

Background Paper The Learning Generation

The Quality of Education Systems and Education Outcomes

John L. Newman, Elizabeth M. King and Husein Abdul-Hamid

This paper was prepared for the International Commission on Financing Global Education Opportunity as a background paper for the report, *The Learning Generation: Investing in education for a changing world*. The views and opinions in this background paper are those of the author(s) and are not endorsed by the Education Commission or its members. For more information about the Commission's report, please visit: report.educationcommission.org.

The Quality of Education Systems and Education Outcomes

John L. Newman, Elizabeth M. King and Husein Abdul-Hamid¹

Revised November 14, 2016

Abstract

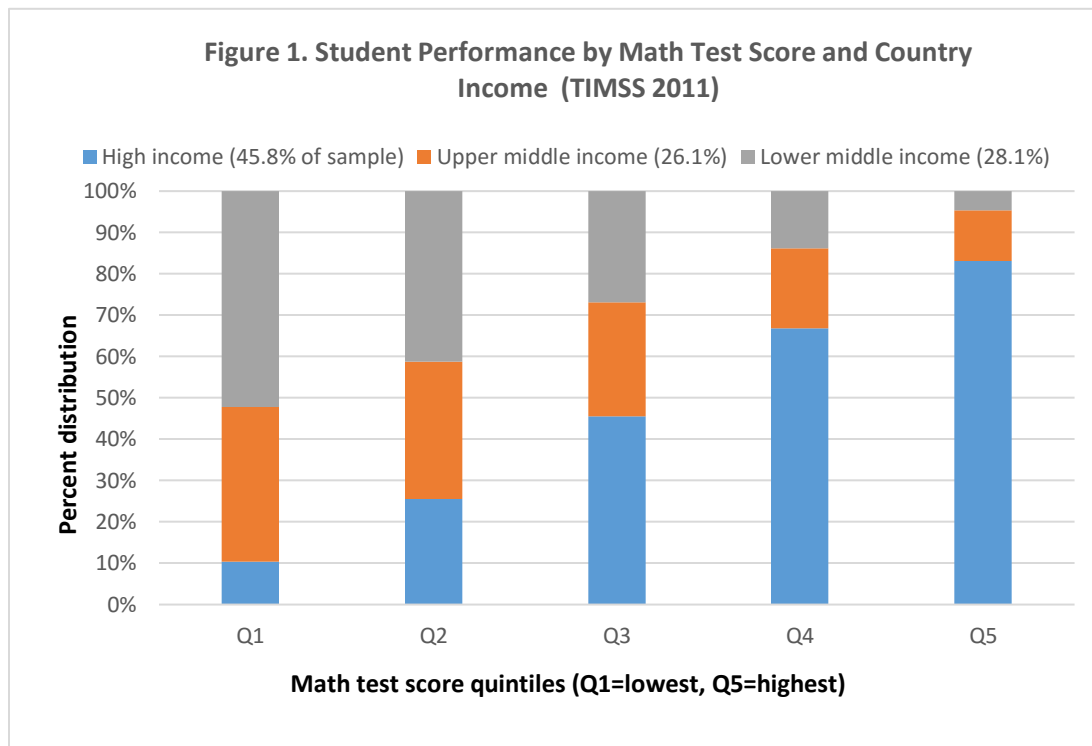
This paper examines the relationship between the quality of a country's educational system and education outcomes. Previous country-level analyses of the factors that affect education outcomes have focused on measures of school inputs such as average pupil-teacher ratios, teacher characteristics, and availability of learning materials. The capacity of the education system to translate these inputs into education outcomes has not been researched in the same way because internationally comparable measures of the quality of the system have not been available. This paper takes advantage of a unique database which has system-level data on different education policy areas for a large number of countries. It also takes advantage of the greater availability of comparable learning data now across countries, in addition to country data on recent completed years of schooling and people's view of their education system. Specifically, the paper sets out to examine whether countries with a better education system have, on average, better education outcomes, controlling for their level of education expenditure and per-capita GDP. Although the data on the quality of the education system can be improved significantly, the findings indicate consistently that better education systems do yield better education outcomes.

¹ KenAct, LLP; The Brookings Institution and the Global Education Commission Research Team; The World Bank. This paper was prepared as a background paper for the International Commission on Financing Global Education, launched at the UN General Assembly in September 2015. We are grateful to Dandan Chen who provided excellent assistance with compiling and organizing the database for this analysis. We also thank Luis Benveniste, Deon Filmer, Halsey Rogers, and other education staff of the World Bank for their useful comments; and Liesbet Steer, Ore Oluwa Badaki, Nicholas Burnett, Bridget Crimpton, Paul Isenman and Babette Wils for their comments on earlier versions of our work. Any remaining errors are entirely ours.

I. WHY A SYSTEMS APPROACH TO EDUCATION REFORM

In the last several years, there has been increased interest in exploring not only how educational inputs but also how the quality of the education system can improve education outcomes. This interest has grown in both developing and advanced countries because of widespread concerns that school systems are falling behind the times and that young job seekers may not be prepared for the labor markets that await them. More cross-country evidence on student learning is now available, revealing that, despite increasing school enrollment and completion rates across the globe, significant numbers of students do not achieve minimum levels of reading comprehension and math skills. Regional and international student assessments show huge differences across countries as well as within countries.

To illustrate, Figure 1 shows that globally more than one-half of students who tested at the lowest levels (Q1) in 2011 are in low middle-income countries, about one-third are in upper middle-income countries, and one-tenth are in high-income countries. Some middle-income countries do have students who perform as well as those in high-income countries (Q5), but they account for less than 20 percent of the top performers—less than 5 percent in low-middle income countries and about 10 percent in the upper-middle income countries. The learning challenge looms large for the education systems of middle-income countries; the challenge is even greater for low-income countries.



Education systems are large, complex organizations. Familiarly, they consist of teachers, school leaders and school administrators, and students—from pre-schooling to tertiary education. Behind the frontline deliverers of instruction and learning are leaders, managers and administrators in national and local education agencies whose roles are to set and implement policies and priorities, allocate budgets, and oversee the schools. The schools (and universities, vocational or technical training programs) include not only state-owned institutions, but also non-state institutions that are owned and operated by the private enterprise sector, faith-based organizations, or private, non-profit organizations and individuals. Further, either as formal units within the system or as private contractors, there are publishers of textbooks and learning materials, builders and maintenance staff of schools, providers of transport or school meals, and so on.

What we emphasize here is that education systems encompass not only various sets of actors and physical plants, but also the connective tissues that allow those actors and units to relate and work together. They include the accountability relationships determined by the standards for selecting teachers and administrators and for developing academic curricula, the rules and incentives guiding the school operation and the behaviors of teachers and learners, the employment contracts defining the responsibilities, the compensation packages and career paths of the education workforce, and the financing and information mechanisms that keep schools operating. When these standards, rules, accountability relationships and financing levels are aligned towards shared education goals, the education system as a whole, in all its complexity and size, is coherent and able to perform well.

These interdependencies within an education system mean that to improve education, it is not enough for system leaders and managers to identify a program or intervention (or even a package of programs) that works; they must also ensure that the system as a whole has what it takes to benefit from that program or intervention. The top education systems in the world are not necessarily the best-financed systems, but the systems that manage their financial resources and talent well towards clear objectives and that inspire leaders, teachers, parents and students to work towards those improvements. In contrast, weak education systems struggle to achieve that alignment and coherence. Their goals, standards, and rules are not clearly defined; the inputs are inappropriate or inadequate; subsystems function poorly because resources are used inefficiently, information about inputs and outcomes are inadequate, and accountability mechanisms are weak; and they are not sufficiently dynamic to adjust to shifts in socioeconomic and political contexts, or to changes in the available financial and management capacities.

There have been recent efforts to assess the quality of education systems and what it takes to improve them. A McKinsey study published in 2010 assessed 20 national and city education systems and used their experience to define an improvement path from low performance to better performance—from poor to fair, to good, to great, and to excellent (Mourshed et al.,

2010). In the *poor to fair* stage, system leaders must choose the set of interventions that support students in achieving the literacy and math basics by providing scaffolding for low-skilled teachers, fulfilling basic student needs, and bringing all the schools up to a minimum quality threshold. In the *fair to good* stage, the interventions must focus on consolidating the system foundations which include the production of high quality performance data, ensuring teacher and school accountability, and creating appropriate financing, organization structure, and pedagogy models. In the *good to great* stage, the interventions must focus on professionalizing the teaching force and school leadership, putting in place clearly defined career paths as those in medicine and law. Finally, in the *great to excellent* stage, the interventions must move the locus of improvement from the center to the schools themselves, by introducing peer-based learning through school-based and system-wide interaction, as well as supporting innovation and experimentation. It's not clear that an education system has to follow these stages in sequence. The study also listed six cross-stage interventions which serve as the fundamentals for any improvement: revising the curriculum and standards, ensuring an appropriate reward and remunerations structure for teachers and principals, building the technical skills of teachers and principals, assessing students, establishing data systems, and facilitating improvement through the introduction of policy documents and education laws (Mourshed et al., 2010).

These fundamentals roughly correspond to the conceptual framework that is driving the ongoing work under the Research on Improving Systems of Education (or RISE) program jointly managed by the Oxford Management Institute and the Blavatnik School of Government in Oxford University, and funded by the UK's DFID. Pritchett's (2015) framework for the RISE program defines education systems as a set of accountability relationships that must be aligned in order to improve learning outcomes—"delegation" which means that the responsible actors are focused on promoting learning; financing the actors and the programs that contribute to learning; using information to measure and monitor learning outcomes; and aligning incentives and rewards with improvements in learning. According to this framework, without a coherent education system built to improve learning, even rigorously proven interventions (such as higher teacher pay, greater autonomy of teachers over classroom practices, more textbooks, smaller class sizes) that work in some contexts will not necessarily produce better learning in other contexts. Discrete interventions or investments, however well designed and executed, will not deliver lasting reform. They are likely to work only if they address a weakness in any of the above accountability relationships.²

² The 6-year RISE program is funding major analytical pieces in several developing countries to examine the effect of different policy levers and investments on education outcomes and to understand how these reforms bring about system-wide change. See the RISE website for details: <http://www.opml.co.uk/projects/rise-improving-education-systems-low-income-countries-0>.

In 2010, as part of a new education strategy centered on learning,³ the World Bank launched its Systems Approach to Better Education Results (or SABER) initiative. The SABER program has been collecting data on the policies and institutions of education systems around the world, and benchmarks them against practices associated with rapid learning.⁴ The metrics are designed principally to identify policy areas which are weak and in need of improvement in a particular country, using a structured questionnaire to underpin that analysis. This ongoing effort has now generated a database that captures different aspects of the education systems in about 100 countries, based on measures that can be easily compared across education systems. The program aims to give all parties with a stake in educational results—from students and teachers and parents to policymakers, business people, and political leaders—an accessible, detailed, objective, and up-to-date snapshot of how well their country's education system is oriented toward delivering learning. This paper analyzes the relationship between the measures of the quality of education systems and education outcomes.

II. OBJECTIVES

This paper is a study, sponsored by the International Commission on Financing Global Education (the Global Education Commission, for short) jointly with the World Bank, to examine the relationship between measures of the quality of a country's education system and education outcome measures that are available for international comparisons. Specifically, our research sought to answer the following question—*whether countries with a better education system do achieve better education outcomes, controlling for their level of per-capita GDP, per-student spending level for education, and average education level of the adult population*. The four education outcomes we examine are:

1. Average years of schooling attained by the age cohort 15-19 (Barro-Lee, 2010)⁵;
2. Percent of students reaching a minimum proficiency level in math and language (using test scores harmonized across several regional and international student assessments by Altinok et al., 2014)⁶;
3. Percent of students reaching an advanced proficiency level (also from Altinok et al. 2014);

³ As the World Bank's *Education Sector Strategy 2020: Learning for All* argues, effective action to promote learning requires a more balanced analysis of the whole education system, aimed at identifying the binding constraints to learning beyond geographic borders. While relaxing those constraints would often demand increasing resources, it may also require better mechanisms to measure student learning, deploy qualified teachers to schools in poor areas, or track education expenditures.

⁴ The SABER program initiated by the World Bank is in partnership with Australia Aid, DFID, Russia Aid and other donor agencies. This ongoing program is in line with the current World Bank's education strategy officially launched in 2011. See <http://saber.worldbank.org/index.cfm?indx=5>.

⁵ <http://www.barrolee.com/data/Barro Lee Human Capital Update 2010April08.pdf>. Data can be found at <http://datatopics.worldbank.org/education/wProjQuery/BPopModel.aspx>.

⁶ Altinok, Nadir, Claude Diebolt and Jean-Luc Demeulemeester, "A new international database on education quality: 1965-2010", *Applied Economics* 46 (11), 2014.

4. Percent of country respondents who say they are satisfied with their education system (Gallup World Poll, 2015).⁷

Our analysis uses education system data from the World Bank's SABER initiative. We explored alternative quantitative databases that have globally comparable metrics on the characteristics of the education system. For example, we are aware of the OECD's Teaching and Learning International Survey (or TALIS) program which collects data from teachers and school leaders, but in 2013, the survey covered just 34, mostly high-income, countries.⁸ We are also aware of systems data contained in reports provided by individual countries to UNESCO's International Bureau of Education (IBE); although the country reports cover similar policy dimensions of the education system, they are difficult to use in a quantitative analysis without developing a set of rubrics to transform the information into comparable data.⁹

As a robustness check to the SABER data, we also analyze the education component of the World Bank's Country Policy and Institutional Assessment (CPIA) index.¹⁰ In contrast to the SABER data which are based on the responses to questionnaires on specific aspects of policies by experts and country respondents and on policy documents, CPIA education ratings are assessments given by World Bank staff on six main dimensions of the quality of the primary and secondary education in each country. These six dimensions are sector strategy, education management and information system, student assessments, teachers, education finance, and school-based management. Annex A describes the criteria used to rate each dimension in the CPIA database. The next section describes the SABER initiative in greater detail and briefly compares the SABER data with CPIA data.

Section IV presents descriptive statistics on education outcomes, the quality of the education system, and country characteristics. Section V displays the results from regression analyses of the relationships among these variables.

⁷ The Gallup survey collects data on a random sample of for more than 130 countries on a large number of questions, including questions about how people regard their education system. With some exceptions, all surveys of the Gallup World Poll (2015), either telephone or face-to-face, are probability based and nationally representative of the resident population aged 15 and above. See http://www.fao.org/fileadmin/templates/ess/voh/Gallup_world_poll_methodology.pdf.

⁸ The TALIS survey collects data regarding six thematic areas for education, including teaching practices, the classroom environment, and school leadership. For more information, see <http://www.oecd.org/edu/school/talis.htm>.

⁹ While the UNESCO country reports followed a guideline about what information to be included, they varied greatly in terms of the available detail on the features of the system. For example, to test the level of effort, we extracted data about the intended instructional time from these country reports and painstakingly transformed those into comparable time units (instruction hours per year), but even this exercise required making assumptions about the time and curriculum structure as not all the countries have this information. For more information, see <http://www.ibe.unesco.org/en/document/world-data-education-seventh-edition-2010-11>.

¹⁰ This program provides ratings on all sectors of the economy, such as agriculture, health, and energy. Our analysis uses the ratings only for the education sector.

III. THE SABER INITIATIVE

The SABER program collects comparable, well-defined, and disaggregated data on policies and institutions.¹¹ It carries out this data collection in education systems at different levels of development, including the systems that have been most successful at achieving high learning levels and those that have not met minimum learning levels, on average. The database covers policies and institutions in areas that, based on formal evidence or experience, appear to be the most important for determining student learning opportunity and outcomes.

The first step of the SABER program was to define the critical elements that countries have to get right in order to achieve the best outcomes from their education system. This step involved extensive reviews of research, global evidence and expert opinion regarding policy domains within an education system—teachers, financing, school governance, workforce development, student assessment, the role of the private sector, information system, and so on. It serves as the basis for the development of an analytical framework that highlights for policymakers and other stakeholders the most important (and actionable) policy choices to spur learning. This analysis led to framework papers on each of the identified policy domains, such as “What Matters for Teacher Policy” (Vegas and Ganimian, 2013) and “What Matters for Student Assessment” (Clarke, 2012), which survey the evidence and experience in the domain and uses that survey to identify the elements of the policy and institutional framework that matter most for improving education outcomes. These analytical framework papers formed the bases for the design of the questionnaires for each policy domain, the retrieval of relevant documents to support the responses to the questionnaires, including legislation, policy documents and regulations, and the development of specific rubrics that have guided the assessment of the country’s progress within each domain. These scoring rubrics are meant to be objective, rather than subjective, to ensure cross-country comparability and replicability.¹² By benchmarking the quality of country-level education policies against international best practices, SABER data highlight areas of strength and weakness, recognizing successful reformers whose experience can inform education policy and practice in other countries.

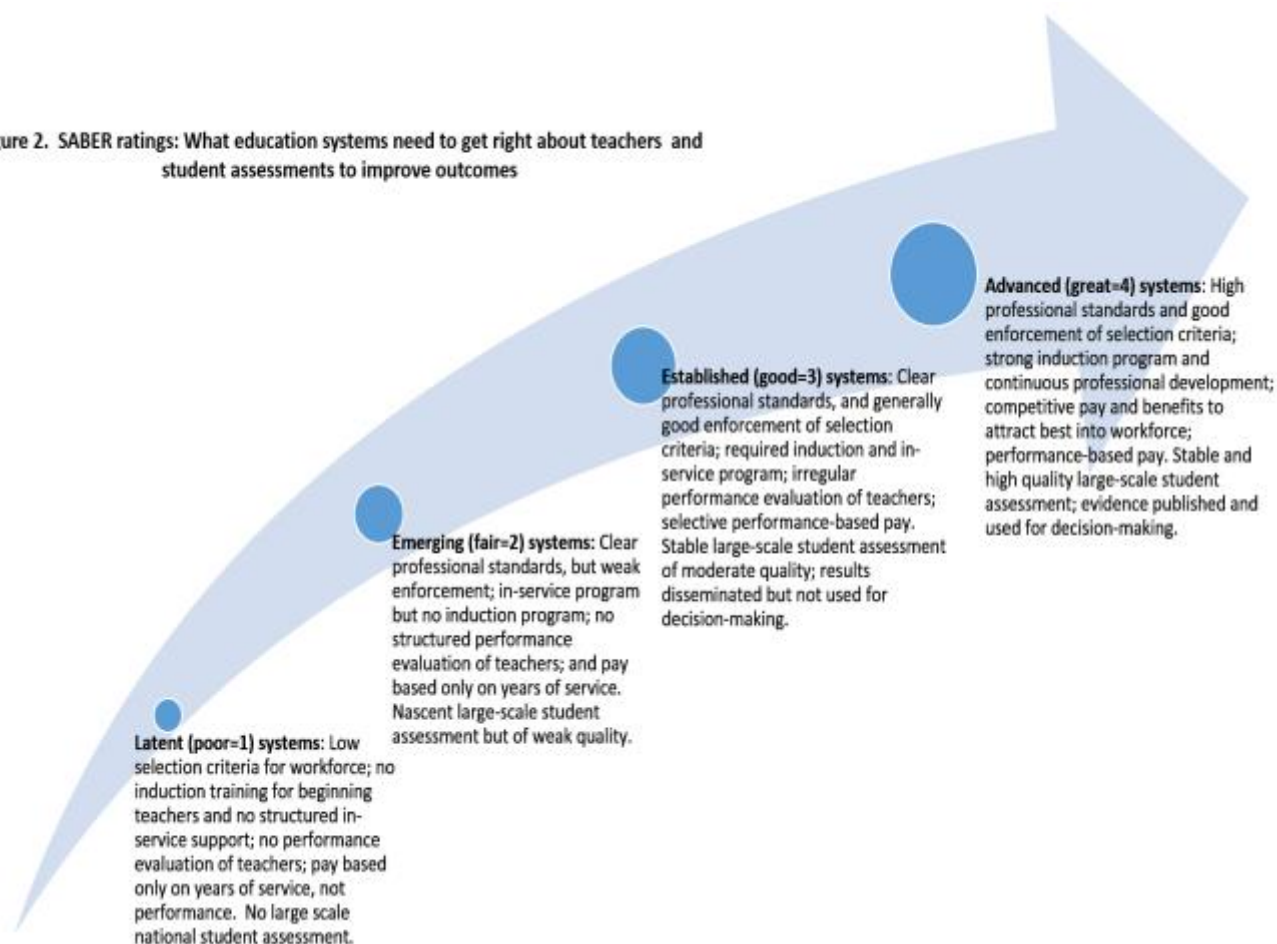
¹¹ For more information, <http://saber.worldbank.org/index.cfm>.

¹² The SABER team has developed a data-collection instrument appropriate for collecting data for the policy and institutional indicators. This instrument is essentially a survey for one respondent—an experienced principal investigator in the country—to fill out using information from key informants, documents, and other sources. In the typical model, an experienced principal investigator collects the policy information and data necessary to fill out the data-collection instrument, by drawing on his or her knowledge of the system and on government contacts. Data collection can usually be completed within a few weeks. An alternative approach used in some domains is to convene a workshop of experts, including government officials, and use that group process to collect the evidence and code data. In either case, data sources are clearly identified and available to the public as the data are posted online (World Bank, 2013).

For each policy domain, SABER data aggregate policy and institutional indicators to rate system development. The ratings are on a 4-level scale for each aspect of the domain: from “Latent” (with an index value of 1), to “Emerging” (value of 2), to “Established” (value of 3), and to “Advanced” (value of 4). To obtain an overall rating for a policy domain, the ratings on all aspects of that policy domain are averaged (unweighted); thus, each country gets several ratings, depending on which policy domains have been covered in that country. No overall country ranking is given, although all countries that have, say, applied the teacher questionnaire can be compared with each other.

A SABER index for a policy domain indicates whether or not the policies in that domain are as high in quality or maturity as what is regarded by global research and policy experts to be best

Figure 2. SABER ratings: What education systems need to get right about teachers and student assessments to improve outcomes



practice. The SABER program is ongoing and its data are not yet available across all policy domains for every country, so we are not able to examine the relationship between specific policy areas and education outcomes. Instead, our analysis is based on a country’s average rating across the domains for which data are available. We interpret this average index as reflecting the overall quality of the country’s education system. Figure 2 illustrates what the four-scale rating mean for two policy domains, teachers and student assessments.

The focus of the current SABER program is to document countries' education policies (*de jure*), although some of the current instruments contain elements about policy implementation (*de facto*) too.¹³ One might expect the quality of implementation to change more frequently, not necessarily in the forward direction in each political administration, while policies tend to be more "sticky" and, in theory at least, to serve as guiding norms for managing and operating the education system. A hard challenge for many countries is to reduce the discrepancy between good policies and actual implementation. A coherent education system is about aligning policies towards learning goals as well as about aligning practice with policy.

IV. DESCRIPTIVE ANALYSES

In the following analyses, we take the average SABER rating across policy domains to arrive at a single value for the quality index of the education system in each country. For some countries, this overall index is computed on the basis of ratings for seven domains, while in others, the index is based on only two or even just one domain, depending on how many SABER assessments have been undertaken for a country at the time of our analysis. In all, we estimate index values for almost 100 countries, but the intersection of countries with SABER data and those with student learning data yields a smaller sample size of 70.¹⁴

Figure 3 presents the position of countries with respect to four education outcomes. Pairing outcomes suggests that education systems that succeed in one outcome tend to do well also in the other outcomes. For example, countries with a larger percentage of students reaching the minimum proficiency level in learning assessments are also more likely to have a larger percentage of students meeting the advanced proficiency level (upper left panel), although this positive relationship is more marked in countries where the percentage of students reaching the minimum proficiency level exceeds 60 percent. As expected, in the countries where the share of students reaching minimum proficiency is below 60 percent, the share of students who achieve advanced competency hovers at 10 percent or lower. Also, although there is a positive relationship between average years of schooling and the student proficiency levels overall, some countries attain higher average years of schooling (e.g., eight years) without the majority of students reaching the minimum proficiency level (lower left panel). This finding suggests the need to focus on student learning, not just enrollment or years of schooling, in order to have a better measure of the effectiveness of the education system. In fact, the data indicate that, on

¹³ According to the SABER site (<http://saber.worldbank.org/index.cfm>), future work to improve the instruments to capture the quality of implementation is planned.

¹⁴ The SABER program is ongoing, so it should be possible to update this analysis with a larger number of countries and use more domains to estimate the quality index for each country. In addition, as more developing countries participate in cross-national student assessments, global or regional, it will be possible to expand the sample size of this study.

average, countries in which less than 10 percent of students achieve advanced proficiency tend to be the countries that attain lower average years of schooling (lower right panel). Additionally, there is a weaker positive relationship between countries' average years of schooling (for the population cohort aged 20-24) and the share of the population who say they are satisfied with their education system (upper right panel). Interestingly, in some countries that have attained the highest average years of schooling (e.g., Korea), the percent of the population is less satisfied with their education system than in countries that have attained much less.

Figure 3. Exploring the education outcomes

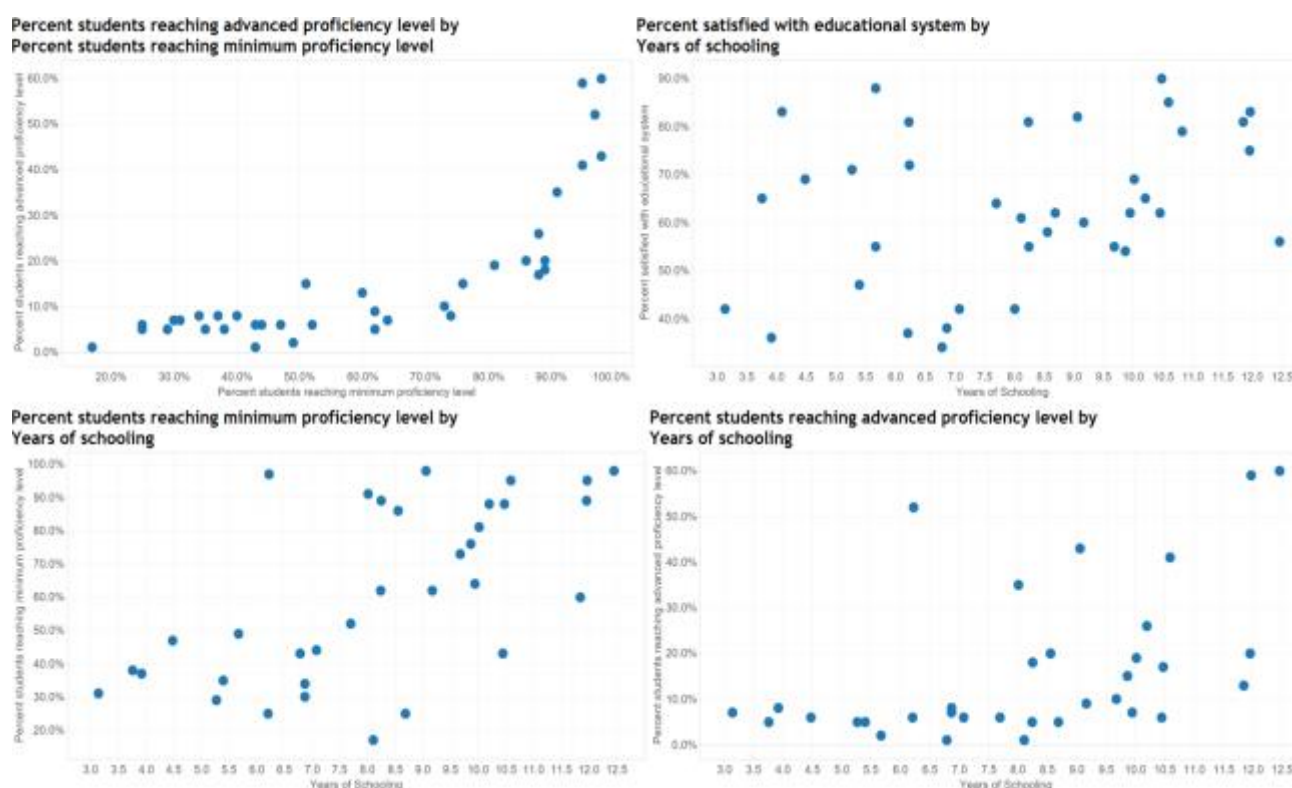


Figure 4 plots the variables that are likely to be associated with education outcomes—a country's income as measured by its GDP per capita, average schooling of the adult population (aged 50-54), and education expenditures as measured by the per-student spending at the primary level. These variables are plotted against the SABER index on the quality of the education system. We note that both per-capita GDP and current per-student expenditures in primary education are positively associated with the SABER index, whereas there is no clear association with the average schooling of the adult population aged 50-54. Countries vary greatly with respect to all three variables. We note, in particular, that many countries spend less than \$1,000 per primary-school student while other countries spend at least five times more; countries with good education policies, as defined as the countries possessing a SABER index value greater than 2.5, tend to be wealthier and tend to spend more per primary student. However, the positive association between expenditures and education quality emerges mainly in countries where education

spending per student exceeds \$1000, implying that a good education system goes hand-in-hand with some minimum level of spending.

Figure 4. Exploring the relationships between the SABER index and control variables

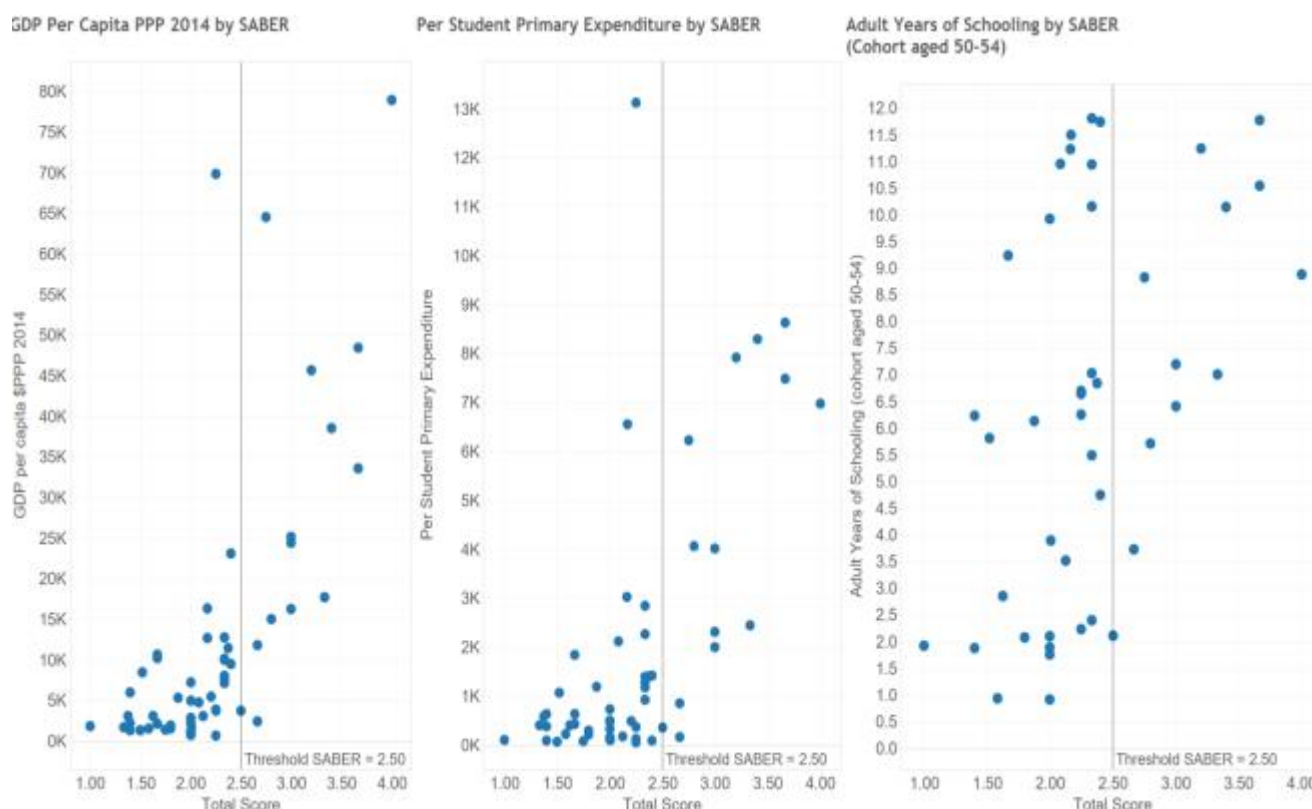
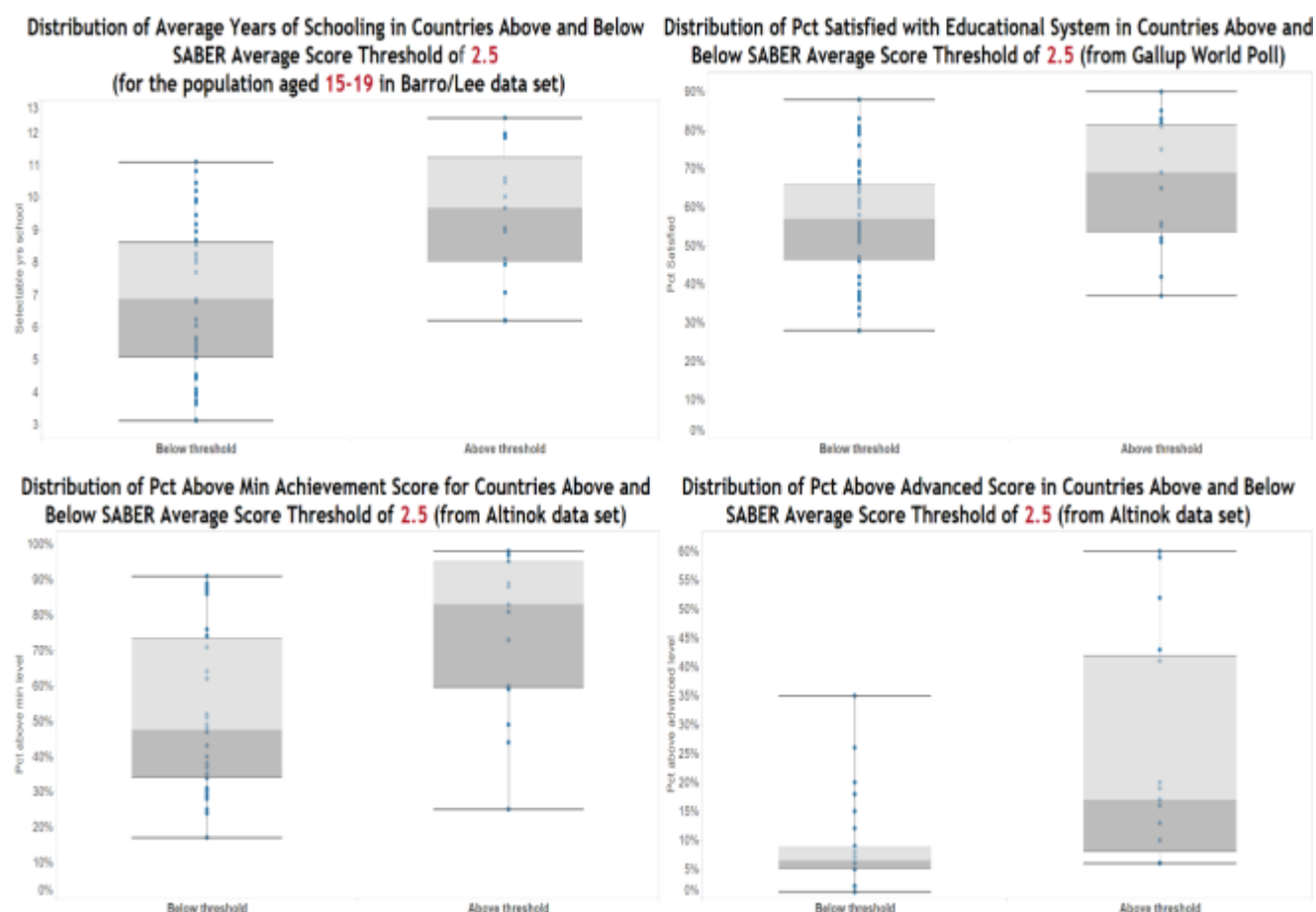


Figure 5 presents the distributions of the four education outcomes by the quality of the education system. The quality measure here is a dummy variable based on whether the SABER index value is greater or less than 2.5. The box graphs show the mean values of the education outcomes, as well as the values corresponding to the 25th and 75th percentiles and also to the 5th and 95th percentiles. In these simple comparisons without controlling for other variables, the countries that meet the SABER threshold value have far better education outcomes. The differences between the countries above and those below the threshold value of 2.5 for the SABER score are notable—over two more years in terms of the average years of schooling, over 30 percentage points in the share of students reaching the minimum proficiency level, over 10 percentage points in the share of students reaching the advanced proficiency level, and 10 percentage points in the satisfaction rate with the national education system.

Figure 5. Distributions of four education outcomes, by quality of the education system, without controls

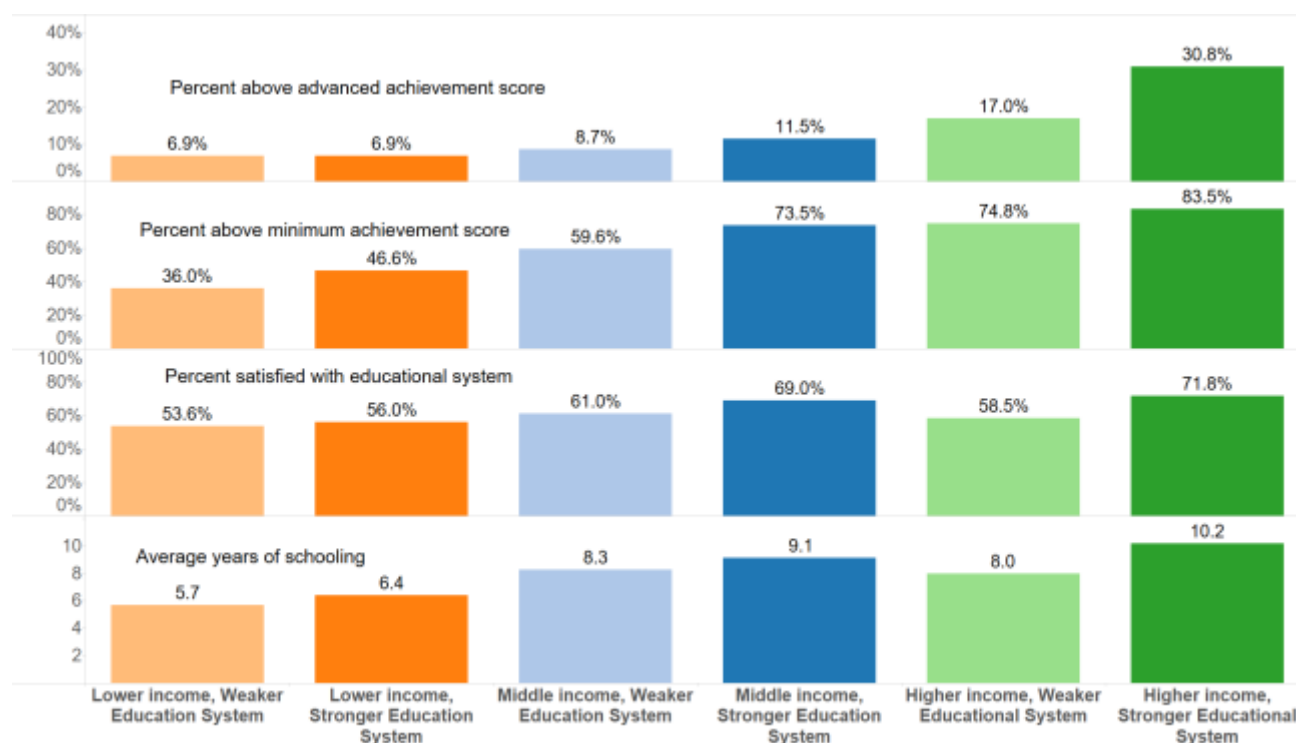


Continuing with simple comparisons, a pattern in the education outcomes emerges when countries are grouped by the quality index and their income level (Figure 6). For middle- and higher-income countries, we define the education system as weaker or stronger if the average SABER rating is below or above the threshold value of 2.5 (corresponding to a level between “Emerging” and “Established”).¹⁵ For lower-income countries, nevertheless, because only a few of these countries garner a SABER score of 2.5 and above, we use the lower threshold value of 2.0. The results suggest that within income groups, as in Figure 6, those countries that have a better education system have better education outcomes. They also suggest that middle-income countries may do as well as higher-income countries with respect to education outcomes if they

¹⁵ The country income groups are defined as follows: lower-income countries are the ones with GDP per capita in constant 2014 \$PPP less than \$ PPP 5,000, middle-income countries are the ones with GDP per capita between \$ PPP 5,000 and \$ PPP 16,000, And higher-income countries are the ones with GDP per capita above \$ PPP 16,000.

have a good education system. In the next section, we examine whether these descriptive results hold up through regression analyses.

Figure 6. Education outcomes by level of GDP and quality of the education system



V. REGRESSION RESULTS

To understand better the relationships suggested by the descriptive analyses above, it is important to control for other factors that are also likely to affect education outcomes so as not to confound the relationship attributed to the quality of the education system with that due to other factors. The controls included in the multivariate analysis are the country's income level, the education level of the adult population aged 50-54, and the per-student public spending for primary education. The first two controls are contextual variables. Previous research has widely shown them to be related to education outcomes, both possibly capturing the demand for education in the economy and the value that parents place on education. Controlling for the country's income, the education spending per student indicates not only the aggregate level of school inputs, but also the country's willingness to allocate sufficient resources for education. In past research that have estimated the production function for education outcomes, school inputs have been measured by the pupil-teacher ratio, average teacher characteristics, and availability of textbooks and learning materials. Per-student expenditures, being an aggregate measure of these inputs, reflect the relative spending on teachers and materials, and also the cost of these inputs. In developing countries, the share of salary costs in education expenditures varies widely, reaching 94% in the case of Togo in 2014, for example. Measures of the quality and efficiency of the education system have been missing in the past estimates of the production function (Hanushek 2003; Glewwe et al. 2014).

Before turning to the regression results, it is useful to summarize briefly the findings from previous studies about the relationships we are estimating. There is no database on a cross-country measure of the quality of the education system quite like the SABER data that we are aware of, so the most relevant literature in this respect are those that have quantitative measures of institutional variables such as quality of teaching based on the TALIS database, school management and governance¹⁶, measures of corruption within the system,¹⁷ and the existence or size of the private sector. These studies also do not estimate the relationships between these measure and all the outcome variables that we examine. In terms of the control variables, most studies consistently show that the level of parents' education (either of one or both parents) has a positive effect on their children's school participation and completed years of schooling.¹⁸ There are fewer studies on the effect of parents' education on student learning than on enrollment, but these tend also to find a positive relationship.¹⁹

The studies of the impact of increased public funding for education portray a mixed picture of that impact, but when that spending goes to building and staffing schools in areas where no school previously existed, then the spending increases enrollment.²⁰ The effect on enrollment is also positive for public spending to reduce school fees, increase textbooks, or reduce pupil-teacher ratios. Fewer studies have used student learning as the outcome variable, rather than enrollment, but experimental and quasi-experimental evidence show that increases in school inputs do raise student learning.²¹ The mechanism by which school expenditures, such as on teacher salaries, affects learning is by determining teaching quality. In general, however, measuring teaching quality directly is relatively challenging,²² so studies tend to skip the mediating effect of inputs on teaching quality and instead estimate the effect of inputs on student learning. Determining the level of resources that is necessary to provide an adequate education is not an easy task.

Countries with better education systems have better education outcomes

Our regression results indicate that countries with a better education system achieve better education outcomes. The results remain the same whether the quality of the system is measured by the SABER index value or by a dummy variable which equals one if the average SABER value is

¹⁶ See Bruns, Filmer and Patrinos (2011) for a review of this literature.

¹⁷ See Ferraz, Finan and Moreira (2012) on Brazil is one example. The study uses data from the auditing of Brazil's local governments to construct measures of corruption involving educational block grants transferred from the central government to municipalities. Students residing in municipalities where corruption in education was detected score 0.35 standard deviations less on standardized tests, and have significantly higher dropout and failure rates.

¹⁸ Orazem and King (2008).

¹⁹ Hanushek and Woessmann (2006).

²⁰ Hanushek (2003); Glewwe et al. (2014).

²¹ See, for example, Clark (2009); Duflo, Hanna, and Ryan (2010); and Muralidharan and Sundararaman (2011).

²² Bruns and Luque (2015).

above a threshold of 2.5 (Table 1).^{23,24} This positive association holds even as controls are added, as shown by the full estimates in Appendix Table B1. When controlling for a country's GDP per capita and the average education level of adults aged 50-54, the results suggest that, of countries with similar levels of GDP per capita or adult education levels, those with a better education system are likely to have more years of schooling for its youth, a higher proportion of students meeting the minimum and advanced proficiency levels in multi-country learning assessments, and more of the general population being satisfied with their education system. With controls, the size of the education system coefficients is smaller, but in general, not dramatically so. Even with controls, in countries that pass the threshold for education system quality (SABER score>2.5), the share of students reaching the minimum proficiency level is 25 percent higher, the share reaching the advanced proficiency level is 14 percent higher, the average schooling is 2.6 years higher, and the satisfaction rate with the education system is greater by 15 percent.

Table 1. Regression Results for Education Outcomes, with Education System Quality and Education Expenditures

	Percent reaching minimum proficiency		Percent reaching advanced proficiency		Years of schooling		Percent satisfied with education system	
System quality								
Average SABER score	0.162**		0.118***		1.935***		0.104*	
	(0.0603)		(0.0382)		(0.530)		(0.0543)	
SABER dummy (=1, if average score >2.5)		0.253***		0.145***		2.623***		0.148**
		(0.0655)		(0.0446)		(0.663)		(0.0658)
Education expenditures	0.00546	0.00532	0.000876	0.000854	-0.327**	-0.336**	-0.00763	-0.00762
	(0.0163)	(0.015)	(0.0103)	(0.0102)	(0.145)	(0.142)	(0.0152)	(0.0150)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Notes: Appendix Table B1 contains the full regression results. The control variables are GDP per capita, average years of schooling of the adult cohort, aged 50-54, and a constant term.

In contrast, the average expenditure level for basic education, even when included by itself and without the system quality variable, is not significantly associated with any of the education outcomes (see Annex Table B.1), and its inclusion does not significantly change the coefficient of system quality. In fact, for years of schooling, the per-student expenditure level has a significantly negative coefficient. Spending more for education per student, by itself, does not appear to benefit education outcomes, and whatever effect it has seems to be dominated by the quality of the education system. We explore this observation further by adding also an interaction term between system quality and expenditures. Results change only in the proficiency-related

²³ We explored alternative threshold values for SABER, including 2 and 2.25. We find that there is a higher positive association with education outcomes with the more stringent rating of 2.5.

²⁴ As mentioned earlier, SABER ratings categorize the level of the quality of education policies in a domain as latent, emerging, established and advanced. To generate the four categories, a total score is calculated based on the rating of individual policies that make up each domain. In our aggregation, the same cut-off points are applied to the total score as are used in aggregating up the individual scores within a policy domain.

outcomes: a loss in statistical significance for system quality with no change in the magnitude of its coefficient, and no positive change in the statistical significance of per-student expenditure.

The regression results suggest some optimism about the ability of low- and middle-income countries to overtake richer countries in terms of the quality of education systems, which is aligned with the findings from the prior descriptive analysis that does not control for other country characteristics. Optimistically, these countries can have as good education outcomes as higher-income countries if they have a good education system, although few low-income countries in the SABER database currently score above 2. Conversely, higher-income countries may do no better than middle-income countries if they do not have a good education system. Per-capita GDP does not figure significantly in any of the regressions, with the exception of one specification for the proportion of students reaching the advanced proficiency level which suggests that higher-income countries may have an advantage.

Consistent with the existing literature, the regression results show the intergenerational effect of education (Annex Table B.1). In all the regressions, the average years of schooling of adults aged 50-54 have a significant positive association with current education outcomes. In countries where the parent generation completed an average of one more year of schooling, the share of students reaching the minimum proficiency is higher by 2.8-5 percent and students reaching advanced competency by 1.2-3.4 percent, depending on the specification of the system quality variable. The average completed years of schooling is up by almost half a year across the specifications, while the share of the population who are satisfied with their education system is higher only by .4-1.5 percent.

Instructional time in the classroom

One factor in the production of education outcomes that we intended to include in our descriptive and regression analyses is the instructional time in the classroom. Countries that spend more on education and have better policies are not likely to achieve better education outcomes unless the system also allocates sufficient time for instruction (Abadzi 2009; Bruns and Luque, 2014). There are not many cross-country quantitative analyses of the relationship between instructional time and enrollment or learning. We suspect that this is because it is not easy to obtain an accurate measure of actual instructional time. We compiled the data on *intended* instructional time from different sources: We calculated the official instructional time based on the data in the country reports from UNESCO-IBE, using the information on the intended number of minutes per subject (usually 40-45 minutes) and the intended number of subjects per school week. The responses to TIMSS's questionnaire from principals and teachers in participating countries in 2011 also contain instructional time data. According to these data, fourth-grade students average about 900 hours per year of instruction, while those in the eighth grade average about 1,000 hours. According to the data from PISA in 2012, the intended instructional hours across countries range from 1,200 hours in Colombia to less than one-half of that in Uruguay, while the average for OECD countries is 942 hours in the typical curriculum for 15-year-olds. There is a slight tendency for richer countries to have a higher intended number of

instructional hours than lower-income countries, but the top-performing countries in the PISA in 2012, except Korea, mandate fewer instructional hours per year than the average OECD country.

But *actual* instructional hours tend to deviate from the official or intended hours in many developing countries, and do so for a variety of reasons. A school may face infrastructure constraints that force the school to shorten teaching hours in order to make space for their enrollees (e.g., shortage of classrooms, limited availability of power or water needed to operate the school), a shortage of teachers, frequent teacher and student absences, poor implementation and monitoring of the school calendar, and possibly security conditions in the school's location. The data collected during spot visits to schools in several countries indicate that the actual classroom time fall considerably short of the intended instructional time because of high teacher absenteeism (Chaudhury et al., 2006; Bold et al., 2016; King et al., 2016). These high absenteeism rates are due to either excused or non-excused reasons. In addition, it is not just how much time is spent by the teacher and students in the classroom but also whether that time is used effectively for student learning. In any case, if the quality of instruction is ineffective, increasing the amount of instruction time will not positively affect learning (Mullis, 2012).

The data in Table 2 as compiled from several papers indicate that the discrepancy between the intended and actual instructional hours vary greatly across countries. Taking into account both the average teacher absence rate and the time spent teaching by those teachers who are present in the classroom, the average effective instructional time as a percentage of the expected teaching time in 13 countries for which we have direct or observational data ranges from 77 percent in Peru to just 12 percent in Mozambique. In a five-day school week, these numbers translate into just 3.8 school days per week in Peru to just 0.6 schools days per week in Mozambique. Given the large variation in the discrepancy between intended and actual hours, we did not include instructional time in our regression analyses.

Table 2. Teacher absenteeism and actual instructional time

	Teacher absence from school		Teacher absence from classroom	Time teaching as % of scheduled time	Effective instructional time as % full schedule
Bangladesh	16			42.9	36.0
Ecuador	14				
India	25				
Indonesia	19				
Peru	11			87	77.4
Pakistan	19				
Uganda	27	30	57	49.4	14.9
Kenya		16	47	40.6	18.1
Nigeria		16	25	45.3	28.5
Togo		23	38	66.9	31.9
Mozambique		46	61	59.5	12.5
Tanzania		14	47	39.3	17.9
Senegal		18	31	50.0	28.3
Brazil (Pernambuco)	7			63	58.6

Ghana	23	38.6	29.7
Morocco	7.5	71.1	65.8
Tunisia	6.5	77.9	72.8

Sources: Abadzi, 2006; Chaudhury et al., (2006, p. 92); Bold et al. (2016); King et al. (2016); UNESCO-IBE (country report for Bangladesh)

VI. A ROBUSTNESS CHECK

In this section, we obtain a different set of regression results utilizing an alternative measure of the quality of the education system, the CPIA index produced by the World Bank. The CPIA index on the education sector has six core components that are graded on a 1-6 scale. Using this index has the advantage that the data related to each dimension of an education system are available for a large number of countries. However, at the low end of quality, the CPIA scores tend to be higher than the SABER scores, with hardly any country being given the lowest rating; at the same time, few countries were given the top rating of 6. Based on this compressed distribution, we rescaled the CPIA score to a 4-point scale (Table 3).

Table 3. Comparing the distributions of the SABER and CPIA scores

	Percent reaching minimum proficiency	Percent reaching advanced proficiency	Average years of schooling	Percent satisfied with education system
A. SABER Score				
Minimum value	1.4	1.4	1.4	1.4
25 th percentile	2.08	2.08	2.0	2.0
50 th percentile	2.33	2.33	2.33	2.33
75 th percentile	2.80	2.80	2.75	2.75
Maximum value	4	4	4	4
Mean	2.46	2.46	2.40	2.43
Standard deviation	.65	.65	.63	.62
N	34	34	39	37
B. Rescaled CPIA score, full sample (SABER countries only)				
Minimum value	2 (2)	2 (2)	2 (2)	2 (2)
25 th percentile	2.44 (2.44)	2.44 (2.44)	2.44 (2.44)	2.44 (2.44)
50 th percentile	2.67 (2.72)	2.67 (2.72)	2.67 (2.78)	2.67 (2.78)
75 th percentile	3 (3)	3 (3)	3 (3)	3 (3.11)
Maximum value	3.78 (3.78)	3.78 (3.78)	3.78 (3.78)	3.78 (3.78)
Mean	2.76 (2.77)	2.76 (2.76)	2.75 (2.78)	2.76 (2.81)
Standard deviation	.40 (.44)	.40 (.44)	.39 (.43)	.39 (.42)
N	53 (28)	53 (28)	65 (33)	65 (31)

As a check on the comparability of the two ratings, we undertake regression analyses on the full CPIA sample as well as on a sample consisting of only those countries with a SABER score. Table 4 shows the results when the control variables are included as in the prior regression analysis, and the data here are expected to be comparable with the results in Table 1.²⁵ Similar to the

²⁵ The full results corresponding to Table 4 are in Annex Table B3.

findings from the regressions using the SABER data, the quality of the education system is significantly associated with learning outcomes, as measured by the percent shares of students reaching the minimum proficiency level and those reaching the advanced proficiency. However, while the coefficients across the two CPIA samples are qualitatively similar, they are not statistically significant in the smaller sample. In contrast to the results using the SABER score, the coefficient of the CPIA score is not statistically significant for either the years of average schooling or the percent of the population satisfied with the education system. Part of the reason for the generally weaker associations of the policy variables with the educational outcomes with the CPIA measure as opposed to the SABER measure may be due to the compressed distribution of the CPIA score (Table 3). As with the SABER specifications, we undertake robustness analyses using the CPIA data; the results vary when alternative threshold values of 2.5 and 3 are set.²⁶

Table 4 Regression analyses using the CPIA score to measure system quality

VARIABLES	Percent reaching minimum competency	Percent reaching advanced competency	Average years of schooling completed	Percent satisfied with education system
A. Rescaled CPIA, full CPIA sample				
Per-student expenditure in primary/1000	0.0142 (0.0179)	0.0175** (0.00821)	-0.167 (0.174)	-0.0151 (0.0178)
Rescaled CPIA average score	0.144*** (0.0535)	0.0898*** (0.0245)	-0.156 (0.463)	-0.0248 (0.0488)
B. CPIA dummy variable, full CPIA sample				
Per-student expenditure in primary/1000	0.0165 (0.0179)	0.0197** (0.00863)	-0.168 (0.173)	-0.0152 (0.0178)
Dummy=1 if Average CPIA score > 2.75	0.0944** (0.0376)	0.0447** (0.0181)	-0.133 (0.335)	-0.0225 (0.0350)
C. Rescaled CPIA, SABER countries only				
Per-student expenditure in primary/1000	0.0286 (0.0215)	0.0146 (0.00956)	-0.185 (0.216)	-0.00942 (0.0232)
Rescaled CPIA average score	0.110 (0.0738)	0.0898** (0.0328)	-0.181 (0.673)	-0.0539 (0.0738)
D. CPIA dummy variable, SABER countries only				
Per-student expenditure in primary/1000	0.0288 (0.0216)	0.0161 (0.0106)	-0.187 (0.217)	-0.00953 (0.0233)
Dummy=1 if Average CPIA score > 2.75	0.0777 (0.0555)	0.0310 (0.0272)	-0.0176 (0.507)	-0.0194 (0.0564)

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

The regression analyses here and in the previous section used average SABER and CPIA scores to measure the quality of the whole education system. Although the SABER program has scored individual components of the system (e.g., teacher policies, financing), as of this writing, SABER

²⁶ See Annex Table B.4 for the full results using alternative threshold values for estimating the CPIA dummy variables.

data on all components are not yet complete. The program was meant primarily to provide systems data and analysis on demand, that is, as needed by countries that are designing and implementing reforms rather than to build a cross-country database. Because of the incomplete availability of data on the dimensions of the education system, we estimated one aggregate index of system quality from whatever SABER ratings are available. Thus far, the paper does not point to which parts of the education system are more or less critical to the overall health of the system and to education outcomes. Believing that such an analysis could be useful, we re-estimated our model using the individual CPIA scores for six dimensions of the education system. The results do not point strongly to any one dimension being dominant, except perhaps having a serious sector strategy (SST) which does appear to be associated with a higher proportion of students reaching the minimum proficiency level (Annex Table B.5). The coefficients for teacher policies (TCH) and school-based management (SBM) are significant only in the regressions with the system dummy variable and only when a higher threshold level is used to define those dummy variables. For a higher percentage of students reaching advanced proficiency level, only meeting a relatively high threshold for school-based management is significant.

VII. CONCLUSIONS

The magnitude of the current call for more education is unprecedented, as exemplified by the UN Sustainable Development Goals for 2030 and the Global Education Commission's report.²⁷ Most striking about this call is the clear shift in focus to quality education and improvements in learning. In response to the Education for All movement in the 1990s, governments dramatically increased the capacity of their school systems to enroll students, mostly by building many more classrooms and recruiting many more teachers than ever before. But learning outcomes have not kept up with this progress in enrollment rates and average years of schooling. Research about how to improve learning point to a large number of factors besides investments in more schools, more classrooms, and even more teachers; these include the quality of teaching, time spent on task in schools, appropriate curriculum, and language of instruction (Bruns et al., 2011). Top school systems in the world pay a great deal of attention to how they select their staff; they work hard to improve the performance of schools, provide an environment in which teachers work together to frame good practice, and establish smart pathways for teachers to grow in their careers. They are able to achieve these improvements because their education systems are organized, adequately resourced, and led by managers who are accountable for their performance. Strong education systems have standards, academic curricula, financing, information and other structured processes that are coherent and aligned towards achieving education goals. In contrast, weak education systems struggle to achieve that alignment and coherence: their standards, goals and rules are not clearly defined; the inputs are inappropriate or inadequate; resources are used inefficiently and accountability mechanisms are weak. Moreover, these systems are not sufficiently dynamic to adjust to shifts in the socioeconomic and political contexts, as well as to changes in the financial and management capacities of the country.

²⁷ The Global Education Commission's report is entitled *The Learning Generation*; <http://report.educationcommission.org/>. See also .

Our analyses of the relationship between measures of the quality of an education system and education outcomes suggest that system quality matters for student test performance, average years of schooling in the country, and people's regard for their education system. These results come across consistently, controlling for country-level factors that may affect these education outcomes. The results indicate that increasing education expenditures is not likely to yield better education outcomes if the education system is weak. Previous cross-country analyses have included system-level measures such as pupil-teacher ratios, percentage of teachers trained, or per-pupil expenditures, but those measures pertain to input levels rather than the quality of the system. Since the SABER initiative which yielded the data used in our analyses is an ongoing enterprise, this unique database can continue to improve in terms of greater coverage of countries as well as the individual components of their education systems. The current SABER instruments focus on policies *de jure*; expanding the scope of these instruments to collect information also about the quality of policy implementation could be very useful for expanding this research.

References

- Abadzi, Helen. "Instructional time loss in developing countries: Concepts, measurement, and implications." *The World Bank Research Observer* (2009): lkp008.
- Altinok, Nadir, Claude Diebolt, and Jean-Luc De Meulemeester. 2013. "A New International Database on Education Quality: 1965-2010."
- Barro, R. J., & Lee, J. W. 2013. "A new data set of educational attainment in the world, 1950–2010," *Journal of Development Economics* 104 (1): 184-198.
- Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Brian Stacy, Christophe Rockmore, Jakob Svensson, and Waly Wane. 2016. "What Do Teachers Know and Do in Primary Schools in Sub-Saharan Africa?" Processed. AERC and World Bank Group.
- Bruns, Barbara, Deon Filmer, and Harry Anthony Patrinos. 2011. *Making schools work: New evidence on accountability reforms*. Washington, DC: World Bank Publications.
- Bruns, Barbara, and Javier Luque. 2014. *Great Teachers: How to Raise Student Learning in Latin America and the Caribbean*. Washington, DC: World Bank.
- Clarke, M. 2012. "What Matters Most for Student Assessment Systems: A Framework Paper." READ/SABER Working Paper Series. Washington, DC: World Bank.
- Ferraz, Claudio, Frederico Finan, and Diana B. Moreira. "Corrupting learning: evidence from missing federal education funds in Brazil." 2012. *Journal of Public Economics* 96, no. 9 (2012): 712-726.
- Glewwe, Paul W., Eric A. Hanushek, Sarah D. Humpage, and Renato Ravina. 2011. *School resources and educational outcomes in developing countries: A review of the literature from 1990 to 2010*. No. w17554. National Bureau of Economic Research.
- Hanushek, Eric A. 2003. "The failure of input-based schooling policies." *The Economic Journal* 113(485): F64-F98.
- Hanushek, Eric and Ludger Woessmann. 2015. *The Knowledge Capital of Nations: Education and the Economics of Growth*. Cambridge, MA: MIT Press.
- Mourshed, M., C. Chijioke, and M. Barber. 2010. *How the World's Most Improved School Systems Keep Getting Better*. McKinsey.
- National Center for Education Studies. 2015. *Trends in International Mathematics and Science Study (TIMSS) report*. <https://nces.ed.gov/TIMSS/>.
- Organisation for Economic Co-operation and Development (OECD). 2013. *PISA 2012 Results in Focus: What 15-Year-Olds Know and What They Can Do with What They Know*. Paris: OECD.
- Pritchett, Lant. 2013. "The Rebirth of Education Why Schooling in Developing Countries Is Flailing; How the Developed World Is Complicit; and What to Do Next." *Center for Global Development Brief*, 1–4
- Pritchett, Lant. 2015. RISE Working Paper 005: "Creating Education Systems Coherent for Learning Outcomes: Making the Transition from Schooling to Learning"

- UNESCO (2016). International Bureau of Education (IBE). *World Data on Education*.
<http://www.ibe.unesco.org/en/document/world-data-education-seventh-edition-2010-11>
- Vegas, Emiliana, and Alejandro Ganimian. 2013. *Theory and evidence on teacher policies in developed and developing Countries*. Inter-American Development Bank, Washington, DC.
- World Bank. 2015. *EdStats*; Education Statistics Dashboard. World Bank, Washington, DC.
<http://datatopics.worldbank.org/education/>.
- World Bank. 2016. *Systems Approach for Better Education Results (SABER)*.
<http://saber.worldbank.org/index.cfm>
- World Bank. 2016. *World Development Indicators*. Retrieved from
<http://databank.worldbank.org/data/reports.aspx?source=2&series=SE.SEC.PRIV.ZS&country=#>

Annex A
CPIA scoring matrix and guidelines

Dimension\Rating	1	2	3	4	5	6
Sector Strategy	There is no sector strategy in written form, and different levels of management in education hold different views on strategic priorities.	There is a written policy but it is outdated, and different levels of management still hold differing views on education priorities	There is a written strategy that needs updating and the different levels of management are aware of the need for a consensus for a new education strategy	There is a written education sector strategy that coincides with the MDG but it has not been discussed and agreed upon with key stakeholders in society.	There is an education strategy that has been consulted with the key stakeholders in society, and it is being implemented while its final configuration is still in progress.	There is an education strategy being implemented that is aligned with the country's own resources and with the policy priorities of national development strategies and in agreement with all key stakeholders in society.
EMIS	Lack of statistical data collection system or institutional framework , and little government commitment and use of data	Basic statistical data collection system is in place, data are collected sporadically and infrequently available to the public.	Basic statistics of varying quality and reliability are collected and reported yearly and made available to the public. However education statistics are rarely used in policy dialogue.	Integrated information system in place; policy and issue relevant data are collected and available to policy makers. Data are of acceptable quality but not always timely.	Education statistics are being used for planning and monitoring sector performance. MOE yearly informs public about the progress of the education sector. Data are reliable.	Education statistics are reliable and widely used in policy dialogue. Information is fed back to schools and community in order to promote accountability and to facilitate school level planning.
Assessment	No assessment of student learning outcomes at country or system level. Policies, human & fiscal resources, institutional setups are nonexistent or grossly inadequate to support assessment of learning outcomes at this level.	Some assessment of student learning outcomes at country or system level, but very ad hoc and donor driven. Policies, human & fiscal resources, institutional setups are inadequate to support assessment of learning outcomes at this level.	Fairly regular (once every 3-5 yrs) assessment of student learning outcomes at country or system level, but data not comparable over time. Policies, human & fiscal resources, institutional setups are adequate to support assessment of learning outcomes at this level.	Regular (at least once every 3-5 yrs) assessment of student learning outcomes in at least one grade/age level at country or system level, and data are comparable over time. Policies, human & fiscal resources, institutional setups are adequate to support assessment of learning outcomes and sustained monitoring and use of these data, although there may be gaps or inconsistencies.	Regular assessment of student learning outcomes in at least two grades/age levels at country or system level, and data are comparable over time. Policies, human & fiscal resources, institutional setups are appropriate and sufficient to support assessment of learning outcomes and sustained monitoring and use of these data, which are increasingly used to guide policy.	Regular assessment of student learning outcomes in at least two grades/age levels and subject areas at country or system level, and data are comparable over time. Policies, human & fiscal resources, institutional setups are appropriate and sufficient to support assessment of learning outcomes and sustained monitoring and use of these data, which are effectively used to guide policy and support schools.

Dim.\Rating	1	2	3	4	5	6
Teachers	Teacher salaries are disbursed irregularly; teachers are not evaluated regularly, and are not required to participate in evaluations; teacher salaries are not differentiated, performance. .	Teacher salaries are disbursed irregularly; Teachers are required to be evaluated but rarely happens; Teacher salaries are differentiated only by one of several factors, such as subject, geographic area, education level, years of service, or demographic characteristics of the student population; salaries are not tied to teacher performance.	Teacher salaries are disbursed irregularly; Teachers are required to be evaluated but that rarely happens;; Teacher salaries are differentiated by at least two factors – subject, geographic area, education level, years of service and/or demographic characteristics of the student population— but not by teacher performance.	Teacher salaries are mostly on time; teacher evaluations are required but that rarely happens; teacher salaries are differentiated by at least three factors –subject, geographic area, education level, years of service and/or demographic characteristics of the student population— but not by teacher performance.	Teacher salaries are disbursed on time; teacher evaluations are required but evaluations are done irregularly; teacher salaries are differentiated by at least three factors – subject, geographic area, education level, years of service and/or demographic characteristics of the student population— as well as by teacher performance.	Teacher salaries are disbursed on time; teacher evaluations are required and performed as per requirements; teacher salaries are differentiated by at least three factors – subject, geographic area, education level, years of service and/or demographic characteristics of the student population—as well as by teacher performance.
Education finance	More than 80% of schools lack need rehabilitation ; more than 95% of public education expenditures go to salaries; large differences in average per pupil expenditures exist between rural and urban schools and/or among provinces/states/regions; budgeted resources for the education sector are unpredictable and unreliable from year to year, and the distribution of financing across educational levels reinforces social inequality	Between 60 and 80 percent of schools lack basic infrastructure and inputs; 85-95 percent of public education expenditures are on personnel salaries; significant differences in average per pupil expenditures exist between rural and urban schools and/or among provinces/states/regions; budgeted resources for the education sector are unpredictable and unreliable from year to year , and the distribution of financing across educational levels reinforces social inequality	Between 40 and 60 percent of schools lack basic infrastructure and inputs; 75-85 percent of public education expenditures are on personnel salaries; government has introduced programs to reduce differences in average per pupil expenditures between rural and urban schools and/or among provinces/states/regions; budgeted resources for the education sector are determined through a transparent process, but they are primarily item-based, and the distribution of financing across educational levels takes into account some social equity	Between 20 and 40 percent of schools lack basic infrastructure and inputs; 65-75 percent of public education expenditures are on personnel salaries; government has introduced programs to reduce differences in average per pupil expenditures between rural and urban schools and/or among provinces/states/regions; budgeted resources for the education sector are determined through a transparent process, but they are primarily item-based, and the distribution of financing across educational levels takes into account some social equity	Between 20 and 40 percent of schools lack basic infrastructure and inputs; less than 75 percent of public education expenditures are on personnel salaries; government has introduced programs to reduce differences in average per pupil expenditures between rural and urban schools and/or among provinces/states/regions; budgeted resources for the education sector are determined through a transparent process, and they are program-based, and the distribution of financing across educational levels is consistent with social equity	Less than 20 percent of schools lack basic infrastructure and inputs; less than 75 percent of public education expenditures are on personnel salaries; government has introduced programs to reduce differences in average per pupil expenditures between rural and urban schools and/or among provinces/states/regions; budgeted resources for the education sector are determined through a transparent process, and they are results-based, and the distribution of financing across educational levels is consistent with social equity

Dim. \Rating	1	2	3	4	5	6
School-Based Management	School funds and personnel are managed directly by the central or national ministry of education. Local school management limited to petty cash and supervision of operations. No parental participation in school management. Academic and financial accountability based on political criteria. School Councils and parents are not informed about school performance	Funding and teacher management are highly centralized, but schools are administered at a sub-national level. Parent participation in school activities generally limited to social events. Accountability is limited to personnel misconduct and is handled centrally. School Councils and parents are not informed about school performance	Education budget is fairly transparent, but based on enrollment and salaries. Teachers are managed at a sub-national level according to collective agreements with the teacher's union. Parents can voice their concerns to the school administration but decisions taken at the central level. School Councils and parents are not informed about school performance	School budget is based on enrollment, salaries, and equity. Schools are able to use discretionary funds for minor capital improvements or for local purchases. Teachers managed at a sub-national level. Parent Councils work with the school but in an advisory role. Parents can raise funds for the school. School Councils and parents are not informed about school performance	School budget is based on enrollment, salaries, and equity and is administered at a sub-national level. Teacher hiring and firing done at the sub-national level. Parent councils can assist schools in budget planning and fund raising. Schools are able to use discretionary funds for minor capital improvements or for local purchases. Teachers managed regionally. School Councils and parents are informed about school and student performance	School budget prepared locally in collaboration with the Parents Council. Central funds are transferred to the school directly and are based on enrollment (on a per capita basis) and accounts for equity (i.e., targeted or compensatory funding). Schools have complete autonomy over the budget, but are subject to government regulations about minimum academic and financial standards. Teacher hiring and firing done at the school level. Parent Councils have legal authority over the budget. Schools are able to raise funds for major capital improvements. School Councils and parents are informed about school and student performance, and about the performance of similar schools.

Annex B

Appendix Table B1. Regression analyses using SABER to measure education quality

	Average SABER score				SABER dummy	
	[1]	[2]	[3]	[4]	[5]	[6]
A. Percent Reaching Minimum Proficiency						
System quality		0.181** (0.069)	0.162** (0.0603)	0.140 (0.0836)	0.264*** (0.0746)	0.253*** (0.0655)
Per-student expenditure in primary education	0.00694 (0.00601)		0.00546 (0.0163)	-0.00946 (0.0419)		0.00532 (0.0150)
System quality x Expenditure				0.00616 (0.0159)		
Adult schooling, aged 50-54	0.0504*** (0.00577)	0.0285*** (0.00933)	0.0390*** (0.00888)	0.0392*** (0.00901)	0.0333*** (0.0086)	0.0429*** (0.00789)
GDP per capita, 2014 \$ PPP	0.000780 (0.000939)	0.00158 (0.00198)	0.000480 (0.00267)	0.000334 (0.00273)	0.000699 (0.00189)	-0.000473 (0.00247)
Constant	0.243*** (0.0397)	-0.0382 (0.146)	-0.0641 (0.127)	-0.0160 (0.179)	0.303*** (0.0563)	0.246*** (0.049)
Observations	99	44	40	40	44	40
R-squared	0.649	0.538	0.678	0.679	0.587	0.728
B. Percent Reaching Advanced Proficiency						
System quality		0.119*** (0.0376)	0.118*** (0.0382)	-0.0390 (0.0355)	0.123*** (0.0439)	0.145*** (0.0446)
Per-student expenditure in primary education	0.00676 (0.00429)		0.000876 (0.0103)	-0.105*** (0.0178)		0.000854 (0.0102)
System quality x Expenditure				0.0436*** (0.00675)		
Adult schooling, aged 50-54	0.0246*** (0.00412)	0.0131** (0.00509)	0.0136** (0.00562)	0.0153*** (0.00383)	0.0163*** (0.00506)	0.0168*** (0.00537)
GDP per capita, 2014 \$ PPP	0.00125* (0.000670)	0.00146 (0.00108)	0.00164 (0.00169)	0.000601 (0.00116)	0.00161 (0.00111)	0.00147 (0.00168)
Constant	-0.0552* (0.0283)	-0.247*** (0.0795)	-0.246*** (0.0804)	0.0944 (0.0759)	-0.0213 (0.0331)	-0.0209 (0.0334)
Observations	99	44	40	40	44	40
R-squared	0.578	0.592	0.638	0.838	0.574	0.646
C. Average years of schooling						
System quality		1.574*** (0.489)	1.935*** (0.530)	2.207*** (0.713)	1.927*** (0.605)	2.623*** (0.663)
Per-student expenditure in primary education	-0.149** (0.0579)		-0.327** (0.145)	-0.115 (0.394)		-0.336** (0.142)
System quality x Expenditure				-0.0844 (0.146)		
Adult schooling, aged 50-54	0.429*** (0.0527)	0.417*** (0.0715)	0.448*** (0.0809)	0.445*** (0.0818)	0.450*** (0.07)	0.500*** (0.0773)
GDP per capita, 2014 \$ PPP	0.0237*** (0.00904)	-0.00871 (0.0147)	0.0190 (0.0209)	0.0198 (0.0211)	-0.00761 (0.0145)	0.0164 (0.0206)
Constant	5.083*** (0.341)	1.710* (1.004)	0.999 (1.104)	0.423 (1.494)	4.610*** (0.445)	4.639*** (0.468)
Observations	117	59	49	49	59	49

R-squared	0.482	0.584	0.629	0.632	0.583	0.644
D. Percent satisfied with education system						
System quality		0.106** (0.0471)	0.104* (0.0543)	0.0742 (0.0712)	0.125** (0.0571)	0.148** (0.0658)
Per-student expenditure in primary education	-0.000233 (0.00576)		-0.00763 (0.0152)	-0.0300 (0.0375)		-0.00762 (0.015)
System quality x Expenditure				0.00924 (0.0142)		
Adult schooling, aged 50-54	0.00434 (0.00460)	0.0122* (0.00655)	0.00703 (0.00802)	0.00754 (0.00811)	0.0148** (0.00638)	0.0103 (0.00758)
GDP per capita, 2014 \$ PPP	0.00163 (0.00107)	0.000191 (0.00147)	0.00164 (0.00249)	0.00137 (0.00255)	0.000264 (0.00147)	0.00121 (0.00249)
Constant	0.575*** (0.0296)	0.270*** (0.0953)	0.313*** (0.111)	0.376** (0.148)	0.465*** (0.0409)	0.508*** (0.0454)
Observations	109	54	45	45	54	45
R-squared	0.112	0.295	0.253	0.261	0.291	0.276

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table B.2 Regression Coefficients of System Variable using Alternative Thresholds for SABER Index

	(1)	(2)	(3)
A. Using a SABER threshold value of 2.0			
% students reaching minimum proficiency	0.227*** (0.0764)	0.122 (0.0768)	0.0626 (0.0829)
% students reaching advanced proficiency	0.0734 (0.0467)	0.00590 (0.0441)	-0.0251 (0.0541)
Average years of schooling completed	1.899*** (0.672)	1.326* (0.710)	1.067 (0.692)
% respondents satisfied with their education system	0.0365 (0.0446)	-0.0194 (0.0450)	0.0198 (0.0645)
B. Using a SABER threshold value of 2.25			
% students reaching minimum proficiency	0.203*** (0.0626)	0.120* (0.0621)	0.0600 (0.0673)
% students reaching advanced proficiency	0.0856** (0.0379)	0.0350 (0.0357)	0.0182 (0.0441)
Average years of schooling completed	2.693*** (0.499)	2.340*** (0.540)	1.715*** (0.516)
% respondents satisfied with their education system	0.0511 (0.0394)	0.00860 (0.0406)	-.0300 (0.0555)

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Specifications: (1) only SABER dummy variable and constant term; (2) SABER dummy variable and control variables; (3) SABER dummy variable, per-student expenditure, and control variables.

Table B.3 Regression analyses using the CPIA score to measure system quality

VARIABLES	% reaching minimum competency	% reaching advanced competency	Average years of schooling completed	% satisfied with education system
A. Rescaled CPIA, full CPIA sample				
Per-student expenditure in primary/1000	0.0142 (0.0179)	0.0175** (0.00821)	-0.167 (0.174)	-0.0151 (0.0178)
Rescaled CPIA average score	0.144*** (0.0535)	0.0898*** (0.0245)	-0.156 (0.463)	-0.0248 (0.0488)
Adult schooling (cohort aged 50-54)	0.0332*** (0.00795)	0.00845** (0.00365)	0.328*** (0.0694)	0.000800 (0.00775)
GDP per capita (2014 \$ PPP/1000)	0.00573 (0.00406)	0.00107 (0.00186)	0.148*** (0.0393)	0.00604 (0.00408)
Constant	-0.124 (0.145)	-0.236*** (0.0667)	5.021*** (1.230)	0.643*** (0.130)
Observations	61	61	77	72
R-squared	0.622	0.580	0.602	0.044
B. CPIA dummy variable, full CPIA sample				
Per-student expenditure in primary/1000	0.0165 (0.0179)	0.0197** (0.00863)	-0.168 (0.173)	-0.0152 (0.0178)
Dummy=1 if Average CPIA score > 2.75	0.0944** (0.0376)	0.0447** (0.0181)	-0.133 (0.335)	-0.0225 (0.0350)
Adult schooling (cohort aged 50-54)	0.0314*** (0.00796)	0.00723* (0.00383)	0.330*** (0.0691)	0.00104 (0.00773)
GDP per capita (2014 \$ PPP/1000)	0.00792* (0.00398)	0.00249 (0.00192)	0.146*** (0.0383)	0.00564 (0.00398)
Constant	0.211*** (0.0442)	-0.0206 (0.0213)	4.670*** (0.365)	0.588*** (0.0390)
Observations	61	61	77	72
R-squared	0.616	0.531	0.603	0.047
C. Rescaled CPIA, SABER countries only				
Per-student expenditure in primary/1000	0.0286 (0.0215)	0.0146 (0.00956)	-0.185 (0.216)	-0.00942 (0.0232)
Rescaled CPIA average score	0.110 (0.0738)	0.0898** (0.0328)	-0.181 (0.673)	-0.0539 (0.0738)
Adult schooling (cohort aged 50-54)	0.0303*** (0.00992)	0.00853* (0.00440)	0.350*** (0.0922)	0.000835 (0.0103)
GDP per capita (2014 \$ PPP/1000)	0.00852 (0.00579)	0.00483* (0.00257)	0.202*** (0.0587)	0.00793 (0.00630)
Constant	-0.0643 (0.208)	-0.250** (0.0924)	4.649** (1.880)	0.688*** (0.208)
Observations	31	31	38	35
R-squared	0.729	0.753	0.721	0.093
D. CPIA dummy variable, SABER countries only				
Per-student expenditure in primary/1000	0.0288 (0.0216)	0.0161 (0.0106)	-0.187 (0.217)	-0.00953 (0.0233)
Dummy=1 if Average CPIA score > 2.75	0.0777 (0.0555)	0.0310 (0.0272)	-0.0176 (0.507)	-0.0194 (0.0564)

Adult schooling (cohort aged 50-54)	0.0304*** (0.0101)	0.00640 (0.00494)	0.357*** (0.0924)	0.00186 (0.0104)
GDP per capita (2014 \$ PPP/1000)	0.0106* (0.00542)	0.00709** (0.00265)	0.196*** (0.0552)	0.00655 (0.00597)
Constant	0.177** (0.0663)	-0.0287 (0.0325)	4.173*** (0.577)	0.555*** (0.0655)
Observations	31	31	38	35
R-squared	0.726	0.697	0.721	0.081

Standard errors in parentheses, *** p<0.01, ** p<0.05, * p<0.1

Table B.4 Regression Coefficients of System Variable using Alternative CPIA specifications

	(1)	(2)	(3)
A. Full CPIA sample			
% students reaching minimum proficiency	0.108*** (0.0383)	0.131*** (0.0430)	0.161*** (0.0480)
% students reaching advanced proficiency	0.0158 (0.0226)	0.0536** (0.0246)	0.0860*** (0.0265)
Average years of schooling completed	0.529 (0.369)	-0.0875 (0.431)	-0.348 (0.490)
% respondents satisfied with their education system	0.0353 (0.0433)	-0.0428 (0.0484)	-0.0519 (0.0557)
B. SABER sample only			
% students reaching minimum proficiency	0.0926* (0.0536)	0.126** (0.0542)	0.127** (0.0599)
% students reaching advanced proficiency	-0.00167 (0.0351)	0.0302 (0.0366)	0.0884** (0.0358)
Average years of schooling completed	0.586 (0.570)	0.391 (0.589)	-0.349 (0.655)
% respondents satisfied with their education system	0.0379 (0.0734)	-0.000158 (0.0113)	-0.0558 (0.0789)

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Specifications: (1) using a CPIA threshold value of 2.5 to define the CPIA dummy variable; (2) using a CPIA threshold of 3.0; (3) using a rescaled CPIA index score. All regressions include the same control variables.

The full CPIA sample consists of all countries with CPIA data and education outcomes data; the SABER sample is the set of countries that have both SABER and CPIA data as well as education outcomes data. The rescaled CPIA index is a simple transformation of the 6-point scale to a 4-point scale because of extremely small number of countries at the bottom and top ratings.

Table B.5 Regression coefficients of system variable using CPIA scores for system dimensions

VARIABLES	% reaching minimum competency	% reaching advanced competency	Average years of schooling completed	% satisfied with education system
A. CPIA index for system dimensions				
Education sector strategy (SST)	0.0488** (0.0238)	0.0164 (0.0113)	0.0293 (0.226)	-0.00266 (0.0247)
Management & Information system (EMS)	0.0165 (0.0270)	0.0178 (0.0128)	-0.0727 (0.229)	-0.0214 (0.0251)
Student assessment (ASS)	-0.0195 (0.0227)	0.00389 (0.0108)	-0.273 (0.203)	0.0210 (0.0208)
Teachers (TCH)	0.0261 (0.0349)	-0.00269 (0.0166)	0.0746 (0.300)	-0.00651 (0.0317)
Financing (FCN)	-0.0118 (0.0292)	0.00454 (0.0139)	0.143 (0.246)	-0.0309 (0.0251)
School management (SBM)	0.0365 (0.0230)	0.00914 (0.0109)	0.0490 (0.206)	0.00899 (0.0218)
B. CPIA dummy variables for system dimensions				
SST >= 4	0.205*** (0.0631)	0.0244 (0.0325)	0.0326 (0.644)	0.0584 (0.0718)
EMS >=4	-0.0500 (0.0503)	0.0316 (0.0259)	0.0172 (0.452)	-0.0455 (0.0492)
ASS >=4	-0.0645 (0.0416)	-0.0127 (0.0214)	-0.0660 (0.396)	0.0678 (0.0416)
TCH >= 4	0.0982** (0.0460)	0.0255 (0.0237)	-0.173 (0.438)	-0.0266 (0.0456)
FNC >=4	-0.0241 (0.0437)	-0.0269 (0.0225)	0.525 (0.422)	-0.0232 (0.0438)
SBM >=5	0.0977** (0.0442)	0.0562** (0.0228)	-0.0352 (0.434)	0.00372 (0.0447)
Observations	60	60	75	70

Note: The variables included in these regressions are per-student education expenditures, GDP per capita, and the average schooling of the adult population aged 50-54; the full CPIA sample is used. Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

the Education Commission

The International
Commission
on Financing Global
Education Opportunity

educationcommission.org

