



IZA DP No. 5401

How Much Do Educational Outcomes Matter in OECD Countries?

Eric A. Hanushek
Ludger Woessmann

December 2010

How Much Do Educational Outcomes Matter in OECD Countries?

Eric A. Hanushek

*Hoover Institution, Stanford University,
NBER and CESifo*

Ludger Woessmann

*University of Munich, Ifo Institute,
CESifo and IZA*

Discussion Paper No. 5401
December 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

ABSTRACT

How Much Do Educational Outcomes Matter in OECD Countries?*

Existing growth research provides little explanation for the very large differences in long-run growth performance across OECD countries. We show that cognitive skills can account for growth differences within the OECD, whereas a range of economic institutions and quantitative measures of tertiary education cannot. Under the growth model estimates and plausible projection parameters, school improvements falling within currently observed performance levels yield very large gains. The present value of OECD aggregate gains through 2090 could be as much as \$275 trillion, or 13.8 percent of the discounted value of future GDP. Extensive sensitivity analyses indicate that, while differences between model frameworks and alternative parameter choices make a difference, the economic impact of improved educational outcomes remains enormous. Interestingly, the quantitative difference between an endogenous and neoclassical model framework – with improved skills affecting the long-run growth rate versus just the steady-state income level – matters less than academic discussions suggest. We close by discussing evidence on which education policy reforms may be able to bring about the simulated improvements in educational outcomes.

JEL Classification: I2, O4

Keywords: education, growth, OECD, cognitive skills, projection

Corresponding author:

Ludger Woessmann
Ifo Institute for Economic Research
at the University of Munich
Poschingerstraße 5
81679 München
Germany
E-mail: woessmann@ifo.de

* Paper presented at the 52nd Panel Meeting of *Economic Policy* in Rome. We thank the editor and four anonymous referees for their comments. Woessmann gratefully acknowledges the support and hospitality provided by the W. Glenn Campbell and Rita Ricardo-Campbell National Fellowship of the Hoover Institution, Stanford University, as well as support by the Pact for Research and Innovation of the Leibniz Association.

1. INTRODUCTION

Despite its surge over the past two decades, research in the economics of growth – both theoretical and empirical – has produced surprisingly few resilient results about policies that might promote long-run growth in developed countries (cf. Aghion and Howitt (2006)). Most of the robust results that exist refer either to the importance of basic economic institutions, with important policy implications for developing countries, or to policies that affect short- to medium-term growth in developed countries. Here we present evidence that improved human capital, measured by cognitive skills, has the potential for substantial improvements in the long-run economic well-being of OECD countries.

The immense variation in the long-run growth experiences of developed countries has largely escaped notice. For example, from 1960 to 2000, GDP per capita grew by less than 1.5 percent on average per year in New Zealand and Switzerland, but by more than 4 percent per year in Ireland, Japan, and South Korea. As a consequence, the average Korean was about 10 times as well off in 2000 as in 1960, and the average Irish and Japanese about 5 times. By contrast, the average New Zealander and Swiss was only 1.6-1.8 times as well off than 40 years before. These stark differences are directly visible when comparing the three fastest-growing and the three slowest-growing countries highlighted (together with the US) in Figure 1, which plots GDP per capita in 1960 and 2000: Korea surpassed several other OECD countries, including Mexico; Japan and Ireland went from 40-45 percent to 131-140 percent of New Zealand's income; and Ireland caught up to Switzerland from an initial level of 35 percent of its GDP per capita.

This paper focuses on these long-run growth differences among OECD countries. Our results show that educational outcomes play a leading role in their understanding, and we estimate how much improvements in educational outcomes could matter for developed economies. The paper builds on existing growth analyses that focus attention on cognitive skills, as opposed to the more commonly measured school attainment or years of school completed (see Hanushek and Woessmann (2008) for a review). But while much of the existing analysis is focused on developing countries and their difference from developed countries, here we focus on whether cognitive skills also matter in understanding growth differences among rich countries.

The basis of assessment of the role of human capital is a series of cross-country growth regressions that characterize variations in growth rates between 1960 and 2000. Our analysis of growth differences relies largely on the 24 OECD countries with consistent data on cognitive skills and economic growth across the period, although we also provide some relevant comparisons with the expanded sample of 50 countries that incorporates non-OECD countries.¹ Section 2 provides the basic estimates of growth models along with investigation of the sensitivity of the results to sample, model

¹ With the limited country samples, there is a distinct trade-off between incorporating the added observations from the full world sample and restricting the economic relationships to being the same across all countries. Throughout the analysis, we provide information on the similarities and differences of developed and developing countries.

specification, and the like. The clear result is that richer measures of human capital based on international math and science tests dramatically increase the ability of the statistical models to explain growth differences across OECD countries. Moreover, the estimated relationship is generally impervious to altered specifications of the model.

Perhaps the leading candidate for being a more fundamental explanation of growth is the quality of a country's economic institutions such as having secure property rights or an open economy. This idea – most fully developed by Daron Acemoglu and his co-authors – does not, however, have clear application in OECD countries where there is limited variation in these broad institutions (see Acemoglu, Johnson, and Robinson (2001, 2005)). Beyond these, considerable attention has been given to variations across the OECD in regulations of labor and product markets, bureaucratic burdens, and the like (see Nicoletti and Scarpetta (2003); Nicoletti and Pryor (2006)). The empirical application has, however, mostly addressed short-term growth effects at the industry level within countries.

Section 3 provides new evidence on the role of various economic institutions in growth. While property-rights and free-trade institutions have been shown to contribute to differences in long-run growth between rich and poor countries, and while labor- and product-market regulations have been shown to contribute to differences in short-run growth among rich countries, we show that neither set of institutions helps us understand differences in long-run growth rates within the group of OECD countries.

In part, this analysis also enters into the debates about the relative importance of institutions and human capital as stressed by Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004). Our results strongly favor cognitive skills, our measure of human capital, over institutions as a fundamental source of OECD-country growth differences.

While cognitive skills can be developed in a variety of ways, we focus on some fundamental questions about school policy. Educational policy differs considerably by country, with different countries pursuing very different investment approaches. At one level, it is seen in very different levels of tertiary education. For example, while Australia currently has 86 percent of a cohort entering tertiary education, Norway has 66 percent, and Italy has 53 percent (Organisation for Economic Co-operation and Development (2009a)).

Indeed, Vandenbussche, Aghion, and Meghir (2006) argue that tertiary schooling is key for developed countries (see also Aghion and Howitt (2006)). They develop a model where countries close to the world technology frontier should invest in colleges and universities in order to move the frontier out through innovation. Developing countries on the other hand should invest in more basic education since they will grow by imitating the technologies of more developed countries. This conclusion represents one of the few pieces of policy advice coming out of growth modeling that apply to developed countries. The evidence, however, comes entirely from consideration of school attainment and not the more refined measures of human capital used here.

The central issue of Section 4 is whether different levels or types of skills have payoffs that differ between developing and developed countries. We separately measure basic and top skills, based on the micro data of the international achievement tests, and then

consider the implications both of high skills and of tertiary schooling on growth. The results suggest that basic skills have substantial growth payoffs in OECD countries, and that if anything, the return to top skills is lower, not higher, in the OECD – i.e., opposite of the innovation/imitation hypothesis. We do not find a specific role of tertiary attainment for OECD growth once the focus is on long-run growth and direct measures of skills are taken into consideration.

Results of the regression analyses suggest a strong and robust effect of cognitive skills on economic growth in OECD countries, but the regression analyses by themselves do not describe the path of benefits from any program of changing the skills of the population. In the next part of the paper (Section 5), we therefore perform projections of the economic value of a set of education reform alternatives for each OECD country under different parameter assumptions. Our focus is on understanding the dynamics of economic impacts of educational programs. Three elements of the dynamics are particularly important for our consideration: programs to improve cognitive skills through schools take time to implement; the impact of improved skills will not be realized until the students with greater skills move into the labor force; and the economy responds over time through the development and implementation of new technologies.

In order to capture these elements, a simple simulation model is developed. Specifically, we evaluate the economic outcomes of a series of plausible policy programs including improving student performance by 25 PISA points (1/4 standard deviation); bringing all countries in the OECD up to the level of Finland; and bringing all students in OECD countries to minimum proficiency (400 points on the PISA tests).

The present value of the reform efforts vary by country, depending on current economic performance and current educational performance. However, the simulation exercises suggest that the aggregate gains across all OECD countries range from \$90-275 trillion for the different policy alternatives. These gains, for example, far exceed the level of stimulus funds in the current global recession.

We also provide detailed sensitivity analyses of the simulations with respect to a range of alternative specifications and parameter choices, including use of endogenous-growth and neoclassical-type specifications of the growth model; consideration of alternative coefficient estimates from the growth regressions; assessment of the lower and upper bound of the significance band around the growth-regression estimates; alternative time horizons for the future effects considered; alternative durations of the reform; alternative discount rates and growth rates of potential GDP; and alternative durations of working life.

Knowledge of the potential gains from improving schools does not, however, indicate what should be done to obtain these results. In fact, school improvement has been high on the policy agenda of a large number of OECD countries, but the results of actions have many times fallen short of expectations. Emerging research results suggest that there are general policies that can promote significantly higher achievement. The final section considers policy options that could support achievement improvements along the lines of the simulations.

2. ECONOMIC MODELS AND BASIC RESULTS

The basic theoretical and empirical framework follows developments of growth modeling over the past two decades. In this work, human capital typically plays a prominent role, but we present evidence that the measurement of human capital dramatically alters the assessment of its importance. Building on our database of international educational outcomes, we provide basic results on the relationship between educational outcomes and long-run growth in OECD countries.

2.1. Conceptual framework

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.² 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 modeling, and it is a standard element of any empirical work.

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 somewhat different empirical specifications for any statistical modeling. 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.

Box 1. Education in theories of economic growth

Theoretical models of economic growth have emphasized different mechanisms through which education may affect economic growth. As a general summary, three theoretical models have been applied to the modeling 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 modeling follows a standard characterization of an aggregate production function where the output of the macro economy is a direct function of the capital and labor 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

² 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).

movement of the economy over time. The sources of this technological change, although central to understanding 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 labor 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).

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 growth rates in GDP per worker (or per capita) to the level of education.

A final view of education in production and growth centers 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.

Our empirical analysis adopts a very simple growth model that combines elements of a general “endogenous growth” framework and a basic “augmented neoclassical” approach. Specifically, we model a country’s growth rate as a function of the skills of workers and other factors 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.

$$growth = \alpha_1 \text{ human capital} + \alpha_2 \text{ other factors} + \varepsilon \quad (1)$$

In this formulation, nations with more human capital tend to continue to make greater productivity gains than nations with less human capital. 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, by including the initial level of income among the control variables, our model does allow for conditional convergence, a leading feature of the “augmented neoclassical” approach, the commonly suggested alternative view.

The empirical 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 human capital aspects of growth models.³ Initial analyses employed school enrollment ratios (e.g., Barro (1991); Mankiw, Romer, and Weil (1992)) as proxies for the human capital of an economy. An important extension by Barro and Lee (1993, 2010) 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 (after an East Asian dummy) on growth in GDP per capita in 1960-1996.

Nevertheless, we believe that these formulations introduce substantial bias into the estimation. Average years of schooling is 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 South Africa is assumed to create the same increase in productive human capital as a year of schooling in Korea. Additionally, formulations relying on this measure assume that formal schooling is the primary (sole) source of education and that variations in non-school factors have negligible effects on education outcomes and skills. 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.

To see this, consider a standard version of an education production function as employed in a very extensive literature (see Hanushek (2002) for a review), where skills are expressed as a function of a range of factors:

$$\text{human capital} = \beta_1 \text{family inputs} + \beta_2 \text{schooling inputs} + \beta_3 \text{individual ability} + \beta_4 \text{other factors} + v \quad (2)$$

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 this literature, see, e.g., Topel (1999); Temple (2001); Krueger and Lindahl (2001); Sianesi and Van Reenen (2003).

In general, human capital combines both school attainment and its quality with the other relevant factors including education in the family, labor market experience, health, and so forth.

Thus, while school attainment has been convenient in empirical work because of its ready availability across countries, its use as a proxy for human capital is very restrictive. Not only does it ignore differences in school quality, but also any other important determinant of people's skills. Still, human capital is a latent variable that is not directly observed. To be useful and verifiable, it is necessary to specify its measurement.

A more satisfying alternative is to incorporate variations in cognitive skills, which can be determined by results of international assessments of mathematics, science, and reading achievement, as a direct measure of the human capital input into empirical analyses of economic growth. The focus on cognitive skills has a number of potential advantages. First, it captures variations in the knowledge and ability that schools strive to produce and thus relates the putative outputs of schooling to subsequent economic success. Second, by emphasizing total outcomes of education, it incorporates 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), it opens the investigation of the importance of different policies designed to affect the quality aspects of schools.⁵ Fourth, it is practical because of the extensive development of consistent and reliable cross-country assessments.

2.2. Data

Our analysis relies on the measures of cognitive skills developed in Hanushek and Woessmann (2009). These 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 people in the labor force.⁶

Between 1964 and 2003, twelve different international tests of math, science, or reading were administered to a voluntarily participating group of countries (see Hanushek and Woessmann (2010) for a review). 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 math of 15-year-olds in 2000 as a part of

⁵ Some recent work has introduced the possibility that noncognitive skills also enter into individual economic outcomes (see importantly Bowles, Gintis, and Osborne (2001), Heckman, Stixrud, and Urzua (2006), and Cunha, Heckman, Lochner, and Masterov (2006)). Hanushek and Woessmann (2008) integrate noncognitive skills into the interpretation of general models such as above and show how this affects the interpretation of the parameter on school attainment and other estimates. While there are no agreed-upon measures of noncognitive skills, at the aggregate level they might well be incorporated in "cultural differences." The importance of such cultural differences in empirical models of growth is considered in Hanushek and Woessmann (2009).

⁶ With few exceptions direct measures of achievement of people in the labor force are unavailable, and analysis instead must rely upon skills measured during the schooling period. The one exception with measures of the cognitive skills of people in the labor force is the 1994-98 International Adult Literacy Survey (IALS), which tested representative samples of people aged 16-65 years. Coulombe, Tremblay, and Marchand (2004) have used these data to construct synthetic cohorts in order to estimate an augmented neoclassical growth model across 14 countries. As discussed below, this construction causes no problems if the relative performance of people in different countries has remained constant, but it could introduce problems if that is not true.

the Programme on International Student Assessment). 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 math and science than in reading, and a majority of the international testing has focused on math and science. Each test is newly constructed, until recently with no effort to link to any of the other tests.

The methodology used to construct consistent measures at the national level across countries relies on empirical calibration of the different tests. By transforming the means and variances of the original country scores (partly based on external longitudinal test score information available for the United States), each is placed into a common distribution of outcomes (see Hanushek and Woessmann (2009) for details).⁷ Each age group and subject is normalized to the PISA standard of mean 500 and individual standard deviation of 100 across OECD countries, and then all available test scores are aggregated to the country level.

We interpret the test scores as an index of the human capital of the populations (and workforce) of each country. This interpretation of our averages over different cohorts is reasonable if a country's scores have been stable over a long period, implying that estimates from the current school-aged population provide an estimate of the older working population. If scores (and skills) do in fact change over time, some measurement error is clearly introduced. We know that scores have changed some (Hanushek and Woessmann (2009)), but within our period of observations it still appears that the differences in levels dominate any intertemporal score changes.⁸ Nonetheless, any measurement error in this case will tend to bias downward the estimates of the impact of cognitive skills on growth, so that our estimates of economic implications will be conservative.

The data on GDP per capita and its growth for our analyses come from the Penn World Tables (Heston, Summers, and Aten (2002)). Data on quantitative educational attainment are taken from the latest version of the Barro and Lee (2010) database. Additional measures of specific control variables will be discussed in the specific sections below.

Table 1 provides basic descriptive statistics on the combined measure of educational performance and the underlying economic data. We have already discussed the wide variation in growth rates across OECD countries. What is also clear from Table 1 is that both school attainment and test scores vary widely, suggesting directly that any impact of these human capital measures on growth differences should be easily detected.

⁷ 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. A look at the underlying data reveals that the distribution of PISA scores within the OECD is normal, even if the distributions in individual countries may not be. In the analysis of minimal and top skills below, the calculations employ the empirical micro distributions for each of the countries and do not assume normality.

⁸⁸ The subsequent simulations that investigate the impact of changing achievement for each country assume that the increments to achievement change for each cohort affects the labor force skills proportionally to their weight in the age distribution.

2.3. Basic results for OECD countries

We use the data on educational outcomes to estimate cross-country regressions that describe long-run growth for OECD countries. This follows a growing literature which, over the past ten years, demonstrates that consideration of cognitive skills dramatically alters the assessment of the role of education and knowledge in the process of economic development. Analyzing growth in 1960-1990 for a sample of 31 countries with available data (including 18 OECD countries), Hanushek and Kimko (2000) first showed a statistically and economically significant positive relationship between cognitive skills and economic growth. This relationship between cognitive skills and economic growth has been subsequently confirmed in a range of studies with different focuses.⁹ Most recently, Hanushek and Woessmann (2009) extend the empirical analysis to incorporate 50 countries that have participated in one or more international testing occasions between 1964 and 2003 and have aggregate economic data for the period 1960-2000. We use that database for our analysis focused on OECD countries.

As a starting point for our analyses, we replicate the basic analysis, only replacing the extended version of the Cohen and Soto (2007) data on years of schooling by the newly available latest version of the Barro and Lee (2010) database on years of schooling. Our sample contains the 24 OECD countries with available data. From the total of 30 OECD countries, the sample misses four countries – the Czech Republic, Hungary, Poland, and the Slovak Republic – because their communist history prevents them from having internationally comparable economic data during the period of analysis in the underlying database. In addition, Germany drops out because of missing economic and test score data for the Eastern parts before 1990, and Luxembourg is left out as a small country with a population of less than one million, as is common in growth analysis (see Mankiw, Romer, and Weil (1992)).

Table 2 presents the basic results on the association between educational outcomes and long-run economic growth in the sample of OECD countries. The inclusion of initial GDP per capita in all specifications simply 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. (This “convergence term” enters into the subsequent projections of economic impacts from school reform, because it suggests that any differences in growth rates from changes in cognitive skills will eventually die out; see Section 5 below).

When the cognitive-skill data are ignored (column (1)), years of schooling in 1960 are significantly associated with average annual growth rates in real GDP per capita in 1960-2000, after controlling for the initial level of GDP per capita. However, once cognitive skills are included in the model (column (2)), the whole explanatory power is taken over by cognitive skills. Cognitive skills are highly significantly associated with economic growth. At the same time, the association between years of schooling and economic growth becomes statistically insignificant and drops to close to zero. Furthermore, the

⁹ Important studies include, for example, Barro (2001), Woessmann (2003b), Bosworth and Collins (2003), and Ciccone and Papaioannou (2009). See Hanushek and Woessmann (2008, 2010) for reviews.

OECD-sample growth variance explained by the model increases from 56 percent to 83 percent when measuring human capital by cognitive skills rather than years of schooling. Note that in the OECD sample, the bivariate association with initial per-capita GDP already accounts for 49 percent of the variance in subsequent growth, making the relative increase in understanding non-convergence growth through cognitive skills substantial.

The estimated coefficient on cognitive skills implies that an increase of one standard deviation in educational achievement (i.e., 100 test-score points on the PISA scale) yields an average annual growth rate over 40 years that is 1.86 percentage points higher. This historical experience suggests a very powerful response to improvements in educational outcomes, particularly when compared to the average 2.2 percent annual growth within the OECD over the past two decades.

Figure 2 depicts the fundamental association graphically, plotting growth in real per-capita GDP between 1960 and 2000 against average test scores after allowing for differences in initial GDP per capita and initial average years of schooling. With the slight exceptions of New Zealand (below the regression line) and the United States (above) – to which we return below – the OECD countries align closely along the regression line that depicts the positive association between cognitive skills and economic growth.

Column (3) of Table 2 reports the same model excluding years of schooling, whose effect could not be significantly differentiated from zero. The point estimate on cognitive skills, as well as the adjusted R^2 , increases slightly in this reduced model.

Columns (4) and (5) break down the analysis into the 20-year sub-periods of 1960-1980 and 1980-2000. The positive association of growth with cognitive skills is clearly visible in both sub-periods, with the point estimate slightly larger in the later period.

To reduce concerns of reverse causality between economic growth and quantitative schooling investments, the basic model uses school attainment in 1960 before the growth period, but combines that with average test scores over the entire period. However, the results are hardly affected by using average years of attainment across the period (column (6)).

Because of the limited sample size, we want to ensure that results are not driven by individual outliers. Columns (7) and (8) show the robustness of results to excluding specific countries. Column (7) excludes Mexico and Turkey, two countries at the bottom of the sample of OECD countries today in terms of measures of GDP per capita, socioeconomic background, and educational spending. While the point estimate on cognitive skills is slightly reduced, the association remains strong and statistically highly significant. The same is true when excluding Korea (column (8)), a country with extraordinary conditional test scores and growth experience (see Figure 2).

Finally, columns (9) and (10) provide tests for differences in the education-growth nexus between OECD and non-OECD countries. The sample now is the full sample of 50 countries with data on test scores and economic growth. In the model without cognitive skills (column (9)), the significant association between years of schooling and economic growth does not differ significantly between OECD and non-OECD countries.

However, the OECD dummy is marginally significant at the 10-percent level and positive, indicating a remaining growth advantage of OECD countries unexplained by the model. But once cognitive skills are included in column (10), neither the OECD dummy nor its interaction with cognitive skills are statistically significant, indicating that the OECD countries actually fit well within the rest of the world on this association.¹⁰ We return to this issue when considering the separate impacts of economic institutions.

3. CAN INSTITUTIONS EXPLAIN DIFFERENCES IN RICH-COUNTRY GROWTH?

To understand to what extent consideration of institutional differences across OECD countries contributes to long-run growth differences and alters the assessment of the importance of educational outcomes, this section focuses on fundamental institutional measures of property rights and openness as well as measures of the regulation of product and labor markets in our OECD-country growth models.

3.1. Property-rights and free-trade institutions

Most economists believe that fundamental economic institutions are important for a well-functioning economy, and by implication for economic growth. The review in Acemoglu, Johnson, and Robinson (2005) strongly supports this premise. Many studies have by now confirmed that institutions such as security of property rights and openness to international trade help to understand long-run differences in economic well-being between developed and developing countries. However, the extent to which such institutions also help to explain long-run growth differences among developed countries is less clear. Indeed, most OECD countries do not differ much in this respect, as they have all reached high levels of property-rights security and international openness. While these institutional factors may have contributed historically to their development, the lack of variation across countries would indicate that they lack power in explaining growth differences with the OECD.

To test whether institutional differences contribute to our understanding of long-run growth differences between OECD countries and whether their consideration alters our result on the importance of educational outcomes, Table 3 enters basic institutional measures into our growth models. Two measures have been most consistently found to be associated with growth and income differences in global country samples. The first is the measure of security of property rights used, among others, by Acemoglu, Johnson, and Robinson (2001). The measure is an index of the protection against expropriation risk, averaged over 1985-1995, from Political Risk Services, a private company which assesses the risk that investments will be expropriated in different countries. The measure is scaled from 0 to 10, with higher values corresponding to higher security of

¹⁰ The OECD dummy and the interactions are also not significant if both interactions – with years of schooling and with cognitive skills – are included together.

property rights. The second is the measure of openness proposed by Sachs and Warner (1995, 1997). The index reflects the fraction of years between 1960 and 1998 that a country is classified as having an economy open to international trade, based on five factors including tariffs, quotas, exchange rate controls, export controls, and whether or not a socialist economy (see also Lucas (2009) for a recent application).

Without considering cognitive skills, protection against expropriation is significantly associated with economic growth across the OECD countries between 1960 and 2000 (column (1) of Table 3). However, when cognitive skills are included in the model, the coefficient on the institutional variable drops substantially in size and becomes statistically insignificant, whereas the coefficient on cognitive skills remains highly significant and close to our previous models without institutional controls (column (2)). The statistical model clearly favors educational outcomes over institutions as a fundamental source of long-run growth differences between OECD countries. This result supports the view expressed in Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004) that human capital may be the more basic source of growth than institutions.¹¹

Conceptually, it is an open question whether any connection between human capital and institutions stems from human capital causing better institutions or the opposite. But at least within the group of OECD countries, that part of institutional variation that is not related to cognitive skills is not related to long-run growth, whereas that part of skill variation that is not related to institutions remains a strong predictor of long-run growth.

Column (3) adds the measure of openness to the model. Property-rights security and openness to trade are individually and jointly insignificant in predicting long-run OECD growth, whereas cognitive skills remain strongly significant. Again, the insignificance in the institutional measures does not mean that institutions are unimportant for long-run growth. Rather, they point to the fact that the OECD countries share broadly similar institutions, so that this kind of institutional variation is unlikely to account for much of the substantial variation in long-run growth in this rich-country sample. For example, while there is some variation in the openness measure among OECD countries, it is very specific and limited. Most OECD countries are coded as open throughout the period of observation. The exceptions are Mexico, New Zealand, and Turkey that had substantial periods of being more closed, but the differences between the openness of these and the remainder of the OECD explain little of OECD growth differences.

This argument is underlined by results of the full-country sample (column (4)). The institutional measures enter jointly significantly in explaining long-run growth differences among the 50 countries with available data. However, as the specification reported in column (5) indicates, there is a significant difference in the institutions-growth nexus between the OECD countries and the non-OECD countries. The interaction between an OECD indicator and protection against expropriation brings the institutional effect close to zero in the OECD sample, and the difference is marginally

¹¹ The analysis of Acemoglu, Johnson, and Robinson (2001) has provided considerable support for the power of economic institutions based on their development of instruments based on historical colonization patterns. An underlying argument is that colonists could bring with them a set of fundamental institutions that can be useful in explaining current institutions. Glaeser, La Porta, Lopez-de-Silanes, and Shleifer (2004) argue that the colonists not only brought knowledge of institutions but also human capital – and that the human capital might be more fundamental to the growth process.

significant (at the 13 percent level). The OECD indicator and the interaction jointly reach statistical significance at conventional levels.¹²

In sum, property-rights and free-trade institutions help us understand long-run growth differences between rich and poor countries, but they do not contribute to our understanding of long-run growth differences within the group of rich countries. By contrast, the significant effect of cognitive skills on long-run growth in the OECD sample is robust to the inclusion of the institutional measures.

3.2. Regulation of product and labor markets

While the variation in fundamental property-rights and free-trade institutions may be limited among OECD countries, a substantial literature has stressed significant differences in how OECD countries regulate their product and labor markets. For example, Nicoletti and Scarpetta (2003) show that short-run growth experiences across OECD countries are related to product market regulations, and Cingano, Leonardi, Messina, and Pica (2010) find that employment protection legislation is associated with firm-level investment and other firm outcomes across sectors and across firms with different financial constraints.

The question is to what extent the effects found for short- to medium-run growth experiences and for between-industry and between-firm growth carry over to the long-run growth experiences of countries. For example, effects of product market regulations on sectoral productivity may lead to structural change and international specialization, thereby reducing any net effects on aggregate growth rates. Similarly, differences in employment protection legislation may lead to differential growth experiences in booms versus recessions that cancel out over the business cycle.

To test the importance of product and labor market regulations for long-run OECD growth, we add a rich set of regulatory measures that the OECD has developed over the past decade to our growth models. Specifically, we employ the latest versions of these far-ranging indicators of regulations of both product and labor markets. These are documented for product market regulations (PMR) in Venn (2009), who updates previous versions since Organisation for Economic Co-operation and Development (1999), and for employment protection legislation (EPL) in Wölfl, Wanner, Kozluk, and Nicoletti (2009), who update previous versions since Nicoletti, Scarpetta, and Boylaud (2000).¹³

The results, reported in Table 4, are unambiguous and telling: Not a single of the large battery of measures derived to depict product and labor market regulations comes close to being significantly related to the variation in long-run growth experiences across OECD countries. At the same time, regulatory practices do not affect the result that

¹² An interaction between openness and the OECD indicator does not reach significance when added to this model.

¹³ For details on the measures of product market regulation and employment protection, see www.oecd.org/eco/pmr and www.oecd.org/employment/protection, respectively.

educational outcomes are a powerful predictor of long-run growth differences among OECD countries.

Specifically, the first six columns of Table 4 employ aggregate and sub-indicators of product market regulation. Each indicator is scaled from 0 to 6, with higher values reflecting more restrictive policies. The first column uses the most aggregate indicator of product market regulation, which is constructed from a total of 18 low-level indicators. The coefficient is insignificant and close to zero. The same is true for the two broad sub-categories of administrative regulation and domestic economic regulation (columns (2) and (3)). The next three columns report results for the specific sub-indicators referring to measures of state control, barriers to entrepreneurship, and barriers to trade and investment, respectively. None of the measures is significantly associated with long-run OECD growth.

Additional specific robustness specifications (not shown) confirm this basic result. For example, the results hold for all underlying sub-indices, including indicators of public ownership, involvement in business operations, regulatory and administrative opacity, administrative burdens on start-ups, barriers to competition, and explicit and other barriers to trade and investment. The available indicators of product market regulation refer to 1998, the first year for which they are available. While earlier measures would be preferable, these should still capture the most basic overall patterns, under the assumption that main institutional variation is in the cross-section. At least, results are unaffected by using the available indicators for 2003 or 2008 instead, or by taking the average over the three observations, indicating that lack of results is not driven by single measurement error. In addition, to align the regulatory measures more closely with the period of growth observations, we performed all regressions for growth between 1980 and 2000. Results were unaffected.

The final five columns of Table 4 employ aggregate and sub-indicators of labor market regulation, which again range from 0 to 6 with higher values reflecting more restrictions. Columns (7) and (8) add the two versions of the aggregate employment protection index suggested by the OECD to the model. The first version combines regulations of regular employment contracts and of temporary contracts, and the second version adds sub-indices of additional regulation of collective dismissal to this. Neither measure enters the model significantly or affects the estimate on cognitive skills. The same is true when the three sub-indicators of protection of permanent workers against (individual) dismissal, strictness of regulation on temporary forms of employment, and specific requirements for collective dismissal separately (columns (9)-(11)).

The indicators of employment regulation are measured as averages of the annual values between 1985 and 2000. Results are similar when the growth period is restricted to 1980-2000, which aligns more closely to the period of observation of the regulatory measures, is used instead of 1960 to 2000 (not shown). Results are also robust to the new, third version of the aggregate OECD employment protection index, available only in 2008, which adds the maximum time to make a claim of unfair dismissal, authorization and reporting requirements for temporary work agencies, and regulations

requiring equal treatment of regular and agency workers at the user firm as three new sub-indicators.

Beyond the regulatory measures reported in Table 4, Nicoletti and Pryor (2006) survey a total of 16 objective and subjective measures of governmental regulation from four different studies, including Botero et al. (2004), Kaufmann, Kraay, and Zoido-Lobaton (1999), Pryor (2002), and earlier versions of the indicators used above. While the above objective indicators are based on extensive reviews and quantifications of laws and regulations, other more subjective indicators are based on observational data of experts familiar with the regulations. The different indicators also cover different sets of regulations on product and labor markets, as well as overall business regulations more generally. Not a single of the regulation measures is significantly associated with long-run growth across OECD countries in our basic model (not shown). These results confirm that the lack of a significant association with long-run OECD growth does not hinge on a specific objective or subjective method of quantifying institutional realities or on a specific area of regulations.

The results of including institutional and regulatory measures in the growth regressions rule out that the strong association of educational outcomes with growth is just driven by omitted causal institutional factors. In fact, the analysis provides no convincing evidence that institutional or regulatory differences can account for differences in long-run growth among rich countries. Instead, cognitive skills emerge as the one strong policy factor underlying growth differences across OECD countries.

4. DIFFERENT LEVELS OF SKILLS AND EDUCATION

An important and recurring policy question is which level of skills and education is most decisive for OECD growth. We analyze several dimensions of this: whether there are differential returns across countries to average skills; whether basic skills are more or less important than top levels of skills; and whether there is a specific role of tertiary attainment for OECD growth.

4.1. Differential returns to average skills

As a first dimension of heterogeneity, we analyze whether the effect of average educational outcomes differs along specific dimensions. First, the graphical plot of Figure 2 suggests that there is no obvious non-linearity in the test score-growth association across OECD countries. Thus, a squared test-score term does not enter the model significantly, and an exponential test-score specification does not improve the fit of the model (not shown).

Second, there is no obvious difference in the effect of average test scores between countries with initially low vs. high income. While the small sample size does not allow for extensive interaction models, an interaction of test scores with an indicator for above-median initial GDP per capita does not enter the model significantly (not shown).

Third, Figure 3 plots quantile regression estimates in 5 percent steps of the effect of average test scores for percentiles of the growth distribution. It is evident that the effect is relatively constant across the whole distribution of growth residuals. In fact, all quantile regression point estimates fall within standard confidence intervals around the OLS estimate. The estimated average effect thus provides a good representation of the effect across the full range of quantiles of the conditional growth distribution and is not driven by specifics in certain parts of the conditional distribution.

The combined results of non-linear specifications, skill interactions, and quantile regressions suggest no obvious patterns of differential returns to average cognitive skills in the sample of OECD countries.

4.2. Basic vs. top skills

A leading policy question refers to the effects of different ranges of the skill distribution. Should developed countries implement relatively egalitarian education policies or rather relatively elitist policies? Should they focus on decent basic skills for the whole population or on nurturing top scientists and engineers? To address such questions, we make use of the micro data of each of the international achievement tests to calculate the population shares in each country that reach a threshold of basic skills and a threshold of top skills (see Hanushek and Woessmann (2009) for details). Using such distributional information allows us to look at different dimensions of the skill distribution separately.

Conceptually, Vandenbussche, Aghion, and Meghir (2006) assume that the innovation process is more intensive in high-skill labor than the imitation process. They then present an endogenous growth model with innovation and imitation where high-skill labor has a greater growth-enhancing effect for countries closer to the technological frontier, whereas countries further from the technological frontier get greater value from what they call “unskilled human capital”. While the untested underlying assumption seems reasonable, there are also reasonable arguments to be made for an opposite assumption – in which case the main prediction of the model would be turned around. In particular, the innovation literature points out that many innovations emerge from lucky coincidences. By contrast, almost by definition purposeful imitation processes require the employment of skilled scientists.

Another conceptual extension starts from the perspective of a high-skilled scientist. If this scientist were to work in a country that produces at the technological frontier, his only option is to use his skill in the innovation of new technologies. If the scientist, by contrast, were to work in a country that produces far below the technological frontier, he also has the option to employ his skills in such innovative activities, but he also has the alternative additional option to employ his skills in imitating the more productive technologies currently employed at the technological frontier. This scientist will tend to choose to employ his skills in the activity that promises the higher benefits. In such a setting, the return to high-skill labor cannot be smaller below the technological frontier

than at it, and it may well be larger. While concentrations of high-skilled labor and spillovers across them may still be important, the alternative perspectives do introduce questions about the underlying assumptions.

Given the conflicting conceptual predictions, the issue warrants a new empirical look. In particular, while the empirical application in Vandenbussche, Aghion, and Meghir (2006) measures growth in 5-year periods, which may be substantially influenced by business cycles and idiosyncratic shocks across 19 OECD countries, we keep our focus on differences in long-term growth experiences across the OECD countries. In addition, while Vandenbussche, Aghion, and Meghir (2006) conceptualize the difference between high- and low-skill labor as school attainment at the tertiary vs. non-tertiary level, our alternative focus on top vs. basic skills may provide an empirical distinction particularly relevant for imitation and innovation processes.

We start by replicating the results in Hanushek and Woessmann (2009) of incorporating both the share of students who reach a basic level of skills (a score of at least 400 on the tests) and the share of students who reach top-level skills (a score of at least 600) in a growth regression of the full 50-country sample of OECD and non-OECD countries (column (1) of Table 5).¹⁴ Both skill dimensions enter the model significantly, but the point estimate on the top-skill dimension is substantially higher. A ten percentage-point increase in the basic-skill share is associated with 0.3 percentage points higher annual growth, and a ten percentage-point increase in the top-skill share is associated with 1.3 percentage points higher annual growth. (Note that this does not necessarily provide an estimate of the relative importance of the two skill dimensions, as it may be much more feasible to increase the basic share than to increase the top share by the same amount; this might be suggested by the fact that the international standard deviation of the basic skill percentage is about four times as large as that of the top-skill level).

When estimating the same model on the OECD sample, though, the point estimate on the top-skill share is only a fourth of the one estimated in the full country sample and loses statistical significance (column (2)). By contrast, the point estimate on the basic-skill share is slightly larger than in the full country sample, and remains highly significant. The specification of column (3) shows that the difference in the estimate on the top-skill share between OECD and non-OECD countries is statistically significant. Of course, the measures of the two skill dimensions are highly collinear (their correlation is 0.73 in the full sample and 0.70 in the OECD sample), limiting precision in the joint specification. Results in columns (4)-(9), however, reveal that the pattern of results is similar when entering one of the two measures at a time.

While small samples and collinearity obviously suggest caution in the interpretation of these results, it seems clear that basic skills are relevant for OECD-country growth. Furthermore, if anything, high-level skills appear more important in non-OECD than in OECD countries, rather than the other way around.

¹⁴ These scores are one standard deviation below and above the OECD mean, respectively. The OECD also identifies five levels of skills on each of its tests (see, for example, Organisation for Economic Co-operation and Development (2004)). The 400 falls in the range of level 1 while the 600 falls in the range of level 5.

4.3. Non-tertiary vs. tertiary schooling

The available test score measures – for average, basic, and top skills – are measured at the primary and secondary level of schooling. However, a lot of policy discussion focuses on another dimension of schooling, namely whether developed countries should place a particular focus on primary, secondary, or tertiary education. Vandebussche, Aghion, and Meghir (2006) suggest that countries close to the technological frontier should rather emphasize tertiary education. To investigate this dimension, we make use of the new, improved Barro and Lee (2010) database which provides average years of schooling separately at the primary, secondary, and tertiary level. With little meaningful variation in the completion of primary education across the bulk of OECD countries, we combine the two basic levels of schooling into one category of non-tertiary schooling.

In the full-country sample, the coefficients on non-tertiary and tertiary schooling are both close to zero when cognitive skills are controlled for (column (1) of Table 6). By contrast, in the OECD sample, the point estimate on years of tertiary schooling becomes larger (column (2)), and reaches marginal significance (at the 10 percent level) when years of non-tertiary schooling are not included in the model (column (3)). However, results in column (4) indicate that this is completely driven by the United States. Once the United States is excluded, the coefficient on tertiary education is much smaller and insignificant again. The United States is well-known for its extensive tertiary education system, and Figure 2 already indicated that the United States has the strongest positive residual in the growth model. While this might be an indication of growth-enhancing effects of its high-quality higher-education system, the lack of robustness in the sample without the United States suggests that it might rather be an indication of the high-skilled immigrant population that it attracts, of a particular set of economic institutions (not captured by our institutional measures), or of any other idiosyncrasy of the US economy.¹⁵

An additional interesting pattern emerges when employing the measures of different skill dimensions. When the basic-skill share is used as an alternative skill measure to the average skills, years of tertiary schooling reach significance (column (5)). When, by contrast, the top-skill share is used, the coefficient on years of tertiary schooling becomes smaller and loses any significance (column (6)). This pattern is indicative that years of tertiary education proxy for the share of students with high-level skills. When the United States is disregarded, the coefficient on years of tertiary schooling declines to close to zero in this specification (column (7)).

While the possibility to distinguish the effects of different dimensions of the skill and schooling distribution again is limited in these small samples, some basic patterns prove clear. First, the significant effect of cognitive skills is extremely robust to consideration of any quantitative measure of different levels of school attainment. Second, the finding of Vandebussche, Aghion, and Meghir (2006) of a particular effect of tertiary attainment in rich countries is not robust once the focus is on long-run growth

¹⁵ See Aghion et al. (2010) for additional analyses of tertiary education in Europe and the United States.

experiences and educational outcome measures are taken into account. Of course, this does not mean that learning beyond the secondary level does not matter. Rather, in the spirit of a lifecycle interpretation where early skills facilitate the development of subsequent skills (Heckman (2006); Cunha, Heckman, Lochner, and Masterov (2006)), it means that outcome measures of learning in school are a good predictor for the accumulation of further skills in life and the capacity to deploy these skills effectively.

5. CALCULATING THE ECONOMIC VALUE OF EDUCATION REFORMS

The results so far indicate that educational outcomes have a strong effect on long-run economic growth of OECD countries. However, they do not tell us directly how much improvements in educational outcomes would actually be worth. In particular, the growth-rate effects do not map linearly into the economic value of any education reform in a country, not least because different time lags are involved between successful reform in the education system today and the improvement of skills in the national workforce.

In this section, we therefore perform simulation analyses that use the estimates from the previous sections to project what the results mean for the economic impact of different scenarios of school improvements in the OECD member states. The projections presume that the estimated impacts of cognitive skills on growth are causal in nature – that is, changing the achievement of a country's population will lead to improved growth. Of course, considerable controversy surrounds such cross-country growth regressions and any causal interpretation of them (see, for example, Levine and Renelt (1992), Levine and Zervos (1993), and Bils and Klenow (2000)). In other work, we have considered a series of analyses aimed at eliminating many of the natural concerns about the identification of the causal impacts of cognitive skills (Hanushek and Woessmann (2009)). Each of the analyses points to the plausibility of a causal interpretation of the basic models. Nonetheless, with our limited international variations, it is difficult to demonstrate identification conclusively. Therefore, here we consider a variety of sensitivity analyses designed to show the impact on outcomes when part of the estimated effects is non-causal.

We consider three specific reform scenarios: first, improving average student performance 1/4 standard deviation, or 25 PISA points, in each country; second, bringing all OECD countries up to the level of the PISA top-performer, Finland; and third, bringing all students in OECD countries to minimum proficiency, defined as 400 points on the PISA test scale. For each scenario, we aim to project the future path of development of GDP in a country to calculate the economic value of the education reform in present-value terms.

One final issue of the projections is important. Because we consider economic outcomes far into the future, the precise form of the underlying growth model potentially makes a noticeable difference. In particular, an endogenous growth model that implies constant growth effects from improved skills will differ from a neoclassical growth model that sees changes in skills as affecting the level of income but not the long-run

growth rate of an economy. Our previously estimated models support analysis of this range of underlying models, and applying the alternatives gives another idea of the bounds on future economic effects from school improvement.

5.1. The projection model

The projection of the total value of any education reform involves several components. First, we calculate the time path of the annual growth rate engendered by education reform designed to move students from their current performance to a given new level. This pattern of economic outcomes represents the confluence of three separate dynamic processes: (1) Changes in schools lead to the progressive improvement in student achievement until students fully reach the new steady-state level of achievement; (2) students with better skills move into the labor force and the average skills of workers increase as new, higher achieving workers replace retiring workers; and (3) the economy responds to the progressive improvement of the average skill level of the workforce. Second, based on the pattern of predicted growth rates, we model the future development of GDP with and without the education reform. Third, based on these projections, we calculate the total value of the reform by aggregating the discounted values of the annual differences between the GDP with reform and the GDP without reform.

Our initial depiction is based on the endogenous-growth framework where higher test scores yield a permanent increase in the long-run growth rate. We subsequently consider a neoclassical framework where the additional convergence term makes the annual growth rate a negative function of the (log) level of GDP reached in the previous period. This alternative implies that the additional growth due to higher test scores is only transitory, leading to a new higher income steady-state path but one where the economy grows in the long run at the same rate as it would have without the reform.

5.1.1. Increase in the annual growth rate in the different phases

For expositional convenience, the description of the projections follows the basic Scenario I where each OECD country begins an educational reform program in 2010 that takes 20 years to be fully implemented and where subsequent students reach the new achievement level. The economic value of the reforms is then traced across an 80-year period (which represents the expected lifetime of somebody born in 2010). The basic set-up is easily modified to consider alternative scenarios.

a) Phase 1 (2010-2030): In the baseline simulations, the education reform program is 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:

$$\Delta^t = \text{growth coefficient} * \Delta PISA * \frac{1}{\text{working life}} * \frac{t - 2010}{20} + \Delta^{t-1} \quad (3)$$

where the *growth coefficient* stems from the regression estimations presented in the previous sections and $\Delta PISA$ is the increase in the average PISA test score due to the respective reform. The working life term indicates that each cohort of new, higher achieving students is only a fraction of the total labor force.

b) Phase 2 (2031-2050): The education reform is now fully enacted, and achievement of all subsequent students remains at the new level. But for the length of a work life from the start of reform, which in the baseline simulations is assumed to last 40 years, there are still workers with initial levels of skills that are being replaced in retirement by higher achieving workers. During this phase, the additional growth in GDP per capita in year t due to the reform is given by:

$$\Delta^t = \text{growth coefficient} * \Delta PISA * \frac{1}{\text{working life}} + \Delta^{t-1} \quad (4)$$

c) Phase 3 (2051-2070): During this phase, the first 20 labor-market cohorts – which only partially profited from the education reform – are replaced by cohorts that profited from the fully enacted education reform:

$$\Delta^t = \text{growth coefficient} * \Delta PISA * \frac{1}{\text{working life}} - (\Delta^{t-40} - \Delta^{t-41}) + \Delta^{t-1} \quad (5)$$

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 :

$$\Delta = \text{growth coefficient} * \Delta PISA \quad (6)$$

5.1.2. Development of GDP with and without reform

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

$$GDP_{no\ reform}^t = GDP_{no\ reform}^{t-1} * (1 + \text{potential growth}) \quad (7)$$

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

$$GDP_{reform}^t = GDP_{reform}^{t-1} * (1 + \text{potential growth} + \Delta^t) \quad (8)$$

In the neoclassical specification, an additional term ensures that the growth rate is negatively affected by the (log) level of GDP reached in the previous period. As a consequence, the annual growth rate without and with reform will converge to the same rate of potential growth in the long run.

5.1.3. Total effect of the reform

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

$$\begin{aligned} & \textit{Total value of the reform} = \\ & \sum_{t=2010}^{t=2090} (GDP_{reform}^t - GDP_{no\ reform}^t) * (1 + \textit{discount rate})^{-(t-2010)} \end{aligned} \quad (9)$$

In the baseline scenario, the time horizon over which future returns will be considered is the lifetime of a child that is born at the beginning of the reform, which equals the year 2090.

5.1.4. Parameter values

Implementing the projection model requires a number of parameter assumptions. Three parameter choices have already been indicated: the duration of the education reform of 20 years, the length of a working life of 40 years, and the overall time horizon of 80 years. In addition, the model requires assumptions about the growth coefficient, the rate of potential growth without the education reform, and the discount rate. Here, we outline the parameter choices for the baseline scenario. Below, we will also provide robustness analyses of results to alternative parameter choices.

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 in the student population 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.

The expected work life is assumed to be 40 years, which implies that each new cohort of workers is 2.5 percent of the workforce. Thus, even after an educational reform is fully implemented, it takes 40 years until the full labor force is at the new skill level. We are not aware of direct estimates of this parameter in the literature. However, estimates for Germany based on pension insurance data suggest values in the range of 35 to 40 years for the length of the average working life (see Woessmann and Piopiunik (2009) for details).

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 (Organisation for Economic Co-operation and Development (2009c)).¹⁶ Therefore, the baseline calculations take a time horizon until 2090, considering all future returns that accrue until then, but neglecting any returns that accrue after 2090.

In order to consider the impacts of improvement on OECD countries, the simulations rely on the estimates of growth relationships derived from the 24 OECD countries with complete data. As indicated in column (2) of Table 2, the coefficient estimate is 1.864, suggesting that, e.g., a 50 point higher average PISA score (i.e., one-half standard deviation higher) would be associated with 0.93 percent higher annual growth in the long run. This estimate clearly includes some uncertainty, a factor that is also included in the robustness analyses below. (Below we also consider a neoclassical alternative where the impact on growth rates is not assumed to be permanent but in fact falls over time according to the estimated convergence parameter until growth rates return to the current level).

The value of improvement in economic outcomes from added growth depends, of course, also 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 percent. This is simply the average annual growth rate of potential GDP per worker of the OECD area over the past two decades: 1.5 percent in 1987-1996 and 1.4 percent in 1997-2006 (Organisation for Economic Co-operation and Development (2009b)).

Finally, because economic benefits accrue at varying times into the future, it is important to recognize 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 percent (e.g., Börsch-Supan (2000), Hagist, Klusen, Plate, and Raffelhüschen (2005)), a precedent that is followed here.¹⁷ By contrast, the influential Stern Review report that estimates the cost of climate change uses a discount rate of only 1.4 percent, thereby giving a much higher value to

¹⁶ 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.

¹⁷ 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 percent for years 0-50, 2.5 percent for years 50-100, 1.5 percent for years 100-200, 0.5 percent for years 200-300, and 0 percent years over 300. (The proper starting value is actually 3.3 percent based on the parameter values they assume for the growth rate in per capita consumption (2.3 percent), the social marginal utility of consumption with respect to per capita consumption (1), and the utility discount rate (1 percent)).

future costs and benefits (Stern (2007)). In our robustness analyses, we will also consider such alternative discount rates.

A number of additional assumptions go into the projections. 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. Third, the projection of simultaneous improvement across countries presumes that all countries can grow faster without detracting from (or benefiting) growth in other countries.

5.2. Results of base scenarios

We start with the results of the baseline projection model in an endogenous-growth framework for the three education reform scenarios. All calculations are in real (inflation-adjusted) terms – 2010 dollars under purchasing power parity.

5.2.1. Scenario I: Increase average performance by 25 PISA points

A simple starting point is to consider the economic impact on OECD countries of a 0.25 standard deviation improvement, equivalent to a 25 point increase on PISA scores. The reform policy is begun in 2010 and on average yields 25 point higher scores in 2030 that remain permanently at that level for all subsequent students.¹⁸

A policy like this is uniform across countries, so the relative improvement is the same for all countries.¹⁹ Figure 4 provides a summary of the marginal 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 labor market, GDP will be more than 3 percent higher than what would be expected without improvements in human capital as early as 2041. (The figure also shows a 95 percent confidence bound of 1.9-4.1 percent higher GDP, based on the relevant 95 percent confidence bounds for the regression coefficient in column (2) of Table 2. By the end of expected life in 2090 for the person born in 2010, GDP per capita would be expected to be over 26 percent above the “education as usual” level.

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 \$3,606 billion in 2041. If on the other hand it achieved the improvement in cognitive

¹⁸ All calculations of PISA scores underlying the following simulations refer to the average performance in math and science (in line with the underlying growth model, as noted above), averaged over the three PISA cycles 2000, 2003, and 2006 (see Organisation for Economic Co-operation and Development (2001, 2003, 2004, 2007)). All underlying measures of Gross Domestic Product (GDP) are in US dollars, 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 (Organisation for Economic Co-operation and Development (2009b)).

¹⁹ 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.

skills that took it from an average PISA score of 505 to 530, total GDP would be expected to be \$3,715 in 2041, or \$108 billion higher.²⁰ These calculations illustrate a simple point: While 3 percent 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 percent improvement in 2041 rises to a 5.9 percent improvement in 2050, 15.3 percent in 2070, and 26.3 percent 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 people in the labor force have obtained the new and improved education (in 2070), annual growth will be 0.47 percentage points higher. This implies that each country that achieves the average improvement of $\frac{1}{4}$ standard deviation of achievement will have a cumulative impact on the economy through 2090 that is equal to 288 percent of current year GDP. The first column of Table 7 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 \$123 trillion in present value.

Because these are put into present value terms, they can be compared to current economic values. For example, these calculations indicate that the value of improvements through long-run growth far outstrips the cost of the current worldwide recession (and are much larger than the worldwide fiscal stimulus efforts).

5.2.2. Scenario II: Bring each country to Finland average level (of 546 PISA points)

The success of Finland on the PISA tests is well-known. In the second scenario, 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 – an average PISA score of 546.

Obviously, the amount of reform necessary varies by where each OECD country ranks on PISA today. Commensurately, the impact on different economies varies by the size of reform (in addition to the size of the economy itself). 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 completely transformed.

Columns (2)-(5) of Table 7 present 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,

²⁰ These calculations assume a constant population size over this period.

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 program to yield such large changes. We model the implications in the robustness analyses below).

The present value for OECD improvements under this scenario is \$275 trillion, or more than six times the current GDP of the OECD countries. The United States itself, which currently falls over 50 points behind Finland, would by historical growth patterns see a present value of improved GDP of over \$112 trillion, or some 40 percent of the OECD total – reflecting both the size of the country and its distance behind Finland. Germany would see a \$17 trillion improvement, or more than five times current GDP.

The rankings of countries according to increases compared to current GDP are shown in Figure 5. One interpretation of this figure is the amount of economic leverage from educational improvements that is possible for different OECD countries.

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

The final scenario considered is a “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, or 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.

In order to understand the implications of changing just one portion of the achievement distribution, we employ the alternative estimation of the underlying economic growth models of column (2) of Table 5. 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.

For these calculations, all OECD countries including Finland have room for improvement. On average, 18 percent of students in the OECD countries score below 400. As might be expected from the average scores, the required improvements are largest in Mexico and Turkey (see column (9) of Table 7).

Columns (6)-(8) of Table 7 display the economic outcomes according to historical growth patterns of bringing all OECD students up to minimum competence levels. The overall OECD change would be an average annual growth rate that was 0.8 percent higher after reform was accomplished and after the full labor force had received the improved education. The total improvements for the OECD countries from achieving universal minimum proficiency would have a present value of \$226 trillion. Again, there is a wide range of outcomes including relatively small improvements of 219 percent of current GDP for Canada as compared to nine OECD countries that would experience a benefit more than five times their current GDP.

The range of outcomes is depicted in Figure 6 that ranks countries by the benefits compared to current GDP. Even Finland could by these calculations over double its current GDP through bringing the relatively modest proportion of low performers (4.7

percent) up to scores of 400. Note also that the effects of these policies on the separate countries differ from the previous scenario, reflecting the differences in the underlying distribution of student performance.

5.3. An alternative neoclassical growth framework

The projections so far assume that higher educational achievement allows a country to keep on growing at a higher rate in the long run. Such a specification captures the basic ideas of endogenous growth theory, where a better-educated workforce leads to a larger stream of new ideas that produces technological progress at a higher rate (see Box 1 above). By contrast, in the augmented neoclassical growth model, changes in test scores lead to higher steady-state levels of income but do not affect the long-run growth path. Our empirical growth model captures the conditional convergence implied by the neoclassical model – but also by a set of endogenous growth models – through including the initial GDP level as a control variable. An alternative approach for the projections is thus to interpret the model in the neoclassical rather than endogenous-growth framework and have educational reforms affect the steady-state level of income but not its long-run growth.

To do so, we re-estimate our growth model with the logarithmic (rather than linear) per-capita GDP as control. The test-score coefficient hardly changes in this specification (1.718 rather than 1.864), and the coefficient on log initial income is -1.835. This estimated convergence rate of 1.8 percent is very close to the one expected under standard parameter assumptions in the augmented neoclassical growth model (Mankiw, Romer, and Weil (1992)). It means that (approximating around the steady state) an economy moves halfway to its steady state in about 38 years. Including this convergence process in our simulations allows us to perform projections that are in line with neoclassical growth theory. In these projections, growth rates with and without education reform will differ only during the transition to the new balanced growth path. In the long run, the economy will grow at the same rate after the reform as without the reform.

We use the estimates from this model to simulate the trajectory from the old to the new balanced growth path in each year during our time horizon. To implement the idea that the world technological frontier grows at 1.5 percent in the absence of education reform in this model, we assume that in the aggregate the three countries with the largest shares of patents in the world – the United States, Japan, and Germany – grow at 1.5 percent without reform. Together, the three countries currently account for over 70 percent of worldwide patents (measured in triadic patent families, Organisation for Economic Co-operation and Development (2008)). We thus choose a constant growth parameter for each future year that has the weighted average of the three countries grow at 1.5 percent

each year, where the weights are each country's share in their combined GDP in the previous year.²¹

Table 8 shows the results of the projections based on the neoclassical model specification. In reform Scenario I, where each country increases by 25 PISA points, the value of the reform – the discounted value of the future increases in GDP – amounts to \$90 trillion in present value terms. While this provides a neoclassical lower bound to our previous projection of \$123 trillion, the noteworthy fact is that over the time horizon of our projections until 2090 (and for a reform that takes until 2070 to take full force in the workforce), the difference between the neoclassical and the endogenous growth model has relatively minor political meaning. In either case, the value of the education reform is very large, at 288 percent or 211 percent of current OECD GDP. Note that in the neoclassical projections, the value of this reform as a share of current GDP differs across countries, because the convergence term means that projected growth rates depend on the level of GDP reached.

To illustrate the specific dynamics of the neoclassical projections, it is worth discussing a few details of the trajectories of the projection. Initial growth rates (with or without reform) in 2010-11 differ across countries, depending on how far away they are from their steady state. The United States is projected to grow at 1.1 percent initially, whereas the simple mean of OECD-country growth rates is higher at 1.9 percent due to many countries' room for catch-up.²² Due to the convergence process, average OECD-country growth is down to 1.6 percent in 2090 without the reform, ranging from 1.0-1.8 percent across countries. With reform, the range is 1.2-2.0 percent, with an average of 1.7 percent. By 2104, the average is down to 1.5 percent, and by 2130, all countries' growth rates have converged to between 1.3 and 1.6 percent without the reform and to 1.4-1.7 percent with reform.

In the neoclassical projections, the difference in average growth rates between the scenarios with and without reform grows to a maximum of 0.28 percentage points in about 2060 and then declines back to 0.18 percentage points in 2090 and 0.06 percentage points in 2150. (Compare this to the endogenous growth model where the long-run growth rate stays 0.47 percentage points higher starting in 2070.) By 2300, the difference in growth rates would be below 0.004 percentage points in each country. Thus, while the difference ultimately converges to zero, the model parameters imply that this convergence process takes a very long time to take full effect.

²¹ We also experimented with alternative ways to implement the growth of the technological frontier, and results fall into the same ballpark. The alternatives include holding the following growth rates constant at 1.5 percent: the (weighted) OECD average growth rate, the U.S. growth rate, a simple or weighted average of the U.S. and Finnish growth rates, and a specification where "technological leadership" (as depicted in per-capita GDP) turns from the U.S. to Finland in the 2030s. The latter specifications reflect the fact that Finland, the country with the highest test scores, is projected to have the highest steady-state level of per-capita GDP in our model. Note that – consistent with such a model – in terms of patenting new technologies, Finland is currently already the country with the largest revealed technological advantage in ICT, as measured by ICT-related patents in total patents (Organisation for Economic Co-operation and Development (2008)).

²² Based on our model which depicts only effects of test scores and reached levels of income, Luxembourg and Norway – the two countries with the highest current levels of GDP per capita – are projected to converge to a lower balanced growth path (both without and with reform) and thus initially have the lowest growth rates. If these countries can keep their current advantage in per-capita relative to the other OECD countries in the future for reasons outside our model, this would increase the projected value of the educational reform in these countries.

In reform Scenario II, where each country improves to the test-score level of Finland, the present value of the reform amounts to \$180 trillion in the neoclassical model, compared to the previous \$275 trillion in the endogenous growth model – or more than four times (rather than six times) current GDP. Note that this scenario implies that in the *very* long run, each country converges to the same steady-state level of per-capita GDP, as test scores are the only variable influencing the steady-state level in our model. However, in 2090 the per-capita GDP of the most advanced country would still be 70.6 percent higher than that of the least advanced country. By 2150, this difference would be down to 19.7 percent, and by 2300 to 1.2 percent.

The present value of reform Scenario III, which brings all students to a minimum level of 400 PISA points, is \$187 trillion in the neoclassical projections, rather than the \$226 trillion of the endogenous-growth type projections.²³ Again, the difference over our time horizon of projections is limited. Across the three reform scenarios, the neoclassical projection value is between 65 and 83 percent of the endogenous-growth projection value.

The difference between the two growth models depends on the time horizon chosen. The further out the projections are carried, the more substantial does the difference between the two growth models get. The final four columns of Table 8 report the present value calculated with the estimated neoclassical version for reform Scenario II with time horizons varying between 2050 and 2150, respectively. The varying time horizons clearly make a huge difference in the estimated returns to reform and illustrate how the neoclassical and endogenous growth specifications differ. Compared to the endogenous-growth projections evaluated over the same time horizons (discussed in Table 9 below), the neoclassical projection is 81 percent of the endogenous-growth projection value for a time horizon until 2050, 65 percent until 2090, and 44 percent until 2150.

Several factors contribute to the closeness of the estimates over our time period for the impact of improvements in cognitive skills based on the two different growth models. First, our reform scenarios gradually introduce changes, due to the lags for the policy to become fully effective and for the new, better-educated workers to change the average skills of the labor force. Our projections from the time after policies are fully felt involve just 20 years, a time too short to have huge differences in the implications of the alternative models. Second, the biggest impacts of the differences across the alternative models occur in the distant future, and thus the impact is lessened by discounting to obtain present values and by disregarding any returns that might accrue after 2090. Third, even ignoring discounting, the estimated convergence parameters imply very long periods before any country returns to its balanced growth path following a perturbation because of policy. Fourth, the present value of any reform keeps increasing even after full convergence has taken place and after growth returns to its pre-reform rates, because the economy is at a higher level due to the reform.

²³ In the underlying growth regression with log per-capita GDP as control variable, the coefficient on the proportion of the population with scores higher than 400 is 6.478 and the coefficient on log initial income is -2.107 (both significant at the 1 percent level).

As a consequence, for the (already long) time horizon taken here, the different dynamics of the neoclassical and endogenous growth model do not make a fundamental difference, regardless of whether human capital increases the long-run growth rate or not. The alternative set of neoclassical projections places a lower bound on the results based on the endogenous growth paradigm but retains the same order of magnitude.

5.4. Sensitivity to alternative parameter choices

The different scenarios indicate very large economic changes from educational improvement, but the estimates come for specific parametric choices. How sensitive are the results to these specific choices? We look at sensitivity in terms of reform Scenario II of bringing each country to the Finish level of PISA scores in our baseline endogenous-growth-type model specification. The qualitative implications for changes in results, however, apply similarly for the other scenarios and growth model. Remember that in the baseline specification, the total value of the reform in the OECD amounts to \$275 trillion, or 645 percent of the current GDP of OECD countries.

The baseline model is based on the coefficient for education outcomes in the most general OECD growth regression framework, i.e., 1.864 percent of additional average annual growth for a one standard deviation increase in test scores (see column (2) of Table 2). However, not least because of the limited sample size of the growth regressions, the precise estimate of the growth effect of a test score increase is an open issue. Therefore, to see how sensitive the simulations are to alternative parameter assumptions for the growth coefficient, the first four columns of Table 9 perform the same projection calculations for the lowest and highest estimated parameters for cognitive skills in the different specifications of Table 2 – i.e., 1.398 and 1.968, respectively. This leads to an estimate of the total discounted value of the education reform for the OECD of \$196 trillion and \$295 trillion, respectively (or 459% and 690%, respectively, of current GDP). Thus, while these alternative parameter choices clearly make a difference, the bottom line of the projection analysis still is the same: No matter which of the coefficient estimates is used, the aggregate value of the education reform is enormous.

An alternative way to account for the imprecision of the growth coefficient estimate, reported in columns (5)-(8) of Table 9, is to use the lower and upper bounds, respectively, of the 95 percent confidence interval around the baseline growth coefficient in the underlying growth regression. Results based on this analysis suggest that, with 95 percent confidence, the net present value of the education reform is between \$164 trillion and \$406 trillion. While these differences are economically significant, the ultimate conclusion remains: Even at the bottom end of the confidence interval, the value of the reform still dwarfs any of the current stimulus packages by several orders of magnitude.

Results so far incorporate any returns that accrue by the year 2090, but disregard any returns beyond that time horizon. Columns (9)-(12) of Table 9 report the net present value to which the reform results aggregate for time horizons to the years 2050 and

2150, respectively. As already indicated by Figure 4, this clearly makes a huge difference. When restricting the time horizon to 2050, the value of the reform is “only” \$36 trillion, and such a horizon neglects much of the very huge long-run growth effects. This indicates how important it is to adopt a long-term horizon in order to see the full effects of any education reform. Still, even by 2050, the net present value of the reform already accumulates to 85 percent of the current GDP of the OECD countries. By contrast, when adopting a time horizon until 2150, the (appropriately discounted) value of the reform sums to a staggering \$948 trillion, or more than 20 times the current GDP.

The baseline scenarios assume that it takes 20 years for the education reform to be fully implemented. The first four columns of Table 10 alternatively assume reform durations of 10 and 30 years, respectively. The faster reform implementation leads to an increase of the reform value to \$341 trillion, whereas the reform value is “only” \$223 trillion if the reform takes 30 years to implement. Thus, while faster reform efforts obviously lead to substantially higher returns, even a slow (but successful) reform, if begun today, would have enormous impact.

Columns (5) and (6) of Table 10 report results under the assumption that the average working life is 35 years, rather than 40 years, which appears to be a more reasonable estimate for many OECD countries. A shorter working life means that the replacement of the workforce with better-educated individuals completes faster, so that the aggregate value of the education reform increases. Assuming a 35 year work life yields a projection estimate of \$304 trillion of the total value of the education reform.

Finally, the rate at which future returns are discounted obviously makes a substantial difference for the net present value of a reform whose main returns do not start to pick for a whole generation. Thus, rather than using the common-practice 3 percent discount rate of the baseline model, columns (7)-(10) of Table instead use discount rates of 2.5 percent and 3.5 percent, respectively. The resulting discounted present values of the projected returns are \$369 trillion and \$207 trillion, respectively. Thus, the precise values of these long-run projections are clearly sensitive to the choice of the discount rate. For a larger band of discount rate of 2 percent to 4 percent, the total discounted reform value would be \$497 trillion and \$157 trillion, respectively (not shown). However, other projections of long-run effects, in the area of climate change, have used much lower rates at which to discount the future. In particular, the influential Stern Review report places a much higher value on future costs and benefits by employing a discount rate of only 1.4 percent (Stern (2007)).²⁴ That report also assumes a slightly lower rate of potential growth of 1.3 percent (rather than 1.5 percent as in our other models), and this is used in our next estimate. Note that what is ultimately relevant for the projections is the difference between discount rate and rate of potential growth, yielding an effective discount rate of 0.1 percent in this scenario. If we were to adopt the discounting practice of the Stern report, the present value of the education reform would

²⁴ Note that this alternative discount rate is highly disputed in the economics literature (e.g., Nordhaus (2007); Tol and Yohe (2006)).

sum to a staggering \$636 trillion, or roughly 15 times the current GDP, by 2090 (as reported in the final two columns of Table 10).

The alternative projections under different parameter choices indicate that, while some of the parameter choices clearly have a big impact on the estimated reform value, even highly conservative parameter choices yield extremely large estimates of the aggregate value of the reform effects across the OECD.

6. POLICIES TO IMPROVE EDUCATIONAL ACHIEVEMENT

The previous discussion has stressed the importance of cognitive skills for economic growth. The evidence indicates a strong impact of skills that can give rise to immense long-term benefits. Table 11 summarizes the projection results for the three scenarios and two underlying models under the baseline parameterization. Normalized against the discounted value of the projected future GDP of the OECD over the same time span (until 2090), the value of the reform amounts to 4.3-13.8 percent of the present value of future GDP. Independent on whether the underlying economic model is specified in endogenous-growth or neoclassical terms, improved educational achievement is projected to have a large impact on future economic well-being of OECD countries.

Yet simply knowing that skill differences are important does not provide a guide to policies that might promote more skills. Indeed, a wide variety of policies have been implemented within various countries without much evidence of success in either achievement or economic terms. We believe that the disappointing results of the past generally reflect pursuing policies for which there is little empirical support.

Substantial research has gone into understanding why achievement differs across students and across countries. While controversies have existed about the interpretation of various individual pieces of evidence, considerable agreement now exists about what kinds of approaches are unreliable avenues for change. There is perhaps less agreement on the precise approaches that might be followed, but even here there is a growing consensus about the general sets of policies that have shown promise based on more credible research approaches.

The work on achievement determinants generally falls under the heading “education production functions.” The extensive work has taken a variety of perspectives and approaches. The general objective is to sort out the causal impacts of school factors (things that can potentially be manipulated through policy) from other influences on achievement including family background, peers, neighborhood influences, and the like (which are less readily amenable to policy change). In this section, we discuss available evidence on the potential for different aspects of school policy – in particular, those focused on school resource, teacher quality, and institutional structures –to improve educational achievement significantly.

6.1. Evidence on school resources

The most extensive generally available evidence relates to the effects of resources. Many policies undertaken involve substantial flows of resources – direct spending, changes in teacher salaries, reductions in class size, and the like – made within the context of current school organization. The empirical evidence clearly documents the difficulties with such policies. Simply providing more resources gives, according to the available evidence, little assurance that student performance will improve significantly. The underlying analyses of resources include studies within individual countries and across different countries and have been extensively reviewed elsewhere (see Hanushek (2003); Woessmann (2007a); Hanushek and Woessmann (2010)).

The easiest way to see the situation is a simple cross-country picture. Figure 7 plots the descriptive association between educational expenditure per student and educational outcomes for recent OECD achievement tests, the PISA 2006 study. Ignoring Mexico and Turkey where cumulative expenditure per student (over the age range of 6 to 15 years) of less than \$20,000, there is no association between educational spending and educational outcomes across OECD countries.²⁵

Table 12 probes the cross-country expenditure-outcome relationship in education in greater detail. Columns (1) to (6) reveal that in all three cycles of the OECD PISA student achievement test – 2000, 2003, and 2006 – there is a very small positive bivariate association between expenditure per student and educational achievement that vanishes once countries below a very basic cumulative spending level of \$25,000 are excluded. In fact, in the latter samples, the adjusted R^2 of a regression of test scores on educational expenditure per student is actually negative in all three PISA cycles. Of course, many more factors such as students' family backgrounds enter the determination of educational outcomes.

The simplest way of addressing bias from unobserved time-invariant country factors is to ignore level differences and restrict the analysis to changes in expenditure and outcomes over time. Thus, column (7) reports a regression in first differences between 2006 and 2000, and column (8) reports a fixed-effects regression that pools all three PISA cycles. In both cases, the association between expenditure per student and educational outcomes is far from being statistically significant at conventional levels, and the point estimates are actually negative. These descriptive patterns suggest that additional resources are not related systematically to improved test scores.²⁶

For policy deliberations, information on the impact of resources from within individual countries is perhaps more appropriate than cross-country information. Researchers have studied the determinants of student achievement for more than 40 years. The work was

²⁵ With the two outliers, there is a weak positive association as long as other effects are ignored. Taken literally, the gray regression line that includes Mexico and Turkey depicts an association where a doubling of expenditure in these two countries is associated with one tenth of a standard deviation in test scores. Note that the test score-growth nexus reported above was robust to dropping these two countries from the growth analysis.

²⁶ These results also address the concern with the growth analyses reported above that there may be a simple simultaneous determination of schooling investments and economic growth. Specifically, if a nation becomes richer through growth, it might put more resources into its schools so as to increase its cognitive skills. However, the current results show that such expenditure increases are not associated with educational outcomes in the first place. Consequently, additional resources in the school system are unlikely to generate important reverse causality from growth to test scores.

begun in the United States in the “Coleman Report” (Coleman et al. (1966)), which introduced the idea of using statistical analysis to relate various inputs of schools to student outcomes. This work also underscored the importance of including non-school factors by demonstrating that family differences were very important in explaining variations in achievement across students. While this original study has been subjected to considerable criticism, it led to an extensive line of research.

The general picture about school resources in developed countries is now well known and has been reviewed in a variety of places (see Woessmann (2005a) for Europe and Hanushek (2002, 2003) for the United States). The available studies concentrate on various common inputs to schools such as teacher experience or class size. These factors are both readily available in both administrative and survey data sets and frequently identified as the focus of policy. The available econometric evidence now includes literally hundreds of separate estimates within the U.S. and other developed countries. Quite uniformly, however, there is little strong evidence that *any* of the following factors has a consistent impact on achievement: the level of teacher education, the pupil-teacher ratio, the characteristics of administration, or the facilities of the school. Specifically, aggregating results across studies, a minority of estimates are statistically different from zero (at the 5 percent level or better), and the studies do not even uniformly indicate improvements in performance with increased resources. A second line of studies focuses on financial inputs. A number of studies simply relate spending per student to achievement or capture teacher differences by teacher salaries. While these studies tend to be lower quality, they also fail to show a consistent relationship between financial resources and achievement.

These results have been controversial. A variety of debates have taken place around the correct interpretation of prior work (see, for example, Burtless (1996)). The most important line of debate has involved study quality and whether or not these works adequately control for various inputs that might complicate the interpretation of resources. For example, the statistical models may not adequately account for other inputs that affect achievement such as the quality of family inputs. The estimates might then erroneously attribute the higher achievement due to better family factors to some of the characteristics of schools.

A simplistic view of this argument – convenient as a straw man in public debates – is that ‘money never matters.’²⁷ The research of course does not say that. Nor does it say that ‘money cannot matter.’ It simply underscores the fact that there has historically been a set of decisions and incentives in schools that have blunted any impacts of added funds, leading to inconsistent outcomes. That is, more spending on schools has not led reliably to substantially better results.

²⁷ For the historical framing of the question, see the exchange between Greenwald, Hedges, and Laine (1996) and Hanushek (1996).

6.2. Teacher quality

The most current research on school inputs and achievement has also led to another set of conclusions – that teacher quality is enormously important in determining student achievement. This work has concentrated on whether some teachers consistently produce more gains in student achievement than other teachers.²⁸ Working with extensive panel data on individual students from different U.S. states, these studies have confirmed large differences among teachers in terms of outcomes in the classroom.

But, they have also shown that the observed differences are not closely related to commonly observed characteristics of teachers (such as amount of teacher education). Some attributes of teachers – such as having one or two years of experience – have explained part of the differences in teacher quality, but these factors are a small part of the overall variance in teacher results.²⁹ This inability to identify specific teacher qualities makes it difficult to regulate or legislate having high-quality teachers in classrooms. It also contributes to our conclusions below that changes in the institutional structure and incentives of schools are fundamental to improving school outcomes.

6.3. Institutional structures and incentives in the school system

Similar to the importance of economic institutions for national economies, it is difficult to have a highly functioning education system without a supportive institutional structure. On this matter, however, there are more different opinions and perhaps a wider divergence in outcomes. Part of the reason for the divergent opinions is simply a lack of sufficient experience, analysis, and evidence.

The evidence does suggest some clear general policies that are important. Foremost among these, the performance of a system is affected by the incentives that actors face. That is, if the actors in the education process are rewarded (extrinsically or intrinsically) for producing better student achievement, and if they are penalized for not producing high achievement, achievement is likely to improve. The incentives to produce high-quality education, in turn, are created by the institutions of the education system – the rules and regulations that explicitly or implicitly set rewards and penalties for the people involved in the education process. Therefore, one might expect that institutional features have important impacts on student learning.

The unifying theme of these institutional studies is that the key to improvement appears to lie in better incentives – incentives that will lead to managerial decisions keyed to student achievement and that will promote strong schools with high-quality teachers. Here, three interrelated policies come to the forefront: promoting more competition, so that parental demand will create strong incentives to individual schools;

²⁸ See, for example, Hanushek (1971, 1992), Rockoff (2004), Rivkin, Hanushek, and Kain (2005), and a number of subsequent studies reviewed in Hanushek and Rivkin (2010).

²⁹ There is some indication that teachers' own academic skills measured by scores on achievement tests may be an important factor (see Wayne and Youngs (2003), Eide, Goldhaber, and Brewer (2004), and Hanushek and Rivkin (2006) for reviews), but methodologically more sophisticated work is needed before conclusive assessments can be given on this dimension.

autonomy in local decision making, so that individual schools and their leaders will take actions to promote student achievement; and, an accountability system that identifies good school performance and leads to rewards based on this.

Choice and Competition. Choice and competition in schools were proposed a half century ago by Milton Friedman (1962). The simple idea is that parents, interested in the schooling outcomes of their children, will seek out productive schools. This demand-side pressure will result in incentives for each school to produce an effective education system. These incentives will also put pressure on schools to ensure high-quality staff in addition to a good curriculum.

In many school systems (with the Netherlands being the most obvious example), a number of privately managed schools provide alternatives for students. These schools, which also often have a religious affiliation, are part of the natural institutional framework. Unfortunately, little thorough evaluation has been done of the choice possibilities, in large part because there is no obvious comparison group (i.e., choice is instituted for an entire country and there is no example of the no-choice alternative). In a cross-country comparison, students in countries with a larger share of privately managed schools tend to perform better on average (cf. Woessmann (2007b, 2009); Woessmann, Luedemann, Schuetz, and West (2009)), and recent evidence corroborates the conclusion that this is due to a causal effect of private-sector competition (West and Woessmann (2010)).³⁰

In the U.S., there are limited examples of private school choice, ranging from the publicly funded school vouchers in Milwaukee, Cleveland, and Washington, DC, to privately financed voucher alternatives.³¹ The evaluations of these generally show that the choice schools do at least as well as the regular public schools, if not better (see Rouse (1998); Howell and Peterson (2002)).³²

In Europe, Bradley and Taylor (2002) and Levačić (2004) find similar positive effects of school competition on the performance of English schools. Sandström and Bergström (2005) and Björklund, Edin, Freriksson, and Krueger (2004) provide evidence on significant positive effects of competition from privately operated schools on the performance of public schools in Sweden. Filer and Munich (2003) show that the introduction of a voucher-type system in the Czech Republic led to the creation of private schools in areas where public schools are doing badly and that the public schools facing private competition improved their performance.

The major issue on choice and competition is still the limited experience. Teachers unions and administrator groups invariably dislike the idea of competition – because it

³⁰ Note that private school management does not mean private school funding; the international evidence suggests that both private school management and *public* school funding are associated with better achievement across countries (Woessmann (2009)).

³¹ The largest U.S. voucher program in the State of Florida provides vouchers for special needs students (Greene (2007)). While there is considerable satisfaction with this program, there is no evaluation available that is based on explicit outcome measures.

³² A special type of choice in the U.S. is charter schools. These schools are discussed next under autonomy. Also, given U.S. residential mobility, individual public school districts compete with each other, and more competition appears to produce better results (Hoxby (2000)), although this finding has been controversial (Rothstein (2007); Hoxby (2007)).

puts pressure on them. Thus, not many examples of operational, large-scale attempts at competition have been evaluated.

Autonomy and Decentralization. Several institutional features of a school system can be grouped under the heading of autonomy or decentralization, including local decision making on different matters, fiscal decentralization, and parental involvement. Almost any system of improved incentives for schools depends upon having school personnel in individual schools and districts heavily involved in decision making. It is difficult to compile evidence on the impact of autonomy, because the degree of local decision making is most generally a decision for a country (or state) as a whole, leaving no comparison group within countries. Across countries, students tend to perform better in schools that have autonomy in personnel and day-to-day decisions (Woessmann (2003a, 2007b); Woessmann, Luedemann, Schuetz, and West (2009)), in particular when there is accountability (see also the review in Hanushek and Woessmann (2010)).

The U.S. states have varying amounts of local autonomy. One systematic form of school autonomy is “charter schools,” which are public schools that are allowed to perform quite autonomously. (Note that these are actually hybrids of choice schools and public-school autonomy, because they survive only if sufficient numbers of students are attracted to them and continue to attend them). These schools are relatively new, a fact that complicates evaluation since many are still in the start-up phase. The evidence on them is mixed but indicates a variety of places where charter schools outperform the regular public schools after the initial start-up phase but also suggests in part that the regulations governing them and the particular competitive public schools they face have an influence.³³

Given the available evidence, support for autonomy also strongly rests on a conceptual basis. A system with strong incentives seems likely to capitalize on local decision making.

School Accountability. Many countries around the world have been moving toward increased accountability of local schools for student performance. The United Kingdom has developed an elaborate system of “league tables” designed to give parents full information about the performance of local schools. The United States has legislated a federal law (“No Child Left Behind”) that all states develop an accountability system that meets certain general guidelines. It also sets into law a series of actions required when a school fails to bring sufficient numbers of students up to proficiency in core subjects.

Evidence on the impacts of these systems has begun to accumulate. While there is some uncertainty given the newness of the overall federal accountability system (introduced in 2002), the best U.S. evidence indicates that strong state accountability

³³ A number of studies have based the analysis on student fixed effects, relying on students moving in and out of charter schools to identify the impact of charters (Booker, Gilpatric, Gronberg, and Jansen (2007); Bifulco and Ladd (2006); Hanushek, Kain, Rivkin, and Branch (2007)). These studies have generally pointed to a range of quality for charter schools but highlight start-up problems. Another set of studies considers charter schools that have more demand than open positions, requiring schools to choose their students by lottery (Hoxby and Murarka (2009); Abdulkadiroglu et al. (2009)). Finally, matching methods have been used to compare public and private school performance, leading to the conclusion that there is wide variation in the quality of charter schools compared to the relevant public schools (CREDO (2009, 2010)).

systems in fact lead to better student performance (Carnoy and Loeb (2002); Hanushek and Raymond (2005); Jacob (2005); Dee and Jacob (2009)).

One institutional set-up that combines accountability with parental choice are systems that give students in schools that repeatedly do badly on the accountability test a voucher to attend private schools. In Florida, the threat of becoming subject to private-school choice if failing on the test has been shown to increase school performance particularly for disadvantaged students (West and Peterson (2006); Figlio and Rouse (2006)).

Curriculum-based external exit exams are another means to introduce some form of accountability into the schooling system. They provide performance information which can hold both students and schools accountable. Students in countries with external exit exam systems tend to systematically outperform students in countries without such systems (Bishop (1997, 2006); Woessmann (2003a, 2007b); Woessmann, Luedemann, Schuetz, and West (2009)). In Canada and Germany, the two national education systems where the existence of external exams varies within the country because some regions feature them and others not, it has similarly been shown that students perform better in regions with external exams (Bishop (1997); Jürges, Schneider, and Büchel (2005); Woessmann (2010b)).

It is difficult to imagine choice or autonomy working well without a good system of student testing and accountability. Thus, the ideas about institutional structure are closely linked together. The international evidence clearly suggests that school autonomy, in particular local autonomy over teacher salaries and course content, is only effective in school systems that have external exams in place (Woessmann (2005b, 2007b); Fuchs and Woessmann (2007); Woessmann, Luedemann, Schuetz, and West (2009)). For example, school autonomy over teacher salaries is negatively associated with student achievement in systems without external exams, but positively associated with student achievement in external-exam systems. This pattern of results has been found in several different TIMSS and PISA studies and in analyses of autonomy in other decision making areas such as school autonomy in determining course content and teacher influence on resource funding. Similar evidence that accountability policies are more effective when there is greater local control has also been found across U.S. states (Loeb and Strunk (2007)).

Finally, given the importance of high teacher quality, a promising candidate for improvement is the specific form of accountability that aims incentives directly at teachers. While convincing evidence on the effects of performance-related teacher pay is scarce, the more rigorous studies in terms of empirical identification tend to find a positive relationship between financial teacher incentives and student outcomes (cf. the surveys in Atkinson et al. (2009) and Podgursky and Springer (2007); see also Figlio and Kenny (2007)). Thus, Atkinson et al. (2009) find that the introduction of performance-related pay had a substantial positive impact on student achievement in England. Similarly, monetary incentives for teachers based on their students' performance have been shown to improve student learning very significantly in Israel and in India (Lavy (2002, 2009); Muralidharan and Sundararaman (2009)). Likewise, the cross-country

variation provides some indication that students perform better in countries that allow for teacher salaries to be adjusted based on performance in teaching (Woessmann (2010a)).

Clearly, research on how school policy can successfully advance educational achievement is an expanding field that still leaves many open questions. At the same time, our reading of the available evidence is that institutional reforms – in particular in the areas of competition, autonomy, and accountability – that create incentives for improving outcomes and focus in particular on teacher quality have substantial potential to create the kinds of learning gains that our results above show to be linked to immense long-term economic benefits.

7. CONCLUSIONS

It is generally the case that national attention to economic policies that deal with current aggregate demand conditions and with business cycles invariably take priority over longer-run policy considerations. Perhaps this has never been as true as today, when the most obvious focus of attention is the worldwide recession. Without minimizing the need to deal with current unemployment conditions, the message of this paper 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 US performance 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))

Our results show that education policy is closely associated with the long-run growth potentials of OECD countries. The regression analyses suggest that direct measures of educational outcomes, in terms of cognitive skills on international achievement tests, emerge as the one strong policy factor underlying growth differences across OECD countries. By contrast, a long battery of institutional and regulatory measures does not add to an explanation of the substantial differences in long-run growth rates that exist across OECD countries, mainly because all OECD countries share a common set of basic institutional structures that ensure a general functioning of market economies. Considering different skill dimensions, basic skills are robustly related to OECD-country growth, whereas the relation of the top-skill dimension with growth is at least substantially smaller than in non-OECD countries. When cognitive skills are accounted for, tertiary attainment is not significantly associated with long-run growth differences across OECD countries.

Our projection analysis suggests that, under plausible parameter assumptions, the real present value of future improvements in GDP due to challenging but achievable educational reform scenarios amounts to \$90-275 trillion. A modest goal of having all OECD countries boost their average PISA scores by 25 points (one-quarter standard deviation) implies an aggregate gain of OECD GDP of \$90-123 trillion dollars. More aggressive goals, such as bringing all students to a level of minimal proficiency for the OECD or bringing all OECD countries to the level reached by Finland today, would

imply aggregate GDP increases beyond \$200 trillion according to historical growth relationships. The precise size of the reform value of such long-run projections is clearly up for debate. Nevertheless, our sensitivity analyses indicate that, while differences between an endogenous and neoclassical model framework and alternative parameter choices clearly make a difference, the estimates of the long-run effect of reasonable education reforms still yield enormous values no matter what. The gains from education reform far exceed the level of stimulus funds in the current global recession.

Our projections do not by themselves indicate how schools should be changed. Nor do they solve the political economy issues of how any change should be achieved politically. They simply underscore the high cost of political inaction or misdirection.

In order to provide some guidance, we review the extensive relevant research on the determinants of educational achievement. Several conclusions appear. First, many of the traditional policies of simply providing more funds for schools or of adding specific resources such as smaller classes do not provide much hope for significant improvements in student achievement. Second, a growing body of research shows that teacher quality is a primary driver of student achievement but that differences in quality are not closely related to teacher education and experience. Because teacher quality is not easily measured and regulated, effective policies to improve quality appear to necessitate more careful attention to the incentives faced by schools and teachers. Here the research on educational institutions suggests productive policy approaches. In particular, evidence from both within and across countries points to the positive impact of competition among schools, of accountability and student testing, and of local school autonomy in decision making.³⁴ Research on these policies, separately and in combination, indicates some continuing uncertainty about the magnitude of any effects but does support more aggressive attention to these in setting school policies.

An important aspect highlighted by the projections is the dynamic nature of human capital and growth. Our basic characterization of growth indicates that higher cognitive skills offer a path of continued economic improvement, so that favorable policies today have growing impacts in the future. However, the full ramifications of schooling outcomes will not become apparent until reasonably far into the future. The economic gains from education reform are surely not reaped within matters of one or two political legislation periods. They rather require a long-run perspective that fully considers the time horizon of a child born today. In the discussion of climate policies, it has become custom to consider expected outcomes that materialize several generations from now. Education policy needs a similar long-term perspective to fully capture the consequences of possible current reforms.

³⁴ 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).

REFERENCES

- Abdulkadiroglu, Atila, Joshua Angrist, Susan Dynarski, Thomas J. Kane, and Parag Pathak. 2009. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." NBER WP15549. Cambridge, MA: National Bureau of Economic Research (November).
- Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2001. "The colonial origins of comparative development: An empirical investigation." *American Economic Review* 91, no. 5 (December): 1369-1401.
- . 2005. "Institutions as a fundamental cause of long-run growth." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf. Amsterdam: North Holland: 385-472.
- Aghion, Philippe, Mathias Dewatripont, Caroline Hoxby, Andreu Mas-Colell, and André Sapir. 2010. "The governance and performance of universities: evidence from Europe and the US." *Economic Policy* 25, no. 61: 7-59.
- Aghion, Philippe, and Peter Howitt. 1998. *Endogenous Growth Theory*. Cambridge, MA: MIT Press.
- . 2006. "Appropriate growth policy: A unifying framework." *Journal of the European Economic Association* 4, no. 2-3: 269-314.
- Atkinson, Adele, Simon Burgess, Bronwyn Croxson, Paul Gregg, Carol Propper, Helen Slater, and Deborah Wilson. 2009. "Evaluating the impact of performance-related pay for teachers in England." *Labour Economics* 16, no. 3: 251-261.
- Azariadis, Costas, and Allan Drazen. 1990. "Threshold externalities in economic development." *Quarterly Journal of Economics* 105, no. 2: 501-526.
- Barro, Robert J. 1991. "Economic growth in a cross section of countries." *Quarterly Journal of Economics* 106, no. 2 (May): 407-443.
- . 2001. "Human capital and growth." *American Economic Review* 91, no. 2: 12-17.
- Barro, Robert J., and Jong-Wha Lee. 1993. "International comparisons of educational attainment." *Journal of Monetary Economics* 32, no. 3 (December): 363-394.
- . 2010. "A new data set of educational attainment in the world, 1950-2010." NBER Working Paper 15902. Cambridge, MA: National Bureau of Economic Research (April).
- Barro, Robert J., and Xavier Sala-i-Martin. 2004. *Economic growth*. Second ed. Cambridge, MA: The MIT Press.
- 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* 34, no. 2: 143-174.
- . 2005. "Human capital and technology diffusion." In *Handbook of Economic Growth*, edited by Philippe Aghion and Steven N. Durlauf. Amsterdam: North Holland: 935-966.
- Bifulco, Robert, and Helen F. Ladd. 2006. "The impacts of charter schools on student achievement: Evidence from North Carolina." *Education Finance and Policy* 1, no. 1 (Winter): 50-90.
- Bils, Mark, and Peter J. Klenow. 2000. "Does schooling cause growth?" *American Economic Review* 90, no. 5 (December): 1160-1183.
- Bishop, John H. 1997. "The effect of national standards and curriculum-based examinations on achievement." *American Economic Review* 87, no. 2: 260-264.
- . 2006. "Drinking from the fountain of knowledge: Student incentive to study and learn - Externalities, information problems, and peer pressure." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 909-944.
- Björklund, Anders, Per-Anders Edin, Peter Freriksson, and Alan B. Krueger. 2004. "Education, equality and efficiency: An analysis of Swedish school reforms during the 1990s." IFAU Report 2004:1. Uppsala: Institute for Labour Market Policy Evaluation.
- Booker, Kevin, Scott M. Gilpatric, Timothy Gronberg, and Dennis Jansen. 2007. "The Impact of Charter School Attendance on Student Performance." *Journal of Public Economics* 91, no. 5-6: 849-876.
- Börsch-Supan, Axel. 2000. "A model under siege: A case study of the German retirement insurance system." *Economic Journal* 110, no. 461: F24-F45.
- Bosworth, Barry P., and Susan M. Collins. 2003. "The empirics of growth: An update." *Brookings Papers on Economic Activity* 2003, no. 2: 113-206.
- Botero, Juan C., Simeon Djankov, Rafael La Porta, Florencio Lopez-De-Silanes, and Andrei Shleifer. 2004. "The regulation of labor." *Quarterly Journal of Economics* 119, no. 4: 1339-1382.
- Bowles, Samuel, Herbert Gintis, and Melissa Osborne. 2001. "The determinants of earnings: A behavioral approach." *Journal of Economic Literature* 39, no. 4 (December): 1137-1176.

- Bradley, Steve, and Jim Taylor. 2002. "The effect of the quasi-market on the efficiency-equity trade-off in the secondary school sector." *Bulletin of Economic Research* 54, no. 3: 295-314.
- Burtless, Gary, ed. 1996. *Does money matter? The effect of school resources on student achievement and adult success*. Washington, DC: Brookings.
- Carnoy, Martin, and Susanna Loeb. 2002. "Does external accountability affect student outcomes? A cross-state analysis." *Educational Evaluation and Policy Analysis* 24, no. 4 (Winter): 305-331.
- Ciccone, Antonio, and Elias Papaioannou. 2009. "Human capital, the structure of production, and growth." *Review of Economics and Statistics* 91, no. 1: 66-82.
- Cingano, Federico, Marco Leonardi, Julián Messina, and Giovanni Pica. 2010. "The effects of employment protection legislation and financial market imperfections on investment: evidence from a firm-level panel of EU countries." *Economic Policy* 25, no. 61: 117-163.
- Cohen, Daniel, and Marcelo Soto. 2007. "Growth and human capital: good data, good results." *Journal of Economic Growth* 12, no. 1 (March): 51-76.
- Coleman, James S., Ernest Q. Campbell, Carol J. Hobson, James McPartland, Alexander M. Mood, Frederic D. Weinfeld, and Robert L. York. 1966. *Equality of educational opportunity*. Washington, D.C.: U.S. Government Printing Office.
- Coulombe, Serge, Jean-François Tremblay, and Sylvie Marchand. 2004. *Literacy scores, human capital and growth across fourteen OECD countries*. Ottawa: Statistics Canada.
- CREDO. 2009. *Multiple choice: Charter school performance in 16 states*. Stanford, CA: Center for Research on Education Outcomes, Stanford University.
- . 2010. *Charter school performance in New York City*. Stanford, CA: Center for Research on Education Outcomes, Stanford University.
- Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the evidence on life cycle skill formation." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: Elsevier: 697-812.
- Dee, Thomas, and Brian Jacob. 2009. "The impact of No Child Left Behind on student achievement." NBER Working Paper 15531. Cambridge, MA: National Bureau of Economic Research (November).
- Denison, Edward F. 1985. *Trends in American economic growth, 1929-1982*. Washington, D.C.: The Brookings Institution.
- Eide, Eric, Dan Goldhaber, and Dominic Brewer. 2004. "The teacher labour market and teacher quality." *Oxford Review of Economic Policy* 20, no. 2: 230-244.
- Figlio, David N., and Lawrence W. Kenny. 2007. "Individual teacher incentives and student performance." *Journal of Public Economics* 91, no. 5-6 (June): 901-914.
- Figlio, David N., and Cecilia Elena Rouse. 2006. "Do accountability and voucher threats improve low-performing schools?" *Journal of Public Economics* 90, no. 1-2 (January): 239-255.
- Filer, Randall K., and Daniel Münich. 2003. "Responses of private and public schools to voucher funding." Paper presented at the Annual Meeting of the American Economic Association. Washington, D.C.
- Friedman, Milton. 1962. *Capitalism and freedom*. Chicago: University of Chicago Press.
- Fuchs, Thomas, and Ludger Woessmann. 2007. "What accounts for international differences in student performance? A re-examination using PISA data." *Empirical Economics* 32, no. 2-3: 433-462.
- Glaeser, Edward L., Rafael La Porta, Forencio Lopez-de-Silanes, and Andrei Shleifer. 2004. "Do institutions cause growth?" *Journal of Economic Growth* 9, no. 3: 271-303.
- Greene, Jay P. 2007. "Fixing special education." *Peabody Journal of Education* 82, no. 4: 703-723.
- Greenwald, Rob, Larry V. Hedges, and Richard D. Laine. 1996. "The effect of school resources on student achievement." *Review of Educational Research* 66, no. 3 (Fall): 361-396.
- Hagist, Christian, Norbert Klusen, Andreas Plate, and Bernd Raffelhüschen. 2005. "Social health insurance: The major driver of unsustainable fiscal policy?" CESifo Working Papers 1574. Munich: CESifo.
- Hanushek, Eric A. 1971. "Teacher characteristics and gains in student achievement: Estimation using micro data." *American Economic Review* 60, no. 2 (May): 280-288.
- . 1992. "The trade-off between child quantity and quality." *Journal of Political Economy* 100, no. 1 (February): 84-117.
- . 1996. "A more complete picture of school resource policies." *Review of Educational Research* 66, no. 3 (Fall): 397-409.
- . 2002. "Publicly provided education." In *Handbook of Public Economics*, edited by Alan J. Auerbach and Martin Feldstein. Amsterdam: Elsevier: 2045-2141.

- . 2003. "The failure of input-based schooling policies." *Economic Journal* 113, no. 485 (February): F64-F98.
- Hanushek, Eric A., John F. Kain, Steven G. Rivkin, and Gregory F. Branch. 2007. "Charter school quality and parental decision making with school choice." *Journal of Public Economics* 91, no. 5-6: 823-848.
- Hanushek, Eric A., and Dennis D. Kimko. 2000. "Schooling, labor force quality, and the growth of nations." *American Economic Review* 90, no. 5 (December): 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, NJ: Princeton University Press.
- Hanushek, Eric A., and Margaret E. Raymond. 2005. "Does school accountability lead to improved student performance?" *Journal of Policy Analysis and Management* 24, no. 2 (Spring): 297-327.
- Hanushek, Eric A., and Steven G. Rivkin. 2006. "Teacher quality." In *Handbook of the Economics of Education*, edited by Eric A. Hanushek and Finis Welch. Amsterdam: North Holland: 1051-1078.
- . 2010. "Generalizations about using value-added measures of teacher quality." *American Economic Review* 100, no. 2 (May): 267-271.
- Hanushek, Eric A., and Ludger Woessmann. 2008. "The role of cognitive skills in economic development." *Journal of Economic Literature* 46, no. 3 (September): 607-668.
- . 2009. "Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation." NBER Working Paper 14633. Cambridge, MA: National Bureau of Economic Research (January).
- . 2010. "The economics of international differences in educational achievement." In *Handbook of the Economics of Education, Vol. 3*, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann. Amsterdam: North Holland.
- Heckman, James J. 2006. "Skill formation and the economics of investing in disadvantaged children." *Science* 312, no. 5782 (June 30): 1900-1902.
- Heckman, James J., Jora Stixrud, and Sergio Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor Economics* 24, no. 3 (July): 411-482.
- Heston, Alan, Robert Summers, and Bettina Aten. 2002. "Penn World Table Version 6.1." Center for International Comparisons at the University of Pennsylvania (CICUP) Philadelphia: University of Pennsylvania.
- Howell, William G., and Paul E. Peterson. 2002. *The education gap: Vouchers and urban schools*. Washington, DC: Brookings.
- Hoxby, Caroline M. 2000. "Does competition among public schools benefit students and taxpayers?" *American Economic Review* 90, no. 5 (December): 1209-1238.
- . 2007. "Does competition among public schools benefit students and taxpayers? Reply." *American Economic Review* 97, no. 5: 2038-2055.
- Hoxby, Caroline M., and Sonali Murarka. 2009. "Charter schools in New York City: Who enrolls and how they affect their students' achievement." NBER Working Paper 14852. Cambridge, MA: National Bureau of Economic Research (April).
- Jacob, Brian A. 2005. "Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago Public Schools." *Journal of Public Economics* 89, no. 5-6 (June): 761-796.
- Jorgenson, Dale W., and Zvi Griliches. 1967. "The explanation of productivity change." *Review of Economic Studies* 34, no. 3 (July): 249-282.
- Jürges, Hendrik, Kerstin Schneider, and Felix Büchel. 2005. "The effect of central exit examinations on student achievement: Quasi-experimental evidence from TIMSS Germany." *Journal of the European Economic Association* 3, no. 5: 1134-1155.
- Kaufmann, Daniel, Aart Kraay, and Pablo Zoido-Lobaton. 1999. "Governance matters." World Bank Policy Research Working Paper 2196. Washington, DC: World Bank.
- Krueger, Alan B., and Mikael Lindahl. 2001. "Education for growth: Why and for whom?" *Journal of Economic Literature* 39, no. 4 (December): 1101-1136.
- Lavy, Victor. 2002. "Evaluating the effect of teachers' group performance incentives on pupil achievement." *Journal of Political Economy* 110, no. 6 (December): 1286-1317.
- . 2009. "Performance pay and teachers' effort, productivity, and grading ethics." *American Economic Review* 99, no. 5 (December): 1979-2011.
- Levačić, Rosalind. 2004. "Competition and the performance of English secondary schools: Further evidence." *Education Economics* 12, no. 2: 177-193.

- Levine, Ross, and David Renelt. 1992. "A sensitivity analysis of cross-country growth regressions." *American Economic Review* 82, no. 4 (September): 942-963.
- Levine, Ross, and Sara J. Zervos. 1993. "What we have learned about policy and growth from cross-country regressions." *American Economic Review* 83, no. 2 (May): 426-430.
- Loeb, Susanna, and Katharine Strunk. 2007. "Accountability and local control: Response to incentives with and without authority over resource generation and allocation." *Education Finance and Policy* 2, no. 1: 10-39.
- Lucas, Robert E. 2009. "Trade and the diffusion of the industrial revolution." *American Economic Journal: Macroeconomics* 1, no. 1: 1-25.
- Lucas, Robert E., Jr. 1988. "On the mechanics of economic development." *Journal of Monetary Economics* 22(July): 3-42.
- . 2003. "Macroeconomic priorities." *American Economic Review* 93, no. 1: 1-14.
- Mankiw, N. Gregory, David Romer, and David Weil. 1992. "A contribution to the empirics of economic growth." *Quarterly Journal of Economics* 107, no. 2 (May): 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* 23, no. 4 (Autumn): 689-812.
- Muralidharan, Karthik, and Venkatesh Sundararaman. 2009. "Teacher performance pay: Experimental evidence from India." NBER Working Paper 15323. Cambridge, MA: National Bureau of Economic Research.
- Nelson, Richard R., and Edmund Phelps. 1966. "Investment in humans, technology diffusion and economic growth." *American Economic Review* 56, no. 2 (May): 69-75.
- Nicoletti, Giuseppe, and Frederic L. Pryor. 2006. "Subjective and objective measures of governmental regulations in OECD nations." *Journal of Economic Behavior & Organization* 59, no. 3: 433-449.
- Nicoletti, Giuseppe, and Stefano Scarpetta. 2003. "Regulation, productivity, and growth: OECD evidence." *Economic Policy* 18, no. 1: 10-72.
- Nicoletti, Giuseppe, Stefano Scarpetta, and Olivier Boylaud. 2000. "Summary indicators of product market regulation with an extension to employment protection legislation." Economic Department Working Paper 226. Paris: OECD (February).
- Nordhaus, William D. 2007. "A review of the Stern review on the economics of climate change." *Journal of Economic Literature* 45, no. 3: 686-702.
- Organisation for Economic Co-operation and Development. 1999. *Employment outlook*. Paris: OECD.
- . 2001. *Knowledge and skills for life: First results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris: OECD.
- . 2003. *Literacy skills for the world of tomorrow: Further results from PISA 2000*. Paris: OECD.
- . 2004. *Learning for tomorrow's world: First results from PISA 2003*. Paris: OECD.
- . 2007. *PISA 2006: Science competencies for tomorrow's world*. Vol. 1 – Analysis. Paris: OECD.
- . 2008. *Compendium of patent statistics*. Paris: OECD.
- . 2009a. *Education at a glance 2009: OECD indicators*. Paris: Organisation for Economic Co-operation and Development.
- . 2009b. *OECD economic outlook, volume 2009/1, No. 85*. Paris: OECD.
- . 2009c. *Society at a glance 2009: OECD social indicators*. Paris: OECD.
- Podgursky, Michael J., and Matthew G. Springer. 2007. "Teacher performance pay: A review." *Journal of Policy Analysis and Management* 26, no. 4: 909-949.
- Pryor, Frederic L. 2002. "Quantitative notes on the extent of governmental regulations in various OECD nations." *International Journal of Industrial Organization* 20, no. 5: 693-714.
- Rivkin, Steven G., Eric A. Hanushek, and John F. Kain. 2005. "Teachers, schools, and academic achievement." *Econometrica* 73, no. 2 (March): 417-458.
- Rockoff, Jonah E. 2004. "The impact of individual teachers on student achievement: Evidence from panel data." *American Economic Review* 94, no. 2 (May): 247-252.
- Romer, Paul. 1990a. "Endogenous technological change." *Journal of Political Economy* 99, no. 5, pt. II: S71-S102.
- . 1990b. "Human capital and growth: Theory and evidence." *Carnegie-Rochester Conference Series on Public Policy* 32: 251-286.
- Rothstein, Jesse. 2007. "Does competition among public schools benefit students and taxpayers? Comment." *American Economic Review* 97, no. 5: 2026-2037.

- Rouse, Cecilia Elena. 1998. "Private school vouchers and student achievement: An evaluation of the Milwaukee Parental Choice Program." *Quarterly Journal of Economics* 113, no. 2 (May): 553-602.
- Sachs, Jeffrey D., and Andrew M. Warner. 1995. "Economic reform and the process of global integration." *Brookings Papers on Economic Activity*, no. 1: 1-96.
- . 1997. "Fundamental sources of long-run growth." *American Economic Review* 87, no. 2: 184-188.
- 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 (September): 813-835.
- Sandström, F. Mikael, and Fredrik Bergström. 2005. "School vouchers in practice: Competition will not hurt you." *Journal of Public Economics* 89, no. 2-3: 351-380.
- Sianesi, Barbara, and John Van Reenen. 2003. "The returns to education: Macroeconomics." *Journal of Economic Surveys* 17, no. 2: 157-200.
- Solow, Robert M. 1957. "Technical change and the aggregate production function." *Review of Economics and Statistics* 39, no. 3 (August): 312-320.
- Stern, Nicholas. 2007. *The economics of climate change: The Stern review*. Cambridge, UK, and New York: Cambridge University Press.
- Temple, Jonathan. 2001. "Growth effects of education and social capital in the OECD countries." *OECD Economic Studies* 33: 57-101.
- Tol, Richard S. J., and Gary W. Yohe. 2006. "A review of the Stern review." *World Economics* 7, no. 4: 233-250.
- Topel, Robert. 1999. "Labor markets and economic growth." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card. Amsterdam: Elsevier: 2943-2984.
- Vandenbussche, Jérôme, Philippe Aghion, and Costas Meghir. 2006. "Growth, distance to frontier and composition of human capital." *Journal of Economic Growth* 11, no. 2 (June): 97-127.
- Venn, Danielle. 2009. "Legislation, collective bargaining and enforcement: Updating the OECD employment protection indicators." OECD Social, Employment and Migration Working Paper 89. Paris: OECD.
- Wayne, Andrew J., and Peter Youngs. 2003. "Teacher characteristics and student achievement gains: A review." *Review of Educational Research* 73, no. 1: 89-122.
- Welch, Finis. 1970. "Education in production." *Journal of Political Economy* 78, no. 1 (January/February): 35-59.
- West, Martin R., and Paul E. Peterson. 2006. "The efficacy of choice threats within school accountability systems: Results from legislatively-induced experiments." *Economic Journal* 116, no. 510: C46-C62.
- West, Martin R., and Ludger Woessmann. 2010. "'Every Catholic child in a Catholic school': Historical resistance to state schooling, contemporary private competition and student achievement across countries." *Economic Journal* 120, no. 546: F229-F255.
- Woessmann, Ludger. 2003a. "Schooling resources, educational institutions, and student performance: The international evidence." *Oxford Bulletin of Economics and Statistics* 65, no. 2: 117-170.
- . 2003b. "Specifying human capital." *Journal of Economic Surveys* 17, no. 3: 239-270.
- . 2005a. "Educational production in Europe." *Economic Policy* 20, no. 43: 446-504.
- . 2005b. "The effect heterogeneity of central exams: Evidence from TIMSS, TIMSS-Repeat and PISA." *Education Economics* 13, no. 2: 143-169.
- . 2007a. "International evidence on expenditure and class size: A review." In *Brookings Papers on Education Policy 2006/2007*. Washington D.C.: Brookings: 245-272.
- . 2007b. "International evidence on school competition, autonomy and accountability: A review." *Peabody Journal of Education*. 82, no. 2-3: 473-497.
- . 2009. "Public-private partnerships and student achievement: A cross-country analysis." In *School choice international: Exploring public-private partnerships*, edited by Rajashri Chakrabarti and Paul E. Peterson. Cambridge, MA: MIT Press: 13-45.
- . 2010a. "Cross-country evidence on teacher performance pay." Program on Education Policy and Governance Working Paper PEPG 10-11. Cambridge, MA: Harvard University (June).
- . 2010b. "Institutional determinants of school efficiency and equity: German states as a microcosm for OECD countries." *Jahrbücher für Nationalökonomie und Statistik / Journal of Economics and Statistics* 230, no. 2: 234-270.
- Woessmann, Ludger, Elke Luedemann, Gabriela Schuetz, and Martin R. West. 2009. *School accountability, autonomy, and choice around the world*. Cheltenham, UK: Edward Elgar.

- Woessmann, Ludger, and Marc Piopiunik. 2009. *Was unzureichende Bildung kostet: Eine Berechnung der Folgekosten durch entgangenes Wirtschaftswachstum*. Gütersloh, Germany: Bertelsmann Stiftung.
- Wölfl, Anita, Isabelle Wanner, Tomasz Kozluk, and Giuseppe Nicoletti. 2009. "Ten years of product market reform in OECD countries: Insights from a revised PMR indicator." OECD Economics Department Working Paper 695. Paris: OECD.

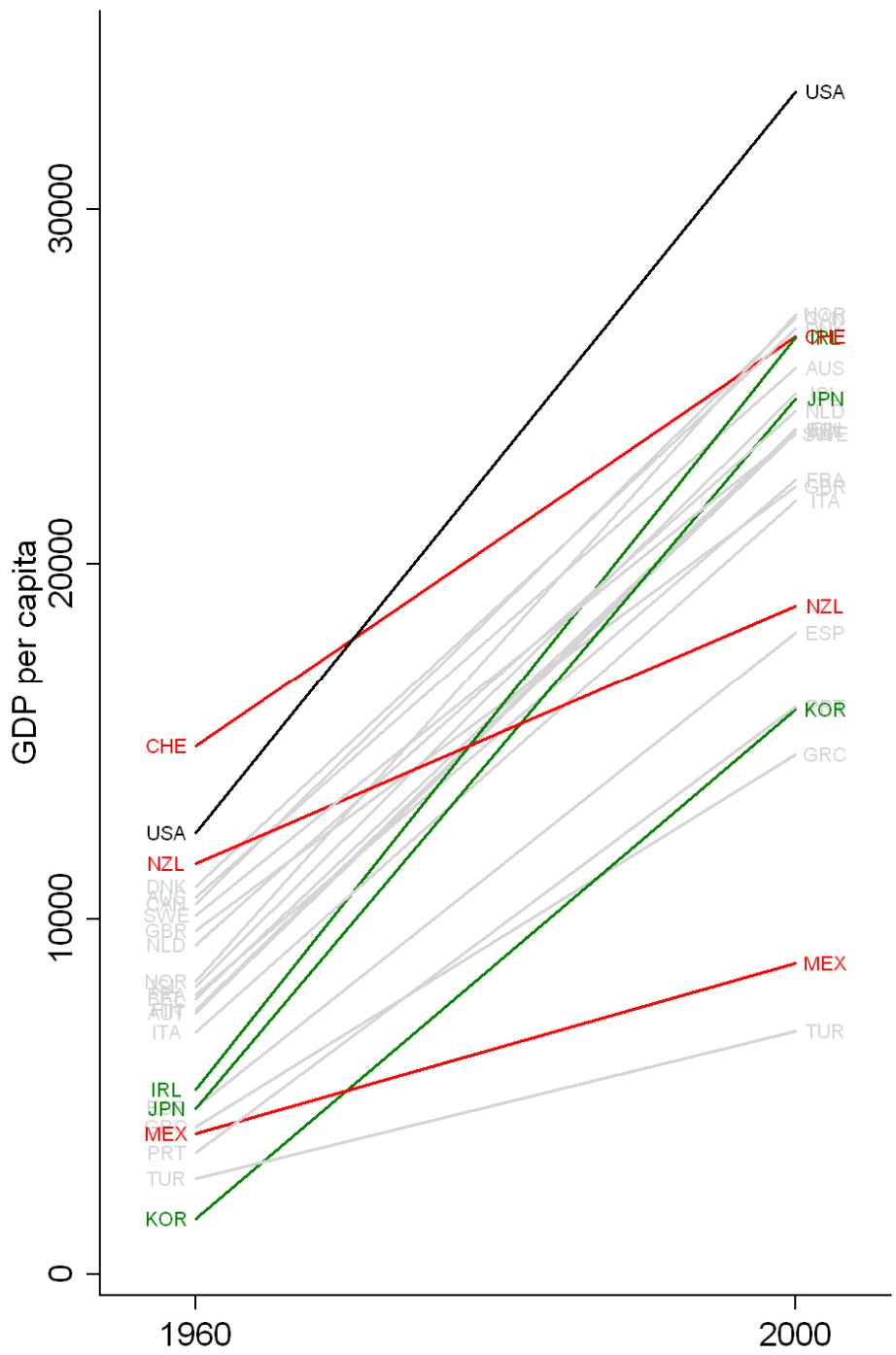


Figure 1. GDP per capita in OECD countries in 1960 and 2000

Notes: GDP per capita in constant international dollars, 24 OECD countries with income and education data for the whole period. Colors for expositional purposes only.

Source: Authors' depiction based on data from Heston, Summers, and Aten (2002).

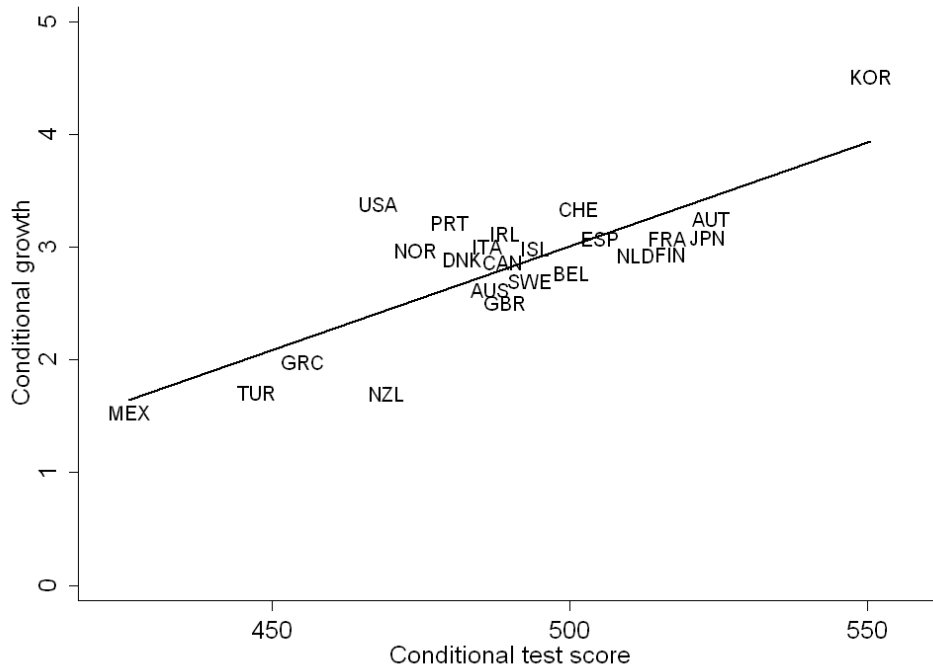


Figure 2. Educational performance and economic growth across OECD countries

Notes: Added-variable plot of a regression of the average annual rate of growth (in percent) 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).

Source: Authors' depiction based on the regression analysis reported in column (2) of Table 2.

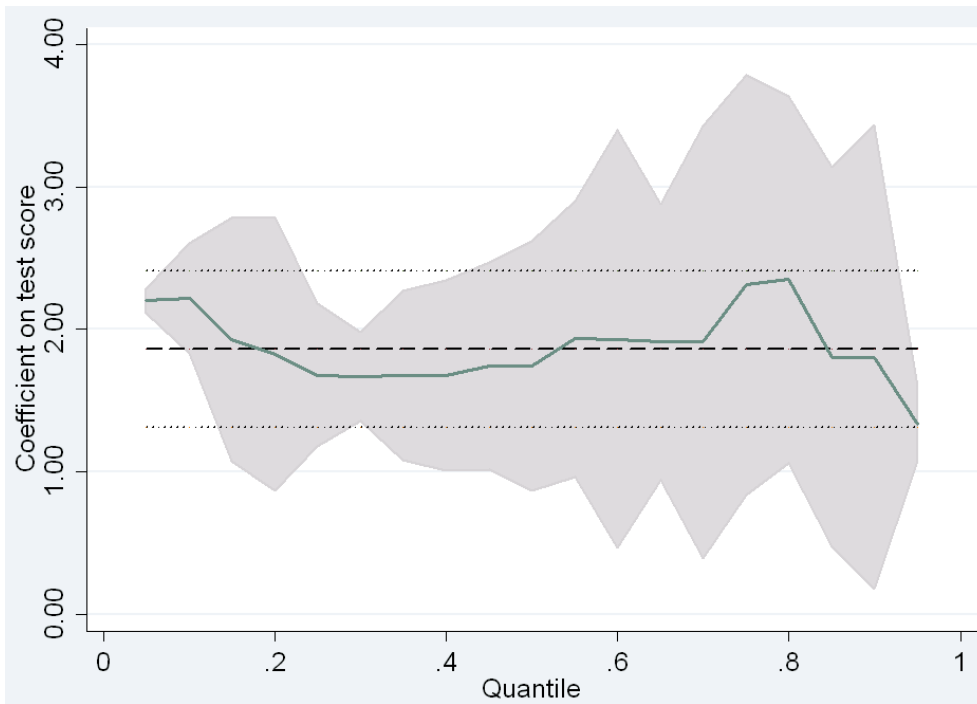


Figure 3. Quantile regression estimates

Notes: The straight line represents quantile regression estimates in 5-percent steps of the coefficient on test scores in a model corresponding to column (2) of Table 2. The shaded area shows the 90-percent confidence interval of the quantile regression estimates. The horizontal dashed line represents the OLS estimate and the dotted line its 90-percent confidence interval.

Source: Authors' depiction based on own regression analysis.

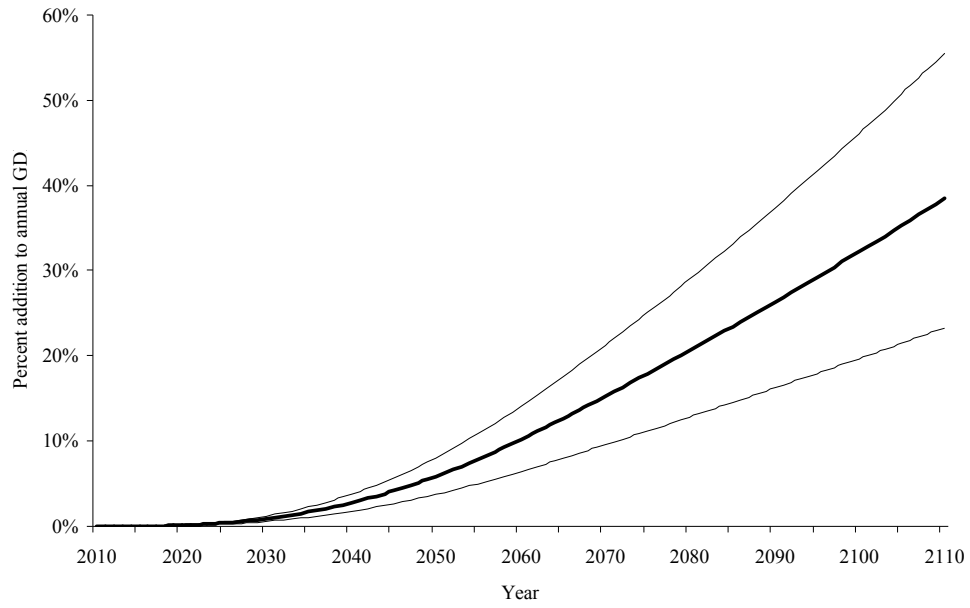


Figure 4. Improvement in annual GDP with Scenario I (increase everyone by $\frac{1}{4}$ std. dev.)

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.

Source: Authors' calculations.

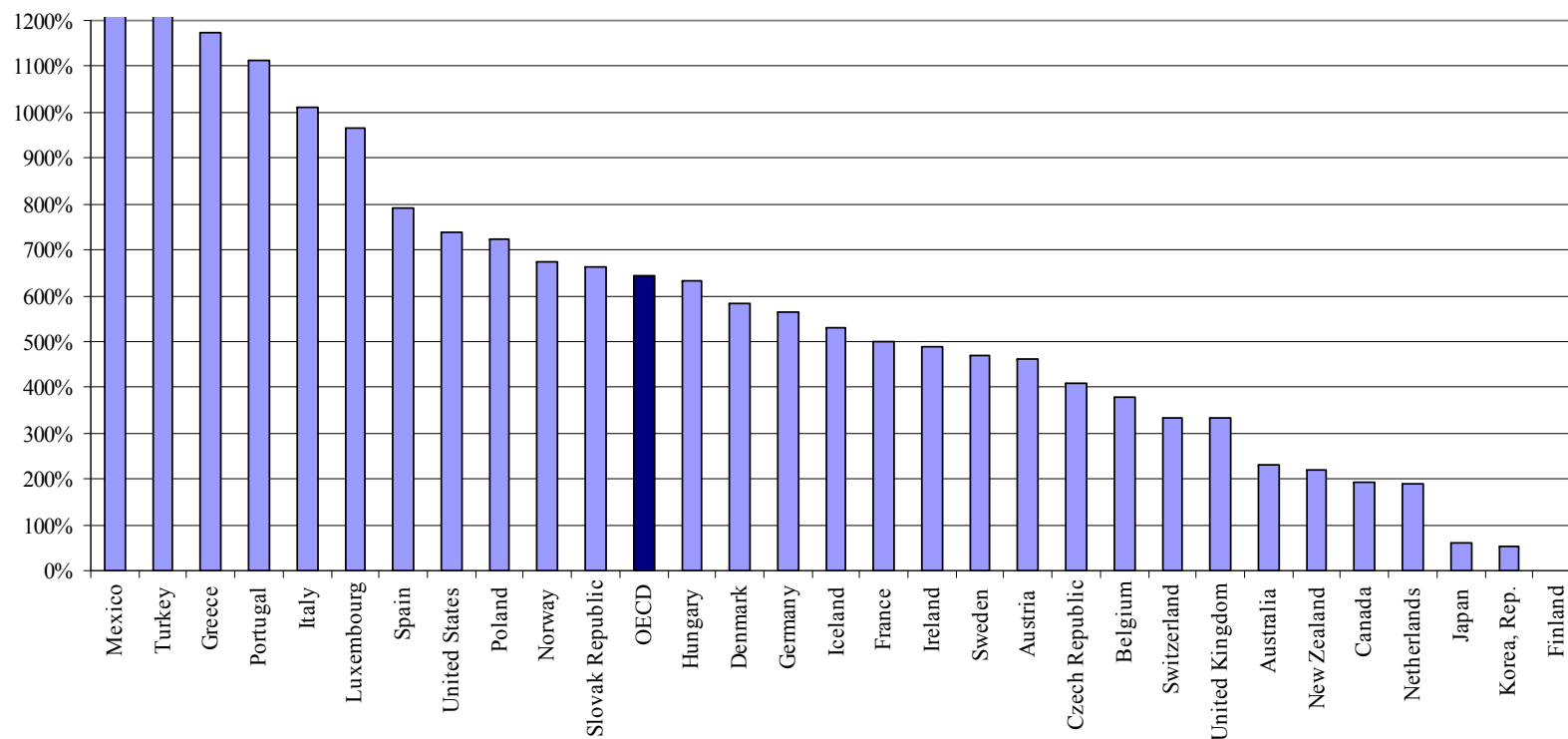


Figure 5. Present value of Scenario II (bring each country to Finish level) in percent of current GDP

Notes: Discounted value of future increases in GDP until 2090 due to a reform that brings each country to the Finish average level of 546 PISA points, expressed as percentage of current GDP. Value is 2389% for Mexico and 1844% for Turkey.

Source: Authors' depiction based on the projection analysis reported in column (3) of Table 7.

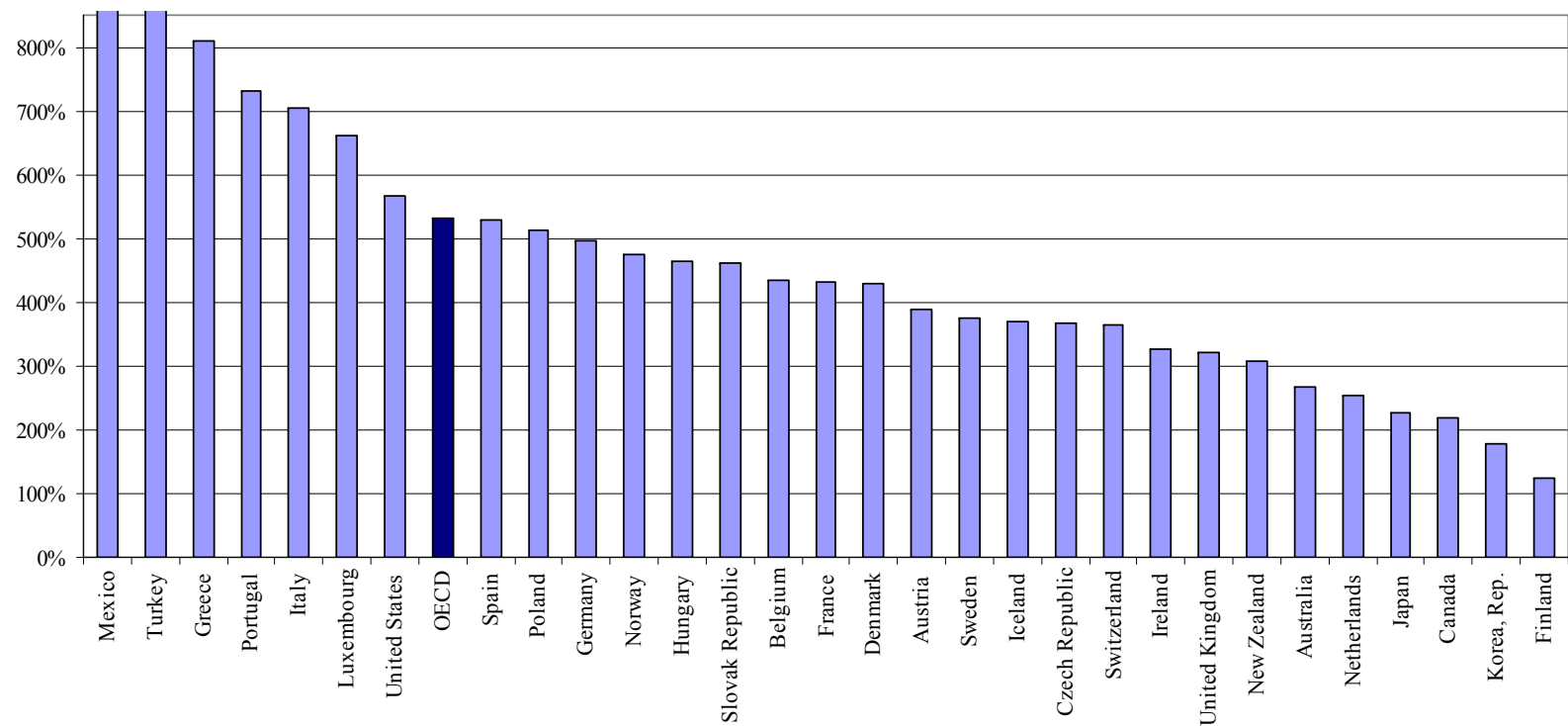


Figure 6. Present value of Scenario III (bring everyone up to minimum skills) in percent of current GDP

Notes: Discounted value of future increases in GDP until 2090 due to a reform that brings every individual in the country up to a minimum skill level of 400 PISA points, expressed as percentage of current GDP. Value is 1794% for Mexico and 1430% for Turkey.

Source: Authors' depiction based on the projection analysis reported in column (7) of Table 7.

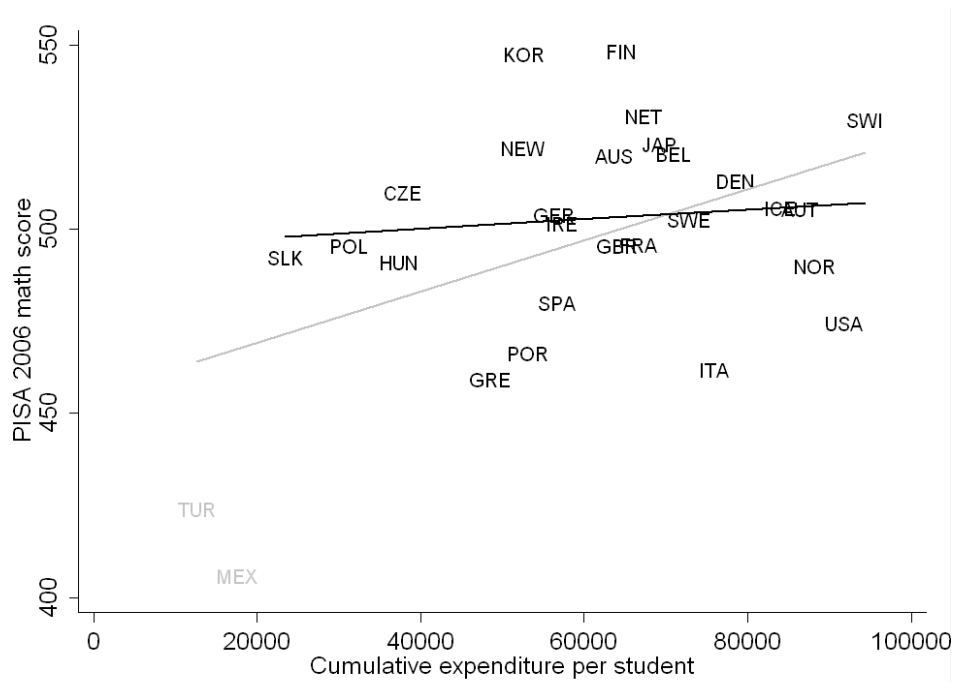


Figure 7. Educational spending and student achievement across countries

Notes: The figure plots the average math achievement in PISA 2006 against cumulative expenditure on educational institutions per student between age 6 and 15, in US dollars, converted by purchasing power parities. Gray line: regression line for full sample. Black line: regression line for countries with expenditure above \$20,000.

Source: Authors' depiction based on Organisation for Economic Co-operation and Development (2007).

Table 1: Income, growth, and schooling in OECD countries, 1960-2000

		Growth			GDP per capita			Years of schooling			Test scores		
		1960-2000	1960-1980	1980-2000	1960	1980	2000	1960	1980	2000	Mean	Basic	Top
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Australia	AUS	2.22	2.41	2.03	10,618	17,092	25,535	9.3	11.5	11.8	509.4	93.8	11.2
Austria	AUT	2.96	3.86	2.07	7,365	15,706	23,681	4.2	7.3	9.0	508.9	93.1	9.7
Belgium	BEL	2.84	3.78	1.91	7,759	16,303	23,784	6.9	8.5	10.2	504.1	93.1	9.4
Canada	CAN	2.40	3.06	1.75	10,419	19,022	26,922	8.3	9.7	11.0	503.8	94.8	8.3
Denmark	DEN	2.25	2.61	1.90	10,917	18,282	26,627	7.9	9.3	9.7	496.2	88.8	8.8
Finland	FIN	2.95	3.73	2.17	7,438	15,484	23,798	5.7	8.3	8.2	512.6	95.8	12.4
France	FRA	2.65	3.68	1.63	7,860	16,201	22,371	4.2	6.0	9.5	504.0	92.6	8.5
Greece	GRE	3.19	5.34	1.09	4,159	11,767	14,625	7.4	7.1	8.9	460.8	79.8	4.2
Iceland	ICE	2.84	4.07	1.61	8,106	18,017	24,794	6.2	7.9	9.6	493.6	90.8	7.4
Ireland	IRE	4.14	3.30	4.98	5,208	9,978	26,379	8.0	9.9	10.9	499.5	91.4	9.4
Italy	ITA	2.95	4.08	1.83	6,817	15,161	21,794	4.9	6.5	9.1	475.8	87.5	5.4
Japan	JAP	4.26	6.24	2.31	4,657	15,631	24,672	8.0	9.3	10.9	531.0	96.7	16.8
Korea, Rep.	KOR	5.95	5.78	6.13	1,571	4,830	15,881	4.3	8.3	11.1	533.8	96.2	17.8
Mexico	MEX	2.00	3.30	0.71	3,970	7,603	8,766	2.8	4.9	7.6	399.8	48.9	0.9
Netherlands	NET	2.44	2.82	2.06	9,263	16,164	24,313	6.3	9.4	10.7	511.5	96.5	9.2
New Zealand	NEW	1.23	1.07	1.38	11,555	14,304	18,824	10.2	11.8	12.2	497.8	91.0	10.6
Norway	NOR	3.02	3.62	2.42	8,239	16,772	27,044	7.6	9.1	11.3	483.0	89.4	5.6
Portugal	POR	3.92	4.95	2.89	3,434	9,024	15,955	3.2	5.5	7.6	456.4	80.3	3.2
Spain	SPA	3.43	4.59	2.27	4,693	11,520	18,055	3.3	6.2	9.3	482.9	85.9	7.9
Sweden	SWE	2.15	2.69	1.61	10,112	17,179	23,662	7.3	9.4	11.1	501.3	93.9	8.8
Switzerland	SWI	1.45	2.05	0.85	14,877	22,320	26,422	7.6	10.3	9.6	514.2	91.9	13.4
Turkey	TUR	2.35	2.38	2.32	2,700	4,325	6,838	1.6	3.6	6.0	412.8	58.2	3.9
United Kingdom	GBR	2.10	1.98	2.21	9,682	14,340	22,188	7.0	8.1	9.0	495.0	92.9	8.8
United States	USA	2.50	2.75	2.25	12,414	21,337	33,308	9.2	12.2	13.0	490.3	91.8	7.3
Mean		2.84	3.51	2.18	7,660	14,515	21,927	6.3	8.3	9.9	490.8	88.1	8.7
Minimum		1.23	1.07	0.71	1,571	4,325	6,838	1.6	3.6	6.0	399.8	48.9	0.9
Maximum		5.95	6.24	6.13	14,877	22,320	33,308	10.2	12.2	13.0	533.8	96.7	17.8

Notes: Sample: 24 OECD countries with income and education data for the whole period. Growth = average annual growth rate in GDP per capita (in percent). GDP per capita: in constant international dollars. Years of schooling: population aged 15 years and above. Test scores: average of math and science, all available tests (1964-2003); basic/top: percent of students over 400/600. *Sources:* Heston, Summers, and Aten (2002); Barro and Lee (2010); Hanushek and Woessmann (2009).

Table 2: Basic results on educational outcomes and long-run economic growth in OECD countries

		1960-1980	1980-2000	Avg. years ^a	GDP>10,000 ^b	Growth<5% ^c	Full country sample			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Cognitive skills		1.864 (5.83)	1.968 (6.72)	1.737 (2.95)	2.350 (3.91)	1.735 (5.26)	1.398 (2.89)	1.525 (4.33)		1.978 (7.98)
Initial years of schooling	0.173 (2.09)	0.046 (0.82)		0.024 (0.23)	0.147 (1.50)	0.096 (1.40)	0.029 (0.50)	0.056 (1.06)	0.396 (2.89)	0.080 (1.07)
Initial GDP per capita	-0.293 (5.20)	-0.303 (8.61)	-0.285 (10.33)	-0.338 (5.23)	-0.250 (6.18)	-0.311 (9.55)	-0.302 (8.58)	-0.273 (7.41)	-0.405 (4.67)	-0.313 (5.61)
OECD									1.437 (1.68)	0.859 (0.32)
OECD x Cognitive skills										-0.203 (0.36)
OECD x Initial years of schooling									-0.110 (0.65)	
No. of countries	24	24	24	24	24	24	22	23	50	50
<i>F</i> (OECD and interaction)									2.37	0.10
<i>R</i> ² (adj.)	0.559	0.828	0.831	0.630	0.658	0.839	0.835	0.738	0.297	0.723

Notes: Dependent variable: average annual growth rate in GDP per capita, 1960-2000 (unless noted otherwise). Sample: OECD countries (unless noted otherwise). Regressions include a constant. *t*-statistics in parentheses. ^a Years of schooling averaged between 1960 and 2000. ^b Sample excludes Mexico and Turkey. ^c Sample excludes Korea.

Source: Authors' calculations based on data in Heston, Summers, and Aten (2002), Hanushek and Woessmann (2009), and Barro and Lee (2010).

Table 3: Economic institutions vs. educational outcomes in OECD-country long-run growth models

	OECD sample			Full country sample	
	(1)	(2)	(3)	(4)	(5)
Cognitive skills		1.527 (3.39)	1.388 (3.05)	1.265 (4.76)	1.266 (4.96)
Initial years of schooling	0.135 (2.11)	0.058 (1.01)	0.067 (1.18)	0.067 (1.02)	0.047 (0.74)
Initial GDP per capita	-0.381 (7.83)	-0.328 (7.75)	-0.322 (7.64)	-0.373 (8.10)	-0.304 (5.65)
Protection against expropriation	0.729 (3.95)	0.224 (1.06)	0.081 (0.36)	0.332 (2.36)	0.492 (3.25)
Openness			0.750 (1.56)	0.485 (1.34)	0.645 (1.83)
OECD					2.589 (1.26)
OECD x Protection against expropriation					-0.372 (1.54)
No. of countries	24	24	23	48	48
F (Protection and Openness)			1.77	6.17	
F (OECD and interaction)					2.76
R^2 (adj.)	0.740	0.830	0.841	0.789	0.806

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

Source: Authors' calculations based on data in Heston, Summers, and Aten (2002), Barro and Lee (2010), Hanushek and Woessmann (2009), Acemoglu, Johnson, and Robinson (2001), and an updated version of Sachs and Warner (1995).

Table 4: Regulation of product and labor markets vs. educational outcomes in OECD-country long-run growth models

	Product market regulation (PMR)						Employment protection regulation (EPR)				
	Product market regulation	Administrative regulation	Domestic economic regulation	State control	Barriers to entrepreneurship	Barriers to trade and investment	EP version 1	EP version 2	Dismissal regular contracts	Regulation temporary contracts	Regulation collective dismissal
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Cognitive skills	1.843 (5.67)	1.897 (5.92)	1.840 (5.60)	1.843 (5.66)	1.859 (5.75)	1.864 (5.67)	1.827 (5.41)	1.788 (5.06)	1.899 (5.39)	1.788 (5.00)	1.913 (5.95)
Initial years of schooling	0.026 (0.41)	0.015 (0.23)	0.036 (0.60)	0.033 (0.55)	0.025 (0.40)	0.047 (0.75)	0.027 (0.41)	0.025 (0.37)	0.034 (0.48)	0.036 (0.58)	0.071 (1.21)
Initial GDP per capita	-0.309 (8.43)	-0.310 (8.69)	-0.309 (8.27)	-0.311 (8.30)	-0.308 (8.52)	-0.303 (8.36)	-0.310 (8.09)	-0.307 (8.25)	-0.308 (7.89)	-0.307 (8.22)	-0.324 (8.52)
Product market regulation	-0.157 (0.68)	-0.178 (1.06)	-0.080 (0.55)	-0.084 (0.67)	-0.157 (0.76)	0.004 (0.02)					
Employment protection legislation							-0.081 (0.62)	-0.109 (0.63)	-0.051 (0.34)	-0.047 (0.58)	0.139 (1.42)
No. of countries	24	24	24	24	24	24	23	23	23	23	23
R^2 (adj.)	0.824	0.829	0.822	0.824	0.825	0.819	0.822	0.822	0.819	0.821	0.836

Notes: Dependent variable: average annual growth rate in GDP per capita, 1960-2000. Sample: OECD countries. Regressions include a constant. *t*-statistics in parentheses.

Source: Authors' calculations based on data in Heston, Summers, and Aten (2002), Hanushek and Woessmann (2009), Barro and Lee (2010), Venn (2009), and Wölfl, Wanner, Kozluk, and Nicoletti (2009).

Table 5: Basic vs. high-level skills

Sample:	full	OECD	full	full	OECD	full	full	OECD	full
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Share of students reaching basic literacy	2.644 (3.51)	4.434 (3.52)	2.146 (2.58)	4.674 (6.71)	5.272 (5.65)	4.549 (5.54)			
Share of top-performing students	12.602 (4.35)	3.234 (0.99)	16.536 (4.90)				18.904 (7.45)	10.933 (3.61)	21.499 (7.19)
Initial years of schooling	0.066 (0.87)	0.035 (0.60)	0.070 (0.94)	0.129 (1.49)	0.039 (0.68)	0.131 (1.47)	0.116 (1.40)	0.086 (1.23)	0.122 (1.53)
Initial GDP per capita	-0.305 (6.43)	-0.305 (8.24)	-0.317 (5.63)	-0.326 (5.85)	-0.314 (8.71)	-0.340 (5.05)	-0.297 (5.62)	-0.275 (6.09)	-0.314 (5.32)
OECD			-0.659 (0.44)			0.150 (0.10)			1.018 (2.17)
OECD x basic literacy			2.074 (0.94)			0.024 (0.01)			
OECD x top-performing			-13.422 (2.08)						-10.733 (2.11)
No. of countries	50	24	50	50	24	50	50	24	50
<i>F</i> (OECD and interaction)			1.62			0.10			2.66
<i>R</i> ² (adj.)	0.724	0.822	0.734	0.616	0.822	0.600	0.655	0.720	0.679

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

Source: Authors' calculations based on data in Heston, Summers, and Aten (2002), Barro and Lee (2010), and Hanushek and Woessmann (2009).

Table 6: Non-tertiary vs. tertiary schooling

	Sample:	full	OECD	OECD	OECD w/o USA	OECD	OECD	OECD w/o USA
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Cognitive skills		1.923 (9.12)	1.888 (6.09)	1.912 (6.83)	2.043 (8.14)			
Share of students reaching basic literacy						5.458 (7.13)		
Share of top-performing students							11.597 (3.93)	12.855 (4.44)
Years of non-tertiary schooling		0.076 (0.94)	0.012 (0.059)					
Years of tertiary schooling		0.198 (0.16)	1.291 (1.58)	1.344 (1.77)	0.543 (0.74)	1.685 (2.31)	1.014 (0.96)	0.149 (0.13)
Initial GDP per capita		-0.325 (6.81)	-0.323 (8.85)	-0.320 (9.78)	-0.323 (11.22)	-0.344 (10.53)	-0.264 (6.08)	-0.263 (6.37)
No. of countries		50	24	24	23	24	24	23
R^2 (adj.)		0.728	0.839	0.847	0.886	0.856	0.712	0.750

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

Source: Authors' calculations based on data in Heston, Summers, and Aten (2002), Barro and Lee (2010), and Hanushek and Woessmann (2009).

Table 7: Baseline projections of the economic value of three education reform scenarios

	Scenario I: Increase avg. performance by ¼ std. dev.	Scenario II: Bring each country to Finish level of 546 points on PISA				Scenario III: Bring all to minimum of 400 points on PISA			
	Value of reform (billion \$)	Value of reform (billion \$)	in % of current GDP	L.-r. growth increase (p.p.)	Note: Increase in PISA score	Value of reform (billion \$)	in % of current GDP	L.-r. growth increase (p.p.)	Note: Share of students below minimum skills
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Australia	2,631	2,092	229%	0.38	20.1	2,430	266%	0.43	9.8%
Austria	969	1,545	460%	0.72	38.4	1,308	390%	0.62	13.9%
Belgium	1,208	1,586	379%	0.60	32.2	1,816	434%	0.68	15.3%
Canada	4,051	2,728	194%	0.32	17.2	3,075	219%	0.36	8.1%
Czech Republic	830	1,177	409%	0.64	34.5	1,054	366%	0.58	13.1%
Denmark	608	1,231	584%	0.88	47.5	908	430%	0.67	15.2%
Finland	594	0	0%	0.00	0.0	255	124%	0.21	4.7%
France	6,557	11,349	499%	0.77	41.3	9,844	433%	0.68	15.3%
Germany	8,822	17,245	564%	0.86	46.0	15,166	496%	0.77	17.3%
Greece	1,047	4,253	1172%	1.59	85.2	2,943	811%	1.17	26.5%
Hungary	603	1,323	633%	0.95	51.0	972	465%	0.72	16.3%
Iceland	36	66	530%	0.81	43.6	46	371%	0.59	13.3%
Ireland	585	995	490%	0.76	40.6	664	327%	0.52	11.8%
Italy	5,526	19,353	1010%	1.41	75.6	13,503	705%	1.04	23.5%
Japan	13,280	2,871	62%	0.11	5.7	10,382	226%	0.37	8.3%
Korea, Rep.	4,120	756	53%	0.09	4.8	2,544	178%	0.30	6.7%
Luxembourg	126	421	963%	1.36	72.7	289	662%	0.99	22.3%
Mexico	4,753	39,363	2389%	2.68	143.9	29,557	1794%	2.19	49.5%
Netherlands	2,032	1,344	191%	0.31	16.9	1,779	253%	0.41	9.3%
New Zealand	361	275	220%	0.36	19.4	385	308%	0.49	11.2%
Norway	844	1,975	675%	1.00	53.9	1,391	476%	0.74	16.6%
Poland	2,119	5,320	724%	1.07	57.2	3,766	513%	0.79	17.8%
Portugal	742	2,860	1112%	1.52	81.7	1,878	730%	1.07	24.2%
Slovak Republic	343	787	661%	0.99	52.9	549	461%	0.72	16.2%
Spain	4,496	12,332	791%	1.15	61.7	8,237	529%	0.81	18.3%
Sweden	1,080	1,761	470%	0.73	39.2	1,406	375%	0.59	13.4%
Switzerland	1,003	1,159	333%	0.53	28.6	1,263	363%	0.58	13.0%
Turkey	3,043	19,450	1844%	2.24	120.1	15,089	1430%	1.85	41.8%
United Kingdom	6,862	7,892	332%	0.53	28.5	7,669	322%	0.52	11.7%
United States	43,835	111,923	737%	1.08	58.1	86,167	567%	0.86	19.4%
OECD	123,108	275,429	645%	0.93	49.8	226,333	530%	0.80	18.0%

Notes: Discounted value of future increases in GDP until 2090, expressed in billion \$ (PPP) and as percentage of current GDP. "Long-run growth increase" refers to increase in annual growth rate (in percentage points) once the whole labor force has reached higher level of educational performance. "Increase in PISA score" refers to the ultimate increase in educational performance due to reform scenario II. "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. See text for reform parameters. Source: Authors' calculations.

Table 8: Projection results with “neoclassical” model specification

	Scenario I:		Scenario II:		Scenario III:		Scenario II,		Scenario II,	
	Increase avg. performance by ¼ std. dev.		Bring each country to Finish level of 546 points on PISA		Bring all to minimum of 400 points on PISA		Time horizon 2050		Time horizon 2150	
	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Australia	2,073	227%	1,656	182%	3,826	419%	276	30%	3,696	405%
Austria	712	212%	1,120	334%	1,471	438%	186	55%	2,511	748%
Belgium	926	221%	1,207	288%	2,334	557%	200	48%	2,711	647%
Canada	3,282	234%	2,227	159%	5,209	371%	368	26%	4,984	355%
Czech Republic	781	271%	1,095	381%	1,631	567%	169	59%	2,570	893%
Denmark	435	206%	859	407%	869	412%	143	68%	1,932	916%
Finland	560	272%	0	0%	651	316%	0	0%	0	0%
France	5,026	221%	8,552	376%	11,090	488%	1,397	61%	19,389	853%
Germany	6,521	213%	12,466	408%	15,347	502%	2,049	67%	28,191	922%
Greece	671	185%	2,548	702%	1,459	402%	417	115%	5,837	1609%
Hungary	618	296%	1,322	632%	1,302	623%	193	93%	3,213	1538%
Iceland	25	204%	46	367%	46	371%	8	61%	102	819%
Ireland	386	190%	645	318%	643	317%	111	55%	1,417	698%
Italy	3,725	194%	12,330	644%	8,102	423%	2,011	105%	28,244	1474%
Japan	12,584	273%	2,772	60%	24,595	534%	442	10%	6,329	137%
Korea, Rep.	4,489	314%	839	59%	7,094	497%	128	9%	1,969	138%
Luxembourg	46	105%	144	330%	93	214%	29	66%	291	665%
Mexico	3,451	209%	24,773	1504%	7,160	435%	3,583	217%	62,461	3791%
Netherlands	1,623	230%	1,082	154%	2,979	423%	180	26%	2,413	342%
New Zealand	360	288%	276	221%	790	632%	42	34%	646	516%
Norway	435	149%	985	337%	855	292%	181	62%	2,088	714%
Poland	2,192	298%	5,322	725%	4,651	633%	770	105%	13,051	1777%
Portugal	579	225%	2,099	816%	1,209	470%	323	126%	4,989	1940%
Slovak Republic	337	283%	749	630%	682	573%	111	93%	1,809	1519%
Spain	3,142	202%	8,281	531%	6,275	403%	1,359	87%	18,825	1208%
Sweden	784	209%	1,259	336%	1,542	412%	210	56%	2,818	753%
Switzerland	722	208%	831	239%	1,598	460%	141	41%	1,837	528%
Turkey	2,699	256%	15,474	1467%	6,363	603%	2,162	205%	39,523	3747%
United Kingdom	5,504	231%	6,308	265%	10,918	459%	1,032	43%	14,243	599%
United States	25,344	167%	62,386	411%	56,407	371%	10,962	72%	135,962	895%
OECD	90,031	211%	179,655	421%	187,191	439%	29,183	68%	414,050	970%

Notes: Discounted value of future increases in GDP until 2090, expressed in billion \$ (PPP) and as percentage of current GDP. See text for reform parameters. *Source:* Authors' calculations.

Table 9: Projection results under alternative parameter assumptions on growth coefficients and time horizon

	Lowest coefficient		Highest coefficient		Lower bound confidence		Upper bound confidence		Time horizon 2050		Time horizon 2150	
	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Australia	1,548	170%	2,217	243%	1,317	144%	2,899	318%	308	34%	5,805	636%
Austria	1,128	336%	1,641	489%	954	284%	2,182	650%	219	65%	4,555	1,357%
Belgium	1,163	278%	1,683	402%	986	235%	2,224	531%	227	54%	4,576	1,093%
Canada	2,022	144%	2,888	206%	1,722	123%	3,767	268%	404	29%	7,497	534%
Czech Republic	862	299%	1,249	434%	730	254%	1,655	575%	168	58%	3,423	1,189%
Denmark	893	423%	1,310	621%	753	357%	1,757	833%	171	81%	3,750	1,778%
Finland	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
France	8,268	364%	12,065	531%	6,989	308%	16,083	708%	1,596	70%	33,815	1,488%
Germany	12,519	409%	18,348	600%	10,567	346%	24,569	803%	2,400	78%	52,262	1,709%
Greece	2,993	825%	4,558	1256%	2,493	687%	6,354	1751%	540	149%	15,071	4,153%
Hungary	957	458%	1,409	674%	807	386%	1,896	907%	182	87%	4,084	1,954%
Iceland	48	385%	70	563%	41	325%	94	753%	9	74%	198	1,591%
Ireland	725	357%	1,057	521%	613	302%	1,409	694%	140	69%	2,957	1,457%
Italy	13,726	717%	20,702	1081%	11,473	599%	28,562	1491%	2,516	131%	65,827	3,436%
Japan	2,145	47%	3,034	66%	1,833	40%	3,920	85%	436	9%	7,618	165%
Korea, Rep.	565	40%	799	56%	483	34%	1,031	72%	115	8%	2,000	140%
Luxembourg	299	685%	450	1030%	250	573%	619	1417%	55	126%	1,415	3,238%
Mexico	26,267	1594%	42,714	2593%	21,416	1300%	63,912	3879%	4,306	261%	186,466	11,318%
Netherlands	996	141%	1,423	202%	849	120%	1,855	263%	199	28%	3,689	524%
New Zealand	204	163%	292	233%	173	139%	381	305%	41	32%	762	610%
Norway	1,425	487%	2,104	719%	1,200	410%	2,839	971%	270	92%	6,162	2,107%
Poland	3,829	521%	5,672	772%	3,221	438%	7,679	1045%	722	98%	16,813	2,289%
Portugal	2,018	785%	3,063	1191%	1,684	655%	4,254	1654%	366	142%	9,983	3,882%
Slovak Republic	568	477%	838	704%	478	402%	1,130	949%	108	91%	2,445	2,054%
Spain	8,845	568%	13,158	844%	7,428	477%	17,893	1148%	1,656	106%	39,653	2,544%
Sweden	1,285	343%	1,872	500%	1,087	290%	2,490	665%	249	66%	5,208	1,391%
Switzerland	852	245%	1,229	353%	723	208%	1,619	466%	168	48%	3,304	950%
Turkey	13,273	1258%	20,994	1990%	10,920	1035%	30,478	2889%	2,265	215%	81,227	7,700%
United Kingdom	5,804	244%	8,372	352%	4,926	207%	11,029	464%	1,141	48%	22,498	946%
United States	80,503	530%	119,336	785%	67,686	445%	161,693	1064%	15,155	100%	354,821	2,335%
OECD	195,731	459%	294,547	690%	163,804	384%	406,270	952%	36,131	85%	947,885	2,221%

Notes: Scenario II: Bring each country to Finish level of 546 points on PISA. Discounted value of future increases in GDP until 2090, expressed in billion \$ (PPP) and as percentage of current GDP.

Source: Authors' calculations.

Table 10: Projection results under alternative parameter assumptions on reform duration, working life, and discount rate

	10-year reform		30-year reform		Working life 35 years		Discount rate 2.5%		Discount rate 3.5%		Disc. rate Stern Report	
	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP	Billion \$	% GDP
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Australia	2,541	279%	1,721	189%	2,287	251%	2,788	306%	1,581	173%	4,763	522%
Austria	1,887	562%	1,263	376%	1,694	504%	2,062	614%	1,165	347%	3,533	1,052%
Belgium	1,933	462%	1,299	310%	1,736	415%	2,115	505%	1,197	286%	3,621	865%
Canada	3,310	236%	2,246	160%	2,980	212%	3,633	259%	2,062	147%	6,204	442%
Czech Republic	1,435	499%	963	335%	1,289	448%	1,570	545%	888	309%	2,689	934%
Denmark	1,508	715%	1,004	476%	1,352	641%	1,645	780%	928	440%	2,823	1,338%
Finland	0	0%	0	0%	0	0%	0	0%	0	0%	0	0%
France	13,875	610%	9,272	408%	12,449	548%	15,152	667%	8,559	377%	25,981	1,143%
Germany	21,116	691%	14,067	460%	18,934	619%	23,034	753%	13,000	425%	39,530	1,293%
Greece	5,279	1455%	3,421	943%	4,705	1296%	5,703	1572%	3,193	880%	9,855	2,716%
Hungary	1,623	777%	1,078	516%	1,454	696%	1,768	846%	997	477%	3,037	1,453%
Iceland	81	648%	54	432%	72	581%	88	707%	50	399%	151	1,213%
Ireland	1,216	599%	813	401%	1,091	538%	1,328	654%	750	370%	2,277	1,122%
Italy	23,939	1250%	15,624	816%	21,367	1115%	25,926	1353%	14,545	759%	44,724	2,335%
Japan	3,472	75%	2,372	52%	3,130	68%	3,820	83%	2,173	47%	6,510	141%
Korea, Rep.	914	64%	625	44%	824	58%	1,005	70%	572	40%	1,713	120%
Luxembourg	520	1191%	340	779%	465	1063%	564	1290%	317	724%	972	2,225%
Mexico	50,015	3036%	30,929	1877%	44,093	2676%	53,101	3223%	29,371	1783%	92,736	5,629%
Netherlands	1,630	231%	1,106	157%	1,468	208%	1,790	254%	1,016	144%	3,056	434%
New Zealand	334	267%	227	181%	301	241%	367	293%	208	166%	627	501%
Norway	2,425	829%	1,607	549%	2,172	742%	2,640	903%	1,488	509%	4,537	1,551%
Poland	6,539	890%	4,323	589%	5,854	797%	7,114	969%	4,006	545%	12,233	1,665%
Portugal	3,546	1379%	2,304	896%	3,161	1229%	3,834	1491%	2,148	835%	6,621	2,575%
Slovak Republic	965	811%	640	538%	865	726%	1,051	883%	593	498%	1,807	1,518%
Spain	15,180	974%	10,005	642%	13,579	871%	16,498	1059%	9,282	596%	28,390	1,822%
Sweden	2,152	575%	1,440	385%	1,931	516%	2,351	628%	1,329	355%	4,030	1,076%
Switzerland	1,411	406%	951	273%	1,268	365%	1,545	444%	875	252%	2,643	760%
Turkey	24,470	2320%	15,436	1463%	21,672	2055%	26,174	2481%	14,550	1379%	45,513	4,315%
United Kingdom	9,609	404%	6,475	272%	8,638	363%	10,523	442%	5,960	251%	18,005	757%
United States	137,603	906%	90,922	598%	123,157	811%	149,676	985%	84,272	555%	257,402	1,694%
OECD	340,528	798%	222,527	521%	303,987	712%	368,864	864%	207,075	485%	635,982	1,490%

Notes: Scenario II: Bring each country to Finish level of 546 points on PISA. Discounted value of future increases in GDP until 2090, expressed in billion \$ (PPP) and as percentage of current GDP.

Source: Authors' calculations.

Table 11: Summary of projection results

	Scenario I: Increase avg. performance by ¼ std. dev. (1)	Scenario II: Bring each country to Finish level of 546 points on PISA (2)	Scenario III: Bring all to minimum of 400 points on PISA (3)
“Endogenous-growth” specification			
in billion \$	123,108	275,429	226,333
in % of discounted future GDP	6.2%	13.8%	11.3%
“Neoclassical” specification			
in billion \$	90,031	179,655	187,191
in % of discounted future GDP	4.3%	8.5%	8.9%

Notes: Discounted value of future increases in OECD GDP until 2090, expressed in billion \$ (PPP) and as percentage of discounted value of all annual projected OECD GDPs until 2090.

Source: Authors' calculations.

Table 12: Educational spending and student achievement

	PISA 2000		PISA 2003		PISA 2006		FD	FE
	<i>Sample:</i> all	exp>\$25,000	all	exp>\$25,000	all	exp>\$25,000	all	all
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expenditure per student	1.079 (2.76)	0.069 (0.13)	0.670 (2.06)	0.203 (0.58)	0.697 (2.70)	0.097 (0.33)	-0.260 (0.44)	-0.125 (0.35)
No. of countries	24	19	26	23	28	25	23	29 (78)
R^2 (adj.)	0.224	-0.058	0.114	-0.031	0.189	-0.039	-0.038	0.027

Notes: Dependent variable: PISA math test score. Expenditure per student = cumulative expenditure on educational institutions per student between age 6 and 15, measured in 1,000 US dollars, converted by purchasing power parities. Sample: all = all OECD countries with available data; exp>\$25,000 = countries with cumulative expenditure per student above \$25,000. FD = first differences, 2006-2000. FE = fixed effects, pooling all three PISA tests, model contains time dummies (R^2 of FE model refers to within R^2). *t*-statistics in parentheses.

Source: Authors' calculations based on Organisation for Economic Co-operation and Development (2003, 2004, 2007).