collins. 2003. "The empirics of growth: An update." *Brookings Papers on Economic Activity* 2003, no. 2:113-206.
- Boyd, Don, Pam Grossman, Hamilton Lankford, Susanna Loeb, and James Wyckoff. 2005. "How Changes in Entry Requirements Alter the Teacher Workforce and Affect Student Achievement." Working Paper 11844, Cambridge, MA, National Bureau of Economic Research December.
- Bradley, Steve, and Jim Taylor. 2002. "The Effect of the Quasi-Market on the Efficiency-Equity Trade-off in the Secondary School Sector." *Bulletin of Economic Research* 54, no. 3:295-314.
- Burtless, Gary, ed. 1996. *Does Money Matter?: The Effect of School Resources on Student Achievement and Adult Success*. Washington, D.C.: The Brookings Institution.
- Card, David. 1999. "Causal effect of education on earnings." In *Handbook of labor economics*, edited by Orley Ashenfelter and David Card. Amsterdam: North-Holland:1801-1863.
- Carnoy, Martin, and Susanna Loeb. 2002. "Does external accountability affect student outcomes? A cross-state analysis." *Educational Evaluation and Policy Analysis* 24, no. 4 (Winter):305-331.

- Castelló, Amparo, and Rafael Doménech. 2002. "Human Capital Inequality and Economic Growth: Some New Evidence." *Economic Journal* 112, no. 478:C187-C200.
- Chaudhury, Nazmul, Jeffrey Hammer, Michael Kremer, Karthik Muralidharan, and F. Halsey Rogers. 2006. "Missing in Action: Teacher and Health Worker Absence in Developing Countries." *Journal of Economic Perspective* 20, no. 1:91-116.
- Ciccone, Antonio, and Elias Papaioannou. 2005. "Human Capital, the Structure of Production, and Growth." Barcelona, Universitat Pompeu Fabra
- Ciccone, Antonio, and Giovanni Peri. 2006. "Identifying human capital externalities: Theory with Applications." *Review of Economic Studies* 73, no. 2 (April):381-412.
- Cohen, Daniel, and Marcelo Soto. 2001. "Growth and human capital: Good data, good results." Technical Paper 179, OECD Development Centre, September.
- Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.
- Coulombe, Serge, and Jean-François Tremblay. 2006. "Literacy and Growth. ." *Topics in Macroeconomics* 6, no. 2:Article 4.
- Coulombe, Serge, Jean-François Tremblay, and Sylvie Marchand. 2004. *Literacy scores, human capital and growth across fourteen OECD countries*. Ottawa: Statistics Canada.
- Cox, Donald, and Emmanuel Jimenez. 1991. "The Relative Effectiveness of Private and Public Schools." *Journal of Development Economics* 34, no. 1:99-121.
- Currie, Janet, and Enrico Moretti. 2003. "Mother's Education and the Intergenerational Transmission of Human Capital: Evidence from College Openings." *Quarterly Journal of Economics* 118, no. 4:1495-1532.
- De Gregorio, José, and Jong-Wha Lee. 2002. "Education and Income Inequality: New Evidence from Cross-Country Data." *Review of Income and Wealth* 48, no. 3:395-416.
- de la Fuente, Angel, and Rafael Doménech. 2001. "Schooling Data, Technological Diffusion, and the Neoclassical Model." *American Economic Review* 91, no. 2:323-327.
- de la Fuente, Angel, and Rafael Doménech. 2006. "Human capital in growth regressions: How much different does data quality make?" *Journal of the European Economic Association* 4, no. 1 (March):1-36.
- Dee, Thomas S. 2004. "Are there civic returns to education?" *Journal of Public Economics* 88, no. 9-10 (August):1697-1720.
- Duflo, Esther. 2001. "Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from an Unusual Policy Experiment." *American Economic Review* 91, no. 4 (September):795-813.

- Dugan, Dennis J. 1976. "Scholastic achievement: its determinants and effects in the education industry." In *Education as an industry*, edited by Joseph T. Froomkin, Dean T. Jamison and Roy Radner. Cambridge, MA: Ballinger:53-83.
- Easterly, William. 2002. *The elusive quest for growth: An economists' adventures and misadventures in the tropics*. Cambridge, MA: The MIT Press.
- Figlio, David N., and Cecilia Elena Rouse. 2006. "Do accountability and voucher threats improve low-performing schools?" *Journal of Public Economics* 90, no. 1-2 (January):239-255.
- Filer, Randall K., and Daniel Münich. 2003. "Responses of Private and Public Schools to Voucher Funding." In *Annual Meeting of the American Economic Association* Washington, D. C.
- Filmer, Deon. 2006. *Educational Attainment and Enrollment around the World. Development Research Group*. The World Bank 2006 [cited 2006]. Available from econ.worldbank.org/projects/edattain.
- Finnie, Ross, and Ronald Meng. 2002. "Minorities, cognitive skills, and incomes of Canadians." *Canadian Public Policy* 28, no.:257-273.
- Friedman, Milton. 1962. *Capitalism and freedom*. Chicago: University of Chicago Press.
- Fuchs, Thomas, and Ludger Wößmann. forthcoming. "What Accounts for International Differences in Student Performance? A Re-examination using PISA Data." *Empirical Economics*.
- Galiani, Sebastian, and Ernesto Schargrotsky. 2002. "Evaluating the Impact of School Decentralization on Educational Quality." *Economia* 2, no. 2:275-314.
- Gemmell, Norman. 1996. "Evaluating the Impacts of Human Capital Stocks and Accumulation on Economic Growth: Some New Evidence." *Oxford Bulletin of Economics and Statistics* 58, no. 1:9-28.
- Gertler, Paul J., Marta Rubio-Codina, and Harry A. Patrinos. 2006. "Empowering Parents to Improve Education: Evidence from Rural Mexico." World Bank Policy Research Working Paper 3935, Washington, World Bank June.
- Glewwe, Paul. 1996. "The relevance of standard estimates of rates of return to schooling for educational policy: A critical assessment." *Journal of Development Economics* 51, no.:267-290.
- Glewwe, Paul. 2002. "Schools and skills in developing countries: Education policies and socioeconomic outcomes." *Journal of Economic Literature* 40, no. 2 (June):436-482.
- Glewwe, Paul, and Michael Kremer. 2006. "Schools, teachers, and educational outcomes in developing countries." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:943-1017.
- Green, David A., and W. Craig Riddell. 2003. "Literacy and earnings: An investigation of the interaction of cognitive and unobserved skills in earnings generation." *Labour Economics* 10, no.:165-184.

- Greenwald, Rob, Larry V. Hedges, and Richard D. Laine. 1996. "The effect of school resources on student achievement." *Review of Educational Research* 66, no. 3 (Fall):361-396.
- Grogger, Jeffrey T., and Eric Eide. 1993. "Changes in college skills and the rise in the college wage premium." *Journal of Human Resources* 30, no. 2 (Spring):280-310.
- Gundlach, Erich, Desmond Rudman, and Ludger Wößmann. 2002. "Second Thoughts on Development Accounting." *Applied Economics* 34, no. 11:1359-1369.
- Gundlach, Erich, and Ludger Wößmann. 2001. "The Fading Productivity of Schooling in East Asia." *Journal of Asian Economics* 12, no. 3:401-417.
- Gundlach, Erich, Ludger Wößmann, and Jens Gmelin. 2001. "The decline of schooling productivity in OECD countries." *Economic Journal* 111, no. 471 (May):C135-C147.
- Hall, Robert E., and Charles I. Jones. 1999. "Why do Some Countries Produce So Much More Output per Worker than Others?" *Quarterly Journal of Economics* 114, no. 1:83-116.
- Hanushek, Eric A. 1971. "Teacher characteristics and gains in student achievement: Estimation using micro data." *American Economic Review* 60, no. 2 (May):280-288.
- Hanushek, Eric A. 1992. "The trade-off between child quantity and quality." *Journal of Political Economy* 100, no. 1 (February):84-117.
- Hanushek, Eric A. 1994. *Making Schools Work: Improving Performance and Controlling Costs*. Washington, D.C.: The Brookings Institution.
- Hanushek, Eric A. 1995. "Interpreting recent research on schooling in developing countries." *World Bank Research Observer* 10, no. 2 (August):227-246.
- Hanushek, Eric A. 1996. "A more complete picture of school resource policies." *Review of Educational Research* 66, no. 3 (Fall):397-409.
- Hanushek, Eric A. 2002a. "Evidence, politics, and the class size debate." In *The class size debate*, edited by Lawrence Mishel and Richard Rothstein. Washington, DC: Economic Policy Institute:37-65.
- Hanushek, Eric A. 2002b. "Publicly provided education." In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein. Amsterdam: Elsevier:2045-2141.
- Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *Economic Journal* 113, no. 485 (February):F64-F98.
- Hanushek, Eric A. 2005. "The Economics of School Quality." *German Economic Review* 6, no. 3 (August):269-286.
- Hanushek, Eric A., John F. Kain, Daniel M. O'Brien, and Steve G. Rivkin. 2005. "The market for teacher quality." Working Paper No. 11154, National Bureau of Economic Research February.

- Hanushek, Eric A., John F. Kain, Steve G. Rivkin, and Gregory F. Branch. forthcoming. "Charter school quality and parental decision making with school choice." *Journal of Public Economics*.
- Hanushek, Eric A., and Dongwook Kim. 1995. "Schooling, labor force quality, and economic growth." NBER, Working Paper 5399, Cambridge, MA, National Bureau of Economic Research December.
- Hanushek, Eric A., and Dennis D. Kimko. 2000. "Schooling, labor force quality, and the growth of nations." *American Economic Review* 90, no. 5 (December):1184-1208.
- Hanushek, Eric A., Victor Lavy, and Kohtaro Hitomi. 2006. "Do students care about school quality? Determinants of dropout behavior in developing countries." Working Paper 12737, National Bureau of Economic Research, Cambridge, MA, December.
- Hanushek, Eric A., and Javier A. Luque. 2003. "Efficiency and equity in schools around the world." *Economics of Education Review* 22, no. 5 (August):481-502.
- Hanushek, Eric A., and Richard R. Pace. 1995. "Who chooses to teach (and why)?" *Economics of Education Review* 14, no. 2 (June):101-117.
- Hanushek, Eric A., and Margaret E. Raymond. 2005. "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24, no. 2 (Spring):297-327.
- Hanushek, Eric A., Steven G. Rivkin, and Lori L. Taylor. 1996. "Aggregation and the estimated effects of school resources." *Review of Economics and Statistics* 78, no. 4 (November):611-627.
- Hanushek, Eric A., and Ludger Wößmann. in process. *The human capital of nations*.
- Hanushek, Eric A., and Lei Zhang. 2006. "Quality Consistent Estimates of International Returns to Skill." National Bureau of Economic Research, WP12664, Cambridge, MA, NBER November.
- Harbison, Ralph W., and Eric A. Hanushek. 1992. *Educational performance of the poor: lessons from rural northeast Brazil*. New York: Oxford University Press.
- Harmon, Colm, Hessel Oosterbeek, and Ian Walker. 2003. "The returns to education: Microeconomics." *Journal of Economic Surveys* 17, no. 2:115-155.
- Heckman, James J., Lance J. Lochner, and Petra E. Todd. 2006. "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:307-458.
- Heckman, James J., and Edward Vytlačil. 2001. "Identifying the role of cognitive ability in explaining the level of and change in the return to schooling." *Review of Economics and Statistics* 83, no. 1 (February):1-12.
- Heston, Alan, Robert Summers, and Bettina Aten. 2002. "Penn World Table Version 6.1." Center for International Comparisons at the University of Pennsylvania (CICUP), Philadelphia, University of Pennsylvania

- Heyneman, Stephen P., Dean T. Jamison, and Xenia Montenegro. 1984. "Textbooks in the Philippines: Evaluation of the pedagogical impact of a nationwide investment." *Educational Evaluation and Policy Analysis* 6, no. (Summer):139-150.
- Heyneman, Stephen P., and William Loxley. 1983. "The effect of primary school quality on academic achievement across twenty-nine high and low income countries." *American Journal of Sociology* 88, no. (May):1162-1194.
- Howell, William G., and Paul E. Peterson. 2002. *The education gap: Vouchers and urban schools*. Washington, DC: Brookings.
- Hoxby, Caroline Minter. 2000a. "Does competition among public schools benefit students and taxpayers?" *American Economic Review* 90, no. 5 (December):1209-1238.
- Hoxby, Caroline Minter. 2000b. "The effects of class size on student achievement: New evidence from population variation." *Quarterly Journal of Economics* 115, no. 3 (November):1239-1285.
- Hoxby, Caroline Minter. 2005. "Competition Among Public Schools: A Reply to Rothstein (2004)." NBER Working Paper 11216, Cambridge, MA, National Bureau of Economic Research
- Hoxby, Caroline Minter, and Sonali Murarka. 2006. "Comprehensive yet simple: Florida's tapestry of school choice programs." In *Reforming Education in Florida: A Study Prepared by the Koret Task Force on K-12 Education*, edited by Paul E. Peterson. Palo Alto: Hoover Institution:167-211.
- Hoxby, Caroline Minter, and Jonah E. Rockoff. 2004. "The impact of charter schools on student achievement." (mimeo), November.
- Hsieh, Chang-Tai, and Miguel Urquiola. 2006. "The effects of generalized school choice on achievement and stratification: Evidence from Chile's voucher program." *Journal of Public Economics* 90, no. 8-9 (September):1477-1503.
- Jacob, Brian A. 2005. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89, no. 5-6 (June):761-796.
- James, Estelle, Elizabeth M. King, and Ace Suryadi. 1996. "Finance, Management, and Costs of Public and Private Schools in Indonesia." *Economics of Education Review* 15, no. 4:387-398.
- Jamison, Eliot A., Dean T. Jamison, and Eric A. Hanushek. forthcoming. "The effects of education quality on mortality decline and income growth." *Economics of Education Review*.
- Jimenez, Emmanuel, and Vicente Paqueo. 1996. "Do Local Contributions Affect the Efficiency of Public Primary Schools?" *Economics of Education Review* 15, no. 4:377-386.
- Jimenez, Emmanuel, and Yasuyuki Sawada. 2001. "Public for private: The relationship between public and private school enrollment in the Philippines." *Economics of Education Review* 20, no. 4 (August):389-399.

- Jolliffe, Dean. 1998. "Skills, schooling, and household income in Ghana." *World Bank Economic Review* 12, no.:81-104.
- Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce. 1993. "Wage inequality and the rise in returns to skill." *Journal of Political Economy* 101, no. 3 (June):410-442.
- Jürges, Hendrik, Kerstin Schneider, and Felix Büchel. 2005. "The Effect of Central Exit Examinations on Student Achievement: Quasi-Experimental Evidence from TIMSS Germany." *Journal of the European Economic Association* 3, no. 5:1134-1155.
- Kane, Thomas J., Jonah E. Rockoff, and Douglas O. Staiger. 2006. "What Does Certification Tell Us About Teacher Effectiveness? Evidence from New York City." Working Paper No. 12155, National Bureau of Economic Research April.
- Katz, Lawrence F., and Kevin M. Murphy. 1992. "Changes in relative wages, 1963-1987: Supply and demand factors." *Quarterly Journal of Economics* 107, no. 1 (February):35-78.
- King, Elizabeth M., Peter F. Orazem, and Darin Wohlgemuth. 1999. "Central Mandates and Local Incentives: The Colombia Education Voucher Program." *World Bank Economic Review* 13, no. 3:467-491.
- Klenow, Peter J., and Andres Rodriguez-Clare. 1997. "The neoclassical revival in growth economics: Has it gone too far?" In *NBER Macroeconomics Annual 1997*, edited by Ben S. Bernanke and Julio J. Rotemberg. Cambridge, MA: MIT Press:83-103.
- Knack, Stephen, and Philip Keefer. 1995. "Institutions and Economic Performance: Cross-Country Tests Using Alternative Measures." *Economics and Politics* 7, no. 3:207-227.
- Knight, John B., and Richard H. Sabot. 1990. *Education, Productivity, and Inequality: The East African Natural Experiment*. 1990: Oxford University Press.
- Knighton, Tamara, and Patrick Bussière. 2006. "Educational Outcomes at Age 19 Associated with Reading Ability at Age 15." Culture, Tourism and the Centre for Education Statistics Research papers, Ottawa, Statistics Canada June.
- Kremer, Michael. 2003. "Randomized evaluations of educational programs in developing countries: Some lessons." *American Economic Review* 93, no. 2 (May):102-104.
- Krueger, Alan B. 2000. "An economist's view of class size research." The class size debate, Washington, DC, Economic Policy Institute
- Krueger, Alan B., and Mikael Lindahl. 2001. "Education for growth: Why and for whom?" *Journal of Economic Literature* 39, no. 4 (December):1101-1136.
- Krueger, Anne O. 1974. "The political economy of the rent seeking society." *American Economic Review* 64, no. 3 (June):291-303.

- Lange, Fabian, and Robert Topel. 2006. "The social value of education and human capital." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:459-509.
- Lau, Lawrence J., Dean T. Jamison, and Frederic F. Louat. 1991. "Education and Productivity in Developing Countries: An Aggregate Production Function Approach." World Bank PRE Working Paper Series 612, Washington, D.C., World Bank
- Lavy, Victor. 2002. "Evaluating the effect of teachers' group performance incentives on pupil achievement." *Journal of Political Economy* 110, no. 6 (December):1286-1317.
- Lavy, Victor. 2004. "Performance Pay and Teachers' Effort, Productivity and Grading Ethics." NBER Working Paper 10622, Cambridge, MA, National Bureau of Economic Research
- Lazear, Edward P. 2003. "Teacher incentives." *Swedish Economic Policy Review* 10, no. 3:179-214.
- Lee, Doo Won, and Tong Hun Lee. 1995. "Human Capital and Economic Growth: Tests Based on the International Evaluation of Educational Achievement." *Economics Letters* 47, no. 2:219-225.
- Levačić, Rosalind. 2004. "Competition and the performance of English secondary schools: Further evidence." *Education Economics* 12, no. 2:177-193.
- Levine, Ross, and David Renelt. 1992. "A sensitivity analysis of cross-country growth regressions." *American Economic Review* 82, no. 4 (September):942-963.
- Levy, Frank, and Richard J. Murnane. 1992. "U.S. earnings levels and earnings inequality: A review of recent trends and proposed explanations." *Journal of Economic Literature* 30, no. 3 (September):1333-1381.
- Lochner, Lance, and Enrico Moretti. 2004. "The Effect of Education on Crime: Evidence from Prison Inmates, Arrests, and Self-Reports." *American Economic Review* 94, no. 1:155-189.
- Lockheed, Marlaine E., and Eric A. Hanushek. 1988. "Improving educational efficiency in developing countries: What do we know?" *Compare* 18, no. 1:21-38.
- Lucas, Robert E. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22, no. (July):3-42.
- Mankiw, N. Gregory, David Romer, and David Weil. 1992. "A contribution to the empirics of economic growth." *Quarterly Journal of Economics* 107, no. 2 (May):407-437.
- Manski, Charles F., and David A. Wise. 1983. *College choice in America*. Cambridge: Harvard University Press.
- McArthur, John W., and Jeffrey D. Sachs. 2001. "Institutions and Geography: Comment on Acemoglu, Johnson and Robinson (2000)." NBER Working Paper 8114, Cambridge, MA, National Bureau of Economic Research

- McEwan, Patrick J. 2000. "The potential impact of large-scale voucher programs." *Review of Educational Research* 70, no. 2 (Summer):103-149.
- McEwan, Patrick J, and Martin Carnoy. 2000. "The effectiveness and efficiency of private schools in Chile's voucher system." *Educational Evaluation and Policy Analysis* 22, no. 3 (Fall):213-240.
- McIntosh, Steven, and Anna Vignoles. 2001. "Measuring and assessing the impact of basic skills on labor market outcomes." *Oxford Economic Papers* 53, no.:453-481.
- Michaelowa, Katharina. 2001. "Primary Education Quality in Francophone Sub-Saharan Africa: Determinants of Learning Achievement and Efficiency Considerations." *World Development* 29, no. 10:1699-1695.
- Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos. 2004. "Does education improve citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics* 88, no. 9-10 (August):1667-1695.
- Mincer, Jacob. 1970. "The distribution of labor incomes: a survey with special reference to the human capital approach." *Journal of Economic Literature* 8, no. 1 (March):1-26.
- Mincer, Jacob. 1974. *Schooling Experience and Earnings*. New York: NBER.
- Mizala, Alejandra, and Pilar Romaguera. 2000. "School performance and choice: the Chilean experience." *Journal of Human Resources* 35, no. 2 (Spring):392-417.
- Mizala, Alejandra, and Pilar Romaguera. 2002. "Equity and Educational Performance." *Economia* 2, no. 2:219-273.
- Mizala, Alejandra, Pilar Romaguera, and Darío Farren. 2002. "The Technical Efficiency of Schools in Chile." *Applied Economics* 34, no. 12:1533-1552.
- Moll, Peter G. 1998. "Primary schooling, cognitive skills, and wage in South Africa." *Economica* 65, no.:263-284.
- Moretti, Enrico. 2004. "Workers' Education, Spillovers, and Productivity: Evidence from Plant-Level Production Functions." *American Economic Review* 94, no. 3:656-690.
- Mulligan, Casey B. 1999. "Galton versus the human capital approach to inheritance." *Journal of Political Economy* 107, no. 6, pt. 2 (December):S184-S224.
- Murnane, Richard J. 1975. *Impact of school resources on the learning of inner city children*. Cambridge, MA: Ballinger.
- Murnane, Richard J., John B. Willett, M. Jay Braatz, and Yves Duhaldeborde. 2001. "Do different dimensions of male high school students' skills predict labor market success a decade later? Evidence from the NLSY." *Economics of Education Review* 20, no. 4 (August):311-320.

- Murnane, Richard J., John B. Willett, Yves Duhaldeborde, and John H. Tyler. 2000. "How important are the cognitive skills of teenagers in predicting subsequent earnings?" *Journal of Policy Analysis and Management* 19, no. 4 (Fall):547-568.
- Murnane, Richard J., John B. Willett, and Frank Levy. 1995. "The growing importance of cognitive skills in wage determination." *Review of Economics and Statistics* 77, no. 2 (May):251-266.
- Murphy, Kevin M., Andrei Shleifer, and Robert W. Vishny. 1991. "The Allocation of Talent: Implications for Growth." *Quarterly Journal of Economics* 106, no. 2 (May):503-530.
- Neal, Derek A., and William R. Johnson. 1996. "The role of pre-market factors in black-white differences." *Journal of Political Economy* 104, no. 5 (October):869-895.
- Nehru, Vikram, Eric Swanson, and Ashutosh Dubey. 1995. "A New Database on Human Capital Stock in Developing and Industrial Countries: Sources, Methodology, and Results." *Journal of Development Economics* 46, no. 2:379-401.
- Neidorf, Teresa S., Marilyn Binkley, Kim Gattis, and David Nohara. 2006. *Comparing Mathematics Content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Program for International Student Assessment (PISA) 2003 Assessments*. Washington: National Center for Education Statistics (May).
- Nelson, Richard R., and Edmund Phelps. 1966. "Investment in humans, technology diffusion and economic growth." *American Economic Review* 56, no. 2 (May):69-75.
- Nickell, Stephen. 2004. "Poverty and worklessness in Britain." *Economic Journal* 114, no. (March):C1-C25.
- North, Douglass C. 1990. *Institutions, Institutional Change and Economic Performance*. United States: Cambridge University Press.
- O'Neill, June. 1990. "The role of human capital in earnings differences between black and white men." *Journal of Economic Perspectives* 4, no. 4 (Fall):25-46.
- Organisation for Economic Co-operation and Development. 2003. *Education at a Glance: OECD Indicators 2003*. Paris: OECD.
- Organisation for Economic Co-operation and Development. 2004. *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD.
- Parente, Stephen L., and Edward C. Prescott. 1994. "Barriers to technology adoption and development." *Journal of Political Economy* 102, no. 2 (April):298-321.
- Parente, Stephen L., and Edward C. Prescott. 1999. "Monopoly rights: A barrier to riches." *American Economic Review* 89, no. 5 (December):1216-1233.
- Peterson, Paul E., and Martin R. West, eds. 2003. *No child left behind? The politics and practice of accountability*. Washington, DC: Brookings.

- Pritchett, Lant. 2001. "Where has all the education gone?" *World Bank Economic Review* 15, no. 3:367-391.
- Pritchett, Lant. 2004. "Access to Education." In *Global Crises, Global Solutions*, edited by Björn Lomborg. Cambridge: Cambridge University Press:175-234.
- Pritchett, Lant. 2006. "Does learning to add up add up? The returns to schooling in aggregate data." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland:635-695.
- Psacharopoulos, George. 1994. "Returns to investment in education: A global update." *World Development* 22, no.:1325-1344.
- Psacharopoulos, George, and Harry A. Patrinos. 2004. "Returns to investment in education: a further update." *Education Economics* 12, no. 2 (August):111-134.
- Rivkin, Steven G. 1995. "Black/white differences in schooling and employment." *Journal of Human Resources* 30, no. 4 (Fall):826-852.
- Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73, no. 2 (March):417-458.
- Rockoff, Jonah E. 2004. "The Impact of Individual Teachers on Student Achievement: Evidence from Panel Data." *American Economic Review* 94, no. 2 (May):247-252.
- Romer, Paul. 1990a. "Endogenous technological change." *Journal of Political Economy* 99, no. 5,pt. II:S71-S102.
- Romer, Paul. 1990b. "Human capital and growth: Theory and evidence." *Carnegie-Rochester Conference Series on Public Policy* 32, no.:251-286.
- Rothstein, Jesse. 2005. "Does Competition Among Public Schools Benefit Students and Taxpayers? A Comment on Hoxby (2000)." NBER Working Paper 11215, Cambridge, MA, National Bureau of Economic Research
- Rouse, Cecilia Elena. 1998. "Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 113, no. 2 (May):553-602.
- Sachs, Jeffrey D., and Andrew Warner. 1995. "Economic Reform and the Process of Global Integration." *Brookings Papers on Economic Activity* 1, no.:1-96.
- Sala-i-Martin, Xavier, Gernot Doppelhofer, and Ronald I. Miller. 2004. "Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach." *American Economic Review* 94, no. 4 (September):813-835.
- Sandström, F. Mikael, and Fredrik Bergström. 2005. "School Vouchers in Practice: Competition Will Not Hurt You." *Journal of Public Economics* 89, no. 2-3:351-380.

- Sapelli, Claudio, and Bernardita Vial. 2002. "The performance of private and public schools in the Chilean voucher system." *Cuadernos de Economia* 39, no. 118 (December):423-454.
- Sass, Tim R. 2006. "Charter Schools and student achievement in Florida." *Education Finance and Policy* 1, no. 1 (Winter):91-122.
- Sianesi, Barbara, and John Van Reenen. 2003. "The returns to education: Macroeconomics." *Journal of Economic Surveys* 17, no. 2:157-200.
- Skoufias, Emmanuel, and Joseph Shapiro. 2006. "Evaluating the Impact of Mexico's Quality Schools Program: The Pitfalls of Using Nonexperimental Data." World Bank Policy Research, Working Paper No. 4036, Washington, World Bank October.
- Temple, Jonathan. 1999. "A Positive Effect of Human Capital on Growth." *Economics Letters* 65, no. 1:131-134.
- Temple, Jonathan. 2001. "Growth effects of education and social capital in the OECD countries." *OECD Economic Studies* 33, no.:57-101.
- Temple, Jonathan, and Ludger Wößmann. 2006. "Dualism and Cross-country Growth Regressions." *Journal of Economic Growth* 11, no. 3:187-228.
- Topel, Robert. 1999. "Labor markets and economic growth." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier:2943-2984.
- Vandenbussche, Jérôme, Philippe Aghion, and Costas Meghir. 2006. "Growth, Distance to Frontier and Composition of Human Capital." *Journal of Economic Growth* 11, no. 2:97-127.
- Vegas, Emiliana. 1999. "School Choice, Student Performance, and Teacher and School Characteristics: The Chilean Case." Policy Research Working Paper No. 2833, Washington, World Bank November.
- Vegas, Emiliana, ed. 2005. *Incentives to improve teaching: Lessons from Latin America*. Washington D.C.: The World Bank.
- Vegas, Emiliana, and Llana Umansky. 2005. *Improving teaching and learning through effective incentives: What can we learn from education reforms in Latin America?* Washington D.C.: World Bank.
- Vijverberg, Wim P.M. 1999. "The impact of schooling and cognitive skills on income from non-farm self-employment." In *The economics of school quality investments in developing countries: An empirical study of Ghana*, edited by Paul Glewwe. New York: St. Martin's Press (with University of Oxford).
- West, Martin R., and Paul E. Peterson. 2006. "The Efficacy of Choice Threats within School Accountability Systems: Results from Legislatively-Induced Experiments." *Economic Journal* 116, no. 510:C46-C62.

- World Bank. 1993. *The East Asian miracle: Economic growth and public policy*. New York: Oxford University Press.
- World Bank. 2004. *World Development Report 2004: Making Services Work for Poor People*. Washington: World Bank.
- World Bank Independent Evaluation Group. 2006. *From Schooling Access to Learning Outcomes: An Unfinished Agenda*. Washington: World Bank.
- Wößmann, Ludger. 2001. "Why students in some countries do better." *Education Matters* 1, no. 2 (Summer):67-74.
- Wößmann, Ludger. 2002. *Schooling and the Quality of Human Capital* Berlin: Springer.
- Wößmann, Ludger. 2003a. "Schooling resources, educational institutions, and student performance: The international evidence." *Oxford Bulletin of Economics and Statistics* 65, no. 2:117-170.
- Wößmann, Ludger. 2003b. "Specifying human capital." *Journal of Economic Surveys* 17, no. 3:239-270.
- Wößmann, Ludger. 2004. "Access to Education: Perspective Paper " In *Global Crises, Global Solutions*, edited by Björn Lomborg. Cambridge: Cambridge University Press: 241-250.
- Wößmann, Ludger. 2005a. "Educational Production in East Asia: The Impact of Family Background and Schooling Policies on Student Performance." *German Economic Review* 6, no. 3:331-353.
- Wößmann, Ludger. 2005b. "Educational production in Europe." *Economic Policy* 20, no. 43:446-504.
- Wößmann, Ludger. 2005c. "The effect heterogeneity of central exams: Evidence from TIMSS, TIMSS-Repeat and PISA." *Education Economics* 13, no. 2:143-169.
- Wößmann, Ludger. 2005d. "Public-Private Partnerships in Schooling: Cross-Country Evidence on their Effectiveness in Providing Cognitive Skills." Program on Education Policy and Governance, Research Paper PEPG 05-09, Cambridge, MA, Harvard University.
- Wößmann, Ludger. forthcoming-a. "International Evidence on Expenditure and Class Size: A Review." In *Brookings Papers on Education Policy*. Washington D.C.: Brookings.
- Wößmann, Ludger. forthcoming-b. "International Evidence on School Competition, Autonomy and Accountability: A Review." *Peabody Journal of Education*.
- Wößmann, Ludger, and Thomas Fuchs. 2005. "Families, Schools, and Primary-School Learning: Evidence for Argentina and Colombia in an International Perspective." World Bank Policy Research, Working Paper 3537, Washington, World Bank.
- Wößmann, Ludger, and Martin R. West. 2006. "Class-Size Effects in School Systems Around the World: Evidence from Between-Grade Variation in TIMSS." *European Economic Review* 50, no. 3:695-736.