The Economic Value of Improving Local Schools

Eric A. Hanushek

Local civic leaders intuitively understand that education is good and that the quality of schools may in one way or another relate to local development. The arguments, however, tend to be general. They are linked only imprecisely to the impacts of schooling on the economy and to ways of improving the schools. This paper discusses what is known about the economic value of better schools and then puts those values into the perspective of school reform actions—particularly actions to improve the quality of teachers.

One important aspect of the discussion is how educational reform fits into notions of local economic development. What we know about the economics of school quality fits more into discussions of national outcomes, which may differ from local outcomes. An attempt is made to put this into the context of a more local economy.

The findings about the importance of school quality are particularly relevant in the context of U.S. accountability policies that emphasize performance on standardized tests in core areas. Some people have suggested that the achievement emphasized by current state accountability systems is not very important and that other aspects of student performance—creativity, the ability to work in teams, or personality traits—should be the focus of attention. While these other aspects are undoubtedly valuable, the analysis here strongly affirms an emphasis on basic cognitive skills by demonstrating its substantial economic returns.

Most consideration of the economic aspects of education has naturally concentrated on school attainment, or the quantity of education. It is easy to calculate the economic return on such an investment—both the costs and benefits are fairly clear. Additionally, until recently, relatively limited data have been available on the quality of schools. Finally, there are great uncertainties about how to change quality and what it costs. Nonetheless, the policy issues today are ones of quality.

Two decades ago, the federal government released a report, *A Nation at Risk* (National Commission on Excellence in Education 1983), which identified some serious problems with school quality. While it precipitated an unbroken period of concern about U.S. schools, it did not lead to any substantial improvements in school quality (Peterson 2003).

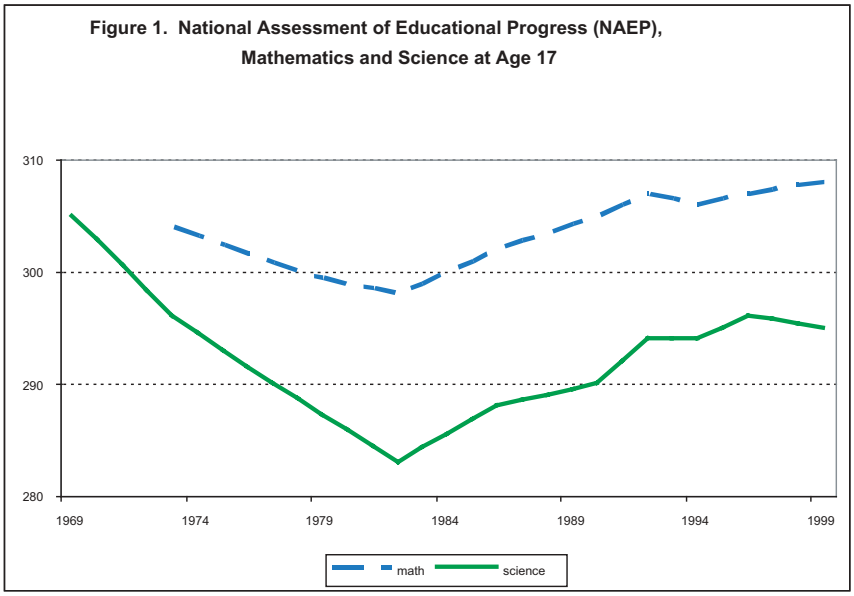
The benefits of reform are generally easier to estimate than the costs, although some information on costs is provided at the end. The central messages are: first, the economic impact of reforms that enhance student achievement will be very large. Second, reform must be thought of in terms of both the *magnitude* of changes and the *speed* with which any changes occur. Third, based on current knowledge, the most productive reforms are almost certainly ones that improve the quality of the teacher force. Fourth, such policies are likely to be ones that improve the hiring, retention, and pay of high quality teachers, that is, selective policies aimed at the desired outcome.

This discussion begins with a consideration of student achievement from varying perspectives. This discussion permits benchmarking the kinds of reforms and economic impacts that are relevant for policy deliberations.

U.S. STUDENT ACHIEVEMENT

The National Assessment of Educational Progress (NAEP) provides direct information on how student achievement has changed over time. It also points to substantial different performance by subgroups.

Figure 1 shows how performance of U.S. students has tracked over the past three decades in the critical areas of mathematics and science. At the end of high school, current students perform slightly better in math than those 30 years ago, but they perform noticeably worse in science. Not shown is the fact that reading scores over the same period are slightly up, and writing scores (only available for a portion of the period) are down. The summary statement is that student performance in the United States has been essentially flat for a long period of time.[1]



A second perspective on achievement is the disparity in scores across racial and ethnic subgroups. Figures 2 and 3 provide pictures of how the math and science performance of African American and Hispanic students compares to the performance of white students. The black–white gap has been very large, although there was some closing during the 1980s. The Hispanic–white gap also closed in the 1980s and went on to show further closing in the 1990s.

The racial and ethnic gaps remain very large. The figures have put the gaps in terms of standard deviations of individual test scores. Blacks fall almost one standard deviation behind whites, while Hispanics fall two-thirds of a standard deviation behind.

It is important to understand what such magnitudes mean, because the subsequent discussion of the economics of quality put scores into standard deviation units. A person who performs one standard deviation below the mean of the distribution will be at the 16th percentile. A person who performs one-half standard deviation below the mean will be at the 31st percentile of the distribution. (Similarly, an improvement of one-half standard deviation will take somebody at the middle of the distribution to the 69th percentile).

A final perspective on current student achievement is found in the distribution of performance across districts. During 2003, NAEP testing provided a finer geographic breakdown for mathematics performance in grade eight. While students in Ohio and the entire midwestern region performed slightly above the national average, performance in Cleveland was almost one standard deviation behind the nation. This partly reflects the heavily minority population in Cleveland, with 72 percent of the NAEP students being black. The white population in Cleveland, however, also scored some two-thirds of a standard deviation below white eighth-graders in the nation as a whole.

The next section translates these scores into economic terms.

BENEFITS OF ENHANCED SCHOOL QUALITY

Economists have devoted considerable attention to understanding how human capital affects a variety of economic outcomes. The underlying notion is that individuals make investment decisions in themselves through schooling and other routes. The accumulated skills that are relevant for the labor market from these investments over time represent an important component of the human capital of an individual. The investments made to improve skills then return future economic benefits in much the same way that a firm's investment in a set of machines (physical capital) returns future production and income. In the case of

Figure 2. White-Black Differences in NAEP Scores, 17-year-olds

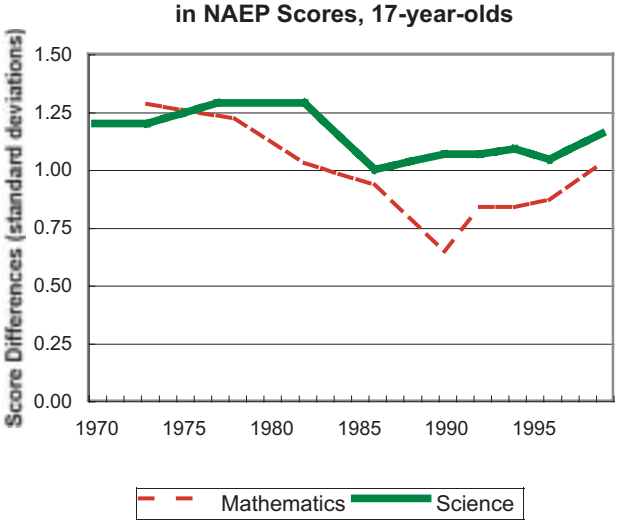
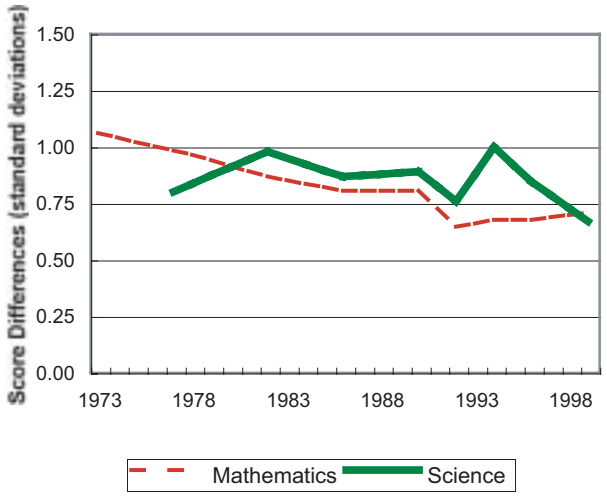


Figure 3. White-Hispanic Differences in NAEP Scores, 17-year-olds



public education, parents and public officials act as trustees for their children in setting many aspects of the investment paths.

In looking at human capital and its implications for future outcomes, economists are frequently agnostic about where these skills come from or how they are produced. Although we will return to that below, it is commonly presumed that formal schooling is one of 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. For this reason, we frequently emphasize the role of schools.

The human capital perspective immediately makes it evident that the real issues are ones of long-run outcomes. Future incomes of individuals are related to their past investments. It is not their income while in school or their income in their first job. Instead, it is their income over the course of their working life.

The distribution of income in the economy similarly involves both the mixture of people in the economy and the pattern of their incomes over their lifetime. Specifically, most measures of how income and well-being vary in the population do not take into account the fact that some of low-income people have low incomes only because they are just beginning a career. Their lifetime income is likely to be much larger as they age, gain experience, and move up in their firms and careers. What is important is that any noticeable effects of the current quality of schooling on the distribution of skills and income will only be realized years in the future, when those currently in school become a significant part of the labor force. In other words, most workers in the economy were educated years and even decades in the past—and they are the ones who have the most impact on current levels of productivity and growth, if for no reason other than that they represent the larger share of active workers.

Individual Incomes

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 were 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 have documented that the earnings advantages to higher achievement on standardized tests are quite substantial. 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 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 a one standard deviation increase in mathematics performance at the end of high school translates into 12 percent higher annual earnings.[3] The impact of one-half standard deviation in test performance is illustrated in figure 4, which builds on the level of median annual earnings for workers in 2001. By way of summary, median earnings, while differing some by age, were about $30,000, implying that a one-half standard deviation increase in performance would boost these by $1,800 for each year of work life. Mean incomes were about $40,000, suggesting that a one-half standard deviation translates into $2,400 per year of average earnings. The full value to individual earnings and productivity is simply the annual premium for skills integrated over the working life. If we accumulate this mean earnings gain over a lifetime and calculate the value at high school graduation, we find that a one-half standard deviation improvement adds an expected $40,000 in earnings for each student.[4]

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.[5] Second, the labor market experiences that are observed begin 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.[6]

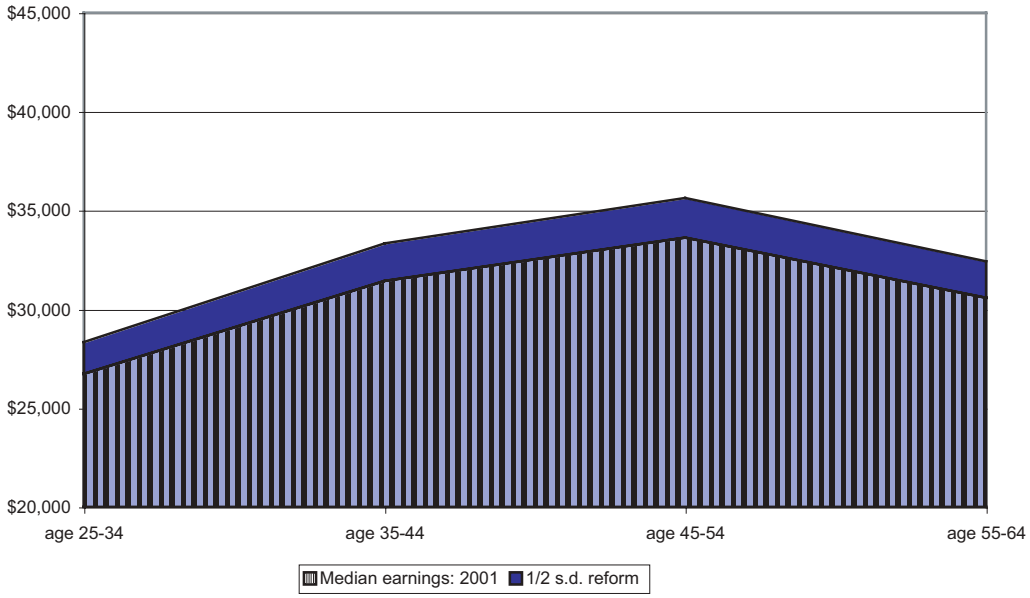
Another part of the return to school quality comes through continuation in school. There is substantial U.S. evidence that students who do better in school, either through grades or scores on standardized achievement tests, tend to go farther in school. Murnane et al. (2000) separate the direct returns 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. (Figure 1 is just the direct effects of skills, not including the indirect effects coming through added 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.

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.

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

Figure 4. Median U.S. Individual Earnings with Moderately Strong Reform



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 current economic position of the United States 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 (see Barro and Sala-I-Martin 1995).

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 the 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.

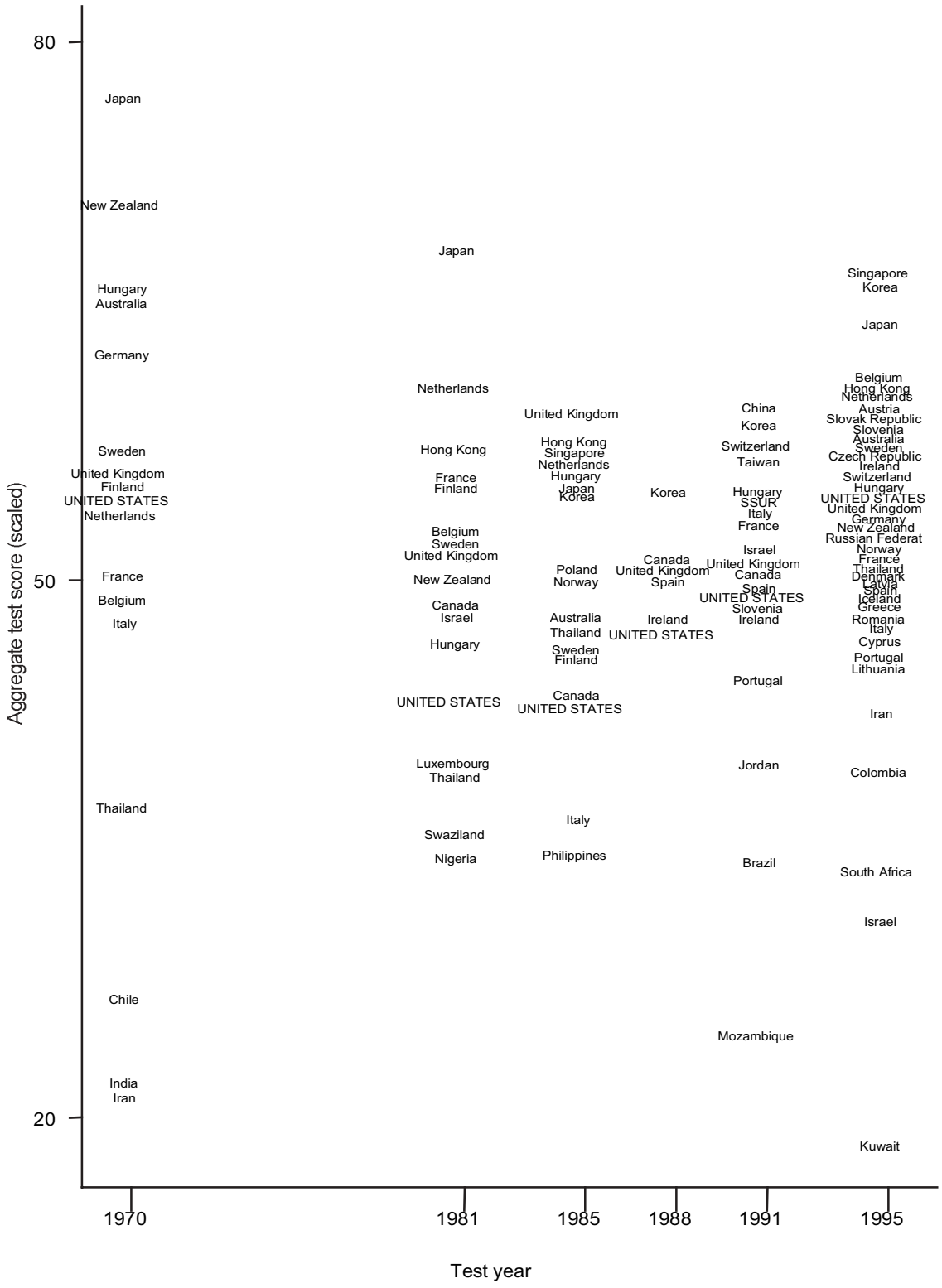
Hanushek and Kimko (2000) go beyond simple quantity of schooling and delve into quality of schooling. We incorporate the information about international differences in mathematics and science knowledge that has been developed through testing over the past four decades, and we find a remarkable impact 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 Educational Achievement (IEA) administered the first of a series of mathematics tests to a voluntary group of countries. These initial tests suffered from a 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.[7]

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. An easy summary of the participating countries and their test performance is found in figure 5. This figure tracks performance aggregated across the age groups and subject area of the various tests and is scaled to a common test mean of 50.[8] 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 is consistent with other data. NAEP performance over this period, shown previously in figure 1, also exhibits a sizable

Figure 5. Normalized test scores on mathematics and science examinations, 1970–1995



dip in the seventies, a period of growth in the eighties, and a leveling off in the nineties.

This figure also highlights a central issue here. The United States has not been competitive on an international level. It has scored below the median of countries taking the various tests. Moreover, this figure—which combines scores across different age groups—disguises the fact that U.S. performance is much stronger at young ages but falls off dramatically at the end of high school (Hanushek 2003).

Kimko 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 to 1990. 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.

Most important, the quality of the labor force as measured by math and science scores is extremely important. A one standard deviation difference on test performance is related to 1 percent difference in annual growth rates of gross domestic product (GDP) per capita.[9] A series of separate tests addresses the issue of whether the effect of quality is causal, a question frequently asked about international growth comparisons. Each test is consistent with a causal interpretation.[10]

This quality effect, while possibly sounding small, is actually very large and significant. Because the added growth compounds, it leads to powerful effects on U.S. national income and on societal well-being.

To underscore the importance of quality, it is possible to simulate the effects of alternative reforms of U.S. schools. As a benchmark, consider a policy introduced in 2005 that leads to an improvement of scores of graduates of one-half standard deviation by the end of a decade. This change, labeled a "moderately strong reform," would be substantial. An improvement of that magnitude would put U.S. student performance closer to that of students in a variety of better-performing European countries, but they still would not be at the top of the world rankings. (It does, however, have a similar lofty goal to that of the governor's summit in 1989 that set a goal of being first in the world in math and science by 2000—a goal that we did not dent during the 1990s.)

Such a path of improvement would not have an immediately discernible effect on the economy, because new graduates are always a small portion of the labor force, but the impact would mount over time. If past relationships between quality and growth hold, GDP in the United States would end up 4 percent higher by 2025 and 10 percent higher by 2035.

This kind of change may or may not be feasible, but the impact on GDP illustrates the real importance of effective school reform. To give some idea of the range of possible outcomes, figure 6 traces out improvements in the national economy from slower and lesser changes in student outcomes.

Figure 6 uses the goal of a one-half standard deviation improvement in performance but aims to achieve this over different time periods ranging from 10 to 30 years. A 30-year reform plan would still yield a gain to the economy in 2035 of 3 percent.

The summary of this analysis is that improvements in schooling outcomes are likely to have very powerful impacts on individuals (the previously identified effect on earnings) and on the economy as a whole. The impact on the aggregate economy will raise the whole economy over and above the individual differences estimated above.

Local Impacts

The prior estimates all place reform in a national context. The gains are not necessarily the same as those that would accrue to the local and regional economy from school quality improvements.

To be concrete, we noted that Cleveland students fell almost one standard deviation below the nation in math performance. If we could increase performance in Cleveland by the moderately strong reform amounts discussed above (that is, by one-half standard deviation), what would we expect to see?

We would expect to see the students leaving the Cleveland public schools to do better over their lifetimes. Today, we expect them to be hurt by the Cleveland schools, and this reform would bring them closer to the average for the nation.

Part of the gains would undoubtedly come through moving to other areas, implying that the overall impact on the Cleveland and Ohio areas might well be below

that of the nation as a whole. Ohio would have contributed to the nation, but it might not directly capture the higher earnings and productivity, because a portion of earnings growth for individuals comes from seeking out areas where they are the most productive.

Nonetheless, recent work on income and productivity differences across cities argues that educated cities have grown more quickly than comparable cities for more than a century (Glaeser and Saiz 2003). This analysis further suggests that the reason for greater growth is that skilled cities become more productive.

No data currently permit analysis of how quality enters into this, but there is every reason to believe that improved quality will confer gains on metropolitan areas and states. As with early work on cross-country growth differences, this analysis (and the others upon which it builds) focuses entirely on years of schooling as a measure of human capital differences across areas. Yet the arguments behind these empirical findings are ones that emphasize how local economies with more skilled workers can adjust to changing circumstances (see Welch 1970; Schultz 1975). These seem to be attributes that, as the individual earnings models and international growth models confirm, are fostered by more skills as directly measured by achievement.

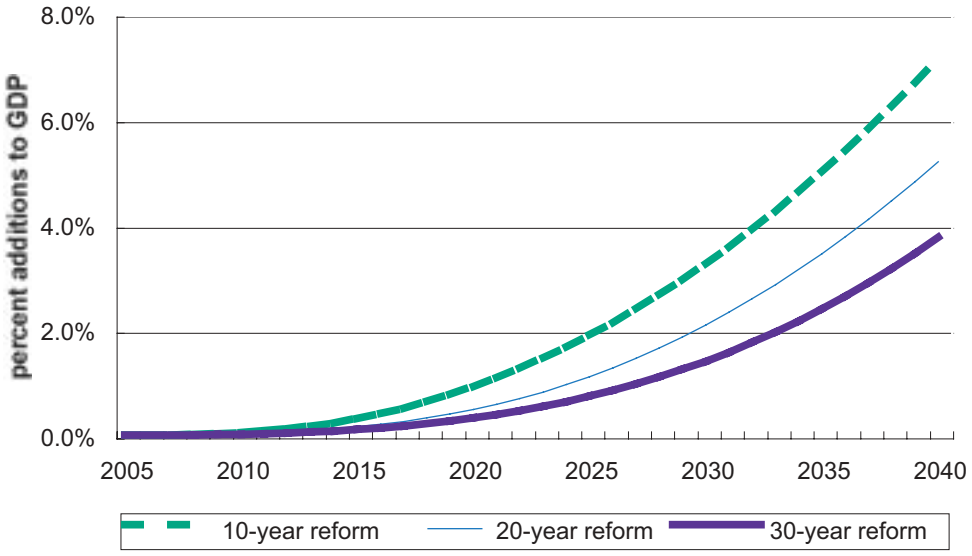
FEASIBLE TEACHER QUALITY POLICIES

The prior analysis has simply projected the benefits of achieving various goals for student achievement. A first question is whether or not achieving such gains could be feasible with realistic reform strategies.

Past reform efforts clearly do not support feasibility. During the two decades since the publication of *A Nation at Risk*, a variety of approaches have been pursued (Peterson 2003). These have involved expanding resources in many directions, including increasing real per pupil spending more than 50 percent. Yet performance has remained unchanged since 1970 when we started obtaining evidence from NAEP (figure 1).

The aggregate picture is consistent with a variety of other studies indicating that resources alone have not yielded any systematic returns in terms of student performance (Hanushek 2003). The character of reform efforts can largely be described as "same operations with greater intensity." Thus, pupil–teacher ratios and class size have fallen dramatically, teacher experience has increased, and teacher graduate degrees have grown steadily—but these have not translated into higher student achievement. On top of these resources, a wide variety of programs have been introduced with limited aggregate success. The experience of the past several decades

Figure 6. Improved GDP with Moderately Strong Knowledge Improvement



vividly illustrates the importance of true reform, that is, reform that actually improves student achievement.

One explanation for past failure is simply that we have not directed sufficient attention to teacher quality. By many accounts, the quality of teachers is the key element to improving student performance. But the research evidence suggests that many of the policies that have been pursued have not been very productive. Specifically, while the policies may have led to changes in measured aspects of teachers, they have not improved the quality of teachers when identified by student performance.[11]

Rivkin, Hanushek, and Kain (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, Kain, and Rivkin 2004). Thus, from an analytical viewpoint, it is difficult to sort out the quality of the teacher from the quality of the students that she has in her classroom. The analysis of teacher performance goes to great lengths to avoid contamination from any such selection and matching of kids and teachers.[12] In the end, it estimates that the differences in annual achievement growth between an average and a good teacher are at least 0.11 standard deviation of student achievement.[13]

Before going on, it is useful to put this estimate of the variation in quality into perspective. If a student had a good teacher as opposed to an average teacher for five years in a row, the increased learning would be sufficient to close entirely the average gap between a typical low-income student and a student not on free or reduced lunch. The earlier discussion also points to the possibility of closing existing ethnic gaps or of bringing our urban centers, such as Cleveland, up to the levels found in the nation.

A reasonable estimate (which is used throughout the following calculations) is actually that differences in quality are twice the lower bound (0.22 standard deviation.). This larger estimate reflects likely differences in teacher quality among schools (plus a series of other factors that bias the previously discussed estimate downwards).

These estimates of the importance of teacher quality permit some calculations of what would be required to yield the reforms discussed earlier. To begin with, consider what kinds of teacher policies might yield a 0.5 or a 1.0 standard deviation improvement in student performance. Obviously an infinite number of alternative hiring plans could be used to arrive at any given end point. A particularly simple plan is employed here to illustrate what is required.

Consider a steady improvement plan where the average new hire is maintained at a constant amount better than the average teacher in any given year. For example, the average teacher in the current distribution is found at the 50th percentile. Consider a policy where the average of the new teachers hired is set at the 56th percentile and where future hires continue to be at this percentile each year of the reform period. By maintaining this standard for replacement of all teachers exiting teaching (6.6 percent annually in 1994–95) but retaining all other teachers, this policy would yield a 0.5 standard deviation improvement in student performance after a 20-year period. If, instead, we thought of applying these new standards to all teacher turnover (exits plus the 7.2 percent who change schools), a 0.5 standard deviation improvement in student performance could be achieved in 10 years.

Figure 7 displays the annual hiring improvement that is necessary to achieve a moderately strong (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 such a 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 demonstrate the challenge of achieving substantial improvements in achievement. It requires significantly upgrading the quality of the current teacher force.

Several aspects of these scenarios deserve note. First, the improvements that are required apply to the teacher distribution that exists each year. In other words, this standard requires continual improvement in terms of the current teachers. The continual improvement comes from the fact that the distribution of teachers improves each year because of the higher-quality teachers hired in prior years. At the same time, it does not imply that all new teachers reach these levels, only that the average teacher does. There will still be a distribution of teachers in terms of quality.

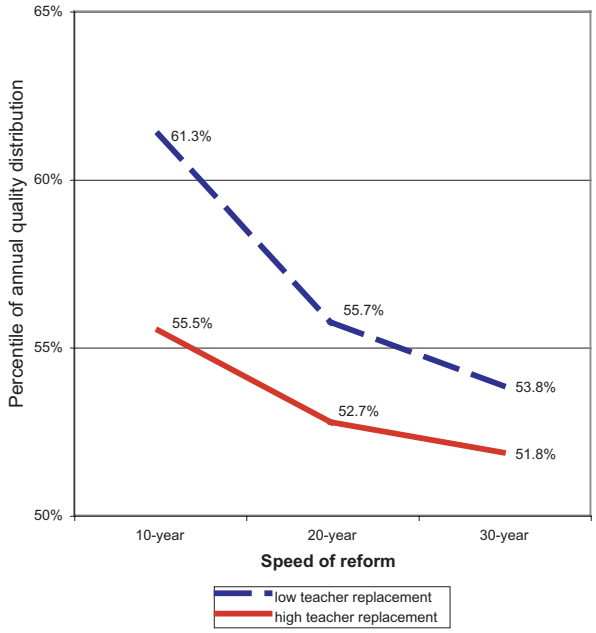
In fact, it is easy to summarize what the distribution of teachers must look like in terms of the current distribution of teachers. In order to achieve a 0.5 standard deviation improvement in student achievement, the average teacher (after full implementation of reform) must be at the 58th percentile of the current distribution. (In order to achieve a 1.0 standard deviation improvement, the average teacher must be at the 65th percentile of the current distribution). The annual adjustments given previously simply translate these quality calculations into the path required for reaching them under different reform periods.

The calculations also freeze many aspects of teaching. They assume no change in teacher turnover. Of course, teacher turnover will be affected by a variety of other policies such as salary policy, tenure, etc.

The calculations also assume that turnover is unrelated to quality—as it largely is with today's passive teacher management approach. An active selection and teacher retention policy could, however, lead to improvements in overall teacher quality would offer relief from the stringency of hiring standards that are required. For example, a policy that retained the best teachers two years longer and dropped the least effective teachers two years sooner would by itself lead to substantial improvements in the average quality of the teacher force.

The required improvements in the teaching force could also be achieved in other ways, at least conceptually. For example, a new professional development program that boosts the quality of current teachers would accomplish the same purpose. However, any such program must be in addition to the current amount of professional development, including obtaining master's degrees and completing in-service training, because the existing professional development activities are already reflected in the current quality distributions.

Figure 7. Teacher Quality Hiring Percentiles for Moderately Strong Improvement in Student Achievement



COST CONSIDERATIONS

Analyzing reform policies directly in terms of their costs is not feasible because we know very little about the supply function for teacher quality. While there has been some work on the cost of hiring teachers with different characteristics (such as experience or advanced degrees), these characteristics do not readily translate into teacher quality (Hanushek and Rivkin 2004).

Much of the current discussion of teacher quality is centered on statements about the overall level of salaries. It seems clear that teacher salaries have slipped relative to alternative earnings of college workers, particularly for women (Hanushek and Rivkin 1997, 2004).[14] For a variety of reasons, however, this does not give much policy guidance for the current discussions. In simplest terms, we do not know how teacher quality responds to different levels of salaries (Hanushek and Rivkin 2004). Moreover, policies that simply raise salaries across the board (even if advanced as a way to increase the attractiveness of the profession) would almost certainly slow any reform adjustments, because they would lower teacher turnover and make it more difficult to improve quality through new hiring.

The aggregate growth numbers suggest that the annual growth dividend from an effective reform plan would cover most conceivable program costs over a relatively short period of time. For example, a 10-year reform plan that yielded a one-half standard deviation improvement in student performance would produce an annual reform dividend that more than covered the *entire* expenditure on K–12 education before 2030.[15] Of course, as shown previously, a reform program of this magnitude and speed would require dramatic changes in hiring of new teachers. But a 20-year reform program with a moderately strong improvement would produce a sufficient dividend to cover all K–12 expenditures by 2035.

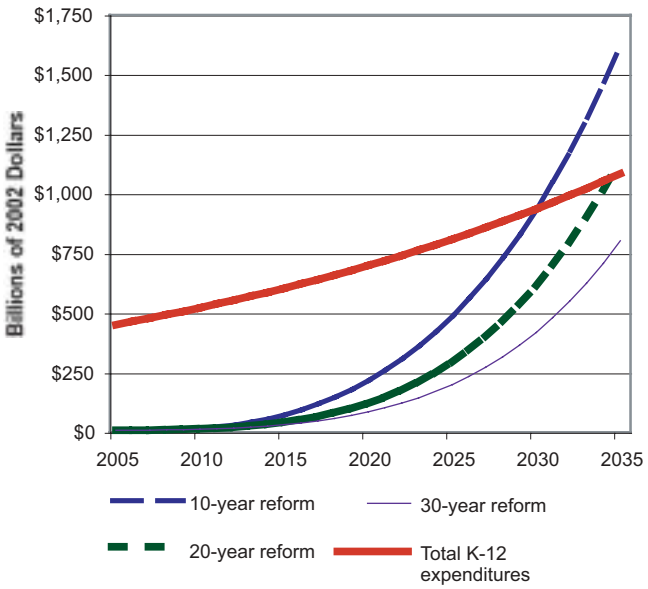
Figure 8 traces out the growth dividend relative to the total education budget for the United States. Educational expenditure for K–12 is calculated to grow at a real 3 percent annually, and the growth dividend of a moderately strong (0.5 standard deviation) reform plan (of varying speed) is plotted against this. This figure shows vividly how true reform (that is, reform that actually yields improvement in student performance) has a cumulative effect on the economy.

The conclusion of the cost considerations is simple. The benefits from quality improvements are very large. Thus, they can support incentive programs that are quite large and expansive *if the programs work*. U.S. schools have in fact expanded in a variety of ways over the past four decades—real expenditures per pupil in 2000 were more than three times those in 1960. It is just that these past programs have not led to significant improvements in student performance. Put another way, the benefits do not justify all types of expenditure. They do justify many conceivable programs if they can be shown to be effective.

CONCLUSIONS

The prior analysis demonstrates that better student outcomes generate considerable benefits. While these benefits have not been previously quantified, the presumption that they exist has surely propelled much of the interest in our schools that has existed at least since the publication of *A Nation at Risk*.

Figure 8. Annual Growth Dividend from Moderately Strong Reform



These findings are particularly relevant to current attention to school outcomes. The federal No Child Left Behind law requires states to institute accountability systems that ensure all students are proficient in core subjects. These accountability systems emphasize measured cognitive skills of just the kind that are shown to have high payoffs in the labor market and for society.[16]

Further, there is substantial reason to believe that improvement in local schools will yield direct benefits to local economies. Local economies with a more educated labor force leads, by existing analysis, to higher local growth. Although not explicitly analyzed in existing work, it is plausible to believe that school quality improvements will lead to local economic gains.

A part of the picture, however, that has not received as much attention is what is required to achieve the student outcome gains. This analysis uses available information about the current distribution of teacher quality to sketch out the kinds of changes that would be required for reform programs of differing magnitude and speed. This analysis highlights the fact that reform will require a significant upgrading of the teaching force. It also discusses feasible timing and speed of reform.

The benefit picture indicates that improvements in student performance have truly substantial impacts on individual productivity and earnings and on the growth and performance of the aggregate economy. The economic gains could in fact cover some substantial changes in expenditure on schools.

Past history, however, provides a key caution. The U.S. has devoted substantial attention to its schools. In just the two decades since *A Nation at Risk*, the nation has increased real spending on schools by over 50 percent. But it has gotten little in terms of student outcomes. We have accumulated considerable experience on things that do not work, but much less on policies that will succeed.

The available evidence does indicate that improvement in the quality of the teacher force is central to any overall improvements. And improving the quality of teachers will almost certainly require a new set of incentives, including selective hiring, retention, and pay.

ENDNOTES

[1]A variety of other factors have changed over this long period. Although it is difficult to assess the importance of these changes, little evidence suggests that these changes have had a large impact on the achievement trends (Hanushek 2003).

[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). The various adjustments for ability differences typically make small differences on the estimates of the value of schooling, and Heckman and Vytlacil (2001) argue that it is not possible to separate the effects of ability and schooling.

[3]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 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. By way of comparison, estimates of the value of an additional year of school attainment are typically 7–10 percent.

[4]These present-value calculations assume that the future is discounted at a real 5 percent rate over a working career of 35 years.

[5]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.

[6]These estimates, as highlighted in figure 4, 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, the impact of improvements in student skills are likely to rise over the work life instead of being constant, as portrayed here.

[7]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.

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

[9]The details of this work can be found in Hanushek and Kimko (2000) and Hanushek (2003). 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.

[10]Questions about causality arise in studies of the quantity of schooling because countries that grow and become richer may decide to spend some of their added income on more schooling. The tests in Hanushek and Kimko (2000) involve (1) investigation of international spending differences and test performance; (2) consideration of performance of immigrants in the U.S. using the test score measures; and (3) exclusion of the high-scoring East Asian countries.

[11]For a review of the existing literature, see Hanushek and Rivkin (2004). This paper describes various attempts to estimate the impact of teacher quality on student achievement.

[12]To do this, it concentrates entirely on differences among teachers within a given school in order to avoid the potential impact of parental choices of schools. Moreover, it employs a strategy that compares grade level performance across different cohorts of students, so that the matching of students to specific teachers in a grade can be circumvented. As such, it is very much a lower-bound estimate on differences in teacher quality.

[13]For this calculation, a teacher at the mean of the quality distribution is compared to a teacher 1.0 standard deviation higher in the quality distribution (84th percentile), labeled a "good teacher."

[14]There is a current debate about how salaries of teachers compare to those in different professions (Podgursky 2003).

[15]These calculations assume K–12 expenditures growth of 3 percent (real), implying that the current $350 billion expenditure would grow the $777 billion in 2025.

[16]Although not the focus of this discussion, there is also evidence that such accountability plans are effective in promoting higher achievement (Hanushek and Raymond 2005).

REFERENCES

Altonji, Joseph G., and Charles R. Pierret. 2001. Employer Learning and Statistical Discrimination. *Quarterly Journal of Economics* 116(1): 313–50.

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

Card, David. 1999. Causal Effect of Education on Earnings. In T*he Handbook of Labor Economics*, ed. Orley Ashenfelter and David Card, 1801–63. Amsterdam: North-Holland.

Glaeser, Edward L., and Albert Saiz. 2003. The Rise of the Skilled City. Working Paper No. 10191, National Bureau of Economic Research.

Hanushek, Eric A. 2003a. The Failure of Input-Based Schooling Policies. *Economic Journal* 113(485):F64–98.

———. 2003b. The Importance of School Quality. In *Our Schools and Our Future: Are We Still at Risk?* ed. Paul E. Peterson, 141–73. Stanford, CA: Hoover Institution Press.

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

Hanushek, Eric A., and Margaret E. Raymond. 2005. Does School Accountability Lead to Improved Student Performance? *Journal of Policy Analysis and Management* 24(2):297–327.

Hanushek, Eric A., and Steven G. Rivkin. 1997. Understanding the Twentieth-Century Growth in U.S. School Spending. *Journal of Human Resources* 32(1): 35–68.

———. 2004. How to Improve the Supply of High Quality Teachers. In *Brookings Papers on Education Policy 2004*, ed. Diane Ravitch, 7–25. Washington, DC: Brookings Institution Press.

Hanushek, Eric A., John F. Kain, and Steve G. Rivkin. 2004. Why Public Schools Lose Teachers. *Journal of Human Resources* 39(2): 326–54.

Heckman, James J., and Edward Vytlacil. 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(1): 1–12.

Lazear, Edward P. 2003. Teacher Incentives. *Swedish Economic Policy Review* 10(2): 179–214.

Mulligan, Casey B. 1999. Galton versus the Human Capital Approach to Inheritance. *Journal of Political Economy* 107(6): S184–224.

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(4): 547–68.

National Commission on Excellence in Education. 1983. *A Nation at Risk: The Imperative for Educational Reform*. Washington, DC: U.S. Government Printing Office.

Peterson, Paul E., ed. 2003. *Our Schools and Our Future: Are We Still at Risk?* Stanford, CA: Hoover Institution Press.

Podgursky, Michael. 2003. Fringe Benefits. *Education Next* 3(3): 71–76.

Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. Teachers, Schools, and Academic Achievement. *Econometrica* 73(2):417–58.

Schultz, Theodore W. 1975. The Value of the Ability to Deal with Disequilibria. *Journal of Economic Literature* 13: 827–46.

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