Chapter 6
The Economic Benefits of Improved Teacher Quality

Eric A. Hanushek[1]

Hoover Institution, Stanford University, Stanford, CA, USA

6.1. Introduction

Most developed countries are acutely aware of how their students do in comparison to those elsewhere in the world. The now frequent scores on PISA and TIMSS provide direct feedback on schools.[1] But, as comparative test scores have become more plentiful, two key questions arise. First, do scores on these tests make any difference? Second, how can they be changed by any governmental policies? This chapter addresses both of these questions.

Economists are now accustomed to looking at issues of human capital. The simplest notion is that individuals make investments in skills that have later payoffs in outcomes that matter. And, in this, it is commonly presumed that formal schooling is one of the several important contributors to the skills of an individual and to human capital. It is not the only factor. Parents, individual abilities, and friends undoubtedly contribute. Schools nonetheless have a special place because they are most directly affected by public policies.

[1] The Programme for International Student Assessment (PISA) has been conducted in 2000, 2003, and 2006; see http://www.oecd.org/pages/0,2966,en_32252351_32235731_1_1_1_1_1,00.html. TIMSS is the Trends in International Mathematics and Science Study (formerly the Third International Mathematics and Science Study) and is a continuation of international testing begun in the 1960s; see http://timss.bc.edu/.

N. C. Soguel and P. Jaccard (eds.), Governance and Performance of Education Systems, 107–135.

Much of the early and continuing development of empirical work on human capital concentrates on the role of school attainment, that is, the quantity of schooling. The revolution in the United States during the twentieth century was universal schooling. This has spread around the world, encompassing both developed and developing countries. Quantity of schooling is easily measured, and data on years attained, both over time and across individuals, are readily available. But quantity proves to be a poor measure of the skills of individuals both within and across countries.

Today, policy concerns in most corners of the world revolve much more around issues of quality than issues of quantity. This brings us back to PISA and TIMSS. Do standardized tests such as these identify qualities that have economic benefits? The next sections assess what we know about the payoff to cognitive skills for individuals and for nations. In short, there are very large payoffs to such skills. Individuals with more measured cognitive skill systematically do better than those with less. And nations with a more skilled population grow faster than those with a less skilled population.

The implications of this for policy have nonetheless been less clear. Simply providing more resources to schools has proved to be very ineffective. On the other hand, mounting evidence suggests that improving teacher quality is the one way in which student outcomes can be systematically improved. The results about the importance of teacher quality are related directly to the economic benefits of improved quality.

6.2. Impacts of Quality on Individual Incomes – Developed Countries

One of the challenges in 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. 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 recently, little comprehensive data have been available to show any relationship between differences in cognitive skills and any related economic outcomes. Such data are now becoming available.

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 and the possible influence of differences in ability. 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.[2]

There is mounting evidence that quality measured by test scores is directly related to individual earnings, productivity, and economic growth. A variety of researchers document that the earnings advantages to higher achievement on standardized tests are quite substantial.[3] 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 country is closely related to individual productivity and earnings.

Three recent US studies provide direct and quite consistent estimates of the impact of test performance on earnings (Mulligan 1999; Murnane et al. 2000; Lazear 2003). These studies employ different nationally representative data sets that follow students after they leave schooling and enter the labor force. When scores are standardized, they suggest that one standard deviation increase in mathematics performance at the end of high schools translates into 12% higher annual earnings.

---

[2] The approaches have included looking for circumstances where the amount of schooling is affected by things other than the student's valuation of continuing and considering the income differences among twins (see Card 1999).

[3] 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 (which are analyzed in Hanushek 2002). (See Bishop 1989, 1991; O'Neill 1990; Blackburn and Neumark 1993, 1995; Grogger and Eide 1993; Murnane et al. 1995, 2000, 2001; Neal and Johnson 1996; Mulligan 1999; Altonji and Pierret 2001; Lazear 2003).

Murnane et al. (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% increase and females a 10% increase per standard deviation of test performance. Lazear (2003), relying on a somewhat younger sample from NELS88, provides a single estimate of 12%. These estimates are also very close to those in Mulligan (1999), who finds 11% for the normalized AFQT score in the NLSY data. By way of comparison, estimates of the value of an additional year of school attainment are typically 7–10%.

There are reasons to believe that these estimates provide a lower bound on the impact of higher achievement. First, these estimates are obtained fairly early in the work career (mid-20s to early 30s), and other analysis suggests that the impact of test performance becomes larger with experience.[4] Second, 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. Third, future general improvements in productivity are likely to lead to larger returns to skill.[5]

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.[6] 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. There is substantial US evidence that students who do better in school, either through grades or scores on standardized achievement tests,

[4] Altonji and Pierret (2001) find that the impact of achievement grows with experience, because the employer has a chance to observe the performance of workers.

[5] These 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. Thus, if the patterns of recent decades continue, the impact of improvements in student skills could likely rise over the work life instead of being constant as portrayed here.

[6] Because they look at discrete levels of skills, it is difficult to compare the quantitative magnitudes directly to the US work.

tend to go further in school.[7] Murnane et al. (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 dropout 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.

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.

The impact of test performance on individual earnings provides a simple summary of the primary economic rewards to an individual. This estimate combines the impacts on hourly wages and on employment/hours worked. It does not include any differences in fringe benefits or nonmonetary aspects of jobs. Nor does it make any allowance for aggregate changes in the labor market that might occur over time.

[7] 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 et al. (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 et al. (2000). Behrman et al. (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.

6.3. Impacts of Quality on Individual Incomes – Developing Countries

Questions remain about whether the clear impacts of quality in the United States 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 6.1 provides a summary of the quantitative estimates available for developing countries. The summary of the evidence 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; 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.[8] 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 et al. (2007) 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 quality concerns are very real for developing countries and that this aspect of schools simply cannot be ignored – a topic that comes up below.

[8] 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).

Table 6.1. 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 from off-farm income with on-farm income unrelated to skills
Ghana	Vijverberg (1999)	?	Income estimates for math and reading with non-farm self-employment; highly variable estimates (including both positive and negative effects) but effects not generally statistically significant
Kenya	Boissiere et al. (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 et al. (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 et al. (forthcoming)	0.25	Estimates of structural model with combined scores for cognitive skill; index significant at 0.01 level
South Africa	Moll (1998)	0.34**–0.48**	Depending on estimation method, varying impact of computation; comprehension (not shown)generally insignificant
Tanzania	Boissiere et al. (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.

6.4. Impacts of Quality on Economic Growth

The relationship between measured labor force quality and economic growth is perhaps even more important than the impact of human capital and school quality on individual productivity and incomes. Economic growth determines how much improvement will occur in the overall standard of living of society. Moreover, the education of each individual has the possibility of making others better off (in addition to the individual benefits just discussed). Specifically, a more educated society may lead to higher rates of invention; may make everybody more productive through the ability of firms to introduce new and better production methods; and may lead to more rapid introduction of new technologies. These externalities provide extra reason for being concerned about the quality of schooling.

The potential effect of differences in growth rates on economic well-being is easy to see. Take the expected growth of a country as given and consider how incomes would change with a marginal improvement. Figure 6.1 begins with the value of gross domestic product (GDP) per capita for a

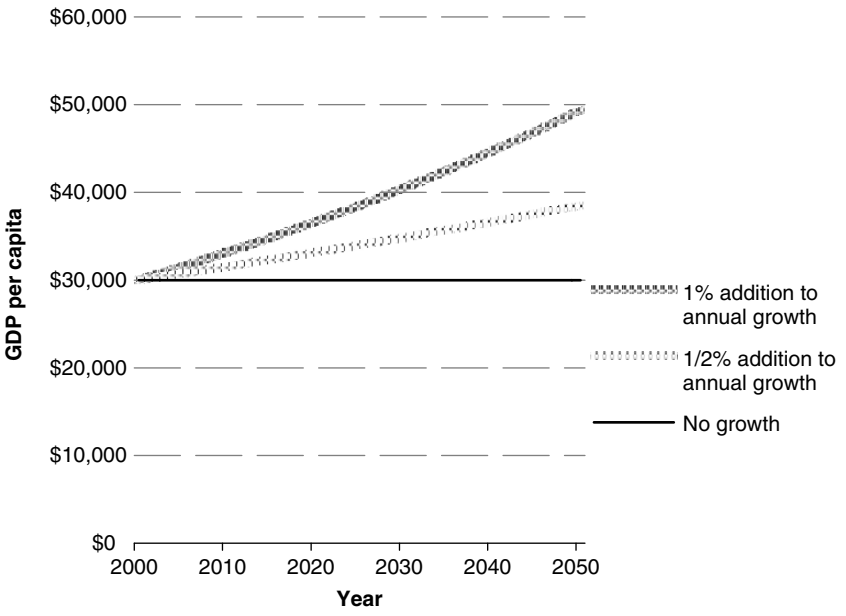

Fig. 6.1. Effect of Economic Growth on Per Capita Income

medium income European country in the year 2000 and shows its value in 2050 under different growth rates (assuming for simplicity that growth would otherwise be zero). If it grows at 1% more each year, this measure (in US dollars) would increase from $30,000 to almost $50,000 – or increasing by almost two-thirds over the period because of this marginal improvement. If it were to grow at 0.5% per year, it would still exceed $38,000 in 2050. Small differences in growth rates have huge implications for the income and wealth of society.

The current economic position of the United States, for example, is largely the result of its strong and steady growth over the twentieth century. Economists have developed a variety of models and ideas to explain differences in growth rates across countries – invariably featuring the importance of human capital.[9]

The empirical work supporting growth analyses has emphasized school attainment differences across countries. Again, this is natural because, while compiling comparable data on many things for different countries is difficult, assessing quantity of schooling is more straightforward. The typical study finds that quantity of schooling is highly related to economic growth rates. But, quantity of schooling is a very crude measure of the knowledge and cognitive skills of people – particularly in an international context.

Hanushek and Kimko (2000) go beyond simple quantity of schooling and delve into quality of schooling.[10] We incorporate the information about international differences in mathematics and science knowledge that has been developed through testing over the past four decades. And we find a remarkable impact of differences in school quality on economic growth.

The international comparisons of quality come from piecing together results of a series of tests administered over the past four decades. In 1963 and 1964, the International Association for the Evaluation of Education al Achievement (IEA) administered the first of a series of mathematics tests to a voluntary group of countries. These initial tests suffered from a

[9] Barro and Sala-i-Martin (2004) review recent analyses and the range of factors that are included.

[10] Barro and Lee (2001) provide an analysis of qualitative differences that also includes literacy. Others have also investigated quality and growth; see Barro (2001), Bosworth and Collins (2003), Wößmann (2002), and Jamison et al. (2006).

number of problems, but they did prove the feasibility of such testing and set in motion a process to expand and improve on the undertaking.[11]

Subsequent testing, sponsored by the IEA and others, has included both math and science and has expanded on the group of countries that have been tested. In each, the general model has been to develop a common assessment instrument for different age groups of students and to work at obtaining a representative group of students taking the tests. Using these test data, it is possible to track performance (aggregated across the age groups and subject area of the various tests) over time.[12] The United States and the United Kingdom are the only countries to participate in all of the testing.

There is some movement across time of country performance on the tests, but for the one country that can be checked – the United States – the pattern on international tests is consistent with other data. The National Assessment of Educational Progress (NAEP) in the United States is designed to follow performance of US students for different subjects and ages and shows a sizable dip in US student performance in the 1970s, a period of growth in the 1980s, and a leveling off in the 1990s – exactly the pattern on international tests.

Kimko's and my analysis of economic growth is very straightforward. We combine all of the available earlier test scores into a single composite measure of quality and consider statistical models that explain differences in growth rates across nations during the period 1960–1990.[13] The basic statistical models, which include the initial level of income, the quantity of schooling, and population growth rates, explain a substantial portion of the variation in economic growth across countries.

---

[11] The problems included issues of developing an equivalent test across countries with different school structure, curricula, and language; issues of selectivity of the tested populations; and issues of selectivity of the nations that participated. The first tests did not document or even address these issues in any depth.

[12] The details of the tests and aggregation can be found in Hanushek and Kimko (2000) and Hanushek and Kim (1995).

[13] We exclude the TIMSS and PISA tests from 1995 on because they were taken outside of the analytical period on economic growth. We combine the test measures over the 1965–1991 period into a single measure for each country. The underlying objective is to obtain a measure of quality for the labor force in the period during which growth is measured.

Most important, the quality of the labor force as measured by math and science scores is extremely important. One standard deviation difference on test performance is related to 1% difference in annual growth rates of GDP per capita.[14]

This quality effect, while possibly sounding small, is actually very large and significant. Because the added growth compounds, it leads to powerful effects on national income and on societal well-being. One needs only to return to the calculations presented in Fig. 6.1 to understand the impact of such skill-based improvements in economic growth.

Extensions of this work by Jamison et al. (2006) to 2000 show a very similar pattern of quality on growth. Importantly, building on the construction of new quality information from recent testing by Hanushek and Woessmann (2007), adds considerably more countries to the sample for the growth analysis – and the results hold.

6.5. Importance of Quality

The frequent focus of governmental programs has been increasing school attainment and expanding on the years of schooling of the population. The previous discussion, however, highlights the central importance of quality. While years of schooling attainment are important, that holds only if quality is maintained.

The impact of improved quality can be calculated from the considerations of how quality affects growth rates for economies. Consider the effects of beginning a successful school improvement program in 2005. Of course, school reform takes time. And, even if successful, it takes some time before the school graduates work their way into the labor force and thus some time before the impact will be felt.

Figure 6.2 illustrates the impact that reform could be expected to have over time if it is successful at achieving moderately strong knowledge improvement (corresponding to a 0.5 standard deviation increase

[14] The details of this work can be found in Hanushek and Kimko (2000) and Hanushek (2003b). Importantly, adding other factors potentially related to growth, including aspects of international trade, private and public investment, and political instability, leaves the effects of labor force quality unchanged.

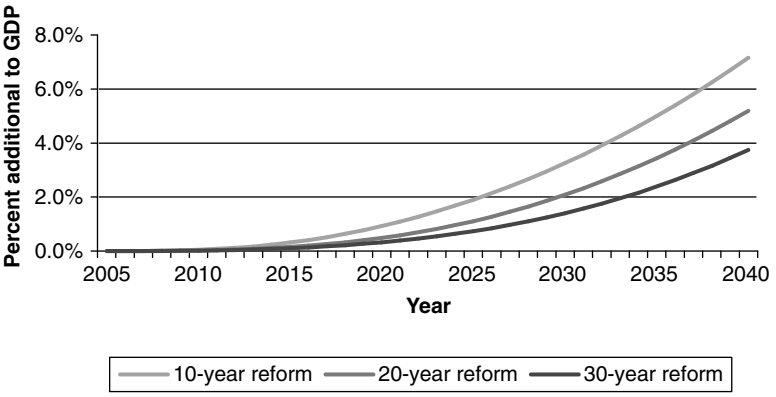

Fig. 6.2. Improved GDP with Moderately Strong Knowledge Improvement

in test score achievement).[15] The curves sketch out the path of GDP improvement that would occur with a reform plan that reaches its improvement goal within 10, 20, or 30 years.

Consider just the slow improvement of schools over a 30-year period. In 2040, the GDP would be almost 4% higher than projected without the schooling reforms. Of course, faster reforms would yield even greater gains in GDP. This magnitude would cover total school spending in most countries of the world.

6.6. Causality

One common concern in analysis 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, and they also had extraordinarily high growth over the 1960–1990 period. It may be

[15] These calculations are calibrated to scores on international mathematics and science exams. The "moderately strong" improvement implies an increase in scores by 0.5 standard deviations across the international comparisons. This is equivalent of bringing a country at the 31st percentile of performance up to the median for the world.

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 might be 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. In order to investigate this, we concentrate on immigrants to the United States who received their education in their home countries. We 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. 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 resources necessary to improve their schools and that better student performance is the result of growth, not the cause of growth. As a simple test of this, we 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, we found relatively better performance in those countries spending less on their schools.

In sum, the relationship between math and science skills on the one hand and productivity and growth on the other comes through clearly when investigated in a systematic manner across countries. This finding underscores the importance of high-quality schooling.

6.7. Why has US Growth been so Strong?

The United States has not been competitive on an international level in terms of tests. It has scored below the median of countries taking the various tests. Moreover, the performance on tests of US students is much stronger at young ages but falls off dramatically at the end of high school

(Hanushek 2003b). Understanding how this matches with growth is important for understanding the broader policy implications.

Earlier, we introduced the discussion of the importance of growth by recounting United States' successful economic growth during the twentieth century. Yet, this is contrasted with the fact that the United States has been at best mediocre in mathematics and science ability. Regardless of the set of countries taking the test, the United States has performed in the middle of the pack or below. Some people find this anomalous. How could math and science ability be important in light of the strong US growth over a long period of time?

The answer 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 demonstrate that schooling is not itself a sufficient engine of growth.

Indeed, some have questioned the precise role of schooling in growth. Easterly (2002), for example, notes that education without other facilitating factors such as functioning institutions for markets and legal systems may not have much impact. He argues that World Bank investments in schooling for less developed countries that do not ensure that the other attributes of modern economies are in place have been quite unproductive. As discussed below, schooling clearly interacts with other factors, and these other factors have been important in supporting US growth. They are also surely relevant for other countries.

It is useful to describe some of the other contributing factors to US growth. This is done in part to understand more fully the character of economic growth, but more importantly to highlight some important related issues that are central to thinking about human capital policies.

Almost certainly the most important factor sustaining the growth of the US 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 (in terms of both labor regulations and overall production), 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 – not only less regulation but also 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 US 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.[16]

Over the twentieth century, the expansion of the education system in the United States also 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. In comparison with other nations of the world, the US labor force has been better educated, even after allowing for the lesser achievement of its graduates. In other words, more schooling with less learning each year has yielded more human capital than found in other nations that have less schooling but learn more in each of those years.

Finally, the analysis of growth rates across countries emphasizes quality of the elementary and secondary schools of the United States. It did not include any measures of the quality of US colleges. By most evaluations, US colleges and universities rank at the very top in the world. No direct measurements of quality of colleges across countries exist. However, there is indirect evidence. Foreign students by all accounts are not tempted to emigrate to the United States to attend elementary and secondary schools – except perhaps if they see this as a way of gaining entry into the country. They do emigrate in large numbers to attend US colleges and universities. They even tend to pay full, unsubsidized tuitions at US colleges, something that many fewer US citizens do.

6.8. Generalizing to Developing Countries

The previous discussions have concentrated considerable attention on the United States and on other developed countries. Most developing countries look very dissimilar to these. Do these results generalize?

The modeling of economic growth in Hanushek and Kimko (2000) relied upon the direct measures of math and science achievement that unfortunately included relatively few developing countries. The analysis did, however, work to extend the modeling to a large number of countries not included in the direct testing. This was done by modeling test scores and

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

then projecting the results to other countries. The analysis did not, however, consider all countries. It excluded countries whose predicted scores fell outside the range of observed tests. This exclusion applied to a number of developing countries.

Within the set of countries with observed or projected test data, the growth models appear rather robust. A variety of tests indicate that the modeling applies to the range of countries. This is reinforced by the additions to the sample by Jamison et al. (2006).

Questions remain, however, about the wider range of countries. Clearly, many of the arguments made by Easterly (2002) obviously apply to the most destitute countries – those which also tend to lack a good structure of laws, which tend to have a variety of restrictions on labor and product markets, and so forth. These countries may not be able to fruitfully use schooling investments if the labor markets will not accommodate skilled workers.

The tentative conclusion would be that the previous results generalize if the other conditions for growth also exist. If they do not, it is much more uncertain. But it is also true in the latter cases that investment in quantity of schooling is unlikely to be productive either.

6.9. Improving Quality

Much of school policy is traditionally thought of as an exercise in selecting and ensuring that the optimal set of resources, somehow defined, is available. Matched with this policy perspective has been a line of research considering the relationship between resource usage and student performance. If the effectiveness of different resources or combinations of resources were known, it would be straightforward to define an optimal set of resources. Moreover, we could often decide about policies that would move us toward such an optimal set of resources. Unfortunately, this alludes us.

Schools in the United States have been the focus of extensive research. Both aggregate data about performance of schools over time and more detailed school and classroom data point to a simple conclusion: There is a lack of any consistent or systematic effect of resources on student achievement. While controversial, partly because of the conflict with existing school policies, the evidence is very extensive (Hanushek 2003a). Most other countries of the world have not tracked student performance over any length of time, making analyses comparable to the US discussion impossible. Nonetheless, international testing over the past four decades permits an overview of spending across countries. The simplest overview

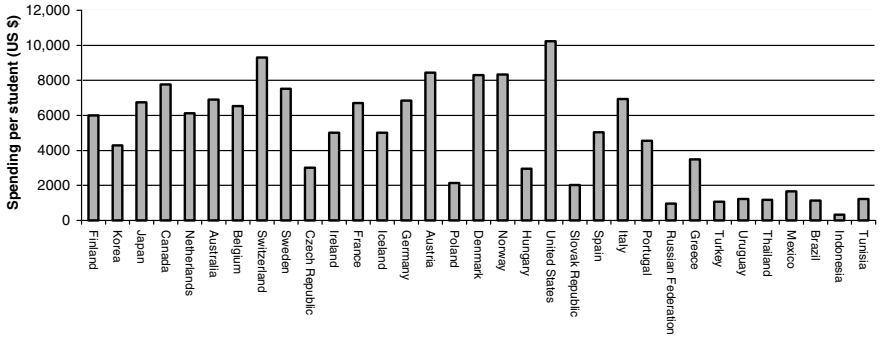

Fig. 6.3. Expenditure per Sutdent at All Levels (countries ranked by combined PISA 2003 scores)

comes from the most recent PISA tests. Figure 6.3 ranks countries by performance on PISA, and the height of the bars gives the spending per pupil in each (on a purchasing power parity basis). Instead of a simple declining pattern, one sees essentially no correlation until reaching the least developed countries.

Seven different mathematics and science tests (the data for the growth analysis) were given between the early 1960s and 1995 to students at different grade levels in a varying set of voluntarily participating nations. Performance bears little relationship to the patterns of expenditure across the countries. Hanushek and Kimko (2000) estimate models that relate spending, family backgrounds and other characteristics of countries to student performance for the tests prior to 1995. This estimation consistently indicates a statistically significant negative effect of added resources on performance after controlling for other influences. Similar findings hold for the OECD countries.

Existing statistical analyses in less developed countries have shown a similar inconsistency of estimated resource effects as that found in the United States (Hanushek 1995). In general, a minority of the available studies suggests much confidence that commonly identified resources – class size, teacher experience, and teacher salaries – positively influence student performance. There is generally somewhat stronger support for these resource policies than that existing in US analyses, hinting that the importance of resources may vary with the level of resources. Nonetheless, the evidence does not indicate that pure resource policies can be expected to have a significant effect on student outcomes.

In sum, a wide range of analyses indicate that overall resource policies have not led to discernible improvements in student performance. It is important to understand what is and is not implied by this conclusion. First, it does not mean that money and resources *never* matter. There clearly are situations where small classes or added resources have an impact. It is just that no good description of when and where these situations occur is available, so that broad resource policies such as those legislated from central governments may hit not only some good uses but also bad uses that generally lead to offsetting outcomes. Second, this statement does not mean that money and resources *cannot* matter. Instead, as described below, altered sets of incentives could dramatically improve the use of resources.

The evidence on resources is remarkably consistent across countries, both developed and developing. Had there been distinctly different results for some subsets of countries, issues of what kinds of generalizations were possible would naturally arise. Such conflicts do not appear particularly important.

Many countries have of course attempted to improve their schools. While some have succeeded, many have not. One explanation for past failure is simply that insufficient attention has been given to teacher quality. By many accounts, the quality of teachers is the key element to improving student performance. But the research evidence also suggests that many of the policies that have been pursued around the world have not been very productive. Specifically, the chosen policies of individual countries may have led to changes in measured aspects of teachers such as degrees or teacher qualifications, but they have not tended to improve the quality of teachers – at least when quality is identified by student performance.[17]

Rivkin et al. (2005) describe estimates of differences in teacher quality on an output basis. Specifically, the concern is identifying good and bad teachers on the basis of their performance in obtaining gains in student achievement. An important element of that work is distinguishing the effects of teachers from the selection of schools by teachers and students and the matching of teachers and students in the classroom. In particular, highly motivated parents search out schools that they think are good, and they attempt to place their children in classrooms where they think the teacher is particularly able. Teachers follow a similar selection process (Hanushek et al. 2004). Thus, from an analytical viewpoint, it is difficult to

[17] For a review of existing US literature, see Hanushek and Rivkin (2004). This paper describes various attempts to estimate the impact of teacher quality on student achievement. Similar studies are currently much less available in other countries.

sort out the quality of the teacher from the quality of the students that he/she has in his/her classroom. The analysis of teacher performance in Rivkin et al. (2005) goes to great lengths to avoid contamination from any such selection and matching of kids and teachers.

Estimates show that the differences in annual achievement growth between an average and a good teacher are large. Within one academic year, a good teacher can move a typical student up at least four percentiles in the overall distribution (equal to a change of 0.12 standard deviations of student achievement).[18] From this, it is clear that having a series of good teachers can dramatically affect the achievement of any student. In fact, a series of good teachers can erase the deficits associated with poor preparation for school.

It is also possible to see what these results imply for improving student achievement in the aggregate. Perhaps the simplest policy is to replace teachers who leave the profession with new, higher-quality teachers. While turnover of teachers differs across countries, a description of the implications for the US school system illustrates the general points. In the United States, around 7% of all teachers exit teaching each year; another 6% change schools. To give some sense of the leverage hiring has on the system, this range (7–13%) is used to identify the replacement possibilities.

Figure 6.4 displays the annual hiring improvement that is necessary to achieve a 0.5 standard deviation improvement under a 10-, 20-, and 30-year reform plan and based on applying it to either just those exiting or the higher turnover rates that include transfers. As is obvious, the stringency of the new hiring is greater when there is a shorter reform period and when fewer new (higher-quality) teachers are brought in each year. Achieving a 0.5 SD boost in achievement in 10 years by upgrading just those who exit each year implies hiring at the 61st percentile, but this declines to the 52nd percentile for a 30-year plan where the higher turnover population is subject to these new hiring standards.

These calculations are meant to illustrate two points. First, existing research into student achievement and teacher quality shows that teachers have significant leverage on performance. By implication, if better teachers can be hired and retained, significant changes in student achievement

[18] In another attempt to estimate the variation in teacher quality, we analyze variations across classrooms within a large school district in Texas (Hanushek et al. 2005). In this, we match individual teachers and students and look at achievement-based quality measures of each teacher compared to all of the teachers in the district or, alternatively, all of the other teachers in each school. On a basis comparable to the prior estimates, we obtain an estimate of teacher quality between 0.15 and 0.18 standard deviations of student achievement.

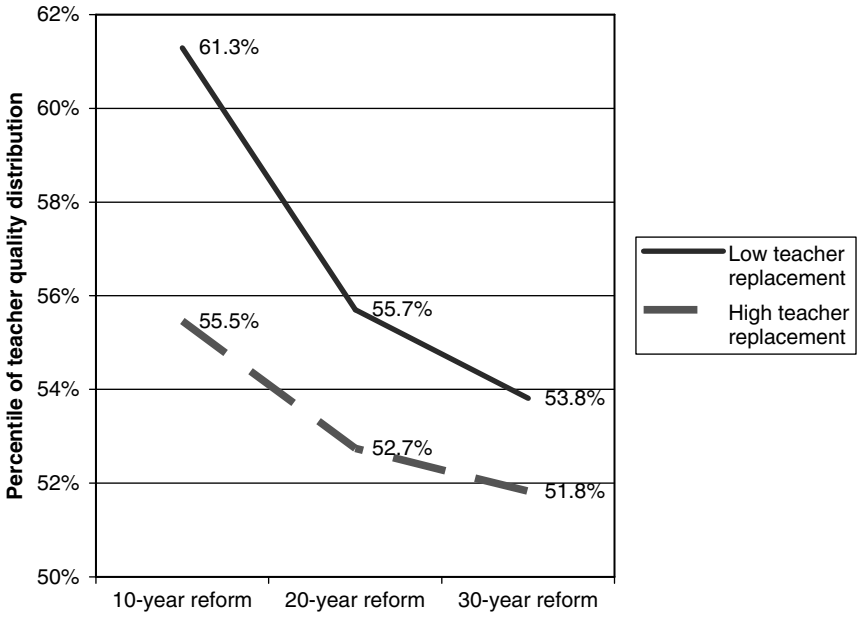

Fig. 6.4. Required Quality Percentile for New Teaches (0.5 s.d. Reform)

can be obtained. Second, without dramatic changes in policies about teacher retention, feasible reform will take a quite long period of time. Specifically, unless larger numbers of current teachers are fired and replaced, changing the character of the teaching force takes time.

6.10. Conclusions

School quality is directly related to decisions about attending schools and to promotion through schools. High-quality schools raise student achievement and speed students through primary (and perhaps secondary) schools, thus conserving on costs. Thus, studies of the rate of return to schooling which only consider quantity of schooling produce a misleading estimate of the potential gains. Estimation of the rate of return to schooling that does not account for quality differences will systematically overstate the productivity gains that are associated with additional years of schooling, because the estimates will include quality differences that are correlated with quantity. If policy simply pushes people to stay in school longer, without changing the fundamental quality of the schools, the newly induced school completers will only get the returns associated with years of

schooling and not with quality. Thus, they will not be able to gain as much as the rate of return estimates suggest.

Policy makers who concentrate on quality of schools are frequently stopped, however, when they begin considering how to improve quality. There has been a huge amount of work on various approaches, but the record of accomplishments is modest.

Recent work underscores the importance of high-quality teachers. While the evidence is limited to US schools, teachers appear to have a very strong impact on student outcomes. Unfortunately, teacher quality is not simply measured by such things as experience or teacher education. Thus, developing policies to implement this finding will take some effort.

Nonetheless, the potential economic gains from improvement also suggest that there is considerable room for aggressive policies to attract and retain good teachers. With a suitable planning horizon, it appears feasible to upgrade the teaching force, yielding truly large gains for students and for nations.

Acknowledgment

Lei Zhang provided valuable research assistance.

References

Alderman H, Behrman JR, David R, Sabot R (1996) The returns to endogenous human capital in Pakistan's rural wage labor market. *Oxford Bulletin of Economics and Statistics* 58:29–55

Altonji JG, Pierret CR (2001) Employer learning and statistical discrimination. *Quarterly Journal of Economics*116:313–350

Angrist JD, Lavy V (1997) The effect of a change in language of instruction on the returns to schooling in Morocco. *Journal of Labor Economics* 15:S48–S76

Barro RJ (2001) Human capital and growth. *American Economic Review* 91:12–17

Barro RJ, Lee J-w (2001) International data on educational attainment: updates and implications. *Oxford Economic Papers* 53:541–563

Barro RJ, Sala-i-Martin X (2004) *Economic Growth* (second edition). The MIT Press, Cambridge, MA

Behrman JR, Kletzer LG, McPherson MS, Schapiro MO (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* 559:12–23

Behrman JR, David R, Sabot R (2007) Improving the quality versus increasing the quantity of schooling: estimates of rates of return from rural Pakistan. *Journal of Development Economics.* Article in press

Bishop J (1989) Is the test score decline responsible for the productivity growth decline? *American Economic Review* 79:178–197

Bishop J (1991) Achievement, test scores, and relative wages. In: Kosters MH (ed) *Workers and Their Wages*. The AEI Press, Washington, DC, pp 146–186

Blackburn, ML, Neumark D (1993) Omitted-ability bias and the increase in the return to schooling. *Journal of Labor Economics* 11:521–544

Blackburn, ML, Neumark D (1995) Are OLS estimates of the return to schooling biased downward? Another look. *Review of Economics and Statistics* 77:217–230

Boissiere MX, Knight JB, Sabot R (1985) Earnings, schooling, ability, and cognitive skills. *American Economic Review* 75:1016–1030

Bosworth BP, Collins SM (2003) The empirics of growth: an update. *Brookings Papers on Economic Activity* 2:113–206

Card D (1999) Causal effect of education on earnings. In: Ashenfelter O, Card D (eds) *Handbook of Labor Economics*. North-Holland, Amsterdam , pp 1801–1863

Dugan DJ (1976) Scholastic achievement: its determinants and effects in the education industry. In: Froomkin JT, Jamison DT, Radner R (eds) *Education as an Industry*. Ballinger, Cambridge, MA, pp 53–83

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

Finnie R, Ronald M (2002) Minorities, cognitive skills, and incomes of Canadians. *Canadian Public Policy* 28:257–273

Glewwe P (1996) The relevance of standard estimates of rates of return to schooling for educational policy: a critical assessment. *Journal of Development Economics* 51:267–290

Glewwe P (2002) Schools and skills in developing countries: education policies and socioeconomic outcomes. *Journal of Economic Literature* 40:436–482

Green DA, Riddell WC (2003) Literacy and earnings: an investigation of the interaction of cognitive and unobserved skills in earnings generation. *Labour Economics* 10:165–184

Grogger JT, Eide E (1993) Changes in college skills and the rise in the college wage premium. *Journal of Human Resources* 30:280–310

Hanushek EA (1995) Interpreting recent research on schooling in developing countries. *World Bank Research Observer* 10:227–246

Hanushek EA——— (2002) Publicly provided education. In: Auerbach AJ, Feldstein M (eds) *Handbook of Public Economics*. Elsevier, Amsterdam, pp 2045–2141

Hanushek EA (2003a) The failure of input-based schooling policies. *Economic Journal* 113:F64–F98

Hanushek EA (2003b) The importance of school quality. In: Peterson PE (ed.) *Our Schools and Our Future: Are We Still at Risk?* Hoover Institution Press, Stanford, CA, pp 141–173

Hanushek EA, Kain JF, O'Brien DM, Rivkin SG (2005) The market for teacher quality. Working Paper 11154, National Bureau of Economic Research

Hanushek EA, Kain JF, Rivkin SG (2004) Why public schools lose teachers. *Journal of Human Resources* 39:326–354

Hanushek EA, Kim D (1995) Schooling, labor force quality, and economic growth. Working Paper 5399, National Bureau of Economic Research

Hanushek EA, Kimko D (2000) Schooling, labor force quality, and the growth of nations. *American Economic Review* 90:1184–1208

Hanushek EA, Pace R (1995) Who chooses to teach (and why)? *Economics of Education Review* 14:101–117

Hanushek EA, Rivkin SG (2004) How to improve the supply of high quality teachers. In: Ravitch D (ed.) *Brookings Papers on Education Policy 2004*. Brookings Institution Press, Washington, DC, pp 7–25

Hanushek EA, Rivkin SG, Taylor L (1996) Aggregation and the estimated effects of school resources. *Review of Economics and Statistics* 78:611–627

Hanushek EA, Woessmann L (2007) The role of education quality for economic growth. Policy Research Working Paper Series 4122, The World Bank. Washington

Jamison EA, Jamison DT, Hanushek EA (2006) The effects of education quality on mortality decline and income growth. International Conference on the Economics of Education, June 20–23, Dijon, France)

Jolliffe D (1998) Skills, schooling, and household income in Ghana. *World Bank Economic Review* 12:81–104

Knight JB, Sabot RH (1990) *Education, Productivity, and Inequality*. Oxford University Press, New York

Krueger AO (1974) The political economy of the rent seeking society. *American Economic Review* 64:291–303

Lazear EP (2003) Teacher incentives. *Swedish Economic Policy Review* 10:179–214

Manski CF, Wise DA (1983) *College Choice in America*. Harvard University Press, Cambridge

McIntosh S, Vignoles A (2001) Measuring and assessing the impact of basic skills on labor market outcomes. *Oxford Economic Papers* 53:453–481

Moll PG (1998) Primary schooling, cognitive skills, and wage in South Africa. *Economica* 65:263–284

Mulligan CB (1999) Galton versus the human capital approach to inheritance. *Journal of Political Economy* 107:S184–S224

Murnane RJ, Willett JB, Braatz MJ, Duhaldeborde Y (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* 20:311–320

Murnane RJ, Willett JB, Duhaldeborde Y, Tyler JH (2000) How important are the cognitive skills of teenagers in predicting subsequent earnings? *Journal of Policy Analysis and Management* 19:547–568

Murnane RJ, Willett JB, Levy F (1995) The growing importance of cognitive skills in wage determination. *Review of Economics and Statistics* 77:251–266

Neal DA, Johnson WR (1996) The role of pre-market factors in black–white differences. *Journal of Political Economy* 104:869–895
O'Neill J (1990) The role of human capital in earnings differences between black and white men. *Journal of Economic Perspectives* 4:25–46
Parente SL, Prescott EC (1994) Barriers to technology adoption and development. *Journal of Political Economy* 102:298–321
Parente SL, Prescott EC (1999) Monopoly rights: a barrier to riches. *American Economic Review* 89:1216–1233
Psacharopoulos G (1994) Returns to investment in education: a global update. *World Development* 22:1325–1344
Psacharopoulos G, Patrinos HA (2004) Returns to investment in education: a further update. *Education Economics* 12:111–134
Rivkin SG (1995) Black/white differences in schooling and employment. *Journal of Human Resources* 30:826–852
Rivkin SG, Hanushek EA, Kain JF (2005) Teachers, schools, and academic achievement. *Econometrica 7*3:417–458
Vijverberg WPM (1999) The impact of schooling and cognitive skills on income from non-farm self-employment. In: Glewwe P (ed) *The Economics of School Quality Investments in Developing Countries: An Empirical Study of Ghana*. St Martin's Press (with University of Oxford), New York
World Bank (1993) *The East Asian Miracle: Economic Growth and Public Policy.* Oxford University Press, New York
Wößmann L (2002) *Schooling and the Quality of Human Capital*. Springer, Berlin

Comments

George Sheldon[1]

[1]*Director of the Industrial Organization and Labor Market Research Unit (FAI), Department of Economics, University of Basel, Basel, Switzerland*

Eric Hanushek's chapter reviews two strands of literature relating to the quality of schooling. One line of research pertains to the impact of school quality on individual earnings and economic growth and uses standardized achievement test scores in place of years of schooling as proxies for human capital in Mincer wage equations and international growth regressions. The second direction of enquiry focuses on the factors determining the level of school quality and employs test scores too, albeit as dependent variables – instead of explanatory variables – in educational production functions intended to explain the variation of achievement test scores across participants. In assessing the research results, Hanushek concludes (1) that the quality of schooling has a strong impact on individual earnings and economic growth and (2) that the amount of educational resources invested in schooling has no systematic effect on the quality of schooling. He qualifies the latter albeit by noting that recent research suggests that at least teacher quality has a statistically significant impact on students' cognitive abilities.

Hanushek provides a very clear and enlightening overview of current research, and I can subscribe to much of what he has to say. Only in two instances do I have any serious reservations. The one case pertains to his claim that no systematic relationship exists between the amount of educational resources invested and the cognitive ability of students, and the other to his thesis that the *quality* of schooling, at least in developed countries, has a greater impact on individual incomes and economic growth than the *quantity* of schooling.

I begin with my first point of contention: the apparent lack of a systematic relationship between the amount of educational resources invested and students' level of cognitive skills. I do not wish to belabor this point, however, as many other authors have already questioned Hanushek's stance on

this issue, as he himself knows. Besides, I generally agree with Hanushek's judgment that throwing money at an educational problem does not necessarily solve it and that educational policy is often economically inefficient. What I wish to bring to mind here are two other points. First of all, the assessment that no systematic relationship exists between the level of educational inputs, and the output of cognitive skills they engender generally rests on a simple comparison of the number of educational production function regressions that yield positive, negative, or statistically insignificant results without considering whether the results are multiple estimates pertaining to the same sample of data and without taking the ranking of the journal of publication into account. Other authors[19] have shown that one can reach quite different conclusions when the latter two factors are regarded.

Secondly, it is important to consider what an apparent failure to find a systematic relationship means or implies. It could indicate that a relationship does in fact not exist or, instead, that the educational process was inadequately modeled, the data were poor, or the empirical methodology was inappropriate. Todd and Wolpin (2003) provide reasons to believe that not only the non-existence of a relationship, but faulty research as well is the cause.

I turn now to my second point of contention, the claim that the *quality* of schooling is the central educational issue today and not the *quantity* of schooling. Hanushek bases his judgment largely on the observation that standardized achievement test scores are better able to explain individual income differences and international economic growth disparities than the number of years spent in school, i.e., than the level of educational attainment. Cognitive skill levels, which standardized achievement test scores are intended to measure, are not solely the product of the quality of schooling, however. In fact, as Hanushek himself notes, cognitive abilities not only depend on the quality of schooling, but among other things also on the level of educational achievement, parental upbringing,[20] cultural differences, and innate ability. Hence it is basically unknown what the determining factors are that lie behind the cognitive skills that enter into these regressions. School quality is but one possibility.

What the better predicative power of achievement test scores really tells us is that years of schooling are a poorer measure of cognitive ability than standardized achievement test scores and that the economic impact of hu-

[19]See for instance Hedges et al. (1994) or Krueger (2003).

[20]Wössmann (2004), for example, finds that the explanatory power of parental background dwarfs the effects of school inputs and institutional features on educational achievement.

man capital increases noticeably when the latter is more accurately measured. That is good news for economists as it underscores the importance of the economic study of education.

The distinction between the quantity and quality of schooling is not merely a matter of semantics either. The critical issue in the United States may indeed be one of school quality, as evidenced by the trend decline in achievement test scores despite increasing spending on education. But in Europe, the quantity of schooling, especially the low educational attainment of foreign youth, is a major issue as well. Many young foreigners in Europe are the children of low-skilled guest workers recruited to perform menial tasks that natives find unattractive to do. Given the low intergenerational educational mobility in Europe, a disproportionate share of young foreigners are thus concentrated in remedial and unchallenging paths of study that offer little opportunity for educational and economic advancement. As a consequence youth unemployment in Europe is particularly high among foreigners.[21] A large stock of low-skilled labor is an educational issue that Europe cannot afford to ignore. Given the skill bias of technical progress and the ongoing export of low-skilled manufacturing jobs to developing countries, it threatens the international competitiveness of Europe's economies.

But why is academic attainment among foreign youth so low in Europe? Recent evidence in Switzerland suggests two possible causes. One of the reasons appears to be the comparatively early selection of students into different educational paths of study. In many cantons in Switzerland, and in most of Germany as well, the decision by the school authorities to allow young persons to pursue a course of studies permitting later entry into college is often based on a student's marks in fourth grade and without the aid of standardized test scores. Bauer and Riphahn (2005) show that early selection significantly lowers the intergenerational educational mobility of foreign youth in Switzerland. In the canton of Ticino, for example, where selection does not occur until after eighth grade, the intergenerational educational mobility among foreign youth is much higher.

A further cause of the low academic path of foreign youth appears to be statistical discrimination. Research presently being carried out at my institute at the University of Basel points in this direction.[22] Statistical discrimination arises in the educational system when the school authorities – for lack of more objective information – use proxies for cognitive ability in placing students of different capabilities into different levels of course study. Nationality could serve as such a proxy as foreign students in Swit-

[21]Cf. OECD (2001).
[22]See Bauer (2006).

zerland score lower on the PISA test on average than observably identical natives. Note that statistical discrimination differs from other forms of discrimination in that it is statistically fair on average. Any injustices arising from incorrect placement result from inadequate information and not from personal prejudices.[23]

Our research indicates that school grades, upon which school placement is based in Switzerland, are a poor predictor of cognitive ability as measured by PISA test scores, showing that the school authorities are indeed subject to informational uncertainty when making their placement decisions on the basis of school grades. Secondly, we find that the predictive ability of grades improves significantly when the nationality of a student is additionally taken into account. Hence it is statistically fairer to include a student's nationality when trying to assess his or her cognitive abilities on the basis of school grades. And finally we discover that teachers do not discriminate against foreigners in setting grades. Thus the lower educational placement of foreigners is not due to personal prejudices, as some Swiss fear.

A simple means of eliminating statistical discrimination of course exists. The solution is to introduce mandatory standardized achievement tests nationwide. Achievement test scores would eliminate quality uncertainty allowing the authorities to place students in accordance with their true cognitive capabilities.

The benefits of standardized achievement test scores are not limited to the elimination of statistical discrimination in school placement, however. As Hanushek's survey clearly points out, achievement test scores also provide the requisite empirical basis for measuring accurately the economic benefits of education, which according to his survey are substantial, and for determining which educational policy instruments are the most effective in which settings. The availability of achievement test scores is also essential for assessing the economic efficiency of the educational system and thus to ensure that scarce educational resources are being put to full use. In short, test score information is vital for forging an educational policy intended to serve the interests of both the instructed and society as a whole. That is to my mind the central message that Hanushek's chapter has for European educational policymakers. Hopefully it will be heeded.

[23]Cf. Phelps (1972) or Aigner and Cain (1977).

References

Aigner D, Cain G (1977) Statistical theories of discrimination in labor markets. *Industrial and Labor Relations Review* 30:175–187

Bauer P (2006) Are school grades a good signal for skills? Industrial Organization and Labor Market Research Unit (FAI), (Mimeo), University of Basel

Bauer P, Riphahn R (2005) Timing of school tracking as a determinant of intergenerational transmission of education. *Economics Letters* 91:90–97

Hedges L, Laine R, Greenwald R (1994) Does money matter? A meta-analysis of studies of the effects of differential school inputs on student outcomes. *Educational Researcher* 23:5–14

Krueger A (2003) Economic considerations and class size. *Economic Journal* 113:F3–F33

OECD (2001) *Employment Outlook*. Paris

Phelps E (1972) The statistical theory of racism and sexism. *American Economic Review* 62:659–661

Todd P, Wolpin K (2003) On the specification and estimation of the production function for cognitive achievement. *Economic Journal* 113:F3–F33

Wössmann L (2004) How equal are educational opportunities? Family background and student achievement in Europe and the United States. CESifo Working Paper 1162, University of Munich