International Journal of
Education Policy & Leadership

IJEPL
Volume 11(4)
2016

School Reform: America's Winchester Mystery House

David Meyer
Jefferson Joint School District

Loredana Werth
University of Pikeville

Abstract This quantitative study examines the correlation between international student achievement test outcomes and national competitiveness rankings. Student achievement data are derived from a variation-adjusted, common-scale metric data set for 74 countries that have participated in any of the international mathematics and science achievement tests since 1964. National competitiveness data are taken from the 2014–15 Global Competitiveness Index (GCI) published by the World Economic Forum. A Spearman's rank-order correlation was run to assess the relationship between student performance on international achievement tests and the competitiveness of nations. For all nations, there was a moderate positive correlation between student performance on international achievement tests and the competitiveness of a nation, $r_s(98)=0.688$, $p<0.001$. However, this relationship disappeared among the 18 most competitive nations, the cohort to which the United States belongs. The relationship also disappeared among the 18 nations with the highest achievement scores on international tests. Student performance on international assessments appears to have no relationship to the competitiveness of the United States. This study has implications for legislators and public education leaders who want to maximize the return on investments in education. Education dollars and reform initiatives should be diverted toward addressing poverty, funding schools equitably, alleviating social stress and violence, and supporting young families and students of immigrant families.

David Meyer is a principal in Jefferson Joint School District No. 251 & **Loredana Werth** is Provost at the University of Pikeville. (2016). *International Journal of Education Policy & Leadership 11*(4). URL: http://journals.sfu.ca/ijepl/index.php/ijepl/article/view/665

Literature review

The Silicon Valley is home to one of California's most unsettling tourist attractions, the Winchester Mystery House. Not long after the death of her husband—the legendary maker of the Winchester rifle—Sarah Winchester moved to California and began a 38-year obsession, building an enormous mansion of architectural enigmas. There was no master plan, but Mrs. Winchester kept workers busy day and night building oddities, including staircases that led to nowhere and doors that opened into walls. Debate continues to this day about what possessed her to build the mysterious house that squandered millions of dollars of her late husband's estate. The bizarre, haphazard construction of the Winchester Mystery House is an apt metaphor for public school reform in America over the past 50 years.

Education reform in America began in earnest after the passage of the National Defense Education Act (NDEA) in 1958. The NDEA was passed in response to the launch of Sputnik and the subsequent perception that America was falling behind the Soviets in the fields of science and technology (Department of Education, 2012b). This inaugurated an era of international student achievement comparisons and discussion about the impact of student achievement on national competitiveness in a global economy. Pons (2011) notes that most of the media attention from international assessment results has focused on the rank order of the countries themselves, rather than a deeper discussion and analysis of what is driving the relative success or failure of participating nations. The concern that the United States might be falling behind the rest of the world has led to reform efforts in public schools such as increased accountability through standardized testing. However, Wu (2010) points out that the statistical complexities of large-scale assessments make it difficult for policymakers to recognize the caveats in the data, leading to misguided conclusions and inappropriate policy decisions. Policymakers who use international student achievement test outcomes as indicators of whether or not American students will be able to compete in the future presume that there is a correlation between student achievement test metrics and economic competitiveness indicators.

In 1964, the International Association for the Evaluation of Educational Achievement (IEA) administered the First International Mathematics Study (FIMS) to 13-year-olds in 12 countries (OECD, 2010). In 1965, Congress passed the Elementary and Secondary Education Act (ESEA), which has been reauthorized numerous times, the most recent being the No Child Left Behind Act of 2001. The ESEA mandates that states receiving federal title money develop academic standards and increase accountability for student achievement (Department of Education, 2014). Not long after the administration of FIMS, the First International Science Study (FISS) followed, and then the Second International Math Study (SIMS) and Second International Science Study (SISS). In 1995, the IEA administered the Third International Mathematics and Science Study (now referred to as Trends in Mathematics and Science Study or TIMSS). Nine- and thirteen-year-olds in 46 countries participated (OECD, 2010). The Organisation for Economic Co-operation and Development (OECD) began assessing the performance of 15-year-old students worldwide in 2000 using the Programme for International Student Assessment (PISA). The PISA has been administered every three years since 2000. For nations that have rankings below the OECD average, this has resulted in alarm and efforts to update or repair apparently broken education systems (Alegre & Ferrer, 2010; OECD, 2010; Tienken, 2008).

The National Commission on Excellence in Education released *A Nation at Risk* (ANAR) in 1983, which claimed that "Our once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world" (National Commission on Excellence in Education, 1983). The first piece of evidence used to make this claim was U.S. performance on international student achievement tests compared with other industrialized nations. The ANAR report referenced Sputnik, claiming that average achievement on the College Board's Scholastic Aptitude Tests for high school students was lower in 1983 than 26 years previous when Sputnik was launched (National Commission on Excellence in Education, 1983). This reignited concern that America's public school system was broken, and reform efforts were redoubled across the nation.

The National Governors Association (NGA) met in 1986 and, among many other recommendations, suggested that the U.S. education system benchmark its performance against international competitiveness (Alexander, 1986). President George H.W. Bush's America 2000: Excellence in Education Bill of 1991 and President Clinton's Goals 2000: Educate America Act of 1994 specifically stated (in goal 4 of *America 2000* and goal 5 of Goals 2000) that American students should score first in the world in science and mathematics by the year 2000 (see Department of Education, 1991, 1994). This goal was not met, and in 2002 President George W. Bush signed into law the No Child Left Behind Act (NCLB). For the next decade, public schools labored under the requirements of NCLB to develop rigorous achievement standards and statewide assessments to measure student progress toward 100 percent proficiency for all students. The outcome of NCLB was a portrait of failure: more than half of America's schools were considered "failing" by 2010, and the rest were not far behind. In Massachusetts, a state considered to have the most rigorous standards, 80 percent of the schools were not making adequate yearly progress as required by NCLB (Karp, 2014).

In 2009, the National Governors Association and the Council of Chief State School Officers met to develop the Common Core State Standards Initiative. In a 2009 press release, the NGA made numerous references to American students falling behind their international counterparts. The press release promoted the need for common standards that could be benchmarked to other top performing nations around the world so that American students would be prepared to compete internationally (National Governors Association, 2009). In 2010, President Obama submitted to Congress his ideas for the reauthorization of ESEA, citing in his introductory letter the moral imperative of creating a world-class education system in the United States (Department of Education, 2011). Facing a gridlocked Congress, President Obama and U.S. Secretary of Education Arne Duncan developed ESEA waivers whereby states could obtain relief from NCLB sanctions (Department of Education, 2012a). In order to obtain an ESEA waiver, states had to meet a number of requirements, one of which was that they develop college- and career-ready standards, defined as "standards that are common to a significant number of states" (Department of Education, 2014).

In accordance with Section 604 of the America COMPETES Reauthorization Act of 2010 (COMPETES), the Secretary of Commerce completed a study of the economic competitiveness and innovative capacity of the United States (Department of Commerce, 2012). The report addresses a number of variables that impact competitiveness, including tax policy, business climate, regional issues, barriers to business startups, trade policy, federal research and development policy, intellectual property in the U.S. and abroad, manufacturing, and science and technology education. The report suggests that only a workforce equipped with skills in science, technology, engineering, and math (STEM) will be capable of generating the innovation necessary to be competitive in the global marketplace (Department of Commerce, 2012).

Finkel (2012) notes that prior to World War II it was common for America to benchmark its academic performance against other nations. Since then, the United States has not adopted the effective learning strategies that other nations use. Meanwhile, America's average performance on PISA 2009 places its students 4th in reading, 17th in science and 25th in mathematics among 70 countries (Finkel, 2012). Nevertheless, Petrilli and Scull (2011) found that the United States produces many more high-achieving students than any other OECD nation. Racial and ethnic segments of the U.S. population rival overall populations in other countries as well (Petrilli & Scull, 2011). Research conducted by Breton (2013) finds that investments in education increase marginal productivity for countries with a highly educated population by 12 percent, while the marginal productivity gains for counties with less-educated populations are higher—greater than 50 percent (Breton, 2013). This finding shows that the economic principle of diminishing returns seems to apply to investments in education as well. As nations become more educated, it becomes more difficult to raise national income through continued investments in education (Breton, 2013). Econometric studies clearly show a relationship between higher education and the productivity of workers; however, it remains unclear as to whether education raises productivity and income, or whether increased levels of productivity and income increase people's demand for more education (Breton, 2013).

Through simultaneous equations modeling (SEM), Jun, Xiao, and Xiaoyu (2009) evaluate the interaction among the endogenous variables of income distribution inequality and education inequality, as well as the instant-impacts and cumulative impacts of these variables. Jun et al. (2009) found that income inequality leads to educational inequality, but that attempts to create education equity do not reduce income inequality. Making a basic education available to more people generally raises the income and education of the population, but there is no causal effect between education equality and income equality (Jun et al., 2009). In other words, efforts to raise the education level of the population through educational policy initiatives does not automatically translate into increased income equality.

Tienken (2008) analyzes the relationship between Global Competitiveness Index (CGI) rankings and international test rankings from three time periods (1957–1982, 1983–2000, and 2001–2006). The Global Competitiveness Report is a publication of the World Economic Forum. Statistical data for the index is obtained from the World Bank, the International Monetary Fund (IMF) the United Nations Educational, Scientific and Cultural Organization (UNESCO), and the World Health Organization (WHO). More than 100 indicators are used to calculate the GCI (Schwab, 2014). The Tienken (2008) analysis is limited to math and science[1] and finds stronger correlations between GCI and ISAT achievement for nations in the bottom 50 percent than for nations in the top 50 percent. This suggests that ISAT performance may predict future economic growth for emerging economies, but not for advanced economies. Tienken (2008) also finds that current ISATs have stronger correlations than older ISATs do. These findings are consistent with Hanushek and Woessmann (2009), who find a strong correlation between variances in cognitive skills and differences in economic growth. Investment in top-performing students has a convergence effect on economic growth that is stronger in underdeveloped countries than in advanced economies (Hanushek & Woessmann, 2009).

A joint project of the Horace Mann League and the National Superintendents Roundtable, *School Performance in Context: Indicators of School Inputs and Outputs in Nine Similar Nations* (Harvey, McKay, Fowler, & Marx, 2015) analyzes 24 indicators in six categories for Canada, China, Finland, France, Germany, Italy, Japan, the United Kingdom, and the United States. The report emphasizes that national rankings on international assessment outcomes lose much of their relevance when they are decoupled from societal factors that contribute to poor student achievement. Policymakers are encouraged to focus on initiatives beyond school accountability—initiatives that tackle problems that directly impact school systems—and to subdue rhetoric that the sky is falling when student performance is lackluster. For example, the report recommends policies that address social inequality, social stress and violence, and support for young families (Harvey et al., 2015). Policies directed specifically at the public school system should address the achievement gap, on-time graduation rates, funding equity, and teacher quality (Harvey et al., 2015). The 2015 Index of Economic Freedom, a publication of the Heritage Foundation, reports that the United States actually trails Hong Kong, Singapore, New Zealand, Australia, Switzerland, Canada, Chile, Estonia, Ireland, Mauritius, and Denmark, due to deteriorations in the rule of law and to the expansion of government during the War on Terror and the Great Recession (Miller & Kim, 2015).

Gaps in the literature

Macroeconomic studies on the topic of international student achievement test (ISAT) outcomes and economic competitiveness have attracted the increasing attention of researchers since the turn of the 21st century, but they are still relatively few in number (Baker, 2007; Hanushek & Kimko, 2000; Tienken, 2008). Research has left gaps in the literature, whether through the exclusion of one or more international student achievement tests, the use of academic attainment instead of an achievement proxy, the lack of a time lag analysis, or an insufficient index to measure economic competitiveness (Hanushek & Woessmann, 2010b; Yu, DiGangi, & Jannasch-Pennell, 2012).

Studies that look for relationships only between one ISAT and other economic indicators do not provide a complete picture, since each test has its own unique format and purpose (Acar, 2012; Hoi Yan & Chan, 2008). Studying relationships between academic attainment and economic indicators assumes that a year of education has the same value for students regardless of where they live (Hanushek & Woessmann, 2007). The lack of a time lag analysis is problematic because, without it, there is no evidence that educational inputs have had a positive impact on economic output. Without a time lag, the students in question will not be old enough to be contributing to the economy (Yu et al., 2012). Most studies look for relationships between academic inputs and gross domestic product (GDP), which is a measure of economic growth, not competitiveness. A reliable measure of competitiveness that incorporates multiple points of data is the Global Competitiveness Index (GCI) produced by the World Economic Forum. This study overcomes these limitations by utilizing the OECD common-scale metric of student achievement on all ISATs over the past 50 years and archival data from the World Economic Forum 2014–15 Global Competitiveness Report.

Up to this point in time, studies have largely focused on analyses between countries *within* individual assessments, and justifiably so. Each test has its own unique format and purpose, making a common ISAT scale elusive (OECD, 2010). Utilizing the International Data on Cognitive Skills common ISAT scale from the publication *The High Cost of Low Educational Performance: The Long-Run Economic Impact of Improving PISA Outcomes*, it is possible to standardize the results for all ISATs since 1964 (OECD, 2010). Filling this gap in the literature provides policymakers with more information about the validity of using ISAT achievement results to make predictions about the United States' competitive status in the world.

Research questions

This study evaluates the effects of cognitive skills on economic competitiveness within the Mankiw, Romer, and Weil (MRW) theoretical framework and the Hanushek-Woessmann education production function and theories of technological diffusion (Benhabib & Spiegel, 2002). This study answers the following research questions:

> What relationship, if any, exists between international student achievement test scores in mathematics and science as measured by the International Data on Cognitive Skills (COG) and the 2014–2015 Global Competitive Index (GCI)?
>
> When the GCI is divided into quartiles, what relationship, if any, exists between international student achievement test scores in mathematics and science as measured by the COG and the 2014–2015 GCI for nations in the top 25 percent of competitiveness?
>
> When the COG is divided into quartiles, what relationship, if any, exists between international student achievement test scores in mathematics and science as measured by the COG and the 2014–2015 GCI for nations in the top 25 percent of achievement?

Data collection

The International Association for the Evaluation of Educational Achievement (IEA) and the Organisation for Economic Co-operation and Development (OECD) are the two organizations that have conducted ISATs and collected student achievement data since 1964 (OECD, 2010). The data file for the International Data on Cognitive Skills is available in the public domain at http://hanushek.stanford.edu/download and is displayed in Appendix A. The international competitiveness data used for regression analysis was the Global Competitiveness Index (GCI) published by the World Economic Forum. This archival data is available in the public domain at the World Economic Forum website and is displayed in Appendix B. The GCI is computed based on successive aggregations of economic indicator scores and survey data, which are grouped into 12 component scores that generate a nation's rank on the GCI (Schwab, 2014). This multi-faceted index is a far more robust measure of competitiveness and captures much more than GDP alone.

Method and data analysis

An explanatory correlation research design is used to evaluate the extent to which ISAT scores co-vary with statistics on international competitiveness for 74 of the 77 countries that have participated in any of the international mathematics and science tests since 1964. Three countries were excluded because they are not part of the 2014–15 Global Competitiveness Index (Liechtenstein, Macao-China, and Palestine). Spearman's *rho* correlations are used to model the relationship between ISAT average test scores, scaled to PISA 2000, in mathematics and science and the Global Competitive Index. Primary grades through the end of secondary school are included in the regression analysis for all years from 1964 through the 2003 cycles of PISA and TIMSS. The majority of the existing literature regarding connections between education and economic growth utilizes years of school attainment as the cognitive skills metric. This is a drawback because a year of education does not necessarily equate to the same increase in knowledge and skills across all education systems around the world (Hanushek & Woessmann, 2007; OECD, 2010). For example, a year of schooling in Kyrgyzstan does not yield the same benefits to a student as does a year of schooling in Finland (OECD, 2010). This study uses the Hanushek-Woessmann education production function, which uses student achievement data as the cognitive skills metric.

Spearman's *rho* correlations are used to model the relationship between the Global Competitive Index and ISAT average test scores in mathematics and science, from primary through the end of secondary school for all years scaled to PISA. Spearman's *rho* is used to determine correlations between ordinal scale variables, or interval scale variables that have been reduced to ordinal scale (Tanner, 2012). The strength and direction of the correlational tests are analyzed. This study incorporates a time-lag analysis of ISAT data from 1964–2003 with 2014–15 GCI data, ensuring that the students who participated in PISA 2003 are now old enough to be contributing to the economy. Interpretations and conclusions are made based on the outcome of the correlational tests.

Participants

Archival data of student achievement on all ISATs over the past 50 years was obtained from the OECD (2010) study and used for this research. The data file is available in the public domain at http://hanushek.stanford.edu/download . Hundreds of thousands of students between the ages of nine and fifteen (depending on the ISAT) from 77 countries participated in one or more of the mathematics or science ISATs. In the OECD (2010) study, Hanushek and Woessmann obtained a standardized comparison across ISATs through a multi-step empirical calibration (OECD, 2010). First, student performance in the United States as found in the National Assessment of Educational Progress (NAEP) since 1969 was used as a pattern against which to scale each international assessment over time (OECD, 2010). Hanushek and Woessmann then selected thirteen countries that had participated in a sufficient number of ISATs and had maintained relatively stable educational systems since 1964 to be a standardization group for the study (OECD, 2010). The mean scores for each ISAT that an OECD country participated in were then calibrated to the variance observed on the PISA 2000 assessment. PISA 2000 was selected as the calibration point because that is the only assessment that all nations participated in together (OECD, 2010). Adjustments in achievement levels based on the NAEP were combined with the standardization group calibration to PISA 2000 to calculate standardized scores for all countries on all ISATs (OECD, 2010).

Limitations

Large-scale assessments like PISA and TIMSS have a number of inherent errors related to measurement, sampling, and equating (OECD, 2010; Hanushek & Woessmann, 2010a; Tienken, 2008; Yu et al., 2012; Yu, 2012). There are a limited number of countries co-operating to produce common data sets (Hanushek & Woessmann, 2010a). When selecting countries to be part of the data standardization group, two criteria were used: stability and secondary population (OECD, 2010). Countries had to be member states of the OECD since 1964, and they needed to have had a significant population of secondary students in 1964. Thirteen countries met both of these criteria: Austria, Belgium, Canada, Denmark, France, Germany, Iceland, Japan, Norway, Sweden, Switzerland, the United Kingdom, and the United States (OECD, 2010). Longitudinal research is impossible due to the cross-sectional design of international assessment instruments (Hanushek & Woessmann, 2010b). Furthermore, the cross-sectional design only shows correlations, not causation (Loveless, 2009).

The PISA assessment, upon which the common-scale metric depends for a variation-adjusted test score, has a number of limitations that should be noted. When disaggregating data by ethnicity, unregistered or undocumented immigrants are not usually included in PISA data (Martin, Liem, Mok, & Xu, 2012). Inconsistencies between what is being taught in schools and what is being tested on PISA may put some nations at a disadvantage (Tienken, 2008). Grouping the United States with other OECD nations is problematic because other OECD nations are smaller and/or more homogeneous (Bracey, 2009; Cavanagh, 2007a; Cavanagh, 2007b; Perry, 2009). Tienken (2008) notes that other nations are selective about the student populations that participate in PISA, resulting in samples that may not reflect the whole nation. According to a 2012 report of the United Nations Children's Fund, only Romania ranks lower than the United States in terms of the percent of children living in households with incomes lower than 50 percent of the national median (UNICEF Innocenti Research Centre, 2012). Poverty and cultural influences have been shown to correlate with achievement on standardized test scores, which may introduce bias (Ammermueller, 2007; Bracey, 2009; Hanushek & Woessmann, 2010b; Tienken, 2008).

Results

COG and GCI, All Nations

The Spearman's *rho* rank-order correlation coefficient value for all nations is r_s=0.688 for *n*=74 paired observations. The two-tailed significance level is *p*<0.001. According to the Cohen and Manion (1994) standard for interpreting the strength of association, there is a moderate, positive, significant correlation between performance on international student achievement tests and the competitiveness of nations, $r_s(98) = 0.688$, *p*<0.001. H_o1 is rejected and H_A1 is accepted. This finding does not mean that stronger performance on achievement tests causes a nation to be more competitive, just that the two factors are associated. To know whether or not the correlation can be applied to the United States, the COG and GCI data were divided into quartiles and a Spearman's *rho* analysis was conducted for the nations in the top 25 percent for each data set.

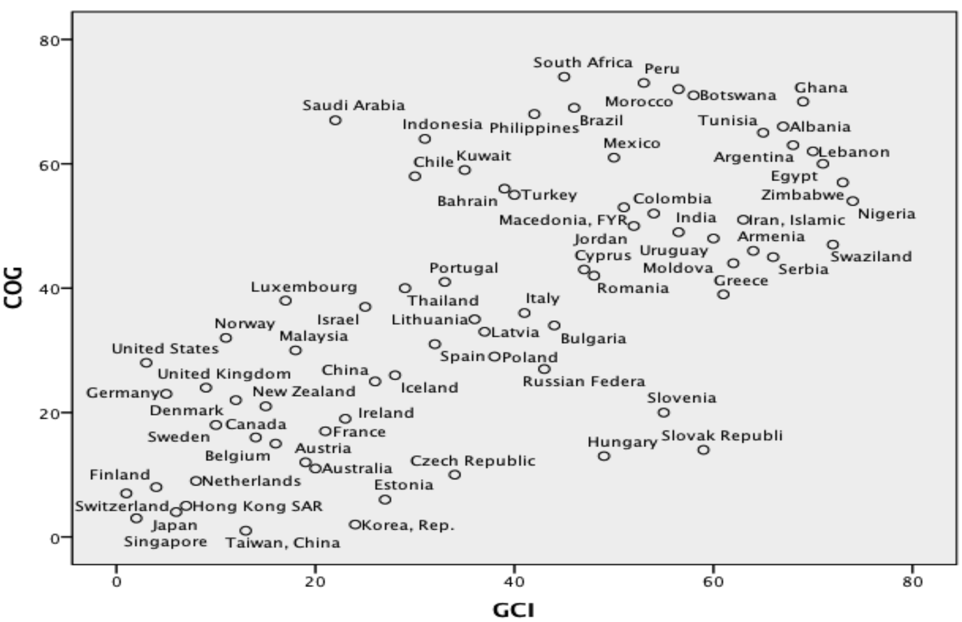

Figure 1. Scatterplot Showing the Monotonic Relationship between ISAT Cognitive Data (COG) and Global Competitiveness Index (GCI)

COG and GCI Top 25 percent most competitive

For (n = 18) nations, $r_s(98) = 0.42$, p = 0.083, indicating no significant relationship. When the GCI is divided into quartiles, there is no relationship between international student achievement test scores in mathematics and science as measured by the International Data on Cognitive Skills and the 2014–2015 Global Competitive Index for nations in the top 25 percent of competitiveness. Performance on international assessments appears to have no relationship to the competitiveness of nations that have the highest competitive rankings on the GCI.

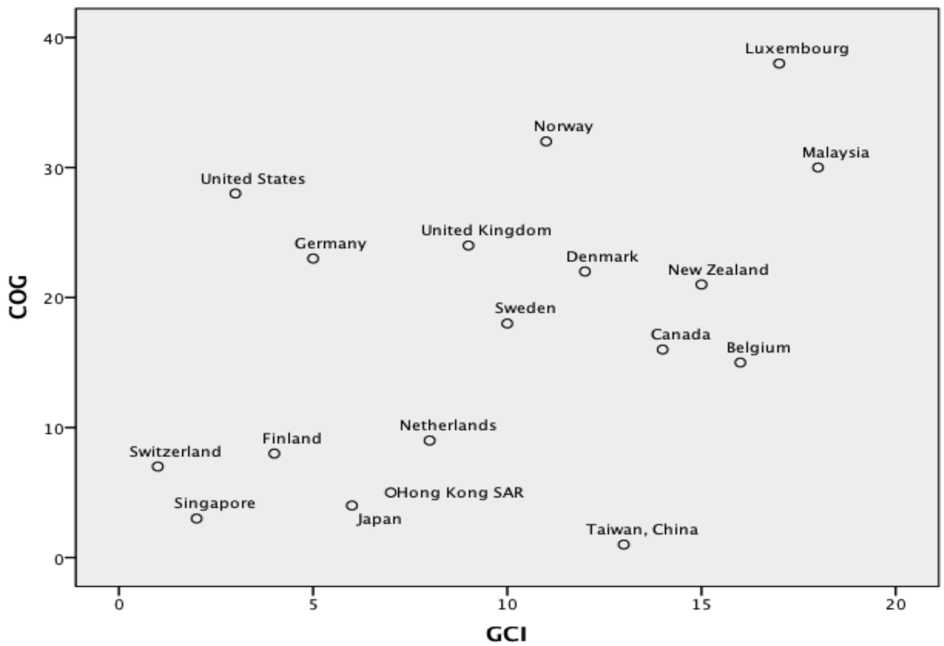

Figure 2. Scatterplot Showing the Relationship between ISAT Cognitive Data (COG) and Global Competitiveness Index (GCI) for Top 25 Percent Most Competitive Nations

COG and GCI top 25 percent cognitive skill

For (n = 18) nations, $r_s(98) = 0.354$, $p = 0.150$, indicating no significant relationship. When the COG is divided into quartiles, there is no relationship between international student achievement test scores in mathematics and science as measured by the International Data on Cognitive Skills and the 2014–2015 Global Competitive Index for nations in the top 25 percent of achievement. Performance on international assessments appears to have no relationship to the competitiveness of nations that rank in top 25 percent on the COG. Again, this finding has important implications. Even though the United States ranks with the cohort of the middle 50 percent in cognitive achieve-

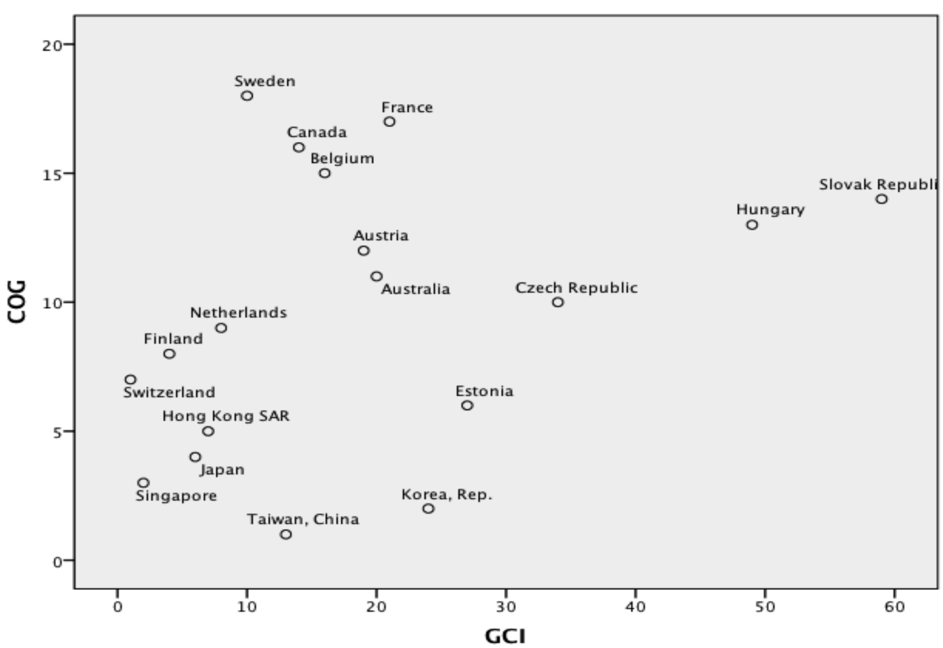

Figure 3. Scatterplot Showing the Relationship between ISAT Cognitive Data (COG) and Global Competitiveness Index (GCI) for the Top 25 Percent of Cognitive Skill

ment, this finding shows that all correlation with competitiveness among the nations in the top 25 percent will vanish if and when the U.S. ever rises into that cohort.

Conclusion and discussion

Since the Colman report in 1966, the U.S. school system has been rife with reform activity, including the end of racial segregation, the publication of *A Nation at Risk*, outcome-based education, Goals 2000, No Child Left Behind, the Common Core State Standards Initiative, and more recently the Every Student Success Act (ESSA). Reform efforts have largely been influenced by concerns about the United States' national competitiveness in a global economy. This study demonstrates that one basis (U.S. international student achievement test performance) for the claim by ANAR in 1983 that the United States' "once unchallenged preeminence in commerce, industry, science, and technological innovation is being overtaken by competitors throughout the world," as it turns out, is faulty (National Commission on Excellence in Education, 1983). Subsequent reform efforts were all based on the same faulty supposition: that is, if the United States did not make cognitive gains as evidenced by improved scores on international achievement tests, the nation would fall behind the rest of the world competitively (Alexander, 1986; National Commission on Excellence in Education, 1983; Department of Education, 1991, 1994, 2011, 2012a, 2014). As Wu (2010) observed, the statistical complexities of large-scale assessments often lead to misguided conclusions and inappropriate policy decisions. At the very least, this finding raises a reasonable doubt about whether policymakers should continue to use international achievement test outcomes as a basis for public school reform initiatives.

The findings of this study show that the law of diminishing returns applies to investments in education. There is a moderate correlation between investments in education and competiveness for developing nations that disappears for advanced economies in the top quartile of competitiveness. The idea that cognitive gains through schooling at the primary and secondary levels have no correlation with the economic competitiveness of the United States seems extraordinarily counterintuitive and may be unsettling to readers. Meyer, Boli, Thomas, and Ramirez (1997) argue that the calamities of World War II and the subsequent Cold War gave rise to economic theories of human progress and development. The idea that nation-states can directly influence socioeconomic development has been deeply institutionalized at a global level and has become the basis for the production and modification of societal structures (Meyer et al., 1997). This idea, however, is frequently at odds with inconvenient realities. For example, the societal belief that mass schooling is necessary and beneficial for economic growth often goes unquestioned, even though the functional correlation between mass education as a societal structure and economic growth is weak and highly conditional (Meyer et al., 1997).

Given the enormous allocation of resources devoted to education reform, policymakers should know if this money will indeed translate into the United States remaining competitive in the world. The findings of this study provide further evidence that student performance on international assessments provides no accurate basis for claims that the global competitiveness of the United States is in jeopardy. Previous research has shown that as nations become more educated, it becomes more difficult to raise national income through continued investments in education (Breton, 2013; Hanushek & Woessmann, 2009; Tienken, 2008). It also appears to be the case that as nations become richer and more educated, it becomes more difficult to justify education reforms that revolve around accountability as measured by high-stakes testing. To maximize the return on investments in education, money and reform initiatives should be channeled toward addressing poverty, funding schools equitably, alleviating social stress and violence, and supporting young families and students of immigrant families (Anil, 2011; Cavanagh, 2007a; Harvey et al., 2015; Perelman & Santin, 2011).

These findings are consistent with Tienken (2008), who found that ISAT performance does not correlate with future economic strength for advanced economies like the United States. They are also consistent with Hanushek and Woessmann (2009), who found the convergence effect of investment in top-performing students on economic growth to be stronger in underdeveloped countries than in advanced economies (Hanushek & Woessmann, 2009), and with Breton (2013), who found that investments in education increased marginal productivity for countries with a highly educated population by just 12 percent, compared to marginal productivity gains of greater than 50 percent for countries with less educated populations.

The rationale for primary and secondary public education reform in the United States is that a failure to repair a supposedly broken education system will cause the U.S. to fall behind the rest of the world competitively (Alexander, 1986; National Commission on Excellence in Education, 1983; Department of Education, 1991, 1994, 2012a; National Governors Association, 2009). The public has become skeptical of federal education reform initiatives. A recent Phi Delta Kappan/Gallup poll reveals that most Americans perceive education reform efforts in the United States to be federally orchestrated and do not support change initiatives they consider to be pushed on local and state education agencies by federal policymakers (Bushaw & Calderon, 2014). From an advocacy standpoint, this study should prompt policymakers, local school board members, and superintendents to critically investigate the research behind K–12 public education reform initiatives. For a nation like the United States to remain competitive in the world, policymakers need to work on enhancing business sophistication and innovations in research and development through quality scientific research institutions, corporate spending on research and development, university–industry collaboration in research and development, government procurement of advanced technology products, availability of scientists and engineers, patent applications, and intellectual property protection (Schwab, 2014).

School reform policies should address social inequalities, violence, support for young and impoverished families, the achievement gap, on-time graduation rates, funding equity, and teacher quality (Harvey et al., 2015). False alarms about a dire future stand in stark contrast to this and other research that show the United States produces many more high-achieving students than any other nation across racial and ethnic lines (Bracey, 2009; Petrilli & Scull, 2011). Any number of justifications for reform may be valid, but they should be researched and proven to have a correlation to the desired outcome (and better yet, be proven to cause the desired outcome) before the public is made to believe they are true and billions of dollars are spent on staircases that lead to nowhere and doors that open into walls.

Note

1. Reading tests were excluded because math and science achievement receives the most attention from policymakers in the United States.

Legislation

America 2000: Excellence in Education Act. H.R. 2460, 102nd Cong. (1991)
Elementary and Secondary Education Act of 1965, Pub. L. No. 89-10 (1965)
Goals 2000: Educate America Act of 1994, Pub. L. No. 103-227 (1994)
No Child Left Behind Act of 2001, Pub. L. No. 107-110 (2001)

References

Acar, T. (2012). The position of Turkey among OECD member and candidate countries according to PISA 2009 results. *Educational Sciences: Theory & Practice, 12*(4), 2567–2572.

Alegre, M.A., & Ferrer, G. (2010). School regimes and education equity: Some insights based on PISA 2006. *British Educational Research Journal, 36*(3), 433–461. doi: 10.1080/01411920902989193

Alexander, L. (1986). Time for results. *Phi Delta Kappan, 68*(3), 202–204.

Ammermueller, A. (2007). Poor background or low returns? Why immigrant students in Germany perform so poorly in the programme for international student assessment. *Education Economics, 15*(2), 215–230. doi: 10.1080/09645290701263161

Baker, K. (2007). Are international tests worth anything? *Phi Delta Kappan, 89*(2), 101–104.

Benhabib, J., & Spiegel, M. (2002). Human capital and technology diffusion. FRBSF Working Paper #2003-02. Retrieved September 12, 2016, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=634128.

Bracey, G. (2009). The big tests: What ends do they serve? *Educational Leadership, 67*(3), 32–37.

Breton, T.R. (2013). The role of education in economic growth: Theory, history and current returns. *Educational Research, 55*(2), 121–138. doi:10.1080/00131881.2013.801241

Bushaw, W., & Calderon, V. (2014). Try it again, Uncle Sam: 46th annual PDK/Gallup poll of the public's attitudes toward the public schools. *Phi Delta Kappan, 96*(1), 8–20. doi: 10.1177/0031721714547856

Cavanagh, S. (2007a). Poverty's effect on U.S. scores greater than for other nations. *Education Week, 27*(15), 1–13.

Cavanagh, S. (2007b). Against other nations, U.S. below par in science. *Education Week, 27*(14), 11.

Cohen, L., & Manion, L. (1994). *Research methods in education* (4th ed.). London and New York: Routledge.

Department of Commerce. (2012). *The competitiveness and innovative capacity of the United States*. Retrieved September 12, 2016 from http://www.esa.doc.gov/reports/competitiveness-and-innovative-capacity-united-states.

Department of Education. (1991). *America 2000: An education strategy*. Sourcebook. Washington, D.C. Retrieved from ERIC database (ED327985).

Department of Education. (1994). *Goals 2000: Educate America Act.* Public Law No. 103-227. Retrieved from http://www2.ed.gov/legislation/GOALS2000/TheAct/index.html .

Department of Education. (2011). *A blueprint for reform: The Reauthorization of the Elementary and Secondary Education Act.* Retrieved September 12, 2016, from http://www2.ed.gov/policy/elsec/leg/blueprint/publicationtoc.html .

Department of Education. (2012a). *President Obama: Our children can't wait for Congress to fix No Child Left Behind, announces flexibility in exchange for reform for ten states.* Retrieved September, 12, 2016, from http://www.ed.gov/news/press-releases/president-obama-our-children-cant-wait-congress-fix-no-child-left-behind-announc.

Department of Education. (2012b). *The federal role in education*. Retrieved September 12, 2016, from http://www2.ed.gov/about/overview/fed/role.html?src=ln.

Department of Education. (2014). *ESEA flexibility*. Retrieved September 12, 2016, from http://www2.ed.gov/policy/elsec/guid/esea-flexibility/index.html.

Finkel, E. (2012). What can U.S. schools learn from foreign counterparts? *District Administration, 48*(2), 30–32.

Hanushek, E.A., & Kimko, D.D. (2000). Schooling, labor-force quality, and the growth of nations. *American Economic Review, 90*(5), 1184–1208.

Hanushek, E.A., & Woessmann, L. (2007). *The role of education quality for economic growth*. The World Bank Policy Research Working Paper Series, No. 4122. Retrieved September 12, 2016, from http://ideas.repec.org/p/wbk/wbrwps/4122.html.

Hanushek, E.A., & Woessmann, L. (2009). *Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation*. NBER Working Paper No. 14633. National Bureau of Economic Research. Retrieved September 12, 2016, from http://www.nber.org/papers/w14633.

Hanushek, E.A., & Woessmann, L. (2010a). *Sample selectivity and the validity of international student achievement tests in economic research*. NBER Working Paper No. 15867. National Bureau of Economic Research. Retrieved September 12, 2016, from http://www.nber.org/papers/w15867.

Hanushek, E.A., & Woessmann, L. (2010b). *The economics of international differences in educational achievement*. NBER Working Paper No. 15949. National Bureau of Economic Research. Retrieved September 12, 2016, from http://www.nber.org/papers/w15949.

Harvey, J., McKay, J., Fowler, C., & Marx, G. (2015). *School performance in context: Indicators of school inputs and outputs in nine similar nations*. Port Ludlow, WA: Horace Mann League and National Superintendents Roundtable. Retrieved September 12, 2016, from http://www.hmleague.org/school-performance-in-context-pdf-versions/.

Hoi Yan, C., & Chan, A.H. (2008). Understanding the relationships among PISA scores, economic growth and employment in different sectors: A cross-country study. *Research in Education, 80*, 93–106.

Jun, Y., Xiao, H., & Xiaoyu, L. (2009). Educational inequality and income inequality: An empirical study on China. *Frontiers of Education in China, 4*(3), 413–434. doi:10.1007/s11516-009-0022-1

Karp, S. (2014). The problems with the common core. *Rethinking Schools*, *28*(2). Retrieved September 12, 2016, from http://www.rethinkingschools.org/archive/28_02/28_02_karp.shtml.

Loveless, T. (2009). How well are American students learning? *The 2008 Brown Center Report on American Education*, *II*(3). Retrieved September 12, 2016, from https://www.brookings.edu/research/2016-brown-center-report-on-american-education-how-well-are-american-students-learning/.

Martin, A.J., Liem, G.A.D., Mok, M.M.C., & Xu, J. (2012). Problem solving and immigrant student mathematics and science achievement: Multination findings from the programme for international student assessment (PISA). *Journal of Educational Psychology, 104*(4), 1054–1073. doi: 10.1037/a0029152

Meyer, J.W., Boli, J., Thomas, G.M., & Ramirez, F.O. (1997). World society and the nation-state. *American Journal of Sociology, 103*(1), 144.

Miller, T., & Kim, A. (2015). *2015 Index of economic freedom*. Washington, DC: The Heritage Foundation and Dow Jones & Company, Inc. Retrieved September 12, 2016 from http://www.heritage.org/index/download.

National Commission on Excellence in Education. (1983). *A nation at risk*. Report. Washington, D.C. Retrieved September 12, 2016, from https://www2.ed.gov/pubs/NatAtRisk/risk.html .

National Governors Association. (2009). *Forty-nine states and territories join Common Core Standards Initiative*. Retrieved August 16, 2016, from http://www.nga.org/cms/home/news-room/news-releases/page_2009/col2-content/main-content-list/forty-nine-states-and-territorie.html .

OECD. (2010). *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes*. OECD Publishing. doi: 10.1787/9789264077485-en

Perelman, S., & Santin, D. (2011). Measuring educational efficiency at student level with parametric stochastic distance functions: An application to Spanish PISA results. *Education Economics, 19*(1), 29-49. doi: 10.1080/09645290802470475

Perry, L. (2009). Characteristics of equitable systems of education: A cross-national analysis. *European Education, 41*(1), 79–100. doi: 10.2753/EUE1056-4934410104

Petrilli, M.J., & Scull, J. (2011). *American achievement in international perspective*. Thomas B. Fordham Institute. Retrieved from ERIC database (ED518219).

Pons, X. (2011). What do we really learn from PISA? The sociology of its reception in three European countries (2001–2008). *European Journal of Education, 46*(4), 540–548. doi: 10.1111/j.1465-3435.2011.01499.x

Schwab, K. (2014). *The global competitiveness report 2014–2015*. Geneva, Switzerland: World Economic Forum. Retrieved September 12, 2016, from http://reports.weforum.org/global-competitiveness-report-2014-2015/.

Tanner, D. (2012). *Using statistics to make educational decisions*. Thousand Oaks, CA: Sage Publications.

Tienken, C. H. (2008). Rankings of international achievement test performance and economic strength: Correlation or conjecture? *International Journal of Education Policy and Leadership*, *3*(4), 1–15.

UNICEF Innocenti Research Centre. (2012). *Measuring child poverty: New league tables of child poverty in the world's rich countries*. Innocenti Report Card 10, UNICEF. Innocenti Research Centre, Florence. Retrieved September 12, 2016, from http://www.unicef-irc.org/publications/pdf/rc10_eng.pdf.

Yu, C. (2012). Beyond gross national product: An exploratory study of the relationship between program for international student assessment scores and well-being indices. *Review of European Studies, 4*(5), 119–134. doi: 10.5539/res.v4n5p119

Yu, C., DiGangi, S., & Jannasch-Pennell, A. (2012). A time-lag analysis of the relationships among PISA scores, scientific research publication, and economic performance. *Social Indicators Research*, 107(2), 317–330.

Wu, M. (2010). Measurement, sampling, and equating errors in large-scale assessments. *Educational Measurement: Issues & Practice, 29*(4), 15–27. doi: 10.1111/j.1745-3992.2010.00190

Appendix A
International data on cognitive skills
(Sorted alphabetically)

Code	Country	gsample	cognitive	lowsec	basic	top
ALB	Albania	0	3.785	3.785	0.424	0.013
ARG	Argentina	1	3.920	3.920	0.492	0.027
ARM	Armenia	0	4.429	4.490	0.745	0.008
AUS	Australia	1	5.094	5.138	0.938	0.112
AUT	Austria	1	5.089	5.090	0.931	0.097
BEL	Belgium	1	5.041	5.072	0.931	0.094
BGR	Bulgaria	0	4.789	4.789	0.765	0.083
BHR	Bahrain	0	4.114	4.114	0.608	0.003
BRA	Brazil	1	3.638	3.638	0.338	0.011
BWA	Botswana	0	3.575	3.575	0.374	0.000
CAN	Canada	1	5.038	5.125	0.948	0.083
CHE	Switzerland	1	5.142	5.102	0.919	0.134
CHL	Chile	1	4.049	3.945	0.625	0.013
CHN	China	1	4.939	4.939	0.935	0.083
COL	Colombia	1	4.152	4.152	0.644	0.000
CYP	Cyprus	1	4.542	4.413	0.825	0.011
CZE	Czech Rep.	0	5.108	5.177	0.931	0.122
DNK	Denmark	1	4.962	4.869	0.888	0.088
EGY	Egypt	1	4.030	4.030	0.577	0.010
ESP	Spain	1	4.829	4.829	0.859	0.079
EST	Estonia	0	5.192	5.192	0.973	0.095
FIN	Finland	1	5.126	5.173	0.958	0.124
FRA	France	1	5.040	4.972	0.926	0.085
GBR	United Kingdom	1	4.950	4.995	0.929	0.088
GER	Germany	0	4.956	4.959	0.906	0.105
GHA	Ghana	1	3.603	3.252	0.403	0.010
GRC	Greece	1	4.608	4.618	0.798	0.042
HKG	Hong Kong	1	5.195	5.265	0.944	0.123
HUN	Hungary	0	5.045	5.134	0.941	0.103
IDN	Indonesia	1	3.880	3.880	0.467	0.008
IND	India	1	4.281	4.165	0.922	0.013
IRL	Ireland	1	4.995	5.040	0.914	0.094
IRN	Iran	1	4.219	4.262	0.727	0.006
ISL	Iceland	1	4.936	4.945	0.908	0.074
ISR	Israel	1	4.686	4.660	0.826	0.053
ITA	Italy	1	4.758	4.693	0.875	0.054
JOR	Jordan	1	4.264	4.264	0.662	0.044
JPN	Japan	1	5.310	5.398	0.967	0.168
KOR	Korea, Rep.	1	5.338	5.401	0.962	0.178
KWT	Kuwait	0	4.046	4.223	0.575	0.000
LBN	Lebanon	0	3.950	3.950	0.595	0.002
LIE	Liechtenstein	0	5.128	5.128	0.860	0.198
LTU	Lithuania	0	4.779	4.694	0.891	0.030
LUX	Luxembourg	0	4.641	4.641	0.776	0.067
LVA	Latvia	0	4.803	4.779	0.869	0.050
MAC	Macao-China	0	5.260	5.260	0.919	0.204
MAR	Morocco	1	3.327	3.243	0.344	0.001
MDA	Moldova	0	4.530	4.419	0.787	0.029
MEX	Mexico	1	3.998	3.998	0.489	0.009

Appendix A (continued)

Code	Country	gsample	cognitive	lowsec	basic	top
MKD	Macedonia	0	4.151	4.151	0.609	0.028
MYS	Malaysia	1	4.838	4.838	0.864	0.065
NGA	Nigeria	0	4.154	4.163	0.671	0.001
NLD	Netherlands	1	5.115	5.149	0.965	0.092
NOR	Norway	1	4.830	4.855	0.894	0.056
NZL	New Zealand	1	4.978	5.009	0.910	0.106
PER	Peru	1	3.125	3.125	0.182	0.002
PHL	Philippines	1	3.647	3.502	0.485	0.006
POL	Poland	0	4.846	4.861	0.838	0.099
PRT	Portugal	1	4.564	4.592	0.803	0.032
PSE	Palestine	0	4.062	4.062	0.571	0.008
ROM	Romania	1	4.562	4.562	0.780	0.046
RUS	Russian Fed.	0	4.922	4.906	0.884	0.081
SAU	Saudi Arabia	0	3.663	3.663	0.331	0.000
SGP	Singapore	1	5.330	5.512	0.945	0.177
SRB	Serbia	0	4.447	4.447	0.718	0.024
SVK	Slovak Rep.	0	5.052	5.052	0.906	0.112
SVN	Slovenia	0	4.993	5.076	0.939	0.061
SWE	Sweden	1	5.013	4.948	0.939	0.088
SWZ	Swaziland	0	4.398	4.398	0.801	0.004
THA	Thailand	1	4.565	4.556	0.851	0.019
TUN	Tunisia	1	3.795	3.889	0.458	0.003
TUR	Turkey	1	4.128	4.128	0.582	0.039
TWN	Taiwan (Chinese Taipei)	1	5.452	5.599	0.958	0.219
URY	Uruguay	1	4.300	4.300	0.615	0.049
USA	United States	1	4.903	4.911	0.918	0.073
ZAF	South Africa	1	3.089	2.683	0.353	0.005
ZWE	Zimbabwe	1	4.107	4.107	0.684	0.010

For details, see	Eric A. Hanushek, Ludger Woessmann. 2009. *Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation.*
Variables	
gsample	Indicator of whether country is in the main sample of 50 countries contained in the growth regressions, for which internationally comparable GDP data is available.
cognitive	Average test score in math and science, primary through end of secondary school, all years (scaled to PISA scale divided by 100).
lowsec	Average test score in math and science, only lower secondary, all years (scaled to PISA scale divided by 100).
basic	Share of students reaching basic literacy (based on average test scores in math and science, primary through end of secondary school, all years).
top	Share of top-performing students (based on average test scores in math and science, primary through end of secondary school, all years).

Appendix B
Global Competitive Index Table

The Global Competitiveness Index 2014-2015 rankings

	GCI 2014-2015		GCI 2013-2014
Country/Economy	Rank (out of 144)	Score	Rank (out of 148)
Switzerland	1	5.70	1
Singapore	2	5.65	2
United States	3	5.54	5
Finland	4	5.50	3
Germany	5	5.49	4
Japan	6	5.47	9
Hong Kong SAR	7	5.46	7
Netherlands	8	5.45	8
United Kingdom	9	5.41	10
Sweden	10	5.41	6
Norway	11	5.35	11
United Arab Emirates	12	5.33	19
Denmark	13	5.29	15
Taiwan, China	14	5.25	12
Canada	15	5.24	14
Qatar	16	5.24	13
New Zealand	17	5.20	18
Belgium	18	5.18	17
Luxembourg	19	5.17	22
Malaysia	20	5.16	24
Austria	21	5.16	16
Australia	22	5.08	21
France	23	5.08	23
Saudi Arabia	24	5.06	20
Ireland	25	4.98	28
Korea, Rep.	26	4.96	25
Israel	27	4.95	27
China	28	4.89	29
Estonia	29	4.71	32
Iceland	30	4.71	31
Thailand	31	4.66	37
Puerto Rico	32	4.64	30
Chile	33	4.60	34
Indonesia	34	4.57	38
Spain	35	4.55	35
Portugal	36	4.54	51
Czech Republic	37	4.53	46
Azerbaijan	38	4.53	39
Mauritius	39	4.52	45
Kuwait	40	4.51	36
Lithuania	41	4.51	48
Latvia	42	4.50	52
Poland	43	4.48	42
Bahrain	44	4.48	43
Turkey	45	4.46	44
Oman	46	4.46	33
Malta	47	4.45	41

Appendix B (continued)

	GCI 2014-2015		GCI 2013-2014
Country/Economy	Rank (out of 144)	Score	Rank (out of 148)
Panama	48	4.43	40
Italy	49	4.42	49
Kazakhstan	50	4.42	50
Costa Rica	51	4.42	54
Philippines	52	4.40	59
Russian Federation	53	4.37	64
Bulgaria	54	4.37	57
Barbados	55	4.36	47
South Africa	56	4.35	53
Brazil	57	4.34	56
Cyprus	58	4.31	58
Romania	59	4.30	76
Hungary	60	4.28	63
Mexico	61	4.27	55
Rwanda	62	4.27	66
Macedonia, FYR	63	4.26	73
Jordan	64	4.25	68
Peru	65	4.24	61
Colombia	66	4.23	69
Montenegro	67	4.23	67
Vietnam	68	4.23	70
Georgia	69	4.22	72
Slovenia	70	4.22	62
India	71	4.21	60
Morocco	72	4.21	77
Sri Lanka	73	4.19	65
Botswana	74	4.15	74
Slovak Republic	75	4.15	78
Ukraine	76	4.14	84
Croatia	77	4.13	75
Guatemala	78	4.10	86
Algeria	79	4.08	100
Uruguay	80	4.04	85
Greece	81	4.04	91
Moldova	82	4.03	89
Iran, Islamic Rep.	83	4.03	82
El Salvador	84	4.01	97
Armenia	85	4.01	79
Jamaica	86	3.98	94
Tunisia	87	3.96	83
Namibia	88	3.96	90
Trinidad and Tobago	89	3.95	92
Kenya	90	3.93	96
Tajikistan	91	3.93	n/a
Seychelles	92	3.91	80
Lao PDR	93	3.91	81
Serbia	94	3.90	101
Cambodia	95	3.89	88
Zambia	96	3.86	93

Appendix B (continued)

	GCI 2014-2015		GCI 2013-2014
Country/Economy	Rank (out of 144)	Score	Rank (out of 148)
Albania	97	3.84	95
Mongolia	98	3.83	107
Nicaragua	99	3.82	99
Honduras	100	3.82	111
Dominican Republic	101	3.82	105
Nepal	102	3.81	117
Bhutan	103	3.80	109
Argentina	104	3.79	104
Bolivia	105	3.77	98
Gabon	106	3.74	112
Lesotho	107	3.73	123
Kyrgyz Republic	108	3.73	121
Bangladesh	109	3.72	110
Suriname	110	3.71	106
Ghana	111	3.71	114
Senegal	112	3.70	113
Lebanon	113	3.68	103
Cape Verde	114	3.68	122
Côte d'Ivoire	115	3.67	126
Cameroon	116	3.66	115
Guyana	117	3.65	102
Ethiopia	118	3.60	127
Egypt	119	3.60	118
Paraguay	120	3.59	119
Tanzania	121	3.57	125
Uganda	122	3.56	129
Swaziland	123	3.55	124
Zimbabwe	124	3.54	131
Gambia, The	125	3.53	116
Libya	126	3.48	108
Nigeria	127	3.44	120
Mali	128	3.43	135
Pakistan	129	3.42	133
Madagascar	130	3.41	132
Venezuela	131	3.32	134
Malawi	132	3.25	136
Mozambique	133	3.24	137
Myanmar	134	3.24	139
Burkina Faso	135	3.21	140
Timor-Leste	136	3.17	138
Haiti	137	3.14	143
Sierra Leone	138	3.10	144
Burundi	139	3.09	146
Angola	140	3.04	142
Mauritania	141	3.00	141
Yemen	142	2.96	145
Chad	143	2.85	148
Guinea	144	2.79	147