THE SCIENCE OF IMPROVING LIVES

EDUCATION POLICY AND DATA CENTER

Making sense of data to improve education

Estimating the Output of Education in Developing Countries

Proposed Methodology for the 2011 International Comparison Program

EPDC Working Paper No. 2011-1

January 2011

Estimating the Output of Education in Developing Countries

Proposed Methodology for the 2011 International Comparison Program*

The **Education Policy and Data Center (EPDC)** is a partnership of the United States Agency for International Development (USAID) and FHI 360. The mission of EPDC is to improve information and policies for education through better access and use of data and policy-oriented evaluation and research.

ABSTRACT

At the request of the World Bank's International Comparison Program Global Office, EPDC developed a methodology for estimating the output of education services in low and middle income countries, by focusing measurement alongside two major elements: volume of services and quality of outcomes. The volume of services is measured by the number of pupils in the formal education system, adjusted for biases resulting from inefficiencies. The quality of outcomes is measured by learning scores, either observed or imputed using a set of predictors. A variety of safeguards and adjustments are presented to minimize the effects of uncertainty and measurement error. As a result, both volume and quality measures are proposed for subsequent transformation into purchasing power parities for all countries participating in the 2011 ICP cycle.

* Working Papers disseminated by the EPDC reflect ongoing research and have received limited review. Views or opinions expressed herein do not necessarily reflect the policy or views of FHI 360 or of any of the EPDC sponsors.

ACKNOWLEDGEMENTS

This paper was written by Babette Wils and Carina Omoeva, with invaluable guidance from David Sprague, and input from Benjamin Sylla, Ania Chaluda, and HyeJin Kim.

EPDC is indebted to members of the International Comparison Program Global Office (Michel Mouyelokatoula, Nada Hamadeh and Yuri Dikhanov), members of the ICP Technical Advisory Group (Alan Heston, Angus Deaton, Kim Zieschang, and others), and experts from the World Bank (Harry Patrinos, Emilio Porta, and Olivier Dupriez) who provided input, ideas, and corrections that improved the proposed methodology. Where possible, the individuals who provided particular contributions are acknowledged in the text of this paper – recognizing that most of the contributions emerged from group discussions.

Contents

INTRODUCTION

This paper outlines the methodology proposed by the Education Policy and Data Center (EPDC) for the measurement of volume and quality of education services in developing countries. It is part of a project aimed at estimating the purchasing power parities (PPP) for the education sector, carried out at the request of the Global Office of the International Comparison Program (ICP) at the World Bank, and under the general guidance of the ICP Technical Advisory Group (TAG).

While PPP's have been used for decades to account for differences in the real and nominal costs of goods and services, capturing the output of such public sectors as education has been a continuous challenge for the ICP. In 2005, the ICP used an input-based approach to PPP conversions for education, based on teacher salaries and other costs, while recognizing that that an output-based approach would, if feasible, be a more direct measure of the value of education services. The ICP Global Office envisioned that such a methodology could be modeled after the method devised by Eurostat, the EU statistics service, which recently developed output-based education PPPs, focusing largely on enrollments and student learning scores on the OECD major achievement study, PISA.

While the Eurostat approach provides the basis for the methodology proposed by EPDC in this paper, we also offer substantial modifications designed to resolve issues surrounding the quality and availability of education data in developing countries. Education systems in non-OECD countries have high variation in access to schooling, attendance during the school year, retention to higher grades, children's backgrounds, and levels of learning -- all of which create biases for measurement of output, even assuming no missing data and consistency across sources of information. In addition, given substantial data gaps and questions about reliability of reported data, additional strategies, such as imputation of missing values, or the validation across multiple sources, are required prior to use of data in estimates of purchasing power parity. These strategies, as well as caveats associated with data use, are discussed at length in this paper.

The conceptual model proposed by EPDC for the measurement of education output consists of two basic elements: 1) the *volume*, or quantity of educational services acquired; and 2) the *quality* of the output acquired as a result of these services (adjusted for the effects of non-education factors and the duration of schooling). At the highest level, the conceptual equation for the calculation of the PPP adjustment factor is:

$$Expenditures = Volume \times Q(adj) \times PPP,$$

where *Expenditures* are collected through existing ICP procedures; *Volume* is a measure of the number of pupils receiving education services and the hours received per year; *Q(adj)* is a measure of the quality adjustments derived from learning scores and corrections for household background and system inefficiencies; and *PPP* are the purchasing power parity adjustments. Because the final PPP indices are derived indirectly through the conceptual equation above, EPDC focused the methodological effort on developing adequate and consistent measurements of volume and quality in education.

The methodology outlined in this paper has not been endorsed as the final method for ICP 2011; rather, it provides a foundation for further work towards the development of PPP estimates for education in developing countries. The methodology described here was tested with data available in late 2010. As more data becomes available in 2011, we expect that the precision of estimates will improve. In addition, a number of issues related to the adaptability of quantitative strategies for use by ICP regional offices, as well as the ultimate functional form for the transformation of volume and quality measures into PPP's remain open for further discussions with the ICP Global Office and the TAG. This paper covers these issues.

The remainder of the paper is organized as follows:

1. Measures of volume
2. Adjustments to volume measures
3. Measures of quality
4. Adjustments to quality measures

1. MEASURES OF VOLUME IN EDUCATION

The primary determinant of the volume of education services is the number of pupils enrolled in an education system – the recipients of these services. The starting point for student volumes will be administrative data reported by the participating countries, not because it is the best measure but rather because it is used throughout the ICP. Household surveys, which collect data on school attendance, may serve as an important alternative source of information, particularly useful when the reliability of administrative data is in question. However, the irregularity with which these surveys are carried out, and their uneven and incomplete geographic coverage make it difficult to use them as the basis for internationally comparable measures of pupil volume.

Many ICP countries have an administrative Education Management Information System (EMIS) in place to take an annual (or semi-annual) census of all pupils in pre-school, primary, secondary and tertiary schools. While much EMIS data can be considered reliable, this is not universally the case. Concern about administrative data arises, at least in part, from the discrepancy between school participation rates as counted by administrative sources (EMIS) and household surveys. Moreover, the actual time that pupils spend in school learning within a given school year can vary considerably due to absenteeism and instructional time loss.

1.1 EMIS pupil counts and data verification

EMIS information is collected typically by school headmasters and/or teachers who fill in standard questionnaires tallying figures such as the number of pupils and teachers at their school (disaggregating by relevant categories, such as sex, grade level, or training and experience), available instructional materials, classrooms and facilities, and periodically, school finances including fees. This information is channeled up the organizational hierarchy and summed at the national level.

Though EMIS data are often the most detailed and timely data available, they do not come without shortcomings. In some cases the available counts are out of date or contain serious errors. In some countries they do not exist at all. Further, schools outside the mainstream education system, such as community schools, part-time and specialized education (vocational and professional training, arts, sports, etc.) are often not included in these questionnaires.

Thus, the EMIS pupil counts need to be verified for accuracy and reliability. Over the years, the UNESCO Institute for Statistics has been working with countries to improve their data collection methods. Education statistics experts have developed, tested and implemented methodologies to deal with measurement error and missing data. Collectively, these methodologies provide well-defined steps to procuring reasonable-to-excellent estimates of the number of pupils who are attending schools in all countries. The methods can be divided into two broad categories: first, steps to ensure the reliability of EMIS data directly; second, comparisons to pupil estimates from other sources, such as household surveys.

Effective EMIS systems mitigate potential sources of error with the following four steps[1].

1. Correct for incomplete coverage of schools (if not all schools report data). Based on the percentage of schools reporting, either the previous year’s data for a school is inserted (since, if the school has not closed, its enrollment the prior year is likely to be close to that for the current year) or a percentage adjustment is made for unreported schools.

2. Correct for incomplete or inaccurate reporting (in case headmasters count incorrectly or falsify enrollments for financial reasons). Inaccurate reporting is countered through a verification process in which, using a 2-5% sample, actual base records are checked by headquarters or regional personnel. This provides both an independent assessment, and a reference to original documents.

3. Count enrolment twice during the year. Enrollment is typically counted at the beginning of the school year, but enrollments often decline over the year. Some countries track enrollment more than once—typically at each term – or track actual weekly or monthly attendance and make adjustments of pupil counts according to actual attendance.

4. Adjust for sectors of the system not counted in any annual census—such as adult education. EMIS analysts determine that all education sub-sectors are reporting, and ensure consistent “composition” of reported education sub-sectors.

This approach to refining data accuracy is supported by recent work conducted in the UIS Statistical Capacity Building Program in seven African countries[2]. Countries participating in the ICP could be asked to document which of these four techniques were in use for which years. Where documentation of the data verification is absent, the ICP should encourage countries to implement it. The information on data verification can be used, along with a cross-comparison with household survey data (discussed below), to decide whether the EMIS pupil counts should be adjusted.

---

1 EPDC is grateful to Kurt Moses, Vice President and Director System Services Center at FHI 360, for his summary of these methods.

2 See the UIS webpage on its Statistical Capacity Building Program http://www.uis.unesco.org/ev.php?ID=5471_201&ID2=DO_TOPIC)

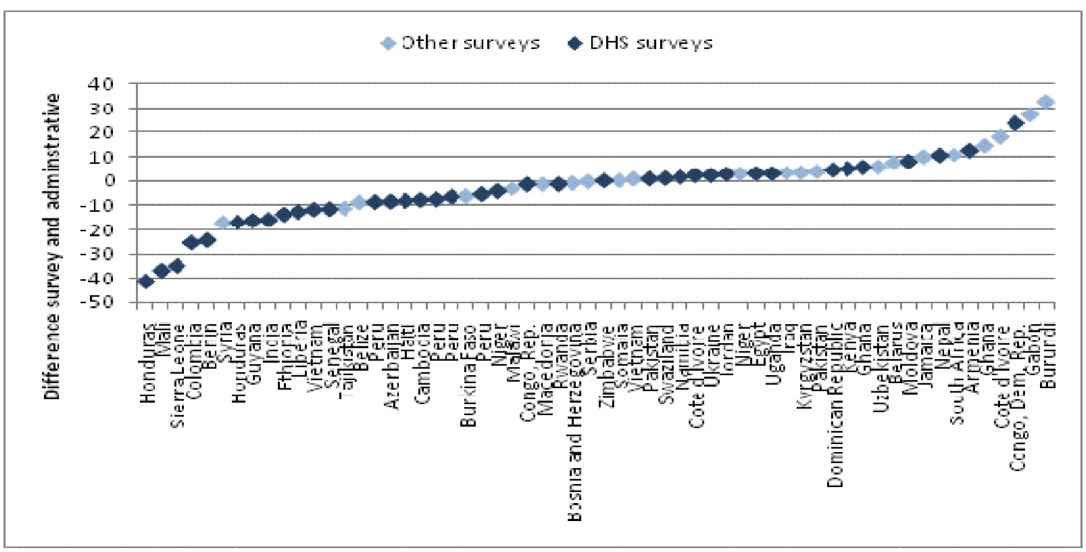
A related approach to obtaining more reliable counts is to use data from the UNESCO Institute of Statistics (UIS). The UIS has several data validation and verification techniques in the electronic survey software that countries use to provide their EMIS data to UIS, which reduce input options and chances for error. In attempting to coordinate the exchange of international data, UIS is also pursuing the standardization of indicator data exchanges through the introduction of specific software.[3]

1.2 Corroborating EMIS counts with data from other sources

A second method to test and improve the accuracy of EMIS pupil counts uses household surveys. It has long been recognized that school participation as measured by EMIS systems and by household surveys differs considerably in many countries. UNESCO, recognizing these differences, developed approaches to analyze the two data sources and identify the number most likely to be accurate. *We recommend that ICP adopt a similar approach in which EMIS pupil counts are compared against household survey counts. When differences are larger than some acceptable margin (10% may be sufficient for the ICP purposes) an expert investigation, along the lines of Stukel and Feroz-Zada (2010) discussed below, can be undertaken to determine which of the sources is more likely to be accurate.*

Figure 1.1-1 shows the difference between primary school gross enrolment rates counted by EMIS systems (pupils enrolled in school/children of primary school age) and primary school gross attendance rates counted by household surveys (children who attended school in the last week or year/children of primary school age) in 60 developing countries post-2005.[4] The DHS surveys are highlighted in dark blue because over the decades, they have earned a reputation of being highly reliable and internationally comparable.

Figure 1.1-1. Difference between primary school gross enrolment rates counted by EMIS systems (pupils enrolled in school/children of primary school age) and primary school gross attendance rates counted by household surveys (children who attended school in last week or year/children of primary school age) in 60 developing countries post-2005.



[3] See UIS website.

[4] The difference is calculated as: (100*(GER/GER-1)), which is the same calculation as used in a UIS paper by Stukel and Feroz, 2010 that analyzes these differences and is referenced later in this section.

The administrative and survey results for enrolment and attendance rates are within 10% of each other in a little over 60% of the countries in the graph – the remaining countries have larger differences. The discrepancies show both higher and lower survey values (relative to administrative), with more discrepancies on the low end. There is a notable difference between DHS and other survey types – DHS surveys tend to produce differences of more than 10 % *below* the administrative data (in 80% of the large EMIS/DHS discrepancies DHS is lower) whereas for the other surveys it is the other way around – three quarters of the survey attendance rates that are 10% or more *above* the administrative data.

When survey rates are lower than administrative rates, the reason could be that more pupils enroll in school than attend; if the converse is true it may be that students entered school later in the year and missed the enrolment counts. Differences could also be due to other errors or problems:

1) The administrative units incorrectly counted (or reported) the number of pupils.
2) The population data underlying administrative enrolment rates is faulty. If enrolment and attendance *rates* differ due to skewing by faulty population estimates, then in principle, the administrative and the survey counts of *pupils* could still be in agreement.
3) Problems with the survey sample, response rates, questionnaire, or implementation of the questionnaire. For example, the survey or EMIS questionnaire might ask only about current attendance (this week), and miss pupils who were out of school because of illness, school break, or another temporary absence.
4) Differences can also be caused by a lack of coherence between the age distribution in the survey sample and the age distribution of the population estimates. This can cause problems because attendance rates are unequally distributed over age. If particular age groups with lower or higher attendance rates are over- or under-represented, this will skew the absolute pupil estimate from the surveys.

UNESCO organizations charged with estimating the global number of children in school and *out of school* have struggled with this discrepancy for years. In 2005, UIS and UNICEF jointly developed a process for corroborating EMIS data with household survey data. The UIS/UNICEF method includes consistent definitions of school levels and school participation. If school levels and defined participation are consistent and if measurement error is minimal, then school attendance rates from a household survey and school enrollment rates from EMIS data for a given country, school level, and year should be very close. UIS and UNICEF consider differences smaller than 5 percentage points to be acceptable, and see such small deviations as an indication that both sources of information are reliable. In such cases, EMIS counts of pupils can be used with reasonable confidence. For cases where a sizeable difference remains (>5 %) it is likely that one or the other data source has an error or an omission that needs to be reconciled. The criteria and a process to locate errors and decide which of the two sources is more reliable is described in a UIS report, and are basically an expert-based analysis (UNESCO 2005).

One UIS publication of particular interest to the ICP presents an in-depth expert analysis of the differences in absolute pupil counts from EMIS and DHS surveys for 10 developing countries (Stukel and Feroz-Zada, 2010). In this analysis, EMIS pupil counts were obtained directly from UIS. Pupil counts from the DHS were obtained by multiplying the number of pupils counted by the survey in the various strata (sub-sections of the country) by the respective strata population weights. The analysis was done for

Bangladesh, Côte d'Ivoire, Egypt, Indonesia, Mozambique, Namibia, Nigeria, Rwanda, Tanzania, and Vietnam. Stukel and Feroz-Zada (2010) analyzed the cause of these differences, and concluded that *when carefully analyzed and adjusted, the two source estimates for pupil numbers are within 10% for 8 out of 10 countries.* More detail on their findings is presented in Appendix A6.

In conclusion, EPDC recommends that the ICP implement an approach similar to that of UIS/UNICEF. When EMIS pupil counts are compared against household survey counts and resulting differences are larger than some acceptable margin (10% may be sufficient for the ICP purposes), an expert investigation, along the lines of Stukel and Feroz, determines which of the sources is more likely to be accurate and those numbers are utilized to estimate pupil volume.

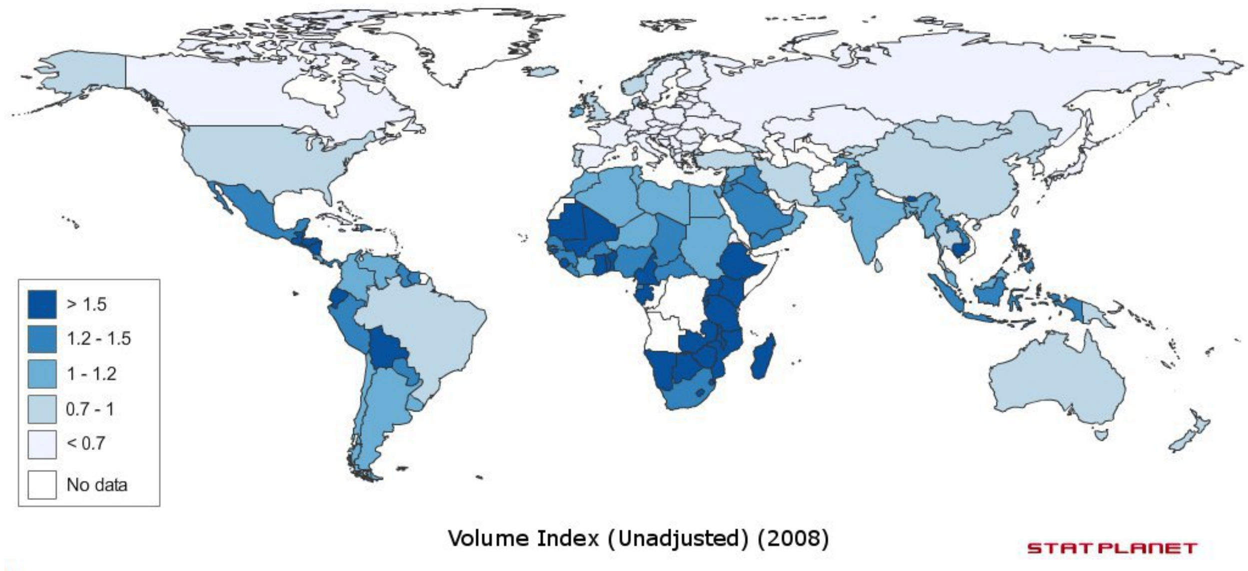
Once the pupil volume estimates have been found and agreed upon, they will be transformed to a volume index that will adjust for population size and can be used in the PPP adjustments. Eurostat (2010) uses an index based on the pupils as a portion of the population, and the volume index is the pupil ratio divided by the geometric mean of all countries' ratios:

$$V_i = \left(\frac{Pupils_i}{Pop_i}\right) \div geomean\left(\frac{Pupils_{i=1-n}}{Pop_{i=1-n}}\right)$$

where $\boldsymbol{V_i}$ is the volume index for country i, and *n* is the number of countries. The map in

Figure 1.1-2. shows the volume index for all countries. Note the higher volume indices localized in the countries with a high proportion of youth in Sub-Saharan Africa and Asia, and the relatively low indices in Eastern Europe. The overall range of the primary pupil volume index is .49 (Serbia) to 4.22 (Uganda), indicating that the volume adjustments to education output measures will be quite large.

Figure 1.1-2. Map of Volume index for primary school pupils based on EMIS pupil counts. The country level data are available in the appendix.



2. ADJUSTMENTS TO VOLUME MEASURES

It is widely acknowledged that pupil counts by themselves do not necessarily reflect educational volume because they do not include information on how many days pupils come to school; nor how many hours of education they receive on the days that they do come to school. These aspects of education are important to capture to the extent possible, because the productive exchange between teacher and student is at the core of education service production (Hill, 1975). The OECD (2007) notes the importance of collecting pupil hours (and also repetition and dropout) by level of education or grade to calculate education volume. One might argue, as Eurostat does for the OECD countries (Eurostat, 2010), that the time spent learning, or the opportunities to learn, per pupil, are already *a component of education quality* and need not be collected separately for the purposes of the education PPPs. Gallais (2006) argues that how much pupils learn (as measured in assessments, discussed next), already includes the pupils-hours component. That said, it is still useful to analyze the distribution of time spent learning, assess the data availability, and the variability in time for those countries where data is available.

2.1. Instructional time loss

The most straightforward measure of time per pupil is the official amount of instruction recommended for each school year. Though this information is not systematically maintained in an international database, many national guidelines or recommendations on instructional time can be found in curricular documentation, much of which has been compiled in the UNESCO International Bureau of Education

World Data on Education Factbook.[5] National guidelines or recommendations on instruction time come in a variety of units of measurement – hours, days, or weeks per academic year. Hours per year is the preferred unit of measurement because it is directly comparable across countries (countries may have a different number of hours in a school day or days in a school week), but instruction time recommendations are more common as measured in days or weeks. No matter what the unit of measurement, there is considerable variation in the recommended duration of instruction time per country. Colombia, for example, recommends 1,000 annual hours of instruction time for primary students while The Gambia recommends 674. Because of this wide variation in the amount of instruction time provided per country, this factor should be taken into account in a measure of volume.

A large body of literature suggests that the actual hours of time in the classroom in developing countries is far lower than the official hours of school per year (Hill, 1975; Atkinson, 2005; Schreyer, 2009; Konijn and Gallais, 2006; Lequiller, 2006; Abadzi 2007; OECD, 2007; Fraumeni, 2008; Schuh Moore et.al. 2010). Furthermore, the literature finds that even in the classroom, time is not always spent on learning tasks (time on task), sometimes because the materials necessary for teaching and learning are not available. For example, Schuh Moore, et. al (2010) find that in over 100 schools surveyed in four countries more than half of the school year was lost due to school closures, teacher absences, student absences, late starts, prolonged breaks and other reasons. Abadzi (2007 and 2009) reports similar findings. This is concerning, because opportunities to learn are an important predictor of how much children learn and thus, of school quality (Schuh Moore et.al., 2010; Abadzi, 2007; Woessmann, 2005).

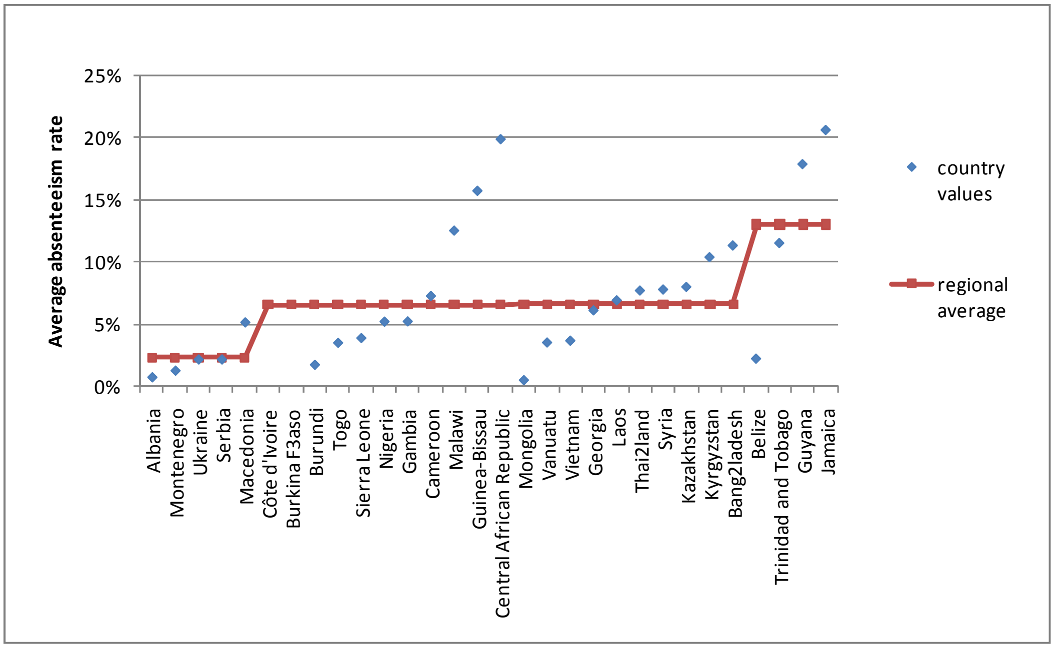
Some countries collect information on some of the time lost due to absenteeism in attendance records. More countries could do the same (suggested by e.g. Chaudary et.al. 2006; Stukel and Feroz-Zada, 2010). Beyond absenteeism, Abadzi (2007) suggests using instructional time surveys for time loss adjustments. Fraumeni et. al. (2008) suggests an aggregation of pupil hours to account for actual service delivery to reflect a time component of educational quantity. Although Fraumeni et. al., (2008) do not recommend methods for collecting actual teacher/pupil interaction time, others assume this component can be created through using official school contact hours (e.g. Eurostat, 2001; OECD, 2007). Lequiller (2006) suggests gathering this information via attendance figures, while another approach is classroom observation (discussed below).

There is no comprehensive data source for absenteeism or time spent learning. The EPDC extracted the average student attendance from 2006 MICS surveys for 30 countries; collected data on teacher absenteeism from reports; and anecdotal evidence on actual in-class time spent on learning activities.

Figure 2.1 shows the average number of days of school missed by students as a percentage of the official school days, in 30 countries with MICS 2006 surveys where the data could be collected. Absenteeism ranges from a minimum of -2% in Cote d'Ivoire, where children on average attend a bit of additional school time (presumably in supplemental, private institutions); to a high of 21% in Jamaica. There is no regional pattern for absenteeism.

[5] http://www.ibe.unesco.org/Countries/WDE/2006/index.html

Figure 2.1. Percent of days of school missed by primary pupils in 30 countries, arranged by region. Country values shown in blue and regional average values shown in red.



Teacher absenteeism for 12 countries was found in Kremer et.al (2006) and Abadzi (2009). The teacher absenteeism range is a little higher than for pupils - from 8% in the Dominican Republic to 30% in Kenya. The effects of pupil and teacher absenteeism are multiplicative, as both have to be present for learning to take place, but the EPDC found both data points for only one country, Bangladesh. There, pupils are absent 5% of the school days; and teachers 16%; suggesting there is a probability of 79% that *both* the pupil and the teacher are in the classroom on the same day.

2.2. Opportunity to learn

In addition to absenteeism, time is lost within the classroom. Abadzi (2009) provides country averages for four countries and the total time lost can range up to 50%. Clearly, these are serious time losses, but there is no empirical way to estimate them for all countries.

Opportunity to learn is a measure of the effective time spent in the classroom – the combination of resources, practices, and conditions available to provide all students with the opportunity to learn material in the national curriculum. To capture the opportunities to learn, Schuh-Moore, DeStefano, and Adelman (2010), use a classroom observation method using several instruments to collect their data including Concepts about Print (CAPs), Early Grade Reading Assessments (EGRA), Stallings Classroom observation protocols, school observations, and interviews with teachers and principals. They collect data on: % of days school is open; student and teacher attendance; % of the school day available for instruction; % of student time- on-task; % of students with a textbook; observed textbook use; % of time spent reading; class size; school support, and, as a measure of output, grade 3 reading fluency. The findings allow educators to diagnose where instructional time is lost and thus, where improvements can be made. That said, time on task data does not present information on pedagogical performance; it simply provides information about how long students and teachers are on-task and what activities are happening in the classroom. Research like this[6] is time and resource intensive. At the scale of a research study, data can be collected from a small number of schools and data collectors can afford to spend an entire day or two at the school collecting the data. At a national level, data collection would need to be done on a sample basis, requiring simplification of all instrumentation.

In conclusion, absenteeism and in-classroom time lost may be significant factors reducing the amount of time that students learn and how much they learn. Unfortunately, the data coverage for these factors is sparse, and even regional averages are not feasible. From the sample of available countries, it looks like absenteeism ranges from 0-30% -- meaning that attendance, which is the relevant indicator for the pupil volume measurement, ranges from 70-100%. In addition to absenteeism, learning time is lost within the classroom. The very sparse data indicate the levels of in-classroom time learning lost can be up to 50%. Nonetheless, we do not recommend using the existing data to impute missing absenteeism rates.

In any case, even if the absenteeism data were available, to include them in a volume adjustment might lead to double-counting its effects. This is because higher rates of instructional time lost, if they are relevant to education, will already be measured in lower levels of learning (see section 4 below).

If the data situation improves, absenteeism would be a useful indicator to collect because it can be used as a predictor of learning when other measures are absent, and can be used to corroborate measures of learning when they are available. While direct school observations may not be feasible for the ICP, sampling schools to collect some of the attendance and day-use measures may be feasible and useful.

2.3. Adjustments for zero-value education – illiterate dropouts and repetition

Even when children are in school, not all are receiving the same benefit. Two instances where students receive no, or much lower value, from their schooling are one, repetition, and two, students who drop out without learning to read. Repeaters can be said to be learning one year's worth of schooling in a second (or third) year, so all of their repeating years should be subtracted from the total pupil volume. The second group, illiterate dropouts, should also be subtracted.

In many of the lower income countries, a significant portion of students drop out of school early. An education system that consistently fails a part of its population would be considered of lower quality, or lower value, even if reasonably good results are achieved with the remaining students. One reason is that dropout is nonrandom - the students who remain in the classrooms are likely to be doing better in school than those who drop out, and the bias created by this distortion increases proportionately with the rate of dropout and repetition.

It is possible to make some adjustments for dropout bias because we know something about primary school dropouts in many countries, namely, whether they learned to read or not. The DHS and MICS surveys give a very brief, one sentence reading test to the respondents. The readers are categorized into three groups "cannot read a sentence"; "can read with difficulty"; and "can read a sentence with ease". The surveys also provide the highest education level, by grade, of the respondents. With these data, it is possible to construct a curve that shows the literacy rates of people who dropped out of school early

[6] Conducted by organizations like FHI 360 and the Research Triangle Institute (RTI)

(measured by their low education attainment). For those people who dropped out of school early, without learning to read, we say the education value was equal to zero.

Figure 2.2 shows the literacy rates of dropouts by highest grade attained for women age 15-24 in Sub-Saharan Africa and Asia. Figure 2.3 shows the survival rates to grade 6 in Africa and Asia, just to illustrate the magnitude of the dropout rates (the inverse of survival). In the calculations for zero-value education services, missing data for literacy by grade and survival were imputed using regional averages. The resulting estimates of zero-value education services for all countries where more than 5% of education services are lost are shown in Figure 2.4. There are 19 (mostly Sub-Saharan Africa) countries where more than 10% of education services are lost; and five where it is greater than 20%. The pupil volume estimates are multiplied by (1- $q0$) to get an adjusted value.

Figure 2.2. Proportion of people who can read a sentence (with difficulty or with ease) by the highest grade attained, shown separately by region. Average values for each region shown in the black curve

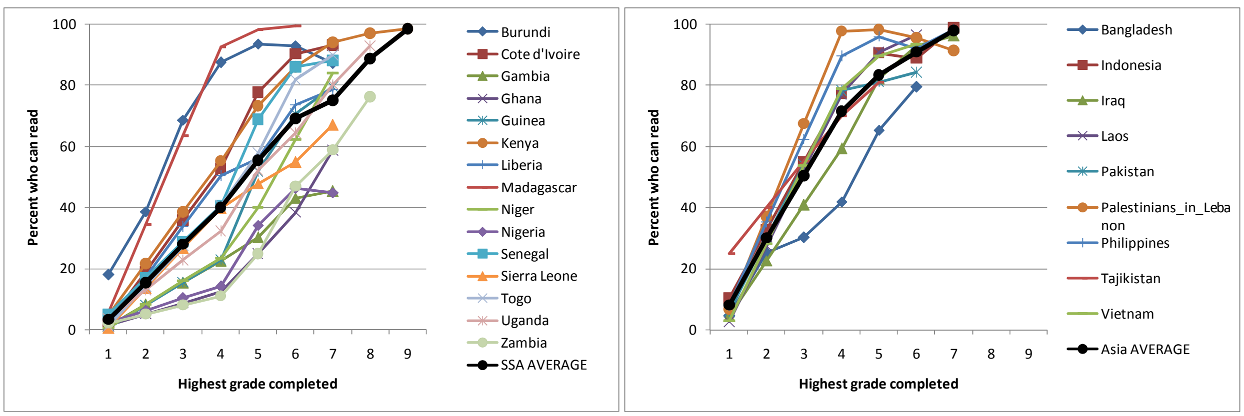
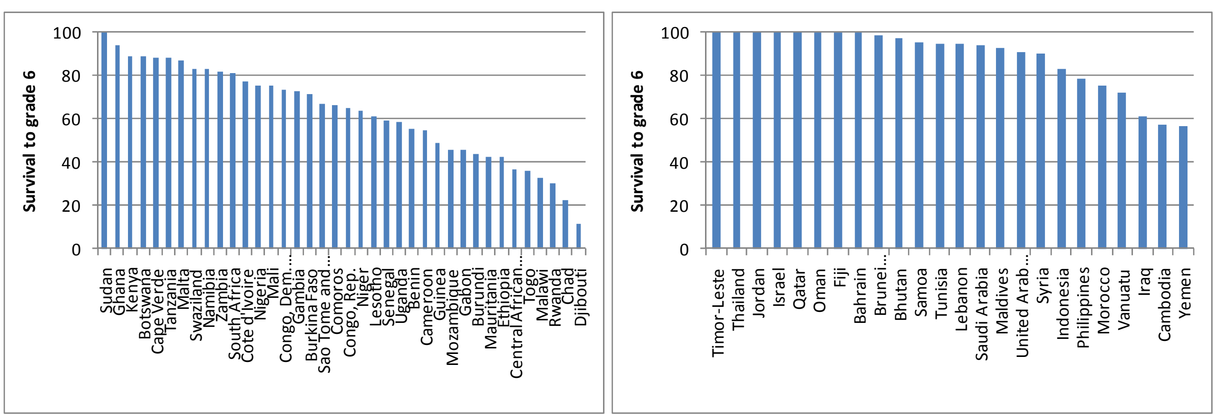


Figure 2.3. Survival rates to grade 6, sub-Saharan Africa, and Asia.



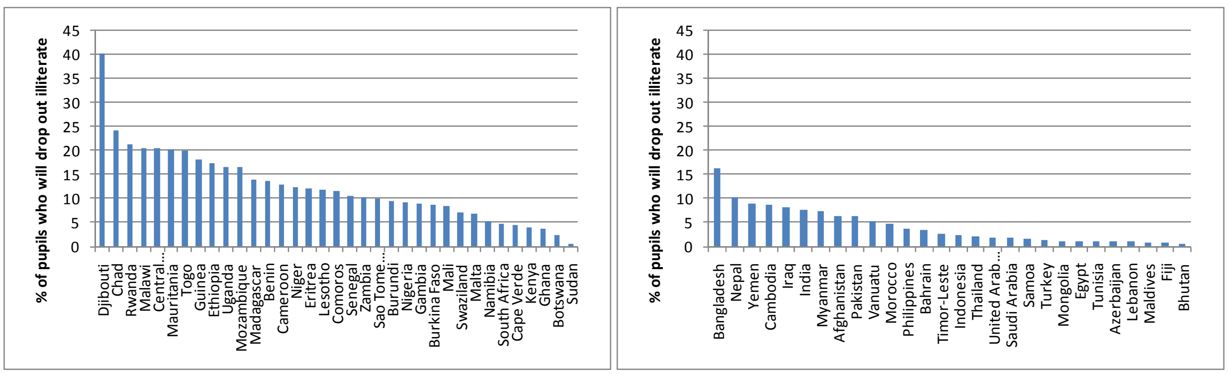
To obtain an estimate of how many children are in (primary) school in any given year, who will drop out before learning to read – and hence whose education will be counted as zero value – we combine the literacy attainment by grade of young women as a proxy for the literacy attainment dropouts today, with the by grade actual dropout rates in primary school.

The proportion of students in school, who will have zero-value education services is equal to the sum of: the percentage of students who drop out in each grade (up to grade 8), multiplied by the percent who are still illiterate when they drop out and the number of years they have been in school (proxied by grade, disregarding repetition), or:

$$q0 = \sum d_g Il_g n_g,$$

Where $q0$ is the proportion of education services that will have a value of zero; d_g is the percent of students who drop out in grade g; Il_g is the percent of drop outs in grade g who are still illiterate; n_g is the grade number – or, the number of years spent in school by students prior to dropping out. The DHS/MICS surveys provide the information on the literacy rates of those who drop out of school by primary school grade.

Figure 2.4. Percent of pupils in school who will drop out without learning to read, or, a proxy of the proportion of education services that have zero value.



2.4 . Volume adjustments: implementation

The overall primary pupil volume in each country, is the primary enrolment, with adjustments for attendance rates; in addition, possibly adjustments for absenteeism; third and fourth, subtractions of the repeaters and the illiterate dropouts:

$$Volume^{*} = Enrolment(adj) * \ Attendance * (1 - repetition) * (1 - illiterate\ dropouts)$$

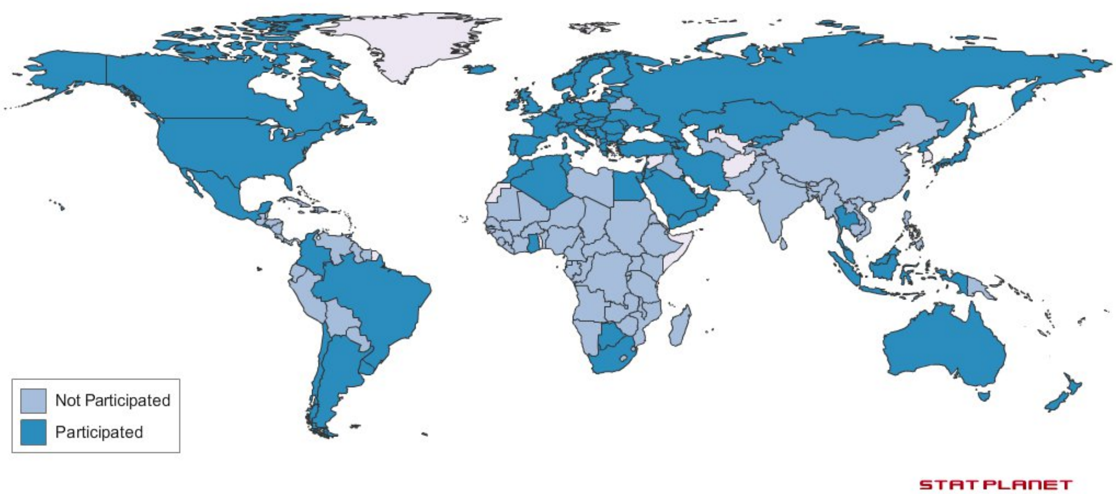
The tables in the appendix provide the unadjusted enrolments; the attendance adjustments; the repetition and the illiterate dropout adjustments for primary school pupils for all ICP countries. The adjusted volumes are shown in Table A2 in the Appendix.

3. MEASURES OF QUALITY: LEARNING OUTCOMES

One of the greatest challenges of estimating the value of education services is the lack of a readily available common measure by which to assess the quality of output. In education, the word "output" encompasses the amount of learning that is transferred from the education system to its recipients – students enrolled in formal schooling. Therefore, the volume of services produced in education must be adjusted for an estimate of quality of learning.

Education professionals agree that learning achievement is but one aspect of quality education, and that standardized tests cannot capture the entirety of the transfer and acquisition of knowledge and skills. Nonetheless, large-scale standardized tests administered across large samples of students in a growing number of countries provide the best available common metric against which to compare learning outcomes.

Figure 3.1. Geographic coverage of major international assessments (PISA 2006, TIMSS 2007, PIRLS 2006)



The geographic coverage of each individual international assessment, such as PISA, TIMSS, and PIRLS, as well as their combined geographic coverage is far from universal, with representation particularly low among low-income countries (see Figure 3.1). How can one put all countries on the same metric when participation is so uneven across tests and missing information so nonrandom? EPDC proposes a methodology to combine available scores from international assessments into a single metric, and impute missing learning scores by exploiting the relationships between quality scores and an array of macro-level indicators. This section explains this methodology.

3.1. Overview: Literature to date

The integration of test scores from various assessments into a common metric has been an area of interest of education and development economists ever since the tests grew in coverage and prominence.

Following their interest in establishing a link between education and economic growth, Hanushek and Kimko (2000) challenged the conventional measure of human capital – mean years of schooling, developed by Barro and Lee (1993; 2000; 2010) – by constructing a unified scale of cognitive skills, based on international assessments since the 1960's, which allowed them to argue that it is quality, rather than quantity of education, that was responsible for economic growth. This argument is relevant to the ICP because it suggests that in terms of the value of education output, it's the quality that counts.

Hanushek and Kimko (2000) first linked available test scores using a single mean and variance (two options were used: one assuming a constant mean and variance linked to OECD countries, and the other linked to the U.S. performance on its domestic longitudinal assessment of educational quality). Regression imputation was then used to predict quality for countries with missing test scores. These measures were later recalibrated and used in Woessmann (2000) and Hanushek and Woessmann (2009), to estimate the value of human capital and argue that there is a causal relationship between education quality and economic growth.

Crouch and Fasih (2004, unpublished, as cited in Steiner-Khamsi & Omoeva, 2009) unified scores from international and regional assessments, such as SACMEQ and SERCE, using regression imputation with a series of univariate models, transforming all of the available test scores into predicted TIMSS scores. When countries obtained more than one predicted TIMSS score, a weighted average of scores was used, with weights equal to the correlation coefficients between the original test and TIMSS. Altinok and Murselli (2006) also constructed a unified metric by directly linking various assessments building on the performance of countries they called "doubloons", whose performance on TIMSS was used as the anchor for bringing other tests into one measurement scale. No imputation of scores for countries with no assessments was performed in either Crouch and Fasih (2004) or Altinok and Murselli (2006). In another approach to predicting missing test scores, a World Bank team of economists built a student-level predictive model for PISA 2006 scores, based on coefficients on student, teacher, and school variables obtained in regressions of earlier PISA scores on these variables (Barrera et al., 2008, unpublished). The method performed reasonably well, suggesting a relative constancy of identified predictor effects on PISA scores across time.

3.2. Creating the common metric: the EPDC method

The method for the construction of the common metric of learning scores proposed by EPDC is similar to one used by Hanushek and Kimko (2000) for the last leg of their scale construction, and to the one used by Crouch and Fasih (2004). This method is regression imputation, or conditional mean imputation[7], which is based on the principle that quality of education can be predicted by a function of several indicators, within a reasonable level of uncertainty. No time series was constructed, and test measures were taken during roughly the same time period (2003 to 2007). This method expands coverage tremendously on previously constructed series, and improves precision by conditioning imputation on several factors that have shown to be highly predictive of test performance. Further, the EPDC method relaxes the need for strong assumptions by running a series of models, and gradually expands the dataset, allowing the imputation process to mitigate sample bias which would have come in effect had one model

[7] EPDC abandoned the originally proposed multiple imputation (MI) both due to the complexity of the missing data pattern and the concerns voiced by the TAG.

been run to fill all of the missing values, as observed scores are predominantly found in the wealthier countries.

The method for the construction of the common metric of learning scores proposed by EPDC is regression imputation[8]. In short, first a set of ordinary least squares (OLS) regression models was fit to the available data, to generate predicted values of a target assessment metric (PISA). The best available predicted values were then used in place of missing learning scores in a series of recursive steps until all missing values were filled.

Table 3.1. Descriptive statistics from international achievement studies

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
PISA 2006 (Science)	54	241.29	322.03	563.32	471.29	55.27	3054.28
TIMSS 2007 8th grade Math	51	291.00	307.00	598.00	452.22	72.47	5252.13
TIMSS 2007 4th grade Math	37	383.00	224.00	607.00	471.89	91.79	8425.02
PIRLS 2007	37	262.00	302.00	564.00	498.34	72.18	5210.65
SERCE 2007 6th grade Math	16	221.83	415.64	637.47	500.00	55.65	3096.63
SACMEQ 2003 Math	13	153.70	430.90	584.60	501.52	52.29	2734.59

For each country, the best available score is either its actual PISA score or a score predicted by a model with the smallest amount of error available for that country. The target of the imputation, the PISA assessment, was selected due both to its wide coverage and its conceptual characteristics (measuring the stock of learning and skills). The key features of the data available from the international and regional assessments are provided in Table 3.1.

In order to minimize the potential bias in predicted values created by the presence of OECD countries with higher PISA scores (higher intercepts in regression models), every effort was made to exclude the 25 high-income OECD countries from the prediction models. This was possible in Stage 1 with Models based on TIMSS 2007 8th grade and 4th grade science tests, and in all of the Stage 3 models. However, with models based on PIRLS 2006, the geographic distribution required the inclusion of all cases in order to compute standard errors and predicted estimates within the plausible range.

In addition, because wealth is generally associated with higher mean scores, we controlled for potential over-prediction in oil-producing countries by including either the percent GDP resulting from oil extraction, or a simple dummy of oil producer (based on Fearon and Laitin 2003, and imputed using the UN Trade % GDP from fuels data). The negative coefficient on this variable, although generally not statistically significant, would account for the lower scores of countries that are resource-rich but have lower levels of achievement in their public schools than countries at the similar level of income.

The imputation process can be broken into three broad stages:

[8] EPDC abandoned the originally proposed multiple imputation (MI) both due to the complexity of the missing data pattern and the concerns voiced by the TAG.

Stage 1. Imputations based on international assessments, and the regional SERCE assessment. The key predictor variable in these OLS models (with PISA as the dependent variable) was a score from 2007 TIMSS 8th grade math test, 2007 TIMSS 4th grade math test, 2007 PIRLS test (reading and literacy in 4th grade), and finally, for a group of Latin American countries, a score on the SERCE test. While the strongest influence for the predicted PISA score was a test score from another assessment, a set of covariates were used in these models to improve the precision of the estimates. The general equation used for regression modeling in this stage is as follows:

$$PISA_i = \alpha + \beta T_i + \gamma S_i + \delta E_i + \varepsilon \quad \text{(equation 3.1)}$$

- where $PISA_i$ is the country *i* score on the PISA assessment in 2006;
- α is the constant term;
- T_i is its score on another assessment (TIMSS, PIRLS, or SERCE);
- S_i is the available indicators of the formal schooling system of country *i* (such as per pupil expenditure, secondary gross enrollment rate, primary completion rate, repetition rate, and pupil-teacher ratio);
- E_i is a group of variables indicating the level of economic development and the demographic features of country *i* (log of household consumption per capita or GDP per capita, percent of youth ages 0-14, percent population living in urban areas, and a dummy designating oil-producing countries);
- and ε is the error term.

The variable for geographic region was not introduced at this stage because a sufficient number of cases with actual PISA scores and information on the other variables were not available in each region. The best predicted values from Stage 1 models were incorporated in the *best value of PISA* variable, or *PISA**, which now included the 56 actual scores, and the 27 imputed scores. For the most part, missing values in the predictor variables were not imputed, and therefore, a large number of models was fit using different combinations of variables to take advantage of all available information for each country. The only exception to this rule was mean years of schooling in the adult population. Because this variable is later used to adjust the *PISA** scores, it was imputed with regional and income group means. Every effort was made to include all of the available indicators, and the choices among a set of predicted values generated for a given country was made in favor of models that showed the least amount of error.

The *PISA** variable was later regressed on the available predictors in Stage 3 to obtain a new set of predicted scores, and then used to further impute missing values.

Stage 2. Imputations based on the SACMEQ assessment. Because there were no 2006 PISA scores for the group of Sub-Saharan African countries that participated in SACMEQ in 2003, the only linkages available for these countries were the Stage 1 imputed PISA scores from other tests - TIMSS 8th grade math in 2007 for Botswana and PIRLS 2006 for South Africa. However, given that only two predicted *PISA** values were available, imputing using regression of *PISA** on SACMEQ with covariates was impossible at this stage. Therefore, in Stage 2, we began by calculating the average index ratio of Botswana and South Africa's SACMEQ scores to their predicted *PISA** scores from Stage 1, and applying this ratio to compute the "starting values" of PISA for subsequent adjustment in Stage 3. The process is similar to Altinok and Murselli (2006), with the exception that in this case we only have two *doubloon* countries. The computation was as follows:

1. $\frac{SACMEQ_B}{PISA_B^*} = Index_B$

2. $\frac{SACMEQ_{SA}}{PISA_{SA}^*} = Index_{SA}$

3. $Index_X = mean(Index_B, Index_{SA})$

4. $PISA_X^{**} = SACMEQ_X \times Index_X$

The resulting starting values were incorporated into a new variable *PISA*** (**denotes a duplicate of *PISA**, with the addition of the starting values of SACMEQ) variable, which by now included actual PISA scores, the scores imputed in Stage 1, and the starting values for SACMEQ countries. This variable was regressed on a set of predictors of quality signified by vectors S and E in equation 3.1 above :

5. $PISA_X^{**} = \alpha + \gamma S_i + \delta E_i + \varepsilon$ (equation 3.2)

As a result of both the greater N of the model and the additional information gained from other variables (see variables in Table A2 in the Appendix), the predicted scores for the SACMEQ participants from these models fit using equation 3.2 were deemed more reliable than the starting values based on the ratio-linking. Because of this, we replaced the starting values *PISA***with the more reliable regression – adjusted predicted values *PISA** for the SACMEQ countries. Table 3.2 shows the SACMEQ countries with the set of their scores on SACMEQ, the transformed starting values of PISA, and the final set of scores after adjustment by other predictors.

Table 3.2 SACMEQ Starting Values and Regression-adjusted PISA scores.

Country	SACMEQ MATH	PISA** Starting Values	PISA*
Mauritius	584.6	377.59	363.76
Kenya	563.3	373.81	311.11
Seychelles	554.3	382.74	359.56
Mozambique	530	352.56	337.58
Tanzania	522.4	359.84	319.15
Swaziland	516.5	352.36	363.15
Botswana	512.9	393.57	393.57
Uganda	506.3	333.02	330.97
South Africa	486.2	298.60	298.6
Lesotho	447.2	302.61	339.7
Malawi	432.9	290.28	314.73
Zambia	432.2	293.82	282.98
Namibia	430.9	296.31	312.81

Stage 3. In this round, first *PISA*** was regressed on a set of predictors that included geographic region as well as other macro-level variables. Once the SACMEQ-based scores were adjusted by regression and incorporated in the *PISA***, that variable served as the dependent for a set of models that exploited the relationships between the now larger pool of PISA values (actual and imputed) and the country's predictors of quality.
Due to the level of missing data on a number of predictors, we first fit a set of models with the most number of predictors, minimizing the mean squared error of the estimates and the residual variance. These values were once again incorporated into *PISA**, prior to fitting more parsimonious models for countries lacking data on all but a few predictors. This was done to strengthen the robustness of the sample that went into the last few models and minimize the residual variance. As in Stage 1, for each country, the value from the model with the smallest amount of error was used to impute a missing score in Stage 3. The general functional form for models in this stage was as follows:

$$PISA_i^* = \alpha + \gamma S_i + \delta E_i + \theta R_i + \varepsilon \qquad \text{(equation 3.3)}$$

where the dependent variable is *PISA** (the "best available value of PISA") , the "other test" variable is no longer present, and a dummy variable R designating geographic region is introduced. The variables are used to the greatest extent possible, but missing data patterns dictate the use of more parsimonious models in order to obtain predicted values (fewer variables in vectors S and E).

Tables 3.3.1 and 3.3.2 provide a brief overview of the models used, the percentage of their contribution to the ultimate *PISA** metric, and the summary statistics for the models used in the imputation. The values obtained in the imputation process are presented in the Appendix A, in Table A1. The table includes an indicator of which model was used for a given imputed score, so that the reader may obtain a sense of the error associated with that predicted score by looking up the model statistics in Table 3.3.2. Regression results from the models selected for imputation are presented in Table A3. The SPSS syntax used to run the models and to incorporate predicted values in *PISA** is available upon request.

Table 3. 3 .1 Contributions to PISA* Scores from EPDC Imputation Models

Model	Test used	Number of scores obtained	Percent	Cumulative Percent
	Actual PISA Score	56	30.4	30.4
A	**TIMSS 2007 8th grade science**	**14**	**7.6**	**38.0**
B	TIMSS 2007 8th grade science	1	.5	38.6
C	TIMSS 2007 8th grade science	1	.5	39.1
D	TIMSS 2007 8th grade science	4	2.2	41.3
E	TIMSS 2007 8th grade science	1	.5	41.8
F	TIMSS 2007 8th grade science	1	.5	42.4
G	TIMSS 2007 4th grade science	2	1.1	43.5
H	PIRLS 2006	4	2.2	45.7
I	SERCE (Latin America)	8	4.3	50.0
J	SERCE (Latin America)	1	.5	50.5
K	SACMEQ (reg adjustment)	8	4.3	54.9
L	None	3	1.6	56.5
M	None	31	16.8	73.4
N	None	1	.5	73.9
O	None	17	9.2	83.2
P	None	23	12.5	95.7
Q	None	5	2.7	98.4
R	None	3	1.6	100.0
	Total	184		

Table 3. 3.2 Summary Statistics from *PISA Imputation Models**

Model	N of the model	F statistic	R-squared	Adj R-Squared	Mean Squared Error	Residual SD	% Residuals < 1 PISA SD	% Residuals < 0.5 PISA SD
A	11.0	3866.5	1.00	1.00	.7	0.3	100.0	100.0
B	12.0	111.9	1.00	0.99	40.0	3.8	100.0	100.0
C	13.0	106.1	0.99	0.98	50.0	5.0	100.0	100.0
D	15.0	121.1	0.99	0.99	43.0	4.3	100.0	100.0
E	12.0	125.1	1.00	0.99	35.8	3.6	100.0	100.0
F	17.0	69.1	0.98	0.97	101.8	7.6	100.0	100.0
G	10.0	35.5	0.99	0.96	175.1	7.6	100.0	100.0
H	24.0	7.3	0.76	0.66	658.4	21.4	95.8	87.5
I	6.0	1.1	0.82	0.10	415.3	9.1	100.0	100.0
J	6.0	0.3	0.58	(1.12)	980.4	14.0	100.0	100.0
K	56.0	25.6	0.90	0.86	675.4	22.4	96.4	83.9
L	75.0	24.8	0.84	0.81	836.2	26.3	97.3	68.0
M	51.0	20.4	0.87	0.82	702.7	23.1	98.0	70.6
N	45.0	25.9	0.91	0.87	523.6	19.5	100.0	82.2
O	60.0	21.6	0.82	0.78	788.9	25.3	98.3	70.0
P	75.0	15.9	0.71	0.67	1394.6	34.7	93.3	61.3
Q	151.0	42.9	0.73	0.72	1176.3	33.3	89.4	65.6
R	156.0	50.6	0.73	0.72	1128.3	32.7	89.7	66.7

Figure 3.2 Distribution of Imputed (Red) vs. Actual (Blue) PISA Scores

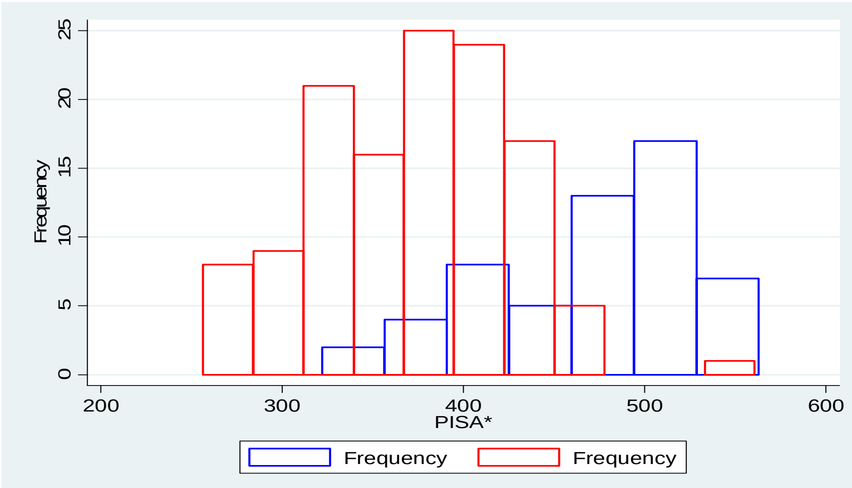
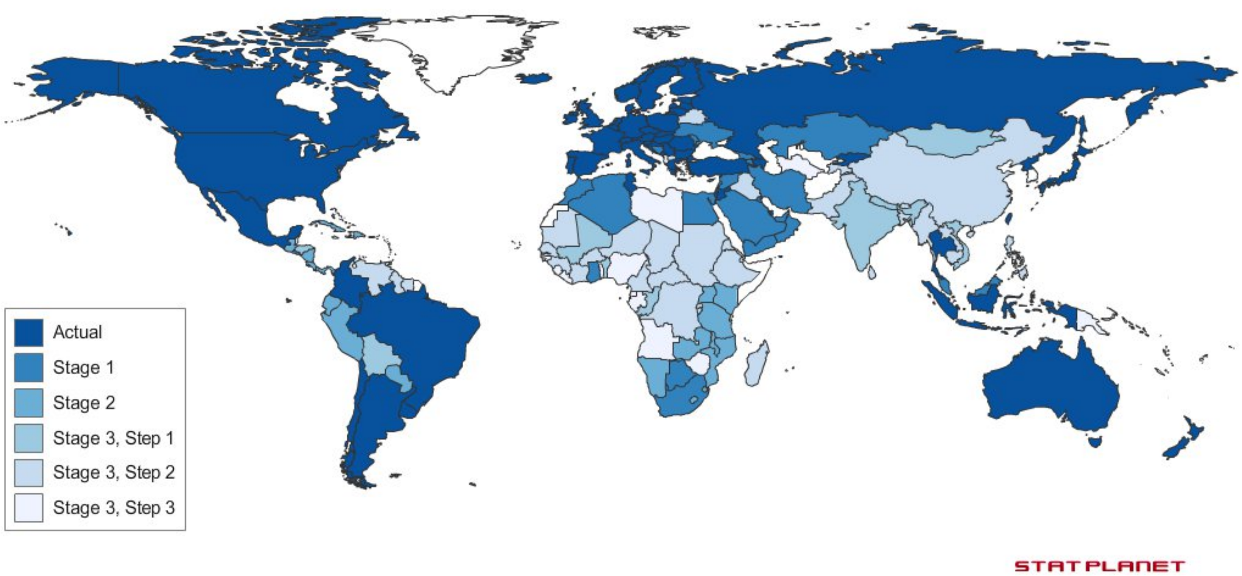


Figure 3.2 plots these imputed values alongside actual PISA values. As could be expected, most of the imputed values are in the lower half of the overall distribution since countries with missing scores tend to be those where the education systems are generally considered weaker.

In terms of geographic representation, subsequent stages of the imputation process gradually expanded coverage until it reached all 184 countries participating in the ICP 2011 round.

Figure 3.4. Geographic representation of the imputation sequence.



3.3. Error and uncertainty in the predicted scores

While it is, by definition, impossible to evaluate the size of the error for countries with no observed PISA test data, one can obtain a gauge of the error by examining the size of the residuals from models predicting scores for countries for which data is available. Generally, the smaller the residuals, relative to the range of the overall score distribution, the better the model is at predicting the outcome of interest. As noted above, for each country with a missing PISA score, the predicted estimate from the model with the smallest residual standard deviation available for that country was used as the "best value.", or *PISA** If a predicted score was not available for a given country from the model with the smallest residual standard error (due to a missing value on one of the predictors), the next best estimate for that country was used, from the model with the next lowest residual standard deviation and mean squared error.

The key assumption made here is that the model is equally good at predicting the scores for observed countries, as it is for unobserved; this assumption may or may not always hold, and it cannot be empirically tested until more learning assessment data becomes available. In other words, if the countries with missing PISA scores are vastly different from countries with observed scores (or, in Stage 3, with *PISA** scores) *in ways that cannot be controlled by the available indicators* (see Table A3 in the Appendix), then the relationship between the variables in the model may produce estimates that are farther from the true (but unobserved) values for these countries.

As Table 3.3.2 demonstrates, the standard deviation of the residuals ranged from 0.3 score points (for PISA metric see Table 3.1) in models with greater numbers of predictors, to almost 35 points. The conventional 95% confidence interval for statistical significance, therefore, may range from roughly 0.6 points above and below the estimate, to about 70 points above and below, depending on the model. This means that if the assumptions hold, the true score should be within 70 points above or below the imputed estimate, but that there is a 5% chance that the true score will be beyond those limits. The choice was made to favor stronger reliability in some of the countries for which more data was available, at the expense of consistency across the entire sample. As a result, all country scores contain a more or less equally large amount of error. As the table shows, most, if not all of the values predicted for countries that have actual PISA scores were within 1 standard deviation from their true score. The same is true for countries with *PISA** scores from Stage 1: most of the Stage 3 predictions are captured within 1 SD of their *PISA** score. There is no guarantee that this will hold for the rest of the countries with unobserved learning scores, but it gives us some degree of confidence in the reliability of the models. Nonetheless, caution must be exercised when comparing countries directly using the imputed scores since small differences (up to 55 points) may be due to random error.

Finally, when examining the variables used in imputation models in Table A3, one must keep in mind that these models are built **for prediction only**. **No causal relationship is assumed between any of these variables and the outcome of interest.** Furthermore, the coefficients might be collinear, endogenous to achievement, and therefore, inconsistent. The sole purpose of inclusion of these variables is to account for some of the variation around the learning scores. Caution must be exercised when examining and interpreting these coefficients.

Figure 3.3 *PISA** Imputed Values and Alternative Predicted Scores

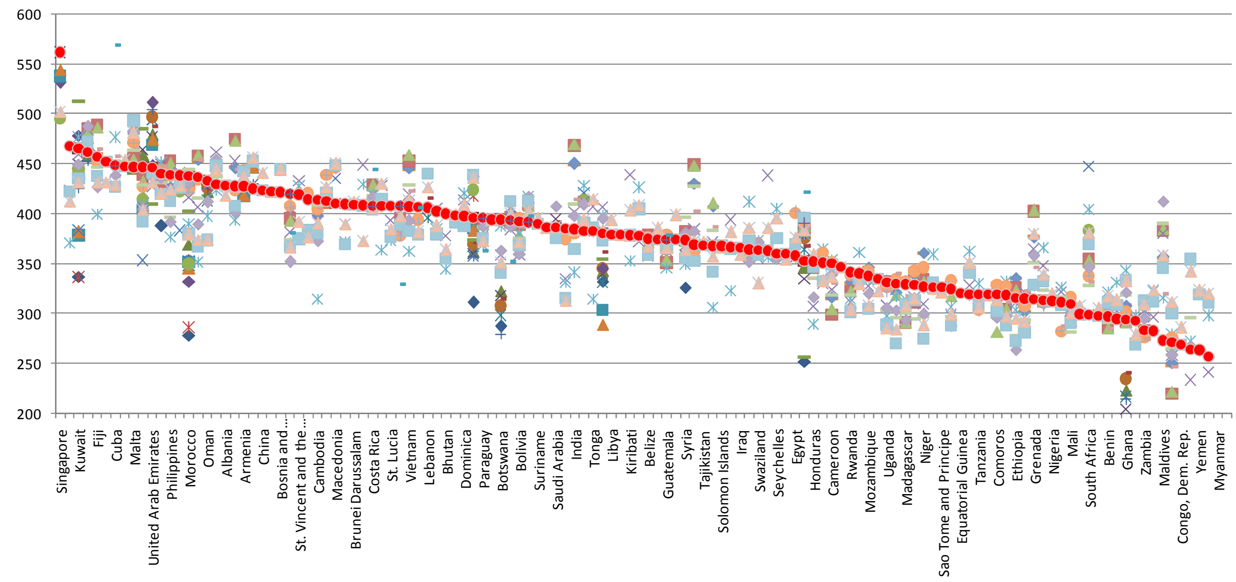


Figure 3.3 plots the predicted values obtained for the countries in the dataset vs. the final imputed values. In this graph, countries are sorted according to their final imputed score plotted in red dots, which is overlaid on top of the available predicted values, which are plotted in lighter-shade markers. As is evident from the graph, there is a considerable degree of fluctuation around the estimates, and the variability differs by country. Generally, predicted estimates are sensitive to model specifications, especially for countries with large missing data rates on many predictors. Information on the fluctuation of predicted scores for each country is provided in Table A2 in the Appendix.

3.4. Dealing with the uncertainty of imputation: Country grouping

Given the error range in the imputed scores, in particular for the countries in the lower half of the scoring range, it may be best not to use the individual imputed scores for each country, but rather to group countries and apply the average score of the group to all countries within the group.

The grouping can be done *a priori* – using a variable that has been found to correlate well with scores, such as GDP per capita, household consumption, or secondary enrolment rate. Countries would be grouped into 5 or 10 categories depending on the value of that variable. Alternatively, the grouping can be done *ex post* – using the model for imputed quality described above. The advantage of the *a priori* grouping is its transparency – one variable, which is known and measured, is used. The disadvantage is that empirically, measured learning scores do not line up exactly with any one indicator, and unmeasured learning is also unlikely to. The *a priori* grouping will result in some countries being placed in higher or lower education quality groups than they "should" have been given their actual measured learning scores or imputed scores based on a more complete model. This disadvantage would not apply to the *ex post* grouping because the countries can be divided exactly according to their measured or imputed score; on the other hand, this division is somewhat less transparent, and there is more opportunity for countries to raise objections to the model and their placement in the education quality groups.

Aside from these conceptual considerations, the grouping method can be informed by an empirical observation of how countries would line up with alternative groupings. Figures 3.5-3.7 below present a selection of three possible alternatives:

1. *A priori* grouping by one (or two) selected indicator(s);
2. *A priori* grouping by one (or two) selected indicator(s), but using actual scores for countries with a PISA test;
3. *Ex post* grouping based on the models for imputed and actual scores described above.

The graphs show the countries arranged by each of these three groupings, with their country-level imputed scores, and the average score for each group overlaid on the country-level scores. The actual indicators used in the *a priori* grouping graphs can be exchanged with others, and different indicators were tested (not shown), but the general insights remained the same. Depending on the indicator(s) used for the *a priori* grouping, however, countries land in different groups.

Figure shows the *a priori* grouping by GDP *per capita*, which is the strongest predictor of observed PISA scores where observation is possible. The countries are in five groups, with GDP *per capita* cut-off points at $1,500, $4,500, $10,000 and $20,000 (cut-offs are illustrative and provide an approximate distribution in quintiles). The figure shows, in blue, the imputed or actual scores, and overlaid in red, the average score for each of the income-determined country groups (the actual country names are not included here). The average values for each group are included in the figure, and range from 307 to 482. The figure shows that while there is some tendency for countries with higher scores to be in the higher average groups (reflected in the ascending line-up of average scores), there is also considerable overlap, that is, countries that have similar actual imputed scores landing in different country groups.

The same exercise, but with different indicators, gives the same picture, but with a different set of countries in each group. Figure 3.6 shows the shifts in groupings by three indicator sets – GDP *per capita*, secondary GER, and the product of both. The countries are arranged by income group (as in Figure) with blue diamonds, but in addition, a dot is shown for the group that each country would be in according to secondary GER (red squares), and the product of both (green triangles). The vertical lines show the shifts for particular countries, depending on each of these three criteria. For less than half (79 out of 184) countries, there are no shifts; half shift by one group (90); and 15 shift two or more groups depending on the indicator chosen.

It seems that both the high level of overlap and the country shift depending on the indicator chosen for the groups outweigh the benefits of transparency of the *a priori* grouping.

Figure 3.5. Imputed country scores and group scores by GDP/capita in 2008 (or most recent data).

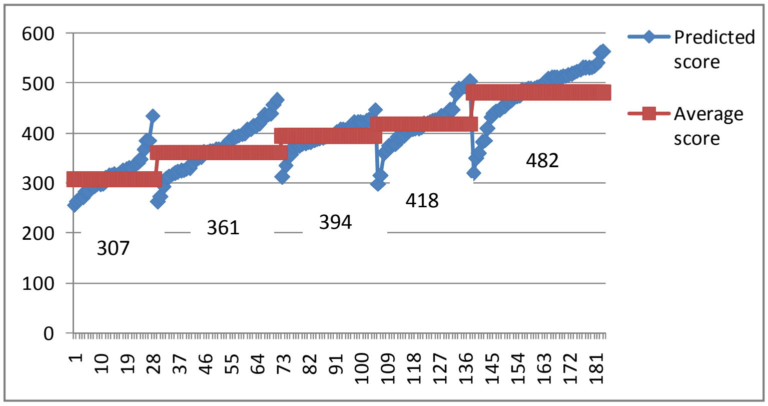
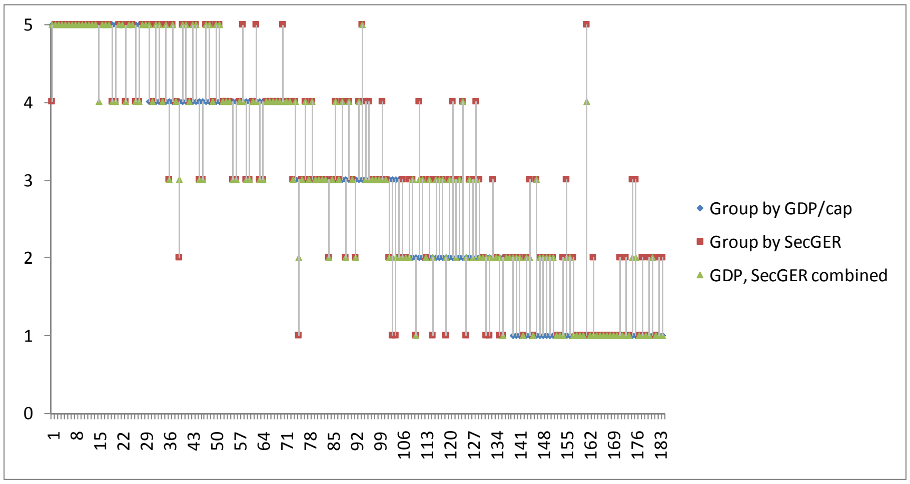
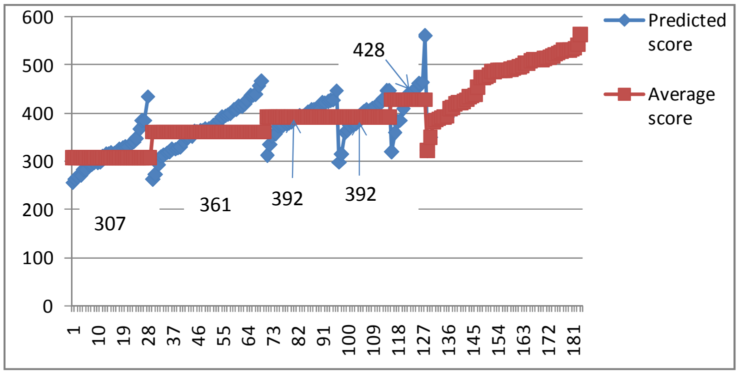


Figure 3.6. Country quintile groups by three criteria: GDP *per capita*, secondary GER, and the product of these two. Data are the most recent available in late 2010. Country names are omitted, as the graph serves only to illustrate the different slots that countries would land in depending on the indicator chosen. Actual country values available in the appendix.



An alternative grouping that would recognize the uncertainty of the imputed scores while using the certainty of the actual PISA scores, would be to use the actual PISA scores when available, and group only the countries with imputed scores. Figure shows one possible distribution with this alternative based on GDP per capita and group cut-off points the same as above (different indicators can be chosen and the cut-off points can vary with the same general result). This alternative does not reduce the level of overlap. For example, there are still a number of countries that would be in the highest average group (group score of 428) having the same country score as countries in the lowest group (group score of 307).

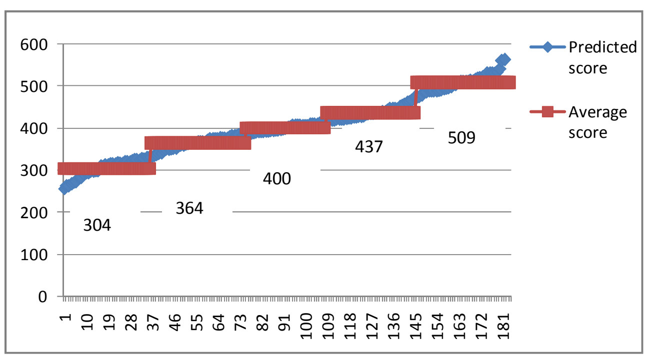
Figure 3.7. Actual PISA scores, and, for non-PISA countries, computed country scores and group scores by GDP/capita in 2008 (or most recent data)



The third grouping would be based on the model that is used to impute scores, described above. By definition, there is no overlap between groups, because the model for selecting the countries is now identical to that which is used to impute the scores. The distribution of average group scores is not much different from the *a priori* grouping by income (a range of 307 – 509 instead of 307 - 483), but the categorization of the countries is now unambiguous and there is no overlap issue.

The EPDC suggests that countries are grouped by the derived model for quality, rather than a single indicator. A possible alternative would be to use actual PISA scores for the quality index where these are available (or imputed PISA scores based on TIMMS and PIRLS) and group only the remaining countries for which uncertainty is higher (Figure 3.8).

Figure 3.8. Actual and imputed country scores and groups by scores.



3.5. Transformation of group (and observed) scores into a measure of education output

Even if the data and the methodology underlying the above quality measures are accepted, then the question remains of how to use the learning scores. How does one transform a learning score into a measure of output in education? Further, how should learning scores be combined with volume measures? These issues were discussed at length by the ICP Global Office, the members of the Technical Advisory Group, and EPDC. One aspect that was raised was the difficulty of combining quality and volume measures in a way that reflects the intuitive understanding of education output. Specifically, if country A had twice as many students as country B, but its learning score, on a conventional metric, was twice as low as that of country B, then a linear combination of these measures would produce a roughly equal level of education output (Dikhanov, email correspondence, 2010). Whether or not this would truly reflect the amount of education output, and more importantly, an adequate estimation of the education purchasing power parity, is a question that remains to be resolved in later stages of this project.

Another challenging aspect in the use of learning scores is defining the functional form of its relationship to the cost of education services. One suggestion is to experiment with functional forms by running regressions of school-level cost on quality, thereby estimating the magnitude of the effect of learning outcomes on expenditures – i.e. how much one would expect to pay for a given level of quality (Zieschang, email correspondence, 2010) While this approach may lead to a better approximation of the relation of education output to learning scores, empirically it is not possible to pursue at this time, because the school-level data on expenditures are not available and the school-level PISA scores (plus TIMMS and PIRLS, which are reasonable proxies) scores are available only for the countries that took these tests. With only country level data to derive the output, the only practical approach at this time appears to be to assume a linear relationship of education output and the learning score indices.

The essence of the email correspondence is provided in Appendix A5. The transformation of learning scores, their combination with volume measures, and the estimation of purchasing power parities remain areas for further discussion of the Technical Advisory Group and should be taken up further in a final phase of the project.

4. ADJUSTMENTS TO QUALITY MEASURES

While the assessments of learning outcomes are often taken as evidence of the contribution of the education system, a substantial portion of student achievement is the result of the enabling educational environment that the students are enjoying at home. It has been shown repeatedly in many studies that students with higher educated parents attain better results in learning, all other factors holding constant.

In this exercise, EPDC attempted to take out the variation in quality that is associated with the level of educational development in the adult population of the ICP countries. The assumption is that the more educated, on average, are the adults in a given country, the more prepared are the students to undertake their studies, and therefore, the higher their average scores. Consequently, teachers in low-income countries not only have to deal with the regular challenges of educating their youth, they also may not have the same support of the parents as that enjoyed in wealthier Western states.

EPDC's methodology is simple: we regress the actual and imputed PISA* scores on the country mean years of schooling, controlling for country wealth (GDP per capita and household consumption per capita), and use this coefficient to downward adjust the scores. Each country loses a few score points, but the loss is greater in countries with higher average level of schooling among adults. Table A2 in the Appendix shows the unadjusted and the adjusted *PISA** scores. The equations are as follows:

1. $PISA^* = \propto + \beta * MeanSchooling + \gamma * HHC + \delta * GDPpercapita + \varepsilon$
2. $Adjusted\ PISA^* = PISA^* - \beta \times MeanSchooling$

The coefficient, β, was equal to roughly 6.2. Therefore, 6.2 PISA* points were subtracted from each country's score for every year of schooling in its adult population.

A criticism of this approach is that by taking out the effect of mean schooling in the adult population, we are not only taking out the bias associated with the contribution of the parents to the learning scores, but also the contribution of the more educated teachers, administrators, and better systems. This functional form may, therefore, be a starting point for further discussion of the most appropriate form of controlling for the contribution of the families.

The quality adjustments to pupil volume will include a function of the imputed scores and the mean adult education factor.

$$Volume^{**} = Volume^{*} * f\ (PISA^{*} - 6.2MeanSchooling)$$

The quality adjustments as well as the final quality adjusted volumes are shown in tables A3 and A4 respectively.

The quality adjustments from the scores in tables A3-A4 is a simple linear function, which is not likely to be the final shape of the quality adjustment. The shape of the function to adjust scores to a quality measure was not finalized and requires further study. To get a general idea of what the function of quality might look like, it would be possible to compare the PPP's as calculated by this method with the PPP's computed on expenditure alone, with the assumption that the two series should look at least somewhat similar.

CONCLUSION

This paper presented the EPDC's proposed methodology for measuring the output of education services in developing countries, for subsequent estimation of the education purchasing power parities. The model proposed by EPDC consists of two major elements: the volume of services, measured primarily by the number of pupils in the system, and quality, measured by assessments of learning outcomes. The main challenges to validity of the volume and quality measures stem from the frequent lack of reliable data in developing countries, discrepancies between multiple sources, and biases resulting from inefficiencies such as high repetition and dropout rates. We reviewed the literature addressing these challenges, proposed ways to alleviate distortions in pupil counts, and developed a methodology for imputing missing learning scores, thereby creating a common metric of learning outcomes as a quality measure. We also suggest a set of adjustments to remove the effects of external, non-education factors on the learning outcomes, such that they only capture the contribution of the education systems.

Measurement of both of these elements involves dealing with some amount of uncertainty and noise in the data, and EPDC has proposed ways of minimizing the effect of uncertainty, by corroborating evidence, as with pupil counts, and aggregating learning scores to the group level (we show a variety of possible country groupings). Further, the transformation of obtained measures of volume and quality into a single measure of the output of education requires further discussion and review of alternatives. We hope that the methodology and models laid out in this paper will advance the understanding of the two major dimensions, and contribute to the estimation of education purchasing power parities for the 2011

International Comparison Program cycle. We expect that the methodology can and will be further refined, possibly simplified, based on more empirical research with actual data and alternative education value computations, combining different components of the methodology described here. The issue of how to functionally translate imputed learning scores into quality measures also remains to be addressed in a final phase of the project.

REFERENCES

Abadzi, H. (2007) Instructional time loss and local-level governance. Prospects. 37(1) 13-16.

Abadzi, H. (2009). Instructional Time Loss in Developing Countries: Concepts, Measurement, and Implications. World Bank Research Observer 24(2): 267-290. Oxford Journals

Altinok, N. (2010) Do School Resources Increase School Quality? Working Paper, Institut de Reserche sur L'Education. Dijon, France.

Altinok, N., Murseli, H. (2007). International Database on Human Capital Quality. *Economics Letters*, 96(2), pp. 237-244.

Atkinson, A. (2005) The Atkinson Review: Final Report. Macmillan: Houndmills

Barrera, F., Garcia-Moreno, V., Metzger, S., MacDonald, K., Patrionos, H.A., and Porta, E. (2008). Using PISA to Understand the Variation in Learning Outcomes and to Predict Country Performance. Draft Report provided by authors.

Barro R., and Lee, J.W. (1993). International comparisons of educational attainment. *Journal of Monetary Economics, 32*(3), 363-394.

Barro R., and Lee, J.W. (2000). International data on educational attainment: Updates and implications. *Oxford Economic Papers, 53*(3), 541-563.

Barro R., and Lee, J.W. (2010) A new data set of educational attainment in the world, 1950-2010. Working Paper 15902. National Bureau of Economic Research. Cambridge, USA.

Chaudhary, N., Hammer, J., Kramer, M., Muralidharan, K. and F. H. Rogers. (2006) Missing in action: Teacher and health worker absence in developing countries. Journal of Economic Perspectives. 20(1) 91-116.

Crouch, L, and Vinjevold, P. (2006). South Africa: Access before Quality, and What to Do Now? Revista de curriculum y formacion del profersorado 10(1). *http://www.ugr.es/local/recfpro/Rev101ART6ing.pdf*

Crouch, L. and Fasih, T. (2004) Patterns in educational development: Implications for further efficiency analysis. Unpublished draft shared by the authors.

Eurostat (2001) Handbook on Price and Volume Measures in National Accounts European Commission. European Commission: Brussels.

Fearon, J.D. and Laitin, D.D. (2003) Ethnicity, Insurgency and Civil War. American Political Science Review, 97(1), 75-90. Data available at: http://www.stanford.edu/group/ethnic/publicdata/publicdata.html

Fraumeni, B., Reinsdorf, M., Robinson, B., and M. Williams. (2008) Price and Real Output Measures for the Education Function of Government: Exploratory Estimates for Primary and Secondary Education. NBER Working Paper #14099. NBER: Cambridge.

Gallais (2006) on pupil learning Consultation Notes on Measuring Educational Output and Quality. unpublished document accessed 4/1/2010 at: *www.ons.gov.uk/.../measuring.../response-to-establishing-the-principles-from-alain-gallais.pdf*

Hanushek, E.A. and Kimko, D.D. (2000) Schooling, Labor Force Quality and the Growth of Nations. *American Economic Review*, 90(5), pp. 1184-1208

Hanushek, E.A. and Woessmann, L (2007). The Role of Education Quality in Economic Growth. World Bank Policy Research Working Paper 4122. World Bank, Washington, DC, USA.

Hanushek, E.A. and Woessmann, L (2010a). Sample selectivity and the Validity of International Student Achievement Tests in Economic Research. Working Paper 15867. National Bureau of Economic Research. Cambridge, USA.

Hanushek, E.A. and Woessmann, L (2010b). The Economics of International Differences in Educational Achievement Working Paper 15949. National Bureau of Economic Research. Cambridge, USA.

Hanushek, E.A., and Woessmann, L. (2009). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Working Paper 14633, National Bureau of Economic Research.* Cambridge, MA, USA.

Hanushek, E.A., Luque, J.A. (2003), Efficiency and Equity in Schools around the World.

Hill, 1975. The production boundary and nature of output (with special emphasis on health and education). accessed April 1, 2010 at: http://siteresources.worldbank.org/ICPEXT/Resources/ICP_2011.html

Konijn, P, and A. Gallais. (2006) Non-market Services: Overview of Eurostat and OECD activities. Powerpoint Presentation for OECD conference in London, October. Accessed April 1 at: www.oecd.org/dataoecd/43/2/37470016.ppt

Kremer, M., Chaudhury, N., Halsey Rogers, F., Muralidharan, K., Hammer, J. (2005). Teacher Absenteeism in India: A Snapshot. Journal of the European Economic Association 3(2-3): 658-667.

Lavy, V. (2010). Do Differences in School's Instruction Time Explain International Achievement Gaps in Math, Science, and Reading? Evidence from Developed and Developing Countries. Working Paper 16227. National Bureau of Economic Research. Cambridge, USA.

Lequiller, F. (2006) Measurement of Non-Market Output. Paper presented as the Fourth meeting of the Advisory Expert Group on National Accounts. February. Frankfurt

OECD (2006) Eurostat-OECD Methodological Manual on Purchasing Power Parities. OECD. Paris, France

OECD (2007) PPP Methodological Manual, Chapter 7: Calculation and Aggregation of PPPS. OECD. Paris, France.

OECD. (2010) *Towards Measuring the Volume Output of Education and Health Services: A Handbook*. Working Paper No. 31. Statistics Directorate, OECD. Paris, France.

Schreyer. P. (2009) *Update on PPP program and new data set for health expenditure*. Power point presentation for working party on national accounts held in November and accessed April 1, 2010 at: www.oecd.org/dataoecd/2/2/44043181.ppt

Schuh-Moore, A., DeStefano, J. and E. Adelman. (2010) *Opportunity to Learn as a Measure of School Effectiveness in Ethiopia, Guatemala, Honduras, and Nepal.* Equip 2 Working Paper. FHI 360: Washington, DC

Steiner-Khamsi, G. and Omoeva, C. , with Van Keuren, C. and Shiotani, A. (2009). *Access and Quality in Basic Education: Is There a Tradeoff?* USAID Education Strategy Development. Washington, DC: USAID.

Stukel, D. and Feroz-Zada, Y. (2010) *Measuring Educational Participation: Analysis of Data Quality and Methodology Based on Ten Studies*. UNESCO Institute for Statistics. Montreal, Canada.

UNESCO-UIS (2005) *Children Out of School: Measuring Exclusion from Primary Education*.

Woessman, L. (2007) *International evidence on expenditures and class size*. Brookings Papers on Educational Policy.

Woessmann, L. (2000). *Specifying human capital: A review, some extensions, and development effects.* Kiel Working Paper No. 1007. Kiel, Germany.

World Bank. (2008) *Global Purchasing Power Parities and Real Expenditures: 2005 International Comparison Program.* International Bank for Reconstruction and Development and World Bank. Washington, DC.

APPENDIX

Progress in International Reading and Literacy Study - PIRLS

PIRLS is administered by the International TIMSS and PIRLS Center under the auspices of the International Association for Evaluation of Educational Achievement (IEA). This study tested 4th grade students for reading and literacy skills in 2001 and 2006. In 2006, 38 countries participated in the PIRLS and the next round is scheduled for 2011. A comprehensive survey of families was administered in both years as part of the study, providing a wealth of information not only on the student background, but also on the household practices contributing to reading and literacy.

Program for International Student Assessment - PISA

PISA tests science, math and reading skills of 15-year olds and has been administered four times since 2000 in three-year cycles (2000, 2003, 2006, 2009). Organized and managed by the OECD, PISA is designed to capture and compare the preparedness of older students for entry into the labor force across countries. PISA has a diversity of academic content often not found in other international assessments .Similarly to other international studies, PISA includes surveys of student and family demographic data which allows for the control of non-school factors during analysis.

Trends in International Mathematics and Science Study - TIMSS

TIMSS is another assessment administered by the International TIMSS and PIRLS Center under the auspices of the IEA. Target groups are $4^{th}$and 8^{th} grade students, and target areas are mathematics and the natural sciences.

Regional Assessments

The major regional assessments of learning outcomes are SAQMEC (Southern Africa), LLECE (Latin America) and PASEC (French-speaking Africa).

Table A1. Actual and Imputed PISA scores, scores adjusted for mean schooling, and the variability across the predicted values by country.

Income group	Country	Actual or Imputed	Unadjusted *PISA** Score OCT 14	*PISA** Adjusted for mean schooling	Model	SD of predicted values
Lower middle income	Albania	Imputed	427.68	365.8	P	5.3
Upper middle income	Algeria	Imputed	379.88	328.48	A	29.17
Lower middle income	Angola	Imputed	334.37	292.65	P	6.77
High income: nonOECD	Antigua and Barbuda	Imputed	374.22	305.09	P	6.27
Upper middle income	Argentina	Actual	391.24	338.09		8.83
Lower middle income	Armenia	Imputed	427.11	364.61	D	9.47
High income: OECD	Australia	Actual	526.88	456.44		
High income: OECD	Austria	Actual	510.84	440.61		5.5
Lower middle income	Azerbaijan	Actual	382.33	320.46		6.08
High income: nonOECD	Bahamas	Imputed	451.81	382.63	P	11.98
High income: nonOECD	Bahrain	Imputed	439.6	382.42	A	11.39
Low income	Bangladesh	Imputed	435.7	405.01	M	37.34
High income: nonOECD	Barbados	Imputed	421.92	352.79	Q	0.01
Upper middle income	Belarus	Imputed	446.55	386.78	P	4.37
High income: OECD	Belgium	Actual	510.36	443.45		7.09
Lower middle income	Belize	Imputed	374.59	327.39	M	6.86
Low income	Benin	Imputed	296.73	276.97	M	13.02
Lower middle income	Bhutan	Imputed	399.41	361.65	O	17.42
Lower middle income	Bolivia	Imputed	391.44	331.44	M	12
Upper middle income	Bosnia and Herzegovina	Imputed	421.12	361.35	F	11.51
Upper middle income	Botswana	Imputed	393.57	348.62	A	26.77
Upper middle income	Brazil	Actual	390.33	352.22		16.55
High income: nonOECD	Brunei Darussalam	Imputed	408.45	357.05	P	2.29
Upper middle income	Bulgaria	Actual	434.08	368.81		21.18
Low income	Burkina Faso	Imputed	318.09	307.07	M	13.09
Low income	Burundi	Imputed	292.6	267.37	O	10.8
Low income	Cambodia	Imputed	413.67	386.91	M	14.12
Lower middle income	Cameroon	Imputed	349.55	312.34	M	14.71
High income: OECD	Canada	Actual	534.47	453.72		
Lower middle income	Cape Verde	Imputed	319.22	277.49	O	14.14
Low income	Central African Republic	Imputed	384.64	359.38	M	30.65
Low income	Chad	Imputed	330.01	314.88	M	25.42
Upper middle income	Chile	Actual	438.18	372.9		7.91
Lower middle income	China	Imputed	422.88	377.74	P	10.19

Income group	Country	Actual or Imputed	Unadjusted *PISA** Score OCT 14	*PISA** Adjusted for mean schooling	Model	SD of predicted values
Upper middle income	Chinese Taipei (Taiwan)	Actual	532.47	466.13		18.96
Upper middle income	Colombia	Actual	388.04	340.6		6.22
Low income	Comoros	Imputed	318.88	291.58	M	12.41
Low income	Congo, Dem. Rep.	Imputed	268.44	243.2	P	10.05
Lower middle income	Congo, Rep.	Imputed	294.2	252.48	O	8.25
Upper middle income	Costa Rica	Imputed	407.7	359.47	I	12.39
Lower middle income	Côte d'Ivoire	Imputed	350.6	328.69	M	8.6
High income: nonOECD	Croatia	Actual	493.2	443.17		9.66
Upper middle income	Cuba	Imputed	447.89	389.26	I	47.12
High income: nonOECD	Cyprus	Imputed	461.34	402.5	A	10.5
High income: OECD	Czech Republic	Actual	512.86	452.66		6.8
High income: OECD	Denmark	Actual	495.89	421.46		5.92
Lower middle income	Djibouti	Imputed	408.05	370.67	P	36.04
Upper middle income	Dominica	Imputed	397.32	349.33	O	10.56
Upper middle income	Dominican Republic	Imputed	393.07	341.85	I	16.88
Lower middle income	Ecuador	Imputed	406.14	356.76	J	11.29
Lower middle income	Egypt	Imputed	357.59	318.05	D	16.65
Lower middle income	El Salvador	Imputed	352.53	308.56	A	20.95
High income: nonOECD	Equatorial Guinea	Imputed	320.07		R	
Low income	Eritrea	Imputed	324.38	297.95	M	15.47
High income: nonOECD	Estonia	Actual	531.39	462.94		9.04
Low income	Ethiopia	Imputed	315.23	305.78	M	21.07
Upper middle income	Fiji	Imputed	456.45	377.39	M	20.91
High income: OECD	Finland	Actual	563.32	483.54		15.56
High income: OECD	France	Actual	495.22	439.72		5.72
Upper middle income	Gabon	Imputed	359.16	312.26	P	2.84
Low income	Gambia, The	Imputed	282.54	257.3	O	18.02
Lower middle income	Georgia	Imputed	395.71	333.84	A	31.28
High income: OECD	Germany	Actual	515.65	436.55		3.02
Low income	Ghana	Imputed	293.56	249.68	A	44.38
High income: OECD	Greece	Actual	473.38	417.83		8.09
Upper middle income	Grenada	Imputed	314.25	302.71	M	29.17
Lower middle income	Guatemala	Imputed	373.91	343.03	I	7.56
Low income	Guinea	Imputed	271.1	256.27	M	32.73
Low income	Guinea-Bissau	Imputed	297.47	272.24	P	4.46
Lower middle income	Guyana	Imputed	368.3	315.86	M	29.07
Low income	Haiti	Imputed	346.6	318.02	Q	0.01
Lower middle income	Honduras	Imputed	351.63	320.62	M	19.32
High income: nonOECD	Hong Kong, China	Actual	542.21	472.88		13.13

Income group	Country	Actual or Imputed	Unadjusted *PISA** Score OCT 14	*PISA** Adjusted for mean schooling	Model	SD of predicted values
Upper middle income	Hungary	Actual	503.93	439.31		15.02
High income: OECD	Iceland	Actual	490.79	426.19		10.45
Lower middle income	India	Imputed	383.91	354.81	M	32.04
Lower middle income	Indonesia	Actual	393.48	350.24		20.01
Lower middle income	Iran	Imputed	446.18	405.34	A	21.77
Lower middle income	Iraq	Imputed	365.59	328.2	P	3.78
High income: OECD	Ireland	Actual	508.33	443.34		5.45
High income: nonOECD	Israel	Actual	453.9	402.51		11.31
High income: OECD	Italy	Actual	475.4	420.04		13.4
Upper middle income	Jamaica	Imputed	401.79	353.8	O	10.9
High income: OECD	Japan	Actual	531.39	462.37		
Lower middle income	Jordan	Actual	421.97	373.78		12.31
Upper middle income	Kazakhstan	Imputed	412.2	346.22	G	9.61
Low income	Kenya	Imputed	311.11	267.82	K	13.34
Lower middle income	Kiribati	Imputed	377.98	328.32	P	24.83
High income: OECD	Korea, Rep.	Actual	522.15	454.66		1.49
High income: nonOECD	Kuwait	Imputed	465.04	413.64	A	36.06
Low income	Kyrgyzstan	Actual	322.03	259.2		0.02
Low income	Laos	Imputed	365.96	336.88	O	12.15
Upper middle income	Latvia	Actual	489.54	419.06		12.36
Upper middle income	Lebanon	Imputed	405.49	354.09	D	12.98
Lower middle income	Lesotho	Imputed	339.7	297.97	L	5.58
Low income	Liberia	Imputed	256.44	231.21	O	37.75
Upper middle income	Libya	Imputed	378.62	327.22	P	8.58
Upper middle income	Lithuania	Actual	487.96	422.91		9.73
High income: OECD	Luxembourg	Actual	486.32	425.93		10.19
High income: nonOECD	Macao, China	Actual	510.84	459.26		9.31
Upper middle income	Macedonia	Imputed	409.52	363	H	17.38
Low income	Madagascar	Imputed	329.03	295.24	M	12.14
Low income	Malawi	Imputed	314.73	287.38	K	17.16
Upper middle income	Malaysia	Imputed	428.7	375.04	E	10.74
Lower middle income	Maldives	Imputed	272.97	235.58	M	42.83
Low income	Mali	Imputed	309.58	299.35	M	6.06
High income: nonOECD	Malta	Imputed	446.23	387.98	A	15.69
Low income	Mauritania	Imputed	328.2	308.63	M	4.73
Upper middle income	Mauritius	Imputed	363.76	319.92	K	11.59
Upper middle income	Mexico	Actual	409.65	361.97		7.04
Lower middle income	Micronesia	Imputed	373.74	324.07	P	14.18
Lower middle income	Moldova	Imputed	437.72	375.84	H	7.12

Income group	Country	Actual or Imputed	Unadjusted *PISA** Score OCT 14	*PISA** Adjusted for mean schooling	Model	SD of predicted values
Lower middle income	Mongolia	Imputed	427.23	366.79	M	23.72
Upper middle income	Montenegro	Actual	411.79	352.02		0
Lower middle income	Morocco	Imputed	437.08	412.37	A	43.45
Low income	Mozambique	Imputed	337.58	321.21	K	14.29
Low income	Myanmar	N/A				
Upper middle income	Namibia	Imputed	312.81	270.67	K	15.33
Low income	Nepal	Imputed	367.44	349.22	M	21.24
High income: OECD	Netherlands	Actual	524.86	457.12		6.07
High income: OECD	New Zealand	Actual	530.38	450.09		
Lower middle income	Nicaragua	Imputed	419.34	369.51	I	21.87
Low income	Niger	Imputed	326.72	318.3	M	21.47
Lower middle income	Nigeria	Imputed	312.07	278.59	P	4.32
High income: OECD	Norway	Actual	486.53	407.42		11.79
High income: nonOECD	Oman	Imputed	432.72	381.33	A	25.31
Lower middle income	Pakistan	Imputed	413.76	389.38	M	17.93
Lower middle income	Palestinian Autonomous Territories	N/A				
Upper middle income	Panama	Imputed	391.14	337.31	I	7.34
Lower middle income	Papua New Guinea	Imputed	378.56	328.9	Q	0.01
Lower middle income	Paraguay	Imputed	395.56	347.96	I	9.71
Upper middle income	Peru	Imputed	407.29	356.53	I	21.26
Lower middle income	Philippines	Imputed	438.61	383.31	M	19.12
Upper middle income	Poland	Actual	497.81	436.55		10.28
High income: OECD	Portugal	Actual	474.31	430.07		
High income: nonOECD	Qatar	Actual	349.31	297.92		43.64
Upper middle income	Romania	Actual	418.39	360.78		15.68
Upper middle income	Russian Federation	Actual	479.47	420.91		12.59
Low income	Rwanda	Imputed	341.05	315.22	M	15.31
Lower middle income	Samoa	Imputed	407.63	357.97	O	10.46
Lower middle income	Sao Tome and Principe	Imputed	325.67	283.94	P	2.41
High income: nonOECD	Saudi Arabia	Imputed	385.71	340.12	B	9.36
Low income	Senegal	Imputed	325.7	300.47	O	2.86
Upper middle income	Serbia	Actual	435.64	375.87		9.11
Upper middle income	Seychelles	Imputed	359.56	314.61	L	18.99
Low income	Sierra Leone	Imputed	299.01	273.77	P	5.2
High income: nonOECD	Singapore	Imputed	561.14	508.64	C	24.5
Upper middle income	Slovakia	Actual	488.43	428.67		14.51
Upper middle income	Slovenia	Actual	518.82	451.28		23.68

Lower middle income	Solomon Islands	Imputed	367.18	317.52	P	0.62
Upper middle income	South Africa	Imputed	298.6	251.7	H	27.14
Income group	**Country**	**Actual or Imputed**	**Unadjusted *PISA** Score OCT 14**	***PISA** Adjusted for mean schooling**	**Model**	**SD of predicted values**
High income: OECD	Spain	Actual	488.42	441.45		7.92
Lower middle income	Sri Lanka	Imputed	397.96	337.9	M	3.76
Upper middle income	St. Kitts and Nevis	Imputed	409.04	361.05	O	14.06
Upper middle income	St. Lucia	Imputed	407.51	359.53	O	11.71
Upper middle income	St. Vincent and the Grenadines	Imputed	418.68	370.69	O	23.26
Lower middle income	Sudan	Imputed	319.06	277.33	P	0.43
Upper middle income	Suriname	Imputed	389.29	341.31	P	3.3
Lower middle income	Swaziland	Imputed	363.15	288.96	K	16.69
High income: OECD	Sweden	Actual	503.33	435.63		4.9
High income: OECD	Switzerland	Actual	511.52	437.65		
Lower middle income	Syria	Imputed	373.09	339.43	A	8.48
Low income	Tajikistan	Imputed	367.54	304.92	P	1.86
Low income	Tanzania	Imputed	319.15	292.83	L	6.37
Lower middle income	Thailand	Actual	421.01	376.73		8.87
Low income	Timor-Leste	Imputed	362.71	333.63	P	31.83
Low income	Togo	Imputed	386.04	360.29	Q	0.01
Lower middle income	Tonga	Imputed	382.08	332.42	O	15.77
High income: nonOECD	Trinidad and Tobago	Imputed	382.49	313.36	H	12.06
Lower middle income	Tunisia	Actual	385.51	348.12		18.23
Upper middle income	Turkey	Actual	423.83	389.33		8.81
Lower middle income	Turkmenistan	Imputed	393.58	329.36	Q	0.01
Low income	Uganda	Imputed	330.97	295.19	K	18.73
Lower middle income	Ukraine	Imputed	424.74	365.86	D	13.2
High income: nonOECD	United Arab Emirates	Imputed	445.4	394.01	A	28.99
High income: OECD	United Kingdom	Actual	514.77	453.73		
High income: OECD	United States	Actual	488.91	416.42		
Upper middle income	Uruguay	Actual	428.13	378.16		13.27
Lower middle income	Vanuatu	Imputed	467	417.33	O	22.82
Upper middle income	Venezuela	Imputed	377.16	329.17	O	18.08
Low income	Vietnam	Imputed	406.7	372.06	M	22.71
Low income	Yemen	Imputed	263.06	225.67	G	72.75
Low income	Zambia	Imputed	282.98	241.94	K	14.25
Low income	Zimbabwe	Imputed	263.56	215.07	N	51.64

Figure A.1 Imputed (Red) vs. Actual (Blue) PISA Scores by Log Household Consumption

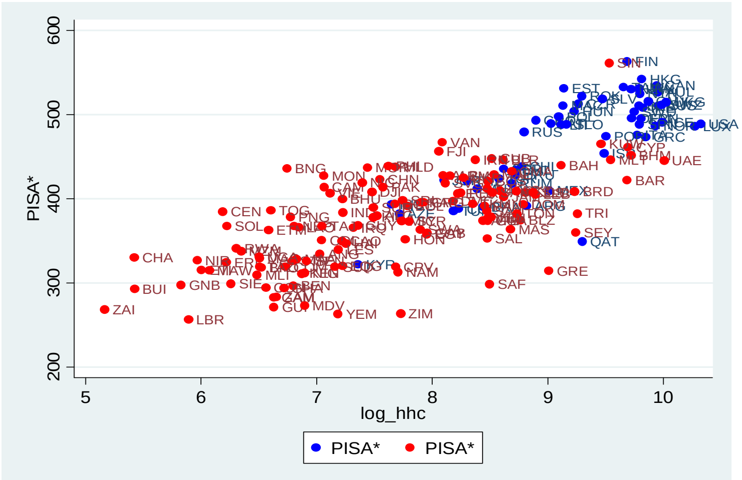


Figure A.2 Imputed (Red) vs. Actual (Blue) PISA Scores by Mean Years of Schooling in Adults 25+

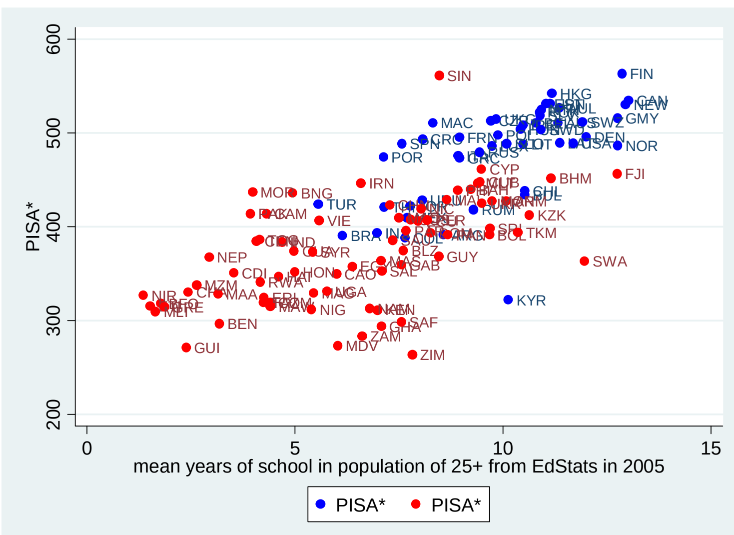


Table A2. Primary pupil volume indicators.

PRIMARY PUPIL VOLUME ESTIMATES

Country	Most recent primary enrolment	Population ~2008	Pupils as % of population	Estimated pupil absenteeism (%)	Estimated pupil attendance (%)	Repetition rate (%)	Non-repeating pupils (%) (repeaters value=0)	Pupils who dropout illiterate (%)	Pupils who learn to read (%)	Effective primary pupils as % of population
Albania	250	3,143	8	1	99		100	0	100	**8**
Algeria	3,942	34,373	11	2	98	8	92	0	100	**10**
Angola	#VALUE!	18,021	#VALUE!	7	93		100	0	100	**#N/A**
Antigua and Barbuda	11	#VALUE!	#VALUE!	13	87	6	94	0	100	**#N/A**
Argentina	4,700	39,883	12	13	87	6	94	0	100	**10**
Armenia	122	3,077	4	2	98	0	100	0	100	**4**
Australia	1,978	21,074	9	2	98		100	0	100	**9**
Austria	337	8,337	4	2	98		100	0	100	**4**
Azerbaijan	490	#VALUE!	#VALUE!	7	93	0	100	1	99	**#N/A**
Bahamas	37	338	11	2	98		100	0	100	**11**
Bahrain	86	776	11	7	93	3	97	3	97	**10**
Bangladesh	16,002	160,000	10	7	93	20	80	16	84	**6**
Barbados	23	#VALUE!	#VALUE!	13	87		100	0	100	**#N/A**
Belarus	362	9,679	4	2	98	0	100	0	100	**4**
Belgium	733	10,590	7	2	98	4	96	1	99	**6**
Belize	52	301	17	2	98	8	92	1	99	**15**
Benin	1,601	8,662	18	7	93	8	92	14	86	**14**
Bhutan	109	687	16	7	93	7	93	1	99	**14**
Bolivia	1,512	9,694	16	13	87	2	98	3	97	**13**
Bosnia and Herzego	182	3,773	5	2	98		100	0	100	**5**
Botswana	330	1,921	17	7	93	4	96	2	98	**15**
Brazil	17,812	191,972	9	13	87	18	82	5	95	**6**
Brunei Darussalam	45	392	12	7	93	1	99	0	100	**11**
Bulgaria	263	7,593	3	2	98	2	98	1	99	**3**
Burkina Faso	1,906	15,234	13	7	93	14	86	9	91	**9**
Burundi	1,603	8,074	20	2	98	37	63	10	90	**11**
Cambodia	2,341	14,562	16	7	93	10	90	9	91	**12**
Cameroon	3,201	19,088	17	7	93	17	83	13	87	**11**
Canada	2,305	33,259	7	2	98		100	0	100	**7**
Cape Verde	76	499	15	7	93	11	89	4	96	**12**
Central African Repu	608	4,339	14	20	80	25	75	20	80	**7**
Chad	1,496	10,914	14	7	93	23	77	24	76	**7**
Chile	1,679	16,804	10	13	87	2	98	0	100	**8**
China	105,951	1,337,411	8	7	93	0	100	0	100	**7**
Chinese Taipei (Taiw	#VALUE!	#N/A	#VALUE!	2	98	#N/A	100	0	100	**#N/A**
Colombia	5,286	45,012	12	13	87	3	97	2	98	**10**
Comoros	111	661	17	7	93	27	73	12	88	**10**
Congo, Dem. Rep.	9,973	#N/A	#N/A	7	93	17	83	0	100	**#N/A**
Congo, Rep.	628	#N/A	#N/A	7	93	22	78	0	100	**#N/A**
Costa Rica	535	4,519	12	13	87	7	93	0	100	**10**
Côte d'Ivoire	2,356	20,591	11	7	93	20	80	0	100	**9**
Croatia	191	4,423	4	2	98	0	100	0	100	**4**
Cuba	0	0	#DIV/0!	#N/A	#N/A	#N/A	100	0	100	**#N/A**
Cyprus	57	862	7	2	98	0	100	0	100	**6**
Czech Republic	460	10,319	4	2	98	1	99	0	100	**4**
Denmark	416	5,458	8	2	98		100	0	100	**7**
Djibouti	56	849	7	7	93	11	89	40	60	**3**
Dominica	8	#VALUE!	#VALUE!	13	87	4	96	1	99	**#N/A**
Dominican Republic	1,306	9,953	13	13	87	3	97	3	97	**11**
Ecuador	2,041	13,481	15	13	87	2	98	2	98	**13**
Egypt	9,988	81,527	12	7	93	4	96	1	99	**11**
El Salvador	994	6,134	16	13	87	6	94	2	98	**13**
Equatorial Guinea	81	659	12	7	93		100	0	100	**11**
Eritrea	314	4,927	6	7	93	14	86	12	88	**4**
Estonia	75	1,341	6	2	98	1	99	0	100	**5**
Ethiopia	12,742	80,713	16	7	93	5	95	17	83	**12**
Fiji	103	844	12	7	93	2	98	1	99	**11**
Finland	357	5,304	7	2	98	0	100	0	100	**7**
France	4,139	62,036	7	2	98		100	0	100	**7**
Gabon	281	1,448	19	7	93		100	0	100	**18**
Gambia, The	221	1,660	13	5	95	#N/A	100	9	91	**11**
Georgia	311	4,307	7	6	94	0	100	0	100	**7**
Germany	3,236	82,264	4	2	98	1	99	1	99	**4**
Ghana	3,625	23,351	16	7	93	7	93	4	96	**13**
Greece	639	11,137	6	2	98	1	99	0	100	**6**
Grenada	14	104	13	13	87	3	97	1	99	**11**

Table A2. Primary pupil volume indicators - continued

PRIMARY PUPIL VOLUME ESTIMATES										
Country	Most recent primary enrolment	Population ~2008	Pupils as % of population	Estimated pupil absenteeism (%)	Estimated pupil attendance (%)	Repetition rate (%)	Non-repeating pupils (%) (repeaters value=0)	Pupils who dropout illiterate (%)	Pupils who learn to read (%)	Effective primary pupils as % of population
Guatemala	2,501	13,686	18	13	87	11	89	5	95	**14**
Guinea	1,364	9,833	14	7	93	17	83	18	82	**9**
Guinea-Bissau	269	1,575	17	16	84		100	0	100	**14**
Guyana	107	763	14	18	82	1	99	0	100	**11**
Haiti	#VALUE!	9,876	#VALUE!	13	87		100	0	100	**#N/A**
Honduras	1,276	7,319	17	13	87	5	95	3	97	**14**
Hong Kong, China	390	6,982	6	7	93	#N/A	100	0	100	**5**
Hungary	394	10,012	4	2	98	2	98	0	100	**4**
Iceland	30	315	9	2	98		100	0	100	**9**
India	140,357	1,181,412	12	7	93	4	96	7	93	**10**
Indonesia	29,498	227,345	13	7	93	3	97	2	98	**11**
Iran	7,028	73,312	10	7	93	2	98	0	100	**9**
Iraq	4,430	30,096	15	7	93	8	92	8	92	**12**
Ireland	487	4,437	11	2	98	1	99	0	100	**11**
Israel	841	7,051	12	7	93	1	99	0	100	**11**
Italy	2,820	59,604	5	2	98	0	100	0	100	**5**
Jamaica	315	2,708	12	21	79	3	97	0	100	**9**
Japan	7,166	127,293	6	2	98		100	0	100	**6**
Jordan	817	6,136	13	7	93	1	99	0	100	**12**
Kazakhstan	951	15,521	6	8	92	0	100	0	100	**6**
Kenya	6,869	38,765	18	7	93	6	94	4	96	**15**
Kiribati	16	#VALUE!	#VALUE!	7	93		100	0	100	**#N/A**
Korea, Rep.	3,680	48,152	8	2	98	#N/A	100	0	100	**7**
Kuwait	209	2,919	7	7	93	1	99	0	100	**7**
Kyrgyzstan	400	5,414	7	10	90	#N/A	100	0	100	**7**
Laos	901	6,205	15	7	93	14	86	0	100	**12**
Latvia	117	2,259	5	2	98	3	97	0	100	**5**
Lebanon	464	4,194	11	7	93	9	91	1	99	**9**
Lesotho	401	2,049	20	7	93	19	81	12	88	**13**
Liberia	540	3,793	14	7	93		100	0	100	**13**
Libya	755	6,294	12	7	93		100	0	100	**11**
Lithuania	136	3,321	4	2	98	1	99	0	100	**4**
Luxembourg	36	481	7	2	98	4	96	0	100	**7**
Macao, China	27	526	5	2	98	5	95	0	100	**5**
Macedonia	101	2,041	5	5	95	0	100	0	100	**5**
Madagascar	4,020	19,111	21	7	93	20	80	14	86	**14**
Malawi	3,198	14,846	22	7	93	19	81	21	79	**13**
Malaysia	3,104	27,014	11	7	93		100	0	100	**11**
Maldives	47	305	15	7	93	3	97	1	99	**14**
Mali	1,926	12,706	15	7	93	15	85	9	91	**11**
Malta	28	407	7	7	93	2	98	7	93	**6**
Mauritania	513	3,215	16	7	93	6	94	20	80	**11**
Mauritius	118	1,280	9	7	93	18	82	0	100	**7**
Mexico	14,699	108,555	14	13	87	4	96	1	99	**11**
Micronesia	19	#N/A	#N/A	7	93		100	0	100	**#N/A**
Moldova	152	#N/A	#N/A	2	98	0	100	0	100	**#N/A**
Mongolia	240	2,641	9	1	99	0	100	1	99	**9**
Montenegro	#VALUE!	622	#VALUE!	1	99		100	0	100	**#N/A**
Morocco	3,879	31,606	12	7	93	11	89	5	95	**10**
Mozambique	4,904	22,383	22	7	93	6	94	16	84	**16**
Myanmar	5,110	49,563	10	7	93	0	100	7	93	**9**
Namibia	407	2,130	19	7	93	18	82	5	95	**14**
Nepal	4,782	#VALUE!	#VALUE!	7	93	14	86	10	90	**#N/A**
Netherlands	1,286	16,528	8	2	98		100	0	100	**8**
New Zealand	348	4,230	8	2	98		100	0	100	**8**
Nicaragua	944	5,667	17	13	87	9	91	7	93	**12**
Niger	1,554	14,704	11	7	93	7	93	12	88	**8**
Nigeria	21,632	151,212	14	5	95		100	9	91	**12**
Norway	430	4,767	9	2	98		100	0	100	**9**
Oman	271	2,785	10	7	93	3	97	0	100	**9**
Pakistan	18,176	176,952	10	7	93	4	96	6	94	**9**
Palestinian Autonom	390	4,147	9	7	93	#N/A	100	0	100	**9**

Table A2. Primary pupil volume indicators – continued

PRIMARY PUPIL VOLUME ESTIMATES										
Country	Most recent primary enrolment	Population ~2008	Pupils as % of population	Estimated pupil absenteeism (%)	Estimated pupil attendance (%)	Repetition rate (%)	Non-repeating pupils (%) (repeaters value=0)	Pupils who dropout illiterate (%)	Pupils who learn to read (%)	Effective primary pupils as % of population
Panama	445	3,399	13	13	87	5	95	2	98	**11**
Papua New Guinea	532	6,577	8	7	93		100	0	100	**8**
Paraguay	894	6,238	14	13	87	4	96	2	98	**12**
Peru	3,855	28,837	13	13	87	7	93	4	96	**10**
Philippines	13,411	90,348	15	7	93	5	95	4	96	**13**
Poland	2,485	38,104	7	2	98	1	99	0	100	**6**
Portugal	754	10,677	7	2	98		100	0	100	**7**
Qatar	78	1,281	6	7	93	1	99	0	100	**6**
Romania	865	21,361	4	2	98	2	98	1	99	**4**
Russian Federation	4,969	141,394	4	2	98	0	100	0	100	**3**
Rwanda	2,190	9,721	23	7	93	17	83	21	79	**14**
Samoa	30	179	17	7	93		100	2	98	**15**
Sao Tome and Princ	34	160	21	7	93	19	81	10	90	**14**
Saudi Arabia	3,211	25,201	13	7	93	3	97	2	98	**11**
Senegal	1,618	12,211	13	7	93	9	91	10	90	**10**
Serbia	290	9,839	3	2	98	1	99	0	100	**3**
Seychelles	9	87	10	7	93		100	0	100	**9**
Sierra Leone	1,322	5,560	24	4	96		100	0	100	**23**
Singapore	300	#VALUE!	#VALUE!	2	98	0	100	0	100	**#N/A**
Slovakia	225	5,400	4	2	98	#N/A	100	0	100	**4**
Slovenia	107	2,015	5	2	98	1	99	0	100	**5**
Solomon Islands	83	511	16	7	93		100	0	100	**15**
South Africa	7,312	49,668	15	7	93	6	94	5	95	**12**
Spain	2,625	44,486	6	2	98	5	95	0	100	**5**
Sri Lanka	1,631	20,061	8	7	93	1	99	0	100	**7**
St. Kitts and Nevis	6	51	13	#N/A	#N/A	2	98	0	100	**#N/A**
St. Lucia	21	170	12	13	87	2	98	0	100	**10**
St. Vincent and the G	15	109	14	13	87	7	93	0	100	**11**
Sudan	4,744	41,348	11	7	93	4	96	1	99	**10**
Suriname	70	515	14	13	87	18	82	0	100	**10**
Swaziland	233	1,168	20	7	93	18	82	7	93	**14**
Sweden	585	9,205	6	2	98		100	0	100	**6**
Switzerland	505	7,541	7	2	98	2	98	0	100	**6**
Syria	2,356	21,227	11	8	92	7	93	0	100	**9**
Tajikistan	692	6,836	10	7	93	0	100	0	100	**9**
Tanzania	8,602	42,484	20	7	93	4	96	0	100	**18**
Thailand	5,371	67,386	8	7	93	9	91	2	98	**7**
Timor-Leste	201	1,098	18	7	93	13	87	3	97	**15**
Togo	1,164	6,459	18	3	97	23	77	20	80	**11**
Tonga	17	104	16	7	93	25	75	0	100	**11**
Trinidad and Tobago	131	1,333	10	11	89	7	93	0	100	**8**
Tunisia	1,036	10,169	10	7	93	8	92	1	99	**9**
Turkey	6,760	73,914	9	7	93	2	98	1	99	**8**
Turkmenistan	#VALUE!	5,044	#VALUE!	7	93		100	0	100	**#N/A**
Uganda	7,964	31,657	25	7	93	11	89	17	83	**17**
Ukraine	1,573	45,992	3	2	98	0	100	0	100	**3**
United Arab Emirate	284	4,485	6	7	93	2	98	2	98	**6**
United Kingdom	4,465	61,231	7	2	98		100	0	100	**7**
United States	24,677	311,666	8	2	98		100	0	100	**8**
Uruguay	359	3,349	11	13	87	7	93	1	99	**9**
Vanuatu	38	234	16	4	96	13	87	5	95	**13**
Venezuela	3,439	28,121	12	13	87	3	97	2	98	**10**
Vietnam	6,872	#VALUE!	#VALUE!	4	96	1	99	0	100	**#N/A**
Yemen	3,282	22,917	14	7	93	5	95	9	91	**12**
Zambia	2,909	12,620	23	7	93	6	94	10	90	**18**
Zimbabwe	2,446	12,463	20	7	93		100	0	100	**18**

Sources: Pupil enrolment, population, repetition rate from UNESCO Institute for Statistics (UIS);

Absenteeism estimates from UNICEF MICS 2006 surveys; Pupils who drop out illiterate from DHS surveys 2005-9 and by grade dropout rates from EPDC.

Table A3. Primary pupil quality indicators

PRIMARY PUPIL QUALITY ESTIMATES

Country	Unadj. score, using with HHC (10/12/10)	Actual PISA (yes=1)	Learning index 500=1	10/12/10 score adj. for adult mean schooling	Ratio of unadj. to adj. score	HHC	Country group by HHC	Group score I by HHC	Group score 1 adj.	Secon dary GER	Country group by SecGER	Group score II by SecGER	Group score II adj.	Country group by score	Group score III by score	Group score III adjusted for mean schooling
Albania	438	1	0.88	382	0.87	3281	3	0.80	0.69	75	3	0.84	0.73	2	0.87	**0.76**
Algeria	354	1	0.71	308	0.87	1845	4	0.73	0.63	79	3	0.84	0.73	4	0.73	**0.64**
Angola	351	1	0.70	314	0.89	1128	4	0.73	0.65	16	5	0.66	0.59	4	0.73	**0.65**
Antigua and Barbuda	420	1	0.84	357	0.85	0	5	0.67	0.57	105	1	0.97	0.82	2	0.87	**0.74**
Argentina	391	0	0.78	343	0.88	6734	2	0.87	0.76	87	3	0.84	0.73	3	0.80	**0.70**
Armenia	455	1	0.91	399	0.88	3466	3	0.80	0.70	88	3	0.84	0.73	2	0.87	**0.77**
Australia	527	0	1.05	463	0.88	21048	1	0.98	0.87	152	1	0.97	0.85	1	1.01	**0.89**
Austria	511	0	1.02	447	0.88	21151	1	0.98	0.86	99	2	0.92	0.80	1	1.01	**0.89**
Azerbaijan	382	0	0.76	326	0.85	2241	3	0.80	0.68	87	3	0.84	0.71	4	0.73	**0.62**
Bahamas	458	1	0.92	395	0.86	16642	1	0.98	0.85	88	3	0.84	0.72	2	0.87	**0.75**
Bahrain	431	1	0.86	379	0.88	9091	2	0.87	0.77	97	2	0.92	0.81	2	0.87	**0.77**
Bangladesh	335	1	0.67	307	0.92	849	5	0.67	0.61	45	4	0.74	0.68	5	0.62	**0.57**
Barbados	465	1	0.93	402	0.87	16132	1	0.98	0.85	100	1	0.97	0.84	2	0.87	**0.76**
Belarus	462	1	0.92	408	0.88	5514	2	0.87	0.77	92	2	0.92	0.81	2	0.87	**0.77**
Belgium	510	0	1.02	450	0.88	17828	1	0.98	0.87	132	1	0.97	0.85	1	1.01	**0.89**
Belize	404	1	0.81	362	0.89	6414	2	0.87	0.78	73	4	0.74	0.66	3	0.80	**0.72**
Benin	315	1	0.63	297	0.94	894	5	0.67	0.63	28	5	0.66	0.62	5	0.62	**0.58**
Bhutan	356	1	0.71	321	0.90	1369	4	0.73	0.66	48	4	0.74	0.67	4	0.73	**0.66**
Bolivia	386	1	0.77	332	0.86	2404	3	0.80	0.68	71	4	0.74	0.64	4	0.73	**0.63**
Bosnia and Herzegovir	431	1	0.86	377	0.87	4831	3	0.80	0.70	89	3	0.84	0.73	2	0.87	**0.76**
Botswana	367	1	0.73	327	0.89	2164	4	0.73	0.65	77	3	0.84	0.74	4	0.73	**0.65**
Brazil	390	0	0.78	356	0.91	4629	3	0.80	0.73	91	2	0.92	0.84	3	0.80	**0.73**
Brunei Darussalam	424	1	0.85	377	0.89	10217	1	0.98	0.88	93	2	0.92	0.82	2	0.87	**0.78**
Bulgaria	434	0	0.87	375	0.86	6132	2	0.87	0.75	100	1	0.97	0.83	2	0.87	**0.76**
Burkina Faso	299	1	0.60	289	0.97	680	5	0.67	0.64	14	5	0.66	0.64	5	0.62	**0.60**
Burundi	278	1	0.56	256	0.92	227	5	0.67	0.61	14	5	0.66	0.61	5	0.62	**0.57**
Cambodia	386	1	0.77	362	0.94	1171	4	0.73	0.68	28	5	0.66	0.62	3	0.80	**0.75**
Cameroon	343	1	0.69	310	0.90	1363	4	0.73	0.66	29	5	0.66	0.59	4	0.73	**0.66**
Canada	534	0	1.07	462	0.86	20834	1	0.98	0.85	104	1	0.97	0.83	1	1.01	**0.88**
Cape Verde	362	1	0.72	324	0.90	2176	4	0.73	0.65	73	4	0.74	0.66	4	0.73	**0.66**
Central African Republ	312	1	0.62	289	0.93	487	5	0.67	0.62	12	5	0.66	0.61	5	0.62	**0.57**
Chad	273	1	0.55	259	0.95	226	5	0.67	0.63	15	5	0.66	0.63	5	0.62	**0.59**
Chile	438	0	0.88	379	0.87	6473	2	0.87	0.75	88	3	0.84	0.72	2	0.87	**0.76**
China	431	1	0.86	390	0.91	1893	4	0.73	0.66	67	4	0.74	0.67	2	0.87	**0.79**
Chinese Taipei (Taiwa	532	0	1.06	473	0.89	15647	1	0.98	0.87	96	2	0.92	0.81	1	1.01	**0.90**
Colombia	388	0	0.78	345	0.89	3749	3	0.80	0.71	79	3	0.84	0.74	3	0.80	**0.71**
Comoros	319	1	0.64	294	0.92	837	5	0.67	0.61	39	5	0.66	0.61	5	0.62	**0.57**
Congo, Dem. Rep.	275	1	0.55	252	0.92	175	5	0.67	0.61	30	5	0.66	0.60	5	0.62	**0.57**
Congo, Rep.	325	1	0.65	288	0.88	706	5	0.67	0.59	39	5	0.66	0.58	5	0.62	**0.55**
Costa Rica	408	1	0.82	364	0.89	6692	2	0.87	0.78	80	3	0.84	0.75	3	0.80	**0.72**
Côte d'Ivoire	336	1	0.67	316	0.94	1147	4	0.73	0.69	25	5	0.66	0.62	5	0.62	**0.58**
Croatia	493	0	0.99	448	0.91	7318	2	0.87	0.79	90	3	0.84	0.76	1	1.01	**0.92**
Cuba	448	1	0.90	395	0.88	0	5	0.67	0.59	91	2	0.92	0.81	2	0.87	**0.77**
Cyprus	486	1	0.97	433	0.89	16253	1	0.98	0.88	97	2	0.92	0.82	1	1.01	**0.90**
Czech Republic	513	0	1.03	458	0.89	10563	1	0.98	0.88	94	2	0.92	0.82	1	1.01	**0.91**
Denmark	496	0	0.99	429	0.86	16746	1	0.98	0.85	125	1	0.97	0.83	1	1.01	**0.88**
Djibouti	358	1	0.72	324	0.91	1768	4	0.73	0.66	21	5	0.66	0.60	4	0.73	**0.66**
Dominica	398	1	0.80	355	0.89	3564	3	0.80	0.71	104	1	0.97	0.86	3	0.80	**0.71**
Dominican Republic	393	1	0.79	347	0.88	6555	2	0.87	0.77	65	4	0.74	0.65	3	0.80	**0.71**
Ecuador	382	1	0.76	337	0.88	3714	3	0.80	0.70	62	4	0.74	0.65	4	0.73	**0.65**
Egypt	387	1	0.77	351	0.91	2840	3	0.80	0.72	78	3	0.84	0.76	3	0.80	**0.73**
El Salvador	368	1	0.74	328	0.89	4799	3	0.80	0.71	53	4	0.74	0.66	4	0.73	**0.65**
Equatorial Guinea	344	1	0.69		0.00	1370	4	0.73	0.00	27	5	0.66	0.00	4	0.73	**0.00**
Eritrea	302	1	0.60	278	0.92	502	5	0.67	0.61	30	5	0.66	0.61	5	0.62	**0.57**
Estonia	531	0	1.06	470	0.88	9321	2	0.87	0.77	97	2	0.92	0.81	1	1.01	**0.90**
Ethiopia	282	1	0.56	273	0.97	404	5	0.67	0.65	24	5	0.66	0.64	5	0.62	**0.60**
Fiji	461	1	0.92	390	0.85	3150	3	0.80	0.67	83	3	0.84	0.71	2	0.87	**0.74**
Finland	563	0	1.13	491	0.87	16109	1	0.98	0.86	118	1	0.97	0.84	1	1.01	**0.88**
France	495	0	0.99	445	0.90	18175	1	0.98	0.88	110	1	0.97	0.87	1	1.01	**0.91**
Gabon	386	1	0.77	344	0.89	2857	3	0.80	0.71	51	4	0.74	0.66	4	0.73	**0.65**
Gambia, The	312	1	0.62	289	0.93	756	5	0.67	0.62	44	4	0.74	0.69	5	0.62	**0.57**
Georgia	389	1	0.78	333	0.86	2762	3	0.80	0.68	83	3	0.84	0.72	3	0.80	**0.69**
Germany	516	0	1.03	444	0.86	19330	1	0.98	0.85	100	1	0.97	0.83	1	1.01	**0.87**
Ghana	294	1	0.59	255	0.87	825	5	0.67	0.58	45	4	0.74	0.64	5	0.62	**0.54**
Greece	473	0	0.95	423	0.89	18886	1	0.98	0.88	97	2	0.92	0.82	2	0.87	**0.78**
Grenada	390	1	0.78	379	0.97	8182	2	0.87	0.85	109	1	0.97	0.94	3	0.80	**0.78**
Guatemala	374	1	0.75	346	0.93	4820	3	0.80	0.74	47	4	0.74	0.69	4	0.73	**0.68**
Guinea	300	1	0.60	287	0.96	757	5	0.67	0.64	27	5	0.66	0.63	5	0.62	**0.59**
Guinea-Bissau	299	1	0.60	277	0.92	339	5	0.67	0.62	20	5	0.66	0.61	5	0.62	**0.57**
Guyana	376	1	0.75	329	0.87	1566	4	0.73	0.64	86	3	0.84	0.73	4	0.73	**0.64**
Haiti	338	1	0.68	312	0.92	1411	4	0.73	0.67	10	5	0.66	0.61	5	0.62	**0.57**
Honduras	353	1	0.71	325	0.92	2361	3	0.80	0.73	64	4	0.74	0.68	4	0.73	**0.67**
Hong Kong, China	542	0	1.08	480	0.88	18281	1	0.98	0.87	82	3	0.84	0.74	1	1.01	**0.90**
Hungary	504	0	1.01	446	0.88	10175	2	0.87	0.77	98	2	0.92	0.81	1	1.01	**0.90**
Iceland	491	0	0.98	432	0.88	21782	1	0.98	0.87	110	1	0.97	0.85	1	1.01	**0.89**
India	351	1	0.70	325	0.93	1373	4	0.73	0.67	67	4	0.74	0.69	4	0.73	**0.68**
Indonesia	393	0	0.79	354	0.90	2099	4	0.73	0.66	64	4	0.74	0.67	3	0.80	**0.72**
Iran	409	1	0.82	372	0.91	4333	3	0.80	0.72	79	3	0.84	0.76	3	0.80	**0.73**
Iraq	351	1	0.70	317	0.90	1506	4	0.73	0.66	38	5	0.66	0.60	4	0.73	**0.66**
Ireland	508	0	1.02	450	0.88	18468	1	0.98	0.87	110	1	0.97	0.85	1	1.01	**0.90**
Israel	454	0	0.91	407	0.90	13175	1	0.98	0.88	93	2	0.92	0.82	2	0.87	**0.78**
Italy	475	0	0.95	425	0.89	17577	1	0.98	0.88	98	2	0.92	0.82	1	1.01	**0.91**
Jamaica	406	1	0.81	362	0.89	5539	2	0.87	0.78	87	3	0.84	0.75	3	0.80	**0.72**

Table A3. Primary pupil quality indicators - continued

PRIMARY PUPIL QUALITY ESTIMATES																
Country	Unadj. score, using with HHC (10/12/10)	Actual PISA (yes=1)	Learning index 500=1	10/12/10 score adj. for adult mean schooling	Ratio of unadj. to adj. score	HHC	Country group by HHC	Group score I by HHC	Group score 1 adj.	Secondary GER	Country group by SecGER	Group score II by SecGER	Group score II adj.	Country group by score	Group score III by score	Group score III adjusted for mean schooling
Japan	531	0	1.06	469	0.88	17949	1	0.98	0.87	101	1	0.97	0.85	1	1.01	**0.89**
Jordan	422	0	0.84	378	0.90	3304	3	0.80	0.71	86	3	0.84	0.75	2	0.87	**0.78**
Kazakhstan	491	1	0.98	431	0.88	4814	3	0.80	0.70	94	2	0.92	0.81	1	1.01	**0.89**
Kenya	400	1	0.80	361	0.90	963	4	0.73	0.66	47	4	0.74	0.67	3	0.80	**0.72**
Kiribati	419	1	0.84	374	0.89	1794	4	0.73	0.65	90	3	0.84	0.75	2	0.87	**0.78**
Korea, Rep.	522	0	1.04	461	0.88	10889	1	0.98	0.87	95	2	0.92	0.81	1	1.01	**0.90**
Kuwait	427	1	0.85	381	0.89	12864	1	0.98	0.88	91	2	0.92	0.82	2	0.87	**0.78**
Kyrgyzstan	322	0	0.64	265	0.82	1567	4	0.73	0.60	95	2	0.92	0.76	5	0.62	**0.51**
Laos	380	1	0.76	353	0.93	937	4	0.73	0.68	42	4	0.74	0.69	4	0.73	**0.68**
Latvia	490	0	0.98	426	0.87	8323	2	0.87	0.76	97	2	0.92	0.80	1	1.01	**0.88**
Lebanon	447	1	0.89	400	0.90	7288	2	0.87	0.78	82	3	0.84	0.75	2	0.87	**0.78**
Lesotho	324	1	0.65	286	0.88	1324	4	0.73	0.64	34	5	0.66	0.58	5	0.62	**0.55**
Liberia	288	1	0.58	265	0.92	362	5	0.67	0.61	32	5	0.66	0.61	5	0.62	**0.57**
Libya	398	1	0.80	352	0.88	4982	2	0.87	0.77	103	1	0.97	0.85	3	0.80	**0.71**
Lithuania	488	0	0.98	429	0.88	9112	2	0.87	0.77	101	1	0.97	0.85	1	1.01	**0.89**
Luxembourg	486	0	0.97	432	0.89	28919	1	0.98	0.87	96	2	0.92	0.81	1	1.01	**0.90**
Macao, China	511	0	1.02	464	0.91	9253	2	0.87	0.79	91	2	0.92	0.83	1	1.01	**0.92**
Macedonia	408	1	0.82	366	0.90	5379	2	0.87	0.78	84	3	0.84	0.75	3	0.80	**0.72**
Madagascar	314	1	0.63	284	0.90	670	5	0.67	0.60	26	5	0.66	0.60	5	0.62	**0.56**
Malawi	310	1	0.62	286	0.92	433	5	0.67	0.61	28	5	0.66	0.61	5	0.62	**0.57**
Malaysia	425	1	0.85	376	0.89	5082	2	0.87	0.77	68	4	0.74	0.66	2	0.87	**0.77**
Maldives	344	1	0.69	310	0.90	990	4	0.73	0.66	69	4	0.74	0.67	4	0.73	**0.66**
Mali	298	1	0.60	288	0.97	656	5	0.67	0.65	18	5	0.66	0.64	5	0.62	**0.60**
Malta	485	1	0.97	432	0.89	14000	1	0.98	0.88	95	2	0.92	0.82	1	1.01	**0.90**
Mauritania	316	1	0.63	298	0.94	5914	2	0.87	0.82	22	5	0.66	0.62	5	0.62	**0.58**
Mauritius	404	1	0.81	364	0.90	922	4	0.73	0.66	84	3	0.84	0.75	3	0.80	**0.72**
Mexico	410	0	0.82	367	0.89	5868	2	0.87	0.78	81	3	0.84	0.75	3	0.80	**0.72**
Micronesia	425	1	0.85	380	0.89	8205	2	0.87	0.78	85	3	0.84	0.75	2	0.87	**0.78**
Moldova	415	1	0.83	359	0.87	2286	3	0.80	0.69	83	3	0.84	0.72	3	0.80	**0.69**
Mongolia	421	1	0.84	367	0.87	2166	4	0.73	0.63	88	3	0.84	0.73	2	0.87	**0.76**
Montenegro	412	0	0.82	358	0.87	1169	4	0.73	0.63	100	1	0.97	0.84	3	0.80	**0.70**
Morocco	332	1	0.66	310	0.93	4397	3	0.80	0.74	46	4	0.74	0.69	5	0.62	**0.58**
Mozambique	377	1	0.75	362	0.96	1701	4	0.73	0.70	12	5	0.66	0.63	4	0.73	**0.70**
Myanmar	403	1	0.81	377	0.93	570	5	0.67	0.62	44	4	0.74	0.69	3	0.80	**0.75**
Namibia	317	1	0.63	279	0.88	2220	3	0.80	0.70	64	4	0.74	0.65	5	0.62	**0.54**
Nepal	322	1	0.64	306	0.95	903	4	0.73	0.69	42	4	0.74	0.70	5	0.62	**0.59**
Netherlands	525	0	1.05	464	0.88	17985	1	0.98	0.87	120	1	0.97	0.85	1	1.01	**0.90**
New Zealand	530	0	1.06	458	0.86	16649	1	0.98	0.85	116	1	0.97	0.83	1	1.01	**0.87**
Nicaragua	419	1	0.84	374	0.89	1625	4	0.73	0.65	63	4	0.74	0.66	2	0.87	**0.78**
Niger	283	1	0.57	275	0.97	391	5	0.67	0.65	9	5	0.66	0.64	5	0.62	**0.60**
Nigeria	339	1	0.68	309	0.91	989	4	0.73	0.66	31	5	0.66	0.60	5	0.62	**0.56**
Norway	487	0	0.97	415	0.85	20582	1	0.98	0.84	113	1	0.97	0.82	1	1.01	**0.86**
Oman	412	1	0.82	365	0.89	5924	2	0.87	0.77	85	3	0.84	0.74	3	0.80	**0.71**
Pakistan	344	1	0.69	322	0.94	1951	4	0.73	0.68	30	5	0.66	0.62	4	0.73	**0.68**
Palestinian Autonomou	355	1	0.71	321	0.90		5	0.67	0.60	89	3	0.84	0.76	4	0.73	**0.66**
Panama	391	1	0.78	343	0.88	4690	3	0.80	0.70	69	4	0.74	0.65	3	0.80	**0.70**
Papua New Guinea	395	1	0.79	350	0.89	876	5	0.67	0.59	20	5	0.66	0.58	3	0.80	**0.71**
Paraguay	396	1	0.79	353	0.89	2708	3	0.80	0.71	65	4	0.74	0.66	3	0.80	**0.71**
Peru	407	1	0.81	361	0.89	3826	3	0.80	0.71	91	2	0.92	0.81	3	0.80	**0.71**
Philippines	427	1	0.85	377	0.88	2040	4	0.73	0.64	82	3	0.84	0.74	2	0.87	**0.77**
Poland	498	0	1.00	442	0.89	8913	2	0.87	0.77	101	1	0.97	0.86	1	1.01	**0.90**
Portugal	474	0	0.95	434	0.92	13419	1	0.98	0.90	103	1	0.97	0.88	1	1.01	**0.93**
Qatar	349	0	0.70	303	0.87	10942	1	0.98	0.85	92	2	0.92	0.80	4	0.73	**0.63**
Romania	418	0	0.84	366	0.88	5928	2	0.87	0.76	85	3	0.84	0.73	2	0.87	**0.77**
Russian Federation	479	0	0.96	427	0.89	6599	2	0.87	0.78	86	3	0.84	0.74	1	1.01	**0.90**
Rwanda	306	1	0.61	283	0.92	548	5	0.67	0.62	15	5	0.66	0.61	5	0.62	**0.57**
Samoa	455	1	0.91	410	0.90	5022	2	0.87	0.79	78	3	0.84	0.75	2	0.87	**0.79**
Sao Tome and Princip	340	1	0.68	302	0.89	889	5	0.67	0.59	44	4	0.74	0.66	5	0.62	**0.55**
Saudi Arabia	380	1	0.76	339	0.89	4785	3	0.80	0.71	93	2	0.92	0.82	4	0.73	**0.65**
Senegal	324	1	0.65	302	0.93	1005	4	0.73	0.68	15	5	0.66	0.61	5	0.62	**0.58**
Serbia	436	0	0.87	382	0.88	5542	2	0.87	0.76	87	3	0.84	0.73	2	0.87	**0.77**
Seychelles	409	1	0.82	369	0.90	10350	1	0.98	0.89	105	1	0.97	0.87	3	0.80	**0.72**
Sierra Leone	310	1	0.62	287	0.93	523	5	0.67	0.62	31	5	0.66	0.61	5	0.62	**0.57**
Singapore	550	1	1.10	502	0.91	13811	1	0.98	0.90	100	1	0.97	0.88	1	1.01	**0.93**
Slovakia	488	0	0.98	434	0.89	9536	2	0.87	0.77	91	2	0.92	0.82	1	1.01	**0.90**
Slovenia	519	0	1.04	458	0.88	12978	1	0.98	0.87	100	1	0.97	0.85	1	1.01	**0.89**
Solomon Islands	380	1	0.76	335	0.88	504	5	0.67	0.59	28	5	0.66	0.58	4	0.73	**0.64**
South Africa	357	1	0.71	315	0.88	4906	2	0.87	0.77	88	3	0.84	0.74	4	0.73	**0.64**
Spain	488	0	0.98	446	0.91	17944	1	0.98	0.90	115	1	0.97	0.88	1	1.01	**0.93**
Sri Lanka	380	1	0.76	326	0.86	2309	3	0.80	0.68	86	3	0.84	0.72	4	0.73	**0.63**
St. Kitts and Nevis	414	1	0.83	371	0.90	6178	2	0.87	0.78	98	2	0.92	0.82	3	0.80	**0.72**
St. Lucia	415	1	0.83	372	0.90	7162	2	0.87	0.78	80	3	0.84	0.75	3	0.80	**0.72**
St. Vincent and the Gr	399	1	0.80	355	0.89	3342	3	0.80	0.71	90	2	0.92	0.82	3	0.80	**0.71**
Sudan	352	1	0.70	315	0.89	1284	4	0.73	0.65	30	5	0.66	0.59	4	0.73	**0.65**
Suriname	373	1	0.75	330	0.88	1792	4	0.73	0.64	73	4	0.74	0.66	4	0.73	**0.65**
Swaziland	377	1	0.75	310	0.82	2683	3	0.80	0.65	46	4	0.74	0.61	4	0.73	**0.60**
Sweden	503	0	1.01	442	0.88	17092	1	0.98	0.86	124	1	0.97	0.85	1	1.01	**0.89**
Switzerland	512	0	1.02	445	0.87	21687	1	0.98	0.86	94	2	0.92	0.80	1	1.01	**0.88**
Syria	371	1	0.74	340	0.92	2439	3	0.80	0.73	59	4	0.74	0.68	4	0.73	**0.67**
Tajikistan	396	1	0.79	340	0.86	1149	4	0.73	0.63	81	3	0.84	0.72	3	0.80	**0.69**
Tanzania	385	1	0.77	361	0.94	669	5	0.67	0.63	13	5	0.66	0.62	4	0.73	**0.69**
Thailand	421	0	0.84	381	0.90	4030	3	0.80	0.72	70	4	0.74	0.67	2	0.87	**0.79**
Timor-Leste	378	1	0.76	351	0.93	722	5	0.67	0.62	46	4	0.74	0.69	4	0.73	**0.68**

Table A3. Primary pupil quality indicators - continued

PRIMARY PUPIL QUALITY ESTIMATES																
Country	**Unadj. score, using with HHC (10/12/10)**	**Actual PISA (yes=1)**	**Learning index 500=1**	**10/12/10 score adj. for adult mean schooling**	**Ratio of unadj. to adj. score**	**HHC**	**Country group by HHC**	**Group score I by HHC**	**Group score 1 adj.**	**Secon dary GER**	**Country group by SecGER**	**Group score II by SecGER**	**Group score II adj.**	**Country group by score**	**Group score III by score**	**Group score III adjusted for mean schooling**
Togo	370	1	0.74	347	0.94	738	5	0.67	0.62	10	5	0.66	0.62	4	0.73	**0.69**
Tonga	455	1	0.91	410	0.90	6202	2	0.87	0.79	104	1	0.97	0.87	2	0.87	**0.79**
Trinidad and Tobago	425	1	0.85	362	0.85	10472	1	0.98	0.84	79	3	0.84	0.71	2	0.87	**0.75**
Tunisia	386	0	0.77	352	0.91	3580	3	0.80	0.73	82	3	0.84	0.76	4	0.73	**0.67**
Turkey	424	0	0.85	393	0.93	5306	2	0.87	0.81	80	3	0.84	0.78	2	0.87	**0.81**
Turkmenistan	434	1	0.87	376	0.87	4228	3	0.80	0.69	95	2	0.92	0.79	2	0.87	**0.76**
Uganda	356	1	0.71	324	0.91	666	5	0.67	0.61	23	5	0.66	0.60	4	0.73	**0.66**
Ukraine	438	1	0.88	385	0.88	3905	3	0.80	0.70	96	2	0.92	0.81	2	0.87	**0.77**
United Arab Emirates	487	1	0.97	440	0.90	22179	1	0.98	0.89	82	3	0.84	0.76	1	1.01	**0.92**
United Kingdom	515	0	1.03	460	0.89	22564	1	0.98	0.88	102	1	0.97	0.86	1	1.01	**0.90**
United States	489	0	0.98	423	0.87	30458	1	0.98	0.85	94	2	0.92	0.79	1	1.01	**0.88**
Uruguay	428	0	0.86	383	0.89	6117	2	0.87	0.78	102	1	0.97	0.86	2	0.87	**0.78**
Vanuatu	445	1	0.89	400	0.90	852	5	0.67	0.60	37	5	0.66	0.59	2	0.87	**0.79**
Venezuela	402	1	0.80	358	0.89	3257	3	0.80	0.71	72	4	0.74	0.66	3	0.80	**0.71**
Vietnam	404	1	0.81	373	0.92	4827	3	0.80	0.73	67	4	0.74	0.68	3	0.80	**0.74**
Yemen	284	1	0.57	250	0.88	1229	4	0.73	0.64	45	4	0.74	0.65	5	0.62	**0.55**
Zambia	314	1	0.63	277	0.88	1321	4	0.73	0.64	34	5	0.66	0.58	5	0.62	**0.55**
Zimbabwe	356	1	0.71	312	0.88	774	5	0.67	0.58	41	4	0.74	0.65	4	0.73	**0.64**

*** Alternative Groups for Group Scores**

A. HHC cutoffs - 1: 10200, 2: 4900, 3: 2200, 4: 900, 5: 0, Score by income group - 1: 0.98, 2: 0.87, 3: 0.80, 4: 0.73, 4: 0.67

B. SecGER quintile cutoffs - 1: 100, 2: 90, 3: 75, 4: 40, 5: 0, Score by SecGER - 1: 0.97, 2: 0.92, 3: 0.84, 4: 0.74, 5: 0.66

C. Score quintile cutoffs - 1: 474, 2: 418, 3: 386, 4: 340, 4: below 340, Group average learning index - 1: 1.01, 2: 0.87, 3: 0.80, 4: 0.73, 5: 0.62

Table A4 – Summary indices of primary pupil volume and quality

FINAL PRIMARY PUPIL ESTIMATES			
Country	Effective primary pupils as % of population	Group score III adjusted for mean schooling	Effective primary pupils adjusted for learning index
Albania	8	0.76	**6**
Algeria	10	0.64	**7**
Angola	#N/A	0.65	**#N/A**
Antigua and Barbuda	#N/A	0.74	**#N/A**
Argentina	10	0.70	**7**
Armenia	4	0.77	**3**
Australia	9	0.89	**8**
Austria	4	0.89	**4**
Azerbaijan	#N/A	0.62	**#N/A**
Bahamas	11	0.75	**8**
Bahrain	10	0.77	**7**
Bangladesh	6	0.57	**4**
Barbados	#N/A	0.76	**#N/A**
Belarus	4	0.77	**3**
Belgium	6	0.89	**6**
Belize	15	0.72	**11**
Benin	14	0.58	**8**
Bhutan	14	0.66	**9**
Bolivia	13	0.63	**8**
Bosnia and Herzegovina	5	0.76	**4**
Botswana	15	0.65	**10**
Brazil	6	0.73	**5**
Brunei Darussalam	11	0.78	**8**
Bulgaria	3	0.76	**2**
Burkina Faso	9	0.60	**6**
Burundi	11	0.57	**6**
Cambodia	12	0.75	**9**
Cameroon	11	0.66	**7**
Canada	7	0.88	**6**
Cape Verde	12	0.66	**8**
Central African Republic	7	0.57	**4**
Chad	7	0.59	**4**
Chile	8	0.76	**6**
China	7	0.79	**6**
Chinese Taipei (Taiwan)	#N/A	0.90	**#N/A**
Colombia	10	0.71	**7**
Comoros	10	0.57	**6**
Congo, Dem. Rep.	#N/A	0.57	**#N/A**
Congo, Rep.	#N/A	0.55	**#N/A**
Costa Rica	10	0.72	**7**
Côte d'Ivoire	9	0.58	**5**
Croatia	4	0.92	**4**
Cuba	#N/A	0.77	**#N/A**
Cyprus	6	0.90	**6**
Czech Republic	4	0.91	**4**
Denmark	7	0.88	**7**
Djibouti	3	0.66	**2**
Dominica	#N/A	0.71	**#N/A**
Dominican Republic	11	0.71	**8**
Ecuador	13	0.65	**8**
Egypt	11	0.73	**8**
El Salvador	13	0.65	**8**
Equatorial Guinea	11	0.00	**0**
Eritrea	4	0.57	**3**
Estonia	5	0.90	**5**
Ethiopia	12	0.60	**7**
Fiji	11	0.74	**8**
Finland	7	0.88	**6**
France	7	0.91	**6**
Gabon	18	0.65	**12**
Gambia, The	11	0.57	**7**
Georgia	7	0.69	**5**
Germany	4	0.87	**3**
Ghana	13	0.54	**7**
Greece	6	0.78	**4**
Grenada	11	0.78	**9**

Table A4 – Summary indices of primary pupil volume and quality – continued

FINAL PRIMARY PUPIL ESTIMATES			
Country	Effective primary pupils as % of population	Group score III adjusted for mean schooling	Effective primary pupils adjusted for learning index
Guatemala	14	0.68	**9**
Guinea	9	0.59	**5**
Guinea-Bissau	14	0.57	**8**
Guyana	11	0.64	**7**
Haiti	#N/A	0.57	**#N/A**
Honduras	14	0.67	**9**
Hong Kong, China	5	0.90	**5**
Hungary	4	0.90	**3**
Iceland	9	0.89	**8**
India	10	0.68	**7**
Indonesia	11	0.72	**8**
Iran	9	0.73	**6**
Iraq	12	0.66	**8**
Ireland	11	0.90	**10**
Israel	11	0.78	**9**
Italy	5	0.91	**4**
Jamaica	9	0.72	**6**
Japan	6	0.89	**5**
Jordan	12	0.78	**10**
Kazakhstan	6	0.89	**5**
Kenya	15	0.72	**11**
Kiribati	#N/A	0.78	**#N/A**
Korea, Rep.	7	0.90	**7**
Kuwait	7	0.78	**5**
Kyrgyzstan	7	0.51	**3**
Laos	12	0.68	**8**
Latvia	5	0.88	**4**
Lebanon	9	0.78	**7**
Lesotho	13	0.55	**7**
Liberia	13	0.57	**8**
Libya	11	0.71	**8**
Lithuania	4	0.89	**4**
Luxembourg	7	0.90	**6**
Macao, China	5	0.92	**4**
Macedonia	5	0.72	**3**
Madagascar	14	0.56	**8**
Malawi	13	0.57	**7**
Malaysia	11	0.77	**8**
Maldives	14	0.66	**9**
Mali	11	0.60	**7**
Malta	6	0.90	**5**
Mauritania	11	0.58	**7**
Mauritius	7	0.72	**5**
Mexico	11	0.72	**8**
Micronesia	#N/A	0.78	**#N/A**
Moldova	#N/A	0.69	**#N/A**
Mongolia	9	0.76	**7**
Montenegro	#N/A	0.70	**#N/A**
Morocco	10	0.58	**6**
Mozambique	16	0.70	**11**
Myanmar	9	0.75	**7**
Namibia	14	0.54	**8**
Nepal	#N/A	0.59	**#N/A**
Netherlands	8	0.90	**7**
New Zealand	8	0.87	**7**
Nicaragua	12	0.78	**10**
Niger	8	0.60	**5**
Nigeria	12	0.56	**7**
Norway	9	0.86	**8**
Oman	9	0.71	**6**
Pakistan	9	0.68	**6**
Palestinian Autonomous Territories	9	0.66	**6**
Panama	11	0.70	**7**
Papua New Guinea	8	0.71	**5**
Paraguay	12	0.71	**8**
Peru	10	0.71	**7**

Table A4 – Summary indices of primary pupil volume and quality – continued

FINAL PRIMARY PUPIL ESTIMATES			
Country	Effective primary pupils as % of population	Group score III adjusted for mean schooling	Effective primary pupils adjusted for learning index
Philippines	13	0.77	**10**
Poland	6	0.90	**6**
Portugal	7	0.93	**6**
Qatar	6	0.63	**4**
Romania	4	0.77	**3**
Russian Federation	3	0.90	**3**
Rwanda	14	0.57	**8**
Samoa	15	0.79	**12**
Sao Tome and Principe	14	0.55	**8**
Saudi Arabia	11	0.65	**7**
Senegal	10	0.58	**6**
Serbia	3	0.77	**2**
Seychelles	9	0.72	**7**
Sierra Leone	23	0.57	**13**
Singapore	#N/A	0.93	**#N/A**
Slovakia	4	0.90	**4**
Slovenia	5	0.89	**5**
Solomon Islands	15	0.64	**10**
South Africa	12	0.64	**8**
Spain	5	0.93	**5**
Sri Lanka	7	0.63	**5**
St. Kitts and Nevis	#N/A	0.72	**#N/A**
St. Lucia	10	0.72	**7**
St. Vincent and the Grenadines	11	0.71	**8**
Sudan	10	0.65	**7**
Suriname	10	0.65	**6**
Swaziland	14	0.60	**9**
Sweden	6	0.89	**6**
Switzerland	6	0.88	**6**
Syria	9	0.67	**6**
Tajikistan	9	0.69	**6**
Tanzania	18	0.69	**12**
Thailand	7	0.79	**5**
Timor-Leste	15	0.68	**10**
Togo	11	0.69	**7**
Tonga	11	0.79	**9**
Trinidad and Tobago	8	0.75	**6**
Tunisia	9	0.67	**6**
Turkey	8	0.81	**7**
Turkmenistan	#N/A	0.76	**#N/A**
Uganda	17	0.66	**12**
Ukraine	3	0.77	**3**
United Arab Emirates	6	0.92	**5**
United Kingdom	7	0.90	**6**
United States	8	0.88	**7**
Uruguay	9	0.78	**7**
Vanuatu	13	0.79	**10**
Venezuela	10	0.71	**7**
Vietnam	#N/A	0.74	**#N/A**
Yemen	12	0.55	**6**
Zambia	18	0.55	**10**
Zimbabwe	18	0.64	**12**

Appendix A5.

Email correspondence of ICP and TAG members on the transformation of learning scores and estimation of education output.

"The main question here, is how to interpret differences in PISA scores. What does it mean if one country's score is 110 and the score for another country is 100 [this would be the normalized scores, PISA publishes scores as centered at 500, with STD of 100, but many users including OECD normalize them to 100%]? The simplest proposal is that the first country's educational output would be raised by 10% when compared to the second one. What about the difference between the scores 70 and 80 then? Would it be the same? Does this mean, for example, that if country A has twice as many students per capita as country B, and the scores are 60 and 120 in countries A and B, respectively, then country A would have the *same* educational output per capita as country B? If the transformations are linear, this would be the case, at first glance, a strange result. But, if not a linear form, then what analytical form should be used to transform the scores into education output? And why? Perhaps we could learn about the analytical form by analyzing education system within one country? It should be able to explain the willingness to pay for the perceived differences in the quality of education.

One proposal is to derive the education output from school-level data. Say we have the education subaggregate of government expenditure whose relative between two countries we are trying to deflate with a purchasing power parity (PPP) to get the relative volume of education services between the two countries (say, A and B):

$$\frac{E^A}{E^B}.$$

We characterize education expenditures as the product

$$E^A = \text{cost per pupil hour in A} \times \text{pupil hours in A} = P^A\left(\text{quality factors}^A\right) \times Q^A.$$

Where the cost per pupil hour (P) is a function of factors affecting the quality of a pupil hour (with the total pupil hours being Q). At the top of these quality indicators is performance on a standardized test such as PISA, or more and more distant proxies for that score in lieu of available PISA information, with the most distant proxy being 'percent pupils having parent with post-secondary education', which explains about 63 percent of the variation in the PISA score for countries where the PISA has been administered.

When we form the PPP, or price relative for pupil hours between countries A and B, we will have to account for the fact that the quality factors differ, and adjust to allow us to compare like with like. Price index-wise, standard practice is to take the following geometric average, simplifying at the end to acknowledge that we have one, results-oriented quality factor, the PISA score:

$$PPP^{AB} = \left[\frac{P^A\left(\text{quality factors}^A\right)}{P^B\left(\text{quality factors}^A\right)} \times \frac{P^A\left(\text{quality factors}^B\right)}{P^B\left(\text{quality factors}^B\right)}\right]^{\frac{1}{2}} = \left[\frac{P^A\left(\text{PISA}^A\right)}{P^B\left(\text{PISA}^A\right)} \times \frac{P^A\left(\text{PISA}^B\right)}{P^B\left(\text{PISA}^B\right)}\right]^{\frac{1}{2}}.$$

Let's suppose to keep it initially simple that we have a PISA dataset over schools for all countries, and we run, for each country, a regression of school cost per pupil hour on school PISA score. The parameters of that regression give us, for each country, say, A, the function $P^A(\text{PISA score}^{Ai})$, where i runs over schools in country A. To compare country A with country B, we'll have to find representative values of PISA score^A and PISA score^B to plug into the above geometric mean

(Fisher-type) price index formula. Ideally, the representative value should be such that when it is plugged into the price function of the same country as that of the representative PISA score, you get the national cost per pupil hour of that country. If there were no other complications, this would be our education PPP.

Note that the function $P^{A}(\text{PISA score}^{Ai})$ need not be linear, and probably should not be, in order to fit the actual school cost per pupil hour data *vis-a-vis* PISA score.

Note also that when we deflate the expenditure ration by this quality adjusted PPP, we will not get the ratio of the Qs between the two countries, but a quality adjusted version of that.

Of course, there is a complication, which is that we do not have PISA scores for all countries, so we have to predict the PISA score based on school data where we have both the PISA score and one of several proxies and/or schools where we have overlaps between proxies, allowing us to link the predicted PISAs or actual PISAs, where available, together."

(Email communications from Yuri Dikhanov and Kim Zieschang, 8/18/10).

Appendix A6. Addressing Discrepancies in Pupil Counts

One UIS publication of particular interest to the ICP presents an in-depth expert analysis of the differences in absolute pupil counts from EMIS and DHS surveys for 10 developing countries (Stukel and Feroz-Zada, 2010). In this analysis, EMIS pupil counts were obtained directly from UIS. Pupil counts from the DHS were obtained by multiplying the number of pupils counted by the survey in the various strata (sub-sections of the country) by the respective strata population weights. The original population weights are not provided with the DHS datasets but were requested from Macro International, the organization that administers DHS surveys. Stukel and Feroz-Zada (2010) were able to obtain the original weights for 11 of the 16 countries they requested (one country was subsequently dropped from their analysis). The analysis was done for Bangladesh, Côte d'Ivoire, Egypt, Indonesia, Mozambique, Namibia, Nigeria, Rwanda, Tanzania, and Vietnam.

The observed pupil count differences were large (>10%) for 9 out of 10 countries - an even lower correspondence than for the attendance ratios shown in Figure 1.1-1. Stukel and Feroz-Zada (2010) comment: "… eight countries (Bangladesh, Côte d'Ivoire, Egypt, Indonesia, Mozambique, Namibia, Rwanda and Tanzania) have values highlighted where the relative percent differences (in absolute numbers) are greater than 10% but less than 25%. For all countries except Indonesia and Tanzania, the values are positive, indicating that the enrolment figures are substantially higher than the attendance figures. In Indonesia and Tanzania, the inverse is true. There is one country where the discrepancy exceeds 25% – Vietnam (47.2%)." (Stukel and Feroz-Zada 2010:15). The numbers are shown in **Error! Reference source not found.**.

Stukel and Feroz-Zada (2010) analyzed the cause of these differences, and conclude that *when carefully analyzed and adjusted, the two source estimates for pupil numbers are within 10% for 8 out of 10 countries.* For one of the two remaining countries, an analysis of the questionnaire suggested which source was more appropriate to choose, leaving just one country out of 10 with unexplained pupil count discrepancies in excess of 10% (Tanzania) where the EMIS pupil counts could need to be adjusted (interestingly in this country, the *rates* are only 6% apart). Specifically, Stukel and Feroz-Zada (2010) findings on the pupil count discrepancies between DHS and administrative sources and the reconciliations of the majority of those discrepancies are as follows:

a) In one country, Vietnam, the pupil count differences between EMIS and DHS were 47% (8.5 million pupils vs. 4.0 million), but the enrolment and attendance rates are identical – 96%. An inquiry to Macro International showed that the pupil count difference is the result of a decision by Macro International to use old population weights. "(T)he Viet Nam DHS for 2002 was based on the previous DHS in 1997, which in turn was a sub-sample of the 1996 Multi-Round Demographic Survey. Nevertheless, Macro International made the decision not to boost the final weights for DHS 2002 by the inverse of the sub-sampling rate since their main interest was to produce ratio estimates and omitting this step would not matter. If the weights had been boosted by the inverse of the sub-sampling rates for Viet Nam, this would have generated survey-based estimates of totals …, which would have been close to the estimate from the corresponding alternate source …" (Stukel and Feroz 2010:16)

b) Most of the pupil count discrepancies can be reduced by aligning the age-distribution of the household survey sample with that of census-based population estimates or counts. Pupil discrepancies can result from differences in the population age-distribution because age-specific attendance varies greatly by age, and an over-or under-representation of a particular age-group can skew the aggregated pupil counts. Often, surveys adjust the population sampling weights using a technique called post-stratification or benchmarking. It is not obligatory however because post-stratification has little effect on ratios, which are the results of interest for surveys. Stukel and Feroz were informed by Macro International that while "DHS countries do perform empirical checks to ensure that there is coherence in age distributions between surveys and national population sources, and that post-stratification takes place when it is deemed necessary (i.e. when there is a lack of coherence), none of the DHS countries considered in this report have used post-stratification" (Stukel and Feroz, 2010:32). When Stukel and Feroz post-stratified the population weights of the DHS surveys to match the UN population estimates used by UIS, the percentage difference of pupil counts was reduced to under 10% for an additional seven of the 10 countries (not including Vietnam, which was discussed in point a).

c) Some pupil count discrepancies arise because some surveys inquired only about attendance in the last week, thus missing pupils who were out of school temporarily due to illness, vacation, or other causes. One country, where the pupil estimate was still significantly lower for household surveys than for EMIS systems even after the post-stratification was Bangladesh – survey pupil estimates were 12.7 million compared to 15.0 million from EMIS. Bangladesh is also a country where the DHS inquired only about attendance in the *last week.* According to Stukel and Feroz: "In Bangladesh, the academic year spans all 12 months of the year. Given that the missing information on past attendance may have constituted a significant portion of attendance for this country, the enrolment figure for Bangladesh may be considered more credible than the attendance figure" (Stukel and Feroz, 2010:38).

Table 1.1. Selected results from Stukel and Feroz, 2010, analysis of discrepancies in school participation and pupil numbers between EMIS and DHS, for 10 countries.

				All pupil numbers in thousands					
	Enrolment rate EMIS)	Attendance rate (DHS)	Relative difference NER/NAR	EMIS enrolment counts (from UIS) in thousands	DHS attendance counts	Unadjusted relative difference to EMIS	Difference within 2 s.d. error of DHS estimate?	DHS attendance counts, post-stratification and sampling adjustments	**Final, adjusted relative difference to EMIS**
Bangladesh	93	79.6	13%	15,020	12,467	-17%	NO	12,728	**-13%**
Cote d'Ivoire	53.7	52.2	2%	1,474	1,304	-11%	YES	1,421	**-4%**
Egypt	93.5	85.5	8%	7,340	6,531	-11%	NO	6,731	**-8%**
Indonesia	100.9	95.3	6%	25,185	29,527	17%	YES	23,588	**-6%**
Mozambique	62.9	59.9	3%	2,318	1,842	-21%	NO	2,243	**-3%**
Namibia	74.2	78.6	-4%	283	226	-21%	NO	298	**5%**
Nigeria	62.1	62	0%	13,211	12,030	-9%	YES	13,299	**1%**
Rwanda	71.1	71.9	-1%	1,046	910	-13%	NO	1,058	**1%**
Tanzania	47.7	53.8	-6%	3,105	3,444	11%	YES	3,492	**12%**
Vietnam	96.1	96.3	0%	8,498	4,487	-47%	NO	8,494	**0%**