The Economic Value of Improved Schools

Eric A. Hanushek¹

April 2019

¹ Hoover Institution, Stanford University; hanushek@stanford.edu.
The objective of this analysis is to describe what is known about the economic impact of a future improvement in U.S. schools. While ideas of improvement have been discussed for a long period of time, little overall improvement has been seen. The message here is simple: Not improving means dramatically different future economic wellbeing than could be obtained by having more competitive schools.

Much of the discussion of the economic impact of education focuses on individual earnings. As is well-known, individuals with more education will on average earn more during their lifetime. Much of this discussion considers just school attainment – i.e., completing high school or college. But this discussion generally ignores the impact of quality and the importance of cognitive skills.

It turns out that the U.S. has some of the highest economic returns to cognitive skills that are observed around the world. From recent analysis of the working age population in varying countries, the economic value to individuals with higher skills is very high in the U.S. (Hanushek, Schwerdt, Wiederhold, and Woessmann (2015)). This reward to skills can be seen also from the other side: the penalties for not having adequate skills in the U.S. are larger than in most other countries.

But this review concentrates more on the aggregate impacts, because this relates most closely to state and national policies. In particular, having a skilled workforce dramatically affects economic growth rates. And, growth rates dictate what the future economic wellbeing of the country will be.

**Aggregate Gains through Economic Growth**

Economists have considered economic growth for much of the last 100 years, but most studies remained as theory with little empirical work. Over the past three decades, economists linked analysis much more closely to empirical observations and by doing so rediscovered the importance of growth. An important element in describing why some countries have grown faster than others is the human capital of the population.

The existing empirical analysis of growth is now quite extensive. While this work has not always been successful, we believe that now-existing analysis resolves the most important uncertainties in understanding long run growth. Specifically, growth is directly and significantly related to the skills of the population, and prior uncertainty about these effects is largely related to measurement issues.

By far the most important determinant of economic growth is the knowledge capital, or the collective cognitive skills, of a country. Virtually all past economic analyses of the long run growth of countries have highlighted a role for human capital, but the validity and reliability of the empirical analysis has been open to question. The instability of many past estimates has been taken as evidence of misspecified relationships where omitted influences of other factors appear likely. Moreover, there is concern about reverse causality; i.e., growth causing schooling rather than the opposite. These prior concerns can now be satisfactorily answered once skills are correctly measured. Then the basic growth relationships can support a detailed analysis of the economic implications of improving on a nation’s knowledge capital.

---


Basic Estimates

This discussion begins with a very general view and then provides some details of how skills relate to growth. Because the subsequent economic analysis relies heavily on the estimates of growth models, it is useful to have an overview of these.

A country’s growth rate can be thought of as a function of the skills of workers and other factors that include initial levels of income and technology, economic institutions, and other systematic factors. Skills are frequently referred to simply as the workers’ human capital stock.

\[
growth = \alpha_1 \text{human capital} + \alpha_2 \text{other factors} + \varepsilon
\]  

(1)

This formulation suggests that nations with more human capital tend to continue to make greater productivity gains than nations with less human capital, although it is possible that the induced growth in productivity disappears over time.4

The empirical macroeconomic literature focusing on cross-country differences in economic growth has overwhelmingly employed measures related to school attainment, or years of schooling, to test the human capital aspects of growth models. It has tended to find a significant positive association between quantitative measures of schooling and economic growth.5

Nevertheless, these formulations are not very useful for consideration of the economic value of school improvement. To begin with, they introduce substantial bias into the picture of economic growth. Average years of schooling is a particularly incomplete and potentially misleading measure of education for comparing the impacts of human capital on the economies of different countries. It implicitly assumes that a year of schooling delivers the same increase in knowledge and skills regardless of the education system. For example, a year of schooling in Peru is assumed to create the same increase in productive human capital as a year of schooling in Japan.

Additionally, and central to this discussion, formulations relying on this measure assume that formal schooling is the only source of education and that variations in non-school factors have negligible effects on education outcomes and skills. This neglect of cross-country differences in the quality of schools and

---

4 A major difference of perspective in modeling economic growth is whether education should be thought of as an input to overall production, affecting the level of income in a country but not the growth rate in the long run (augmented neoclassical models as in Mankiw, Romer, and Weil (1992)) or whether education directly affects the long-run growth rate (endogenous growth models as, importantly, in Lucas (1988), Romer (1990), and Aghion and Howitt (1998)); see Acemoglu (2009), Aghion and Howitt (2009), Barro and Sala-i-Martin (2004), and Jones and Vollrath (2013) for textbook introductions. In terms of these major theoretical distinctions, our formulations combine key elements of both competing models. The fact that the rate of technological change and productivity improvement is directly related to the stock of human capital of the nation makes it an endogenous growth model. At the same time, by including the initial level of income among the control variables, our model does allow for conditional convergence, a leading feature of the augmented neoclassical approach.

5 To give an idea of the robustness of this association, an extensive empirical analysis by Sala-i-Martin, Doppelhofer, and Miller (2004) of 67 explanatory variables in growth regressions on a sample of 88 countries found that primary schooling was the most robust influence factor (after an East Asian dummy) on growth in GDP per capita in 1960-1996.
in the strength of family, health, and other influences is probably the major drawback of such a quantitative measure of schooling.

To see this, consider a standard version of an education production function as employed in a very extensive literature, where skills are expressed as a function of a range of factors:

\[
\text{human capital} = \beta_1 \text{schools} + \beta_2 \text{families} + \beta_3 \text{ability} + \beta_4 \text{health} + \beta_5 \text{other factors} + \nu
\] (2)

In general, human capital combines both school attainment and its quality with the other relevant factors including education in the family, labor market experience, health, and so forth.

Thus, while school attainment has been convenient in empirical work because of its ready availability across countries, its use ignores differences in school quality in addition to other important determinants of people’s skills. A more satisfying alternative is to incorporate variations in cognitive skills, which can be determined through international assessments of mathematics, science, and reading achievement, as a direct measure of the human capital input into empirical analyses of economic growth.

The focus on cognitive skills has a number of potential advantages. First, it captures variations in the knowledge and ability that schools strive to produce and thus relates the putative outputs of schooling to subsequent economic success. Second, by emphasizing total outcomes of education, it incorporates skills from all sources – including schools, families, and ability. Third, by allowing for differences in performance among students with differing quality of schooling (but possibly the same quantity of schooling), it opens the investigation of the importance of different policies designed to affect the quality aspects of schools. Fourth, it is practical because of the extensive development of consistent and reliable cross-country assessments.

The growth analysis summarized here relies on the measures of cognitive skills developed in Hanushek and Woessmann (2015a). Between 1964 and 2003, twelve different international tests of math, science, or reading were administered to a voluntarily participating group of countries. These include 36 different possible scores for year-age-test combinations (e.g., science for students of grade 8 in 1972 as part of the First International Science Study or math of 15-year-olds in 2000 as a part of the first PISA test). The assessments are designed to identify a common set of expected skills, which were then tested in the local language. Each test is newly constructed, until recently with no effort to link to any of the other tests. Hanushek and Woessmann (2015a) describe the construction of consistent measures at the national level across countries through empirical calibration of the different tests. These measures of

---

7 See Hanushek and Woessmann (2011) for a review. Note that there have been seven major international assessments since 2003. We emphasize the early assessments because they fit into our analysis of long run growth. In the analysis of economic impacts for countries, we rely on the subsequent testing.
8 By transforming the means and variances of the original country scores (partly based on external longitudinal test score information available for the United States), each is placed into a common distribution of outcomes. Each age group and subject is normalized to the PISA standard of mean 500 and individual standard deviation of 100 across OECD countries, and then all available test scores are averaged at the country level.
Knowledge capital for nations rely on the average (standardized) test scores for each country’s historical participation in the tests.

The test scores can be interpreted as an index of the human capital of the populations (and workforce) of each country. This interpretation is reasonable if a country’s scores have been stable across time, implying that estimates from recent school-aged populations provide an estimate of the older working population. If scores (and skills) do in fact change over time, some measurement error is clearly introduced. In fact, scores have changed some, but within the relevant period of observations through 2003 differences in levels across countries dominate any intertemporal score changes.9

Using the aggregate test scores for each country – its knowledge capital – equation 1 can be directly estimated with a refined measure of human capital. The estimation used here for assessing the economic value of school improvement looks at the association between educational outcomes and long-run economic growth in the sample of 50 countries for which we have both economic growth data and our measure of knowledge capital.10 The basic estimates include initial GDP per capita for each country simply to reflect the fact that it is easier to grow when one is farther from the technology frontier, because one just must imitate others rather than invent new things. The estimation also includes school attainment (average years of schooling) in 1960 and the test measure of human capital.

If knowledge capital is ignored, years of schooling in 1960 is significantly associated with average annual growth rates in real GDP per capita in 1960-2000.11 However, once the test measure of human capital is included, cognitive skills are highly significant while school attainment becomes statistically insignificant and its estimated coefficient drops to close to zero.

The estimated coefficient on cognitive skills implies that an increase of one standard deviation in educational achievement (i.e., 100 test-score points on a PISA scale) yields an average annual growth rate over 40 years of observation that is two percentage points higher. This historical experience suggests a very powerful response to improvements in educational outcomes, particularly when compared to the average 2.3 percent annual growth within sampled countries over the past two decades.

Perhaps the easiest way to see the relationship is to plot the marginal impact of knowledge capital on long run growth. Figure 1 depicts the fundamental association graphically, plotting growth in real per-capita GDP between 1960 and 2000 against average test scores after allowing for differences in initial GDP per capita and initial average years of schooling. Countries align closely along the regression line that depicts the positive association between cognitive skills and economic growth.

---

9 For the 50 countries in our growth analysis, 73 percent of the variance in scores lies between countries (Hanushek and Woessmann (2012)). The remaining 27 percent includes both true score changes and any measurement error in the tests. Any measurement error in this case will tend to bias downward the estimates of the impact of cognitive skills on growth, so that our estimates of economic implications will be conservative.

10 See Hanushek and Woessmann (2012), (2015a) for a more complete description of both the data and the estimation, which extends previous work by Hanushek and Kimko (2000).

11 To avoid the 2008 global recession, its aftermath, and any potential bubbles building up beforehand, the growth analysis stops in 2000, but results are very similar when extending the growth period to 2007 or 2009; see Hanushek and Woessmann (2015a), Appendix 3A.
Causality

The fundamental question is: should we interpret this tight relationship between cognitive skills and economic growth as a causal one that can support direct policy actions? In other words, if achievement were raised, would we really expect growth rates to go up by a commensurate amount? Knowing that the relationship is causal, and not simply a byproduct of some other factors, is very important from a policy standpoint. It is essential to be confident that, if a country managed to improve its achievement in some manner, it would see a corresponding improvement in its long-run growth rate. Said differently, if the relationship between test scores and growth rates simply reflects other factors that are correlated with both test scores and growth rates, a change in test scores may have little or no impact on the economy.

Work on differences in growth among countries, while extensive over the past two decades, has been plagued by legitimate questions about whether any truly causal effects have been identified, or whether the estimated statistical analyses simply pick up a correlation that emerges for other reasons.

The early studies that found positive effects of years of schooling on economic growth may have, indeed, been suffering from what is known as reverse causality, that is, improved growth was leading to more schooling rather than the reverse. If a country gets richer, it tends to buy more of many things, including more years of schooling for its population.

There is less reason to think that higher student achievement is caused by economic growth. For one thing, scholars have found little impact of additional education spending on achievement outcomes, so it is unlikely that the relationship comes from growth-induced resources lifting student achievement. Still, it remains difficult to develop conclusive tests of causality with the limited sample of countries included in our analysis.

Hanushek and Woessmann (2012) consider causality in detail, and the approaches are just summarized here. Suffice it to say, while some doubt could naturally be raised, there are strong reasons to believe that improve a country’s knowledge capital leads to higher long run growth.

First, the estimated relationship is little affected by including other possible determinants of economic growth. In an extensive investigation of alternative model specifications, Hanushek and Woessmann (2012) employ different measures of cognitive skills, various groupings of countries (including some that eliminate regional differences), and specific sub-periods of economic growth. These efforts show a consistency in the alternative estimates, in both quantitative impacts and statistical significance, that is uncommon in cross-country growth modeling. Moreover, measures of geographical location, political stability, capital stock, and population growth do not significantly affect the estimated impact of cognitive skills. These specification tests rule out some basic problems attributable to omitted causal

---

12 This section summarizes the detailed analysis found in Hanushek and Woessmann (2015a).
13 See, for example, Bils and Klenow (2000).
14 See the review in Hanushek and Woessmann (2011).
factors that have been noted in prior growth work. Of course, there are other possible omitted factors, leading to further analyses into the details of international differences.

Second, by focusing on key elements of school systems such as the degree of competition, the use of accountability systems, and relative salaries, it is possible to estimate the impact on growth of just achievement differences across countries that arise because of characteristics of the school systems. The results are the same as the base case.

Third, in a very direct test, the changes in test scores that have occurred over the full testing period (1964-2003) can be related to changes in annual growth rates for a subset of countries. This analysis shows a clear positive relationship.

Each approach to determining causation is subject to its own uncertainty. Nonetheless, the combined evidence consistently points to the conclusion that differences in cognitive skills lead to significant differences in economic growth. Moreover, even if issues related to omitted factors or reverse causation remain, it seems very unlikely that these cause all of the estimated effects.

Since the causality tests concentrate on the impact of schools, the evidence suggests that school policy can, if effective in raising cognitive skills, be an important force in economic development. While other factors – culture, health, and so forth – may affect the level of cognitive skills in an economy, schools clearly contribute to the development of human capital. More years of schooling in a system that is not well designed to enhance learning, however, will have little effect.

The Impact of School Improvement

The growth relationship portrayed in Figure 1 permits analysis of the likely economic impacts that would accrue from improving school performance. Assuming that the historical relationship of knowledge capital and growth holds into the future, it is possible to trace out the implications of improving schools. Of course, any improvements take time to be accomplished, and the economic impacts come over an even longer time horizon. Thus, it is important to pay attention to the dynamics of human capital improvement and to the time path of economic impacts.

The following calculations consider two changes that are assumed to occur over fifteen years; i.e., current students only move fully to higher achievement after 15 years of reform. The changes are that: 1) all 15-year-olds in each country are in school and brought up to minimal skill levels; and 2) the average student in each country is brought up to the level of the average student in Finland. The pace of student improvement is assumed to be linear.

Of course, improvement of students is also not the same as improvement in the labor force. The labor force improves only as young, more skilled students replace retiring less skilled workers. Calculations show how the average quality of the labor force changes by assuming that 2.5 percent of the labor force retires each year and is replaced by better educated workers. This implies that the labor force does not fully reach its ultimate quality for 55 years (15 years of reform followed by 40 years of retirements).

The annual growth of each nation is calculated in each year based on the average quality of the labor force in each given year. The analysis looks at the implications for GDP growth over an 80 year period,
reflecting the expected lifetime of somebody born at the beginning of reform. Given the extended period of labor force reform, the largest impacts clearly appear in the more distant future.

A simple graphical display of what improvement means for the U.S. economy can be seen by going back to prior ideas of improvement in the early goals of President George H.W. Bush. In 1989, the President and governors set the goal of being first in the world in math and science performance by 2000. From the relationship in Figure 1, it is possible to trace what that would have meant for added growth and higher GDP. Figure 2 shows projections of how much larger GDP would have been in each year had our schools improved and had our student achievement risen. With school improvement, the economic impacts take time to accrue – because the new, more skilled students must enter the labor force and must become a substantial enough portion to bring about overall improvements. But then the higher growth rate pushes the path of GDP up. The dashed horizontal line at 4.5 percent represents the total amount of GDP that goes to K-12 schools. It shows that – had the U.S. followed through on its achievement pledge – the added GDP would be sufficient by 2015 to cover all K-12 expenditures. Even if the period of improvement stretched out to 2010, the added GDP cover total K-12 expenditures by 2019.

It is also possible to consider a more general set of improvements. The U.S. falls slightly below the average scores found in the major developed countries of the world, those in the Organization for Economic Cooperation and Development or OECD. Since 2000, the OECD has been testing representative samples of 15-year-olds in math, science, and reading. These assessments, the PISA tests, provide a picture of the cognitive skills of different countries. Similar to the projections for being “first in the world,” the basic growth relationship can be used to see what bringing U.S. students up to the level of those in other countries would mean for the U.S. economy according to historical growth patterns.

In these projections the future impact on the economy is put into dollar terms and into how much average GDP for the U.S. would be lifted. In recognition of the fact that these play out over a long period of time, early gains are weighted more heavily than later gains by calculating present values using a discount rate for future years of three percent per year (implying that gains after 80 years, are weighted only 9 percent as heavily as initial year gains).

The projections in Table 1 assume that schools start an improvement plan in 2015 that takes 15 years to complete. The simple idea is to bring the average U.S. student up to the level of the average student in Germany, Canada, and Finland. Bringing all students up to minimal levels of skill is also considered in an added projection. This is similar to accomplishing the goals of No Child Left Behind, except that

15 PISA stands for the Programme for International Student Assessment. See, for example, OECD (2013).
16 A standard value of the social discount rate used in long-term projections on the sustainability of pension systems and public finance is 3 percent (e.g., Börsch-Supan (2000)), a precedent that is followed here. As a practical value for the social discount rate in cost-benefit analysis (derived from an optimal growth rate model), Moore et al. (2004) suggest using a time-declining scale of discount rates for intergenerational projects that do not crowd out private investment, starting with 3.5 percent for years 0-50, 2.5 percent for years 50-100. By contrast, the influential Stern Review report that estimates the cost of climate change uses a discount rate of only 1.4 percent, thereby giving a much higher value to future costs and benefits (Stern (2007)).
improvements would not be accomplished until 2030 instead of the 2014 anticipated in the NCLB Act. In all cases, these goals seem reasonable if there is a concerted move to improve the schools in the U.S.

The simplest policy scenario considers bringing U.S. students up to the level of Germany.\textsuperscript{17} This improvement (about 25 PISA points, or one-quarter standard deviation) would, by the historical growth relationships in Figure 1, yield present value of gains in GDP of $43.8 trillion. This sum provides the additions to GDP that would be expected compared to the normal economic growth with no improvement in skills. This value exceeds the 2015 total level of GDP of $16 trillion by a factor of almost three. Such values are, nonetheless, hard to make sense of. Thus, translating this into how much the level of average GDP for the remainder of the century would be increased compared to the no-change GDP one gets a 6.2 percent increase. The gains would, in other words, be large enough to take care of current U.S. deficits and could be used to solve both the Social Security and the Medicare imbalances.

Going further to the levels of Canada or of Finland shows extraordinarily large gains. Getting to the level of Canadian students would, for example, imply average gains of 11.4 percent in GDP – an amount roughly equal to average wage and salary increases for all workers over the remainder of the century of over 20 percent. Getting to the levels of Finland yields almost 50 percent larger economic gains than getting to Canadian levels.

Interestingly, while NCLB was derided as a foolhardy policy, bringing up the bottom of the distribution to minimal skills by 2030 would yield long run gains exceeding those of getting to Canadian levels. While it may be possible to question the precise structure of NCLB, there is no doubt that the goals would have had a tremendous economic impact had they been met.

\textbf{Conclusions}

At least from the publication of \textit{A Nation at Risk} in 1984 (The National Commission on Excellence in Education (1984)), the U.S. has discussed the need to improve the schools. At the same time a variety of forces has resisted the kinds of changes that might actually have accomplished change. The argument has essentially been “yes, improvement is needed, but it is so hard to change the schools.” The message of the prior economic calculations is that the cost of not changing is truly huge.

As Condoleezza Rice and Joel Klein have argued, the future of the U.S. as a world leader is highly dependent upon improving U.S. schools (Klein, Rice, and Levy (2012)). It is of course not an easy task both because of uncertainty about the best policies and because of resistance to change. But, if we do not improve the schools, the United States is likely to be left in a very compromised position.

\textsuperscript{17} Note that these comparisons are based on PISA in 2012. There have been some changes of the countries since then.
References


Figure 1. Cognitive Skills and Economic Growth, 1960-2000

Source: Hanushek and Woessmann (2015a)
Figure 2. “First in the World by 2000”

Source: Hanushek (2009)
Table 1. Value of U.S. Improvement From 2015 Baseline

<table>
<thead>
<tr>
<th></th>
<th>Equaling Germany</th>
<th>Equaling Canada</th>
<th>Equaling Finland</th>
<th>Achieving NCLB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present value (trillions of 2015 $’s)</td>
<td>43.8</td>
<td>82.2</td>
<td>111.9</td>
<td>86.2</td>
</tr>
<tr>
<td>% of future GDP</td>
<td>6.2</td>
<td>11.4</td>
<td>15.8</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Source: Hanushek, Peterson, and Woessmann (2013)