The Impact of Education on Economic Growth

by

Emil Chen

An honors thesis submitted in partial fulfillment

of the requirements for the degree of

Bachelor of Science

Undergraduate College

Leonard N. Stern School of Business

New York University

May 2014

Professor Marti G. Subrahmanyam	Dean Peter Henry
Faculty Adviser	Thesis Adviser

Abstract

Historically, many cross-country growth regressions have often incorporated some measurement of human capital as a determinant of economic growth. However, most previous literature uses attainment rates as a measure of education, which has proven to be inaccurate and statistically insignificant. Thus, this thesis develops an alternative measurement of human capital by incorporating a variable that controls for the *quality* of a country's education, not the *quantity*. This variable is measured by academic achievement, or international test scores. To prove the economic and statistical significance of this measure, this thesis first builds upon Hanushek (2012), one of the first publications that uses "cognitive skills," or international test scores, in cross-country growth regressions. We extend Hanushek (2012) to present-day data, which requires a series of replication studies and extension studies. After our work, we find that cognitive skills not only provide a more economically and statistically significant measure of education than attainment rates, but the close relationship between test scores and GDP per capita growth holds relatively closely over an extended period of time. Given this statistical analysis, we finally incorporate test scores into a cross-country regression conditional on a number of macroeconomic factors, thus estimating the true impact of education on economic growth. Our results indicate that an improvement in academic achievement can lead to a statistically and economically significant increase in GDP growth. Additionally, from past improvements in international test scores, we find it feasible for a country to significantly improve their economic growth prospects by focusing on policies that improve academic achievement.

Acknowledgements

I would first like to extend a big thank you to my thesis advisor, **Dean Peter Henry**, for his guidance throughout this past year. Not only was your advice integral to my growth in this process, but your intelligence, patience, generosity, and kindness have been truly inspirational. Thank you for always making time to meet with me each week, despite your busy schedule and enormous responsibilities. You have been an impeccable role model, and I am unbelievably grateful to have worked under your guidance.

To **Professor Marti Subrahmanyam,** for your dedication to the Honors Program and letting students explore their interests and passions.

To **my friends**, for their constant support throughout my undergraduate college career. All of you have been my motivation, pushing me to be the best I can be. Thank you all, and cheers to the past four years.

And lastly, to **my family**, for always being there for me. It goes without saying that I would not have ever had the opportunity to accomplish these things without you.

Table of Contents

I. Introduction

In today's modern world, education is often seen as an investment on both an individual and societal level. On the individual, micro level, the benefits of education are numerous. Education helps equip individuals with the necessary skill sets to pursue more technical jobs. It can play a significant role in training and expanding mental capacity, thus helping people increase their output in the workplace. In the most general sense, education as an investment helps an individual become a more productive member of society, which theoretically should lead to personal income growth.

The focus of this thesis, however, is more on the macro level and analyzing how education can impact society as a whole. If education can help an individual become more productive in the workplace and increase personal income, then how can an education system help an entire country become more productive and increase national income? As Aghion, Boustan, Hoxby, and Vandenbussche (2009) detail, policymakers often assert their willingness to invest in education. These policymakers rationalize that if their state spends more on improving their education system, incomes can grow sufficiently to more than recover that investment. In addition, their investment in education can also create several positive externalities, such as technological innovation. By educating a workforce to be more innovative, they can make capital and labor more productive, thus increasing overall output and generating greater income growth.

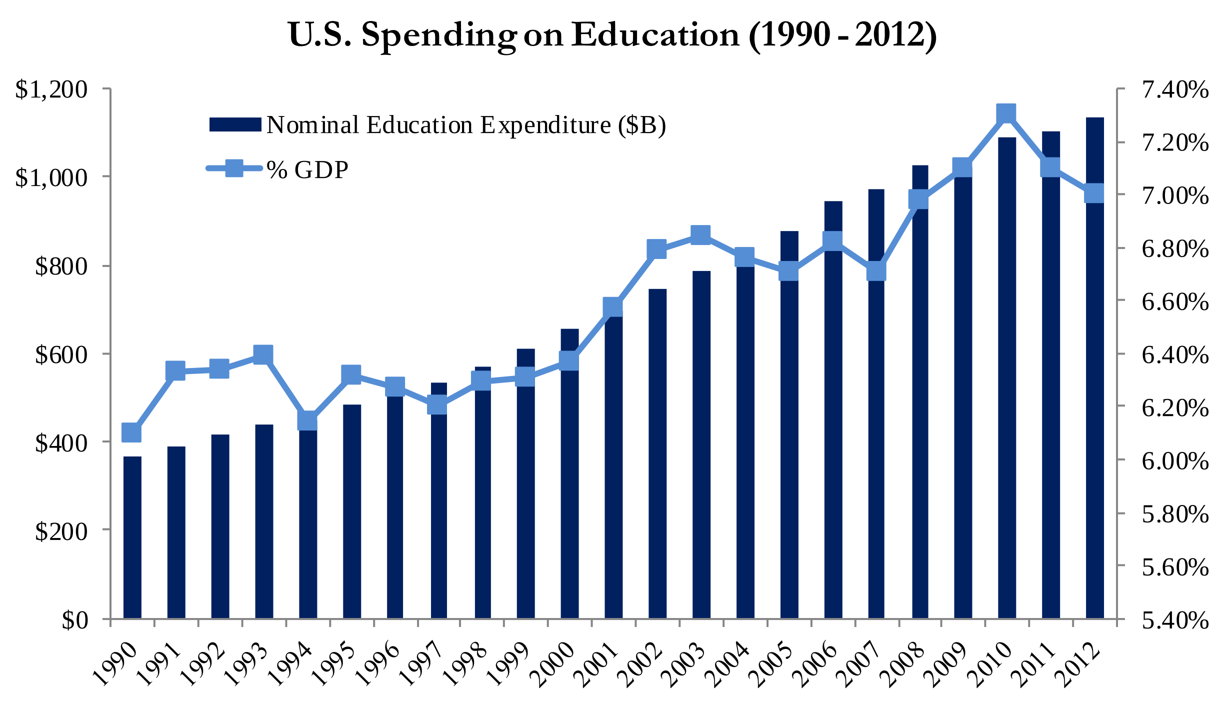
Hanushek and Woessmann (2010) build on this concept and outline three broad impacts that education can have on economic growth:

1. Education can increase the human capital inherent in the labor force, thus increasing labor productivity and output;

2. Education can increase the innovative capacity of the economy and lead to new technologies, products, and processes that can impact growth (akin to the positive externalities mentioned by Aghion, Boustan, Hoxby, and Vandenbussche [2009]); and
3. Education can facilitate the diffusion of knowledge to successfully implement new technologies devised by others, which again can promote economic growth.

Given this view that education drives growth, governments have responded by pouring their resources into education reform. Figure 1 below shows the amount of education spending in the U.S. alone since 1990 in both nominal figures and as a percentage of GDP.

Figure 1

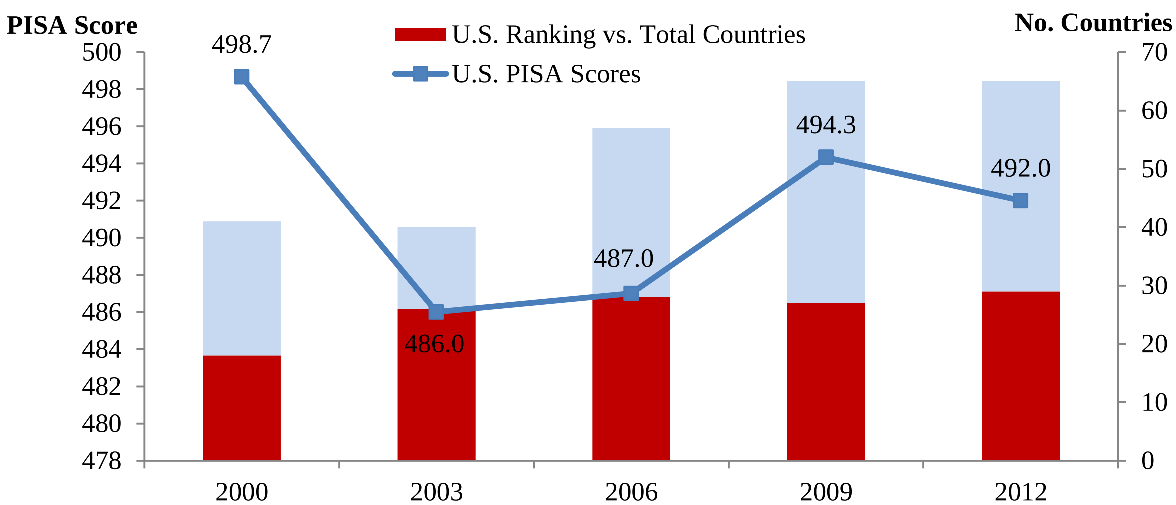


Source: OECD Education at a Glance 2013, U.S. Office of Management and Budget

As Figure 1 demonstrates, there is a clear upward trend in the amount of focus the United States has been placing on education (more than $8 trillion invested from 2003 – 2012). To put these numbers into perspective, the U.S. spent 7.3% of its GDP on education in 2010. Only five other countries topped them in this category – Denmark (8.0%), Iceland (7.7%), Korea (7.6%), Norway (7.6%), and Israel (7.4%). In addition, 7.3% of GDP is well above the OECD average of 6.3%. In pure nominal dollar terms, this equates to the highest spending of any country in the world. Even on a per student basis, the United States spends $15,171 USD, which is the highest of any country.[1]

Yet, despite spending more than any other country on a per student basis, the United States has shown flat, if not declining, trends in education results. Figure 2 depicts the United States' overall performance on the OECD-administered Programme for International Student Assessment (PISA), an international exam taken by 15 year old students that measures their abilities in mathematics, science, and reading.

Figure 2



Source: OECD PISA

[1] OECD "At a Glance" Study 2013

The scale on the left relates to the line graph, which represents the United States' average score on the PISA since 2000. The right-hand side scale relates to the bar graphs, which show the U.S. ranking on the PISA in out of the total number of participating countries. As Figure 2 shows, the United States' rank amongst participating nations has remained relatively flat over the years but recently declined to its lowest ranking yet on the 2012 exam (29 out of 65 participating countries). In addition, the U.S. has yet to break above the OECD average score of 500.

These bleak results coincide with the fact that the U.S. economy has seemed to stall since 2005, struggling to even break 3% year-over-year real GDP growth.[2] To compound these issues, unemployment continues to hover around a relatively high 7%, with increasing rhetoric that new graduates simply lack the skills that employers seek. This so called "skills gap" raises a few concerns. As mentioned previously, education should theoretically help individuals gain the proper skills to be more productive or perform more technical jobs. Yet, in reality, there still seems to be a mismatch where education is not accomplishing enough.

This begs the question, what is the true impact of education on economic growth? This thesis attempts to analyze this question in a few distinct parts. First, this thesis will explore different methodologies of measuring the impact of education on growth. To do so, it will delve into the previously published literature on the subject and examine what factors have previously been tested for causality. After analyzing this literature, this paper will then attempt to replicate and extend those studies with updated, present day data. Following a brief analysis, this thesis will conduct a regression of human capital on GDP growth, conditional on a number of control factors, to determine the true impact of education on growth. Lastly, given the results of our statistical analysis, we will conclude with a few policy recommendations.

---

[2] However, this statement must come with an asterisk. The financial crisis of '07-'08 hindered economic growth prospects for half a decade.

II. Education and Growth: Measurement Methodologies

II. i) Simple Growth Model

The link between education and economic growth first started as a small component of a much larger body of research. Beginning in the 1980s and 1990s, macroeconomists sought to explain why countries around the world experienced different rates of growth. To test their hypotheses, many of the initial experiments were performed using cross-country regressions under a simple growth model:

$$g_i = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \dots \beta_n x_{n,i} + \varepsilon \quad (1)$$

where g_i represents growth of a certain country, i, as a function; α represents the intercept term; $x_{1,i}$, $x_{2,i}$,… $x_{n,i}$ represent explanatory variables (again, of a certain country i); β_1, β_2… β_n represent unknown parameters to be estimated; and ε represents some estimation error.

Various studies ultimately chose to model g_i using a wide range of different explanatory variables. However, these studies would typically account for some measure of education. Robert Barro, who pioneered much of the initial work in this area of research, labeled this education factor as “initial human capital.” In his first paper, Barro (1991) substituted human capital with 1960 school-enrollment rates and measured its impact on real per capita GDP as a proxy for economic growth. The results demonstrated a strong, positive relationship between increases in human capital and subsequent GDP growth.[3]

II. ii) Attainment Rates

Although Barro used school-enrollment rates as an initial proxy for human capital, an ever increasing body of work chose to approximate human capital with a different measure:

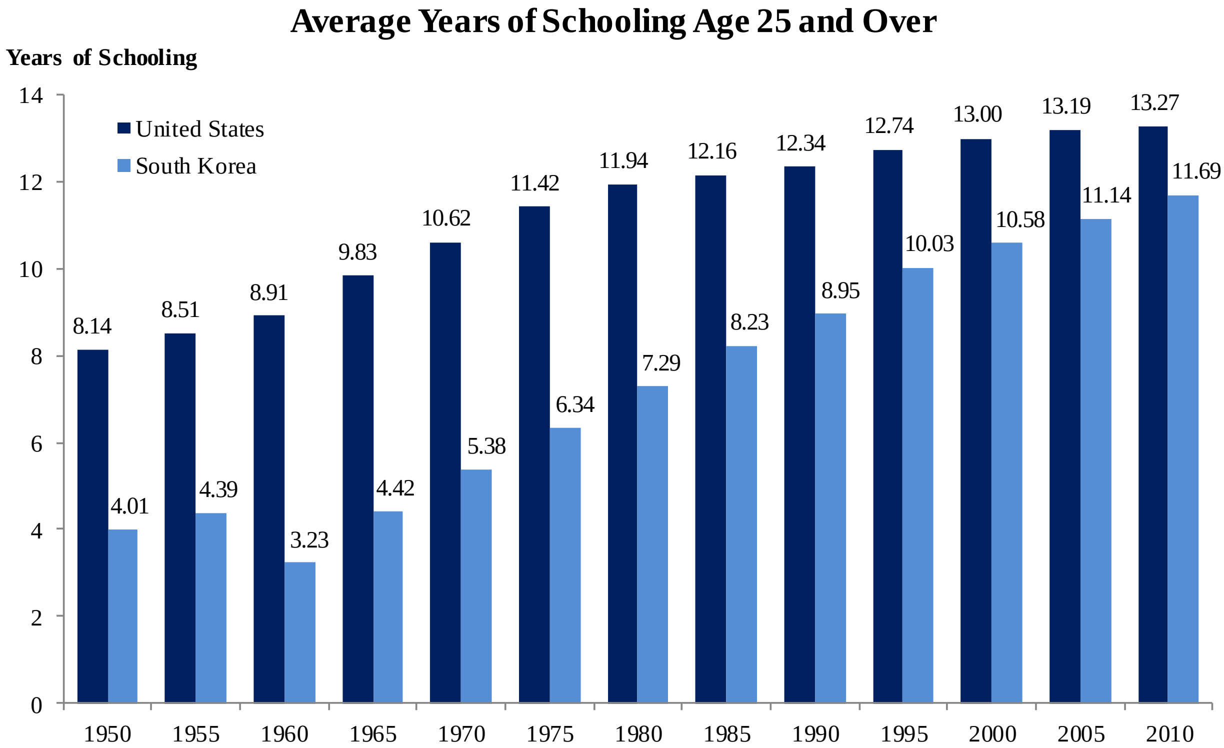
[3] In Barro (1991), as with all cross-country regressions in subsequent studies, regressions of explanatory factors against real GDP growth were performed with initial GDP held fixed. In this case, Barro (1991) held 1960 GDP fixed.

attainment rates. Jacob Mincer, who sought to explore human capital's effect on wage differentials, laid much of the foundation in this area of research. Mincer argued that a primary motivation for education was to develop the general skills required for work. Therefore, it would make sense to measure human capital not necessarily by a country's total enrollment, but rather by the amount of schooling completed by individuals (Mincer 1974). In his study, Mincer ultimately showed how wage differentials could be significantly explained by attainment rates, thus providing justification for their use. Attainment rate data also happened to be readily available across multiple countries, making them a very convenient variable to use in subsequent studies. Over time, researchers would continue to replicate Mincer's work across other countries – to the point where attainment rates eventually became synonymous with "human capital."

Barro, after his initial study, would produce similar papers on the determinants of economic growth; however, in his subsequent studies (e.g. Barro [1996] and Barro [1998]), Barro uses attainment rates as his proxy for human capital (or more specifically, attainment rates for males over the age of 25). Again, through cross-country regressions, Barro (1998) shows a significant, positive relationship between attainment rates and GDP growth. The estimated coefficient implies that an additional year of schooling could raise the growth rate by 0.7 percentage points per year. Published research from this point onwards mostly follows this existing empirical framework. For example, UNESCO conducted its own independent study in 2011 and found that an additional year of schooling could raise annual GDP by 0.37 percentage points.

To put Barro's figures into perspective, Figure 3 below highlights attainment rates for the population of the United States and South Korea aged 25 and over.

Figure 3



Source: Barro and Lee Data Set v 1.3

Looking at the United States at the time of Barro's study in 1998, the average years of schooling increased from 8.14 years in 1950 to approximately 13.00 years (in 2000). This means that over the course of 50 years, the United States was able to raise its average attainment rate by less than 5 total years, or less than a year per decade. By Barro's research, this equates to an increase in GDP per capita growth of approximately 0.7 percentage points per decade. While policymakers would certainly advocate for a 0.7 percentage point increase in GDP per capita, they must realize the time and difficulty it would take to raise the average attainment rate by just one year. In addition, increasing attainment rates only becomes more difficult as a country develops. The United States, as a well-developed country, has seen its attainment rates plateau over the past decade, increasing very slightly from 13.00 to 13.27 years between 2000 and 2010.

By Barro's numbers, this equates to less than a 0.2 percentage point increase in per capita GDP growth over an entire decade.

By comparison, even in a country such as South Korea, which grew from an underdeveloped country in the mid-1900s to an economic powerhouse by the 2000s, the impact of attainment rates is only slightly more pronounced. As shown in Figure 3, the average number of years of schooling for South Korea increased from 4.01 in 1950 to 10.58 in 2000, or approximately 1.3 years of schooling per decade. This leads to an approximate increase of 0.9 percentage points in GDP per capita growth per decade. However, South Korea has also seen its attainment rate growth start to slow as it transitions into a developed country. From 2000 to 2010, the country "only" increased its average attainment rate by a little over a year, which still translates to an approximate 0.7 percentage point increase in GDP per capita.

Therefore, while attainment rates can certainly have an impact on economic growth from Barro's data, it takes a significant amount of time to witness an increase of just one year. In addition, attainment rate growth eventually approaches a limit as a country develops. Once a country becomes well-developed, their level of human capital stock is already so large that it becomes difficult to raise the average number years of schooling by yet another year.

II. iii) Test Scores – A Measurement of Cognitive Skills

Despite the time it takes to increase attainment rates by a single year, many countries have chosen to place an emphasis on quantity of schooling and expansion of school attainment, as noted by Hanushek (2013). This is due to the fact that average attainment rates still more or less have become synonymous with "human capital." However, a few additional issues have recently been raised about their use. Much of the criticism has come from Eric Hanushek, who asserts that the *quality* of education ultimately matters, not necessarily the *quantity*. Hanushek

(2010) argues that the use of attainment rates implicitly assumes that a year's worth of schooling is quantitatively equivalent across countries. In other words, an additional year of schooling in a strong education system like South Korea's can deliver the same increase in knowledge as an additional year of schooling in a poor education system like Indonesia's. This type of logic creates several points of contention.

On the one hand, a year's worth of schooling in South Korea, given the country's more robust infrastructure and developed economy, would presumably be of higher quality than a year's worth of schooling in Indonesia, which has a more fragmented education system. In this regard, attainment rates are likely to reflect a greater increase in knowledge and skill in South Korea in absolute terms. However, on the other hand, it can be argued that the marginal year of schooling in a developing country like Indonesia would be more valuable than the marginal year of schooling in a developed country like South Korea. Since the average Indonesian only receives 7 years of schooling by age 25 (compared to over 11 years for South Koreans), an extra year may provide a significantly greater increase in income growth. While both arguments make valid points, the underlying issue remains unchanged: measuring education through attainment rates is inexact and needs to be adjusted for some degree of quality.

To further Hanushek's argument, attainment rates also ignore the difference in schooling at separate levels of education. For example, an additional year of schooling at the PhD level is equal to an additional year of schooling in primary school. While both levels of education are important in their own respects, they should not be considered equivalent measurements of human capital. By aggregating a country's total years of attainment into one average number, it becomes impossible to determine whether growth rates are ultimately impacted by an additional year of higher education or primary education.

Therefore, Hanushek has been a major proponent of measuring human capital using some basis of education quality, or "cognitive skills." To account for this new variable, Hanushek (2012) uses average test scores on internationally administered exams as a determinant of economic growth.

To give a brief overview of Hanushek's methodology, his study aggregates a variety of international test scores dating back to the 1960s. There are a total of 12 different exams between 1964 and 2003 that tested students in math, reading, and science across countries.[4] Hanushek takes the results from these various exams and adjusts them for difficulty and variance across time.[5] He then normalizes each test score to reflect a mean score of 500 with a standard deviation of 100. This scale essentially replicates the scale of the Programme for International Student Assessment (PISA), which has an OECD average score of 500 and standard deviation of 100.[6] By normalizing scores on every historical exam, Hanushek is then able to compare and aggregate all test results across any time period.

Using these aggregated scores as a new measure of human capital, Hanushek can then test their impact on economic growth, as per formula (1), by running cross-sectional regressions. As a measurement of economic growth, Hanushek uses per capita GDP growth in $US. In his first test, Hanushek chooses to aggregate every country by region and run a cross-region

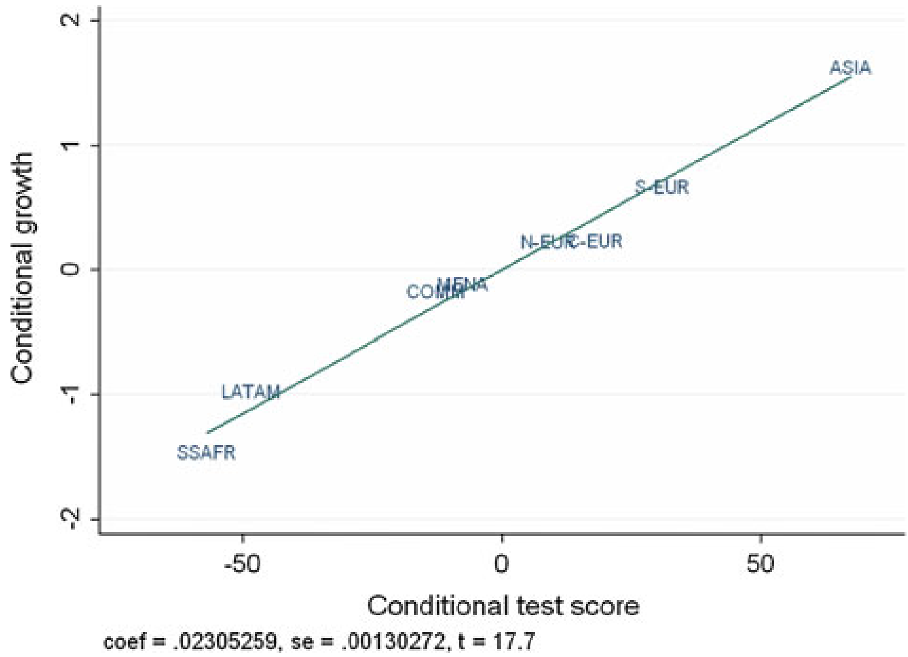
[4] These international assessments include Trends in International Mathematics and Science Study (TIMSS); the Programme for International Student Assessment (PISA); and the First, Second, and Third International Math, Science, and Reading Studies, administered by the International Association for the Evaluation of Educational Achievement (IEA).

[5] Since Hanushek (2012) compiled different sets of exams across multiple years, the difficulty of the exam would inevitably vary over time. Taking a simple average would produce an inaccurate and inconsistent measurement of human capital. Fortunately, the United States had taken part in every international exam and also participated in its own national exam, the National Assessment of Education Progress (NAEP). By taking the patterns in performance on the NAEP and applying them to the United States' score on the international exams, Hanushek is able to adjust the scores accordingly to account for the difficulty of the exam across time.

[6] Hanushek also divides the final score by 100 in the end to make the data easier to work with. This means that the test score data has a mean of 5 and a standard deviation of 1.

regression. Figure 4 below shows the plot of regional per capita GDP growth between 1960 and 2000 against average test scores while controlling for initial GDP per capita in 1960.

Figure 4



Source: Hanushek (2012)

Looking at Figure 4, the growth rates of every region seem to fall on a straight line with test score measurements. According to Hanushek (2012), the R^2 on the regression amounts to a strong 0.985. This result suggests that, conditional on initial GDP in 1960, regional growth from 1960-2000 can be almost completely described by differences in “human capital,” or cognitive skills. Hanushek also adds that incorporating school attainment rates into the regression has no impact on the results. In fact, attainment rates have little explanatory power and are unrelated to the differences in growth rates across regions. It is also worth noting the positive coefficient of 0.023 from the regression, which suggests that a 100 point increase in PISA-scaled assessment scores, or one standard deviation, would result in a 2.3 percentage point increase in per capita GDP growth.

In addition to regional per capita GDP growth rates, Hanushek also examines the impact of cognitive skills on country per capita GDP growth rates from 1960-2000. The full results of his regression are displayed in Table 1 below:

Table 1

	(1)	(2)	(3)	(4)[a]	(5)[b]	(6)[c]	(7)[d]	(8)[e]	(9)[f]
Cognitive skills		2.015	1.980	1.975	1.933	1.666	1.265	1.239	1.985
		(10.68)	*(9.12)*	*(8.28)*	*(8.29)*	*(5.09)*	*(4.06)*	*(4.12)*	*(7.83)*
Years of schooling 1960	0.369		0.026	0.024	0.025	0.047	0.004	−0.049	−0.090
	(3.23)		*(0.34)*	*(0.78)*	*(0.29)*	*(0.54)*	*(0.05)*	*(0.66)*	*(1.02)*
GDP per capita 1960	−0.379	−0.287	−0.302	−0.298	−0.298	−0.255	−0.351	−0.310	−0.879
	(4.24)	*(9.15)*	*(5.54)*	*(6.02)*	*(5.04)*	*(3.12)*	*(6.01)*	*(5.73)*	*(3.39)*
No. of countries	50	50	50	50	52	50	47	45	50
R^2 (adj.)	0.252	0.733	0.728	0.728		0.706	0.784	0.797	0.637

Notes Dependent variable: average annual growth rate in GDP per capita, 1960–2000. Regressions include a constant. Test scores are average of math and science, primary through end of secondary school, all years. Absolute t-statistics in parentheses

[a] Measure of years of schooling refers to the average between 1960 and 2000

[b] Robust regression including the two outliers of Botswana and Nigeria (using *rreg* robust estimation command implemented in Stata)

[c] Specification includes dummies for the eight world regions depicted in Fig. 1

[d] Specification includes additional controls for openness and property rights

[e] Specification includes additional controls for openness, property rights, fertility, and tropical location

[f] GDP per capita 1960 measured in logs

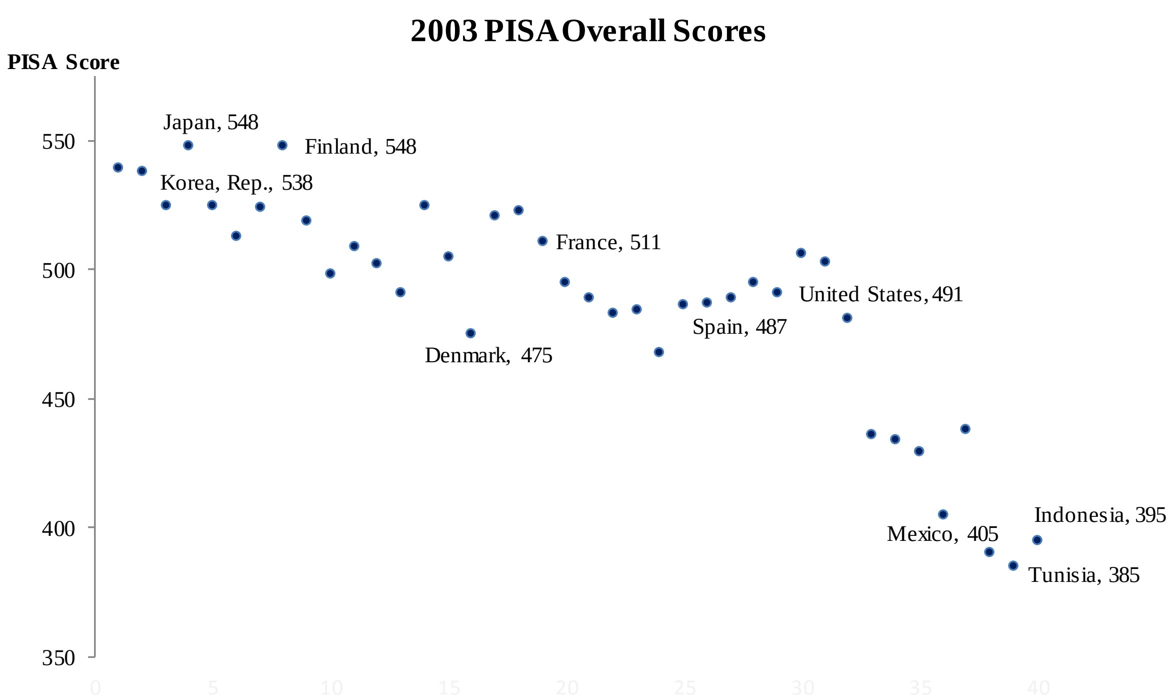
Source: Hanushek (2012)

One of the key points to note in Table 1 is the impact of cognitive skills vs. attainment rates on the variance in per capita growth. Looking at column (1), attainment rates only explain approximately 25% of the variance in growth rates. However, as depicted in column (2), cognitive skills play a much more significant role, explaining nearly 75% of the variation in growth rates. This relationship does not change once attainment rates are re-included in column

(3), as the R^2 and coefficient on cognitive skills are nearly equal. Even as Hanushek incorporates different control variables in columns (4) through (9), cognitive skills still explain a high fraction of the variance in growth. The second point to note is the point estimate of cognitive skills vs. attainment rates. Column (3) highlights the stark difference between the two, as the coefficient on cognitive skills is 2.015 while the coefficient on attainment rates is only 0.026, which is practically zero. In general, the coefficient on cognitive skills fluctuates around 2.00, which suggests that a 100 point increase on PISA-scaled assessment scores, or one standard deviation, can lead to, approximately, a 2.0 percentage point increase in GDP per capita growth. Attainment rates, on the other hand, have a minimal economic impact on growth.

To put Hanushek's results in perspective, the differences in test scores on the 2003 PISA exam are shown in Figure 5 below, with a few select countries highlighted:

Figure 5



Source: OECD PISA Exam 2003

The difference between the top performing countries, Japan and Finland, and the worst performing country, Tunisia, is a wide 163 point margin. By Hanushek's estimates, this equates to a difference of 3.26 percentage points in GDP per capita growth. Even looking at the United States, a well-developed country, they lag behind Japan and Finland by a 57 point margin, which equates to an estimated 1.14 percentage point difference in GDP per capita growth, all other factors held equal. Thus, from the gap in these test scores alone, there can be a significant amount of deviation in GDP per capita growth.

However, the other factor to consider is the feasibility of a country to significantly increase its test score results. In this regard, there is less evidence that a country can markedly improve. Looking back at Figure 2, we see that the United States' PISA scores actually declined 6 points from 2003 – 2012. By Hanushek's growth model estimates, this means that the U.S. actually lost 0.12 percentage points of GDP growth due to its decline in cognitive skills. However, this does not preclude the notion that countries can improve their test score results over time. Take a country like Germany, which has grown to become Europe's economic powerhouse in the 21st century. Germany's overall PISA score increased from 487 in 2003 to 524 in 2012. This 37 point increase equates to a 0.74 percentage point increase in per capita GDP by Hanushek's estimates. Other countries, such as Liechtenstein and Luxembourg, have seen even greater increases in their overall PISA scores (of 49 and 48 points respectively).

Thus, the data shows that increases in test scores are certainly feasible, but may take a long span of time to complete (much like we saw with Barro's work and attainment rates). However, as Hanushek (2013) remarks, the focus of policymakers historically has been on raising attainment rates, not necessarily on improving academic achievement and cognitive skills. Given new literature, such as Hanushek (2012), that shows the strong connection between cognitive skills and growth, policymakers may choose to shift their attention from the *quantity* of education to the *quality* of education, thus leading to more significant increases in test scores.

Therefore, this thesis will build upon Hanushek (2012) to further test whether cognitive skills provide a better measure of human capital than attainment rates when estimating economic growth. Specifically, this thesis will explore whether the relationship between test scores and economic growth still hold over an extended period of time. We hypothesize that a statistically and economically significant relationship between cognitive skills and economic growth will still hold using the most recent data. The next section will lay the foundation of our work and go into the methodology behind our study.

III. Study Replication

To determine whether the relationship between cognitive skills and economic growth still holds over time, our statistical analyses require the most updated data. In our study, the data has updated metrics as of May 2014, which is nearly 10 years more than some of Hanushek's data. However, by using updated data points, we end up using different data sets and slightly modified metrics from Hanushek's original study. This means that the methodology used to create these data could be different, thus producing a different outcome in comparison to Hanushek (2012).

Therefore, it is important to use our updated data sets to first replicate Hanushek's original work over the same time period. By replicating the study in the same time period, we can attribute any differences in results to the differences in data, not necessarily due to extra data points. If the outcomes are too divergent, then it would question the use of these updated data sets in an extension of Hanushek's work. Thus, we have chosen to conduct a series of replication studies that were detailed in Hanushek (2012) using our updated data sources. The differences in data are noted below.

III. i) Data Sources

There are three main variables that need to be collected for the purposes of repeating Hanushek's study: per capita GDP, test scores, and attainment rates. Per capita GDP numbers are taken from the Penn World Tables (PWT) v. 7.1, an updated version since Hanushek's study, which uses v 6.1. In addition, Hanushek's study reports all real GDP numbers in $US. Given that the updated 7.1 PWT include 10 extra years' worth of data, real GDP numbers in $US would slightly change because they would be adjusted off a different base year. Therefore, to keep measurements as stable as possible, the study replications use PPP converted (i.e. real) per capita

GDP in international dollars ($I). The growth rate in per capita GDP is then calculated using differences in natural logs.

Test score data is obtained from the original publication of Hanushek (2012). Although Hanushek makes his test score data set publicly available, he only provides the final average test scores from 1964-2003. The original raw data from each individual test is intentionally omitted. However, for the purposes of replication, the final average test score is sufficient.

For attainment rates, Hanushek uses an extended version of the Cohen and Soto (2007) attainment data set. However, we choose to use Barro and Lee's attainment rate data set for our study replications instead. Barro and Lee, who initially pioneered the use of attainment rates as a human capital proxy, produce a publicly available data set of attainment rates every few years. Their latest release, version 1.3, was done in 2013. Using Barro and Lee's data set provides more up-to-date information that can be readily used in extended studies. In addition, the differences between the Barro and Lee and Cohen and Soto data sets ultimately have a negligible effect on statistical estimates. In a section of Hanushek (2012), Hanushek investigates the impact of using Barro and Lee's data in place of Cohen and Soto's and concludes that the change has minimal impact on the results. Therefore, it seems reasonable and safe to substitute Barro and Lee's data in the study replications as well.

Taking these data sources, we choose to replicate three of Hanushek's original regressions: 1) test scores against regional per capita GDP; 2) initial attainment rates against per capita GDP; and 3) initial attainment rates and test scores against per capita GDP. The hypotheses I will test will be detailed in the sections below, along with the results of all three replications.

III. ii) Cross-Region Regression (1960 – 2000): Test Scores Against Per Capita GDP

The first statistical analysis we choose to replicate is Hanushek's cross-region regression of test scores on per capita GDP. As mentioned previously, Hanushek found that test scores explain a significant fraction of the variance in economic growth ($R^2 = 0.985$). We expect that the estimate of our growth model, as per Eq. 1, should be relatively similar with Hanushek's results and show a significant and positive relationship between test scores and economic growth. Taking our data sources, we compile the necessary information for our cross-region regression, as shown in Table 2 below:

Table 2

Region[a]	# Countries	Total Score	Avg. Score	GDP per Capita 1960 ($I)	GDP per Capita 2000 ($I)	GDP Growth
Asia	11	52.8	4.8	85.5	3794.5	9.48%
OECD Commonwealth	4	20.0	5.0	2479.8	25689.6	5.84%
Central Europe	7	35.4	5.1	2553.6	29619.1	6.13%
Northern Europe	5	24.9	5.0	2380.7	34060.0	6.65%
Southern Europe	5	23.3	4.7	698.3	24059.8	8.85%
Latin America	7	27.2	3.9	597.3	7381.3	6.29%
Middle East and Northern Africa	7	29.2	4.2	931.7	7514.9	5.22%
Sub Saharan Africa	3	10.8	3.6	526.1	800.2	1.05%

[a] Regions are grouped by the following countries: Asia (11): China, Hong Kong, Indonesia, India, Japan, Republic of Korea, Malaysia, Philippines, Singapore, Thailand, Taiwan; OECD Commonwealth (4): Australia, Canada, New Zealand, United States; Central Europe (7): Austria, Belgium, Switzerland, France, United Kingdom, Ireland, Netherlands; Northern Europe (5): Denmark, Finland, Iceland, Norway, Sweden; Southern Europe (5): Spain, Greece, Italy, Portugal, Romania; Latin America (7): Argentina, Brazil, Chile, Colombia, Mexico, Peru, Uruguay; Middle East and Northern Africa (7): Cyprus, Egypt, Iran, Israel, Jordan, Morocco, Turkey; Sub Saharan Africa (3): Ghana, South Africa, Zimbabwe

Sources: Hanushek (2012), Penn WDT v. 7.1

The region's average score is achieved by taking a simple average of PISA scores from each observed country contained in the region. Regional GDP per capita is calculated by aggregating every observed country's GDP ($I) in the region and dividing by total population. From this data, we estimate Eq. 1 using average test scores as the sole explanatory variable and initial 1960 per capita GDP held constant. The results of this statistical analysis are displayed in Table 3 below:

Table 3

	Coefficients	Standard Error	t Stat	P-value
Intercept	-15.413	4.8659	-3.168	0.025
Average Score	5.383	1.1837	4.548	0.006
GDP per Capita 1960 ($I)	-0.002	0.0007	-3.246	0.023

No. Observations	8
R Square (adj.)	0.806

From these results, a few key points can be observed. First, the adjusted R^2 of our regression is lower than Hanushek's result of 0.985. However, these results still suggest that 80.6% of the variation in regional per capita GDP growth can be explained by test scores, which remains a relatively strong predictor. In addition, a coefficient of 5.383 on our average test scores variable is of a considerably higher magnitude than the 2.31 achieved in Hanushek's original study. This suggests that an average increase in 100 points on the PISA across an entire region can lead to a 5.38 percentage point increase in regional per capita GDP.

Since our test score data remained exactly the same from Hanushek (2012), we know that the variation between the study replication and Hanushek's original statistical analysis stems mainly from our use of real GDP per capita in international dollars. However, since our regression seeks to estimate *growth rates* in GDP per capita, differences in GDP levels should not ultimately affect our data. The only point of difference in our growth model would be the initial level of 1960 GDP per capita, which acts as a control variable. Despite this difference, we maintain that using PPP converted ($I) remains an accurate method of measuring per capita GDP consistently across countries.

Although the exact outcome of our statistical analysis is different, we still find the R^2 of the model and point estimate on test scores to demonstrate a significant, positive relationship between cognitive skills and GDP growth. Thus, the ultimate conclusion remains the same as Hanushek's original work. Given that we reach similarly strong linkages between test scores and growth, even while using slightly modified data, we believe this offers a strong reaffirmation of Hanushek's cross-region statistical analysis.

III. iii) Cross-Country Regression (1960 – 2000): Attainment Rates Against Per Capita GDP

The next replication centers on Hanushek's cross-country regression of initial attainment rates against per capita GDP. In Hanushek (2012), this regression was produced to help prove the minimal effect that attainment rates have on economic growth. From Hanshek's statistical analysis, the adjusted R^2 of the model was decisively small (0.252) and attainment rates only had a slight positive relationship in the growth estimation. The coefficient from his study was 0.369. We expect the results from our study replication here to remain similar to these outcomes.

Using the same data sources as mentioned above, we gather initial attainment rates from 1960 on all countries with test score data. We then run cross-country regressions of initial attainment rates in 1960 against GDP per capita growth from 1960 – 2000, conditional on initial GDP per capita in 1960. The results of the regression are displayed in Table 4 below:

Table 4

	Coefficients	Standard Error	t Stat	P-value
Intercept	6.951	0.435	15.992	0.000
Initial Attainment 1960	0.160	0.140	1.141	0.259
GDP per Capita 1960 ($I)	-0.001	0.000	-2.016	0.049

No. Observations	51
R Square (adj.)	0.050

From these results, we see that initial attainment rates as a sole explanatory variable have very little impact on GDP growth. In fact, we observe that initial attainment rates only explain 5.0% of the variation in growth rates between countries. This result is weaker than Hanushek's study (25.2%) and much weaker than the 80.6% we observed in our cross-regional model replication with cognitive skills. Additionally, the coefficient for attainment rates in our estimated growth model is a weak 0.160, which is of an even smaller magnitude than the coefficient estimate achieved in Hanushek's original study (0.369). While these figures differ slightly, the immediate conclusion from our replication study remains similar to Hanushek's original statistical analysis. Attainment rates have significantly less explanatory power and economic significance on a country's economic growth, especially when compared to cognitive skills.

III. iii) Cross-Country Regression (1960 – 2000): Attainment Rates and Test Scores Against Per Capita GDP

The last replication we produce focuses on the core of Hanushek (2012), which is the series of cross-country regressions of per capita GDP growth on both initial attainment rates and test scores, conditional on some initial level of GDP per capita. From Hanushek's publication, we expect to see a significant increase in explanatory power (R^2 = 75%) compared to when attainment rates were isolated as the only independent variable. In addition, the bulk of the impact in our estimated growth model should come from test scores instead of attainment rates.

We run a regression of initial attainment rates in 1960 and test scores against per capita GDP growth from 1960 – 2000. The results of the regression are displayed in Table 5 below:

Table 5

	Coefficients	Standard Error	t Stat	P-value
Intercept	-0.862	1.348	-0.639	0.526
Average Test Score	1.967	0.329	5.979	0.000
Initial Attainment 1960	-0.044	0.112	-0.391	0.697
GDP per Capita 1960 ($I)	-0.001	0.000	-3.330	0.002

No. Observations	51
R Square (adj.)	0.449

From the results of this statistical analysis, we observe a few noticeable discrepancies. First, our data suggest an R^2 of only 44.9%, which is significantly lower than the R^2 of 75% achieved in Hanushek (2012). This presents a slight caveat to any conclusions we draw from this replication, since less than half the variation in GDP growth can be explained by our data on attainment rates and test scores.

However, despite this discrepancy, our study replication does not necessarily diverge from the conclusions of Hanushek (2012). While a low R^2 of .449 tells us that we are explaining a significantly lower fraction of the variation, it still remains higher than the R^2 we achieve of just attainment rates against per capita GDP (0.050). In fact, we see a relatively similar jump in explanatory power as we did in Hanushek (2012), which had an R^2 increase from approximately 25% to 75%. Therefore, despite the fact that our model explains just a fraction of growth variations, our statistical analysis still demonstrates that the addition of cognitive skills significantly improves the explanatory power of the regression overall. This provides an upgrade from prior studies which only used initial attainment rates.

The next and perhaps more important differences to note are the point estimates on our explanatory variables. From Table 5, we witness that the coefficient on attainment rates is actually negative at -0.044. This outcome may run contrary to expectations, especially since past research has seemed to show at least some positive relationship between attainment rates and economic growth. However, this negative coefficient does not meaningfully impact the results of our replication. The first thing to recognize is that the coefficient (-0.044) is relatively close to 0. In other words, an additional year of initial attainment in 1960 would only lead to a marginal decrease of 0.04 percentage points in GDP per capita. Additionally, we cannot draw much from the relationship between attainment rates and economic growth anyways given that the statistical result is insignificant at the 10%, 5%, and 1% level, given its p-value of 0.697.

What is more important about the growth model is the economically and statistically positive impact that cognitive skills still play. The coefficient on our test score variable is 1.967, which is comparable to the results from Hanushek (2012) and the rest of our study replications. Revisiting the case of Germany, these estimates suggest that Germany's 37 point increase produces a 0.73 percentage point increase in GDP growth. In addition, with a p-value extremely close to 0, we can say with strong conviction that relationship between cognitive skills and GDP growth is significant at all statistical thresholds.[7]

III. iv) Brief Summary of Study Replications

In an attempt to replicate Hanushek (2012) with our own data sources, we ultimately find that Hanushek's implications, for the most part, hold true across different data sets. Despite using

[7] All thresholds entail the conventional 10%, 5%, and 1% level.

a different data set for attainment rates and an alternate measurement for GDP per capita, we still find that cognitive skills play a much greater role in economic growth than attainment rates. In general, the move to measure GDP per capita in international dollars seemed to hurt the explanatory power of our variables across all replications. In each respective replication, the R^2 of the regression decreased (in some cases, more significantly than others).

However, the estimated coefficients of our growth model remained consistent throughout. In all cases, the connection between attainment rates and economic growth remained insignificant at all levels of statistical significance. There was also very little economic significance, given that the coefficient was relatively close enough to 0 to be negligible. In addition, test scores continually had a statistically and economically significant relationship with GDP per capita growth. From all of our regressions, our test score variable was significant at all statistical thresholds. The coefficient was relatively consistent throughout as well, typically fluctuating around 2.00 and 2.50. By our previous analysis of PISA score increases, this coefficient range suggests that some countries, such as Germany, could experience a 0.7 to 0.9 percentage point increase in real per capita GDP growth just based on cognitive skill improvements. This provides enough evidence that our test score variable is also economically significant. Therefore, we believe for the purposes of replication, our statistical analyses ultimately confirm Hanushek's assertions on the importance of cognitive skills over attainment rates.

IV. Extended Study

Given that our replication studies reach similar implications to Hanushek (2012), extend our replications to present-day data to see if the conclusions still hold. We use the same data sets as our replication studies, but only with the most recent data points available. The Penn World Table v 7.1 provides GDP per capita and population figures up until 2010. Barro and Lee v 1.3 contain attainment rate data up until 2010 as well.

The methodology to calculate new average test scores is simplified from Hanushek's original technique. As mentioned previously, Hanushek's test score data incorporates a variety of different tests up until 2003. He then normalizes those scores to the same scale of the PISA, which was first administered in 2000. Since Hanushek's work up to 2003, there have been three additional PISA studies, the last of which was administered in 2012. To calculate a new aggregate score, we take Hanushek's test score data from 1963 – 2003 and assume that to be an accurate representation of four decades' worth of cognitive skills. We then take a simple average of PISA scores from 2003 – 2012 and consider that an accurate representation of another decade's worth of cognitive skills. Since both data sets are already adjusted to the same PISA scale, we can weight the exam scores by the length of their time frame to produce one average measurement of cognitive skills from 1963 – 2012.[8]

Although the last decade's worth of scores consist solely of the PISA (and not any of the other exams in Hanushek's original study), the PISA results are widely considered to be more thorough and reliable. The PISA study includes all of the OECD countries, whereas other exams such as TIMSS tend to be more weighted towards developing countries. In terms of the test itself, the PISA emphasizes more real-world applications instead of simple, straightforward

[8] Essentially, Hanushek's data set is weighted by 0.8 and the additional PISA scores from 2003 – 2012 are weighted by 0.2. The weight-adjusted scores are then just added to produce one average measurement of test scores from 1963 to 2012.

question-and-answer material. Therefore, we believe that only including PISA results in the last decade's worth of data does not materially impact the measurement of cognitive skills for our extended studies.

Given these data sources and calculations, we take the same replication studies and extend them to present day data. The extended studies will explore the following three statistical analyses: 1) test scores against regional per capita GDP; 2) initial attainment rates against per capita GDP; and 3) initial attainment rates and test scores against per capita GDP. Some of these extended studies will also look at average attainment rates from 1960 – 2010 instead of initial attainment rates to gauge if that modification has any impact on our growth estimations.

IV. i) Cross-Region Regression (1960 – 2010): Test Scores Against Per Capita GDP

Using the methodology as the replication study described above, we aggregate each country by region as shown in Table 5 below:

Table 5

Region[a]	Code	# Countries	Total Score	Avg. Score	GDP per Capita 1960 ($I)	GDP per Capita 2000 ($I)	GDP Growth
Asia	Asia	9	44.9	5.0	78.9	9795.2	9.64%
OECD Commonwealth	OECD	4	20.1	5.0	2479.8	36413.8	5.37%
Central Europe	CE	7	35.4	5.1	2553.6	39208.6	5.46%
Northern Europe	NE	5	24.9	5.0	2380.7	38972.3	5.59%
Southern Europe	SE	5	23.3	4.7	698.3	30329.4	7.54%
Latin America	LA	7	27.4	3.9	597.3	15831.5	6.55%
Middle East and Northern Africa	ME	4	17.6	4.4	1011.2	12342.9	5.00%

[a] Regions are grouped by the following countries: Asia (9): China, Hong Kong, Indonesia, Japan, Republic of Korea, Malaysia, Singapore, Thailand, Taiwan; OECD Commonwealth (4): Australia, Canada, New Zealand, United States; Central Europe (7): Austria, Belgium, Switzerland, France, United Kingdom, Ireland, Netherlands; Northern Europe (5): Denmark, Finland, Iceland, Norway, Sweden; Southern Europe (5): Spain, Greece, Italy, Portugal, Romania; Latin America (7): Argentina, Brazil, Chile, Colombia, Mexico, Peru, Uruguay; Middle East and Northern Africa (4): Cyprus, Israel, Jordan, Turkey

Sources: Hanushek (2012), Penn WDT v. 7.1, OECD PISA Examinations (2003 – 2012)

A few key differences should be noted with this data. A few countries included in the previous study did not partake in the PISA from 2003 – 2012 (Asia: Philippines, India; Middle East and Northern Africa: Egypt, Iran, Morocco; Sub-Saharan Africa: Ghana, South Africa, Zimbabwe). As a result, they were dropped completely from this extended study. This means that an entire region, Sub-Saharan Africa, becomes eliminated from the statistical analysis. While the region only consisted of 3 countries, which is relatively small amount compared to the 41 total countries still included, it ultimately eliminates a whole observation from the regression which can significantly impact the results.

Bearing this in mind, we run a cross-region regression of average test scores against GDP per capita, conditioned on initial GDP per capita in 1960. The results of the regression are displayed in Table 6:

Table 6

	Coefficients	Standard Error	t Stat	P-value
Intercept	-2.949	3.881	-0.760	0.490
Average Score	2.508	0.873	2.873	0.045
GDP per Capita 1960 ($I)	-0.002	0.000	-4.816	0.009
No. Observations	7			
R Square (adj.)	0.853			

Given this statistical analysis, the strong, positive relationship between cognitive skills and economic growth still holds with the extended time frame. In fact, the results strongly confirm the results of Hanushek (2012) and our replication study. Approximately 85% of the variation in per capita GDP can be explained by test scores, which, although still not as strong as Hanushek's analysis, is actually stronger than our replication study's results (80.6%). In addition, the estimated coefficient on our test score variable is 2.51, which is similar to the 2.30 coefficient that Hanushek originally achieved in his cross-regional regression. This result not only confirms the economic impact that cognitive skills can have on a region's economic growth, but it also suggests that the estimated magnitude of that impact remains consistent and stable through time.

The accuracy of this measurement comes with a few cautions, however. As previously mentioned, the elimination of an entire data point in our regression certainly may have played a role in a higher explanatory power. In addition, with a p-value of 0.0495, the relationship between test scores and GDP growth is not significant at the 1% level. But despite these caveats, the results of our statistical analysis still reach the same conclusions as our replication study, which offers strong support that the relationship between cognitive skills and economic growth holds over time.

III. ii) Cross-Country Regression (1960 – 2010): Attainment Rates Against Per Capita GDP

Our second analysis looks at the impact of initial attainment rates on economic growth, as measured by real per capita GDP. As mentioned previously, our replication study and Hanushek's study both found that attainment rates have little impact on economic growth when isolated as the sole explanatory variable in our estimated growth model. Since our cross-region regression supports the results from Hanushek (2012) and our replication study, we hypothesize that a similar outcome should appear in our extended study here as in our replication.

Taking the 42 countries that took the PISA exams from 2003 – 2012, we first ran a cross-country regression of their initial attainment rates in 1960 against real per capita GDP in international dollars. We then compared those results to that of Hanushek (2012) and our replication study. The outcomes are summarized in Table 7 below:

Table 7

	Hanushek (2012)	Study Replication (1960 - 2000)	Extended Study (1960 - 2010)
Initial Attainment 1960	0.369***	0.160	0.078
	(3.23)	*(1.14)*	*(0.73)*
GDP per Capita 1960 ($I)		-0.001***	-0.001***
		(2.02)	*(3.23)*
GDP per Capita 1960 ($US)	-0.379***		
	(4.24)		
No. Observations	50	51	42
R Square (adj.)	0.252	0.050	0.274

Absolute t-stats are provided in parentheses
* - Statistically significant at 10% level
** - 5% level
*** - 1% level

From the regression outcome, there are a few discrepancies in the results when compared to the other studies. First, the explanatory power of our regression increased significantly to 27.4%, which is greater than any of the other studies. While we could attribute this again to the computation of GDP per capita in international dollars, the more likely cause here is the exclusion of certain countries.

As the table shows, the number of observations in our regression drops from 50 countries to 42 countries. Two of the countries that dropped from our data set due to a lack of data are Nigeria and Botswana. These two data points end up being highly significant because they were acknowledged as outliers in Hanushek (2012). The data that is currently reported in the Hanushek (2012) results exclude these two outliers, which results in a relatively low 25.2% for an R^2. When we include Botswana back into our data in our study replication,[9] we see that the R^2 declines significantly, as expected with the inclusion of an outlier.[10] However, once our extended

[9] Nigeria had no attainment data in the Barro and Lee data set, thus leading to its eventual exclusion regardless.
[10] Some of the decline in the R^2 is also due to the calculation of GDP per capita in international dollars, as mentioned previously. However, we can also reasonably contribute some of the decline in explanatory power to the inclusion of an outlier such as Botswana.

data set gets rid of this outlier, as well as a few other observations, the R^2 increases more towards the level once observed in Hanushek (2012). Therefore, the higher explanatory power of our extended study does not necessarily signal an error in our data; rather, it seems to be a consequence of excluding certain data points.

Another discrepancy of even greater significance is the coefficient on initial attainment rates in our extended study. Despite an increase in explanatory power in our statistical analysis, we find that the impact of attainment rates on growth decreases significantly from 0.160 in our study replication to 0.078. While this decrease in the coefficient could be due to the weakening impact of attainment rates on economic growth over time, it could also be due to the fact that our extended study has fewer observations. Therefore, a lot of uncertainty arises due to the decrease in the number of observations.

To help clarify these discrepancies, we modified the study replication to only include the 42 countries available in our extended study. The results are shown below in Table 8:

Table 8

	Study Replication (1960 - 2000)	**Extended Study (1960 - 2010)**	**Extended Study (1960 - 2010)**
Initial Attainment 1960	0.116	0.078	
	(0.87)	*(0.73)*	
Average Attainment 1960 - 2010			0.156
			(1.45)
GDP per Capita 1960 ($I)	-0.001***	-0.001***	-0.001***
	(2.65)	*(3.23)*	*(3.88)*
No. Observations	42	42	42
R Square (adj.)	0.165	0.274	0.302

Absolute t-stats are provided in parentheses
* - Statistically significant at 10% level
** - 5% level
*** - 1% level

A few key points can be highlighted from Table 8. First, from column (1), we see that the R^2 in the replication study increases significantly from 0.050 to 0.165 just by excluding certain countries from our statistical analysis. This confirms that some of the increase in our extended study's explanatory power is simply due to the smaller sample size. Also, there is a slight decrease in the coefficient on initial attainment rates (from 0.160 to 0.116).

However, Table 8 more importantly allows us to distinguish how much of the decrease on the coefficient is attributable to the smaller sample size and the extended time period. Since the samples between columns (1) and (2) are exactly the same, any differences between the two are attributable to changes over time. Thus, we see over the additional span of 10 years that the impact of attainment rates on growth has actually declined slightly. From our growth model estimates, an additional year of attainment only leads to a 0.078 percentage point increase in per capita GDP, as opposed to 0.116 percentage points 10 years prior. Additionally, our model explains a greater percentage of variation in growth rates solely due to the changing relationship between attainment rates and growth over time.

Column (3) introduces average attainment rate as a new explanatory variable. Arguably, if one were to measure the true impact of attainment rates on growth, initial attainment rates from 1960 may not be the most accurate measure. Initial attainment rates only capture a country's "human capital" at one point in time. When considering data as recent as 2010, using an arbitrary starting point such as 1960 may be outdated. It does not take into consideration whether an improvement in attainment rates ultimately led to the economic growth during a specific time frame. Therefore, just as we chose to measure the impact of *average* test scores on growth over a specific time period, it may be sensible to also use *average* attainment rates as an explanatory variable on growth.

However, as the outcome in column (3) shows, the difference in using average attainment rate from 1960 – 2010 is rather minimal. The R^2 of the statistical analysis increases slightly, but remains low overall at only 30.2%. Average attainment also seems to have a greater impact on per capita GDP growth due to its larger coefficient. However, a slight increase to 0.156 is still a relatively weak relationship, especially since it remains a lot lower than Hanushek's original coefficient of 0.369.

Overall, our work shows that the impact of attainment rates (both initial rate in 1960 and average rate from 1960 – 2010) on economic growth actually decreases as a result of the period of the sample. Yet, despite the weakening relationship between attainment rates and economic growth, our estimated growth model still confirms much of what we originally hypothesized based on Hanushek's work and our study replication – that attainment rates simply do not have a significant impact on real GDP growth.

III. iii) Cross-Country Regression (1960 – 2010): Attainment Rates and Test Scores Against Per Capita GDP

The last extended study pairs attainment rates with test scores to see whether the impact of cognitive skills still outweighs the impact of attainment rates. From our previous estimate of Eq. (1) in the study replication, we found that the relationship between cognitive skills and per capita GDP growth was positive and significant at all statistical levels. By comparison, the impact of attainment rates on growth was close to zero, albeit slightly negative. This result was a strong indictment for focusing on the quality of education to drive growth, not necessarily the quantity. With an extended time frame, we hypothesize that a similar relationship will hold.

Using the same methodology as before, we estimate our growth model in Eq. (1) with average test scores from 1960 – 2013 and initial 1960 attainment rates as the explanatory variables. Initial 1960 GDP per capita is held constant once again. Table 9 below shows a comparison of the results from our statistical analysis.

Table 9

	Hanushek (2012)	**Study Replication (1960 - 2000)**	**Extended Study (1960 - 2010)**
Average Test Score	1.98***	1.967***	1.682***
	(9.12)	*(5.98)*	*(6.71)*
Initial Attainment 1960	0.026	-0.044	-0.058
	(0.34)	*(0.39)*	*(0.76)*
GDP per Capita 1960 ($I)		-0.001***	-0.001***
		(3.33)	*(5.23)*
GDP per Capita 1960 ($US)	-0.302***		
	(5.54)		
No. Observations	50	51	42
R Square (adj.)	0.733	0.449	0.659

Absolute t-stats are provided in parentheses
* - Statistically significant at 10% level
** - 5% level
*** - 1% level

The results from our statistical analysis support our original hypothesis. First, looking at the point estimate on test scores in our regression, there continues to be an economically significant relationship between test scores and GDP per capita growth, albeit at a slightly smaller magnitude. Also, given the high absolute t-stat (6.71), the relationship between cognitive skills and economic growth is significant at all standard statistical levels. Looking next at attainment rates, the point estimate suggests that the impact of attainment rates on economic growth continues to be marginal. Not only is the point estimate statistically insignificant, but with a magnitude of -0.044, it remains negligibly close to zero.

What also jumps out about our extended study is the increase in R^2. From Table 9, the increase from 0.449 in our study replication to 0.659 is quite significant. The study's results are comparable to Hanushek's original work (with its R^2 of 0.733). However, much like our previous extended study, this raises the question of how much of that increase is attributable to the period of the sample rather than to the size of the sample. Thus, we reexamine our study replication using only the 42 countries sampled with the extended time frame. The results are shown in Table 10.

Table 10

	(1) Study Replication (1960 - 2000)	(2) Extended Study (1960 - 2010)	(3) Extended Study (1960 - 2010)
Average Test Score	2.199***	1.682***	1.719***
	(7.82)	*(6.71)*	*(6.48)*
Initial Attainment 1960	-0.066	-0.058	
	(0.76)	*(0.76)*	
Average Attainment 1960 - 2010			-0.067
			(0.81)
GDP per Capita 1960 ($I)	-0.001***	-0.001***	-0.001***
	(4.75)	*(5.23)*	*(5.41)*
No. Observations	42	42	42
R Square (adj.)	0.671	0.659	0.659

Column (1) of Table 10 shows the results of our study replication using only the 42 countries available in our extended study. By narrowing the sample size, the R^2 of our growth model increases significantly to match what we observe in Hanushek (2012) and our extended study, as per column (2). This suggests that a lot of the differences we previously observed between Hanushek's study and our study replication stem from some potential outliers.

However, more importantly, Table 10 strips out the impact of these outliers and allows us to compare how our growth model estimates change over time. Comparing column (1) and column (2), we see little differences except for a drop in the Average Test Score coefficient. This suggests that the impact of test scores on GDP per capita growth has weakened over time. However, in general, what we witness is a consistent pattern. Average test scores continue to maintain a direct relationship with GDP growth while attainment rates remain relatively inconsequential. In addition, our model consistently explains a significant fraction of the difference in growth rates over time.

Factoring in average attainment rates over initial attainment rates also has little impact on our results. The R^2 of our model actually remains exactly the same while the coefficient on attainment remains negative and close to zero. Most importantly, the coefficient of test scores in our regression remains statistically significant and steady across all analyses. Therefore, we believe the outcome of our extended regressions continue to support our original hypotheses – that the quality of education ultimately has a greater bearing and effect on economic growth than the quantity of education.

III. iv) Brief Summary of Extended Studies

Given the outcomes of Hanushek (2012) and our replication studies, we originally hypothesized that the significant, positive relationship between cognitive skills and economic growth would still hold over an extended period of time. In addition, we hypothesized that the explanatory power and statistical relationship of attainment rates on GDP growth would remain weak and relatively insignificant. Table 11 below provides an overall summary of our findings, which support our hypotheses.

Table 11

	(1)	**(2)**	**(3)**	**(4)**
Average Test Score		1.630***	1.682***	1.719***
		(6.78)	*(6.71)*	*(6.48)*
Initial Attainment 1960	0.078		-0.058	
	(0.73)		*(0.76)*	
Average Attainment 1960 - 2010				-0.067
				(0.81)
GDP per Capita 1960 ($I)	-0.001***	-0.001***	-0.001***	-0.001***
	(3.23)	*(8.23)*	*(5.23)*	*(5.41)*
No. Observations	42	42	42	42
R Square (adj.)	0.274	0.662	0.659	0.659

Starting with column (1), we see that attainment rates explain only 31% of the variation in economic growth. By comparison in column (2), the R^2 of our regression jumps to 66.2% once we use average test scores as our sole explanatory variable. It is also important to note the difference in point estimates, as the coefficient on attainment rates is a low 0.078 while the coefficient on test scores is 1.630. As previously noted, both test scores and attainment rates take an equally long time period to increase, so the comparison of coefficients is essentially one-to-one. Thus, we see that cognitive skills overall explain more of the variation in economic growth differences while also having a greater marginal impact on growth.

Columns (3) and (4) continue to support our findings by testing the impact of both attainment rates and test scores on GDP growth per capita. As Table 11 shows, these tests are minimally different from column (2). The explanatory power remains steady around 65.9% and the coefficient on test scores fluctuates around 1.65 – 1.70. What this suggests is that cognitive skills are still explaining most of the variation in economic growth, despite the extra explanatory variable. Therefore, we conclude that, even over an extended time frame, cognitive skills remain a better measurement of human capital than attainment rates.

V. The Impact of Education – Controlling for Other Variables

Now that we have determined that test scores, or cognitive skills, are a more suitable measure of human capital and education than attainment rates, we aim to determine education's true impact on economic growth by controlling for other explanatory variables. To do so, we estimate our growth model in Eq. 1 again with a larger set of country-specific variables related to economic growth. Through this process, we can isolate the relationship between test scores and economic growth by stripping out the impact of other extraneous variables that may also impact differences in growth rates between countries. In deciding which variables to include, we looked at several previous cross-country regressions, such as Barro (1996) and Barro (1998). In general, the explanatory factors Barro includes fall into three general categories: economic environment, political landscape, and demographics.

Economic environment is a broad term to encapsulate the economic freedom of a country. This may include factors such as corporate taxation laws, credit availability, free trade agreements, foreign direct investment restrictions, and labor market regulations amongst others. In past publications, such as Barro (1996), Barro only included a measurement for trade restrictions and terms of trade. However, there have been several new databases since then that encompass more aspects of a country's economic freedoms. Thus, in our study, we choose to measure a country's economic environment using The Fraser Institute's *Economic Freedom of the World* index. The *Economic Freedom of the World* index draws from over 42 variables to measure how supportive countries are of economic freedom. The latest report was released in 2013 based on 2011 economic conditions.

Overall, the *Economic Freedom of the World* index measures a multitude of factors such as monetary policy, property rights, credit market regulations, bureaucracy costs, and the ease of starting a new business, and compiles them into one overall score between 1 and 10. For example, Hong Kong, which is well-known for its stable business environment for corporations, is ranked #1 overall with an index score of 8.97. On the other hand, Venezuela, which is saddled in corruption, military influence, and poor property rights, is ranked #152 with an index score of 3.93. The *Economic Freedom of the World* index provides a highly comprehensive measure of a country's overall economic environment and, thus, we choose to include the index scores in our growth model estimation.

Political landscape is yet another broad term that includes the institutional and regulatory impact of a country's public sector. There are multiple measurements of a country's political freedom that are publicly available for use. For Barro, he included measurements for rule of law and a democratic political system. For the purposes of our study, we incorporate two different variables which we find to be a more thorough and accurate representation of a country's political landscape: Henisz's POLCON index and Transparency International's (TI) Corruption Index.

Henisz's POLCON index incorporates factors such as a democratic political system, but also looks at how well the different branches of government align with each other. In other words, the POLCON index looks at how well the different players in a country's political system interact with each other. Henisz also includes a number of other variables, such as the number of branches in a political system with veto power over policy change. After aggregating all of these factors, the POLCON index ranks every country on a scale from 0 (most hazardous) to 1 (most constrained). To put these numbers in perspective, Argentina's POLCON score in 2012 was a low 0.166. By contrast, Belgium, a very institutionalized, complex federation of multiple communities and regions, has one of the highest POLCON scores of 0.711. The United States falls in the middle of the pack around 0.414.

Given this dispersion, it is not immediately clear what relationship a POLCON score will have with economic growth. If a government operates very efficiently without bureaucracy, it might have a positive impact on economic growth policymakers push for beneficial economic policies. However, this situation can easily be flipped, such as in Argentina, where the government backs poor economic policies that mire the country in unresolved debt. Thus, a country with built-in checks and balances, such as the United States, might have better growth prospects because poor economic policies are less likely to be approved. While it is difficult to determine the exact relationship between political efficiency and economic growth, we still find this to be an important factor to at least control for in our statistical analysis.

In addition, while the POLCON index provides a picture of the political system in a given country, it does not necessarily speak to the effectiveness of that system in terms of its governance and policy decisions. Some countries that have a well aligned democratic system may not necessarily have a stable political system if the government fails to uphold the rule of law or basic property rights. Therefore, we also choose to include TI's Corruption Index as part of our growth model estimations to account for the effectiveness and reliability of a country's public sector. Based on TI's 2013 results, we find that a lot of countries with strong economic environments, such as Hong Kong, Singapore, and Switzerland, score well on the Corruption Index. Therefore, we find it prudent to include both the POLCON index and the Corruption Index to strip out any effect of regulatory bodies on our economic growth estimations.

The last factor we choose to incorporate in our growth model estimation is a measurement of a country's demographics. In Barro's initial studies, he uses factors such as life expectancy and fertility rates to capture the demographic state of a country. This ultimately becomes a very important variable to control for. There has been a wide array of literature that shows an inverse correlation between fertility and GDP between countries. In general, countries with lower fertility rates are observed to have a much higher GDP per capita levels. This is despite the fact that a wealthier population can support more children, hence creating a so-called "demographic-economic paradox."

Barro (1996) provides a few possible explanations for this paradox. First, a growing population means that a country's total pool of capital must be spread across a greater number of people. In effect, a greater portion of economic investments must be used to provide capital to new workers, rather than to raise capital per worker. In addition, with more children being born, a greater number of resources must be devoted to childcare and childrearing, rather than to the production of goods and GDP related factors. Another common theory is that improved female literacy and independence comes with increased GDP per capita. Thus, a wealthier nation contains people more knowledgeable and willing to use contraceptives. Given this strong link between fertility rates and GDP per capita, we choose to include fertility rates (as of 2014), as estimated by the CIA World Factbook, in our growth model estimation to control for its impact on GDP per capita growth.

Assembling all of these variables, we estimate Eq. (1) using average test scores, Economic Freedom Index data (2013), POLCON data (2012), Corruption Index data (2013), and fertility rates (2014 est.) as explanatory variables, conditional on initial 1960 GDP per capita. The results are shown in Table 12 below.

Table 12

	Coefficients	Standard Error	t Stat	P-value
Intercept	-0.862	1.348	-0.639	0.526
Average Test Score	1.379	0.325***	4.233	0.000
Economic Freedom	0.006	0.269	0.024	0.981
POLCON	-0.159	0.676	-0.235	0.816
Corruption	-0.011	0.013	-0.847	0.403
Fertility	-0.828	0.295***	-2.801	0.008
GDP per Capita 1960 ($I)	-0.001	0.000***	-4.622	0.000

No. Observations	42
R Square (adj.)	0.702

From the results of our growth model estimation, there are a few key points to note. First, percent variation explained in our statistical analysis is around 70%. Our previous regressions of just test scores on per capita GDP growth obtained an adjusted R^2 of approximately 65%. Thus, controlling for a few additional factors increased the explanatory power of our model by 5 percentage points. This suggests that some of the other explanatory variables also have an impact on explaining differences in GDP growth across countries, aside from test scores and human capital measurements.

However, the more important aspect to note is the coefficient on our average test scores variable. Our coefficient of 1.379 suggests that an increase of 100 points on PISA-scaled exam scores would lead to an increase of 1.379 percentage points in GDP per capita growth. While this impact is at a lesser magnitude than our previous statistical analysis (see Table 11), the relationship between cognitive skills and economic growth remains significant and positive at all statistical levels. The decrease in the coefficient is also expected with the addition of extra explanatory variables. If the additional factors have an impact on the explanatory power of the model, then they would likely take away some of the effect of test scores. We see a similar phenomenon in past studies as well. Looking at columns (7) and (8) from Hanushek's study in Table 1, the additional controls for openness, property rights, fertility, and tropical location helped increase the explanatory power of the model while also reducing the coefficient of cognitive skills significantly.

To summarize, since this estimate of our growth model encompasses control variables for key economic, political, and demographical factors, the results of the regression present an accurate representation of the impact of education on economic growth. Overall, the quality of a country's education, as measured by test scores, shows a significant a positive relationship with economic growth at all statistical levels. By our estimates, an increase of 100 points on PISA-scaled exams will result in a 1.379 percentage point increase in per capita GDP growth. To put this figure in perspective using Germany as an example again, Germany's 37 point increase on the PISA exam from 2003 – 2012 led to a 0.51 percentage point increase in per capita GDP growth over the past decade, if all other factors were held equal. This proves to be quite economically significant, given that Germany's total per capita GDP growth over that same time period was approximately 3.56%. Thus, given these results in addition to Hanushek's work, we believe that countries should start focusing more of their attention on finding effective ways to improve the quality of their education system. These ideas will be discussed further in the next section.

VI. Policy Implications and Ideas for Further Study

Given the economically and statistically significant impact that test scores have on economic growth, policymakers should seek additional measures to improve the quality of education in their school systems. Previous policies that aimed to keep children in school longer and reduce dropout rates, while still important and helpful, may not have produced the desired economic effect that policymakers were seeking. These policies focused largely on increasing the *quantity* of schooling, but did not target specific measures to improve the *quality* of this extra time in school.

Therefore, one potential solution for policymakers to consider is to set minimum standards of academic achievement for all students at certain ages. This policy was already given some thought in the United States through the No Child Left Behind (NCLB) Act of 2001. Under the NCLB Act, all states were required to set their own achievement standards in order to receive federal education funding. States would then test to see whether schools met these standards by administering a state-wide standardized test. For schools to receive funding, they must show that students have made “adequate yearly progress” on these exams. Should a school not reach these minimum standards, they ultimately face reductions in funding and mandatory restructuring.

While the premise of testing children on academic achievement is well-founded, the implementation and punishments under the NCLB Act have come under heavy criticism. The most prominent criticism is the fact that underperforming schools face heavy reductions in funding. Thus, if a school is already struggling, reducing its level of funding would only exacerbate the problem.

However, would decreasing a school’s funding have a significant impact on academic achievement? This brings up a natural question to further explore: what are the key determinants that have an impact on the variation in test scores? We have found that test scores can have a statistically and economically significant impact on growth, but what factors have a statistically significant impact on test scores? In this regard, perhaps attainment rates would be a better determinant of test scores, not necessarily economic growth. Other factors to consider would be amount of federal funding, teacher salaries, and cultural values. Therefore, while the NCLB Act could potentially be poorly executed by cutting funds to underperforming schools, there could be more important factors that ultimately determine students' performance on tests.

Lastly, it is important to keep in mind the context of our study's statistical analysis. While it is encouraging to suggest that test scores and cognitive skills have an economically and statistically significant impact on economic growth, it is also important to realize that there are several moving parts to the equation. Our regression analysis was created with a set of conditional estimates that consider a country's economic, political, and demographic environments. Thus, while a country like Germany improved drastically on the PISA over the past decade, much of their economic growth is also due to other moving variables, such as their favorable economic environment under the Euro. A country such as the United States, which has historically performed poorly on internationally administered exams, still operates as a dominant global economy due to its strong political and economic infrastructure. Therefore, while education is clearly an important factor in a country's development, it will not singlehandedly lead to significant economic growth. A country should consider a wide variety of institutions and policies when improving its growth prospects – education just being one of the very important factors.

Works Cited

Aghion, P., L. Boustan, C. Hoxby, and J. Vandenbussche. "The Causal Impact of Education on Economic Growth: Evidence from U.S." Web.

Barro, Robert J. *Determinants of Economic Growth: A Cross-Country Empirical Study*. Working paper no. 5698. Cambridge: National Bureau of Economic Research, 1996. Print.

Barro, Robert J. "Economic Growth in a Cross Section of Countries." *The Quarterly Journal of Economics* 106.2 (1991): 407-43. Print.

Barro, Robert J. *Human Capital and Growth in Cross-Country Regressions*. Rep. 1998. Print.

"Corruption Perceptions Index 2013." Transparency International, 2013. Web. 09 May 2014.

"Country Comparison: Total Fertility Rate." *Central Intelligence Agency*. Central Intelligence Agency. Web. 11 May 2014.

Fiscal Year 2014: Historical Tables - Budget of the U.S. Government. Rep. Washington, D.C.: Office of Management and Budget, 2014. Print.

Gwartney, James, Robert Lawson, and Joshua Hall. *Economic Freedom of the World: 2013 Annual Report*. Publication. Fraser Institute, 2013. Print.

Hanushek, E. A. (2013). “Economic growth in developing countries: The role of human capital.” Economics of Education Review. 37: 204-212.

Hanushek, E. A., & Wöessmann, L. (2008). “The role of cognitive skills in economic development.” Journal of Economic Literature. 46 (September (3)): 607–668.

Hanushek E A and Wöessmann L. (2010). "Education and Economic Growth." In: Penelope Peterson, Eva Baker, Barry McGaw, (Editors), International Encyclopedia of Education. 2: 245-252. Oxford: Elsevier.

Hanushek E A and Wöessmann L. (2012). "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." Journal of Economic Growth. 17: 267-321.

Henisz, Witold J., and Bennet A. Zelner. *Measures of Political Risk*. Rep. Print.

Heston, A., Summers, R., & Aten, B. (2011). Penn World Table Version 7.1. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, Nov 2012

Indicators of Governance and Institutional Quality. Rep. World Bank. Print.

Organisation for Economic Co-operation and Development. (1998). Education at a glance: OECD Indicators. Paris: Organisation for Economic Co-operation and Development.

Organisation for Economic Co-operation and Development. (2003). Literacy Skills for the World of Tomorrow - Further results from PISA 2000. Paris: OECD

Organisation for Economic Co-operation and Development. (2004). Learning for tomorrow's world: First results from PISA 2003. Paris: OECD.

Organisation for Economic Co-operation and Development. (2007). PISA 2006: Science Competencies for Tomorrow's World: Volume 1: Analysis. Paris: OECD.

Organisation for Economic Co-operation and Development. (2010). PISA 2009 Results: Executive Summary. Paris: OECD.

Organisation for Economic Co-operation and Development. (2013). Education at a glance 2013: OECD Indicators. Paris: OECD.

Organisation for Economic Co-operation and Development. (2013). PISA 2012 Results in Focus. Paris: OECD.

UNESCO. *Snapshot from UNESCO on Enhancing Access to Education in the LDCs, with a Special Focus on Technical and Vocational Education and Training (TVET). Www.unohrlls.org*. UN Office of the High Representative for the Least Developed Countries, Landlocked Developing Countries and Small Island Developing States. Web.