The High Cost of Low Educational Performance

THE LONG-RUN ECONOMIC IMPACT OF IMPROVING PISA OUTCOMES



Programme for International Student Assessment

OECD

The High Cost of Low Educational Performance

THE LONG-RUN ECONOMIC IMPACT OF IMPROVING PISA OUTCOMES

Programme for **I**nternational **S**tudent **A**ssessment



ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where the governments of 30 democracies work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, the Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The Commission of the European Communities takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

This work is published on the responsibility of the Secretary-General of the OECD. The opinions expressed and arguments employed herein do not necessarily reflect the official views of the Organisation or of the governments of its member countries.

ISBN 978-92-64-07748-5 (PDF)

Photo credits: Cover © Jack Hollingsworth/Photodisc/Getty Images.

Corrigenda to OECD publications may be found on line at: *www.oecd.org/publishing/corrigenda*.

PISA™, OECD/PISA™ and the PISA logo are trademarks of the Organisation for Economic Co-operation and Development (OECD). All use of OECD trademarks is prohibited without permission from the OECD.

© OECD 2010

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgment of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre français d'exploitation du droit de copie (CFC) at *contact@cfcopies.com*.



Foreword

In response to the need for cross-nationally comparable evidence on student performance, the Organisation for Economic Co-operation and Development (OECD) launched the OECD Programme for International Student Assessment (PISA) in 1997. PISA represents a commitment by governments to monitor the outcomes of education systems in terms of student achievement on a regular basis and within an internationally agreed common framework. It aims to provide a new basis for policy dialogue and for collaboration in defining and implementing educational goals, in innovative ways that reflect judgements about the skills that are relevant to adult life.

Results of the three-yearly PISA surveys reveal wide differences in the performance of education systems in terms of the learning outcomes achieved by students. For some countries, the results from PISA are disappointing, showing that their 15-year-olds' performance lags considerably behind that of other countries, sometimes by the equivalent of several years of schooling and sometimes despite high investments in education. However, PISA also shows that other countries are very successful in achieving strong and equitable learning outcomes. Moreover, some countries have been able to significantly improve their learning outcomes, in the case of Poland by almost three-quarters of a school year between 2000 and 2006 alone.

This report uses recent economic modelling to relate cognitive skills – as measured by PISA and other international instruments – to economic growth. The relationship indicates that relatively small improvements in the skills of a nation's labour force can have very large impacts on future well-being. Moreover, the gains, put in terms of current Gross Domestic Product (GDP), far outstrip the value of the short-run business-cycle management. This is not to say that efforts should not be directed at issues of economic recession, but it is to say that the long-run issues should not be neglected.

The report was written by Prof. Eric. A. Hanushek from the Hoover Institution at Stanford University and CES ifo and by Prof. Ludger Woessmann from the Ifo Institute for Economic Research, CES ifo, and the University of Munich, in consultation with members of the PISA Governing Board as well as Andreas Schleicher, Romain Duval and Maciej Jakubowski from the OECD Secretariat. The report was produced by the Indicators and Analysis Division of the OECD Directorate for Education and is published on the responsibility of the Secretary-General of the OECD.



Table of contents

Foreword 3

Executive summary 5

Introduction 6

The effect of education on economic growth 11
- Empirical growth analyses using school attainment data 13
- Empirical growth analyses considering cognitive skills 13
- Measuring cognitive skills 14
- Basic facts of cognitive skills and economic growth 16
- Evidence on causality 17

The economic costs of low educational achievement 20
- Simulation approach 20
- Scenario I: Increase average performance by 25 PISA points 22
- Scenario II: Bring each country to the average level of Finland 24
- Scenario III: Bring everyone up to a minimum skill level of 400 PISA points 25

Policy conclusions 27

References 29

Notes 31

Annex A: Deriving a combined historical measure of cognitive skills 35

Annex B: Underlying cross-country growth regressions 45

Annex C: Projection of the economic value of education reforms 49

Notes to Annexes 52



Executive summary

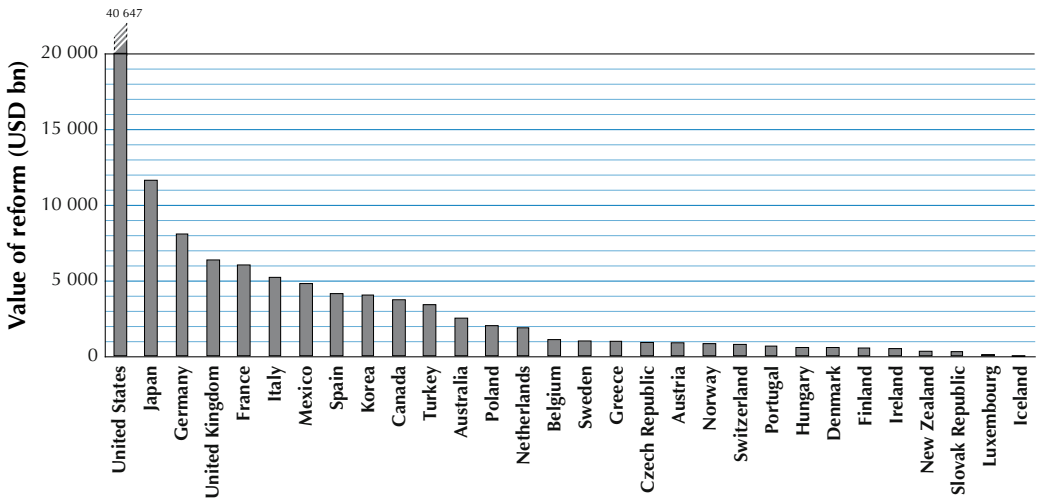
While many nations express a commitment to improved educational quality, education often slips down on the policy agenda. Because the benefits of educational investments are seen only in the future, it is possible to underestimate the value and the importance of improvements.

This report uses recent economic modelling to relate cognitive skills – as measured by PISA and other international instruments – to economic growth. This relationship indicates that relatively small improvements in the skills of a nation's labour force can have very large impacts on future well-being. Moreover, the gains, put in terms of current GDP, far outstrip today's value of the short-run business-cycle management. This is not to say that efforts should not be directed at issues of economic recession, but it is to say that the long-run issues should not be neglected.

A modest goal of having all OECD countries boost their average PISA scores by 25 points over the next 20 years – which is less than the most rapidly improving education system in the OECD, Poland, achieved between 2000 and 2006 alone – implies an aggregate gain of OECD GDP of USD 115 trillion over the lifetime of the generation born in 2010 (as evaluated at the start of reform in terms of real present value of future improvements in GDP) (Figure 1). Bringing all countries up to the average performance of Finland, OECD's best performing education system in PISA, would result in gains in the order of USD 260 trillion (Figure 4). The report also shows that it is the quality of learning outcomes, not the length of schooling, which makes the difference. Other aggressive goals, such as bringing all students to a level of minimal proficiency for the OECD (*i.e.* reaching a PISA score of 400), would imply aggregate GDP increases of close to USD 200 trillion according to historical growth relationships (Figure 2).

Figure 1

Present value of Scenario I (improve student performance in each country by 25 points on the PISA scale) in billion USD (PPP)



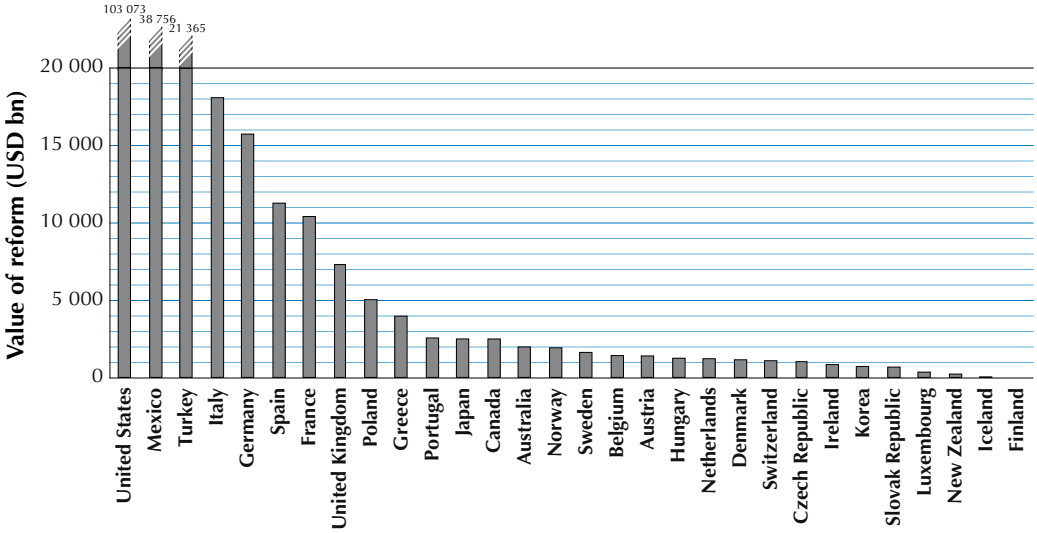
Note: Discounted value of future increases in GDP until 2090 due to reforms that improve student performance in each country by 25 points on PISA, or by ¼ standard deviation, expressed in billion USD (see also Table 1).

© OECD 2010



Figure 2

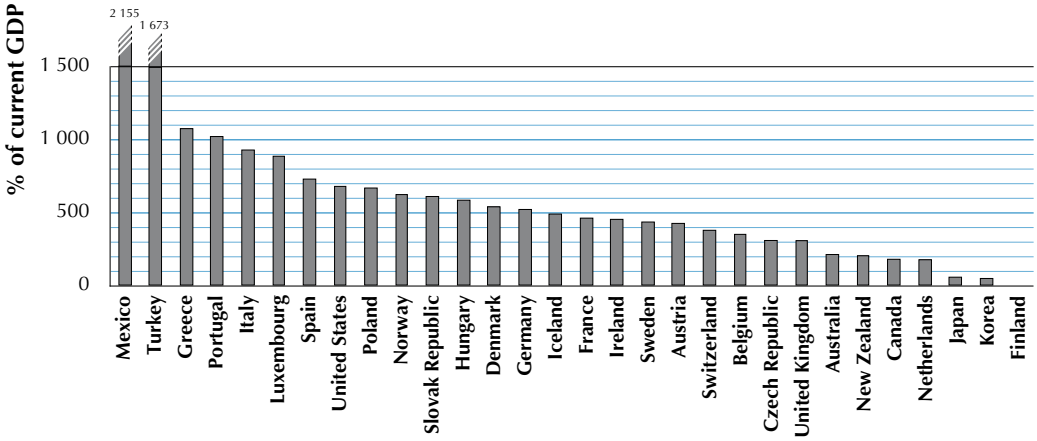
Present value of Scenario II (improve student performance in each country to reach the level achieved by Finland, the country with the highest performance in PISA) in billion USD (PPP)



Note: Discounted value of future increases in GDP until 2090 due to reforms that improve student performance in each country to reach the level achieved by Finland, at 546 points on the PISA 2000 scale (average of mathematics and science in 2000, 2003 and 2006), expressed in billion USD (see also Table 2).

Figure 3

Present value of Scenario II (improve student performance in each country to reach the level achieved by Finland, the country with the highest performance in PISA) in percent of current GDP

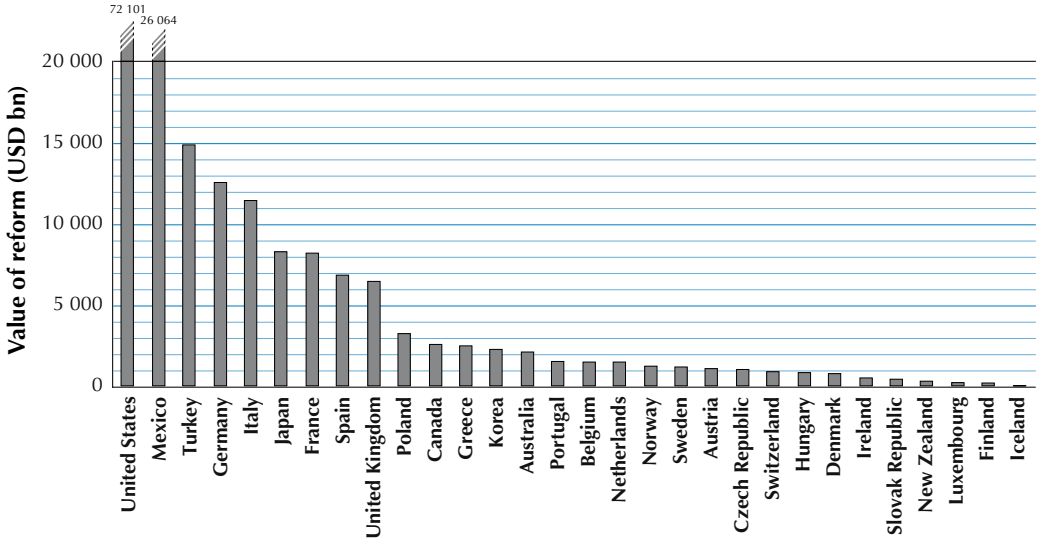


Note: Discounted value of future increases in GDP until 2090 due to a reform that improves student performance in each country to reach the level achieved by Finland, at 546 points on the PISA 2000 scale (average of mathematics and science in 2000, 2003 and 2006), expressed as percentage of current GDP (see also Table 2).



Figure 4

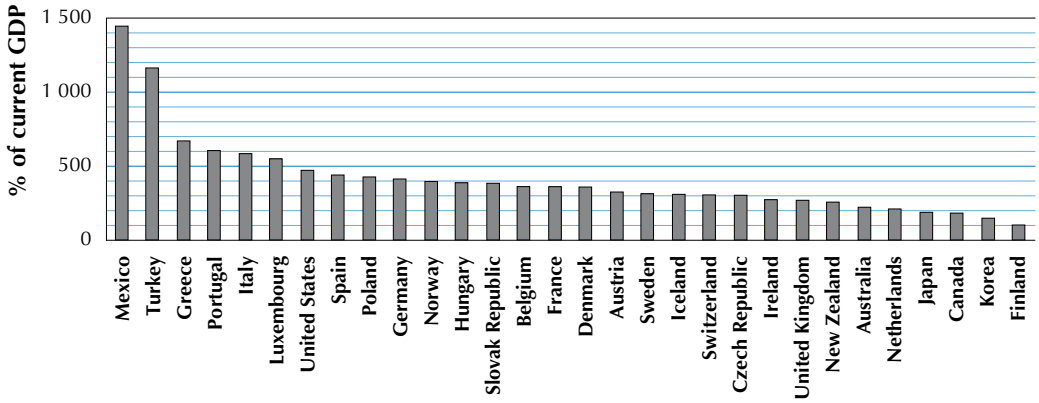
Present value of Scenario III (ensure that all students perform at a minimum of 400 points on the PISA scale) in billion USD (PPP)



Note: Discounted value of future increases in GDP until 2090 due to a reform that ensures that all students perform at a minimum of 400 points on the PISA scale, expressed in billion USD (see also Table 3).

Figure 5

Present value of Scenario III (ensure that all students perform at a minimum of 400 points on the PISA scale) in percent of current GDP



Note: Discounted value of future increases in GDP until 2090 due to a reform that ensures that all students perform at a minimum of 400 points on the PISA scale), expressed as percentage of current GDP (see also Table 3).

© OECD 2010



There is uncertainty in these projections as there is in all projections. The first issue is whether the statistical models used to characterise OECD growth between 1960 and 2000 accurately reflect the underlying determinants of growth. Economists disagree about the most appropriate way to model economic growth, and these estimates are based upon the specific form of endogenous growth models. Moreover, the execution of the estimation, including the measurement of cognitive skills and the allowance for other growth factors, incorporates additional elements of uncertainty. The second issue is the economic reality that is being projected. These projections trace the economy for 80 years into the future. A changing impact of cognitive skills on technological change and economic growth would clearly directly affect the specific estimates (although there is little reason to presume that it is more likely that the role of cognitive skills decreases as opposed to increases). Similarly, the present value of improved growth depends on the general health and growth of individual economies, which again is simply projected according to the historic patterns of the OECD nations. Other details, including how heavily future incomes are discounted and the time span for the calculations, also enter the specific projections.

Nonetheless, even reducing the projections substantially to allow for plausible minimal estimates suggests very large implications of improved cognitive skills and human capital. If the estimated impacts of cognitive skills were twice as large as the true underlying causal impact on growth, the resulting present value of successful school reform still far exceeds any conceivable costs of improvement.

Changing schools and educational institutions is, of course, a difficult task. Moreover, countries that have attempted reforms of schools have often found that the results in terms of student achievement are relatively modest. At the same time, the results from countries achieving high and equitable learning outcomes in PISA – like Finland in Europe, Canada in North America or Japan and Korea in East Asia – or from those that have seen rapid improvements in the quality of schooling (like Poland) underline that doing better is possible. Concluding that change is "too difficult" would imply foregoing enormous gains to the well-being of OECD nations.



INTRODUCTION

Nations around the world seek to improve their schools in order to enhance the skills and employability of their youth or to reduce inequalities in economic outcomes found within their societies. But there are also countervailing forces, because changing schools is politically difficult. If the gains from change are not too large, it may not make sense to politicians and decision-makers to take any moves unpopular with the existing educational establishment. The evidence presented here, however, indicates that this reasoning is flawed. The potential gains from improving schools within developed countries appear truly enormous.

This conclusion draws upon prior work delving into the determinants of economic growth. Work over the past two decades on why some countries have succeeded economically while others have not now provides a much clearer picture of the role of human capital in economic development. The human capital influence on growth is best characterised by the relationship between direct measures of cognitive skills and long-term economic development. The evidence points to differences in cognitive skills as an explanation of a majority of the differences in economic growth rates across OECD countries. Moreover, the available historical evidence indicates that differential skills have a very powerful and continuing impact.

The historical growth relationships provide a means for projecting out how improvements in schools would translate into economic results. Based on the historic patterns, it is possible to estimate both the time pattern and the ultimate impact of school quality improvements. The performance deficits of countries, measured by average scores on PISA tests and other international tests of mathematics and science, identify serious shortfalls in economic performance relative to economic possibilities.

An important aspect highlighted by these projections is the dynamic nature of human capital and growth. The basic characterisation of growth indicates that higher cognitive skills offer a path of continued economic improvement, so that favourable policies today have growing impacts in the future. The underlying idea is that economies with more human capital (measured by cognitive skills) innovate at a higher rate than those with less human capital, implying that nations with larger human capital in their workers keep seeing more productivity gains. Characterising the full ramifications of schooling outcomes requires tracing future developments far into the future.

These projections do not indicate how schools should be changed. Nor do they solve the political-economy issues of how any change should be achieved politically. They simply show the cost of inaction.

Projections into the future of course contain uncertainties. Important facets of uncertainty come both from structural aspects of world economies and from potential analytical problems. On the structural side, all of the projections assume that historical economic growth patterns of OECD countries from 1960 to 2000 give a good indication of how growth will evolve across the 21st century. On the analytical side, the projections rest on a series of statistical models that are assumed to accurately characterise the underlying factors that are most important for national economic growth – but it is difficult to rule out fully alternative explanations for the observed differences in growth across countries.

Obviously, learning does not end with school and an assessment of schooling outcomes such as PISA cannot reflect the skills which individuals acquire subsequently in their lives. However, evidence from longitudinal surveys in Australia, Canada and Denmark shows that performance in PISA is a strong and consistent predictor for subsequent educational experiences.

Plausible scenarios for educational improvement yield estimates of economic impact reaching USD 260 trillion when aggregated across OECD countries, suggesting that the importance of action must be taken seriously. Even if the historic patterns of growth are not fully realised in the future or even if the estimates of how cognitive skills affect growth are too optimistic, the potential impact on economic well-being remains

© OECD 2010



enormous. If the true impact of improved cognitive skills in the OECD countries were just half of these estimates, the impact on the well-being of OECD societies would still be considerably larger than, for example, the aggregate gains from smoothing out all future business cycles.

The next section describes analyses of human capital and economic growth. It provides a short review of alternative models of economic growth and links these to the progression of empirical analysis. From this, it summarises the best evidence on how cognitive skills affect economic growth. The third section then presents a series of simulations for OECD countries of various amounts of improvement in cognitive skills. The fourth section concludes and discusses some of the political-economy issues involved. Technical issues, including the development of measures of cognitive skills and the underlying statistical modelling, are presented in annexes.

THE EFFECT OF EDUCATION ON ECONOMIC GROWTH

At any point in time, attention to economic policies that deal with current demand conditions and with business cycles always seems to take priority. Perhaps this has never been as true as in 2009, when the most obvious focus of attention was the worldwide recession. Without minimising the need to deal with current unemployment conditions, the message of this analysis is that considering issues of longer-run economic growth may be more important for the welfare of nations. Nobel Laureate Robert Lucas, in his presidential address to the American Economic Association, concluded that "Taking performance in the United States over the past 50 years as a benchmark, the potential for welfare gains from better long-run, supply-side policies exceeds by far the potential from further improvements in short-run demand management" (Lucas, 2003).

Economists have considered the process of economic growth for much of the last 100 years, but most studies remained as theory with little empirical work.[1] Over the past two decades, economists linked analysis much more closely to empirical observations and in the process rediscovered the importance of growth.

The analysis here particularly concentrates on the role of human capital. Human capital has been a central focus of much of the recent growth modelling, and it is a standard element of any empirical work. Its importance from a policy perspective is clear and unquestionable.

The prior analytical work has nonetheless diverged in important ways. Economists have developed a number of alternative models designed to highlight important determinants of economic growth. These theoretical views about the determinants of growth have gone in a variety of directions (see Box 1). Two aspects of theoretical investigations stand out for the discussion here. First, each of the approaches suggests different empirical specifications for any statistical modelling. Second, while each of the approaches has some conceptual appeal, it has been difficult to test the validity of the alternatives in an adequate manner. The restricted variation of experiences across countries plus general data limitations have made it difficult to distinguish among the competing models of growth – and such is the case here.

This analysis adopts a general "endogenous growth" framework, for both conceptual and data reasons. Specifically, in this formulation nations with more human capital tend to continue to make greater productivity gains than nations with less human capital.[2] The fact that the rate of technological change and productivity improvement is directly related to the stock of human capital of the nation makes it an endogenous growth model. The relationship between cognitive skills on the one hand and innovations and technology on the other seems to be a natural view of the role of education. At the same time, it is impossible within available data to test this approach against alternatives that do not have this linkage.



Box 1. **Theories of economic growth**

Theoretical models of economic growth have emphasised different mechanisms through which education may affect economic growth. As a general summary, three theoretical models have been applied to the modelling of economic growth, and each has received support from the data. At the same time, it has been difficult to compare the alternative models and to choose among them based on the economic growth data.

The most straightforward modelling follows a standard characterisation of an aggregate production function where the output of the macro economy is a direct function of the capital and labour in the economy. The basic growth model of Solow (1957) began with such a description and then added an element of technological change to get the movement of the economy over time. The source or determinants of this technological change, although central to understanding economic growth, were not an integral part of the analysis. Augmented neoclassical growth theories, developed by Mankiw, Romer and Weil (1992), extend this analysis to incorporate education, stressing the role of education as a factor of production. Education can be accumulated, increasing the human capital of the labour force and thus the steady state level of aggregate income. The human capital component of growth comes through accumulation of more education that implies the economy moves from one steady state level to another; once at the new level, education exerts no further influence on growth. The common approach to estimating this model is to relate changes in GDP per worker to changes in education (and capital). This view of the role of human capital is fairly limited, because there are natural constraints on the amount of schooling in which a society will invest. It also fails to explain patterns of education expansion and growth for many developing countries (cf. Pritchett, 2006).

A very different view comes from the "endogenous growth" literature that has developed over the past two decades. In this work, a variety of researchers (importantly, Lucas, 1988, Romer, 1990a and Aghion and Howitt, 1998) stress the role of education in increasing the innovative capacity of the economy through developing new ideas and new technologies. These are called endogenous growth models because technological change is determined by economic forces within the model. Under these models, a given level of education can lead to a continuing stream of new ideas, thus making it possible for education to affect growth even when no new education is added to the economy. The common way to estimate these models is to relate changes in GDP per worker (or per capita) to the level of education.

A final view of education in production and growth centres on the diffusion of technologies. If new technologies increase firm productivity, countries can grow by adopting these new technologies more broadly. Theories of technological diffusion such as Nelson and Phelps (1966) , Welch (1970), and Benhabib and Spiegel (2005) stress that education may facilitate the transmission of knowledge needed to implement new technologies. In tests involving cross-country comparisons, Benhabib and Spiegel (1994) find a role for education in both the generation of ideas and in the diffusion of technology.

All approaches have in common that they see education as having a positive effect on growth. The latter two stress its impact on long-run growth trajectories.



Empirical growth analyses using school attainment data

The macroeconomic literature focusing on cross-country differences in economic growth has overwhelmingly employed measures related to school attainment, or years of schooling, to test the predictions of growth models.[3] Initial analyses employed school enrolment ratios (*e.g.* Barro, 1991; Mankiw, Romer and Weil, 1992; Levine and Renelt,1992) as proxies for the human capital of an economy. An important extension by Barro and Lee (1993, 2001) was the development of internationally comparable data on average years of schooling for a large sample of countries and years, based on a combination of census and survey data.

The vast literature of cross-country growth regressions has tended to find a significant positive association between quantitative measures of schooling and economic growth.[4] To give an idea of the robustness of this association, an extensive empirical analysis by Sala-i-Martin, Doppelhofer, and Miller (2004) of 67 explanatory variables in growth regressions on a sample of 88 countries found that primary schooling was the most robust influence factor on growth in GDP per capita in 1960-96 (after allowing for the faster growth in East Asia).[5]

However, average years of schooling are a particularly incomplete and potentially misleading measure of education for comparing the impacts of human capital on the economies of different countries. It implicitly assumes that a year of schooling delivers the same increase in knowledge and skills regardless of the education system. For example, a year of schooling in Kyrgyzstan (the country with the lowest performance in the PISA 2006 science assessment) is assumed to create the same increase in productive human capital as a year of schooling in Finland (the country with the highest performance in the PISA 2006 science assessment).[6] Additionally, this measure assumes that formal schooling is the primary (sole) source of education and, again, that variations in non-school factors have a negligible effect on education outcomes. This neglect of cross-country differences in the quality of education and in the strength of family, health, and other influences is probably the major drawback of such a quantitative measure of schooling.

Empirical growth analyses considering cognitive skills

Over the past ten years, empirical growth research demonstrates that consideration of cognitive skills dramatically alters the assessment of the role of education and knowledge in the process of economic development. Using data from international student achievement tests, Hanushek and Kimko (2000) demonstrate a statistically and economically significant positive effect of cognitive skills on economic growth in 1960-90. Their estimates suggest that one country-level standard deviation higher test performance would yield around one percentage point higher annual growth rates. The country-level standard deviation is equivalent to 47 test-score points in the PISA 2000 mathematics assessment. Again, in terms of the PISA 2000 mathematics scores, 47 points would be roughly the average difference between Sweden and Japan (the best performer among OECD countries in 2000) or between the average Greek student and the OECD average score. One percentage point difference in growth is itself a very large value, because the average annual growth of OECD countries has been roughly 1.5%.

Their estimate stems from a statistical model that relates annual growth rates of real GDP per capita to the measure of cognitive skills, years of schooling, the initial level of income and a wide variety of other variables that might affect growth including in different specifications the population growth rates, political measures, or openness of the economies. The general concern, described in more detail below, is that things other than human capital are the real causes of some or all of the observed growth and that ignoring them in the statistical analysis artificially inflates the importance of cognitive skills. One solution to this is inclusion of the other factors in the statistical model.

The relationship between cognitive skills and economic growth has now been demonstrated in a range of studies. As reviewed in Hanushek and Woessmann (2008), these studies employ measures of cognitive skills that draw upon the international testing of PISA and of TIMSS (Trends in International Mathematics and



Science Study) (along with earlier versions of these).[7] The uniform result is that the international achievement measures provide an accurate measure of the skills of the labour force in different countries and that these skills are closely tied to economic outcomes.[8]

This analysis follows in the lines of more recent work that measures human capital by cognitive skills instead of school attainment.

Measuring cognitive skills

For the analysis here, a central issue is construction of a measure of the skills of a nation's workforce that can be matched with economic outcomes. With few exceptions, however, direct measures of achievement of individuals in the labour force are unavailable, and analysis instead must rely upon skills measured during the schooling period.[9] The analytical approach that underlies this analysis is to combine data from international tests given over the past 45 years in order to develop a single comparable measure of skills for each country that can be used to index skills of individuals in the labour force. As discussed below, this construction causes no problems if the relative performance of individuals in different countries has remained constant, but it could introduce problems if that is not true.

While the PISA tests are now well-known throughout the OECD, the history of testing is less understood. Between 1964 and 2003, 12 different international tests of mathematics, science, or reading were administered to a voluntarily participating group of countries (see Annex Tables A1 and A2). These include 36 different possible scores for year-age-test combinations (*e.g.* science for students of grade 8 in 1972 as part of the First International Science Study or mathematics of 15-year-olds in 2000 as a part of the Programme on International Student Assessment). Only the United States participated in all possible tests.

The assessments are designed to identify a common set of expected skills, which were then tested in the local language. It is easier to do this in mathematics and science than in reading, and a majority of the international testing has focused on mathematics and science. Each test is newly constructed, until recently with no effort to link to any of the other tests. While the analysis here focuses on mathematics and science, these scores are highly correlated with reading test scores and employing just mathematics and science performance does not distort the growth relationship that is estimated; see Hanushek and Woessmann (2009). Also, the narrower measures do not imply that other skills are irrelevant, only that they tend to be closely related to mathematics and science skills at the national level.

The goal here is construction of consistent measures at the national level that will allow comparing performance across countries, even when they did not each participate in a common assessment. This section sketches the methodology. The details of this construction along with the final data are presented in Annex A.

The method of construction of aggregate country scores employed here focuses on transformations of the means and variances of the original country scores in order to put each into a common distribution of outcomes.[10]

Test-score levels across assessments. Comparisons of the difficulty of tests across time are readily possible in the United States because the country has participated in all assessments and because there is external information on the absolute level of performance of students in the United States of different ages and across subjects. The United States began consistent testing of a random sample of students around 1970 under the National Assessment of Educational Progress (NAEP). By using the pattern of NAEP scores for the United States over time, it is possible to equate student performance in the United States across each of the international tests.

Test-score variance across assessments. Each assessment has varying country participation and has different test construction so that the variance of scores for each assessment cannot be assumed to be constant. The

© OECD 2010



Box 2. **Empirical growth models with cognitive skills**

The endogenous growth models that are estimated here relate growth rates of GDP per capita to the initial level of GDP per capita, the years of school attainment, and the level of cognitive skills measured by mathematics and science scores on available international exams. Inclusion of initial income reflects the fact that lower income countries just have to imitate more developed countries and will find this easier than innovating with new products, technologies, or production techniques (often called conditional convergence).

The basic estimation employs a sample of 23 OECD countries for which appropriate economic data are available for the period of 1960-2000. (Hanushek and Woessmann (2008) provide estimates for an expanded sample of 50 countries that are very similar to those presented below).

Ideally, one would want the level of test performance for the workers in the economy, and not just the test performance of students who range in age from roughly 10-18 years old. The analysis assumes that the average scores observed for students are a good proxy of labour-force skills. This assumption would clearly be satisfied if the educational outcomes within countries remain roughly constant. There is some indication that this is not the case (see Figure 6), which would tend to introduce some error into these measures. Nonetheless, in one set of tests, scores before 1984 are linked to growth from 1980-2000, thus getting the timing closer to ideal, and the estimated effects are somewhat larger than found for the full period (Hanushek and Woessmann, 2009). In general, this kind of measurement error will tend to lead to estimates of the impact of skills that is biased downward.

The basic model estimated for the 23 OECD countries is:

$$g = \underset{(2.0)}{-3.54} - \underset{(5.8)}{0.30}\, GDP/capita_{1960} + \underset{(4.2)}{1.74}\, C + \underset{(0.3)}{0.025}\, S \qquad R^2=0.83$$

where g is the average annual growth rate in GDP per capita between 1960 and 2000, $GDP/capita_{1960}$ is initial national income, C is the composite measure of cognitive skills, and S is years of schooling (measured in 1960, but qualitative results are the same when measured as average over 1960-2000). Absolute values of t-statistics are reported in parentheses below coefficients. (The sources of data are found in Annex B along with alternative estimation models.)

The estimated coefficient on cognitive skills implies that an increase of one standard deviation in performance (*i.e.*, 100 on the PISA scale) would yield an annual growth rate that is 1.74 percentage points higher.

As discussed in the text, the estimates presume that $GDP/capita_{1960}$, C, and S are the systematic determinants of growth rates and that other factors that might explain growth are uncorrelated with these. Moreover, C is assumed to cause g, and not the other way around. See text for analyses supporting these assumptions.

approach here is built on the observed variations of country means for a group of countries that have well developed and relatively stable educational systems over the time period.[11] An "OECD Standardisation Group" (OSG) is created by using the 13 OECD countries that had half or more of the relevant population attaining a secondary education in the 1960s (the time of the first tests). For each assessment, the variance



in country mean scores for the subset of the OECD Standardisation Group participating is calibrated to the variance observed on the PISA tests in 2000 (when all countries of the OECD Standardisation Group participate). While it is difficult to judge the accuracy of this estimation approach, it seems plausible that variations in test score construction are more significant than changes in country performance over time.

By combining the adjustments in levels (based on the NAEP scores in the United States) and the adjustment in variances (based on the OECD Standardisation Group), it is possible to calculate standardised scores for all countries on all assessments. Each age group and subject is normalised to the PISA standard of mean 500 and individual standard deviation of 100 across OECD countries.

Basic facts of cognitive skills and economic growth

Combining the measures of achievement from the international tests with basic economic statistics permits analysing how skills affect subsequent outcomes. The basic elements of the underlying models are described in Box 2 and in the overview of the outcomes below. More details of the underlying analysis are found elsewhere.[12]

The extended empirical analysis relates long-term growth to cognitive skills and other aspects of national economies, relying upon an international dataset for 50 countries. These countries have participated in one or more the international testing occasions between 1964 and 2003 and have aggregate economic data for the period 1960-2000.[13] The underlying statistical model relates average annual growth rates in real GDP per capita over the 1960-2000 period to GDP per capita in 1960, various measures of human capital (including the cognitive skills measure), and other factors that might influence growth. The inclusion of initial GDP per capita simple reflects the fact that it is easier to grow when one is farther from the technology frontier, because one just must imitate others rather than invent new things. Real GDP is called on a purchasing power parity basis.

The empirical approach is consistent with a basic endogenous growth model. Those models are based on the generation of ideas and new technologies, which seems consistent with the perspective and measurement of cognitive skills. Nonetheless, it is not possible to adequately distinguish among alternative forms of growth models within the limited cross-country data employed here.

The simplest overview of the relationship is found in Figure 6 that plots regional growth in real per capita GDP between 1960 and 2000 against average test scores after allowing for differences in initial GDP per capita in 1960.[14] Regional annual growth rates, which vary from 1.4% in Sub-Saharan Africa to 4.5% in East Asia, fall on a straight line.[15] But school attainment, when added to this regression, is unrelated to growth-rate differences. Figure 6 suggests that, conditional on initial income levels, regional growth over the last four decades is completely described by differences in cognitive skills.

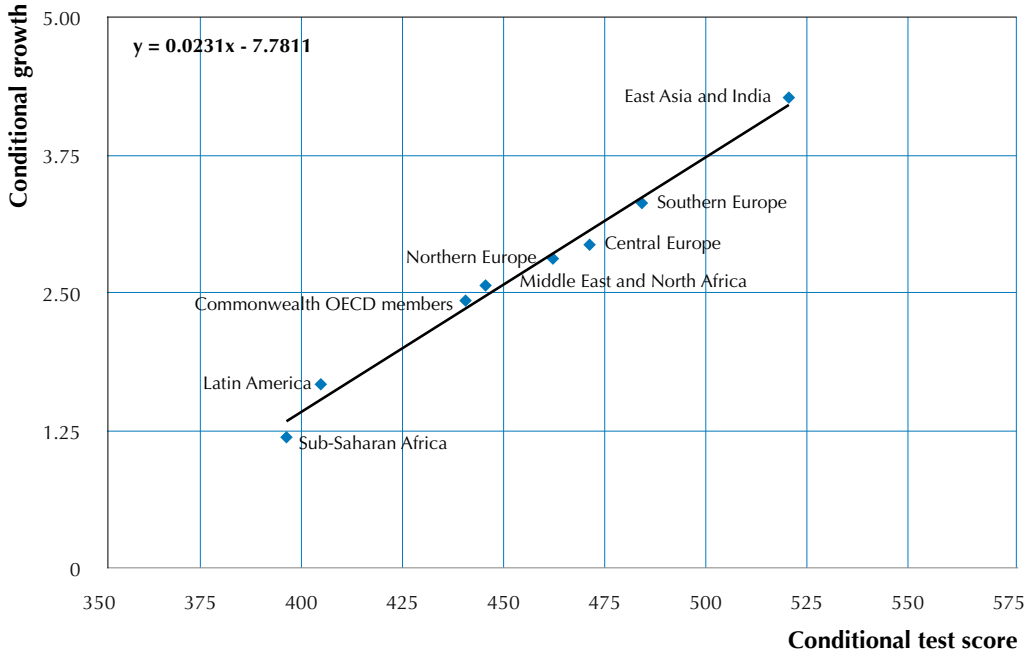
This aggregate regional analysis can be done over the same 1960-2000 time period for the 50 countries of the world with both test scores and economic data. Figure 7 also identifies individual OECD countries.[16] There are two important messages from this. First, test scores are closely related to growth across the world. Second, the OECD countries fit well within the rest of the world on this plot.

The consistency of growth with test scores seen here with the previous regional picture is quite remarkable. Moreover, once information is included on cognitive skills, school attainment bears no relation to economic growth. In other words, added years of schooling do not affect growth unless they yield greater achievement.[17] Of course, much of the observed cognitive skill is developed in schools, so this does not say that schools are irrelevant. It does say that the quality of schools, as determined by increases in student achievement, is very important.

© OECD 2010



Figure 6
Educational performance and economic growth across world regions



Notes: Added-variable plot of a regression of the average annual rate of growth (in percentage) of real GDP per capita in 1960-2000 on the initial level of real GDP per capita in 1960 and average test scores on international student achievement tests (mean of the unconditional variables added to each axis). Own depiction based on the database derived in Hanushek and Woessmann (2009).

Figure 7 also shows the ability of achievement differences to explain differences in growth rates just within the OECD countries. Some have argued that the differences in performance within the OECD are not large enough to have much impact on other outcomes – but the evidence suggests the opposite.

The underlying statistical estimates indicate a powerful effect of cognitive skills on growth. An improvement of one-half standard deviation in mathematics and science performance at the individual level implies, by historical experience, an increase in annual growth rates of GDP per capita of 0.87%. While more detail is provided about these improvements below, suffice it to say that Finland was approximately one-half standard deviation above the OECD average over the 2000-06 period. This historical impact suggests a very powerful response to improvements in educational quality.

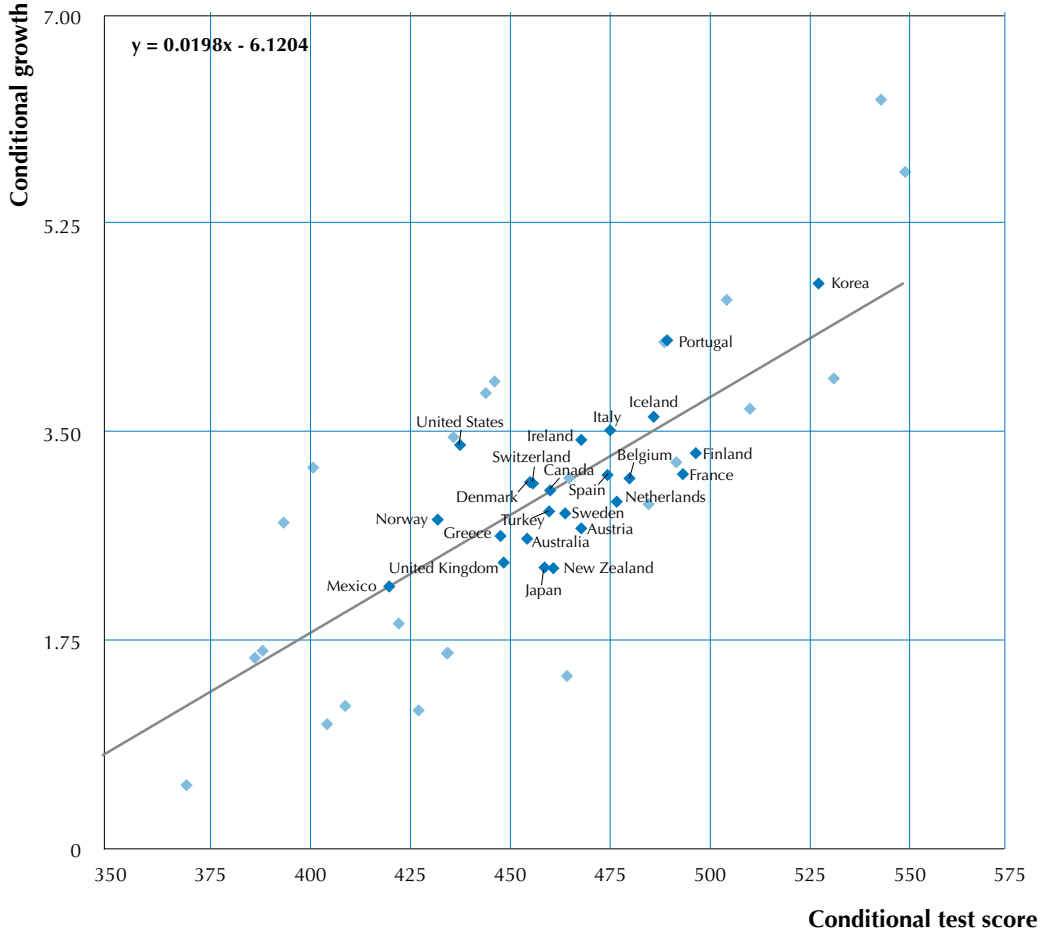
Evidence on causality

Before going into any analysis of the implications of these differences, it is important to know whether to interpret the tight relationship between cognitive skills and growth as reflecting a causal relationship that can support direct policy actions. Work on cross-country growth analysis has been plagued by legitimate questions about whether any truly causal effects have been identified, or alternatively whether the estimated statistical analyses simply pick up a correlation without causal meaning. Perhaps the easiest way to see the problems is early discussion of how sensitive estimated growth relationships were to the precise factors that were included in the statistical work and to the country samples and time periods of the analyses (Levine and Renelt, 1992, Levine and Zervos, 1993). The sensitivity of the estimated models provided *prima facie* evidence that various factors were omitted from many of the analyses.



Figure 7

Educational performance and economic growth in the full sample



Notes: Added-variable plot of a regression of the average annual rate of growth (in percentage) of real GDP per capita in 1960-2000 on the initial level of real GDP per capita in 1960, average test scores on international student achievement tests, and average years of schooling in 1960 (mean of the unconditional variables added to each axis). OECD countries are labeled by country codes for better readability and non-OECD countries by symbols only. Own depiction based on the database derived in Hanushek and Woessmann (2009).

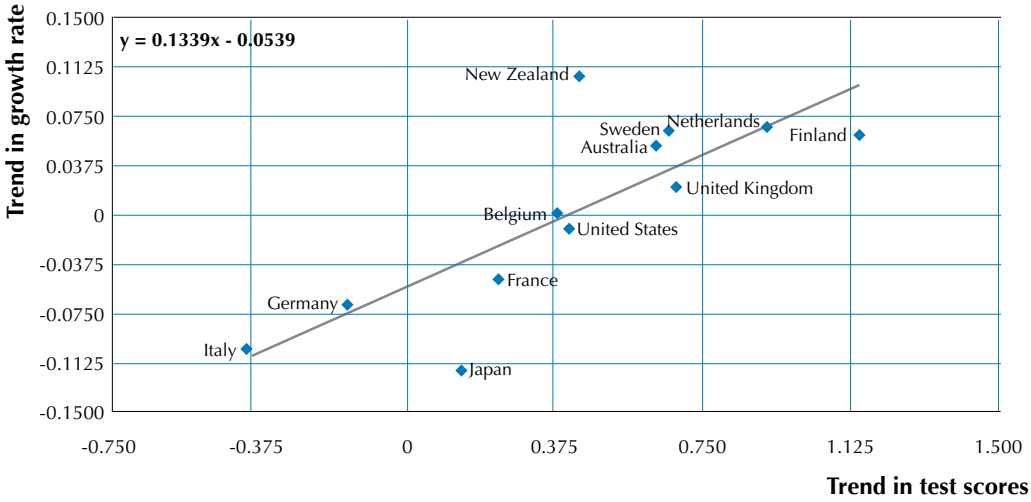
Whether or not this is a causal relationship is indeed a very important issue from a policy standpoint. It is essential to know that, if a country managed to improve its achievement in some manner, it would see a commensurate improvement in its long-run growth rate. Said differently, if the figures simply reflect other factors that are correlated with test scores, a change in test scores may have little or no impact on the economy (unless the other factors also changed). Indeed, analysis of prior estimates of school attainment have been identified as possibly reflecting reverse causality; *i.e.*, improved growth leads to more schooling rather than the reverse (Bils and Klenow, 2000).

It is difficult to develop conclusive tests of causality issues within the limited sample of countries included in the analysis. Nonetheless, Hanushek and Woessmann (2009) pursue a number of different approaches to ruling out major factors that could confound the results and that could lead to incorrect conclusions about the potential impact. In the end, none of the approaches addresses all of the important issues. Each approach fails to be conclusive for easily identified reasons. However, the combination of approaches, with similar

© OECD 2010



Figure 8
Trends in educational performance and trends in economic growth rates



Notes: Scatter plot of trend in the growth rate of GDP per capita from 1975 to 2000 against trend in test scores for countries whose test scores range back before 1972. Own depiction based on the database derived in Hanushek and Woessmann (2009).

support for the underlying growth models, provides some assurance that the most obvious problematic issues are not driving the results. First, this estimated relationship is little affected by including other possible determinants of economic growth. In an extensive investigation of alternative model specifications, different measures of cognitive skills, various groupings of countries including eliminating regional differences, and specific sub-periods of economic growth, Hanushek and Woessmann (2009) show a consistency of the alternative estimates – both in terms of quantitative impacts and statistical significance – that is uncommon to most cross-country growth modelling. Moreover, these estimates complement prior findings that measures of geographical location, political stability, capital stock, population growth, and school inputs (pupil-teacher ratios and various measures of spending) do not significantly affect the estimated impact of cognitive skills.[18] The only substantial effect on the estimates is the inclusion of various measures of economic institutions (security of property rights and openness of the economy) which reduces the estimated impact of cognitive skills by 15%.[19] These specification tests rule out some basic problems of omitted causal factors that have been seen in other work, but of course there are other possible omitted factors.[20]

Second, to tackle the most obvious reverse-causality issues, Hanushek and Woessmann (2009) separate the timing of the analysis by estimating the effect of scores on tests conducted until the early 1980s on economic growth in 1980-2000. In this analysis, available for a smaller sample of countries only, test scores pre-date the growth period. The estimate shows a significant positive effect that is about twice as large as the coefficient used in the simulations here. In addition, reverse causality from growth to test scores is also unlikely because additional resource in the school system (which might become affordable with increased growth) do not relate systematically to improved test scores (*e.g.* Hanushek, 2002).

Third, the analysis traces the impact on growth of just the variations in achievement that arise from institutional characteristics of each country's school system (exit examinations, autonomy, and private schooling).[21] This estimated impact is essentially the same as previously reported, lending support both to the causal impact of more cognitive skills and to the conclusion that schooling policies can have direct economic returns. Nonetheless, countries that have good economic institutions may have good schooling



institutions, so that this approach, while guarding against simple reverse causality, cannot eliminate a variety of issues of omitted factors in the growth regressions.[22]

Fourth, one major concern is that countries with good economies also have good school systems – implying that those that grow faster because of the basic economic factors also have high achievement. To deal with this, immigrants to the United States who have been educated in their home countries are compared to those who were educated in the United States. Since it is the single labour market of the United States, any differences in labour-market returns associated with cognitive skills cannot arise because of differences in the economies of their home country. Looking at labour-market returns, the cognitive skills seen in the immigrant's home country lead to higher incomes – but only if the immigrant was educated at home. Immigrants from the same home country schooled in the United States see no economic return to home-country quality, thus pinpointing the value of better schools.[23] While not free from problems, this difference-in-differences approach rules out the possibility that test scores simply reflect cultural factors or economic institutions of the home country.[24] It also provides further support to the potential role of schools to change the cognitive skills of citizens in economically meaningful ways.

Finally, perhaps the toughest test of causality is reliance on how *changes* in test scores over time lead to *changes* in growth rates. This approach eliminates country-specific economic and cultural factors. Figure 8 simply plots trends in educational performance and trends in growth rates over time for OECD countries.[25] This investigation provides more evidence of the causal influence of cognitive skills. The gains in test scores over time are related to the gains in growth rates over time.[26] As with the other approaches, this analysis must presume that the pattern of achievement changes has been occurring over a long time, because it is not the achievement of school children but the skills of workers that count. Nonetheless, the consistency of the patterns and the similarities of magnitudes of the estimates to the basic growth models is striking (see Hanushek and Woessmann, 2009).

Again, each approach to providing a deeper look at the issue of causation is subject to its own uncertainty. Nonetheless, the combined evidence consistently points to the conclusion that differences in cognitive skills lead to economically significant differences in economic growth. Moreover, even if some issues of omitted factors or reverse causation remain, it seems very unlikely that these cause *all* of the estimated effect – something that enters into the interpretation of the projections below.

Since the tests concentrate on the impact of schools, the evidence also suggests that school policy can, if effective in raising cognitive skills, be an important force in economic development. While other factors – cultural, health, and so forth – may affect the level of cognitive skills in an economy, schools also contribute to the relevant human capital.

THE ECONOMIC COSTS OF LOW EDUCATIONAL ACHIEVEMENT

The historical record on the relationship between cognitive skills and economic growth provides a means of directly evaluating the benefits from any educational reform programmes. Or, read the other way, it can provide an indication of the cost of not improving schools. Without taking a position on how much school improvement is possible, desirable, or likely, the analysis uses several alternative benchmarks to provide country-specific information about the economic impact of change.

Simulation approach

The prior analysis provides an indication of the long-run impact on growth rates of a labour force with varying skills as measured by mathematics and science scores of students enrolled in school. This long-run relationship does not, however, describe the path of benefits from any programme of changing the skills of the population. A variety of programmes, as noted previously, could improve the cognitive skills of

© OECD 2010



the population – including health programmes, schooling programmes, the introduction of new teaching technologies, and the like. For this analysis, however, the entire focus is schooling programmes, because schools are the locus of a large share of governmental policies today.

It is important to understand the dynamics of economic impacts of programmes. Three elements of the dynamics are particularly important for consideration: first, programmes to improve cognitive skills through schools take time to implement and to have their impact on students. It is simply not possible to change learning over night. Second, the impact of improved skills will not be realised until the students with greater skills move into the labour force. Third, the economy will respond over time as new technologies are developed and implemented, making use of the new higher skills.

In order to capture these elements, a simple simulation model is employed (the details of the simulation are shown in Annex C). The underlying idea is that moving from one quality level to another of the workforce depends on the shares of workers with different skills. As such, the impact of skills on GDP at any point in time will be proportional to the average skill levels of workers in the economy. The expected work life is assumed to be 40 years, which implies that each new cohort of workers is 2.5% of the workforce. Thus, even after an educational reform is fully implemented, it takes 40 years until the full labour force is at the new skill level.

In order to consider the impacts of improvement on OECD countries, the simulations rely on the estimates of growth relationships derived from the 23 OECD countries with complete data. These estimates suggest that a 50-point higher average PISA score (*i.e.*, one-half standard deviation higher) would be associated with 0.87% higher annual growth. This estimate clearly includes some uncertainty, a factor that is also included in the simulations below (see Box 2 and Annex B for details of the underlying models).

The simulations are conducted for all of the OECD countries and assume that each country can simultaneously grow faster. In other words, the higher levels of human capital in each country allow it to innovate, to improve its production, and to import new technologies without detracting from the growth prospects for other countries.[27] Further, the estimates ignore any other aspect of interactions such as migration of skilled labour across borders. Of course, one way that a country could improve its human capital would be by arranging for its youth to obtain schooling in another country with better schools – as long as the more educated youth return to their home country to work.

The simulation does not adopt any specific reform package but instead focuses just on the ultimate change in achievement. For the purposes here, reforms are assumed to take 20 years to complete, and the path of increased achievement during the reform period is taken as linear. For example, an average improvement of 25 points on PISA is assumed to reflect a gain of 1.25 points per year. This might be realistic, for example, when the reform relies upon a process of upgrading the skills of teachers – either by training for existing teachers or by changing the workforce through replacement of existing teachers. This linear path dictates the quality of new cohorts of workers at each point in time. To gauge the magnitude of such changes, Poland, the country that displayed the largest improvement in PISA, improved its performance in reading by 29 points between 2000 and 2006.

The dynamic nature of reform on the economy implies that the benefits to the economy from any improvement continue to evolve after the reform is completed. This characteristic, again, is an outgrowth of the growth models that are estimated, where improvements in technology and productivity are related to the level of skills of the workforce.

It is possible to summarise these changes in different ways, and it is important to understand the meaning of each. Perhaps the simplest way to see the impact of any improvement in cognitive skills is to trace out the increased GDP per capita that would be expected at any point in the future. The prior estimates of the effect on economic growth of differences in cognitive skills yield a path of relative gains in GDP per capita. Thus,



for example, it is possible to say what percentage increase in GDP per capita would be expected in 2050, given a specific change in skills started today. These changes are relative to the GDP in 2050, since the prior work indicates the marginal changes in growth rates that would be expected from higher skills.

An alternative approach is to summarise the economic value of the entire dynamic path of improvement in GDP per capita. Doing this is more difficult than the previous evaluation because the results will be dependent on a variety of additional factors. The value of improvement in economic outcomes from added growth depends, of course, on the path of economies that would be obtained without educational improvement. The analysis here takes the annual growth of OECD economies in the absence of education reform to be 1.5%. This is simply the average annual growth rate of potential GDP per worker of the OECD area over the past two decades: 1.5% in 1987-96 and 1.4% in 1997-2006 (OECD, 2009a).

The length of the time period over which gains are calculated is somewhat arbitrary and depends in part on the use of the analysis for any policy decisions. The benchmark here considers all economic returns that arise during the lifetime of a child that is born at the beginning of the reform in 2010. According to the most recent data (that refer to 2006), a simple average of male and female life expectancy at birth over all OECD countries is 79 years (OECD, 2009b).[28] Therefore, the calculations will take a time horizon until 2090, considering all future returns that accrue until then, but neglecting any returns that accrue after 2090.

Finally, because economic benefits accrue at varying times into the future, it is important to recognise that more immediate benefits are both more valuable and more certain than those far in the future. In order to incorporate this, the entire stream is converted into a present discounted value. In simplest terms, the present discounted value is the current dollar amount that would be equivalent to the future stream of returns calculated from the growth model. If we had that amount of funds and invested it today, it would be possible to reproduce the future stream of economic benefits from the principal amount and the investment returns. Thus, this calculation of present discount value allows a relevant comparison for any other current policy actions.

In doing so, the discount rate at which to adjust future benefits becomes an important parameter. A standard value of the social discount rate used in long-term projections on the sustainability of pension systems and public finance is 3% (*e.g.* Börsch-Supan, 2000, Hagist, Klusen, Plate and Raffelhüschen, 2005), a precedent that is followed here.[29] By contrast, the influential Stern Review report that estimates the cost of climate change uses a discount rate of only 1.4%, thereby giving a much higher value to future costs and benefits (Stern, 2007). If this practice were followed here, the discounted values of the considered education reforms would be substantially bigger than reported here.[30]

Scenario I: Increase average performance by 25 PISA points

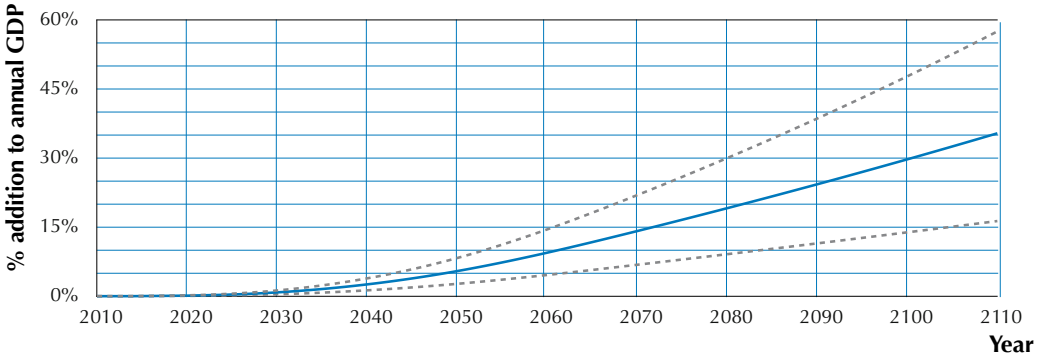
A simple starting point is to consider the economic impact on OECD countries of a 25-point increase on PISA scores (the country with the largest performance increase in PISA scores between 2000 and 2006 was Poland, with an increase of 29 score points in the reading assessment.)[31] The economic models (presented in Annex B) relate this gain (0.25 standard deviation of improvement) to economic growth. (The precise estimates consider a reform policy that is begun in 2010 and that on average yields 25 point higher scores in 2030 that remain permanently at that level for all subsequent students.)[32]

A policy like this is uniform across countries, so the relative improvement is the same for all countries.[33] Figure 9 provides a summary of the impact on GDP for each year into the future. While there are no impacts initially until higher-achieving students start becoming more significant in the labour market, GDP will be more than 3% higher than what would be expected without improvements in human capital as early as 2042. (The figure also shows a 95% confidence bound of 1.5 - 4.6% higher GDP, based on the relevant bounds for the regression coefficient in Box 2.) By the end of expected life in 2090 for the person born in 2010, GDP per capita would be expected to be about 25% above the "education as usual" level.

© OECD 2010



Figure 9
Improvement in annual GDP with Scenario I (improve student performance by 25 points on the PISA scale)



Notes: GDP with reform relative to GDP without reform in each year after the reform starts. Main line: point estimate of Scenario I. Gray dotted lines: 95% confidence interval of the point estimate of the growth regression. Authors' calculations.

Table 1 Effect on GDP of Scenario I: Increase average performance by 25 points on PISA, or by ¼ standard deviation

	Value of reform (USD bn)
Australia	2 527
Austria	899
Belgium	1 108
Canada	3 743
Czech Republic	918
Switzerland	792
Germany	8 088
Denmark	586
Spain	4 147
Finland	553
France	6 043
United Kingdom	6 374
Greece	996
Hungary	587
Ireland	514
Iceland	40
Italy	5 223
Japan	11 640
Korea	4 054
Luxembourg	116
Mexico	4 812
Netherlands	1 889
Norway	841
New Zealand	338
Poland	2 029
Portugal	680
Slovak Republic	311
Sweden	1 019
Turkey	3 416
United States	40 647
OECD	**114 930**

Notes: Discounted value of future increases in GDP until 2090, expressed in billion USD (PPP). For reform parameters, see Annex Table C1.



Again, a number of assumptions go into these calculations. First, they assume that skills play the same role in the future as they have in the past, so that the evidence of past results provides a direct way to project the future. Second, while the statistical analysis did not look at how economies adjust to improved skills, the calculations assume that the experience of other countries with greater cognitive skills provide the relevant insight into how the new skills will be absorbed into the economy.

The magnitude of such a change is best understood with an example. In the absence of changes in educational policy, France would be expected to have a GDP (in 2010 USD) of USD 3 638 billion in 2042. If, on the other hand, it achieved the improvement in cognitive skills that took it from an average PISA score of 505 to 530, total GDP would be expected to be USD 3 749 in 2042, or USD 111 billion higher.[34] These calculations illustrate a simple point: while 3% may at first seem like a small change, it is a very large number when applied to the entire GDP of any of the OECD countries.

These calculations are by themselves misleading, because the impacts of improved cognitive skills continue to occur far into the future. The 3.0% improvement in 2042 rises to a 5.5% improvement in 2050, 14.2% in 2070, and 24.3% in 2090. These dynamic improvements in the economy yield on-going gains to society, and the appropriate summary of the impact of educational improvements accumulates the value of these annual gains.

Importantly, after all individuals in the labour force have obtained the new and improved education (in 2070), annual growth will be 0.43 percentage points higher. This implies that each country that achieves the average improvement of ¼ standard deviation of achievement will have a cumulative impact on the economy through 2090 that is equal to 268% of current year GDP. Table 1 provides these discounted values of all of the future increases through 2090 for each OECD country. The dollar value for each country varies by the level of GDP in 2010 – but the total impact across the OECD is USD 115 trillion in present value.

These calculations reinforce the argument made earlier: the value of improvements through long-run growth far outstrip the costs of the current worldwide recession.

Scenario II: Bring each country to the average level of Finland

The success of Finland on the PISA tests is well-known. For the policy purposes here, the performance of Finnish students is taken as a benchmark for the performance levels that are possible. The economic impact calculated is found from projecting the impact on growth for each OECD country under the assumption that it could bring itself to the top of the rankings as identified by Finland, at an average PISA score of 546 (average of mathematics and science in 2000, 2003 and 2006).

Quite obviously, the amount of reform necessary varies by where each OECD country ranks today. And, commensurately, the impact on different economies also varies. Finland, for example, under this scenario would neither change its schools nor see any long-term economic changes. On the other side, Mexico and Turkey would require enormous changes in their educational achievement, and, if the changes were feasible, would see their economies virtually transformed.

Table 2 presents the country-by-country impacts of these changes. On average, the OECD countries would see a nearly 50-point increase in performance (one-half standard deviation). While the change in Japan or Korea amounts to about 5 points, the change in Mexico is 144 points – an almost inconceivable change given current knowledge of how to transform schools or cognitive skills in general. (Again, the calculations assume that adjustment is complete within 20 years. An alternative view would be that a number of countries would actually require more than 20 years for a reform programme to yield such large changes.)

The present value for OECD improvements under this scenario is USD 260 trillion, or six times the current GDP of the OECD countries. The United States itself, which currently falls over 50 points behind Finland,

© OECD 2010



Table 2 Effect on GDP of Scenario II: Bring each country to Finish level of 546 points on PISA

	Value of reform (USD bn)	% of current GDP	Long-run growth increase (percentage points)	Increase in PISA score
Australia	2 011	213%	0.35	20.1
Austria	1 430	425%	0.67	38.4
Belgium	1 452	351%	0.56	32.2
Canada	2 524	180%	0.30	17.2
Czech Republic	1 060	309%	0.50	28.6
Switzerland	1 120	378%	0.60	34.5
Germany	15 743	521%	0.80	46.0
Denmark	1 181	539%	0.82	47.5
Spain	11 289	728%	1.07	61.7
Finland	0	0%	0.00	0.0
France	10 424	461%	0.72	41.3
United Kingdom	7 326	307%	0.49	28.5
Greece	3 996	1 073%	1.48	85.2
Hungary	1 282	584%	0.88	51.0
Ireland	870	453%	0.71	40.6
Iceland	74	489%	0.76	43.6
Italy	18 094	927%	1.31	75.6
Japan	2 526	58%	0.10	5.7
Korea	746	49%	0.08	4.8
Luxembourg	383	884%	1.26	72.7
Mexico	38 756	2 155%	2.50	143.9
Netherlands	1 251	177%	0.29	16.9
Norway	1 956	622%	0.94	53.9
New Zealand	258	204%	0.34	19.4
Poland	5 061	667%	0.99	57.2
Portugal	2 588	1 019%	1.42	81.7
Slovak Republic	709	609%	0.92	52.9
Sweden	1 657	435%	0.68	39.2
Turkey	21 365	1 673%	2.08	120.1
United States	103 073	678%	1.01	58.1
OECD	**260 204**	**606%**	**0.87**	**49.8**

Notes: Discounted value of future increases in GDP until 2090, expressed in billion USD (PPP) and as percentage of current GDP. "Long-run growth increase" refers to increase in annual growth rate (in percentage points) once the whole labour force has reached higher level of educational performance. "Increase in PISA score" refers to the ultimate increase in educational performance due to the reform (of bringing each country to the Finish average level of 546 PISA points on the 2000 reading scale). For reform parameters, see Annex Table C1.

would by historical growth patterns see a present value of improved GDP of over USD 100 trillion, or some 40% of the total – reflecting both the size of the country and its distance behind Finland. Germany would see a USD 16 trillion improvement, or more than five times current GDP. All of these calculations are in real, or inflation-adjusted, terms.

The rankings of countries according to absolute dollar increases and to increases compared to current GDP are shown in Figures 2 and 3, respectively. One interpretation of these figures is the amount of economic leverage from educational improvements that is possible for different OECD countries.

Scenario III: Bring everyone up to a minimum skill level of 400 PISA points

The final scenario considered is the "compensatory" improvement in education where all students are brought up to a minimal skill level – which is defined here as obtaining a score of 400 on the PISA tests (one standard deviation below the OECD average). While the previous simulations could be thought of as displaying the results of shifting the entire achievement distribution, this scenario considers the implications of bringing up the bottom of the distribution.



Table 3 Effect on GDP of Scenario III: Bring all to minimum of 400 points on PISA

	Value of reform (USD bn)	% of current GDP	Long-run growth increase (percentage points)	Share of students below minimum skills
Australia	2 127	225%	0.37	9.8%
Austria	1 102	328%	0.52	13.9%
Belgium	1 509	364%	0.58	15.3%
Canada	2 594	185%	0.31	8.1%
Czech Republic	1 050	306%	0.49	13.0%
Switzerland	913	308%	0.50	13.1%
Germany	12 576	416%	0.65	17.3%
Denmark	793	362%	0.57	15.2%
Spain	6 865	443%	0.69	18.3%
Finland	217	105%	0.18	4.7%
France	8 222	364%	0.58	15.3%
United Kingdom	6 481	272%	0.44	11.7%
Greece	2 508	673%	1.00	26.5%
Hungary	857	390%	0.62	16.3%
Ireland	530	276%	0.45	11.8%
Iceland	47	312%	0.50	13.3%
Italy	11 465	587%	0.89	23.5%
Japan	8 306	191%	0.32	8.3%
Korea	2 288	151%	0.25	6.7%
Luxembourg	239	552%	0.84	22.3%
Mexico	26 064	1449%	1.87	49.5%
Netherlands	1 508	214%	0.35	9.3%
Norway	1 254	399%	0.63	16.6%
New Zealand	328	260%	0.42	11.2%
Poland	3 260	430%	0.67	17.8%
Portugal	1 545	608%	0.92	24.2%
Slovak Republic	450	387%	0.61	16.2%
Sweden	1 205	316%	0.51	13.4%
Turkey	14 895	1167%	1.58	41.8%
United States	72 101	475%	0.74	19.4%
OECD	193 301	450%	0.68	18.0%

Notes: Discounted value of future increases in GDP until 2090, expressed in billion USD (PPP) and as percentage of current GDP. "Long-run growth increase" refers to increase in annual growth rate (in percentage points) once the whole labour force has reached higher level of educational performance. "Share of students below minimum skills" refers to the share of students in each country performing below the minimum skill level of 400 PISA points. For reform parameters, see Annex Table C1.

In order to understand the implications of changing just one portion of the achievement distribution, an alternative estimation of the underlying economic growth models is employed. Specifically, instead of relying on just average cognitive skills in the growth models, the proportion of the population with scores less than 400 and the proportion with scores over 600 are included in the growth models (see Annex Table B1, column 5).

For these calculations, all OECD countries including Finland have room for improvement. On average, 18% of students in the OECD countries score below 400. And, as might be expected from the average scores, the required improvements are largest in Mexico and Turkey.

© OECD 2010



Table 3 displays the adjustments that would be required to bring all OECD students up to minimum competency levels along with the economic benefits that would flow from those improvements.

The overall change would be an average annual growth rate that was 0.7% higher after reform was accomplished and after the full labour force had received the improved education. The total improvements for the OECD countries from achieving universal minimum proficiency would have a present value of USD 193 trillion. Again, there is a wide range of outcomes including relatively small improvements of 185% of current GDP for Canada as compared to six OECD countries that would experience a benefit more than five times their current GDP.

The range of outcomes is depicted in Figures 4 and 5 that rank countries by the absolute benefits and by the benefits compared to current GDP. Even Finland could by these calculations more than double its current GDP through bringing the relatively modest proportion of low performers (4.7 %) up to scores of 400.

POLICY CONCLUSIONS

There is one message from these calculations: past experiences suggest that there are enormous economic gains to be had by OECD countries that can improve the cognitive skills of their populations. Moreover, the gains, put in terms of current GDP, far outstrip the value of short-run business-cycle management. This is not to say that efforts should not be directed at current issues of economic recession, but it is to say that the long-run issues should not be neglected.

The implications for the OECD countries as a whole are dramatic. A modest goal of having all OECD countries boost their average PISA scores by 25 points over the next 20 years – which is less than the most rapidly improving education system in the OECD, Poland, achieved between 2000 and 2006 alone – implies an aggregate gain of OECD GDP of USD 115 trillion over the lifetime of the generation born in 2010 (as evaluated at the start of reform in terms of real present value of future improvements in GDP) (see Figure 1). More aggressive goals, such as bringing all students to a level of minimal proficiency for the OECD (a PISA score of 400), would imply aggregate GDP increases of close to USD 200 trillion according to historical growth relationships (see Figure 4). Bringing all countries up to the OECD's best performing education system in PISA, Finland, would result in gains in the order of USD 260 trillion (see Figure 2). The report also shows that it is the quality of learning outcomes, not the length of schooling, which makes the difference.

There is uncertainty in these projections as there is in all projections, although it is perhaps greater here than in many other situations. The first issue is whether the statistical models used to characterise OECD growth between 1960 and 2000 accurately reflect the underlying determinants of growth. Economists disagree about the most appropriate way to model economic growth, and these estimates are based upon the specific form of endogenous growth models. Moreover, the execution of the estimation, including the measurement of cognitive skills and the allowance for other growth factors, incorporates additional elements of uncertainty. The second issue is the economic reality that is being projected. These projections trace the economy for 80 years into the future. A changing impact of cognitive skills on technological change and economic growth would clearly directly affect the specific estimates (although there is little reason to presume that it is more likely that the role of cognitive skills decreases as opposed to increases). Similarly, the present value of improved growth depends on the general health and growth of individual economies, which again is simply projected according to the historic patterns of the OECD nations. Other details, including how heavily future incomes are discounted and the time span for the calculations, also enter the specific projections.

Nonetheless, even reducing the projections substantially to allow for plausible minimal estimates suggests very large implications of improved cognitive skills and human capital. If the estimated impacts of cognitive



skills were twice as large as the true underlying causal impact on growth, the resulting present value of successful school reform still far exceeds any conceivable costs of improvement. The scenarios analysed using the best estimates (those in Box 2) have an economic value that is three to six times the aggregate GDPs of OECD countries. Half the impact remains a remarkably important potential change in the economic welfare of OECD countries.

Changing schools and educational institutions is a generally difficult task. Moreover, countries that have attempted reforms of schools often find that the results in terms of student achievement are relatively modest. This analysis does not provide answers in terms of the reforms that might be most productive.[35] Instead, it stops at providing an indication of the potential gains from true reforms that lead to improvements.

The political-economy issues are that the real impacts on OECD economies come sometime into the future, because it takes time for schools to improve student performance and for students to become a substantial part of a country's labour force. Thus, countries must make substantial changes now to reap the future benefits. On the other hand, simply saying that change is "too difficult" implies foregoing enormous gains to OECD nations.

© OECD 2010



References

Acemoglu, **Daron**, **Simon Johnson** and **James A. Robinson** (2005), "Institutions as a Fundamental Cause of Long-run Growth", in *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf , North Holland, Amsterdam, pp. 385-472.

Aghion, **Philippe** and **Peter Howitt** (1998), *Endogenous Growth Theory*, MIT Press, Cambridge, MA.

Azariadis, **Costas** and **Allan Drazen** (1990), "Threshold Externalities in Economic Development", *Quarterly Journal of Economics*, Vol.105, No. 2, pp.501-526.

Barro, **Robert J.** (1991), "Economic Growth in a Cross Section of Countries", *Quarterly Journal of Economics,* Vol. 106, No. 2, pp.407-443.

Barro, **Robert J.** (2001), "Human Capital and Growth", *American Economic Review,* Vol. 91, No.2, pp.12-17.

Barro, **Robert J.** and **Jong-Wha Lee** (1993) "International Comparisons of Educational Attainment", *Journal of Monetary Economics*, Vol. 32, No.3 pp. 363-394.

Barro, **Robert J.** and **Jong-Wha Lee** (2001), "International Data on Educational Attainment: Updates and Implications", *Oxford Economic Papers,* Vol. 53, No.3, pp. 541-563.

Barro, **Robert J.** and **Xavier Sala-i-Martin** (2004), *Economic Growth* (second edition), The MIT Press, Cambridge, MA.

Benhabib, **Jess** and **Mark M. Spiegel** (1994), "The Role of Human Capital in Economic Development: Evidence from Aggregate Cross-Country Data", *Journal of Monetary Economics,* Vol. 34, No.2, pp.143-174.

Benhabib, **Jess** and **Mark M. Spiegel** (2005), "Human Capital and Technology Diffusion", *in Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf, North Holland, Amsterdam, pp. 935-966.

Bils, **Mark** and **Peter J. Klenow** (2000), "Does Schooling Cause Growth?", *American Economic Review,* Vol. 90, No.5, pp. 1160-1183.

Börsch-Supan, **Axel** (2000), "A Model under Siege: A Case Study of the German Retirement Insurance System", *Economic Journal,* Vol. 110, No.461, pp. F24-F45.

Cohen, **Daniel** and **Marcelo Soto** (2007), "Growth and Human Capital: Good Data, Good Results", *Journal of Economic Growth,* Vol. 12, No.1, pp. 51–76.

Coulombe, **Serge** and **Jean-François Tremblay** (2006), "Literacy and Growth", *Topics in Macroeconomics,* Vol. 6, No. 2, Article 4.

Coulombe, **Serge**, **Jean-François Tremblay** and **Sylvie Marchand** (2004), *Literacy Scores, Human Capital and Growth across Fourteen OECD Countries*, Statistics Canada, Ottawa.

Denison, **Edward F.** (1985), *Trends in American Economic Growth, 1929-1982*, The Brookings Institution, Washington, D.C.

Glaeser, **Edward L.**, **Rafael La Porta**, **Forencio Lopez-de-Silanes** and **Andrei Shleifer** (2004), "Do Institutions Cause Growth?", *Journal of Economic Growth,* Vol. 9, No.3, pp. 271-303.

Goldin, **Claudia** and **Lawrence F. Katz** (2008), *The Race between Education and Technology*, Harvard University Press, Cambridge.

Gundlach, **Erich**, **Ludger Woessmann** and **Jens Gmelin** (2001), "The Decline of Schooling Productivity in OECD Countries", *Economic Journal,* No. 111, pp. C135-C147.

Hagist, **Christian**, **Norbert Klusen**, **Andreas Plate** and **Bernd Raffelhüschen** (2005), "Social Health Insurance: The Major Driver of Unsustainable Fiscal Policy?", *CESifo Working Papers*, Munich, p. 1574.

Hanushek, **Eric A.** (2002), "Publicly Provided Education", in *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein, Elsevier, Amsterdam, pp. 2045-2141.

Hanushek, **Eric A.** and **Dennis D. Kimko** (2000), "Schooling, Labour Force Quality, and the Growth of Nations", *American Economic Review,* Vol. 90, No.5, pp.1184-1208.

Hanushek, **Eric A.** and **Alfred A. Lindseth** (2009), *Schoolhouses, Courthouses, and Statehouses: Solving the Funding-Achievement Puzzle in America's Public Schools*, Princeton University Press, Princeton, NJ.

Hanushek, **Eric A.** and **Ludger Woessmann** (2008), "The Role of Cognitive Skills in Economic Development", *Journal of Economic Literature,* Vol. 46, No.3, pp. 607-668.

Hanushek, **Eric A.** and **Ludger Woessmann** (2009), "Do Better Schools Lead to More Growth? Cognitive Skills, Economic Outcomes, and Causation", *NBER Working Paper No. 14633*, National Bureau of Economic Research, Cambridge, MA.

Hanushek, **Eric A.** and **Ludger Woessmann** (forthcoming), "How Much do Educational Outcomes Matter for Developed Countries", Mimeo, Hoover Institution, Stanford University Stanford, CA.

Hanushek, **Eric A.** and **Lei Zhang** (2009), "Quality-Consistent Estimates of International Schooling and Skill Gradients", *Journal of Human Capital,* Vol. 3, No.2, pp.107-143.

Heston, **Alan**, **Robert Summers** and **Bettina Aten** (2002), "Penn World Table Version 6.1.", Center for International Comparisons at the University of Pennsylvania (CICUP), University of Pennsylvania, Philadelphia.



Jorgenson, **Dale W.** and **Zvi Griliches** (1967), "The Explanation of Productivity Change", *Review of Economic Studies,* Vol. 34, No.3, pp. 249-282.

Krueger, **Alan B.** and **Mikael Lindahl** (2001), "Education for Growth: Why and for Whom?", *Journal of Economic Literature,* Vol. 39, No.4, pp. 1101-1136.

Levine, **Ross** and **David Renelt** (1992), "A Sensitivity Analysis of Cross-Country Growth Regressions", *American Economic Review,* Vol. 82, No.4, pp. 942-963.

Levine, **Ross** and **Sara J. Zervos** (1993), "What We Have Learned about Policy and Growth from Cross-Country Regressions?", *American Economic Review,* Vol. 83, No.2, pp. 426-430.

Lucas, **Robert E., Jr.** (1988), "On the Mechanics of Economic Development." *Journal of Monetary Economics* (22 July), pp. 3-42.

Lucas, **Robert E., Jr.** (2003), "Macroeconomic Priorities", *American Economic Review* Vol. 93, No.1, pp. 1-14.

Mankiw, **N. Gregory**, **David Romer** and **David Weil** (1992), "A Contribution to the Empirics of Economic Growth", *Quarterly Journal of Economics,* Vol. 107, No.2, pp. 407-437.

Moore, **Mark A.**, **Anthony E. Boardman**, **Aidan R. Vining**, **David L. Weimer** and **David Greenberg** (2004), "'Just Give me a Number!' Practical Values for the Social Discount Rate", *Journal of Policy Analysis and Management,* Vol. 23, No.4, pp. 689-812.

Neidorf, **Teresa S.**, **Marilyn Binkley**, **Kim Gattis** and **David Nohara** (2006), *Comparing Mathematics Content in the National Assessment of Educational Progress (NAEP), Trends in International Mathematics and Science Study (TIMSS), and Programme for International Student Assessment (PISA) 2003 Assessments*, National Center for Education Statistics, Washington.

Nelson, **Richard R.** and **Edmund Phelps** (1966), "Investment in Humans, Technology Diffusion and Economic Growth", *American Economic Review,* Vol. 56, No.2, pp. 69-75.

Nicoletti, **Giuseppe** and **Frederic L. Pryor** (2006), "Subjective and Objective Measures of Governmental Regulations in OECD Nations", *Journal of Economic Behavior & Organization,* Vol. 59, No.3, pp. 433-449.

Nicoletti, **Giuseppe** and **Stefano Scarpetta** (2003), "Regulation, Productivity, and Growth: OECD Evidence", *Economic Policy,* Vol. 18, No.1, pp. 10-72.

OECD (2003), *Education at a Glance 2003: OECD Indicators*, OECD, Paris.

OECD (2009a), *OECD Economic Outlook, Volume 1, No. 85*, OECD, Paris.

OECD (2009b), *Society at a Glance 2009: OECD Social Indicators*, OECD, Paris.

Pritchett, **Lant** (2006), "Does Learning to Add up Add up? The Returns to Schooling in Aggregate Data", in *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch, North Holland, Amsterdam, pp. 635-695.

Romer, **Paul** (1990a), "Endogenous Technological Change", *Journal of Political Economy,* Vol. 99, No.5, part. II, pp. S71-S102.

Romer, **Paul** (1990b), "Human Capital and Growth: Theory and Evidence", *Carnegie-Rochester Conference Series on Public Policy,* No. 32, pp. 251-286.

Sala-i-Martin, **Xavier**, **Gernot Doppelhofer** and **Ronald I. Miller** (2004), " Determinants of Long-Term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach", *American Economic Review* 94, No.4, pp. 813-835.

Sianesi, **Barbara** and **John Van Reenen** (2003), "The Returns to Education: Macroeconomics", *Journal of Economic Surveys,* Vol. 17, No.2, pp. 157-200.

Solow, **Robert M.** (1957), "Technical Change and the Aggregate Production Function", *Review of Economics and Statistics,* Vol. 39, No.3, pp. 312-320.

Stern, **Nicholas** (2007), *The Economics of Climate Change: The Stern Review*, Cambridge University Press, Cambridge, UK, and New York.

Temple, **Jonathan** (2001), "Growth Effects of Education and Social Capital in the OECD Countries", *OECD Economic Studies No. 33*, pp. 57-101.

Topel, **Robert** (1999), "Labor Markets and Economic Growth", in *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, Elsevier, Amsterdam, pp. 2943-2984.

U.S. Department of Education, Institute of Education Sciences (2008), *National Assessment of Educational Progress – The Nation's Report Card, http://nces.ed.gov/nationsreportcard/aboutnaep.asp.*

Welch, **Finis** (1970), "Education in Production", *Journal of Political Economy,* Vol. 78, No.1, pp. 35-59.

Woessmann, **Ludger**, **Elke Luedemann**, **Gabriela Schuetz** and **Martin R. West** (2009), *School Accountability, Autonomy, and Choice Around the World*, Edward Elgar, Cheltenham, UK.



Notes

1. For an account of the historical development, see Barro and Sala-i-Martin (2004). The associated empirical work concentrated on within-country analyses such as Solow (1957), Jorgenson and Griliches (1967), or Denison (1985).

2. The commonly suggested alternative view, an "augmented neoclassical" approach, identifies human capital as another input to production in the economy, subject to the same diminishing returns as capital and labour. It suggests that changes in the human capital of a nation will lead to growth as a nation moves from one steady-state level of income to another but that it will not keep growing after it adjusts to the new level on aggregate income. See Box 1.

3. The earliest studies used adult literacy rates (*e.g.* Azariadis and Drazen, 1990; Romer, 1990b) but these data cover a limited number of countries and are error-prone.

4. For extensive reviews of the literature, see, *e.g.* Topel (1999); Temple (2001); Krueger and Lindahl (2001); Sianesi and Van Reenen (2003).

5. In the statistical analyses, they simply allow for differences in average growth in East Asian economies by including a dummy variable for being an East Asian country into the regression analysis.

6. Note that there are also problems within individual countries if school quality changes over time. For the sample of countries participating in the International Adult Literacy Survey (IALS), there is evidence of considerable change in quality within countries; see Hanushek and Zhang (2009).

7. Hanushek and Kimko (2000) also use data from the International Assessment of Educational Progress (IAEP), but those results have not been employed in most subsequent studies because they derive specifically from the NAEP tests and thus the curriculum in the United States. See Annex A.

8. Note that this does not mean that individuals learn nothing after age 15, but rather that what they have learned in school is a good predictor for the accumulation of further skills in life and the capacity to deploy these skills effectively.

9. The one exception with measures of the cognitive skills of individuals in the labour force is the International Adult Literacy Survey (IALS) conducted in 1994-98. This survey, conducted by the OECD in 23 countries (or regions of separate countries), measured cognitive skills of a representative sample of individuals aged 16-65. One set of studies has used these data to construct synthetic cohorts in order to estimate an augmented neoclassical growth model using panel data techniques across 14 countries. See Coulombe, Tremblay, and Marchand (2004) and Coulombe and Tremblay (2006).

10. Transforming scores based on just the mean and variance is appropriate if the underlying score distribution is normal but potentially introduces errors for other underlying distributions. As shown in Annex Figure A2, the distribution of PISA scores within the OECD is normal, even if the distributions in individual countries may not be. The transformation is designed to put national means on a common scale and does not apply to the subnational level. In the analysis of minimal skills (below), the calculations employ the empirical distributions for each of the countries and do not assume normality.

11. The development of aggregate scores by Hanushek and Kimko (2000) and by Barro (2001) assumed that the test variances across assessments were constant, but there is no reason for this to be the case. The approach employed here is in the spirit of Gundlach, Woessmann and Gmelin (2001).

12. The summary regression statistics are found in Annex B. A more complete overview of the results is found in Hanushek and Woessmann (2008). Extensive sensitivity tests and discussion of causation are found in Hanushek and Woessmann (2009).

13. International economic data come from the Penn World table (Heston, Summers and Aten, 2002). During this period, countries under a communist regime are not included.

14. Regional data come from averaging all countries with available data in a region. The division of Europe into three regions illustrates the heterogeneity within OECD countries, but a combined Europe also falls on the line in Figure 6.

15. In statistical terms, the regression line behind Figure 6 has an R^2=0.985.

16. The sample with complete data includes 23 OECD countries.



17. The basic statistical models use school attainment in 1960 but combine that with average test scores over the entire period. However, the results are unchanged by using average years of attainment across the period.

18. The prior estimates rely not just on OECD countries but on the full set of countries in Hanushek and Kimko (2000) and Hanushek and Woessmann (2008).

19. Including measures of economic institutions, suggested for example by Acemoglu, Johnson and Robinson (2005), does lead to some reduction in the estimated impact of cognitive skills – reducing the coefficient from 1.74 to 1.47 (although the institutional measures do not enter significantly in the OECD-sample estimation). However, as Glaeser, La Porta, Lopez-de-Silanes and Shleifer (2004) argue, there is a good case that human capital causes better institutions as opposed to the opposite. Thus, one could consider the estimate of 1.47 as a lower bound on any achievement effects. As seen below, even evaluating the simulations at the lower bound has little impact about the overall conclusions of the importance of cognitive skills for economic well-being.

20. For example, there are significant variations across the OECD in regulations of labour and product markets, bureaucratic burdens, and the like (see Nicoletti and Scarpetta, 2003, Nicoletti and Pryor, 2006) that have yet to be investigated in terms of long-run growth.

21. The statistical analysis employs an instrumental variables strategy that relies upon changes in achievement induced by school structure. Its major limitation is that the instruments tend to be weak, given the small number of countries that is included.

22. One frequent argument is that countries that grow faster will have more resources, allowing them to increase their spending on schools and thus affecting quality. By this argument, growth would cause better schooling rather than the opposite. But this situation is just what the instrumental-variables procedure will guard against when the educational institutions of testing, private schooling, and autonomy of decision-making are not themselves a result of growth.

23. These results also hold when Mexicans (the largest immigrant group) are excluded and when only immigrants from English-speaking countries are included.

24. Three potential problems arise in this analysis. First, it just looks at the labour-market returns for the individual and not the aggregate impact on the economy of achievement differences. Second, those who migrate at a young enough age to be educated in the United States might differ from those who migrate at later ages. Third, employers may treat individuals with a foreign education differently from those who received their education in the United States. The second two potential problems, however, can only affect the results in complicated ways, because the identification of the impact of cognitive skills is based on a comparison across the home countries. As long as the impact of these is similar for the different origin countries, the results would remain. Any problems would come from different patterns of these factors that are correlated with test scores across countries.

25. Only 12 OECD countries have participated in international tests over a long enough period to provide the possibility of looking at trends in test performance over more than 30 years. The analysis simply considers a bivariate regression of test scores on time for countries with multiple observations. The trends in growth rates are determined in a similar manner: annual growth rates are regressed on a time trend. The plot provides the pattern of slopes in the test regression to slopes in the growth rate regression. Hanushek and Woessmann (2009) consider more complicated statistical relationships, but the overall results hold.

26. It is possible but unlikely that the changes in growth rates suffer the same reverse causality concerns suggested previously. A change in growth rates can occur at low levels of growth and in lower-income economies.

27. No attempt is made to consider how technological change occurs and the impact on wages and earnings. Obviously, different patterns of productivity improvements will play out differently in the labour market as seen in the United States over time (Goldin and Katz, 2008).

28. Note that these life expectancy numbers are based on age-specific mortality rates prevalent in 2006, and as such do not include the effect of any future decline in age-specific mortality rates. Life expectancy at birth has increased by an average of more than 10 years since 1960.

29. As a practical value for the social discount rate in cost-benefit analysis (derived from an optimal growth rate model), Moore *et al.* (2004) suggest using a time-declining scale of discount rates for intergenerational projects that do not crowd out private investment, starting with 3.5% for years 0-50, 2.5% for years 50-100, 1.5% for years 100-200, 0.5 % for years 200-300 and 0% years over 300. (The proper starting value is actually 3.3% based on the parameter values they assume for the growth rate in per capita consumption (2.3%), the social marginal utility of consumption with respect to per capita consumption (1), and the utility discount rate (1%).)



30. Hanushek and Woessmann (forthcoming) present projections based on several alternative model parameters, time horizons, and discount rates.

31. Throughout this section and each of the subsequent simulations, it is assumed that the increase in scores is permanent.

32. All calculations of PISA scores underlying the following simulations refer to the average performance in mathematics and science (in line with the underlying growth model, as noted above), averaged over the three PISA cycles 2000, 2003, and 2006 (see OECD, 2001, 2003b, 2004, 2007). All underlying measures of Gross Domestic Product (GDP) are in US dollars (USD), measured in purchasing power parities (PPP), expressed in prices of 2010. The GDP measures were calculated from the most recent measure of GDP in current prices and current PPPs available for all countries (2007, extracted from *http://stats.oecd.org* on 10 August 2009), projected to 2010 using OECD estimates of annual changes in potential GDP and in GDP deflators (OECD, 2009a).

33. Note that the calculations also assume that the top-ranked countries can feasibly improve their scores. The relatively flat performance of countries such as Japan and Korea that have been at the top for a number of years raises the question about whether there is room for further improvement or whether there is some sort of ceiling effect in the existing tests. As an alternative, the next scenario will only assume improvements that do not go beyond the current top performer.

34. These calculations assume a constant population size over this period.

35. In a variety of other work, the importance of different educational institutions is investigated. See the international study of Woessmann, Luedemann, Schuetz and West (2009) and the United States analysis of Hanushek and Lindseth (2009).



Annex A

Deriving a Combined Historical Measure of Cognitive Skills



A key element of the work is developing a measure that can equate knowledge of individuals across countries. In many ways this is an extension of notions of human capital that have been developed over the past half century. But it is a specific refinement that, while important in a variety of applications within nations, becomes a necessity when comparing different countries. Within a country, human capital is often proxied by quantity of schooling. This is partly necessitated by commonly available data but partly justified on the idea that differences in knowledge between levels of schooling are greater than those within levels of schooling. This development was originally presented in Hanushek and Woessmann (2009).

Until recent publicity, most individuals were unaware of the international student testing that could provide direct comparisons of student knowledge across countries. In fact, international assessments of student achievement, aimed largely at mathematics and science, were begun over four decades ago. Although national participation has been voluntary, recent expansions to all OECD countries and more have led to increasingly valid and reliable indicators of cognitive skills.

Internationally comparable student achievement tests have been conducted since the First International Mathematics Study (FIMS), which tested students in 1964. The latest international studies used in the analyses are the 2003 cycles of the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA). From FIMS to the latest TIMSS and PISA, a total of 12 international student achievement tests (ISATs) were conducted.[1] Although varying across the individual assessments, testing covers mathematics, science, and reading for three age/grade groups: primary education (age 9/10), lower secondary education (age 13 to 15), and the final year of secondary education (generally grade 12 or 13). (The 2006 PISA results were not included because the results were unavailable when the test series was constructed and because they represent scores noticeably after the period of economic observations studied in this report).

Table A1 International tests by period, subject, and age group

1964-72

	Mathematics	Science	Reading
Primary		FISS	
Lower secondary	FIMS	FISS	FIRS
Final secondary	FIMS	FISS	

1982-91

	Mathematics	Science	Reading
Primary		SISS	SIRS
Lower secondary	SIMS	SISS	SIRS
Final secondary	SIMS	SISS	

1995-2003

	Mathematics	Science	Reading
Primary	TIMSS TIMSS 2003	TIMSS TIMSS 2003	PIRLS
Lower secondary	TIMSS TIMSS-Repeat PISA 2000/02 TIMSS 2003 PISA 2003	TIMSS TI-Repeat PISA 2000/02 TIMSS 2003 PISA 2003	PISA 2000/02 PISA 2003
Final secondary	TIMSS	TIMSS	

Note: For abbreviations and details, see Table A2.

Source: Hanushek and Woessmann (2009).

© OECD 2010



Given this 3x3 grade-by-subject matrix, Table A1 summarises the specific ISATs that have been conducted in three periods of time: late 1960s-early 1970s (1964-72), 1980s (1982-91), and late 1990s/early 2000s (1995-2003). Several features of the emerging pattern are worth noting. First, mathematics and science have been tested at all three grade levels, while reading has not been tested in the final grade of secondary school. Second, all subjects are available in all periods, although coverage is more extensive in mathematics and science than in reading. Third, in each period, the lower secondary level has been tested in all three subjects; thus, there is no primary or final-secondary study that would add a subject not already tested at lower secondary in the period. Fourth, each cell available in 1964-91 has at least one counterpart in 1995-2003.

Table A2 provides additional detail on each ISAT. A total of 77 countries have participated in at least one of the ISATs in mathematics or science, but several of the countries have participated at only one or a few points in time. Even within the same assessment, countries do not always participate at all grade levels. The largest number of countries tends to have participated at the lower secondary level.

Table A2 The international student achievement tests

	Abbr.	Study	Year	Subject	Age[a,b]	Countries[c]	Organisation[d]	Scale[e]
1	FIMS	First International Mathematics Study	1964	Mathematics	13,FS	11	IEA	PC
2	FISS	First International Science Study	1970-71	Science	10,14,FS	14,16,16	IEA	PC
3	FIRS	First International Reading Study	1970-72	Reading	13	12	IEA	PC
4	SIMS	Second International Mathematics Study	1980-82	Mathematics	13,FS	17,12	IEA	PC
5	SISS	Second International Science Study	1983-84	Science	10,13,FS	15,17,13	IEA	PC
6	SIRS	Second International Reading Study	1990-91	Reading	9,13	26,30	IEA	IRT
7	TIMSS	Third International Mathematics and Science Study	1994-95	Mathematics/ Science	9(3+4), 13(7+8),FS	25,39,21	IEA	IRT
8	TIMSS-Repeat	TIMSS-Repeat	1999	Mathematics/ Science	13(8)	38	IEA	IRT
9	PISA 2000/02	Programme for International Student Assessment	2000+02	Reading/ Mathematics/ Science	15	31+10	OECD	IRT
10	PIRLS	Progress in International Reading Literacy Study	2001	Reading	9(4)	34	IEA	IRT
11	TIMSS 2003	Trends in International Mathematics and Science Study	2003	Mathematics/ Science	9(4),13(8)	24,45	IEA	IRT
12	PISA 2003	Programme for International Student Assessment	2003	Reading/ Mathematics/ Science	15	40	OECD	IRT

Notes:
a. Grade in parentheses where grade level was target population.
b. FS = final year of secondary education (differs across countries).
c. Number of participating countries that yielded internationally comparable performance data.
d. Conducting organisation: International Association for the Evaluation of Educational Achievement (IEA); Organisation for Economic Co-operation and Development (OECD).
e. Test scale: percent-correct formal (PC); item-response-theory proficiency scale (IRT).
Source: Hanushek and Woessmann (2009).

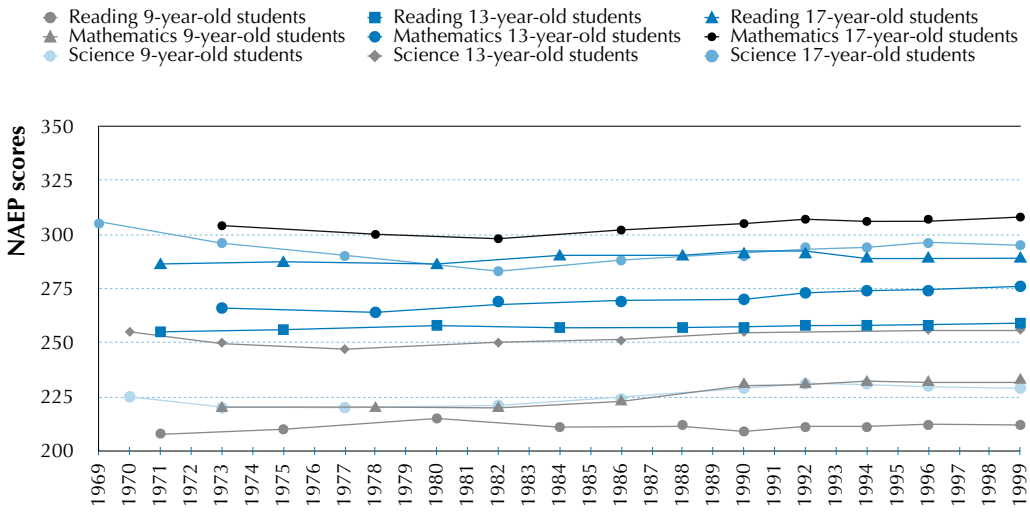
Obtaining a common measure of cognitive skills calls for drawing on as much internationally comparable information as possible. This raises the issue whether the different ISATs with their different participating countries, student samples, and perspectives on what should be tested (see Neidorf, Binkley, Gattis and Nohara (2006)) are measuring a common dimension of cognitive skills. For example, the TIMSS tests are related to elements of the school curricula common to participating countries, while the PISA tests are designed to be applied assessments of real-world problems, irrespective of specific curricula. However, the fact is that the TIMSS tests with their curricular focus and the PISA tests with their applied focus are highly correlated at the country level. For example, the correlation between the TIMSS 2003 tests of 8th graders and the PISA 2003 tests of 15-year-olds across the 19 countries participating in both is as high as 0.87 in mathematics and 0.97 in science. It is also 0.86 in both mathematics and science across the 21 countries participating both in the TIMSS 1999 tests and the PISA 2000-02 tests. Thus, ISATs with very different foci and perspectives tend, nonetheless, to be highly related, lending support to the approach of aggregating different ISATs for each country.

The general idea behind the approach to aggregation is that of empirical calibration. It relies upon information about the overall distribution of scores on each ISAT to compare national responses. This contrasts with the psychometric approach to scaling that calibrates tests through common elements on each test. In reality, the international testing situations are separate events with no general attempt to provide common scaling across tests and across the full time period.

The fact that the scales of their test-score results are not directly equated across tests is a major drawback in comparative uses of the various ISATs. They do not use the same test questions; nor do they even use the same technique and scale of mapping answers into test scores.[2] The early tests mainly used aggregate scores in "percent correct" format, but with questions of varying difficulty in the different tests, these scores will not be comparable across tests. The later tests use a more sophisticated scale, constructed using Item Response Theory (IRT). Among other things, IRT weights different questions by their revealed difficulty and then maps answers onto a pre-set scale set to yield a given international mean and standard deviation among the participating countries. However, the questions on which the mapping is based are not the same in the different tests. Even more, the set of participating countries varies considerably across tests, making the separately developed scales incomparable across ISATs.

Figure A1

Student achievement in the United States over time: The National Assessment of Educational Progress (NAEP)



Source: U.S. Department of Education (2008)

© OECD 2010



Therefore, to compare performance on the ISATs across tests and thus over time, the performance of different countries on different tests is projected onto a common metric. For that, it is necessary to develop a common metric both for the *level* and for the *variation* of test performance.

Comparable level. To make the level of ISATs comparable, the key needed information is test performance that is comparable over time. Such information is available in the United States in the form of the National Assessment of Educational Progress (NAEP), which tests the mathematics, science, and reading performance of nationally representative samples of 9-, 13-, and 17-year-old students in the United States in an intertemporally comparable way since 1969. This is the only information on educational performance that is consistently available for comparisons over time. The United States is also the only country that participated in every ISAT. Given the time-series evidence on the performance of students in the United States, the level of each ISAT is scaled relative to the known intertemporally comparable test performance of the United States. Figure A1 shows the available NAEP results in the three subjects and age groups. Despite some notable changes, the performance of students in the United States has been relatively flat over the period 1969-1999.

Start by calculating the performance difference of students in the United States between 1999 and any earlier point in time and express it in standard deviations (standard deviation) of the international PISA 2000 assessment:

$$U^{US}_{a,s,t} = \left(NAEP^{US}_{a,s,t} - NAEP^{US}_{a,s,1999}\right)\frac{SD^{US,PISA}_{s}}{SD^{US,NAEP}_{a,s}} \qquad \text{(A1)}$$

where U is the standardised performance difference of students in the United States at age a in subject s at time t relative to 1999, $NAEP$ is the age-, subject-, and time-specific NAEP test score, $SD^{US,PISA}$ is the subject-specific standard deviation of students in the United States on the PISA test, and $SD^{US,NAEP}$ is the age- and subject-specific standard deviation of the NAEP test.[3] NAEP scores are available at 2-4 year intervals over the period; values for non-NAEP years are obtained by linear interpolation between available years.[4]

This alone does not yet yield a common scale for all the countries on the different tests. While it is known for each participating country whether it performed above or below the respective performance of the United States on each specific test, the international variation in test scores comparable across the different ISATs to determine "how much" above or below is not known.

Comparable variation. Developing a common metric for the variation of test scores in the different ISATs is harder to achieve than for the level. There is no explicit external information available on trends in the cross-country performance variation, and the diversity of original tests and participating countries precludes a direct comparison across tests. One way to achieve comparability, though, would be to have a group of countries across which it is reasonable to assume relative constancy in the size of the cross-country variation in test scores and whose members participated in sufficient number in the different tests. This group could only include relatively stable countries with relatively stable education systems over time, which should not have experienced major changes in overall enrolment across the ISATs.

Thus, this suggests two criteria for a group of countries to serve as a standardisation benchmark for performance variation over time. First, the countries have been member states of the relatively homogenous and economically advanced group of OECD countries in the whole period of ISAT observations, that is, since 1964. Second, the countries should have had a substantial enrolment in secondary education already in 1964. Given data constraints, this is implemented by applying a somewhat arbitrary rule of dropping all countries where more than half of the 2001 population aged 45-54 (the cohort roughly in secondary school in the first ISAT) did not attain upper secondary education (OECD 2003). There are 13 countries that meet both of these measures of stability, which is termed the "OECD Standardisation Group" (OSG) of countries.[5]

Under the assumption that the cross-country variation among the OSG countries did not vary substantially since 1964, the OSG countries provides a means to develop a comparable scale for the variation on the different ISATs. This proceeds by projecting the standard deviation among those of the OSG countries that participated in any particular ISAT from the

subject-specific PISA test onto the particular ISAT. That is, the original test score *O* of country *i* (specific for each age *a* and subject *s*) at time *t* is transformed into test score *X* according to:

$$(A2) \qquad X^{i}_{a,s,t} = \left(O^{i}_{a,s,t} - \overline{O^{OSG}_{a,s,t}}\right)\frac{SD^{OSG}_{s,PISA}}{SD^{OSG}_{a,s,t}}$$

The test score *X* has the following distributional characteristics for each ISAT. First, it has a mean of zero among the OSG (attained by subtracting the OSG mean $\overline{O^{OSG}}$ from each country's original test score). Second, it has a between-country standard deviation among the OSG that is the same as the standard deviation of the very same countries on the PISA test in the specific subject (attained by dividing through the standard deviation among the OSG countries in the specific test and multiplying by the standard deviation of these same countries in the relevant PISA test). In effect, this rescaled test score now has a metric whose variation is comparable across tests.

Performance on a common metric. Finally, the time-series evidence on educational performance of students in the United States derived above is used to put a common level to the intertemporally comparable metric for the different ISATs. This is achieved in the standardised test score *I*:

$$(A3) \qquad I^{i}_{a,s,t} = X^{i}_{a,s,t} - X^{US}_{a,s,t} + O^{US}_{s,PISA} + U^{US}_{a,s,t}$$

which adjusts the variation-adjusted test score *X* so that the performance level of the United States on each test equals the performance of the United States on the PISA test in the specific subject plus the age- and subject-specific adjustment factor *U* based on NAEP as derived in equation (A1) above.

Equation (A3) yields measures of the performance of the participating countries in each ISAT on a common scale that is comparable across ISATs. In effect, the internationally and intertemporally standardised test score *I* projects the PISA scale onto all other tests.

While the comparisons of the standardised scores within the OECD countries fully tested in recent years seems reasonable, there is less certainty about countries that are far from the measured OECD performance. In particular, countries far off the scale of the original test scores – *e.g.* two standard deviations below the mean – may not be well represented because the tests may be too hard and thus not very informative for them. The linear transformations are susceptible to considerable noise for these countries.

The main measure of cognitive skills is a simple average of all standardised mathematics and science test scores of the ISATs in which a country participated. Table A3 reports the basic combined measure for the 77 countries that have ever participated in any of the mathematics and science tests.[6] The full sample for growth regressions contains 50 of these countries (see Figure 5).[7]

© OECD 2010



[Part 1/2]

Table A3 International data on cognitive skills

Country	Cognitive	Basic	Top
Albania	3.785	0.424	0.013
Argentina	3.920	0.492	0.027
Armenia	4.429	0.745	0.008
Australia	5.094	0.938	0.112
Austria	5.089	0.931	0.097
Bahrain	4.114	0.608	0.003
Belgium	5.041	0.931	0.094
Botswana	3.575	0.374	0.000
Brazil	3.638	0.338	0.011
Bulgaria	4.789	0.765	0.083
Canada	5.038	0.948	0.083
Chile	4.049	0.625	0.013
China	4.939	0.935	0.083
Colombia	4.152	0.644	0.000
Cyprus	4.542	0.825	0.011
Czech Rep.	5.108	0.931	0.122
Denmark	4.962	0.888	0.088
Egypt	4.030	0.577	0.010
Estonia	5.192	0.973	0.095
Finland	5.126	0.958	0.124
France	5.040	0.926	0.085
Germany	4.956	0.906	0.105
Ghana	3.603	0.403	0.010
Greece	4.608	0.798	0.042
Hong Kong	5.195	0.944	0.123
Hungary	5.045	0.941	0.103
Iceland	4.936	0.908	0.074
India	4.281	0.922	0.013
Indonesia	3.880	0.467	0.008
Iran	4.219	0.727	0.006
Ireland	4.995	0.914	0.094
Israel	4.686	0.826	0.053
Italy	4.758	0.875	0.054
Japan	5.310	0.967	0.168
Jordan	4.264	0.662	0.044
Korea, Rep.	5.338	0.962	0.178
Kuwait	4.046	0.575	0.000
Latvia	4.803	0.869	0.050
Lebanon	3.950	0.595	0.002
Liechtenstein	5.128	0.860	0.198
Lithuania	4.779	0.891	0.030
Luxembourg	4.641	0.776	0.067
Macao-China	5.260	0.919	0.204
Macedonia	4.151	0.609	0.028
Malaysia	4.838	0.864	0.065
Mexico	3.998	0.489	0.009
Moldova	4.530	0.787	0.029
Morocco	3.327	0.344	0.001
Netherlands	5.115	0.965	0.092
New Zealand	4.978	0.910	0.106

[Part 2/2]

Table A3 International data on cognitive skills

Country	Cognitive	Basic	Top
Nigeria	4.154	0.671	0.001
Norway	4.830	0.894	0.056
Palestine	4.062	0.571	0.008
Peru	3.125	0.182	0.002
Philippines	3.647	0.485	0.006
Poland	4.846	0.838	0.099
Portugal	4.564	0.803	0.032
Romania	4.562	0.780	0.046
Russian Fed.	4.922	0.884	0.081
Saudi Arabia	3.663	0.331	0.000
Serbia	4.447	0.718	0.024
Singapore	5.330	0.945	0.177
Slovak Rep.	5.052	0.906	0.112
Slovenia	4.993	0.939	0.061
South Africa	3.089	0.353	0.005
Spain	4.829	0.859	0.079
Swaziland	4.398	0.801	0.004
Sweden	5.013	0.939	0.088
Switzerland	5.142	0.919	0.134
Taiwan (Chinese Taipei)	5.452	0.958	0.219
Thailand	4.565	0.851	0.019
Tunisia	3.795	0.458	0.003
Turkey	4.128	0.582	0.039
United Kingdom	4.950	0.929	0.088
United States	4.903	0.918	0.073
Uruguay	4.300	0.615	0.049
Zimbabwe	4.107	0.684	0.010

Notes: A data file is available at *www.cesifo.de/woessmann#data*.
Cognitive: Average test score in mathematics and science, primary through end of secondary school, all years (scaled to PISA scale divided by 100).
Basic: Share of students reaching basic literacy (based on average test scores in mathematics and science, primary through end of secondary school, all years).
Top: Share of top-performing students (based on average test scores in mathematics and science, primary through end of secondary school, all years).
Source: Based on Hanushek and Woessmann (2009).

Note that countries participate in varying numbers of assessments. Thus, it is plausible that the combined test scores for countries with many observations are better representations of the skills of the labour force than those calculated on fewer observations. Nonetheless, there is no obvious way to correct the statistical models to reflect this.

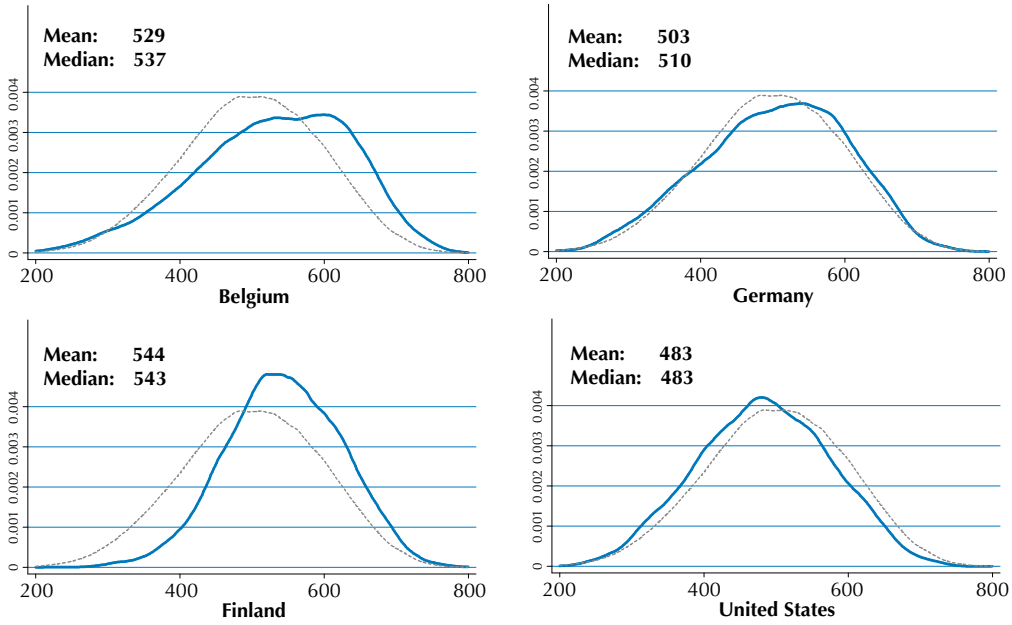
Distributional measures. Apart from the mean scores, the distribution of test scores in each country can be analysed by accessing the micro data of all ISATs.[8] The kernel density plots for mathematics achievement on the 2003 PISA in Figure A2 show that countries vary significantly in their patterns of test-score distributions. The depicted selected examples of OECD countries reveals that it is possible to achieve relatively high median performance both with a relatively equitable spread (Finland) and with a relatively unequal spread (Belgium) in the test scores at the student level.

A straightforward way to depict both ends of the distribution is to calculate both the share of students reaching a basic level of literacy in the different subjects equivalent to 400 test-score points on the PISA scale (one student-level standard deviation below the OECD mean) and the share of students reaching a top performance level equivalent to 600 test-score points on the PISA scale.

© OECD 2010



Figure A2
Examples of the distribution of student performance in selected OECD countries



Notes: Kernel densities of student performance on the PISA 2003 mathematics test. Bold solid line: specified country; thin dotted line: OECD countries.

Source: Hanushek and Woessmann (2009).

To do so, the above score transformation is used to translate these two thresholds into the specific metric of each ISAT. Using the micro data of each ISAT, the share of students in each country reaching the thresholds in the overall distribution of the ISAT is then calculated. The information from the different ISATs is again combined by taking a simple average of the shares across tests.

One possible concern with combining the different tests into a single measure is that enrolment shares have changed to different extents over time, especially at the secondary level. To test the extent to which this affects the cognitive-skill measures, the correlation between the measure of trend in test scores and changes in enrolment rates is calculated. It turns out that the two are virtually orthogonal to each other, diluting concerns that differential changes in enrolment bias the results reported in this paper.



Annex B

Underlying Cross-Country Growth Regressions



The underlying models of economic growth and their derivation are described in Hanushek and Woessmann (2008, 2009). This is a summary of the approach and provides the basic estimates employed.

Consider a very simple growth model: a country's economic growth rate (g) is a function of the skills of workers (H) and other factors (X) that include initial levels of income and technology, economic institutions, and other systematic factors. Skills are frequently referred to simply as the workers' human capital stock. For simplicity in equation (B1), assume that H is a one-dimensional index and that growth rates are linear in these inputs, although these are not important for the purposes here.[9]

$$(B1) \qquad g = \gamma H + \beta X + \varepsilon$$

As discussed in the extensive educational production function literature (Hanushek, 2002), these skills (H) are affected by a range of factors including family inputs (F), the quantity and quality of inputs provided by schools (qS), individual ability (A), and other relevant factors (Z) which include labour market experience, health, and so forth as in:

$$(B2) \qquad H = \lambda F + \phi(qS) + \eta A + \alpha Z + \nu$$

The schooling term combines school attainment (S) and its quality (q).

Human capital is nonetheless a latent variable that is not directly observed. To be useful and verifiable, it is necessary to specify the measurement of H. The vast majority of existing theoretical and empirical work on growth begins – frequently without discussion – by taking the quantity of schooling of workers (S) as a direct measure of H.

A more satisfying alternative is to focus directly on the cognitive skills component of human capital and to measure H with test-score measures of mathematics, science, and reading achievement.[10] The use of measures of cognitive skills has a number of potential advantages. First, they capture variations in the knowledge and ability that schools strive to produce and thus relate the putative outputs of schooling to subsequent economic success. Second, by emphasising total outcomes of education, they incorporate skills from any source – families, schools, and ability. Third, by allowing for differences in performance among students with differing quality of schooling (but possibly the same quantity of schooling), they open the investigation of the importance of different policies designed to affect the quality aspects of schools.

The basic growth model in equation (B1) is estimated for all of the 50 countries and for the 23 OECD countries with cognitive-skill and economic data over the period 1960-2000.[11] Cognitive skills are measured by the simple average of all observed mathematics and science scores between 1964 and 2003 for each country.

As a comparison to prior cross-country analyses, the first column of Table B1 presents estimates of a simple growth model with school attainment.[12] While this model explains one-quarter of the variance in growth rates, adding cognitive skills increases this to three-quarters of the variance. The test score is strongly significant with a magnitude that is unchanged by excluding school attainment (column 2) or including initial attainment (column 3). School attainment is never statistically significant in the presence of the direct cognitive-skill measure of human capital.

In the 50-country sample, one standard deviation in test scores (measured at the OECD student level) is associated with a two percentage points higher average annual growth rate in GDP per capita across 40 years.

© OECD 2010



Table B1 Educational performance and long-run economic growth across OECD countries

	(1)	(2)	(3)	(4)	(5)
Cognitive skills		2.015	1.980	1.736	
		(10.68)	(9.12)	(4.17)	
Years of schooling 1960	0.369		0.026	0.025	0.035
	(3.23)		(0.34)	(0.26)	(0.35)
GDP per capita 1960	-0.379	-0.287	-0.302	-0.301	-0.310
	(4.24)	(9.15)	(5.54)	(5.81)	(5.60)
Share of students reaching basic literacy					3.783
					(2.72)
Share of top-performing students					3.676
					(1.02)
Constant	2.785	-4.827	-4.737	-3.539	1.328
	(7.41)	(6.00)	(5.54)	(1.96)	(1.36)
No. of countries	50	50	50	23	23
R2 (adj.)	0.252	0.733	0.728	0.830	0.829

Notes: Dependent variable: average annual growth rate in GDP per capita, 1960-2000. *t*-statistics in parentheses.

When restricting the sample to the 23 OECD country with relevant data, the estimate is slightly reduced, to1.74 percentage points higher average annual growth per standard deviation in test scores (column 4). This is the estimate underlying the simulations presented in this report.

Instead of measuring human capital by the average cognitive skills of a population, a final growth specification enters the share of students reaching basic literacy and the share of top-performing students as separate growth determinants. To do so, the analysis employs measures of the share of students in each country that reaches a certain threshold of basic skills (taken as 400 test-score points on a PISA-equivalent international scale), as well as the share of students that surpasses an international threshold of top performance (taken as 600 test-score points). When estimated on the OECD sample, the coefficients on both shares capture similar magnitudes, but only the one on the share of students reaching basic literacy reaches statistical significance at standard levels (column 5). This estimate is used for the simulations of Scenario III in this report.

The extensive investigation of alternative model specifications, different measures of cognitive skills, various groupings of countries, inclusion of regional fixed effects, and specific sub-periods of economic growth in Hanushek and Woessmann (2009) shows very consistent estimates in terms of quantitative impacts and statistical significance. Moreover, prior estimates of similar models find that measures of geographical location, political stability, capital stock, population growth, and school inputs (pupil-teacher ratios and various measures of spending) do not significantly affect the estimated impact of cognitive skills.

The only substantial effect on the OECD estimates is the inclusion of various measures of economic institutions (security of property rights and openness of the economy) which reduces the estimated impact of cognitive skills by 15%. Including measures of economic institutions, suggested for example by Acemoglu, *et al.* (2005), does lead to some reduction in the estimated impact of cognitive skills – reducing the coefficient from 1.74 to 1.47 (although the institutional measures do not enter significantly in the OECD-sample estimation). However, as Glaeser, *et al.* (2004) argue, there is a good case that human capital causes better institutions as opposed to the opposite. Thus, one could consider the estimate of 1.47 as a lower bound on any achievement effects.



Annex C

Projection of the Economic Value of Education Reforms



This annex provides a formal description of the different steps of the projection of the growth effects of the education reform. The parameter values underlying the simulations are shown in Table C1.

Increase in the annual growth rate in the different phases:

a) Phase 1 (2010-2030): The education reforms are assumed to take 20 years to complete, and the path of increased achievement during this phase is taken as linear. The additional growth in GDP per capita due to the reform in year t is given by:

$$(C1) \qquad \Delta^t = growth\ coefficient * \Delta PISA * \frac{1}{working\ life} * \frac{t-2010}{20} + \Delta^{t-1}$$

where the *growth coefficient* stems from the underlying regression estimation and $\Delta PISA$ is the increase in the average PISA test score due to the reform.

b) Phase 2 (2031-2050): The education reform is now fully enacted. After the assumed length of work life of 40 years, the current workforce is fully replaced. During this phase, the additional growth in GDP per capita due to the reform in year t is given by:

$$(C2) \qquad \Delta^t = growth\ coefficient * \Delta PISA * \frac{1}{working\ life} + \Delta^{t-1}$$

c) Phase 3 (2051-2070): Then, the first 20 labour-market cohorts – which had not yet fully profited from the education reform – are replaced by cohorts that profited from the fully enacted education reform:

$$(C3) \qquad \Delta^t = growth\ coefficient * \Delta PISA * \frac{1}{working\ life} - \left(\Delta^{t-40} - \Delta^{t-41}\right) + \Delta^{t-1}$$

d) Phase 4 (after 2070): Finally, the whole workforce has gone through the reformed education system. The annual growth rate is now increased by the constant long-run growth effect Δ:

$$(C4) \qquad \Delta = growth\ coefficient * \Delta PISA$$

Development of GDP with and without reform:

a) Without reform, the economy grows at the constant growth rate of potential GDP:

$$(C5) \qquad GDP^t_{no\ reform} = GDP^{t-1}_{no\ reform} * \left(1 + potential\ growth\right)$$

b) With reform, the annual growth rate is additionally increased by the growth effect Δ^t:

$$(C6) \qquad GDP^t_{reform} = GDP^{t-1}_{reform} * \left(1 + potential\ growth + \Delta^t\right)$$

Total effect of the reform:

The total value of the reform is given by the sum of the discounted values of the annual differences between the GDP with reform and the GDP without reform:

$$(C7) \qquad Total\ value\ of\ the\ reform = \sum_{t=2010}^{t=2090} \left(GDP^t_{reform} - GDP^t_{no\ reform}\right) * \left(1 + discount\ rate\right)^{-(t-2010)}$$

© OECD 2010



Table C1 Parameter values for simulations

Parameter	Value
Start of reform	2010
Reform duration (years)	20
Horizon of future returns considered	2090
Working life (years)	40
Coefficient of average test score in growth regression	1.736
Coefficient of basic literacy in growth regression	3.783
Discount rate	3%
Growth of potential GDP (w/o reform)	1.5%



Notes to Annexes

1. In this study, there are two tests excluded, both conducted by the International Assessment of Educational Progress (IAEP) in 1988 and 1991. These used the NAEP test as their testing instrument, which is geared to the curricula in the United States and may thus introduce bias to the international testing. By contrast, the tests included here are not associated with the curriculum in any particular country, but have been devised in an international co-operative process between all participating countries.

2. Recent testing in both TIMSS and PISA has involved overlapping test items that permit test calibration, but these do not provide any benchmarks across the two testing regimes or linked with earlier testing.

3. The standard deviation of the NAEP tests in reading for 1984-1996 and in mathematics and science in 1977/78-1996 are reported in U.S. Department of Education (2008). Since no standard deviation information is available for the earlier and the 1999 NAEP tests, and since the available standard deviation is relatively stable over time, the simple mean of the available standard deviation in each subject at each age level over time is used. PISA tested only 15-year-olds, but has the same three subjects as the NAEP test.

4. For the PISA tests after 2000, the linking of the PISA tests is applied.

5. The OSG countries are: Austria, Belgium, Canada, Denmark, France, Germany, Iceland, Japan, Norway, Sweden, Switzerland, the United Kingdom, and the United States. The Netherlands also meets both criteria, but does not have the internationally comparable PISA 2000 data required for the standardisation.

6. The sources of the underlying international test data are official releases of the conducting organisation, the International Association for the Evaluation of Educational Achievement (IEA) and the OECD, as well as own calculations based on the micro data of the early tests; see Annex B of Hanushek and Woessmann (2009) for a complete list.

7. Twenty-five of the total of 77 countries with cognitive-skill data are not included in the growth database due to lack of data on economic output or because they drop out of the sample for a standard exclusion criterion in growth analyses (15 former communist countries, 3 countries for which oil production is the dominant industry, 2 small countries, 3 newly created countries, 2 further countries lacking early output data). In addition, two strong outliers are excluded in most models (see Hanushek and Woessmann, 2009). There are four countries with cognitive-skill data which have a few years of economic data missing at the beginning or end of the 1960-2000 period. Data for Tunisia start in 1961, and data for Cyprus (1996), Singapore (1996), and Taiwan (1998) end slightly earlier than in 2000. These countries were included in the growth regressions by estimating average annual growth over the available 36-to-39-year period.

8. Unfortunately, the micro data from the FIMS test do not seem to be available in an accessible way any more, so that the distributional measures only draw on the remaining ISATs.

9. The form of this relationship has been the subject of considerable debate and controversy. As noted, this formulation of an endogenous growth model is retained, in part because it is not possible to adequately distinguish among alternative forms.

10. Some researchers have suggested that test scores should be thought of as a measure of school quality (q), leading to use of test scores times years of schooling as a measure of H, but this ignores the influence of family factors and other elements of equation (B2) that have been shown to be very important in determining cognitive skills.

11. See Hanushek and Woessmann (2009) for details on the country sample. The source of the income data is version 6.1 of the Penn World Tables (cf. Heston, *et al.*, 2002). The data on years of schooling are an extended version of the Cohen and Soto (2007) data.

12. While not the focal point of this analysis, all specifications include GDP per capita in 1960, which provides consistent evidence for conditional convergence, *i.e.*, countries with higher initial income tend to grow more slowly.

© OECD 2010

OECD PUBLISHING, 2, rue André-Pascal, 75775 PARIS CEDEX 16
PRINTED IN FRANCE
(98 2010 02 1 E) – No. 57207 2010

The High Cost of Low Educational Performance

THE LONG-RUN ECONOMIC IMPACT OF IMPROVING PISA OUTCOMES

While governments frequently commit to improving the quality of education, it often slips down the policy agenda. Because investing in education only pays off in the future, it is possible to underestimate the value and the importance of improvements.

This report uses recent economic modelling to relate cognitive skills – as measured by PISA and other international instruments – to economic growth, demonstrating that relatively small improvements to labour force skills can largely impact the future well-being of a nation. The report also shows that it is the quality of learning outcomes, not the length of schooling, which makes the difference. A modest goal of all OECD countries boosting their average PISA scores by 25 points over the next 20 years would increase OECD gross domestic product by USD 115 trillion over the lifetime of the generation born in 2010. Other aggressive goals could result in gains in the order of USD 260 trillion.

Even if there is some uncertainty in such projections the gains, put in terms of current gross domestic product, far outstrip today's value of the short-run business-cycle management. While efforts should be directed at issues of economic recession, the long-run issues can no longer be neglected.

FURTHER READING

PISA 2006: Science Competencies for Tomorrow's World (OECD, 2007).

THE OECD PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA)

PISA is a collaborative process among the 30 member countries of the OECD and nearly 30 partner countries and economies. It brings together expertise from the participating countries and economies and is steered by their governments on the basis of shared, policy-driven interests. Its unique features include:

– *The literacy approach*: PISA defines each assessment area (science, reading and mathematics) not mainly in terms of mastery of the school curriculum, but in terms of the knowledge and skills needed for full participation in society.

– *A long-term commitment*: It enables countries to monitor regularly and predictably their progress in meeting key learning objectives.

– *The age-group covered:* By assessing 15-year-olds, *i.e.* young people near the end of their compulsory education, PISA provides a significant indication of the overall performance of school systems.

– *The relevance to lifelong learning*: PISA does not limit itself to assessing students' knowledge and skills but also asks them to report on their own motivation to learn, their beliefs about themselves and their learning strategies, as well as on their goals for future study and careers.

The full text of this book is available on line via this link:
www.sourceoecd.org/education/9789264077485

Those with access to all OECD books on line should use this link:
www.sourceoecd.org/9789264077485

SourceOECD is the OECD online library of books, periodicals and statistical databases.
For more information about this award-winning service and free trials ask your librarian, or write to us at **SourceOECD@oecd.org.**

OECDpublishing
www.oecd.org/publishing

98 2010 02 1 E