International Database on Human Capital Quality: An Update

Nadir ALTINOK*, Claude DIEBOLT, Jean-Luc DE MEULEESTER

Incomplete version – Please do not quote
This version: 19/01/2010

Abstract.

The aim of this paper is to propose an evaluation of the performance of schooling systems. Based on a previous work, our approach tries to evaluate the performance of each level of education, namely the primary and the secondary levels. We use international test on pupil achievement, by employing a unique methodology which enables us to obtain a database containing four complementary databases. The first database gives the possibility to compare schooling performance between countries for each level of education, by computing an average score for schooling quality between 1965 and 2007 (Database 1). This database includes quality scores for 108 countries in primary education and 93 countries in secondary education. The second database permits to compare the global *change* in schooling quality over time, for a long term period. This long term table includes quality scores from 1970 to 2007 for 13 countries (Database 2). Moreover, we propose a complementary database which allows to analyze short term trends – between 1995 and 2007 – on schooling quality for 79 countries (Database 3). Finally, we present an analysis of trends of schooling quality since 1995 by combining all existing data, in order to find significant changes of pupils' achievement in international tests (Database 4). The file containing these four database is available on request.

Key words: Quality, Human Capital, Education, Database, PISA, TIMSS.

J.E.L. Classification: I2.

Author contacts.
*Corresponding author, Nadir Altinok, ID2 (Metz University), BETA (Strasbourg University), IREDU (University of Bourgogne). Adress: BETA (Bureau d'Economie Théorique et Appliquée), 61 avenue de la Forêt Noire, 67085 Strasbourg Cedex, France. Tel: +33 387 54 71 38 Fax: +33 390 24 20 70. Email nadir.altinok@gmail.com.
Claude Diebolt, BETA (Bureau d'Economie Théorique et Appliquée), 61 avenue de la Forêt Noire, 67085 Strasbourg Cedex, France. Tel: +33 368 85 20 69 Fax: +33 368 85 20 71. Email: cdiebolt@unistra.fr
Jean-Luc De Meulemeester, DULBEA (University of Bruxelles), Avenue F.D. Roosevelt 50, 1050 Bruxelles. Email: jldemeul@ulb.ac.be

Acknowledgments. We are most grateful to Geeta Kingdon (London University) and Jean Bourdon (University of Bourgogne, CNRS) for helpful comments on earlier drafts of this paper

Introduction

The debate around the contribution of human capital to economic growth has been especially tense during the last two decades. The renewal of interest in growth theories since the end of the 80s was first mainly theoretical and led to the development of various endogenous growth models. At this stage, there was already an intense theoretical debates between the partisans of enlarged Solow-style exogenous growth models with an extensive definition of capital (including human capital – viewed here as another factor of production whose accumulation matters for growth, at least in a transitory way), and the partisans of endogenous growth models. Among the latter, a distinction had to be made between partisans of accumulation models (see Lucas, 1988) where human capital is considered as other factors of production (its accumulation matters, i.e. the growth rate of human capital impacts permanently on the growth rate of the economy per capita) and the partisans of innovation models (see Romer, 1990; Aghion and Howitt, 1992) where the stock of human capital (e.g. the portion of it dedicated to the R&D activities) matters for growth. Here again one should add that some economists stress innovation activities whereas others (in the spirit of the pioneering paper by Nelson and Phelps, 1966) stress imitation and adaptation (diffusion of existing knowledge). In both cases, the stock itself matters more than the flow (growth rates of human capital). At first, these debates were first of all theoretical (maybe at the exception of the partisans of exogenous models, as the contribution of Mankiw, Romer and Weill in 19921 was mainly empirical). But if economics pretends to be a science it had to confront, sooner or later, to the empirical test. The problem was the measurement of human capital. The early empirical attempts relied upon measure of school participation (in primary and secondary schools for example). These very imperfect measures were used as proxies for the growth rate of human capital and tended to support the accumulation view (as in Barro, 1990). In other words, it seemed that the growth rate of human capital was central for the growth rate of GDP per head. However, once international databases became available (as Summer and Heston, 1993), and more precise measures of human capital, the accumulation model does not seem as solid as it appeared first. When one used the average year of schooling in a population, the education variable (proxying the growth rate of human capital) did not appear significant any more (Benhabib and Spiegel, 1994; Pritchett, 2001). These disappointing results led to huge debates concerning the role of human capital and especially education for economic growth. Results showing that the contribution of education to growth was not especially clear in developed countries led to tense policy debates (see its echo in publications as Myths about education and growth by Alison Wolf, 2002). If there were theoretical and econometric debates (see Krueger and Lindahl, 2001, for a good survey), debates also pertained to the quality of the databases. Inconsistency and imperfect comparability of figures in those cross-country databases suggest that the education variable was subject to considerable measurement errors. Some economists as Gurgand (2004) stressed that once one used such imperfect data to compute growth rates, errors are compounded and this tends to reduce the significance of its estimated coefficient

in regression analyses. It was maybe why, after all, accumulation models appeared to perform less well than mere innovation or adaptation (i.e. stock) models. On top of that, debates were also launched concerning the pertinence of considering homogenous effects of education whatever the level of economic development of the countries analyzed. For example, Benhabib and Spiegel (1994) not only showed that a stock measure of human capital perfomed better, but also that in developing countries a catch-up variable (a measure of the stock of education times the ratio of the country GDP versus the American one) was more significant than the mere stock variable, the reverse being true for developed countries. Even if some errors were later found in Benhabib and Spiegel work, it nevertheless launched an interesting debate and questioned the "one size fits all" policy of the time regarding education and human capital. A neo-schumpeterian approach has been developed following this intuition and earlier developments on endogenous innovation models (with a Schumpeterian touch, i.e. models where innovators hope to displace existing producers and enjoy for a time the monopoly rents, before being displaced themselves by new innovators). This literature (Aghion and Howitt, 2005) introduced a difference between an innovation economy and an imitation economy (or accumulation one) and demonstrated that the optimal policies for growth are different in the two types of economy (the difference being measured by the distance to a technological frontier). This model has been tested on data for 19 OECD countries by Vandenbussche, Meghir and Aghion (2006) and seemed to confirm these intuitions. Countries near the technological frontier should rely on innovation and hence on the development of their higher education sector (research-oriented) whereas countries below the technological frontier (at 25% below the total factor productivity of the leader in their paper) should rely on imitation and by the way on the development of their primary and secondary education systems. As far as innovation is concerned, one could also think that quality matters most. This is another way of debating the role of human capital to economic growth. A large proportion of the literature has indeed focused on the role of the quantity of human capital and not that much on its quality. This paper is mainly within this latter framework of analysis. It will provide a new and unique database enabling researchers to carry out long run econometric analyses on the contribution of the quality of education to economic growth.

In this paper, we propose to evaluate the potential differences in the quality of education systems. The main reason for referring to learning achievement surveys relates to the now acknowledged impact of educational quality on wages (this is the micro level) and economic growth.

Until very recently, most studies sought to demonstrate that the quantity of education – generally measured by the number of years of study – had an impact on income. Drawing on the pioneering work of Mincer (1974), such research showed that, on average, an additional year of education was correlated with a 10% increase in income, but that the returns on education differed significantly from one country and income level to another (Psacharopoulos and Patrinos, 2004). During the 90s, a pioneering paper by Card

(1996) - Does School quality matters? – introduced the importance of (secondary) school quality in determining future income paths. This study pointed to the resources devoted to schools (e.g. teacher-pupil ratio) (?? As well as peer-group effects?). More recent studies tended to highlight direct measures of the quality of pupils achievements and its role in the determination of incomes (see Coulombe and al.?). For some years now, fresh studies have pointed to a positive link between learning achievement and income (see Unesco, 2004, for a review of bibliographic references). These studies have generally been based on pupils' scores in tests held to assess their skills in mathematics, science and reading. Three studies in the United States of America have revealed that an increase of one standard deviation in pupils' mathematics scores is linked to a 12% increase in income (Mulligan, 1999; Murname et al., 2000; Lazear, 2003). On the basis of data from the International Adult Literacy Survey (IALS)[1] covering 15 countries including Canada, Chile, the United States of America and 12 European countries, Leuven, Oosterbeek and van Ophen (2004) have shown that the workers' cognitive abilities – measured by their scores in reading – have a significant impact on their earnings. This impact is still apparent when the number of years spent in education is taken into account, indicating that greater importance should be attached to the quality of education. In regard to the developing countries, Hanushek and Woessmann (2007) point out that their return on education appears to be higher than in the developed countries. Furthermore, Sakellariou (2006) uses the IALS data on Chile to show that an increase of one standard deviation in workers' reading levels leads to a 15% to 20% rise in their income.

The influence of education on economic development is equally worthy of note. The relation between education and economic growth has often been evaluated on the basis of quantitative education indicators. Most of the research concerned has highlighted that it was a key factor of economic development. Pritchett (2001) has shown, however, that the effect of education on growth was not straightforward. Emphasizing the importance of educational quality to economic development, he demonstrated that poor educational quality could be linked to weak economic development. In itself, the qualitative aspect of education may basically spur economic growth. Recent research correlating educational quality with economic growth tends to highlight the importance of quality (Lee and Lee, 1995; Hanushek and Kimko, 2000; Barro, 2001; Coulombe and Tremblay, 2006; Hanushek and Woessmann, 2007, 2009). The most influential study shows, in particular, that the increase of one standard deviation in pupils' scores correlates with a 1% rise in the annual growth rate of per capita GDP (Hanushek and Kimko, 2000).

The goal of this paper is to provide a means of measuring the quality of schooling for the purposes of international comparison. Its main contribution is its provision of comparable standardized indicators of achievement in 108 countries for primary education and 93 countries for secondary education. This is, to

[1] IALS is a seven-country initiative first conducted in the fall of 1994. Its goal was to obtain comparable literacy profiles across national, linguistic and cultural boundaries. A second and a third round of data collection of IALS were conducted in an additional 16 countries in 1996 and 1998. See for more information: http://www.statcan.gc.ca/dli-ild/data-donnees/ftp/ials-eiaa-eng.htm

our knowledge, the most complete database on human capital quality and the first one which permits to tracks trends in education quality. Many developing countries are included, mostly due to the fact that we take into account regional learning assessments, in Latin America and in Subsaharan Africa. Previous studies often neglected these continents, and focused only on developed countries. However, in the context of the 'Education for All' initiative, many billion of dollars are currently engaged to promote more education. Hence, our database can be useful to analyze to what extent the promotion of *more* education is linked with a *better* education.

We obtain four complementary databases, which are available on request by email. The first database gives the possibility to compare schooling performance between countries for each level of education, by computing an average score for schooling quality between 1965 and 2007 (Database 1). This database includes quality scores for 108 countries in primary education and 93 countries in secondary education. The second table permits to compare the global *change* in schooling quality over time, for a long term period. This long term table includes quality scores from 1970 to 2007 for 13 countries (Database 2). Moreover, we propose a complementary analysis which allows to analyze short term trends – between 1995 and 2007 – on schooling quality for 79 countries (Database 3). Finally, we present an analysis of trends of schooling quality since 1995 by combining all existing data, in order to find significant changes of pupils' achievement in international tests (database 4). All this data is available on request.

2. Data and methodology

2.1. A Brief Presentation of International and Regional Learning Achievement Surveys

This section describes international learning achievement surveys that may be used to devise indicators on the quality of educational achievement. They comprise seven groups of international surveys in which about 105 countries have participated worldwide. These groups may be divided into two subgroups: the first consists of international assessments, while the second comprises regional assessments. Full information on these assessments is provided in Table 1. Here, we make a quick presentation of learning assessments. More information can be obtained in Appendix 1 and in Altinok (2009).

The International Association for the Evaluation of Educational Achievement (IEA) was the first body to measure individual learning achievement for international comparative purposes in the early 1960s. The surveys include the highly regarded "Trends in International Mathematics and Science Study" (TIMSS) and "Progress in International Reading Literacy Study" (PIRLS). If we only take into account recent surveys, four rounds are available for TIMSS (1995, 1999, 2003, 2007), while only two rounds are available for PIRLS (2001 and 2006). Another international assessment is PISA. The Organisation for Economic Co-operation and Development (OECD) launched its Programme for International Student Assessment (PISA) in 1997. More generally, PISA has assessed the skills of 15-year-olds every three years since 2000 in

countries that together account for almost 90% of the global economy. Until now, three rounds of PISA ara available (2000, 2003 and 2006).

Two other international assessments are available. Drawing on the experience of the National Assessment of Educational Progress (NAEP), the International Assessment of Educational Progress (IAEP) comprises two surveys first conducted in 1988 and in 1991. Under a joint UNESCO and UNICEF project, learning achievement is being assessed as part of the Monitoring Learning Achievement (MLA) programme on a vast geographical scale in more than 40 countries with a view to building national assessment capacity (Chinapah, 2003).

Three major regional assessments have been conducted in Africa and Latin America. The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) grew out of a very extensive national investigation into the quality of primary education in 15 african countries in 1995-1999 and 2000-2002. Following a different approach, surveys under the "Programme d'Analyse des Systèmes Educatifs" (PASEC, or "Programme of Analysis of Education Systems") of the Conference of Ministers of Education of French-Speaking Countries (CONFEMEN) have been conducted in the French-speaking countries of sub-Saharan Africa between 1993 until now. Finally, the network of national education systems in Latin American and Caribbean countries, known as the Latin American Laboratory for Assessment of the Quality of Education (LLECE), was formed in 1994 and is coordinated by the UNESCO Regional Bureau for Education in Latin America and the Caribbean. Assessments conducted by the LLECE thus focused on learning achievement in reading and mathematics in grades 3 and 4 in 13 countries of the subcontinent in 1998 and 2006

All surveys undertaken and the main information relating to them are summarized in Table 1. The methodology used to adjust them in order to yield comparable indicators is breafly presented below.

2.2. A methodology for the consolidated study of surveys

This paper is an update of a previous publication (Altinok and Murseli, 2007). The international surveys discussed in the previous section will now be placed on common ground. As learning achievement surveys vary in nature, methods must be devised for inter-survey adjustment, otherwise comparison between findings from countries that have taken part in two or more surveys would be impossible. For example, a given country's score may be higher in one survey than in another because the first survey has recorded higher absolute scores. If the scores obtained by the same country in the two surveys are not adjusted, its performance would appear to have deteriorated if the first assessment were conducted shortly before the second. Survey findings must therefore be adjusted to permit comparisons between countries' performance and between trends over time. Some modifications in the methodology occurred since the first version of our database.

Under the method used here, some surveys are adjusted to the score obtained by the United States of America, which has been used as the particular country "base" for this purpose (called 'survey series A'). Since a separate measurement of the performance of pupils in the United States of America is available, the level of difficulty of each international survey may be evaluated on the basis of the differential between the US score in the international survey and in its own national survey. This adjustment makes it possible to evaluate the extent to which a country's scores in different surveys have been overestimated or underestimated and thus to obtain comparable data over time. However, a specific procedure has been used in recent international surveys to avoid adjusting them all solely in accordance with performance in the United States of America by devising them in such a way as to permit analysis of trends in country scores over time (called 'survey series B'). Moreover, for some achievement tests, the USA did not participated. This is especially the case for regional surveys such as LLECE or SACMEQ. This 'survey series C' are taken into account by using countries' performance which participated both in regional and international surveys.

Three different adjustments may be required in order to obtain comparable indicators. The first involves recurrent reference to scores obtained by the United States of America. Readily comparable indicators may be obtained by anchoring data from other surveys to these scores (a procedure known as "the first NAEP anchoring methodology"). These "series A" surveys account for most international learning achievement surveys conducted until 1995. All international surveys of reading until PIRLS 2001 inclusive and of mathematics and science until TIMSS 1995 inclusive, together with PISA 2000, the very first OECD survey, are thus "series A" surveys. The other surveys (PIRLS 2006, TIMSS 1999, TIMSS 2003, TIMSS 2007, PISA 2003, PISA 2006) fall into "series B".

This second series entails adjustment by means of another procedure known as the "repeated NAEP anchoring methodology". recent surveys have been designed to permit analysis of country score trends. Pupils are given test pieces at the same level of difficulty in survey series conducted by the same body, which makes it possible to analyse trends in pupil performance directly over time. Accordingly, the same adjustment coefficients as those calculated for the series A surveys are used in order to achieve a single linear conversion of country scores, and this procedure does not compromise the comparability of the scores obtained by countries participating in the same survey series. It should be noted that this adjustment is no more than a constant linear translation of the recent surveys. This means that the variance within surveys as well as the variation in country scores in separate surveys by the same body are exactly the same as those calculated on the basis of unadjusted data.

Finally, the "series C" surveys include regional surveys, that is surveys in which the USA did not participated. We choose a reference survey and look for countries which participated to this survey and the regional survey. For instance, the reference survey for mathematics is TIMSS. From the TIMSS indexes, we tried to track countries which had participated in the TIMSS and at least one other survey. It will be noted that all the surveys include countries which participated in the TIMSS and at least one other

survey. Except for the PASEC survey which does not include the countries of the TIMSS survey[2]. We then proceeded a matching based on the means of the countries which participated in at two surveys (called "doubloons"); each survey was re-adjusted according to the level of comparison related to the referee survey of IEA (TIMSS). The reference survey of reading is PIRLS. In appendix B, we present a detailed version of our methodology.

3. An analysis of schooling quality

3.1. Cross-section analysis

Before examining trends in pupils' performance over time, it may be desirable to study possible differences in the averages of countries' performance in international learning achievement surveys. A single indicator of countries' performance is obtained by calculating the arithmetic mean of the different scores obtained by countries at each educational level. For example, if a country has taken part in three different surveys of secondary education, its mean score for the three will be calculated. This procedure may be used to obtain indicators of the quality of student achievement at both primary and secondary levels (*QIHC-PRI* and *QIHC-SEC*, respectively) and to avoid some of the inconsistencies in student achievement levels from one survey to the other. By calculating the mean of a country's scores between 1970 and 2007, it is possible to evaluate its mean performance for the last 37 years, without relying solely on one survey that might overestimate or underestimate its pupils' actual achievement level. However, it should be noted that available data change from country to country, and for some countries, we only have one single score. We hope to correct this drawback in the future with new surveys. The database contains information on primary education in 108 countries and on secondary education in 93 (see appendix C for this database). In order to achieve comparable scores across different educational levels, the former have been standardized to obtain an arithmetic mean of 50 and a standard deviation of 10 for primary and secondary education.

Table 2 shows that the quality of learning achievement varies with the economic level of countries and the educational level concerned. The World Bank classification has been used to group countries in accordance with their economic level. We distinguish high income, middle income and low income countries[3]. There is a fairly clear positive correlation between countries' economic level and their pupils' performance in international tests at any educational level. In primary education, the mean deviation between the developed and developing countries is more than 22 points: whereas on average the *QIHC* is about 60 score points at primary level for high income countries (*HIC*), it is about 37 score points for the

[2] As for the re-adjustment of the PASEC survey to the referee survey, we have used the re-adjusted MLA-survey.
[3] Economies are divided according to 2008 Gross National Income per capita, using the World Bank Atlas Method. The groups are Low Income Countries (LIC), $975 or less, Middle Income Countries (MIC), $976-$11,905, and High Income Countries (HIC), $11,906 or more.

low income ones (*LIC*). On the other hand, the difference between high income countries and middle income countries (*MIC*) is less important, but remains substantial (60 and 48 points respectively). The difference is much the same in the case of secondary education: here, the arithmetic mean for the high income countries is 59 points compared to 35 points for the low income ones. The deviation in performance between the *HIC* and the *MIC* is approximately the same in secondary education in comparison with the primary level (59 and 45 points respectively).

The variability in scores may be measured by means of standard deviations. The greater the standard deviation, the greater also the variability within the sample under consideration. While pupils' scores for the developed countries are fairly uniform (with a standard deviation of around 3 at primary and secondary levels), they are less so for midlle income countries (standard deviation close to 7.5 at primary and secondary level). Concerning the low-income countries, one have to be cautious with the standard deviations in the secondary level, because our database include only 7 countries. However, it seems to be quite similar in the primary level, but it can be due to the lack of data concerning this group of countries.

In order to determine each country's score, attention is focused in Table 3 on the countries' performance in primary education with respect to their economic level. Classifying countries in this way is preferable in order to establish comparable benchmarks. It should be remembered that it is learners' average performance in overall terms – at national level – which is measured and not the whole dimension relating to the quality of learning achievement. The classification of countries by economic level implies that if a developed country's performance ranking is poor, its performance is not necessarily worse than that of a favourably ranked developing country. The developed countries in Asia and Northern Europe seem to perform best (Figure 1, and Table 3). The top two countries in this group are Republic of Korea and Finland, followed by Japan and Taiwan of China. Conversely, the countries with the lowest performance scores in the group for which data are available are Trinidad & Tobago, Portugal, Cyprus, and Iceland. The relatively modest scores of these countries in TIMSS or PISA type surveys are logically reflected in the present classification.

It is clear from the analysis of the performance of the middle income countries that pupils in Kazakhstan, the Russian Federation, Lithuania, and Latvia have demonstrated high levels of achievement (see Table 3 and Figure 2). It may thus be concluded that, among the middle income countries, those in Eastern Europe and Central Asia perform best. In the same group, two African countries appear to perform least well. (Namibia, Lesotho and Swaziland). By way of comparison, the score achieved by Namibia is less than half that of the Republic of Korea. Concerning the low income countries, Madagascar seems to outperform other countries. The level of education quality in this country appear to be higher than other middle income countries, such as Swaziland or Lesotho. On the contrary, some countries such as Mauritania and Niger appear to have a very low level of schooling quality.

Analysis of the countries' performance in secondary education shows little change compared to the primary education ranking (see Table 3 and Figure 2). The developed countries in Asia perform best,

whereas those in Eastern Europe – Cyprus and Greece – score lowest. The performance levels of developing countries in terms of learning achievement are again highest in Eastern Europe. The four countries that perform best are all in this region, namely Poland, the Russian Federation, Croatia and Lativa. Conversely, most countries with the lowest scores for secondary education are in Africa (South Africa, Botswana). Moreover India's very low score should be placed into perspective given that the country has not taken part in international surveys since 1970.

3.2. An Analysis of trends on schooling quality

3.2.1. Long term analysis

It appears crucial to deeply analyze the global trends on schooling quality. As it has been highlighted in the introduction, comparable data relative to schooling quality is very difficult. When we analyze the quality of schooling data already existing (such as the Barro and Lee (2001) database), it seems important to obtain alternative indicators which partially take into account the qualitative dimension of education.

Moreover, we propose to analyze the global trends on schooling quality, by using the scores computed by combining mathematics and science scores for secondary education. To our knowledge, this is the first study who try to analyze trends in schooling quality[4]. After each survey has been adjusted, as some of them obviously concern roughly the same educational level and the same skill, they may be considered jointly by calculating the arithmetic mean of the scores obtained by the countries surveyed. For instance, the TIMSS 1999 and PISA 2000 surveys for secondary-school mathematics and science have been grouped together in this way, as have the TIMSS 2003 and PISA 2003 surveys for secondary-school mathematics and science. In final, we obtain data on human capital quality based on results from tests in mathematics and science in secondary education. When a score is not available between two years, we make a linear interpolation by using informations relative to the variation of performance for other years, in mathematics and science. It is clear that this first version is still incomplete, since it concerns only 19 countries. However, we hope to update it with the next surveys, such as PISA 2009 and TIMSS 2011.

As we can see in the table 4, comparable data is available for very few countries from 1970 to 2007. While, only 13 countries have comparable data in the long run (from 1970 to 2007), our database include 19 countries for which we can obtain comparable data between 1982 and 2007. Countries are ranked by the global variation of their schooling quality between 1980 and 2007.

The first interpretation which can done is a global convergence between countries on schooling quality since the 1970s. For instance, whereas Hungary was ranked higher than Finland in 1970 (respectively 62 points and 49 points score), these two countries tend to converge their quality of secondary education

[4] Although Hanushek and Woessmann (2009) try to make similar analysis, their study is only graphical and very few informations are available regarding to the amplitude of trends in schooling quality.

across time. Indeed, in 2007, the quality of secondary schooling is approximately equal to 55 for Hungary and 60 for Finland. The comparison of standard deviation of scores between 1970 and 2007 for the 13 countries where data is available confirm the convergence hypothesis. The standard deviation for schooling quality in 1970 is equal to around 13, where it is equal to approximately 3.6 in 2007 for the same group of countries. The figure 3 highlight this convergence. This figure presents the long term trends on schooling quality for selected countries. Although for some countries, a significant growth of schooling quality is present (for instance for Chile, Hong-Kong and Luxembourg), a global decline of schooling quality can be observed in countries like Japan. While the level of schooling quality was the highest in this country in 1970 (80 score points compared to the 1970 mean performance of 52 score points, that represents approximately 30 score points higher than the average), a continious decline is present until 2007. Indeed, the global decline between 1970 and 2007 for Japan is approximately equal to -26%. On the opposite side, Hong-Kong has strongly improved the quality of its secondary schools since 1980 (from 44 to 58 score points, which represents a global evolution of +31%).

3.2.2. Short term analysis for secondary schools, combining mathematics and science scores

Due to the scarcity of data for the long term, we propose to undertake an analysis for short term (below 15 years). We follow the same methodology compared to the long term analysis, however now we use recent surveys (between 1995 and 2007). In final, we obtain a panel database for schooling quality, in secondary level, and combining mathematics and science scores. Our database – called Database n°3 – include complete data from 1995 to 2007 for 48 countries, and incomplete data for approximately 70 countries[5].

Table 6 presents trends of schooling quality for 48 countries. The full database, including data for the 70 countries can be obtained on request. The last three columns show respectively the absolute change in scores, the relative variation and the average annual growth rate of schooling quality between 1995 and 2007. Countries are presented according to the AAGR. Here, we restrict schooling quality to the skills in mathematics and science in secondary schools. Schooling quality can be tracked by other channels, such as reading in primary schools. However, due to the scarcity of data, we are forced to restrict to mathematics and science in secondary education. In the next section, we pool all possible information in order to check for the mean variation of schooling quality.

Short term variations show no general trend. While for some countries, scores tend to be higher, in others, schooling quality seems to decrease significantely. This divergence clearly prove the usefulness of analyzing schooling quality. For 25 countries, schooling quality appears to increase during the 1995-2007

[5] For some countries however, we only have one single score. This is the reason why we preferred to present here only countries where all scores between 1995 and 2007 are available. Again, for a minority of countries, a linear interpolation have been used in order to compute the combined scores. For instance, when we only have data for 2006, based on PISA 2006 survey, we make an interpolation, based on the average annual growth rate of the PISA score of the considered country.

period. This is especially the case of some developing countries, such as Brazil and Chile. The AAGR for Brazil is particularly high (+1.80%), compared to Luxembourg, which is the first developed country where we observe a high AGGR (+0.73%). Trends of improve appear to be higher than +0.5% for AGGR in 5 other developed countries (Belgium, Finland, Denmark, New Zealand and Iceland). A lower but significant increase is present in United Kingdom, Germany and the United States. On the opposite side, trends are negative in countries such as Israel, Thailand, Bulgaria or the Syrian Arab Republic. In the case of Israel, schooling quality follow a decrease from 53.7 points to 48.9 points, which represents an AAGR of -0.78% between 1995 and 2007. In other countries as France, Ireland or Cyprus, schooling quality appear to be quite stable over the 12 years period considered here.

3.2.3. Short term trends combining all data

It may be interesting to compare the variation of schooling quality since 1995 and to see which countries improved their quality of schooling. Table 7 presents results for the short term trends on schooling quality. While in previous section, we only focused in mathematics and science in secondary schools, now we include all possible combinaisons. We choose this way in order to check for the robustness of trends of schooling quality in countries for which we have several participations. However, it should be noted that the period between two participations for a country is not always the same compared to other countries. For instance, the United States took part in 1995, 1999, 2003 and 2007 to the TIMSS survey. Then, it is possible to compute the growth rate of schooling quality between 1995 and 2007, 1999 and 2007, 2003 and 2007. Other countries, such as France took part to the PISA survey between 2000 and 2006. Here, the interval is lower compared to the 1995-2007 period. For controlling these differences, we calculated the average annual growth rate (AAGR) of schooling quality. While periods differ between countries and surveys, it is then possible to compare trends on schooling quality. When we have several possibilities to compare trends for the same country, we compute a mean of the AAGR. Firstly, we distinguish the primary and the secondary education (columns 1 to 4 in Table 7). Then, we agregate the two levels in order to check for the global variation. For instance, 11 possible comparisons over time are available for Latvia. If we compute the mean of all 11 AAGR, we obtain an average annual growth rate equal to 0.43. This digit means that on average results in tests for Latvia increase about 0.43 % a year since its first recent participation (here 1995). Among these 11 series, 9 show a positive or constant evolution, whereas only 2 show a negative evolution. Hence, we can conclude that there is a high probability that the quality of education in Latvia improved in recent years (since 1995).

In order to classify short term trends, we adopt the following classification: when a specific evolution appears twice compare to the other trend, we conclude that there is a significant trend. For instance, for Latvia, we find 9 positive trends against 2 negative trends. Here, the number of positive trends are more than the double of the negative trends. On the column "Trend", we can see "++". The number of "+" depends on the amplitude of the mean AAGR: "+" when the AAGR is lower than 0.1; "++" than

0.1<AAGR>0.5 and "+++" when AAGR>0.5. We adopt the same rule for negative trends. When we observe approximately the same number of variations in both sides, we conclude that there is a stagnation of schooling quality. For instance, for Austria, we observe 5 negative trends and 3 positive trends. As 5/3 is lower than 2, we suppose that there is no significant variation of schooling quality. A stagnation is showed as a "o". The number of "o" depends on the mean of AAGR: "ooo" when the AAGR is lower than 0.1; "oo" than 0.1<AAGR>0.5 and "o" when AAGR>0.5. For countries where there is only 1 possible variation, a single "+", "-" or "o" is noted, depending on the sign of this variation.

Data on short term trend is available for 79 countries. 474 different variations have been calculated in order to obtain these trends. Hence, there are approxately 6 different series for each country. For developped countries, such as Hong-Kong-China, Hungary and Norway, there are 15 series of variation, which can permit to increase the probability to conclude for an increase or a decrease of the quality of schooling. There is a significant increase of schooling quality for 27 countries, where we observe no significant evolution for 23 countries and a significant decrease for 29 countries. Countries where the mean AAGR of quality of schooling is the highest are mainly developing countries. A strong increase is observed in Ghana (mean AAGR equal to +3.63%), Armenia (+1.86%) and Chile (+1.18%). The evolution is higher than +0.5% for 11 countries, including only 3 developed countries (Luxembourg, Germany and Finland). A significant decrease of schooling quality is present for 29 countries. The decrease is higher than -0.5% for 11 countries. All these countries are developing countries, excepting the case of Kuwait. Schooling quality tends to decrease in some African countries, such as Cameroon, Burkina Faso and Madagascar.

The comparison of quality of schooling between primary and secondary levels tends to show a high correlation: when we observe an increase in the primary education quality, trends in secondary level show an increase in quality too. Nevertheless, for some countries, contradictory evolution appears. While the primary education quality is increasing in countries like Islamic Republic of Iran, Italy or Japan, the quality of secondary schools tend to significantely decrease. In some countries, the evolution is more contrasted: we observe no significant change in secondary schools, when primary school quality show an improving (see for example the case of United Kingdom, Slovenia and Cyprus). This is quite interesting to notice that when the primary education quality is diminishing, the quality of secondary schools show a significant decrease too.

4. Comparisons with alternative measures of human capital stock

In this section, we compare our concept of human capital quality with alternative measures of the stock of human capital. Our estimates of international measures of cognitive skills provide a reasonable proxy for human capital quality. However, these data have a number of shortcomings. First, our measure of human capital quality does not take into account all dimensions of schooling quality, including for

instance skills in geography or foreign langage. Then, it could be interesting to compare our database with other works in this area. Second, the measure should be compared to human capital stock, in order to analyze the possible correlation between a quantitative and a qualitative view of the education sector. Third, as we present a methodology to check for the human capital quality trends, we propose to compare this evolution with trends in human capital stock.

4.1. Alternative measures of human capital quality

Three major research papers propose an alternative measure of human capital quality[6]. Firstly, Hanushek and Kimko (H&K) (2000), concerned with finding a better measurement of the quality of human capital, measure it using scores obtained from students participating in international assessments in science and mathematics. Starting from these test scores, they construct a unique (normalised) labour force quality measure for 31 countries covering the period from 1960 to 1990. They computed a quality measure for each country's labour force using the weighted average over all harmonised test scores where each country's weight is calculated as the normalised inverse of its standard error. They then performed a single cross-country regression for the 31 countries over the 1960-1990 period. The authors used the investigations of the IEA (International Association of the Evaluation of Educational Achievement) and of the IEAP (International Assessment of Educational Progress). In total, twenty-six series of educational performances were taken into account (by distinguishing according to age, the field of competence, namely mathematics and sciences, and year). Two series are available. The first series sets the world mean on each of the tests used equal to 50. The second series adjust all scores based on the U.S. international performance modified for the national time pattern of scores on the National Assessment of Educational Progress.

Another contribution to the field of study which directly includes the measurement of the quality of education in a model of growth is that of Barro (2001). These data come from the same sources as those of Hanushek and Kimko. Barro however builds different indicators according to skill, namely mathematics, sciences and reading. These indicators are available for one period only and are introduced into a panel regression. Because of the restricted number of countries for which qualitative indicators of education are available, the sample is smaller and involves only 43 countries. Moreover, the methodology used by Barro does not take into account for the possible differences in test variances across assessments. In final, Barro only scaled all possible data and computed the average.

Recently, Hanushek and Woessmann (H&W) (2009) extended measures from H&K (2000). They add new international tests, more countries and changed the methodology in order to make available the

[6] Other papers deal with this topic, however, we think that the databases proposed suffer from either lack of enough data or huge methodological issues. See for instance, Lee and Lee (1995), Ciccone and Papaioannou (2005) and Coulombe and Tremblay (2006)

possibility of tracking evolution in cognitive skills. The methodology used combine the adjustments in levels (based on the U.S. NAEP scores) and the adjustment in variances (based on the OECD Standardization Group). Then, they calculate standardized scores for all countries on all assessments. Each age group and subject is normalized to the PISA standard of mean 500 and individual standard deviation of 100 across OECD countries. Cognitive skills are measured by the simple average of all observed math and science scores between 1964 and 2003 for each country. Standardized scores are available for 77 countries.

The correlation between our data and database of H&K (2000) is high (r = 0.82 ; 57 observations). We include imputed data from H&K (2000) in this correlation. The comparison with the recent database of H&W (2009) is higher than the previous paper while it includes more countries (r = 0.89 ; 63 observations). However, as our database is larger than other databases, some countries are not included in this comparison. Figure 4 show the correlation between our database and the database of H&W (2009). Although a clear correlation appear, we tend to observe some significant differencies, in particular for countries having a lower score for human capital quality. This is due to the fact that we take into account more surveys than H&W, especially LLECE, PASEC and SACMEQ achievement tests. Informations used by H&W for countries such as India, Swaziland or Jordan are not recent and are only based on one single score of achievement tests in 1980s. On the contrary, our database include new scores for these countries (except for India), based on regional achievement tests.

4.2. Data on Educational Attainment

Three major databases are available for educational attainment. The first database is from Barro and Lee (2000), which is an update of a previous paper (Barro and Lee, 1993). Their paper presents a data set on educational attainment for the population over age 15 and over age 25 at five-year intervals between 1960 and 2000 for a broad number of countries. The data set comprises at least one observation for 142 economies, of which 109 have complete information at five-year intervals from 1960 to 2000. We take account of this variation by using information on the typical duration of each level of schooling within countries. With these data we can also construct measures of average years of schooling at all levels for each country. The main source of data is from Unesco's database on educational attainment, which is itself based on national censuses and sample surveys. In years when censuses or surveys are not available, Barro and Lee estimate the educational attainment using enrolment rates.

Cohen and Soto (2007) propose a new set of data on human capital. It is originally based upon data released at the OECD for a subgroup of 38 member and non-member countries. Cohen and Soto expanded this data set to other developing countries. The key of their methodology is to minimise extrapolations and keep the data as close as possible to those directly available from national censuses (in

the spirit of the work of De la Fuente and Domenech (2002, 2006) for OECD countries[7]). The data set consists of 95 countries and have been computed for the beginning of each decade from 1960 to 2000, plus a projection for 2010. This projection is based on population projections by age taken from the U.S. Census Bureau web site and the estimated of educational attainment for the year 2000.

The World Population Program of the International Institute for Applied Systems Analysis (IIASA) in Laxenburg, Austria and the Vienna Institute for Demography, Austrian Academy of Sciences developed the IIASA/VID educational attainment model for reconstruction of educational attainment distributions by age and sex from 1970 to 2000 and projections to 2050. This database is downloadable from the website of Edstats[8]. We use data from 1970 to 2005 by combining two models proposed by the IIASA/VID: the Reconstruction model and model of Projections. Projections have been carried out for 120 countries. More information concerning this database can be obtained in Lutz et al. (2007) and K.C. et al. (2008).

The correlation between our adjusted students' test scores and data on human capital quantity is not very high. The correlation of the Qualitative Indicators of Human Capital in Secondary Education with the Net Enrollment Ratio in Secondary level, for High Income Countries, in 1995 is 0.25 and concerns 25 observations. While some countries have approximately the same NER in secondary education, the QIHC tend to be different (figure 5). This is for instance the case of Hungary and Greece: the former has a higher level on schooling quality, while the NER in 1995 is equal in these countries. The comparison with database of B&L (2001) show a positive and quite high significance: the correlation is equal to 0.70 (72 observations). However, this correlation is mainly due to a "structural" effect, indicating that the more a country is developed higher is its human capital quality. When we restrict this comparison to High income countries, the correlation shrunk to 0.31 (34 observations). As the new data set of C&S appear to be more precise than the one of B&L, we present in the figure 6 its correlation with our data. When we include all countries, the correlation is very high ($r = 0.80$; 72 observations). However, for high income countries, the correlation decreases to 0.47 (27 observations). We found similar results when we compare our data with the IIASA/VID database. The correlation decrease from 0.74 to 0.20 when we include only High income countries (30 observations against 59 for the full sample).

Moreover, it appear interesting to compare the trends of years of schooling with quality of schooling variation between 1995 and 2007. Although data in available for only 33 countries, this correlation is slightly negative: $r=-0.20$. Figure 7 show this negative correlation, which implies that improving years of schooling does not significantly imply an increase of schooling quality. This is especially the case of some countries, such as Morocoo, Tunisia or Cuba. Hence, the absence of strong correlation between the basic

[7] We do not test the correlation between our data and the data base proposed by De la Fuente and Doménech, since the database of Cohen and Soto (2001) is quite similar and inlude more countries.

[8] Database can be downloaded at: http://go.worldbank.org/47P3PLE940.

quantitative indicator of human capital and our data tend to prove the importance of taking into account the qualitative dimension of human capital.

Conclusion

It is now common practice to measure the performance of education systems. The findings from international surveys of learning achievement are increasingly being used to measure the effectiveness of education. Of course, such surveys measure only one aspect of the quality of learning achievement. Yet they provide a solid foundation with which to do so.

The purpose of this paper is to devise an indicator for the measurement of learning achievement by grouping together all international and regional learning assessments. Where other research has focused solely on one survey, this paper has sought to group them together and thus obtain a more comprehensive range of data on pupils' performance. The methodology used is based largely on countries' participation in several surveys simultaneously, thus permitting the computation of inter-survey equivalence indices. Comparable indicators on learning achievement can thus be devised on the basis of a variety of surveys.

Bibliography

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

Altinok, N. (2009). An International perspective on trends in the quality of learning achievement (1965-2007). Background Paper for the Unesco GMR EFA 2009 Annual Report, Unesco, Paris.

Barro, R.J. (2001). Education and Economic Growth. in Helliwell, J.F. (Ed.), *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being* (pp. 14-41). Paris: OECD Press.

Barro, R.J., Lee, J.W. (1993). International Comparisons of Educational Attainment. *Journal of Monetary Economics*, 32, pp.363-394.

Barro, R.J., Lee, J.W. (2001). International Data on Educational Attainment: Updates and Implications. Center for International Development Working Paper no. 45, Harvard University.

Ciccone A., Papaioannou E. (2005). Human Capital, the Structure of Production, and Growth. Barcelona, Universitat Pompeu Fabra.

Cohen, D., Soto, M. (2007). Growth and Human Capital: Good Data, Good Results. *Journal Of Economic Growth*, 12(1), pp.51-76.

Coulombe, S., Tremblay, J.-F. (2006). Literacy and Growth, *Topics in Macroeconomics*, 6(2). Berkeley Electronic Press, http://www.bepress.com/bejm /topics/vol6/iss2/art4/.

De la Fuente, A., Domenech, R. (2002). Educational Attainment in the OECD, 1960-1995. *CEPR DP 3390*.

De la Fuente, A., Domenech, R. (2006). Human Capital in growth regression: How much difference does quality data make?. *Journal of the European Economic Association*, 4(1), pp.1-36.

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

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

Hanushek, E.A., Woessmann, L. (2009). Do Better School Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation. IZA Working Paper, No. 4575.

K.C., S., Barakat, B., Goujon, A., Skirbekk, V., Lutz, W. (2008). Projection of Populations by Level of Educational Attainment, Age and Sex for 120 Countries for 2005-2050. International Institute for Applied Systems Analysis, *Interim Report IR-08-038*, Austria.

Lazear, E.A. (2003). Teacher Incentives. *Swedish Economic Policy Review*, 10(3), pp. 179-214.

Lee, D.-W., Lee, T.-H. (1995). Human Capital and Economic Growth: Tests Based on the International Evaluation of Educational Achievement. *Economics Letters*, 47(2), pp. 219-225.

Leuven, E., Oosterbeek, H., van Ophen, H. (2004). Explaining International Differences in Male Skill Wage Differentials by Differences in Demand and Supply of Skill. *Economic Journal*, 114(495), pp. 466-486.

Lutz, W., Goujon, A., K.C., S., Sanderson, W. (2007). Reconstruction of Populations by Age, Sex and Level of Educational Attainment for 120 Countries for 1970-2000. International Institute for Applied Systems Analysis, *Interim Report IR-07-002*, Austria.

Mincer, J. (1974). *Schooling, Experience, and Earnings*. New York: National Bureau of Economic Research Press.

Mulligan, C.B. (1999). Galton versus the Human Capital Approach to Inheritance. *Journal of Political Economy*, 107(6), S184-S224.

Murname, R.J., Willet, J.B., Duhaldeborde, Y., Tyler, J.H. (2000). How Important Are the Cognitive Skills of Teenagers in Predicting Subsequent Earnings?. *Journal of Policy Analysis and Management*, 19(4), pp. 547-568.

Pritchett, L. (2001). Where has all the education gone?. *World Bank Economic Review*, 15(3), pp. 367-391.

Psacharopoulos, G., Patrinos, H. (2004). Returns to Investment in Education: A Further Update. *Education Economics*, 12(2), pp. 111-134.

Ross, K.N., Postlethwaite, T.N. (1991). *Indicators of the Quality of Education: A Study of Zimbabwean Primary Schools*. Harare: Ministry of Education and Culture; Paris: International Institute for Educational Planning.

Sakellariou, C. (2006). Cognitive Ability and Returns to Schooling in Chile. Background Paper for Vegas, E., Petrow, J. (Eds.) (2008), *Raising Student Learning in Latin America. The Challenge for the 21st Century*. World Bank. Latin American Development Forum series.

UNESCO (2004). *EFA Global Monitoring Report 2005: The Quality Imperative*. Paris.

Vegas, E., Petrow, J. (2008). *Raising Student Learning in Latin America. The Challenge for the 21st Century*. World Bank, Latin American Development Forum series.

.

APPENDIX A: A SYNOPSIS OF INTERNATIONAL AND REGIONAL TESTS ON EDUCATIONAL ACHIEVEMENT

A.1. International learning assessments

The International Association for the Evaluation of Educational Achievement (IEA) was the first body to measure individual learning achievement for international comparative purposes in the early 1960s. The surveys include the highly regarded "Trends in International Mathematics and Science Study" (TIMSS) and "Progress in International Reading Literacy Study" (PIRLS). Both will be highlighted in this paper.

The major survey series from IEA are the "Trends in International Mathematics and Science Study" (TIMSS). The central goal of TIMSS is to assess pupils' performance in both subjects and to describe the environment in which they learnt. With the second goal in view, those who launched TIMSS firmly took a policy-oriented approach, since the pupils' scores were correlated with the various factors involved in teaching. Four TIMSS surveys have been held to date: the first, in 1994-1995, covered 45 education systems and three groups of learners[9] (in grades 3 and 4, 7 and 8, and the final year of secondary education respectively); the second survey covered 38 education systems in 1999, examining grade 8 only; the third covered 50 education systems in 2003, focusing on grades 4 and 8; and finally, the fourth covered grades 4 and 8 and more than 66 education systems. The next round will be in 2011 and will cover 64 countries. The content of the questionnaires is quite varied and each topic is given a special weighting (examples are numbers, algebra and geometry in mathematics and life sciences, physical sciences and the history of science in science).

Other major IEA survey is the "Progress in International Reading Literacy Study", also known as PIRLS. Two major rounds of PIRLS have been done until 2009: in 2001 and in 2006. PIRLS survey tests grade 4 pupils in reading literacy. For instance, the 2006 PIRLS survey involved 41 countries or education systems, only two of which were African countries (Morocco and South Africa). In all, 15 developing countries took part in PIRLS 2006 (Bulgaria, Georgia, Indonesia, Islamic Republic of Iran, Latvia, Lithuania, The former Yugoslav Republic of Macedonia, Moldova, Morocco, Poland, Romania, Russian Federation, Slovakia, South Africa and Trinidad and Tobago). The next PIRLS round will be in 2011 and will include 54 countries, including 21 developing countries (Armenia, Azerbaijan, Bosnia-Herzegovina, Botswana, Bulgaria, Colombia, Egypt, Georgia, Honduras, Indonesia, Iran, Libya, Mongolia, Morocco, Oman, Poland, Romania, Russian Federation, Serbia, South Africa, and Ukraine).

This paper has drawn on virtually all IEA studies in the three skills. Data for 1964 to 1990 have been taken from Lee and Barro (2001) and the scores for other years from official reports (Harmon, Smith et al., 1997; Martin, Mullis et al., 2000; Mullis, Martin et al., 2000; Mullis, Martin et al., 2003; Mullis, Martin et al., 2004; Martin, Mullis et al., 2004; Mullis, Martin et al., 2007, Mullis, Martin et al., 2009).

[9] Some Canadian provinces or states in the United States of America have occasionally taken part in the IEA surveys. For the sake of simplicity, these regions are not included in the number of countries participating in the surveys.

The Organisation for Economic Co-operation and Development (OECD) launched its Programme for International Student Assessment (PISA) in 1997 to meet the need for data on student performance that would be readily comparable at international level. The key principles underlying PISA are the concept of "literacy", which concerns the pupils' capacity to extrapolate from what they have learnt and apply their knowledge in novel settings, its emphasis on lifelong learning and its regularity. More generally, PISA has assessed the skills of 15-year-olds every three years since 2000 in countries that together account for almost 90% of the global economy. PISA concentrates on three key areas, namely mathematics, science and reading literacy. Each PISA cycle focuses on one of these areas, thus gathering more information on the area assessed. The focus was on reading in 2000 and on mathematics in 2003. The third survey in the series was carried out in 2006 with science as the main field of assessment. In a way similar to recent IEA surveys, PISA is above all a monitoring tool: it evaluates pupils' knowledge and skills in mathematics, science and reading once every three years. The basic assessment model remains unchanged so that one cycle's findings can be compared to those of others[10]. The next PISA round is currently ongoing and results will be published at the end of 2010. Unlike the IEA surveys, PISA assesses only 15-year-olds, whatever their grade, whereas grade is the main criterion in selecting pupils for IEA assessments.

Drawing on the experience of the National Assessment of Educational Progress (NAEP), the International Assessment of Educational Progress (IAEP) comprises two surveys first conducted in 1988. IAEP statistical methodology and assessment procedures are derived from those of the NAEP, which became the main instrument for measuring the quality of learning achievement in the United States of America from 1970 onwards. As a result, the IAEP surveys have been strongly influenced by the American curriculum. The first one was held in 1988 and assessed 13-year-olds in two subjects (mathematics and science) in six countries. The second IAEP survey took place in 1991 and tested learners aged 9 and 13 in 19 countries in mathematics and science only. The results are all contained in Lee and Barro (2001).

Under a joint UNESCO and UNICEF project, learning achievement is being assessed as part of the Monitoring Learning Achievement (MLA) programme on a vast geographical scale in more than 40 countries with a view to building national assessment capacity (Chinapah, 2003). From data gathered at the end of primary education (the sixth year), countries should be able to identify factors that encourage or inhibit learning in primary school and to analyse problems, devise adjustments to education policies and suggest novel practices that will improve the quality of education. The scope of the more recent MLA II project has been expanded to include the first few grades of secondary education (grade 8). In comparison with the PASEC and SACMEQ surveys, which test only knowledge acquired at school, MLA also deals

[10] As it is explained in the PISA 2006 technical report, this is only the case for reading literacy between 2000, 2003 and 2006 and for science between 2003 and 2006. See methodological section for more details.

with practical and safety skills. In all, 72 countries have taken part in MLA assessments of achievement. However, all of the data have not been published. Supplementing national reports, a separate report on MLA I was drafted for 11 African countries (Botswana, Madagascar, Malawi, Mali, Morocco, Mauritius, Niger, Senegal, Tunisia, Uganda and Zambia; see UNESCO, 2000).

A.2. Regional learning assessments

Three major regional assessments have been conducted in Africa and Latin America.

The Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) grew out of a very extensive national investigation into the quality of primary education in Zimbabwe in 1991, supported by the UNESCO International Institute for Educational Planning (IIEP) (Ross and Postlethwaite, 1991). The 15 SACMEQ-member education ministries are those of Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, United Republic of Tanzania, United Republic of Tanzania (Zanzibar), Uganda, Zambia and Zimbabwe. The first SACMEQ round took place between 1995 and 1999. SACMEQ I thus covered seven different countries and assessed performance in reading at grade 6. In the second round, which was held between 2000 and 2002 and covered 14 countries and one territory (Zanzibar), performance in mathematics and reading was assessed. The target cohort consisted of grade 6 pupils, as under SACMEQ I. Several SACMEQ II items were replicated from the TIMSS survey to secure comparable results. The questionnaires were used to collect information on educational inputs, the educational environment and issues relating to the fair allocation of human and material resources. More generally, SACMEQ II included items selected from four previous surveys, namely the *Indicators of the Quality of Education* (Zimbabwe) study, SACMEQ I, TIMSS and the 1985-94 IEA *Reading Literacy Study*. As with TIMSS, the SACMEQ surveys provide estimates of mean national scores in tests. Data are available on the SACMEQ website (www.sacmeq.org). The third SACMEQ round (SACMEQ III) is covering the same countries as in 2002. However, data relative to this round is not yet available.

Surveys under the "Programme d'Analyse des Systèmes Educatifs" (PASEC, or "Programme of Analysis of Education Systems") of the Conference of Ministers of Education of French-Speaking Countries (CONFEMEN) have been conducted in the French-speaking countries of sub-Saharan Africa. In both CP2 (the second class in primary school) and CM1 (the fifth), between 2,000 and 2,500 young learners in about 100 schools, along with their teachers and school heads, have been surveyed in each of the 11 countries. Some countries have taken part in the PASEC survey on several occasions. The following is a list of participating countries in chronological order: Djibouti (1993-1994), Congo (1993-1994), Mali (1994-1995), Central African Republic (1994-1995), Senegal (1995-2000), Burkina Faso (1995-1998), Cameroon (1995-1996), Côte d'Ivoire (1995-1998), Madagascar (1997-1998), Guinea (1997-1998),

Togo (2000-2001), Mali (2001-2002), Niger (2001-2002), Chad (2003-2004), Mauritania (2003-2004), Guinea (2003-2004), Benin (2004-2005), Cameroon (2004-2005), Madagascar (2005-2006), Mauritius (2006), Congo (2006-2007), Senegal (2006-2007) and Burkina Faso (2006-2007). It should be noted that the Senegal (1995-2000) and Côte d'Ivoire (1995-1998) surveys were cohort follow-up assessments, whereas the others were "diagnostic surveys". Furthermore, the findings of the first four assessments are not available because the surveys were not conducted under acceptable circumstances. The distinctive feature of the survey procedures is that they comprise two assessments of achievement, the first at the beginning and the second at the end of the year. The surveys are unique in that they have assessed performance in terms of "added value" (see CONFEMEN, 2004). The data have been derived from national reports downloadable from the CONFEMEN website (http://www.confemen.org/).

The network of national education systems in Latin American and Caribbean countries, known as the Latin American Laboratory for Assessment of the Quality of Education (LLECE), was formed in 1994 and is coordinated by the UNESCO Regional Bureau for Education in Latin America and the Caribbean. Assessments conducted by the LLECE thus focused on learning achievement in reading and mathematics in grades 3 and 4 in 13 countries of the subcontinent (Casassus et al., 1998, 2002), namely Argentina, Bolivia, Brazil, Chile, Columbia, Costa Rica, Cuba, Dominican Republic, Honduras, Mexico, Paraguay, Peru and the Bolivarian Republic of Venezuela. Data for only 11 countries were collated in the official report (Casassus et al., 1998). In each country, samples of about 4,000 pupils in grade 3 (ages 8 and 9) and grade 4 (ages 9 and 10) were assembled. These surveys covered over 50,000 children, amounting to at least 100 classes per country. In 2006, the second part of the LLECE survey (named SERCE) was initiated in the same countries as LLECE I (. However, it differs from the latter in that it includes science in addition to mathematics and reading (only for 6 countries).

Data sources

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

Casassus, J., Froemel, J.E., Palafox, J.C., Cusato, S. (1998*). First International Comparative Study of Language, Mathematics, and Associated Factors in Third and Fourth Grades. First Report*. Santiago, Chile: Latin American Laboratory for Evaluation of the Quality of Education.

Casassus, J., Froemel, J.E., Palafox, J.C., Cusato, S. (2002). *First International Comparative Study of Language, Mathematics, and Associated Factors in Third and Fourth Grades. Second Report.* Santiago, Chile: Latin American Laboratory for Evaluation of the Quality of Education.

Harmon, M., Smith, T.A., Martin, M.O., Kelly, D.L., Beaton, A.E., Mullis, I.V.S., Gonzales, E.J., Orpwood, G. (1997). *Performance Assessment in IEA's Third International Mathematics and Science Study (TIMSS).* Boston: Boston College Press.

Lee, J.W., Barro, R.J. (2001). Schooling Quality in a Cross-Section of Countries. *Economica*, 38(272), pp. 465-488.

LLECE-UNESCO (2008). *Student Achievement in Latin America and the Caribbean. Results of the Second Regional Comparative and Explanatory Study (SERCE)*, Unesco, Paris.

Martin, O.M., Mullis, I.V.S., Gonzalez, E.J., Gregory, K.D., Garden, R.A., O'Connor, K., Chrostowski, S.J., Smith, T.A. (2000). *TIMSS 1999 International Science Report*. Boston: Boston College Press.

Martin M.O., Mullis, I.V.S., Gonzales, E.J., Chrostowski, S.J. (2004). *TIMSS 2003 International Science Report*. Boston: Boston College Press.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Gregory, K.D., Garden, R.A., O'Connor, K., Chrostowski, S.J., Smith, T.A. (2000). *TIMSS 1999 International Mathematics Report*. Boston: Boston College Press.

Mullis, I.V.S., Martin, M.O., Gonzalez, E.J., Kennedy, A.M. (2003). *PIRLS 2001 International Report: IEA's Study of Reading Literacy Achievement in Primary Schools*. Chestnut Hill, MA: Boston College.

Mullis, I.V.S., Martin, M.O., Gonzales, E.J., Chrostowski, S.J. (2004). *TIMSS 2003 International Mathematics Report*. Boston: Boston College Press.

Mullis, I.V.S., Martin, M.O., Kennedy, A.M., Foy, P. (2007). *IEA's Progress in International Reading Literacy Study in Primary School in 40 Countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., & Foy, P. (2009). *TIMSS 2007 International Mathematics Report: Findings from IEA's Trends in International Mathematics and Science Study at the Fourth and Eighth Grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

OECD (2001). *Knowledge and Skills for Life. First Results From the OECD Programme for International Student Assessment (PISA) 2000*. Organisation for Economic Co-operation and Development. Paris.

OECD (2004). *Learning for Tomorrow's World. First Results PISA 2003*. Organisation for Economic Co-operation and Development. Paris.

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

OECD, UNESCO Institute for Statistics. (2003). *Literacy Skills for the World of Tomorrow: Further Results from PISA 2000*. Paris and Montreal. OECD and UNESCO Institute for Statistics.

UNESCO (2000). *With Africa for Africa. Toward quality education for all*. UNESCO. Paris: Human Sciences Research Council.

Walker, D.A. (1976). *The IEA Six-Subject Survey: An Empirical Study of Education in Twenty-One Countries*. Stockholm: Almquist & Wiksell; New York: John Wiley & Sons.

Chinapah, V. (2003). Monitoring Learning Achievement (MLA) Project in Africa. *ADEA Biennal Meeting 2003*. Grand Baie, Mauritius, 3-6 December.

APPENDIX B: DETAILED METHODOLOGY FOR THE MAIN DATABASE

Below, we present the methodology used to obtain the main database for Qualitative Indicators of Human Capital. Some differences have been done since the first version of our database (Altinok and Murseli, 2007). The main change is the different procedure of adjustment of recent international and regional tests (such as PISA 2003, TIMSS 2003, 2007, PIRLS 2006 and LLECE 2006). See Stage 3 for more details. Moreover, we updated our database by including new data such as PISA 2006, PIRLS 2006, TIMSS 2007, LLECE 2006 and some new results from the PASEC survey.

Stage 1: Distinction between surveys

Under the method used here, some surveys are adjusted to the score obtained by the United States of America, which has been used as the particular country "base" for this purpose (called 'survey series A'). However, a specific procedure has been used in recent international surveys to avoid adjusting them all solely in accordance with performance in the United States of America by devising them in such a way as to permit analysis of trends in country scores over time (called 'survey series B'). Moreover, for some achievement tests, the USA did not participated. This is especially the case for regional surveys such as LLECE or SACMEQ. This 'survey series C' are taken into account by using countries' performance which participated both in regional and international surveys.

Stage 2: Series A surveys – the specific NAEP anchoring methodology

The methodology used in "series A" surveys involved recurrent anchoring of data to the American National Assessment of Educational Progress (NAEP) survey in the way described by Hanushek and Kimko (2000). The NAEP has been the main instrument used to measure the learning achievement of pupils in the United States of America since 1969, and the IAEP is its international equivalent. The assessment procedure is thus based on curriculum in the United States of America. At various times since 1970, pupils aged 9, 13 and 17 in the United States of America have been tested on their achievement in science, mathematics and reading. These tests may be regarded as an absolute benchmark of achievement levels in the United States of America. In order to process both IEA and IAEP survey data, Hanushek and Kimko (2000) used scores obtained in the United States of America as reference points. They thus altered the mean values obtained from the IEA surveys so that they equalled the closest NAEP mean values (for age, year and subject). More specifically, the procedure involves tracking the level of "difficulty" of a survey in accordance with the successive scores recorded in the United States of America. This adjustment has been made to all surveys conducted before 1995 in mathematics and science and before 2001 in reading. A different adjustment is made to surveys conducted after those years (see stage 3).

This kind of adjustment has been made to all surveys in which the United States of America took part and which were conducted no later than 1995 in the case of the IEA mathematics and science surveys in 2001 in regard to the reading surveys and 2000 in regard to PISA. Surveys held after those dates, namely TIMSS 1999, TIMSS 2003, PISA 2003, PISA 2006 and PIRLS 2006, have been adjusted on the basis of coefficients obtained from the 1995, 2000 or 2001 surveys, depending on the skill tested (we apply "the repeated NAEP anchoring methodology" for series B surveys).

Stage 3: Adjustment of series B surveys

If recent surveys such as PISA 2006 or PIRLS 2006 were to be adjusted according to the above procedure, all survey scores would be based on scores obtained in the United States of America. However, recent surveys have been designed to permit analysis of country score trends. Pupils are given test pieces at the same level of difficulty in survey series conducted by the same body, which makes it possible to analyse trends in pupil performance directly over time. Any adjustment of the series B surveys to the NAEP survey may thus distort the analysis of country performance trends. For example, if the United States of America score increases in the NAEP survey but decreases in another survey, such as PISA, the adjustment may lead to a fairly significant distortion. Yet the level of difficulty in pupil performance assessments may vary significantly from one type of survey to another – for example, marking under TIMSS may be stricter than under PISA. Any adjustment should thus be such as to result in comparable scores under both types of survey. Accordingly, the same adjustment coefficients as those calculated for the series A surveys are used in order to achieve a single linear conversion of country scores, and this procedure does not compromise the comparability of the scores obtained by countries participating in the same survey series. It is possible to compare trends in the United States of America's scores in the NAEP survey and in international surveys (see Table A.1.).

While trends in the United States of America are fairly similar in some of these surveys (as in PIRLS between 2001 and 2006), in others they are clearly contradictory. For example, while pupil attainment in primary-school mathematics appears to have increased by 5% between 1995 and 2003 according to the NAEP survey, exactly the opposite (a 5% decrease) seems to have occurred according to the TIMSS surveys for those years. Two explanations may account for such differences. The first is the likelihood of differences in the level of difficulty within a given survey series, which would warrant successive adjustment to the NAEP. A 5% fall in the United States of America's scores under the TIMSS primary education survey may be attributed mainly to the fact that the 2003 assessment was stricter than the 1995 one. Should this be the case, then an upward adjustment of all scores in TIMSS 2003 would be necessary (adjusting therefore as from NAEP 2003). The second explanation relates more to the likelihood of a difference in the measurement of achievement under NAEP and under international surveys. While the NAEP assessment of learning achievement is based on sound principles, such a difference may arise from the greater emphasis in international surveys on measuring generic skills and knowledge (and thus applicable to virtually all countries), while NAEP assesses skills and knowledge specific to the United

States of America. If there is any difference in the content of the test piece assessed, adjustments based only on results obtained in the United States of America may undermine the validity of the trends shown in the quality of learning achievement over time (while remaining, in principle, valid for the long term). This is so all the more because PISA focuses more on assessing skills, while NAEP assesses acquired knowledge and facts specific to the United States of America's education system. Unfortunately, the assessment questionnaires used for the various surveys cannot be examined in detail because they are not published by the bodies concerned.[11] While it is necessary to adjust surveys conducted until the 1990s to the NAEP survey, there is no need to adjust recent surveys because they have been designed to show the trend in country scores. The adjustment to latter survey series will therefore be similar to the one made to the first survey series concerned. The trend in country scores will thus be roughly similar whether the data compared are adjusted or not.

Table A.1. Comparison of trends in the scores obtained by the USA in the NAEP survey and in international surveys

Survey	Skill	Level	Year		Score		Evolution	
			Year 1	Year 2	Score 1	Score 2	Absolute	%
IEA Surveys								
TIMSS	Mathematics	Primary	1995	2003	54.5	51.8	-2.7	-5%
NAEP	Mathematics	Primary	1995	2003	44.8	47.0	2.2	5%
TIMSS	Science	Primary	1995	2003	56.5	53.6	-2.9	-5%
NAEP	Science	Primary	1995	2003	46.0	45.8	-0.2	0%
TIMSS	Mathematics	Secondary	1995	2003	50.0	50.4	0.4	1%
NAEP	Mathematics	Secondary	1995	2003	54.4	54.6	0.2	0%
TIMSS	Science	Secondary	1995	2003	53.4	52.7	-0.7	-1%
NAEP	Science	Secondary	1995	2003	51.2	51.2	0.0	0%
PIRLS	Reading	Primary	2001	2006	54.2	54.0	-0.2	0%
NAEP	Reading	Primary	2001	2006	43.6	44.0	0.4	1%
OECD surveys								
PISA	Mathematics	Secondary	2000	2006	49.3	47.4	-1.9	-4%
NAEP	Mathematics	Secondary	2000	2006	55.0	56.0	1	2%
PISA	Science	Secondary	2000	2006	49.9	48.9	-1.0	-2%
NAEP	Science	Secondary	2000	2006	51.2	51.2	0.0	0%
PISA	Reading	Secondary	2000	2003	50.4	49.5	-0.9	-2%
NAEP	Reading	Secondary	2000	2003	52.8	52.6	-0.2	0%

Source: For the NAEP findings, see *Conditions of Education* between 1993 and 2007. For international surveys, see section 3.

All surveys in which the United States of America has taken part after 1995 in the case of TIMSS, 2000 in the case of PISA and 2001 in the case of PIRLS have therefore been adjusted using the same adjustment coefficient as was used at those cut-off years. Greater accuracy is thus achieved in showing trends in the different countries' scores. Even if the scores are readjusted, the range of variation is very

[11] The main reason for not publishing the questionnaires is to leave open the possibility of using some of their items for several surveys, so that trends in the quality of learning achievement can indeed be studied over time.

slight. The results obtained from adjustments based on NAEP scores only are also given for comparative purposes. However, in this paper, only the results obtained by making similar adjustments to surveys of the same kind will be addressed.

It should be noted that the stage 3 adjustment is no more than a constant linear translation of the recent surveys. This means that the variance within surveys as well as the variation in country scores in separate surveys by the same body are exactly the same as those calculated on the basis of unadjusted data. This observation is crucial to a proper understanding of the need to examine short-term pupil performance trends in international learning achievement surveys. Once stage 3 has been completed, short-term trends in the quality of learning achievement can be analysed (see Section 7). The last two stages may be followed to make other analyses.

Stage 4 : Adjustment of series C surveys :

For the last group of surveys (called « Series C »), we can't use a simple anchoring on the NAEP, since the United States did not participated to some surveys, mainely the regional surveys, such as SACMEQ, LLECE, PASEC and MLA. Firstly, we compute a single score for each country which participated to these surveys between 0 and 100. Hence, we can compare results between surveys.

We selected countries which participated in at least two different surveys so as to establish a comparison between the surveys. The IEA surveys have been chosen as reference survey as they cover most of the countries and as the economic levels of participating countries is the most heterogeneous.

For instance, from the TIMSS indexes, we tried to track countries which had participated in the TIMSS and at least one other survey. We then proceeded a matching based on the means of the countries which participated in at two surveys (called "doubloons"); each survey was re-adjusted according to the level of comparison related to the referee survey of IEA (TIMSS).

APPENDIX C: International Database on Human Capital Quality

Countries	Abbreviation	Id	Primary	Secondary	General
Albania	ALB	2		42.73	
Algeria	DZA	3	47.92	41.10	41.26
Argentina	ARG	6	48.91	43.92	43.36
Armenia	ARM	7	57.58	49.54	51.30
Australia	AUS	8	60.29	60.44	58.78
Austria	AUT	9	61.12	59.04	58.48
Azerbaijan	AZE	10		44.75	
Bahrain	BHR	11		45.95	
Belgium (FI)	BEL	14	62.69	60.57	60.20
Belgium (FR)	BFL	15	57.38	56.76	55.13
Benin	BEN	16	34.81		
Bolivia	BOL	18	44.43		
Bosnia & Herzegovina	BIH	19		48.04	
Botswana	BWA	20	37.42	34.52	31.76
Brazil	BRA	21	48.36	40.53	41.20
Bulgaria	BGR	22	59.84	52.71	54.30
Burkina Faso	HVO	23	39.52		
Cameroon	CMR	26	51.07		
Canada	CAN	27	60.49	61.22	59.31
Chad	TCD	29	35.64		
Chile	CHL	30	47.06	43.31	41.98
China	CHN	31	43.55		
Chinese Tapei	TWN	32	63.26	63.81	62.29
Colombia	COL	33	49.13	41.09	41.94
Congo, Rep.	COG	36	39.77		
Costa Rica	CRI	37	50.68		
Cote d'Ivoire	CIV	38	39.04		
Croatia	HRV	39		57.30	
Cuba	CUB	40	58.99		
Cyprus	CYP	41	56.14	51.10	51.34
Czech Republic	CZE	42	60.72	59.44	58.47
Denmark	DNK	43	59.63	57.29	56.68
Dominican Republic	DOM	46	42.30		
Ecuador	ECU	47	45.25		
Egypt, Arab Rep.	EGY	48		41.96	
El Salvador	SLV	49	47.89	38.12	39.62
Estonia	EST	53		60.68	
Finland	FIN	56	64.80	62.32	62.34
France	FRA	57	60.10	59.39	58.10
Gabon	GAB	58	40.45		
Georgia	GEO	60	53.51	42.82	45.34

APPENDIX C (continued)

Countries	Abbreviation	Id	Primary	Secondary	General
Germany	DEU	61	60.71	58.15	57.77
Ghana	GHA	62		24.36	
Greece	GRC	63	56.72	53.72	53.10
Dominica	DMA	45			
Guatemala	GTM	65	44.67		
Honduras	HND	70	44.32		
Hong-Kong China	HKG	71	61.75	61.46	60.16
Hungary	HUN	72	61.60	62.77	60.79
Iceland	ISL	73	56.71	57.58	55.20
India	IND	74	39.61	21.65	25.96
Indonesia	IDN	75	46.30	43.43	41.62
Iran, Islamic Rep.	IRN	76	50.24	37.15	40.41
Ireland	IRL	78	59.43	58.76	57.38
Israel	ISR	80	59.70	54.04	54.95
Italy	ITA	81	61.63	55.98	57.10
Japan	JPN	83	63.98	66.10	63.95
Jordan	JOR	84	31.55	45.77	34.60
Kazakhstan	KAZ	85	63.04		
Kenya	KEN	86	40.64		
Korea, Rep.	KOR	89	66.94	66.12	65.62
Kuwait	KWT	90	45.31	41.08	39.78
Kyrgyz Republic	KGZ	91		36.08	
Latvia	LVA	93	60.96	56.23	56.86
Lebanon	LBN	94		44.65	
Lesotho	LSO	95	32.83		
Liechenstein	LIE	98		60.26	
Lithuania	LTU	99	61.12	55.00	56.27
Luxembourg	LUX	100	63.30	53.36	56.61
Macao China	MAC	101		60.00	
Macedonia, FYR	MKD	102	50.56	45.80	45.31
Madagascar	MDG	103	44.97		
Malawi	MWI	104	37.04	33.52	31.00
Malaysia	MYS	105		53.86	
Mali	MLI	107	36.92		
Mauritania	MRT	109	30.26		
Mauritius	MUS	110	39.99		
Mexico	MEX	112	48.06	46.98	44.55
Moldova	MDA	114	57.54	48.18	50.54
Morocco	MAR	116	42.89	37.48	36.45
Mozambique	MOZ	117	39.40	28.60	29.65
Namibia	NAM	119	32.13		
Netherlands	NLD	121	61.12	61.47	59.81
New Zealand	NZL	122	58.85	59.49	57.45
Nicaragua	NIC	123	46.19		

APPENDIX C (continued)

Countries	Abbreviation	Id	Primary	Secondary	General
Niger	NER	124	33.67		
Nigeria	NGA	125		38.57	
Norway	NOR	126	57.33	56.75	55.10
Oman	OMN	127	39.60	42.45	37.32
Palestinian Nat'I Auth	PSE	130		43.70	
Panama	PAN	131	46.21		
Paraguay	PRY	133	46.33		
Peru	PER	134	46.45	36.47	37.91
Philippines	PHL	135	48.12	38.97	40.21
Poland	POL	136	57.64	60.11	57.11
Portugal	PRT	137	54.98	55.03	52.84
Romania	ROU	139	55.95	48.98	50.08
Russian Federation	RUS	140	62.97	57.34	58.59
Saudi Arabia	SAU	144		38.72	
Scotland	SCO	145	59.36	56.21	55.94
Senegal	SEN	146	37.00		
Serbia	SCG	147		50.02	
Seychelles	SYC	148	39.72		
Singapore	SGP	150	62.55	63.45	61.69
Slovak Republic	SVK	151	59.89	57.33	56.86
Slovenia	SVN	152	59.52	58.90	57.50
South Africa	ZAF	155	37.37	25.15	26.62
Spain	ESP	156	60.36	55.99	56.39
Swaziland	SWZ	160	36.27	38.30	33.18
Sweden	SWE	161	61.35	58.87	58.52
Switzerland	CHE	162	57.56	60.61	57.34
Syrian Arab Republic	SYR	163		51.22	
Tanzania (Main Land)	TZA	165	39.16		
Tanzania (Zanzibar)	ZAN	166	37.53		
Thailand	THA	167	55.59	51.09	51.03
Togo	TGO	168	40.18		
Trinidad and Tobago	TTO	170	53.30	52.17	50.33
Tunisia	TUN	171	45.90	44.12	41.77
Turkey	TUR	172	51.78	49.46	47.99
Uganda	UGA	174	39.25		
Ukraine	UKR	175	56.78	49.94	51.07
United Kingdom	GBR	176	61.17	59.62	58.82
United States	USA	177	61.23	57.17	57.52
Uruguay	URY	178	49.43	50.17	47.06
Venezuela, RB	VEN	181	43.98	43.99	40.62
Yemen, Rep.	YEM	184	34.16		
Yugoslavia, Fed. Rep.	YUG	185		49.74	
Zambia	ZMB	186	34.53		
Zimbabwe	ZWE	187	37.09	39.86	34.49

Table 1. International Surveys of Learning Achievement since 1964

Year	Abbreviation	Skill	Countries	Age of pupils[a]	Grade tested[a]
1964	IEA-FIMS	M	12	13, Fin sec.	8,12
1970-72	IEA-SRC	R	15	10,14, Fin. Sec.	4,8,12
1970-72	IEA-FISS	S	19	10,14, Fin sec.	4,8,12
1982-83	IEA-SIMS	M	19	13, Fin sec.	7,12
1984	IEA-SISS	S	23	10,14, Fin sec.	4,8,12
1988	IAEP	M ; S	6	13	7
1991	IEA-RLS	R	32	9,14	5,8
1990/91	IAEP[a]	M ; S	19	9,13	5,8
1993-98	IEA-TIMSS	M ; S	45	9,13, Fin sec.	3,4,7,8,12
1992-97	UNESCO-MLA[b]	M ; S ; R	72	9,10,13	3,4,7
1997	UNESCO-LLECE	M ; R	11	9,10	3,4
1999	UNESCO-SACMEQ I	R	7	10	6
1999	IEA-TIMSS	M ; S	38	14	8
1995-2009	CONFEMEN-PASEC	M ; R	13	7,8,9,10	2,5
2000	OECD-PISA	M ; S ; R	43	15	7,8,9
2001	IEA-PIRLS	R	35	9,10	4
2002	UNESCO-SACMEQ II	M ; R	14	10	6
2003	IEA-TIMSS	M ; S	26,48	10,14	4,8
2003	OECD-PISA	M ; S ; R	41	15	7,8,9
2006	IEA-PIRLS	R	41	9-10	4
2006	OECD-PISA	M ; S ; R	57	15	7,8,9
2007[c]	UNESCO-SACMEQ III	M ; R	15	11-12	6
2006	UNESCO-LLECE II (SERCE)	M ; S ; R	17 or 6 for S	8, 11	3,4,6
2007	IEA-TIMSS	M ; S	66	10,14	4,8
2009[c]	OECD-PISA	M ; S ; R	66	15	7,8,9

Note: Fin sec. denotes the final year of secondary school. Abbreviations: IEA (International Association for the Evaluation of Educational Achievement); FIMS (First International Mathematics Study); SRC (The Study of Reading Comprehension); FISS (First International Science Study); SIMS (Second International Mathematics Study); SISS (Second International Science Study); IAEP (International Assessment of Educational Progress); TIMSS (Trends in International Mathematics and Science Study); PIRLS (Progress in International Reading Literacy Study); PISA (Programme for International Student Assessment); CONFEMEN (Conference of Ministers of Education of French-Speaking Countries); PASEC (Programme of Analysis of Education Systems of the CONFEMEN); SACMEQ (Southern and Eastern Africa Consortium for Monitoring Educational Quality).

(a) The age of pupils does not always correspond to that indicated in the table when studies are more concerned with their grade than with their age. Thus age within a given survey may vary between countries. Concerning the PISA survey, there is sometimes a grade variation, since this survey is based on the age of pupils.

(b) The number of countries taking part in the MLA survey varies according to its scope. Furthermore, not all countries have evaluated the three school levels. See Chinapah (2003) for further information.

(c) Studies for which data are not available when this paper was written (October 2009).

Source: See the text to identify the source of each study.

Table 2. International variation of schooling quality, by economic level

	Number of Countries	Mean	Standard Deviation	Minimum	Maximum
Primary level					
High-income countries	36	59.94	2.73	53.29	66.97
Middle-income countries	52	47.58	7.95	31.02	62.87
Low-income countries	18	37.10	3.08	32.46	43.35
Total	**106**	**50**	**10**	**31.02**	**66.97**
Secondary level					
High-income countries	39	58.81	3.46	51.10	66.12
Middle-income countries	47	44.82	7.60	21.66	60.11
Low-income countries	7	34.96	6.70	24.37	43.70
Total	**93**	**50**	**10**	**21.66**	**66.12**

Note : Country classification comes from the World Bank.

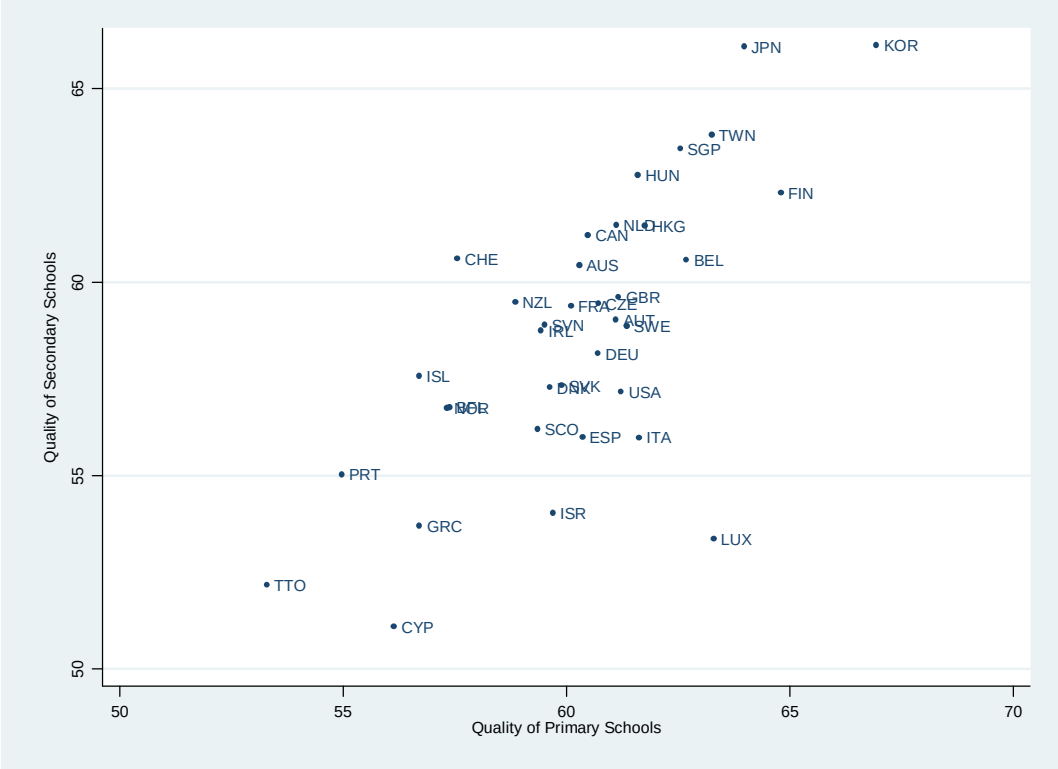

Figure 1. Quality in Primary and Secondary Schools, HIC (r = 0.73)

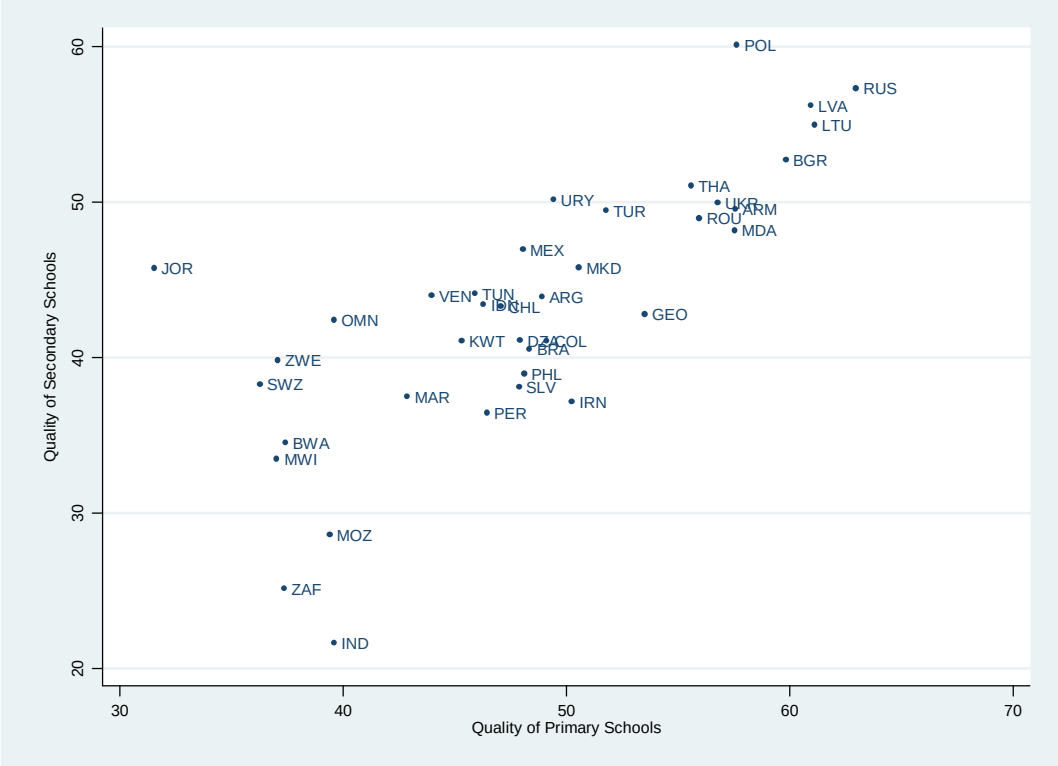

Figure 2. Quality in Primary and Secondary Schools, MIC & LIC (r = 0.76)

Table 3. Comparison of schooling education in primary schools (ISCED 1)

Regions	High Income Countries		Middle Income Countries		Low Income Countries	
Best scores by economic level	**Country**	**Score**	**Country**	**Score**	**Country**	**Score**
	Korea, Rep.	66.94	Kazakhstan	63.04	Madagascar	44.97
	Finland	64.80	Russian Federation	62.97	Kenya	40.64
	Japan	63.98	Lithuania	61.12	Togo	40.17
	Chinese Tapei (Taïwan)	63.26	Latvia	60.96	Burkina Faso	39.52
Lowest scores by economic level	**Country**	**Score**	**Country**	**Score**	**Country**	**Score**
	Trinidad & Tobago	53.30	Jordan	31.55	Mauritania	30.26
	Portugal	54.98	Namibia	32.13	Niger	33.67
	Cyprus	56.14	Lesotho	32.83	Yemen	33.24
	Iceland	56.71	Swaziland	36.27	Zambia	34.53

Note : Country classification comes from the World Bank.

Table 4. Comparison of schooling education in secondary schools (ISCED 2A)

Regions	High Income Countries		Middle Income Countries		Low Income Countries	
Best scores by economic level	**Country**	**Score**	**Country**	**Score**	**Country**	**Score**
	Korea, Rep.	66.12	Poland	60.11	Palestinian Nat'l Auth.	43.70
	Japan	66.10	Russian Federation	57.34	Zimbabwe	39.86
	Chinese Tapeï (Taïwan)	63.81	Croatia	57.30	Nigeria	38.57
	Singapore	63.45	Latvia	56.23	Kyrgyz Rep.	36.08
Lowest scores by economic level	**Country**	**Score**	**Country**	**Score**	**Country**	**Score**
	Cyprus	51.10	India	21.65	Ghana	24.36
	Trinidad & Tobago	52.17	South Africa	25.15	Mozambique	28.60
	Luxembourg	53.36	Botswana	34.52	Malawi	33.52
	Greece	53.72	Peru	36.47	n.a.	-

Note : Country classification comes from the World Bank. "n.a." means "not available"

Table 5. Long-Term Trends on Secondary Schools Quality

Countries	1970	1975	1980	1985	1990	1995	1999	2003	2007	Variation (1980-2007)		
										Abs.	%	AAGR
Chile	23.78	26.51	29.23	32.28	35.77	39.44	42.64	42.01	48.74	19.51	+67%	1.91
Iran, Islamic Rep.			30.32	34.76	40.07	46.24	45.69	45.33	45.19	14.87	+49%	1.49
Luxembourg			36.91	45.25	46.19	48.53	50.49	52.63	52.97	16.06	+43%	1.35
Hong-Kong China			44.49	54.34	55.75	56.89	58.63	60.15	58.07	13.58	+31%	0.99
Canada			46.88	56.28	56.46	54.43	55.96	56.58	57.28	10.40	+22%	0.75
Finland	48.97	51.77	54.57	55.88	55.96	55.68	55.45	58.54	59.81	5.24	+10%	0.37
Israel			44.35	55.26	57.34	53.78	49.12	51.78	48.96	4.60	+10%	0.34
New Zealand	56.20	54.98	53.76	53.52	53.60	53.19	52.60	53.27	56.81	3.05	+6%	0.20
Sweden	47.65	49.16	50.67	52.63	54.15	54.48	55.10	53.74	52.65	1.98	+4%	0.14
United States	52.79	52.42	52.05	52.00	52.50	52.80	53.46	54.17	54.04	1.98	+4%	0.14
Belgium	52.57	53.36	54.15	53.54	52.41	51.86	53.18	54.50	55.82	1.67	+3%	0.11
France			52.14	61.31	57.45	53.62	55.05	54.93	53.72	1.58	+3%	0.11
Thailand	37.15	41.90	46.66	50.15	52.41	53.99	49.87	49.65	47.89	1.23	+3%	0.10
United Kingdom	55.11	55.19	55.27	55.80	55.58	54.13	54.28	54.68	55.42	0.15	0	0.01
Italy	53.21	53.39	53.56	54.51	57.85	54.09	51.09	51.26	51.24	-2.32	-4%	-0.16
Australia	58.14	57.49	56.85	56.21	55.58	54.96	55.98	54.22	53.13	-3.72	-7%	-0.25
Netherlands	47.51	54.74	61.96	62.75	59.43	56.26	57.06	56.39	57.14	-4.82	-8%	-0.30
Hungary	62.88	61.90	60.91	61.55	61.33	55.90	56.96	56.35	55.49	-5.42	-9%	-0.34
Japan	80.26	75.40	70.54	66.45	63.05	59.80	59.46	59.06	59.16	-11.38	-16%	-0.65

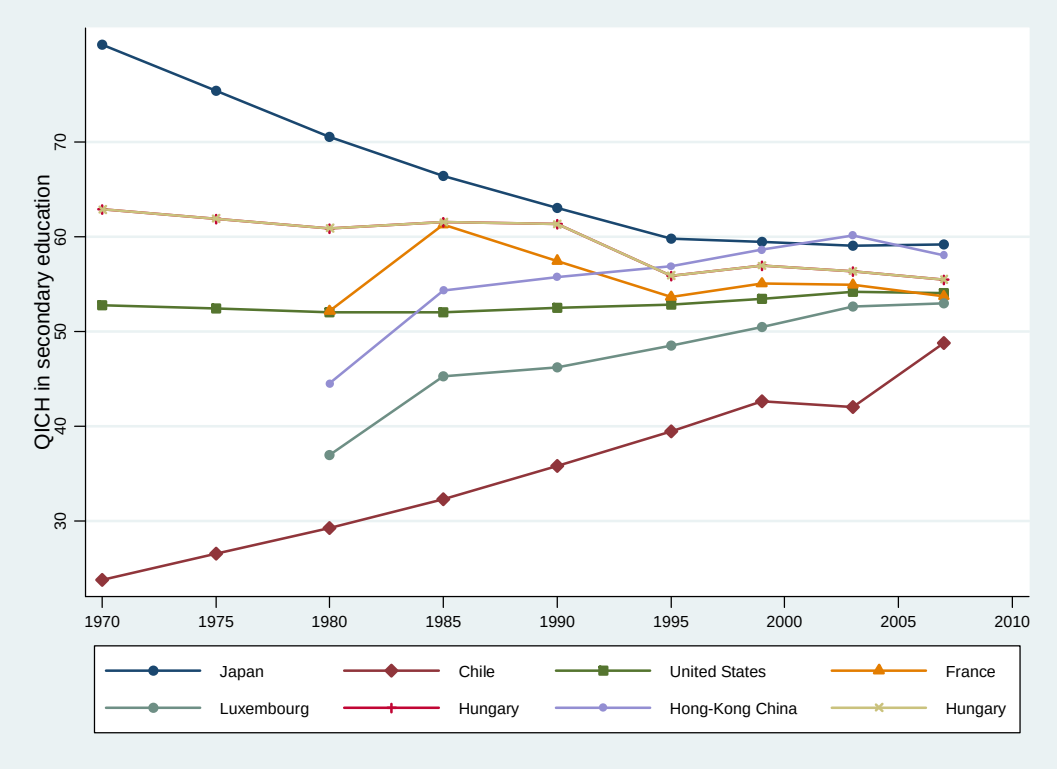

Figure 3. Long term trends on secondary schools quality, selected countries

Table 6. Short term trends on secondary schools quality (combined mathematics and science), 1995-2007, 48 countries

Countries	1995	1999	2003	2007	Variation (1995-2007)		
					Absolute	Relative (%)	AAGR (%)
Brazil	35.85	35.88	44.67	44.41	8.56	23.88	1.80
Chile	39.44	42.64	42.01	48.74	9.30	23.58	1.78
Lithuania	49.22	51.00	53.66	53.88	4.65	9.46	0.76
Luxembourg	48.53	50.49	52.63	52.97	4.44	9.15	0.73
Latvia	50.76	53.02	53.64	54.76	4.00	7.89	0.63
Belgium (FR)	51.86	53.18	54.50	55.82	3.96	7.64	0.62
Finland	55.68	55.45	58.54	59.81	4.13	7.43	0.60
Denmark	51.02	54.39	53.63	54.79	3.77	7.39	0.60
Jordan	43.96	46.12	47.15	47.12	3.15	7.17	0.58
New Zealand	53.19	52.60	53.27	56.81	3.62	6.80	0.55
Iceland	50.92	54.81	54.52	54.15	3.23	6.35	0.51
Colombia	39.53	40.29	41.06	41.82	2.29	5.80	0.47
Canada	54.43	55.96	56.58	57.28	2.85	5.24	0.43
Portugal	48.55	51.27	50.30	51.00	2.45	5.05	0.41
Switzerland	54.88	55.44	56.02	56.51	1.63	2.97	0.24
Germany	53.61	53.56	54.05	55.12	1.51	2.81	0.23
United Kingdom	54.13	54.28	54.68	55.42	1.28	2.37	0.20
United States	52.80	53.46	54.17	54.04	1.24	2.34	0.19
Korea, Rep.	59.34	59.85	60.41	60.61	1.26	2.13	0.18
Hong-Kong China	56.89	58.63	60.15	58.07	1.18	2.08	0.17
Netherlands	56.26	57.06	56.39	57.14	0.88	1.57	0.13
Spain	51.82	53.09	52.29	52.45	0.63	1.22	0.10
Kuwait	40.33	40.36	40.40	40.44	0.11	0.26	0.02
France	53.62	55.05	54.93	53.72	0.10	0.18	0.02
Ireland	54.53	54.84	54.17	54.59	0.06	0.10	0.01
Cyprus	48.41	49.28	47.39	48.27	-0.14	-0.29	-0.02
Greece	50.73	51.03	49.51	50.53	-0.20	-0.40	-0.03
Russian Federation	55.03	55.48	53.74	54.76	-0.27	-0.50	-0.04
Hungary	55.90	56.96	56.35	55.49	-0.42	-0.75	-0.06
Japan	59.80	59.46	59.06	59.16	-0.64	-1.06	-0.09
Scotland	52.28	52.68	53.09	51.68	-0.61	-1.17	-0.10
Belgium (FI)	56.98	57.55	55.44	55.82	-1.16	-2.04	-0.17
Austria	56.17	55.50	53.78	54.96	-1.20	-2.15	-0.18
Iran, Islamic Rep.	46.24	45.69	45.33	45.19	-1.05	-2.26	-0.19
Romania	49.69	49.65	49.72	48.55	-1.15	-2.31	-0.19
Singapore	62.61	61.74	62.29	61.09	-1.52	-2.43	-0.21
Poland	56.99	57.67	54.59	55.58	-1.41	-2.47	-0.21
Slovenia	56.35	55.90	53.21	54.55	-1.80	-3.20	-0.27
Australia	54.96	55.98	54.22	53.13	-1.84	-3.34	-0.28
Sweden	54.48	55.10	53.74	52.65	-1.84	-3.38	-0.29
Slovak Republic	56.07	56.23	53.89	53.18	-2.89	-5.16	-0.44
Italy	54.09	51.09	51.26	51.24	-2.85	-5.27	-0.45
Czech Republic	57.86	55.65	55.74	54.77	-3.09	-5.34	-0.46
Norway	53.23	54.30	50.14	50.23	-3.00	-5.64	-0.48
Israel	53.78	49.12	51.78	48.96	-4.82	-8.97	-0.78
Thailand	53.99	49.87	49.65	47.89	-6.10	-11.30	-0.99
Bulgaria	56.63	54.10	50.23	49.11	-7.52	-13.29	-1.18
Syrian Arab Republic	52.57	49.69	46.97	44.40	-8.17	-15.54	-1.40

Table 7. Short term trends on schooling quality, Mean variations

Country	Primary			Secondary			Total (Primary + Secondary education)					
	Score	Mean	Trend	Score	Mean	Trend	Score	Mean	Num-ber	Posi-tive	Nega-tive	Trend
Ghana				24.36	3.63	+++		3.63	2	2	0	+++
Armenia	57.58	2.46	+++	49.54	1.26	+++	51.30	1.86	4	4	0	+++
Chile	47.06	0.96	+++	43.31	1.33	+++	41.98	1.18	5	5	0	+++
Lebanon				44.65	1.11	+++		1.11	2	2	0	+++
Luxembourg	63.30			53.36	0.89	+++	56.61	0.89	6	5	1	+++
Mexico	48.06	1.61	+++	46.98	0.61	+++	44.55	0.86	8	6	2	+++
Indonesia	46.30			43.43	0.75	+++	41.62	0.75	8	6	2	+++
Poland	57.64			60.11	0.71	+++	57.11	0.71	6	6	0	+++
Germany	60.71	0.33	+	58.15	0.65	+++	57.77	0.60	7	7	0	+++
Finland	64.80			62.32	0.56	+++	62.34	0.56	6	6	0	+++
Colombia	49.13	0.61	+++	41.09	0.47	++	41.94	0.54	4	4	0	+++
Denmark	59.63			57.29	0.45	++	56.68	0.45	6	4	2	++
Latvia	60.96	0.38	++	56.23	0.48	++	56.86	0.43	11	9	2	++
Paraguay	46.33	0.39	++					0.39	2	2	0	++
Hong-Kong China	61.75	0.94	+++	61.46	0.11	oo	60.16	0.38	15	11	4	++
Iran, Islamic Rep.	50.24	0.79	+++	37.15	-0.14	--	40.41	0.38	9	6	3	++
Russian Federation	62.97	0.95	+++	57.34	0.12	oo	58.59	0.32	13	9	4	++
Portugal	54.98			55.03	0.31	++	52.84	0.31	6	4	2	++
Switzerland	57.56			60.61	0.28	++	57.34	0.28	6	6	0	++
Chinese Tapei	63.26	0.40	++	63.81	0.06	ooo	62.29	0.23	4	3	1	++
Lithuania	61.12	-0.09	oo	55.00	0.43	+++	56.27	0.21	7	5	2	++
Korea, Rep.	66.94			66.12	0.20	++	65.62	0.20	10	7	3	++
Tunisia	45.90	-0.29	oo	44.12	0.41	++	41.77	0.13	5	4	1	++
Moldova	57.54	0.32	+	48.18			50.54	0.32	1	1	0	+
Kenya	40.64	0.12	+					0.12	1	1	0	+
Ireland	59.43			58.76	0.08	+	57.38	0.08	6	4	2	+
Hungary	61.60	-0.01	ooo	62.77	0.05	+	60.79	0.03	15	10	5	+
Saudi Arabia				38.72	0.04	ooo		0.04	2	1	1	ooo
Austria	61.12	-0.30	--	59.04	0.15	oo	58.48	0.03	8	3	5	ooo
Belgium (FI)	62.69			60.57	0.03	ooo	60.20	0.03	6	3	3	ooo
United Kingdom	61.17	0.24	++	59.62	-0.12	oo	58.82	0.03	12	7	5	ooo
New Zealand	58.85	-0.07	ooo	59.49	0.10	o	57.45	0.02	11	4	7	ooo
United States	61.23	0.15	ooo	57.17	-0.09	o	57.52	0.00	13	8	5	ooo
Jordan	31.55			45.77	-0.02	ooo	34.60	-0.02	2	1	1	ooo
Turkey	51.78			49.46	-0.04	ooo	47.99	-0.04	3	2	1	ooo
Brazil	48.36	0.04	ooo	40.53	0.59	+++	41.20	0.45	8	5	3	oo
Slovenia	59.52	0.58	o	58.90	0.18	ooo	57.50	0.41	9	5	4	oo
Liechenstein				60.26	0.30	oo		0.30	6	3	3	oo
Cyprus	56.14			51.10	0.22	oo	51.34	0.22	4	2	2	oo
Singapore	62.55	0.67	+++	63.45	-0.35	--	61.69	0.22	9	5	4	oo
Canada	60.49			61.22	0.19	oo	59.31	0.19	6	3	3	oo
Greece	56.72			53.72	0.14	oo	53.10	0.14	6	3	3	oo
Serbia				50.02	-0.07	ooo		-0.07	5	3	2	oo
Italy	61.63	0.37	+	55.98	-0.16	oo	57.10	-0.10	9	4	5	oo
Japan	63.98	0.07	+	66.10	-0.34	--	63.95	-0.22	14	6	8	oo
Spain	60.36			55.99	-0.32	oo	56.39	-0.32	6	3	3	oo
Israel	59.70	0.12	+	54.04	-0.43	--	54.95	-0.36	8	3	5	oo
Argentina	48.91	-0.14	oo	43.92	-0.56	o	43.36	-0.39	5	3	2	oo
Bahrain				45.95	0.71	-		0.71	2	1	1	o
Macedonia, FYR	50.56	0.00	o	45.80			45.31	0.00	1	0	1	o

Table 7 (continued)

Country	Primary			Secondary			Total (Primary + Secondary education)					
	Score	Mean	Trend	Score	Mean	Trend	Score	Mean	Num-ber	Posi-tive	Nega-tive	Trend
Tanzania (Zanzibar)	37.53	-0.45	-					-0.45	1	0	1	-
Mauritius	39.99	-0.51	-					-0.51	1	0	1	-
Uruguay	49.43			50.17	-0.63	-	47.06	-0.63	1	0	1	-
Namibia	32.13	-1.04	-					-1.04	1	0	1	-
Malawi	37.04	-1.50	-	33.52			31.00	-1.50	1	0	1	-
Zambia	34.53	-1.62	-					-1.62	1	0	1	-
Norway	57.33	0.25	oo	56.75	-0.27	--	55.10	-0.10	15	5	10	--
Slovak Republic	59.89	0.50	+	57.33	-0.32	--	56.86	-0.11	4	1	3	--
Australia	60.29	0.22	ooo	60.44	-0.26	--	58.78	-0.12	14	4	10	--
Scotland	59.36	-0.03	ooo	56.21	-0.39	--	55.94	-0.19	9	2	7	--
Netherlands	61.12	-0.18	--	61.47	-0.21	oo	59.81	-0.19	8	1	7	--
Czech Republic	60.72	-0.57	---	59.44	-0.17	--	58.47	-0.25	10	2	8	--
Iceland	56.71	-0.04	-	57.58	-0.30	--	55.20	-0.26	7	2	5	--
Sweden	61.35	-0.43	-	58.87	-0.27	--	58.52	-0.29	11	2	9	--
Botswana	37.42			34.52	-0.41	--	31.76	-0.41	2	0	2	--
Thailand	55.59			51.09	-0.47	--	51.03	-0.47	8	0	8	--
France	60.10	-0.11	-	59.39	-0.53	---	58.10	-0.47	7	1	6	--
Romania	55.95	-0.92	-	48.98	-0.39	--	50.08	-0.49	5	0	5	--
Morocco	42.89	-0.87	---	37.48	-0.01	ooo	36.45	-0.52	5	1	4	---
Bulgaria	59.84	-0.11	-	52.71	-0.73	---	54.30	-0.65	8	0	8	---
Egypt, Arab Rep.				41.96	-0.86	---		-0.86	2	0	2	---
Cameroon	51.07	-0.87	---					-0.87	4	1	3	---
Cuba	58.99	-1.03	---					-1.03	2	0	2	---
Dominican Republic	42.30	-1.30	---					-1.30	2	0	2	---
Kuwait	45.31	-1.58	---	41.08			39.78	-1.58	3	0	3	---
Palestinian Nat'I				43.70	-1.67	---		-1.67	2	0	2	---
Malaysia				53.86	-1.84	---		-1.84	2	0	2	---
Burkina Faso	39.52	-1.96	---					-1.96	4	0	4	---
Madagascar	44.97	-3.61	---					-3.61	4	0	4	---

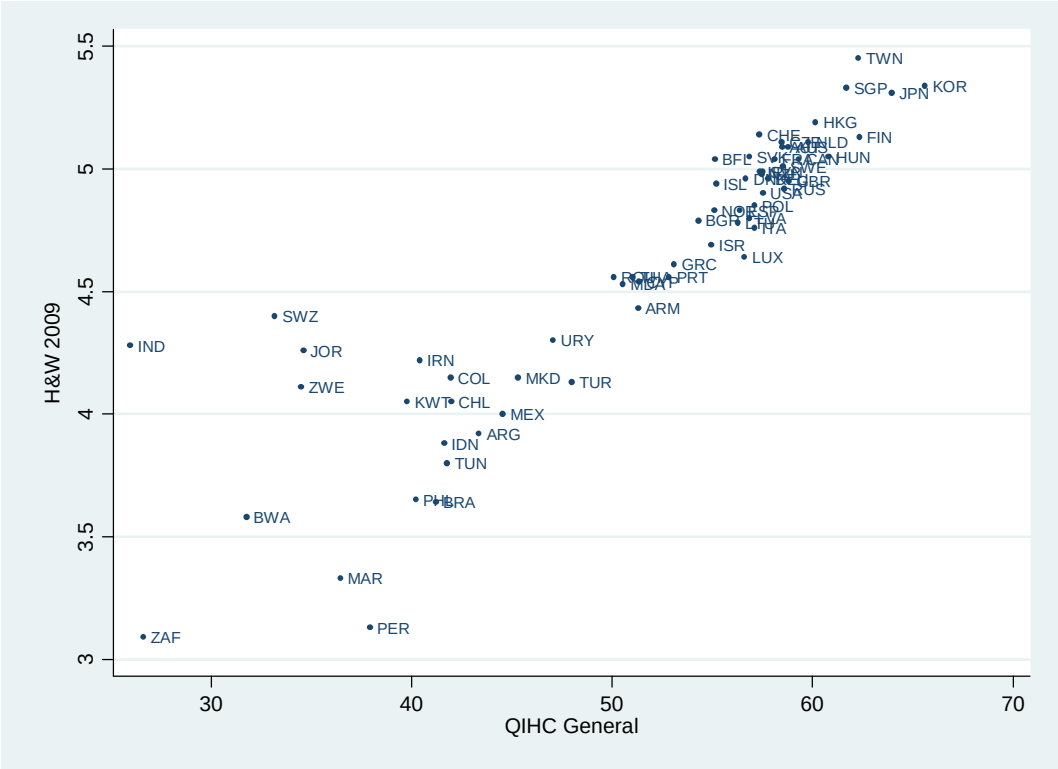

Figure 4. Hanushek and Woessman (2009) vs our data (R^2 = 0,89 ; 63 Countries)

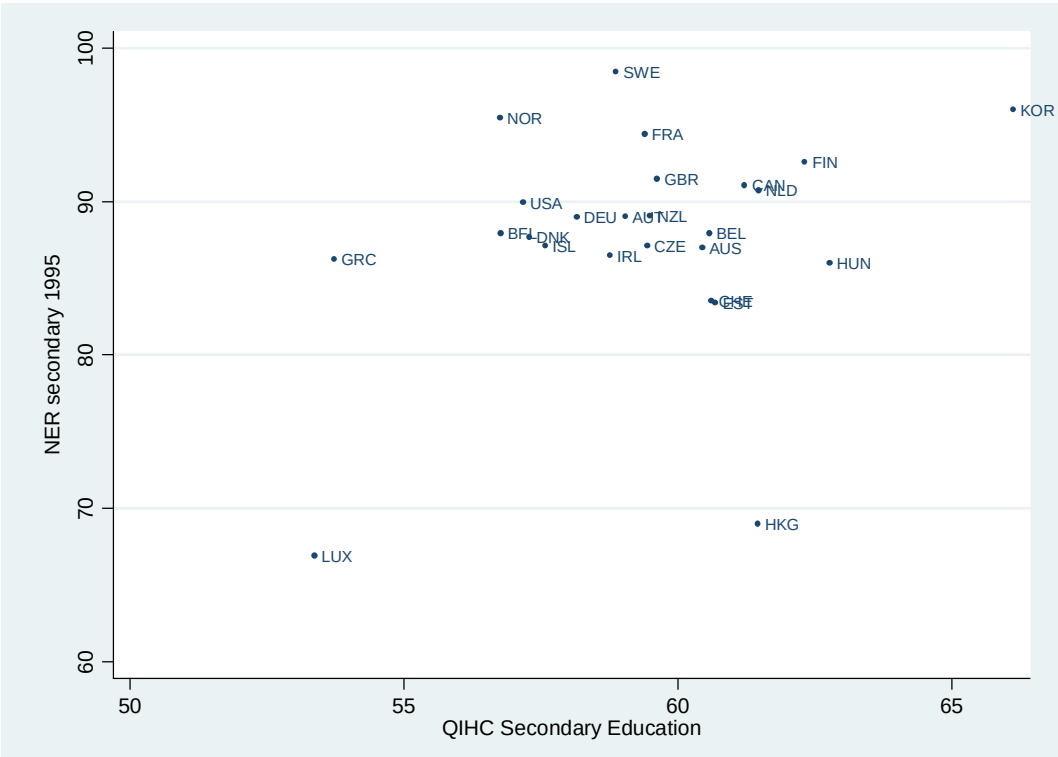

Figure 5. NER in Secondary (1995) vs our data for secondary education, High Income Countries (R^2 = 0,28 ; 25 Observations)

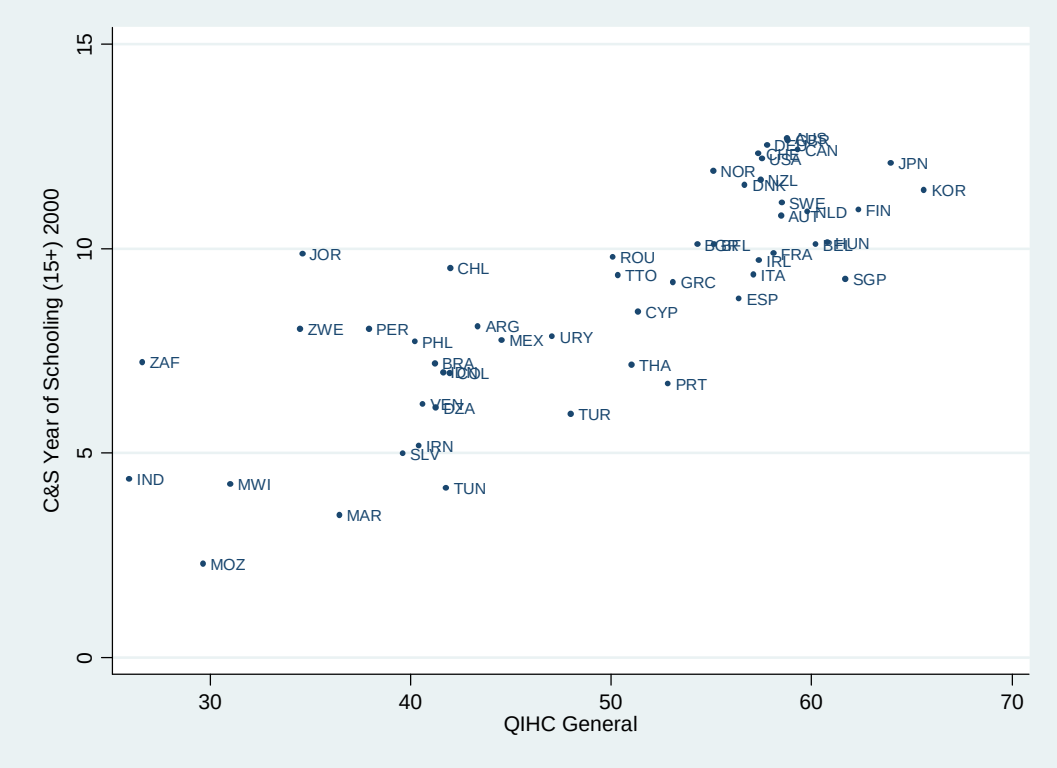

Figure 6. Cohen and Soto data vs. our data ($R^2 = 0{,}80$; 72 Observations)

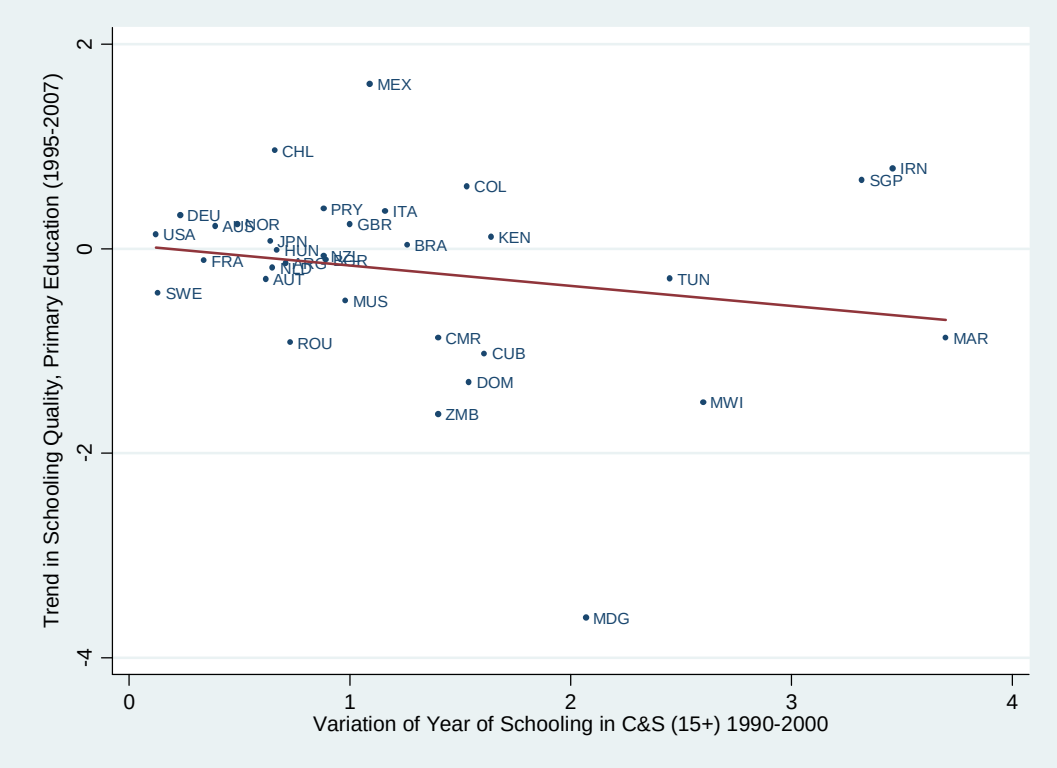

Figure 7. Variation in years of schooling and trends in schooling quality ($R^2 = -0{,}20$; 33 Observations)