Journal of Development Economics 99 (2012) 497–512

Contents lists available at SciVerse ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Schooling, educational achievement, and the Latin American growth puzzle[☆]

Eric A. Hanushek [a,*], Ludger Woessmann [b,1]

[a] Hoover Institution, Stanford University, CESifo and NBER, Stanford, CA 94305-6010, United States
[b] University of Munich, CESifo and Ifo Institute for Economic Research, Poschingerstr. 5, 81679 Munich, Germany

ARTICLE INFO

Article history:
Received 2 June 2009
Received in revised form 11 June 2012
Accepted 12 June 2012

JEL classification:
O4
I2
H4
N16

Keywords:
Economic growth
Latin America
Cognitive skills
Schooling

ABSTRACT

Latin American economic development has been perceived as a puzzle. The region has trailed most other world regions over the past half century despite relatively high initial development and school attainment levels. This puzzle, however, can be resolved by considering educational achievement, a direct measure of human capital. We introduce a new, more inclusive achievement measure that comes from splicing regional achievement tests into worldwide tests. In growth regressions, the positive growth effect of educational achievement fully accounts for the poor growth performance of Latin American countries. These results are confirmed in a number of instrumental-variable specifications that exploit plausibly exogenous achievement variation stemming from historical and institutional determinants of educational achievement. Finally, a development accounting analysis finds that, once educational achievement is included, human capital can account for between half and two thirds of the income differences between Latin America and the rest of the world.

1. Introduction

If transported back to 1960, one might well have expected Latin America to be on the verge of significant economic growth. Both its level of school attainment and its income level were well ahead of East Asia and of the Middle East and North Africa (MENA) region (Table 1). But by 2000, growth in East Asia had moved that region far ahead of Latin America. While not going as far, the MENA region also jumped ahead, leaving only Latin America and Sub-Saharan Africa at the bottom with very low growth rates and commensurate low income per capita.[2] This outcome remains a puzzle by conventional thinking. Why did Latin America have such a poor growth performance relative to Asia and even MENA, given its high schooling level in 1960? While much attention has been given to institutional and financial factors,[3] we suggest that the level of educational achievement (or cognitive skills, which we use interchangeably here) is the crucial component of the long-run picture.

In simplest terms, while Latin America has had reasonable school attainment, the skills of students remain comparatively very poor. In terms of student achievement on international tests, both Latin America and Sub-Saharan Africa are near the bottom of the international rankings, while MENA and especially East Asia are much higher. As Fig. 1 reveals, consideration of the low level of educational achievement appears sufficient to reconcile the poor growth performance of Latin America with outcomes in the rest of the world over

☆ We benefited from helpful comments from Paul Romer and the participants of the Stanford Conference on Latin America and the Caribbean and gratefully acknowledge support by the Inter-American Development Bank (IADB), CESifo, and the Pact for Research and Innovation of the Leibniz Association.

* Corresponding author. Tel.: +1 650 736 0942.
E-mail addresses: hanushek@stanford.edu (E.A. Hanushek), woessmann@ifo.de (L. Woessmann).
URL's: http://www.hanushek.net (E.A. Hanushek), http://www.cesifo.de/woessmann (L. Woessmann).
1 Tel.: +49 89 9224 1699.

2 Even there a mystery remains, because Latin America has considerably higher levels of school attainment in 2000 than does Sub-Saharan Africa. Of course, the recent spurt in growth in Latin America might represent a turnaround, but that would require a very uncertain extrapolation.

3 See, for example, Edwards et al. (2007) and Fernández-Arias et al. (2005). Cole et al. (2005) state unequivocally that "Latin America's TFP gap is not plausibly accounted for by human capital differences" (p. 69). Similarly, in a recent high-level forum on the puzzle of Mexico's disappointing growth performance, schooling gets only side mention (Hanson, 2010) or no mention at all (Kehoe and Ruhl, 2010). In contrast to these macroeconomic studies, several microeconomic studies highlight the high labor-market returns to years of schooling in Latin America (see Psacharopoulos and Patrinos (2004) for an overview). Over the past decades, these returns also tended to increase in Latin America (cf. Behrman et al., 2007; Pritchett, 2004). Apart from the returns to education quantity, labor-market returns to cognitive skills in the one Latin American country that participated in the International Adult Literacy Survey, Chile, are the second-highest of all participating countries after the United States (Hanushek and Zhang, 2009).

0304-3878/$ – see front matter
doi:10.1016/j.jdeveco.2012.06.004

Table 1
Latin American income and education in a global perspective.

	GDP per capita 1960 (1)	Growth of GDP per capita 1960–2000 (2)	GDP per capita 2000 (3)	Years of schooling 1960 (4)	Test score (5)
Asia	1891	4.5	13,571	4.0	479.8
Sub-Saharan Africa	2304	1.4	3792	3.3	360.0
Middle East and North Africa	2599	2.7	8415	2.7	412.4
Latin America	4152	1.8	8063	4.7	388.3
Europe	7469	2.9	21,752	7.4	491.5
Commonwealth OECD	11,252	2.1	26,147	9.5	500.3
Note: Asia w/o Japan	1614	4.5	12,460	3.5	474.7

Underlying country sample: All countries with internationally comparable data on GDP that ever participated in a worldwide student achievement test; see Hanushek and Woessmann (forthcoming) for details. The country observations contained in the six regions are: Asia (11): China, Hong Kong, India, Indonesia, Japan, Rep. of Korea, Malaysia, Philippines, Singapore, Taiwan, Thailand; Commonwealth OECD members (4): Australia, Canada, New Zealand, USA; Europe (17): Austria, Belgium, Denmark, Finland, France, Greece, Iceland, Ireland, Italy, Netherlands, Norway, Portugal, Romania, Spain, Sweden, Switzerland, United Kingdom; Latin America (7): Argentina, Brazil, Chile, Colombia, Mexico, Peru, Uruguay; Middle East and North Africa (8): Cyprus, Egypt, Iran, Israel, Jordan, Morocco, Tunisia, Turkey; Sub-Saharan Africa (3): Ghana, South Africa, Zimbabwe. Sources: Own calculations based on Penn World Tables (Cohen and Soto, 2007; Hanushek and Woessmann, forthcoming; Heston et al., 2002.

the past four decades. Our interpretation is simple: Even though many things enter into economic growth and development, the educational achievement of the population are extremely important for long-run growth. Moreover, in the presence of measures of educational achievement, school attainment does not even have a significant relationship with growth. This finding corroborates the stylized fact discussed in the literature that performance on years of schooling data is largely inconsistent with growth performance (Bils and Klenow, 2000; Easterly, 2001; Pritchett, 2001, 2006), suggesting that considering acquired skills rather than time in school provides an explanation for this inconsistency. A crucial missing link in explaining why Latin America went from reasonably rich in the early post-war period to relatively poor today is its low educational achievement.

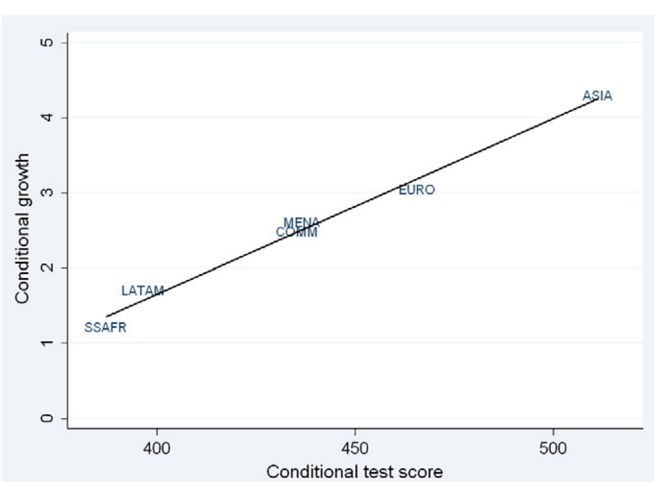

Fig. 1. Educational achievement and economic growth across world regions. Added-variable plot of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960–2000 on the initial level of real GDP per capita in 1960 and average scores on international student achievement tests (mean of the unconditional variables added to each axis). See Table 1 for a list of countries contained in each world region. Region codes: Asia (ASIA), Commonwealth OECD members (COMM), Europe (EURO), Latin America (LATAM), Middle East and North Africa (MENA), Sub-Saharan Africa (SSAFR).

Focusing on the relationship between educational achievement and economic development in Latin America introduces two main analytical concerns. First, prior work using worldwide achievement tests has relatively few observations from Latin America (seven of the available 50 countries in the analysis in Hanushek and Woessmann (2008)), making it difficult to analyze patterns of within-region economic outcomes. Second, the international assessments of math and science skills may simply be too difficult for the typical Latin American student, making the comparisons across Latin American countries unreliable.

The performance of Latin American countries on the worldwide student achievement tests has been truly dismal. Because test efficiency requires the international assessments to focus testing time on discriminating performance in the vicinity of the international mean, tests developed for the Organisation for Economic Co-operation and Development (OECD) may not have sufficient test questions to identify performance at the level of most Latin American countries reliably.[4]

This paper contributes in a variety of ways to the growing literature revealing the central role of educational achievement in economic development.[5] It introduces a new set of test scores from regional achievement tests that cover all 16 Latin American countries available for long-run growth analyses,[6] permitting the first comprehensive analysis of the role of skills in Latin American growth. Once educational achievement is taken into account, the analysis demonstrates that the pattern of Latin America's growth is indistinguishable from growth elsewhere in the world and that intraregional variations in Latin America can be consistently explained by the same factors. A number of instrumental-variable models add confidence that the relationships capture a causal impact of educational achievement. Similarly, a complementary development accounting analysis with the new educational achievement data adds support to the validity of the underlying growth regressions.

The expanded skill measures incorporate regional assessments of achievement that were designed specifically for Latin America. While Latin American countries participated only sporadically in the worldwide student achievement tests, the Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) conducted two regional tests of student achievement in math and reading that together cover the full usable set of 16 Latin American countries. The first LLECE assessment tested third and fourth grade students in 1997, the second survey – the Segundo Estudio Regional Comparativo Explicativo (SERCE) – tested third and sixth grade students in 2006. Neither of these is perfect, because they measure performance just in early grades and because both are very recent — with the second test actually occurring outside of the period for which growth is analyzed. Nonetheless, their regional test designs and broad coverage of countries hold promise for regional analyses. To our knowledge, neither of the tests has been used before in models of economic outcomes. We suggest a simple method to splice the regional educational assessments into the worldwide assessments. From an analytical perspective, this analysis demonstrates the feasibility of linking different assessments for the analysis of economic outcomes. Not only do the regional tests provide greater and more reliable detail on country differences at the low end of the economic

[4] Note that, while Mexico joined the OECD in 1994 and Chile in 2010, we place them within the Latin American set of countries throughout this analysis.

[5] For reviews of economic analyses of the role of educational achievement in international comparisons, see Hanushek and Woessmann (2008, 2011a).

[6] The criteria are having populations greater than one million and no communist background. We do not include Caribbean island countries in the analyses of this paper, as only two of them ever participated in the tests. Cuba lacks internationally comparable income data, and the remaining country – the Dominican Republic – proves a significant outlier in the analyses of the Latin American mainland countries.

distribution, but they also point to how comparisons can be made to the rest of the world.

Our results using the regional test data support the important role that educational achievement plays in understanding long-run Latin American growth. These test scores are significantly related to differences in economic growth within the Latin America region over the period 1960–2000. Furthermore, the new comprehensive dataset allows us to perform cross-country growth regressions on an extended sample of 59 countries that now includes 16 Latin American countries. While a Latin American dummy consistently proves significantly negative in standard growth models, it is statistically insignificant and close to zero once differences in educational achievement are controlled for.[7] Educational achievement is significantly associated with economic growth in the worldwide growth regressions. It increases the explanatory power of standard growth models considerably and renders the effect of years of schooling insignificant. Years of schooling appears relevant for economic growth only insofar as they actually raise the knowledge that students gain as depicted in achievement tests. Finally, when modeling a curvilinear relationship between the standard international skill metric and growth, the test score–growth nexus does not differ significantly between Latin America and the rest of the world.

Cross-country growth regressions are subject to concerns about possible endogeneity bias due to omitted variables and reverse causality. An additional innovation of this paper is the introduction of instrumental-variable estimation of the growth model that is designed to deal with major potential endogeneity issues. By instrumenting educational achievement with fundamental aspects of school quality – in particular, school attainment levels in 1960, historical Catholic shares in 1900 that predict modern levels of competition in the school system, and the relative position of teacher salaries in the income distribution of a country – we can estimate the growth equations using just variation in test scores that is plausibly exogenous. This estimation, while necessarily less precise than the simple cross-sectional estimation, provides strong support for the basic skills model of growth.

Finally, we complement our regression analysis with a development accounting analysis that extends human capital measurement to include our achievement measures. Instead of estimating the parameters in macro regressions, development accounting relies on parameters from the microeconometric literature to assess the importance of educational achievement in a standard functional form of the macroeconomic production function. In particular, we map years of schooling and educational achievement into aggregate human capital using consistent estimates of their micro returns on the U.S. labor market. Results show that human capital can account for between half and two thirds of the variation in current levels of per-capita income between Latin American countries and the rest of the world. In contrast, human capital accounts for only slightly more than a quarter when relying just on school attainment without consideration of differences in achievement. These results corroborate the major relevance of educational achievement in understanding Latin American growth performance.

This paper begins with a conceptual framework for the relationship between years of schooling, educational achievement, and economic growth. Section 3 provides descriptive evidence on the low levels of educational achievement in Latin American countries both from worldwide test data and from regional achievement tests and introduces a new test-score dataset that splices the regional Latin American data into the worldwide data. Section 4 uses these data in economic growth regressions to provide evidence on the role of educational achievement in understanding economic growth both within Latin America and relative to the rest of the world. Section 5 reports several instrumental-variable models to address fundamental endogeneity concerns. Section 6 turns to a development accounting exercise that calculates the role of human capital in accounting for Latin American levels of development relative to other regions.

2. Schooling, achievement, and growth: a conceptual framework

Theoretical models of economic growth have emphasized different mechanisms through which education may affect economic growth, stressing respectively the role of education as a production factor that can be accumulated (Mankiw et al., 1992), its role in increasing the innovative capacity of the economy (Aghion and Howitt, 1998; Romer, 1990), or its role in facilitating the transmission of knowledge needed to implement new technologies (Benhabib and Spiegel, 2005; Nelson and Phelps, 1966). What all approaches have in common is that they predict that education has a positive effect on growth, and in particular the latter two stress its impact on long-run growth trajectories.

An increasing wave of empirical growth research, following the seminal contributions by Barro (1991, 1997) and Mankiw et al. (1992), tries to estimate why some countries grow faster than others. The recent literature, involving cross-country growth regressions and invariably considering the impact of education, relies mostly on the important internationally comparable data on average years of schooling provided by Barro and Lee (1993, 2010) and its refinements (Cohen and Soto, 2007) as the proxy for the human capital of an economy.[8]

Here, we take an alternative perspective, originating in the work of Hanushek and Kimko (2000) and applied in a series of studies surveyed in Hanushek and Woessmann (2008), which concentrates directly on educational achievement and relies on the following model:

$$g = \gamma H + \beta X + \varepsilon \tag{1}$$

where g is the growth rate of real GDP per capita over an extended period, H is human capital, X is the other factors affecting growth, and ε is a stochastic term where it is assumed that $E(H,X|\varepsilon) = 0$.

The typical growth analysis simply substitutes a measure of school attainment for H when estimating Eq. (1), but this requires two very strong assumptions that each lack prima facie validity. First, it must be the case that a year of schooling produces the same knowledge and skills, or human capital, regardless of the country. For example, a year of schooling in Peru must be equivalent to a year in Japan, a difficult position to argue from the aggregate data below. Second, schooling must be the only systematic factor influencing skills, something that is refuted in virtually all individual-level analyses of achievement (Hanushek, 2002). The central issues for growth modeling are easily seen by considering additional sources of human capital accumulation:

$$H = \delta_1(qS) + \delta_2 F + \delta_3 A + \nu. \tag{2}$$

This formulation builds on the extensive literature of educational production functions. The components determining H include years of schooling (S) and schooling quality (q), family factors (F), and other attributes (A) including health, ability, and peer influences of the country's population. Eq. (2) suggests how inputs into the formation of human capital, such as schooling levels, could be used as a proxy for human capital when direct measures are unavailable. But, it also indicates how the interpretation is affected when only an imperfect set of measures is available.

[7] In our regression analysis across 50 countries, the seven Latin American countries unconditionally had an average growth rate over the period that was 1.3 percentage points lower than the rest of the sampled countries, and 1.4 percentage points after conditioning on initial income, years of schooling, and physical capital. Quite similarly, the three Sub-Saharan African countries had 1.6 percentage points slower growth, and conditionally even 1.9 percentage points. Suggestive evidence, which of course is very limited given the small number of participants from Sub-Saharan African in the international achievement tests, suggests that this African "growth tragedy" (Easterly and Levine, 1997) can also be accounted for by its low levels of educational achievement; detailed results are available from the authors on request.

[8] For extensive reviews of the literature, see Krueger and Lindahl (2001), Pritchett (2006), and Topel (1999).

Instead of estimating the components of Eq. (2), however, we turn to direct measures of cognitive skills as indicators of H.[9] Specifically, we rely on measures of educational achievement across countries that have been developed through international testing initiatives. Although human capital is a latent variable that is not directly observed, the use of achievement measures 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), this formulation opens the investigation of the importance of different policies designed to affect the quality aspects of schools.[10]

3. Educational achievement in Latin America

3.1. A description of educational achievement in Latin America based on worldwide data

The existing data from worldwide student achievement tests paint a bleak picture of performance in Latin America.[11] While Latin American countries have not participated frequently in the existing testing, their performance is uniformly uncompetitive either with developed countries or with many developing countries.

Between 1964 and 2003,[12] international agencies developed and deployed a total of 42 different international student achievement tests in math, science, or reading on 14 separate international testing occasions (several of which tested more than one subject and age level).[13] Only seven Latin American countries ever participated in any of the international math or science tests: Argentina, Brazil, Chile, Colombia, Mexico, Peru, and Uruguay.[14]

Before 2000, only Chile and Colombia participated in math or science tests based on an international curriculum. Their performance was at the bottom (between the second- and fourth-last ranks on five different occasions that included between 12 and 39 participating countries), and they only outperformed a handful of countries such as India, Iran, Malawi, and South Africa.[15] In IEA assessments after 2000, other Latin American countries also established positions near the bottom. Argentina and Colombia, for example, were fifth and sixth from the bottom (with only Belize, Morocco, Kuwait, and Iran below) in the 2001 Progress in International Reading Literacy Study (PIRLS) of 4th graders.

International testing expanded considerably in 2000 when the OECD started the Programme for International Student Assessment (PISA), which tests 15-year-old students in mathematics, science, and reading every 3 years. Yet, by 2003, only six Latin American countries participated in any of the PISA rounds, and the results mirrored the earlier testing. In 2000 and 2003, Indonesia and Tunisia were the only countries to keep Brazil and Mexico off the bottom of the 31 participants in the three tested subjects. In 2002, an additional ten countries took the 2000 test. Peru came out last, at an amazing distance, among the combined sample of 41 countries, whereas Argentina and Chile performed between sixth and eighth from the bottom on the three subjects (followed only by Albania, Indonesia, and Macedonia outside Latin America).

As a simple summary, for the 40 occasions on which a Latin American country participated in an international student achievement test until 2003 (counting different subjects and age groups separately), the average rank was 31.8 among an average of 34.5 participants (where a significant portion of the ranks below were taken up by other Latin American countries).

For our growth analysis, however, we need a description not just of the rank but of the magnitude of score differences. Comparing the level of performance across tests is difficult, because no attempt is made to calibrate the tests across time and because a varying group of countries has voluntarily participated in each of the existing international assessments. In order to make performance on the international mathematics and science tests comparable and usable to analyzing growth, Hanushek and Woessmann (forthcoming) develop a common metric for the tests between 1964 and 2003. The development of a common metric involves adjusting both the level of test performance and its variation across the different assessments. First, each of the separate international tests is benchmarked to a comparable level by calibrating the U.S. international performance over time to the external standard of the available U.S. longitudinal test (the National Assessment of Educational Progress, NAEP). Second, the dispersion of the tests is standardized by holding the score variance constant within a group of 13 OECD countries with relatively stable secondary school attendance rates over time. This empirical calibration puts all the international tests on the metric of the PISA test, which has a mean performance across the OECD countries of 500 and a standard deviation (at the student level) of 100.

Fig. 2 depicts the average performance between 1964 and 2003 on the standardized tests for the 50 countries contained in our growth analyses below, that is, all countries that have both participated in one of the tests and comparable income data. There is a clear performance gap between the best country in Latin America and the worst OECD country except Turkey, or any country in East Asia with the exception of Indonesia and the Philippines. In fact, the latter two countries, together with the African participants, are the only countries that consistently perform worse than any of the Latin American countries. Even the best-performing Latin American country, Uruguay, on average performs a full 0.70 standard deviations below the OECD

[9] Hanushek and Woessmann (2008) elaborate on this model to consider an imperfect measurement of H, particularly the consideration of noncognitive skills. Because schooling is likely to be correlated with the other determinants of human capital and we do not separately identify their effects, we see our results as measuring the effect of cognitive skills combined with that part of other human-capital components, including noncognitive skills, which are correlated with cognitive skills.

[10] An important extension of this framework would explore how the effect of schooling quantity S depends on its quality q. Note, however, that ours is a measure of total cognitive skills H, rather than of the quality of schools q, meaning that this question cannot be addressed by a simple interaction specification between the available measures of S and H.

[11] Throughout the paper, our analysis focuses on Latin American countries with greater than one million population. (The Latin American countries of Belize, French Guiana, Guyana, and Suriname all have a population of less than one million). We exclude Nicaragua from the economic analysis because of its extended period under communist rule and nonmarket conditions. Caribbean countries, while sometimes put together with Latin American countries, are not included in this analysis. No Caribbean country ever participated in the worldwide testing of math and science.

[12] Throughout the paper, we use the worldwide tests conducted until 2003 only, so as to remain close to our period of growth observations. In fact, six Latin American countries participated in the PISA 2006 cycle: four of them are among the bottom ten in math and science of the 57 participating countries. The only Latin American country ever making it to the "top 40" of the 57 countries is Chile (with rank 39 in reading).

[13] The available tests emanate from two main organizations — the International Association for the Evaluation of Educational Achievement, or IEA, and the OECD (see Hanushek and Woessmann (2011a) for details). The IEA introduced international testing in 1964 and has conducted periodic assessments up to the current TIMSS (Trends in Mathematics and Science Study). The OECD began international testing in 2000 with the Programme for International Student Assessment, or PISA. Both continue on a periodic schedule, and both the IEA and OECD have added reading assessments.

[14] As discussed below, all Latin American countries (with more than one million population) have participated in one or both regional testing programs conducted in 1997 and 2006 — a fact that we exploit below. On the international tests, Venezuela did participate in a 1991 reading test, and their student scores only exceeded those in Botswana, Nigeria, and Zimbabwe on the test for 13-year-olds and no other country on the test for 9-year-olds.

[15] All worldwide testing considered in this paper is based on an international collaboration designed to capture the typical curricular elements found across countries. An exception is the International Assessment of Educational Progress (IAEP) study which mirrors the U.S. curriculum. Brazil participated in the IAEP study in 1991, coming out second from the bottom (followed only by Mozambique) among 19 countries in math and last among 18 countries in science.

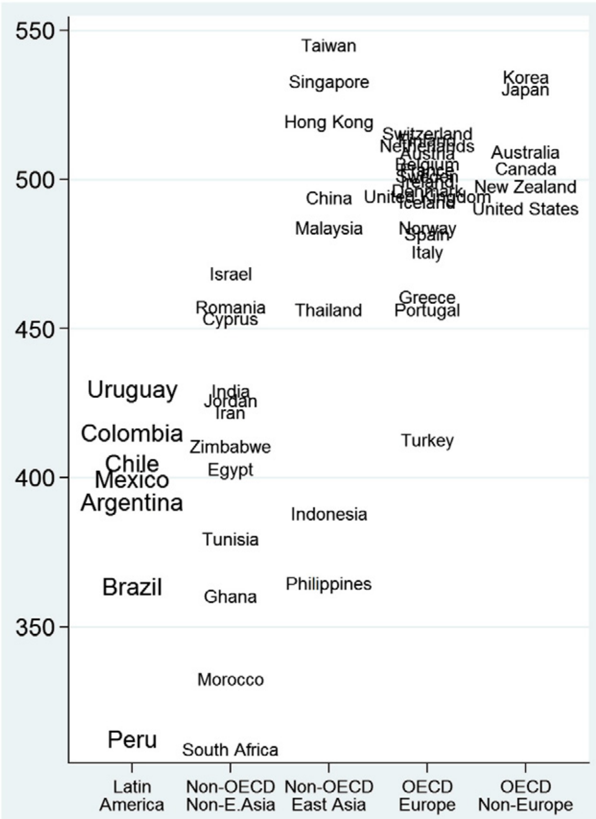

Fig. 2. Latin American performance on international student achievement tests. Simple average of mathematics and science scores over all international tests in 1964–2003, using the re-scaled data by Hanushek and Woessmann (forthcoming) that puts performance at different international tests on a common scale.

mean. Peru, the worst-performing country in Latin America, is nearly two standard deviations below the OECD mean (see also column (5) of Appendix A Table A1 for the Latin American data).

Nonetheless, such a comparison of the performance of those in school will even understate the true gap in average educational achievement between full cohorts. Enrollment in secondary school has not been universal in Latin American countries, leading to more selective test taking in these countries compared to most others in Fig. 2.[16] Assuming that those children who dropped out of school before ninth grade did not reach functional literacy, and taking a test-score performance of one standard deviation below the OECD mean (400 points on the PISA score) as depicting a basic level of functional literacy in mathematics and science, Hanushek and Woessmann (2008) provide a rough measure of the share of a cohort who really reach basic literacy. Less than 5% of the tested students fall below this threshold of basic literacy in developed countries such as Japan, the Netherlands, Korea, Taiwan, and Finland. But, *of those who stayed in school until age 15*, as many as 82% in Peru and 66% in Brazil do not reach such a level of basic literacy. Combined with information on the educational attainment of 15-to-19-year-olds, this means that in Brazil and Peru, the share of recent cohorts that can be termed functionally literate is as small as 8% and 12%, respectively — a number smaller only in Ghana and South Africa among the countries with available data. The remaining roughly 90% of the population in Brazil and Peru have to be viewed as illiterate — because they never enrolled in school, dropped out of school at the primary or early secondary level, or completed lower secondary education without even a basic level of skills. In Colombia, the share of functionally literates in a cohort in their late teens is greater but still only 30%.

The bottom line of the performance of Latin American countries on the worldwide tests is truly dismal: The average educational achievement of Latin American students is consistently near or at the bottom of the international distribution, and only a very small fraction of each young cohort reaches a level of even the most basic skills by international standards.

3.2. Regional achievement tests in Latin America

The poor performance of Latin American countries on the worldwide tests poses a severe problem for the accuracy of intra-regional analyses of educational achievement. The international tests that are designed primarily for developed countries (who support the testing in general) can accurately place student performance near the OECD mean but are thin in questions that would allow discriminating among performance in the tails of the distribution. As a result, the worldwide tests may be unable to distinguish reliably among varying levels of learning in the region of Latin American students. At the very least, the differences recorded among Latin American countries undoubtedly contain considerable noise, even though several thousand students in each country take the tests.

The limitations of worldwide tests in discriminating at the level of Latin American performance leads us to turn to two regional achievement tests specifically designed for the Latin American countries. Starting in the 1990s and aided by UNESCO, Latin American countries developed tests of math and reading skills that could be applied across the region. In 1997, the Latin American Laboratory for the Assessment of Quality in Education – Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación (LLECE) – carried out the "First International Comparative Study in Language, Mathematics, and Associated Factors in the Third and Fourth Grades of Primary Education" (Primer Estudio Internacional Comparativo) specifically designed to test educational achievement in Latin American countries (see Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 1998, 2001, 2002 for details). For ease of reference, we will refer to this study as "LLECE" throughout this report. LLECE provides data on educational performance for nine Latin American countries that also have internationally comparable GDP data.

LLECE tested the performance in math and reading of representative samples of students in each participating country in primary schools. The study released country medians in each grade and subject; in our analyses, we use the performance of the older (fourth-grade) students (see column (6) of Appendix A Table A1).[17] The LLECE scores are standardized to have an international mean of 250 test-score points and a standard deviation of 50 among participating countries. The Median math performance ranges from 226 in Venezuela to 269 in Argentina and Brazil, and the median reading performance from 233 in Bolivia to 286 in Chile. In other words, student performance across countries differs by around one standard deviation on the tests — a huge within-region variation.

In 2006, the Latin American bureau of the UNESCO also conducted the "Second Regional Comparative and Explanatory Study" (Segundo Estudio Regional Comparativo Explicativo, or SERCE) designed for Latin American countries (see Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2005, 2008a, 2008b). It covers 13 countries usable in our growth analyses. Combining the LLECE and SERCE studies, a total of 16 Latin American countries[18] – all Latin American countries with

[16] See Hanushek and Woessmann (2011b) for a sensitivity analysis of the growth regressions to differences in the extent to which the tests cover the full cohort of children in a country.

[17] Scaling is based on a Rasch model that allows for differences in question difficulty. Results of growth analyses that use third-grade scores are similar to those reported below.

[18] Bolivia, Honduras, and Venezuela participated in LLECE but not in SERCE, while Costa Rica, Ecuador, El Salvador, Guatemala, Panama, Peru, and Uruguay participated only in SERCE. Six countries (Argentina, Brazil, Chile, Colombia, Mexico, and Paraguay) participated in both tests.

populations greater than one million and without communist background – can be used in our regional growth analyses.[19]

SERCE tested the performance in math and reading of representative samples of students in third and sixth grades, reporting country medians in each grade and subject.[20] In our analyses, we again use the performance of the older (sixth-grade) students (see column (7) of Appendix A Table A1). The SERCE scores are standardized to have an international mean of 500 test-score points and a standard deviation of 100 among participating countries. Across the 13 countries, median performance (averaged across math and reading) ranges from about 454 in Ecuador and Guatemala to 560 in Uruguay, again revealing a within-regional difference of median performance of more than one standard deviation.

3.3. Splicing the regional tests into the worldwide tests

In order to place the whole Latin American region in the worldwide analysis, it is necessary to place the regional tests on the scale of the worldwide tests. As is apparent from Fig. 2, the performance of the seven Latin American countries that ever participated in a worldwide test is very far down on the worldwide tests such as TIMSS and PISA. It thus seems questionable whether the variation in performance on these tests across individuals and schools in each country, as well as across countries in the region, is informative at all. Across the five Latin American countries that participated both in LLECE and in some worldwide achievement tests, there is no significant correlation between the LLECE score and the score on the global tests. However, the range of the average international test scores across these five countries is 364 to 415 points. Remember that only the lowest 16% of students in OECD countries perform below 400. Our new achievement information using tests designed for the region suggests that it might be possible to improve upon the information on Latin American performance on the global scale by using the more reliable information about intra-regional variations in performance. This also allows us to expand the sample of Latin American countries used in the worldwide growth analysis.

Splicing the regional tests into the world picture involves a number of steps. As a first step, we combine the two regional tests on a common scale using the sample of countries taking both assessments. Specifically, using the average of math and reading performance of the older cohort in both LLECE and SERCE, we first standardize both tests to have mean zero and standard deviation one among the six countries participating in both tests. The combined score on the two regional tests is then given by the simple mean of a country's performance on these two re-scaled test metrics.

Second, we presume that the regional mean observed for the Latin American participants on the worldwide tests provides a reasonable scale for the level of the regional performance on the global scale. Therefore, we re-scale the mean of the combined regional test so that the seven Latin American countries that also participated in the worldwide tests have the same average performance that they have on the Hanushek and Woessmann (forthcoming) combined scale of the worldwide tests (see above). Third, we take the view that the within-region placement of individual Latin American countries found in the worldwide tests is not reliable. This leads us to take the individual country performance in Latin America from the regional rather than the worldwide tests (even for the seven countries participating in the worldwide tests). Finally, splicing of the regional Latin American tests into the worldwide test metric requires an assumption about the size of the within-regional variation on the worldwide scale. Here, we again take the broad metric from the global scale and re-scale the combined regional test so that the seven Latin American countries that also participated in the worldwide tests have the same cross-country standard deviation that they have on the worldwide tests. This method effectively superimposes the distributional information from the international tests onto the Latin American regional tests, but uses the regional test information to more accurately place the individual Latin American countries within the region. The result is an expanded sample that includes all large mainland Latin American countries.

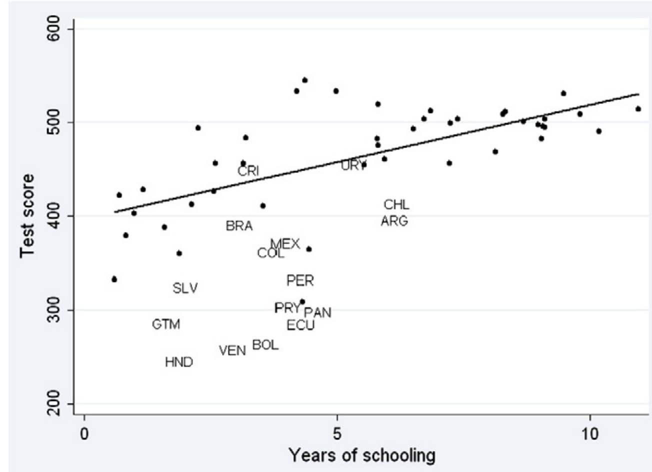

Fig. 3. Years of schooling and educational achievement in Latin America and the world. Scatter plot of average years of schooling in 1960 against average scores on international student achievement tests (extended with regional test measures). Latin American countries are labeled by acronyms, and non-Latin American countries by dots. Regression line relates to non-Latin American countries only. Country codes: Argentina (ARG), Bolivia (BOL), Brazil (BRA), Chile (CHL), Colombia (COL), Costa Rica (CRI), Ecuador (ECU), Guatemala (GTM), Honduras (HND), Mexico (MEX), Panama (PAN), Peru (PER), Paraguay (PRY), El Salvador (SLV), Uruguay (URY), Venezuela (VEN).

The ensuing regional-test-based performance measure of the Latin American countries, expressed on the Hanushek and Woessmann (forthcoming) combined scale of the worldwide tests, is shown in column (8) of Appendix A Table A1.

We have emphasized the measures of educational achievement, but clearly achievement will be related to years of schooling. This combined test metric allows us to compare the schooling relationship for Latin American countries with the rest of the world. The scatter plot in Fig. 3 reveals one of the central messages of this analysis: In virtually all Latin American countries, the average student seems to get much less learning, as depicted in the test scores, for each year of schooling than the average student in the rest of the world. This is the crucial element in our resolution of the Latin American growth puzzle.

Note that considering differences in educational achievement in Latin America, even if they were to arise solely from schooling quality, is not inconsistent with various estimates of high rates of return to years of schooling in Latin America.[21] Under common assumptions, most importantly that the cost of schooling is just the foregone earnings from being out of the labor market, the rate of return to additional schooling is simply the proportionate increase in earnings with an

[19] Both tests included Cuba and the Dominican Republic, and SERCE also included Nicaragua. Nicaragua and Cuba are excluded because of their history of nonmarket economies, although Cuban students scored dramatically higher than students in the Latin American countries. The Dominican Republic was excluded as the sole remaining Caribbean country, which turns out to be a strong outlier if included in the growth analysis (see Hanushek and Woessmann (2009)).

[20] SERCE also tested a much more limited sample of countries in science in sixth grade.

[21] In their survey, Psacharopoulos and Patrinos (2004) report an average Mincer estimate of the return to years of schooling of 12.0% for Latin America, which is the highest across world regions (and roughly similar to Sub-Saharan Africa). A possible explanation for the high Mincer returns in Latin America is likely to be that skills are relatively scarce, thereby increasing the market return to schooling.

Table 2
Latin America in worldwide growth regressions: Worldwide test measures.

	(1)	(2)	(3)	(4)[a]	(5)[a]
Latin America	−1.328**	−1.379***	−0.080	−0.043	−0.491
	(2.47)	(3.00)	(0.23)	(0.14)	(0.48)
Test score			1.954***	2.140***	2.152***
			(7.72)	(9.13)	(9.04)
Test score × Latin America					−0.728
					(0.46)
Initial years of schooling		0.349***	0.023	−0.013	−0.014
		(3.22)	(0.28)	(0.17)	(0.18)
Initial physical capital per capita		0.001	0.014	0.012	0.009
		(0.02)	(0.49)	(0.46)	(0.34)
Initial GDP per capita		−0.380***	−0.332***	−0.302***	−0.294***
		(3.13)	(4.14)	(4.15)	(3.92)
Constant	3.089***	3.082***	4.289***	4.351***	4.351***
	(15.34)	(8.32)	(14.80)	(16.71)	(16.56)
N	50	50	50	50	50
R^2	0.113	0.403	0.747	0.794	0.795
Adj. R^2	0.094	0.350	0.718	0.770	0.766
F (Latin America and interaction)					0.11
Prob > F					(0.893)

Dependent variable: Average annual growth rate in GDP per capita, 1960–2000. *t*-Statistics are in parentheses. See Table 1 for a list of the six world regions and the countries contained in each, as well as data sources. In the country-level analyses, all variables are mean-subtracted.
[a] Test scores measured in exponential terms.
** Statistical significance at 5%.
*** Statistical significance at 1%.

additional year spent in school.[22] While overall earnings at all levels of schooling may be low in a country with low quality schools, the typical earnings gradient with respect to years of schooling would be unaffected by the average quality of a country's schools (see Appendix B). Even with measures of differences in educational achievement at the individual level – which can be the result of both school and nonschool factors – the typical estimation of rates of return to schooling is unaffected.[23]

4. Schooling, achievement, and Latin American growth

4.1. Analyzing Latin American growth with worldwide achievement measures

Can the low levels of educational achievement of Latin American countries account for their poor record of economic growth? To answer this question, we start with the sample of 50 countries that have ever participated in a worldwide student achievement test and have internationally comparable data on GDP growth. (This sample, as noted, includes just seven Latin American countries).

Throughout our analysis, we take the framework of a macroeconomic production function with human capital and physical capital as input factors. The basic specification of the growth model in Eq. (1) regresses the average annual growth in GDP per capita in 1960–2000 on measures of human capital, the initial stock of physical capital per capita,[24] and the initial level of GDP per capita (see Table 1 for data sources).

The “Latin American growth puzzle” motivating our analysis is evident from the first two columns of Table 2. Economic growth in 1960–2000 was significantly lower in Latin America than in the rest of the world. There is a significant negative bivariate correlation between growth and an indicator for Latin American countries which is hardly affected by controlling for initial income, initial physical capital, and initial education in the form of average years of schooling. Over this period, Latin American countries on average grew by 1.3 percentage points per year less than the rest of the sampled countries, even after allowing for starting income levels.

This puzzle is resolved once differences in educational achievement, our more encompassing measure of human capital, are taken into account. Column (3) of Table 2 shows that, conditional on educational achievement, the Latin American countries do not show significantly lower long-run growth than the rest of the world. In this basic specification, test scores that are larger by one standard deviation (measured at the student level across all OECD countries in PISA), or 100 PISA points, are associated with an average annual growth rate in GDP per capita that is about two (1.95) percentage points higher over the whole 40-year period.[25] The share of growth variance explained by the model increases from 0.35 to 0.72 when the achievement measure is included. Furthermore, the coefficient on years of schooling becomes insignificant and drops to close to zero. Conditional on test scores – that is, conditional on what students know – additional years of schooling does not show a significant association with economic growth. In other words, school attainment, which is positively correlated with test scores (see Fig. 3), only matters for economic growth inasmuch as it is related to better skills of the students.

[22] As Heckman et al. (2006) discuss in detail, interpreting the gradient on schooling in a Mincer earnings function as a rate of return requires a number of stringent assumptions that typically are not met, at least in the U.S.

[23] See the summary of U.S. estimates of the impact of cognitive skills across studies in Hanushek (2011) and of differences in returns to schooling across countries in Hanushek and Zhang (2009).

[24] Physical capital stocks are calculated by the perpetual inventory approach from income and investment data in the Penn World Tables in the standard way (taken from Jamison et al. (2007)).

[25] Hanushek and Woessmann (2008, forthcoming) provide additional sensitivity analyses on this. They show that the association between educational achievement and economic growth proves extremely robust to including obvious additional control variables such as economic institutions, geography, and fertility, to different sets of sub-samples of countries, to restricting the test-score measure to specific periods, subjects, or age groups, and to applying the analysis to immigrants from different countries on the same U.S. labor market. Much discussion of Latin America emphasizes different economic institutions, frequently referring to the fact that many Latin American countries have followed import substitution policies over significant parts of the period. Considering trade policies does not compromise the importance of skills: Both openness and educational achievement enter significantly positive in a joint country-level specification. In addition, Hanushek and Woessmann (2008) do show that the effect of educational achievement interacts positively with the openness of the economy: While skills have positive effects even in closed economies, they are significantly bigger in open economies. Hanushek and Woessmann (forthcoming) also show that several instrumental variables and differences-in-differences specifications support a causal interpretation of the association; we will return to this topic below.

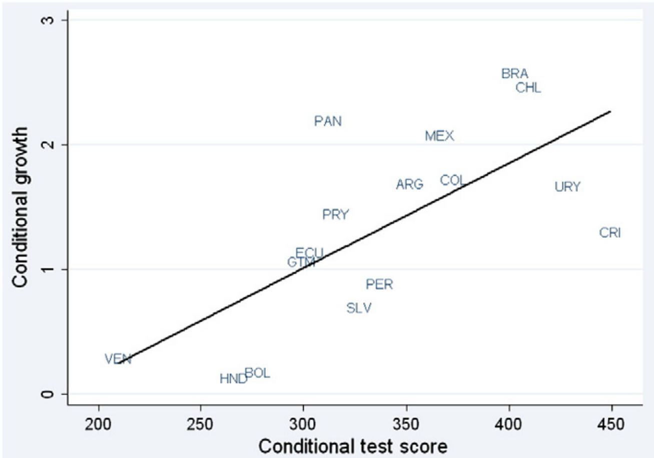

Fig. 4. Educational achievement and economic growth in Latin America. Added-variable plot of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960–2000 on the initial level of real GDP per capita in 1960 and average scores on Latin American student achievement tests (mean of the unconditional variables added to each axis). See Fig. 3 for country codes.

The previous specification assumes that there is a linear relationship between growth and the international test scores, whose scaling is based on a procedure that yields a normal shape of the test-score distribution of individual students across the OECD countries. The shape of the scaling of test performance is ultimately arbitrary, and casual inspection of the conditional association between growth and the test score measure indicates some curvilinear pattern. Therefore, the specification of column (4) enters our test score measure in exponential form. The results reveal that such a specification provides a considerably better fit to the data, while the Latin American dummy continues to be insignificant.[26] The exponential form of the test score–growth relation ultimately reflects arbitrary scaling choices for the underlying achievement tests, but it means that the same absolute increase in test scores – when measured on the standard international PISA scale – is related to larger absolute increases in economic growth rates at higher levels than at lower levels of the test metric. For example, according to this specification, a 10-point increase from 400 to 410 on the PISA scale is related to a 0.13 percentage point increase in the economic growth rate, whereas a 10-point increase from 500 to 510 is related to 0.35 percentage points higher growth.

In the final column, we allow the test-score effect to differ between the seven Latin American countries and the rest of the world. The interaction term, either individually or jointly with the Latin American dummy, is not statistically significant, suggesting that the relationship between educational achievement and growth is not different in Latin America compared to the rest of the world. (This result holds if interactions of the Latin American dummy with years of schooling, physical capital, and initial GDP are included, which are strongly insignificant). However, this analysis is restricted by the limited number of Latin American countries participating in the worldwide tests, as well as the latter's relatively low informational content for Latin American countries. We therefore turn to our extension to the regional Latin American tests next.

4.2. Intra-regional analyses of educational achievement and growth in Latin America

We start with a descriptive look at the relationship between educational achievement and economic growth within the sample of Latin American countries. As indicated, there is considerable variation in both skill levels and economic performance across the 16 Latin American countries in our analysis (see Appendix A Table A1). For example, even on the worldwide tests, more than a whole standard deviation distinguishes the average achievement in Peru (312) from Uruguay (430). Average years of schooling in 1960 was a mere 2 years in Guatemala, Honduras, and El Salvador, as opposed to more than 6 years in Chile and Argentina. Likewise, economic performance shows substantial variation among the Latin American countries. The level of GDP per capita in 1960 ranges from below $2000 in Honduras and Ecuador to more than $7000 in Venezuela and Argentina — close to the mean of European countries. The growth experience between 1960 and 2000 ranges from negative in Venezuela to almost 3% per year in Brazil. As a consequence of the differing initial income levels and growth experiences, GDP per capita in 2000 ranges from about $2000 in Honduras to more than $10,000 in Argentina.

As summarized by Fig. 4, this difference in growth performance within the Latin American region is closely related to educational achievement. Using the LLECE and SERCE skill measures confirms that Latin American countries that have higher educational achievement have experienced faster economic growth over the long run.[27] This pattern clearly illustrates the potential of educational achievement in explaining growth differences within the region. There are, of course, concerns with such an analysis. While the LLECE and SERCE tests provide a reliable measure of performance differences among Latin American countries, they refer to performance at relatively early grades, necessitating an assumption that such early performance is a reasonable index of performance throughout the schooling system. In addition, they were administered toward the end of the observed growth period, necessitating an assumption that cognitive performance differences have been relatively stable over the prior decades.

4.3. Latin America in the World: combining regional and worldwide tests

Using the additional regional test score information, we can return to the question of how the Latin American evidence fits into the worldwide analysis. The noisiness of the worldwide test measures at low levels means that estimates of the effect of worldwide test scores for Latin American countries are likely to suffer from measurement error. Therefore, we now employ the new test score dataset that splices the regional Latin American tests into the global test score measure in our cross-country growth regressions.

Results are reported in Table 3, which retains the exponential form of the test score measure derived above. The first four columns mirror the analyses on just the worldwide tests from Table 2: Also in the expanded 59-country sample that includes the entire set of 16 Latin American countries, a Latin American dummy is strongly negative in models without educational achievement but becomes very small and statistically insignificant once achievement is controlled for. The point estimate on educational achievement is slightly smaller than in the previous analyses. Column (5), which reverts to the original 50-country sample except with the new test information, shows that this is not due to the additional Latin American countries in the expanded sample but instead due to the substitution of the regional for the global test data for the Latin American countries. Again, there is no significant interaction between test scores and the indicator for Latin American countries.

The fact that the test score–growth nexus does not differ between Latin America and the rest of the world is clearly visible in Fig. 5: Together with the other countries, the Latin American countries fall around a straight line that captures the conditional association between educational achievement and economic growth. Given the

[26] We describe this exponential relationship in terms of test scaling, but there are alternative interpretations. First, larger measurement error in the tests at the lower achievement levels compared to higher levels could flatten out the relationship at the bottom. Second, it could be that at very low levels there are simply lower returns to skills. We cannot distinguish among the alternative explanations.

[27] Although analyses within the limited sample of 16 Latin American countries are limited by their degrees of freedom, regression analyses reported in Hanushek and Woessmann (2009) show that educational achievement as measured by the LLECE and SERCE data enters significantly in within-regional growth regressions, substantially increases their explanatory power, and is robust to controlling for differences in openness and property rights security.

Table 3
Latin America in worldwide growth regressions: Extension with regional test measures.

	(1)	(2)	(3)	(4)	(5)[a]	(6)
Latin America	−1.744***	−1.711***	−0.351	−0.590	−0.559	
	(4.70)	(5.22)	(1.36)	(1.28)	(0.87)	
Test score (exponential)			1.596***	1.627***	1.706***	1.739***
			(8.93)	(8.73)	(9.05)	(11.94)
Test score (exp.) × Latin America				−0.330	−0.837	
				(0.63)	(0.87)	
Initial years of schooling		0.347***	0.041	0.039	−0.013	0.025
		(3.75)	(0.61)	(0.57)	(0.17)	(0.37)
Initial physical capital per capita		0.002	0.013	0.012	0.007	0.018
		(0.04)	(0.53)	(0.47)	(0.26)	(0.74)
Initial GDP per capita		−0.372***	−0.332***	−0.328***	−0.290***	−0.341***
		(3.41)	(4.77)	(4.66)	(3.91)	(4.87)
Constant	3.089***	3.044***	3.928***	3.927***	4.057***	3.900***
	(15.99)	(8.93)	(16.45)	(16.35)	(16.42)	(16.26)
N	59	59	59	59	50	59
R^2	0.280	0.524	0.810	0.811	0.795	0.803
Adj. R^2	0.267	0.489	0.792	0.790	0.766	0.789
F (Latin America and interaction)				1.11	0.40	
Prob > *F*				(0.337)	(0.671)	

Dependent variable: Average annual growth rate in GDP per capita, 1960–2000. *t*-Statistics are in parentheses. See Table A1 for data sources.
[a] Country sample as in Table 2.
*** Statistical significance at 1%.

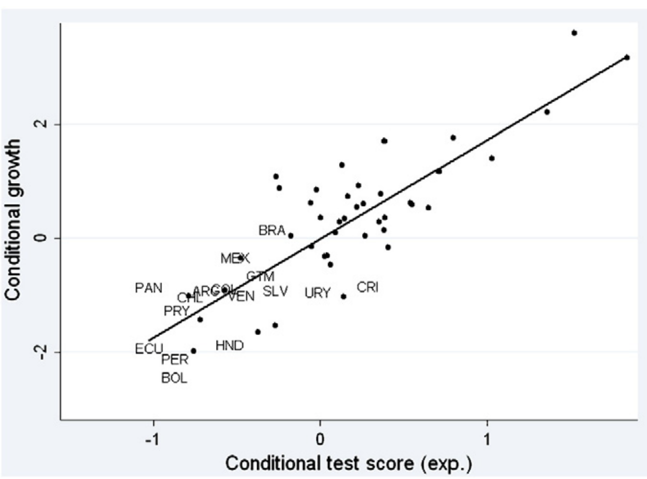

Fig. 5. Educational achievement and economic growth in Latin America and the world. Added-variable plot of a regression of the average annual rate of growth (in percent) of real GDP per capita in 1960–2000 on average (exponential) scores on international student achievement tests, the initial level of real GDP per capita in 1960, physical capital per capita in 1960, and average years of schooling in 1960. Latin American countries are labeled by acronyms (see Fig. 3 for country codes), non-Latin American countries by dots. Based on Table 3, column (6).

exponential specification, the same absolute improvements in terms of the international PISA scale appear to be related to stronger growth improvements at higher compared to lower test score levels, although this could be just an artifact of the particular scale. An interpretation that skill improvements have lower growth impact at very low skill levels would carry substantial arbitrariness, given that the usual scaling of tests is anchored only in the underlying assumption that skill distributions are roughly normal. But our results refute the possibility that this curvilinear pattern is specific to Latin America.[28]

5. Instrumental-variable estimates

In terms of econometric identification, the extent to which associations found in cross-country growth regressions can be interpreted as causal effects has long been the subject of controversy. Beginning with the analysis of Levine and Renelt (1992), evidence of the sensitivity of results to model specification has been plentiful. In terms of years of schooling, Bils and Klenow (2000) provide convincing evidence of the endogeneity of school attainment in growth models.

In prior work, Hanushek and Woessmann (forthcoming) use several econometric techniques to address the issue of causality between educational achievement and growth.[29] While not conclusive, the combined approaches provide general support for a causal interpretation of the consistent relationship between educational achievement and growth. First, a differences-in-differences approach focuses on the earnings of immigrants to the U.S. and finds that the international test scores for their home country significantly explain U.S. earnings for those educated in their home country but not for those educated in the U.S. Second, in a model that eliminates the levels of achievement and growth, changes in test scores over time are found to be systematically correlated with changes in growth rates over time. Third, instrumenting by institutional features of school systems does not change the growth results in samples that are mostly restricted to OECD countries because of limited data availability. Finally, extensive sensitivity analysis shows that the results are insensitive to the sample of countries, to the specific tests employed, to estimation within separate regions, or to restricting the test score measure to international tests conducted before the observed growth period. Another analysis reducing endogeneity concerns stemming from country-specific effects is the cross-sector analysis of Ciccone and Papaioannou (2009), who employ country and industry fixed effects to show that countries with higher educational achievement have higher growth in skill-intensive industries.

Here, we propose three new instrumental-variable (IV) strategies that complement the prior analyses but also deal with additional concerns in the study of Latin America. They are designed to provide information about the extent of possible reverse causation and omitted country variables in the specific comparisons between countries in Latin America and the rest of the world. Each of the three IV models addresses a particular concern with the OLS models, and their joint consideration provides further support that common worries about limitations of the OLS models do not appear to be driving our results.

[28] The consistency between our expanded Latin America sample and other low scoring countries does indicate that additional test errors due to the splicing of worldwide and regional tests do not appear to be driving the observed growth patterns of Latin American countries.

[29] Aghion et al. (2005) use within-country variation in the United States to address causality issues in the relationship between investments in college education and growth.

Table 4
Instrumental-variable models.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Second stage:								
Test score (exponential)	1.847***	1.812***	2.568***	2.676***	1.938***	1.931***	1.793***	1.693**
	(7.65)	(6.12)	(3.25)	(2.79)	(8.10)	(6.90)	(3.27)	(2.23)
Latin America		−0.167		0.546		−0.024		−0.242
		(0.48)		(0.83)		(0.06)		(0.39)
Catholic share in 1970			0.378	0.033	−0.055	−0.042		
			(0.62)	(0.08)	(0.20)	(0.13)		
Years of schooling in 1960			−0.162	−0.166			−0.144	−0.130
			(0.86)	(0.83)			(1.02)	(0.79)
Physical capital p.c. in 1960	0.017	0.015	0.022	0.021	0.013	0.013	−0.002	−0.003
	(0.72)	(0.62)	(0.69)	(0.66)	(0.53)	(0.54)	(0.10)	(0.14)
GDP per capita in 1960	−0.333***	−0.327***	−0.314***	−0.307***	−0.332***	−0.332***	−0.221***	−0.223***
	(5.33)	(5.07)	(3.65)	(3.41)	(4.87)	(4.92)	(2.83)	(2.91)
Constant	4.003***	4.047***	4.539***	4.523***	4.064***	4.065***	4.586***	4.595***
	(21.89)	(24.61)	(7.65)	(7.66)	(23.49)	(23.75)	(8.34)	(8.89)
First stage:								
Years of schooling in 1960	0.233***	0.192***	0.236***	0.207***	0.236***	0.207***	0.226***	0.200***
	(4.30)	(4.27)	(4.88)	(4.56)	(4.88)	(4.56)	(3.64)	(3.85)
Catholic share in 1900			2.312*	1.965*	2.314*	1.965*		
			(1.84)	(1.70)	(1.84)	(1.70)		
Relative teacher salary							0.213**	0.152*
							(2.05)	(1.73)
Latin America		−0.852***		−0.759***		−0.759***		−0.694***
		(5.35)		(3.31)		(3.31)		(3.72)
Catholic share in 1970			−3.107**	−2.162*	−3.107**	−2.162*		
			(2.35)	(1.74)	(2.35)	(1.74)		
Physical capital p.c. in 1960	0.009	−0.007	−0.012	−0.010	−0.012	−0.010	0.003	−0.001
	(0.38)	(0.38)	(0.57)	(0.51)	(0.57)	(0.51)	(0.14)	(0.04)
GDP per capita in 1960	−0.070	−0.025	−0.020	−0.026	−0.020	−0.026	−0.028	−0.026
	(1.09)	(0.47)	(0.34)	(0.49)	(0.34)	(0.49)	(0.42)	(0.46)
Constant	−0.949***	−0.554***	−0.662***	−0.540***	−0.662***	−0.540***	−1.492***	1.039***
	(5.23)	(3.35)	(3.78)	(3.28)	(3.78)	(3.28)	(4.10)	(3.18)
Fuller (1) modification of LIML:								
Test score (exponential)	1.842***	1.801***	2.350***	2.383***	1.938***	1.931***	1.738***	1.580**
	(7.84)	(6.26)	(3.68)	(3.20)	(8.10)	(6.90)	(3.57)	(2.46)
N	59	59	59	59	59	59	34	34
Centered R^2 (second stage)	0.801	0.805	0.696	0.679	0.796	0.797	0.755	0.768
Centered R^2 (first stage)	0.426	0.625	0.571	0.645	0.571	0.645	0.561	0.706
Instrument F-stat. in 1st stage	18.51	18.20	3.40	2.90	12.45	10.78	4.20	2.98
Sargan statistic					1.098	1.099		
p-Value					(0.295)	(0.294)		
Durbin–Wu–Hausman χ^2 test	0.151	0.406	2.182	2.265	1.358	1.509	0.247	0.345
p-Value	(0.698)	(0.524)	(0.140)	(0.132)	(0.244)	(0.219)	(0.619)	(0.556)

Dependent variable in the second stage: Average annual growth rate in GDP per capita, 1960–2000. Dependent variable in the first stage: Test score (exponential). t-Statistics are in parentheses. See text for details.
* Statistical significance at 10%.
** Statistical significance at 5%.
*** Statistical significance at 1%.

We begin with the concern that our test scores are observed toward the end of our growth period, a particularly acute potential problem given that we spliced in the end of period tests for Latin America. A simple reverse causality notion suggests that improved economic outcomes could lead to schooling investments that have a positive effect on students' cognitive achievement. The use of average years of schooling in the adult population in 1960 as an instrument for our measure of educational achievement rules out such reverse causality (although not all endogeneity fears). Such an IV specification uses only a part of the variation in the achievement measure that is related to a measure observed in 1960, before the growth period. It thus rules out the possibility that the association between test scores and growth depicted in our growth regressions just captures a reverse effect of improved economic conditions on test score performance. Conceptually, years of schooling in 1960 can be justified as capturing the education level of the parent and teacher generation (see Bils and Klenow, 2000), which – after conditioning on per-capita income in 1960 – may reflect an ingredient of the "education production function" that is exogenous to the growth model. The identifying assumptions of this IV model are that initial years of schooling is correlated with cognitive skills but does not belong in the second-stage (growth) equation once educational achievement is included. The two main features of our analysis in the last section are that years of schooling is clearly (although far from perfectly) related to achievement and that they are not associated with economic growth beyond their association with achievement.[30] These features make years of schooling in 1960 a potentially attractive instrument.

The first column of Table 4 shows the result of this IV specification. The first-stage model shows that years of schooling in 1960 is a strong instrument for our test score measure. The F-statistic of the instrument in the first stage is 18.5. In this (and all of the subsequent first-stage models), per-capita GDP in 1960 and per-capital physical capital in 1960 are not significant predictors of the test scores, suggesting that achievement is actually unrelated to these measures of initial economic performance of the country. The second-stage

[30] While it is not possible to test the validity of an instrument directly in the just-identified case, we supply information on this from the over-identified models below.

model confirms a strong and highly significant effect of the instrumented test score measure on economic growth. The point estimate of the test score effect is 1.8, close to the OLS estimate of the same specification of 1.7 (column (6) of Table 3).

Column (2) adds an indicator for Latin American countries to this model. In confirmation of our previous results, the first stage underscores that Latin American countries have significantly lower educational achievement (conditional on observed years of schooling and measured determinants of achievement), whereas the second stage confirms that the growth experience of Latin American countries is no different from the rest of the world once their low level of educational achievement is taken into account.

While addressing basic concerns of reverse causality, using years of schooling in 1960 as an instrument of course does not rule out all concerns of bias from unobserved country characteristics, which may be associated with schooling in 1960 (past its correlation with per-capita income in 1960). To deal with these further concerns, we incorporate ideas deriving from two very recent studies on factors influencing school quality differences across countries to deal further with endogeneity issues in our growth modeling.

Our second IV strategy employs variations in the productivity of school systems that are rooted very deep in history. Specifically, we investigate whether competitive forces set in the historical development of different countries' schooling systems provide variations in educational achievement that are related to current growth. West and Woessmann (2010) show that the resistance of the Catholic Church in the 19th century against the emerging state schooling system created private school sectors in many countries that remain until today, with higher Catholic shares in a country's population in 1900 significantly increasing the share of privately operated schools today. This increased competition in a country's school system in turn raises educational achievement. While internationally consistent data on private school shares are not available for our purposes, we can use the "reduced form" of their model to obtain arguably exogenous variation in test scores. That is, we use the share of Catholics in a country in 1900 as an instrument for our test score measure. Barrett et al. (2001) provide data on the Catholic population share of all countries in our sample for 1900, as well as for 1970.

Of course, religious affiliation may itself be related to educational achievement and to economic growth. Consequently, this IV specification conditions on Catholic shares in the modern population; in the empirical analysis, we use the 1970 value available in the Barrett et al. (2001) data.[31] Thus, the identifying assumption of this IV model is that holding constant the effects of modern religious affiliation, historical religious affiliation is not otherwise related to modern growth, apart from the indirect effect through its bearing on competition and thus productivity in the modern school system.

The IV estimation reported in column (3) of Table 4 provides new support for our underlying model of educational achievement and growth. The share of the Catholic population in 1900 is indeed positively related to our measure of educational achievement in the first stage. At the same time, the share of the Catholic population in 1970 is negatively related to achievement, which is in line with the previous literature.[32] At the same time, the 1970 Catholic share does not enter the second-stage growth model significantly. But, importantly, the variation in test scores that is related to historical Catholic shares does indeed have a significant positive effect on economic growth. The point estimate is in fact even larger than in the OLS model, although the Durbin–Wu–Hausman test does not reject the null that the estimates from the OLS and IV models are the same. Adding the Latin America indicator in column (4) does not change the main result.[33]

However, this estimate may be subject to a weak instrument problem, as the *F*-statistic of the instrument in the first stage is only 3.4. As a first response, the bottom part of Table 4 also reports results on the Fuller (1977) modification of the limited information maximum likelihood (LIML) estimator, which is more robust than 2SLS in the presence of weak instruments and performs relatively well in the simulation exercise of Hahn et al. (2004). While point estimates of the LIML estimator are again closer to the OLS estimate, all LIML estimates in Table 4 remain highly significant. In columns (5) and (6), we combine the historical Catholic share and years of schooling in 1960 as instruments.[34] The level of joint significance of the two instruments in the first stage is substantially higher (*F*-statistic of 12.5), easing concerns about weak instruments. Furthermore, the Sargan test does not reject the overidentifying restrictions implied in this model, suggesting that as long as one of the two instruments is accepted as valid, the other one is, too. Finally, this result is confirmed when using robust methods for inference in response to possible biases from weak instruments. In particular, the 95-percent confidence band of the conditional likelihood ratio test proposed by Andrews et al. (2007) and Moreira (2003) ranges from 1.450 to 2.596 for the column (5) estimate (1.358 to 2.751 for the column (6) estimate), which is significantly different from zero at the 1 percent level. This suggests that the presented results are not driven by weak instrument problems.

Our third IV specification aims at identifying a variation that stems explicitly from the education system in order to reduce concerns that the OLS estimates capture omitted variables related to the overall economy. A key result of the economics of education literature in recent years is that the quality of the teaching force is a leading observable determinant of student test scores (see Hanushek and Rivkin, 2010; Rivkin et al., 2005; Rockoff, 2004). In a cross-country perspective, Dolton and Marcenaro-Gutierrez (2011) suggest that relative teacher salaries provide a useful proxy for the overall quality of the teaching force.[35] We therefore use their measure of teacher salary, expressed relative to the per-capita income of each country, as an instrument for our achievement measure.[36] By expressing teacher salary *relative* to the earnings distribution in a country, we focus on the point in the overall "ability" distribution from which a country is likely to draw its population of teachers and avoid just capturing overall income levels that would be correlated with growth.[37] The exclusion restriction of this specification is that relative teacher pay is independent of the error term of the growth equation, over and above its relation with educational achievement. This is likely met for relative teacher pay — in contrast to other possible ingredients of the education production function such as expenditure per student (which could obviously be a function of the economic growth process) and socio-economic background of families (which could co-move with the economic growth process).

[31] Results are fully robust – with even larger point estimate – when, because of this, the growth model is estimated for the 1970–2000 period.

[32] These findings are entirely consistent with prior work on Catholic education. Becker and Woessmann (2009) show that historically Protestant Christians were significantly more educationally advanced than Catholic Christians because of their emphasis on the individual ability to read the Bible. West and Woessmann (2010) provide further evidence that Catholicism is unlikely to have an own positive effect on educational achievement.

[33] These results are fully robust in the sample of 36 countries where Catholic and Protestant Christians together formed the majority of the population in 1900.

[34] Note that the specifications in (3) and (4) include years of schooling in 1960 as a control variable in the second stage instead of using it as an instrument.

[35] The teacher salary data come from surveys conducted by OECD and by UNESCO. Specifically, we use teacher salaries at the top of the experience scale. A downside of these data is that they are observed only at the end of our growth period (in 2003), but joint use with the two historical instruments confirms that the results are not just driven by identifying a variation that occurred during the growth period.

[36] Results are very similar when using the Dolton and Marcenaro-Gutierrez (2011) proxy of the percentile position of teachers' salaries in the earnings distribution, which, however, is missing for three additional countries in our sample.

[37] Similarly, a recent McKinsey analysis (Auguste et al., 2010) highlights relative teacher salaries as an important determinant of why some countries do better than others on the latest international tests.

Even though the estimation sample is reduced because of lack of data on teacher salaries, this IV estimation again confirms the independent role of educational achievement on growth. (The salary data are available for a limited number of 34 countries in our sample only, but this does include six Latin American countries (Argentina, Brazil, Chile, Mexico, Peru, and Uruguay)). As shown in columns (7) and (8) of Table 4, teacher salaries relative to per-capita income are indeed significantly related to our test score measure in the first-stage model. Again, the point estimates on test scores in the second stage are close to the OLS estimates and are statistically highly significant. The instrument is again somewhat weak (*F*-statistic of 4.2), but the LIML estimates also support the result. Furthermore, results are confirmed in models that use relative teacher salaries together with years of schooling (where the joint instruments are much stronger again), or all three instruments together (not shown). Again, Sargan tests do not reject the over-identification restrictions of the models with several instruments.

In sum, the IV results of all three strategies strongly support our basic finding of a significant effect of educational achievement on economic growth. Moreover, the IV estimates leave its order of magnitude essentially unaffected. All three IV models also underscore that Latin American countries have significantly lower quality of schooling (conditional on observed determinants including years of schooling), but that their growth experience is no different from the rest of the world once their low level of educational achievement is taken into account. Reassuringly, over-identification tests corroborate the validity of the three separate IV specifications. While it is impossible to put conclusively to rest all of potential issues of causality in cross-country growth regressions, and while IV identification in a cross-country application will always suffer from limited statistical power, the IV results provide some strongly supportive analysis to our interpretation of the Latin American growth puzzle.

6. Development accounting

We also pursue a different, complementary approach to address the endogeneity concerns of cross-country macro regressions. These concerns arise from the fact that the estimated growth parameters may suffer from reverse causality and omitted variables in the specified model. Development accounting, an alternative approach, does not depend on estimating the growth parameters but rather takes them from the microeconomic literature and derives structure in the macroeconomic analysis by assuming a particular functional form of the macroeconomic production function (see Caselli (2005) and Hsieh and Klenow (2010) for details on the underlying concept).[38] It also does not require the assumption that (unmeasured) total factor productivity is orthogonal to factor inputs, in particular to human capital.

Consider a standard Cobb–Douglass macroeconomic production function that has the well-known per-capita form:

$$y = h\left(\frac{k}{y}\right)^{\frac{\alpha}{1-\alpha}}A. \tag{3}$$

Taking logs on both sides allows the cross-country variation in per-capita output y to be decomposed into three components: A first one that can be attributed to cross-country variation in human capital per capita h, a second one that can be attributed to cross-country variation in physical capital-output ratios k/y, and a third one that is effectively left unattributed to the measured factor inputs (usually attributed to total factor productivity, A). Since we are interested in the extent to which differences in human capital can account for differences in economic output between Latin America and other regions of the world, we will mostly focus on a decomposition of differences between two groups of countries. In the spirit of the decomposition provided in Hall and Jones (1999) (see also Gundlach et al., 2002), the share of income variation between Latin American countries and other countries in a comparison group that can be accounted for by differences in human capital is given by $\frac{\ln(h_{LatAm}/h_{other})}{\ln(y_{LatAm}/y_{other})}$, based on geometric averages of the countries in each group. In addition, we will report results of the "covariance measure" proposed by Klenow and Rodriquez-Clare (1997) which allows decomposing the total variance in log per-capita output in the total 59-country sample into shares attributed to its covariance with the two factor input components — in the case of human capital, $\frac{\mathrm{cov}(\ln(y),\ln(h))}{\mathrm{var}(\ln(y))}$. Here, we are only interested in the share attributed to variation in human capital, for which no additional parameter assumptions are required.[39]

Table 5
Development accounting: Share of variation in GDP per capita attributed to variation in human capital.

	Considering only years of schooling	Considering years of schooling and educational achievement		
		$w=.20$	$w=.15$	$w=.25$
	(1)	(2)	(3)	(4)
Latin America vs. non-Latin America	0.28	0.59	0.51	0.67
Latin America vs. OECD countries	0.28	0.50	0.45	0.56
Latin America vs. East Asia	0.29	0.80	0.68	0.93
Covariance measure (global sample)	0.24	0.39	0.35	0.43

Results of development accounting analyses for GDP per capita in 2000 in the 59-country sample. The parameter w refers to the income return to a standard deviation in test scores in Mincer earnings functions. Decompositions between Latin America and a second group of countries are based on $\frac{\ln(h_{LatAm}/h_{other})}{\ln(y_{LatAm}/y_{other})}$. The covariance measure is based on the decomposition $\frac{\mathrm{cov}(\ln(y),\ln(h))}{\mathrm{var}(\ln(y))}$. See text for details.

To measure human capital, we adopt a variant of the by-now standard Mincer specification that weighs education by its micro labor-market returns. We extend the typical specification from the usual one that considers only years of schooling to one that additionally considers test-score performance:

$$h = e^{rS+wT} \tag{4}$$

where S is years of schooling, T is test scores, and r and w come from the micro literature on Mincer log earnings functions. Specifically, we assume a standard rate of return of $r=.10$ for each year of schooling. (Some variation of this assumed value, moreover, does not make any substantial difference for our pattern of results).

Crucially for our interest in educational achievement, we require a micro estimate of the return to educational achievement on the labor market. Using U.S. data for a representative sample spanning the entire work life from the International Adult Literacy Survey (IALS), Hanushek and Zhang (2009) estimate a standard Mincer equation that adds the IALS test score measure of educational achievement as a regressor to the usual years of schooling. Their estimate indicates that each individual-level standard deviation in test scores is associated with .193 higher earnings.[40] For our calibration, we therefore use a return of $w=.20$ per standard deviation in test scores, but also present versions for lower and higher values.

[38] A parallel growth accounting analysis is precluded by the fact that there is no consistent data on variation in educational achievement over time.

[39] Assuming a production elasticity of physical capital of $\alpha=1/3$, we find results on physical capital that are very similar to the literature: The share of the variation in GDP per capita attributed to differences in the capital–output ratio k/y is calculated as .19 for the variation of Latin America vs. Non-Latin America and as .17 in the covariance measure of the global sample.

[40] Hanushek (2011) reviews available estimates. Several additional U.S. studies using different datasets confirm this general result, although with somewhat lower estimates in the range of .10 to .15, which likely reflects their restriction to samples of workers fairly early in their work career (Lazear, 2003; Mulligan, 1999; Murnane et al., 2000). Chetty et al. (2011) find a return to measured achievement in young workers of 0.18.

Table 5 reports the results on the share of the cross-country variation in GDP per capita in 2000 that is attributable to variation in human capital in various development accounting comparisons. The sample is composed of the 59 countries with test-score and economic data, including all 16 Latin American countries. On average, log per-capita output is 2.2 times as high in non-Latin American countries than in Latin American countries. When human capital is measured only by years of schooling (i.e., $w=0$), 28% of this income variation between non-Latin American countries and Latin American countries is attributed to human capital in the development accounting analysis. However, once differences in educational achievement are included as part of the human capital variation, the share attributed to human capital increases to 59%. That is, based on the microeconomic returns to years of schooling and to test scores observed on the U.S. labor market, a human capital measure that combines these two aspects is able to account for a majority of the lag of income in Latin America relative to the rest of the world.

We also consider two specific groups of countries as comparison groups for Latin America: The OECD countries and the East Asian countries. Half of the income difference between Latin America and the OECD and four-fifths of the income difference between Latin America and East Asia are attributed to differences in human capital in this decomposition. The final row of Table 5 reports results of the covariance measure, calculated for the overall income variance in the 59-country sample (with no specific focus on Latin America). Again, consideration of educational achievement in the human capital measure increases the share of income variance attributed to human capital from 24% to 39%.[41]

The parameterization so far assumes that the microeconomic returns to skills on the labor market capture their full macroeconomic effect. However, the social returns to skills may be higher than the private returns if there are externalities, for example in the spirit of the innovation effects of endogenous growth models (e.g., Aghion and Howitt, 1998; Romer, 1990). This may mean that the macro return to skills is substantially higher than indicated by our parameter value of $w=.20$. On the other hand, some labor market studies suggest that the return to test scores may be somewhat lower. Therefore, columns (3) and (4) show results of our development accounting analyses for parameter values of $w=.15$ and $w=.30$, respectively. If the macroeconomic return to skills lies in this range, the share of variation in GDP per capita between Latin American and non-Latin American countries that can be attributed to human capital differences lies between one half and two thirds.

These development accounting results, which derive structure from assuming a production function and taking well-identified parameters on the income returns to educational achievement from the micro literature, indicate that differences in educational achievement are indeed significant enough to account for major parts of the income differences between Latin America and the rest of the world. The approach eliminates major sources of endogeneity bias (although obviously not all sources) because it is not subject to the kind of bias from omitted variables or reverse causation that has appeared to plague prior cross-country growth regressions. The evidence is consistent with and complementary to our growth regression results above.

7. Conclusions

Economic growth in Latin America over the past half century has been disappointing and puzzling. The region historically has had relatively high levels of school attainment. In 2001 expected school attainment in Latin America was 13 years, compared to 8.6 in South and West Asia, 7.1 in Sub-Saharan Africa, and 9.5 across all developing countries (UNESCO, 2005). Yet these human capital investments have not translated into clear patterns of growth and development.

The growth puzzle is reconciled by consideration of the level of educational achievement across Latin America. The average achievement of Latin American students on international tests is substantially lower than in East Asia and MENA and nearly as dismal as in Sub-Saharan Africa.[42] Despite their relative success in achieving high levels of school attainment, the knowledge that students in Latin America have gained by their mid-teens is staggeringly low. We show that in countries such as Brazil and Peru that provide the necessary data, only about one in ten children of each cohort can be termed functionally literate in their late teens.

Not only does the low level of educational achievement account for the lack of growth of Latin America relative to the other world regions, but it also provides much of the explanation for variations in economic performance within the group of Latin American countries participating in the Latin American LLECE and SERCE tests of student achievement. This conclusion is based, however, on limited measures of relatively recent differences in mathematics and reading at the primary-school level, and it will be important to confirm this with expanded assessments.[43] Instrumental-variable models that restrict the identifying variation in test scores to parts that derive from historical and institutional sources of the quality of schooling, as well as development accounting analyses that take the economic return to differences in skills from micro labor regressions, confirm the importance of educational achievement in understanding the Latin American lag in economic performance.

By ignoring differences in what students actually know, the existing literature very significantly misses the true importance of human capital for economic growth in Latin America. Our results reveal that school attainment is associated with economic growth only insofar as it entails educational achievement — which in Latin America has been the case to a much lesser extent than elsewhere.

We stop short of any considerations of how countries should go about improving their schools, something that is discussed in a growing number of international studies (see the review in Hanushek and Woessmann (2011a)). However, the importance of educational achievement has not been missed by some Latin American countries that have shown strong improvements in recent years. In terms of achievement trends of 15-year-olds over the most recent PISA cycles, Brazil and Chile in particular show considerable improvements (see Hanushek et al. (2012) for detailed international trend estimates). These changes support our basic perspective that it is possible to improve achievement through governmental policies. They also suggest that Latin American countries are not necessarily destined to remain at the bottom of the economic distribution.

[41] The smaller share attributed to generalized human capital in the covariance approach is quite consistent with previous comparisons specific to Latin America. From Fig. 3, we see that years of schooling is correlated with educational achievement for most of the sampled countries. The covariance measure implicitly gives larger weight to the relationship in these non-Latin American countries than the simple development accounting in the other comparisons.

[42] Sub-Saharan Africa is actually very similar to Latin America from an analytical standpoint. The international tests, which have been taken by a relatively small number of African countries, also appear too difficult for the typical student. On the other hand, regional testing initiatives such as the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and the Program on the Analysis of Education Systems (PASEC) in Francophone Sub-Saharan Africa possibly provide a platform both for better linkages to economic comparisons with the rest of the world and for intra-regional analysis.

[43] An implication of the limited informational content of worldwide tests such as TIMSS and PISA for developing countries is that it is important to develop ways to expand the tests at the lower levels so that both the between- and within-country variations are more reliable. It may be advisable to develop tests geared toward the performance levels relevant for these countries that allow a clear diagnosis of performance at levels equivalent to, say, 200 to 400 points on the PISA test, while at the same time including link items that allow anchoring these tests to the global assessments. The evolving capacity for adaptive testing that can adjust test content to students' ability levels seems particularly promising in this context.

Appendix A. Descriptive statistics

Table A1
Income and education of Latin American countries.

	GDP per capita 1960	Growth of GDP per capita 1960–2000	GDP per capita 2000	Years of schooling 1960	Measures of educational achievement			
					Combined test score from worldwide tests	LLECE (avg. math/ reading, 4th grade)	SERCE (avg. math/ reading, 6th grade)	Combined test score from regional tests
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Argentina	7395	1.0	10,995	6.1	392.0	275.5	509.7	395.3
Bolivia	2324	0.4	2722	3.6	–	239	–	264.0
Brazil	2395	2.8	7185	3.1	363.8	273	509.9	390.2
Chile	3818	2.4	9920	6.2	404.9	275.5	531.7	412.7
Colombia	2525	1.9	5380	3.7	415.2	261.5	503.8	361.4
Costa Rica	3480	1.3	5863	3.3	–	–	556.3	448.6
Ecuador	1974	1.4	3467	4.3	–	–	453.5	285.2
El Salvador	3306	0.7	4435	2.0	–	–	478.1	324.3
Guatemala	2354	1.3	3914	1.6	–	–	453.6	285.5
Honduras	1705	0.5	2054	1.9	–	234.5	–	245.3
Mexico	3970	2.0	8766	4.0	399.8	254	535.8	371.2
Panama	2340	2.4	6066	4.6	–	–	461.8	298.5
Paraguay	2437	1.6	4682	4.0	–	249.5	461.8	303.1
Peru	3118	1.0	4583	4.3	312.5	–	483.1	332.4
Uruguay	5840	1.3	9613	5.3	430.0	–	560.3	454.9
Venezuela	7751	−0.5	6420	2.9	–	237.5	–	257.8

Sample: All Latin American countries with populations greater than one million and without communist background; see footnote 11 for details. Measures of educational achievement: Column (5) shows the combined score from all international tests conducted between 1964 and 2003 on the metric developed in Hanushek and Woessmann (forthcoming). Columns (6) and (7) show the scores from the regional Latin American LEECE and SERCE tests on the scale of the original tests. Column (8) shows the combined measure of LLECE and SERCE performance mapped on the worldwide metric, derived in this paper (see Section 3.3 for details).
Sources: Penn World Tables (Cohen and Soto, 2007; Hanushek and Woessmann, forthcoming; Heston et al., 2002; Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 1998, 2008a).

Appendix B. School quality and Mincer returns across countries

The rates of return to years of schooling are estimated for different countries using a variety of approaches. The most common is a simple Mincer earnings function, but at times people also use direct calculations of internal rates of return (see, for example, the discussion in Psacharopoulos and Patrinos (2004)). In all of these, what is really estimated is the schooling gradient of earnings, and interpretation of this as a rate of return relies on a variety of assumptions (as partially detailed in Heckman et al. (2006)). The most basic assumption is that the cost of schooling is entirely the foregone earnings from being out of the labor market — but assumptions about the length of working life, the constancy of rates of return, the separability of schooling and on-the-job training investments, and the independence from ability differences also enter. For this, we ignore these latter complications and just concentrate on the role of school quality.

The simplest version of an earnings model employs a basic investment model where the return r on the investment of spending one more year in school (rather than in the labor market) is the parameter that equates the following investment identity:

$$Y_S(\phi^c) = Y_{S-1}(\phi^c) + r_S I_S \tag{B1}$$

where $Y_\tau(\phi^c)$ is the earnings with τ years of schooling, ϕ^c is the school quality in country c, I_S is the investment in the Sth year of schooling, and r_S is the return to investment in that year.

The standard simplification is that the cost of investing in another year of schooling is just foregone earnings, such that $I_S = Y_{S-1}(\phi^c)$. With this we have:

$$Y_S(\phi^c) = Y_{S-1}(\phi^c)(1 + r_S) \tag{B2}$$

and then by recursion and assuming that the return, r_τ, is constant:

$$Y_S(\phi^c) = Y_0(\phi^c)\prod_0^S(1+r) = Y_0(\phi^c)(1+r)^S \tag{B3}$$

where S is years of schooling.[44] By taking logs, we have:

$$\begin{aligned} \ln Y_S(\phi^c) &= \ln Y_0(\phi^c) + \ln(1+r)S \\ &\approx \ln Y_0(\phi^c) + rS \end{aligned} \tag{B4}$$

(where the approximation holds for small r). This is the standard log-linear human capital earnings function that, after allowing for experience, becomes the "Mincer" earnings model that is most frequently used in empirical analysis. Note that the return to years of schooling r is independent of the average quality of schooling ϕ^c in a country.

What this says simply is that the micro estimates of the returns to a year of schooling are a function of the proportional differences in earnings at different levels of schooling. Those earnings can be a direct function of ϕ^c, the school quality in country c, but the Mincer earnings function in Eq. (B4) still holds as a representation of earnings at different levels of schooling in any country.

An even easier way to see this is to presume that school quality in country c has a constant impact across schooling levels with the simple functional form of $Y_\tau(\phi^c) = \phi^c Y_\tau$. Then, rearranging Eq. (B2), we

[44] While this indicates that the intercept could be interpreted as the earnings with no schooling, in any actual empirical application the estimation samples generally consider those with at least a few years of primary schooling because the foregone earnings at no schooling are not really observed in the data. Further, the return parameter, r, is frequently estimated to differ by level of schooling — primary, secondary, and tertiary.

can see that the rate of return to one additional year of schooling can be expressed as:

$$r = \frac{\phi^c Y_S - \phi^c Y_{S-1}}{\phi^c Y_{S-1}} = \frac{Y_S - Y_{S-1}}{Y_{S-1}}. \tag{B5}$$

In other words, the rate of return, r, is simply the proportionate increase in earnings when going from $S-1$ to S years of schooling, and the quality of schooling, ϕ^c, drops out.

If one estimated the standard Mincer Eq. (B4) with a pooled, cross-country sample (which has never been done), it would be necessary to adjust the earnings in each country for ϕ^c, but this term within a country is subsumed in the base level of income (the constant) in each country's earnings function. Another way of saying this is that earnings before and after getting additional schooling can be lower or higher in one country or another but the parameter estimated (r) just provides the gradient of earnings across schooling levels.

References

Aghion, Philippe, Howitt, Peter, 1998. Endogenous Growth Theory. MIT Press, Cambridge, MA.

Aghion, Philippe, Boustan, Leah, Hoxby, Caroline M. and Vandenbussche, Jérôme. 2005. "Exploiting states' mistakes to identify the causal impact of higher education on growth." Mimeo, Department of Economics, Harvard University.

Andrews, Donald R., Moreira, Marcelo J., Stock, James H., 2007. Performance of conditional Wald tests in IV regression with weak instruments. Journal of Econometrics 139 (1), 116–132.

Auguste, Byron, Kihn, Paul, Miller, Matt, 2010. Closing the Talent Gap: Attracting and Retaining Top-Third Graduates to Careers in Teaching. McKinsey and Company. September.

Barrett, David B., Kurian, George T., Johnson, Todd M., 2001. World Christian Encyclopedia, 2nd ed. Oxford University Press, Oxford.

Barro, Robert J., 1991. Economic growth in a cross section of countries. Quarterly Journal of Economics 106 (2), 407–443 (May).

Barro, Robert J., 1997. Determinants of Economic Growth: a Cross-Country Empirical Study. MIT Press, Cambridge, MA.

Barro, Robert J., Lee, Jong-Wha, 1993. International comparisons of educational attainment. Journal of Monetary Economics 32 (3), 363–394.

Barro, Robert J., Lee, Jong-Wha, 2010. A new data set of educational attainment in the world, 1950–2010. NBER Working Paper 15902. National Bureau of Economic Research, Cambridge, MA. April.

Becker, Sascha O., Woessmann, Ludger, 2009. Was Weber wrong? A human capital theory of Protestant economic history. Quarterly Journal of Economics 124 (2), 531–596.

Behrman, Jere R., Birdsall, Nancy, Székely, Miguel, 2007. Economic policy changes and wage differentials in Latin America. Economic Development and Cultural Change 56 (1), 57–97.

Benhabib, Jess, Spiegel, Mark M., 2005. Human capital and technology diffusion. In: Aghion, Philippe, Durlauf, Steven N. (Eds.), Handbook of Economic Growth. North Holland, Amsterdam, pp. 935–966.

Bils, Mark, Klenow, Peter J., 2000. Does schooling cause growth? The American Economic Review 90 (5), 1160–1183 (December).

Caselli, Francesco, 2005. Accounting for cross-country income differences. In: Aghion, Philippe, Durlauf, Steven N. (Eds.), Handbook of Economic Growth. North Holland, Amsterdam, pp. 679–741.

Chetty, Raj, Friedman, John N., Hilger, Nathaniel, Saez, Emmanuel, Schanzenbach, Diane Whitmore, Yagan, Danny, 2011. How does your kindergarten classroom affect your earnings? Evidence from Project Star. Quarterly Journal of Economics 126 (4), 1593–1660 (November).

Ciccone, Antonio, Papaioannou, Elias, 2009. Human capital, the structure of production, and growth. The Review of Economics and Statistics 91 (1), 66–82 (February).

Cohen, Daniel, Soto, Marcelo, 2007. Growth and human capital: good data, good results. Journal of Economic Growth 12 (1), 51–76 (March).

Cole, Harold L., Ohanian, Lee E., Riascos, Alvaro, Schmitz Jr., James A., 2005. Latin America in the rearview mirror. Journal of Monetary Economics 52 (1), 69–107 (January).

Dolton, Peter, Marcenaro-Gutierrez, Oscar D., 2011. If you pay peanuts do you get monkeys? A cross country analysis of teacher pay and pupil performance. Economic Policy 26 (65), 7–55 (January).

Easterly, William, 2001. The Elusive Quest for Growth: an Economist's Adventures and Misadventures in the Tropics. The MIT Press, Cambridge, MA.

Easterly, William, Levine, Ross, 1997. Africa's growth tragedy: policies and ethnic divisions. Quarterly Journal of Economics 112 (4), 1203–1250 (November).

Edwards, Sebastian, Esquivel, Gerardo, Márquez, Graciela (Eds.), 2007. The Decline of Latin American Economies: Growth, Institutions, and Crises. University of Chicago Press, Chicago, IL.

Fernández-Arias, Eduardo, Manuelli, Rodolfo, Blyde, Juan S. (Eds.), 2005. Sources of Growth in Latin America: What Is Missing? Inter-American Development Bank, Washington, DC.

Fuller, Wayne A., 1977. Some properties of a modification of the limited information estimator. Econometrica 45 (4), 939–954.

Gundlach, Erich, Rudman, Desmond, Woessmann, Ludger, 2002. Second thoughts on development accounting. Applied Economics 34 (11), 1359–1369.

Hahn, Jinyong, Hausman, Jerry A., Kuersteiner, Guido, 2004. Estimation with weak instruments: accuracy of higher-order bias and MSE approximations. The Econometrics Journal 7 (1), 272–306.

Hall, Robert E., Jones, Charles I., 1999. Why do some countries produce so much more output per worker than others? Quarterly Journal of Economics 114 (1), 83–116.

Hanson, Gordon H., 2010. Why isn't Mexico rich? Journal of Economic Literature 48 (4), 987–1004 (December).

Hanushek, Eric A., 2002. Publicly provided education. In: Auerbach, Alan J., Feldstein, Martin (Eds.), Handbook of Public Economics, vol. 4. North Holland, Amsterdam, pp. 2045–2141.

Hanushek, Eric A., 2011. The economic value of higher teacher quality. Economics of Education Review 30 (3), 466–479 (June).

Hanushek, Eric A., Kimko, Dennis D., 2000. Schooling, labor force quality, and the growth of nations. The American Economic Review 90 (5), 1184–1208 (December).

Hanushek, Eric A., Rivkin, Steven G., 2010. Generalizations about using value-added measures of teacher quality. The American Economic Review 100 (2), 267–271 (May).

Hanushek, Eric A., Woessmann, Ludger, 2008. The role of cognitive skills in economic development. Journal of Economic Literature 46 (3), 607–668 (September).

Hanushek, Eric A.,Woessmann, Ludger, forthcoming. Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. Journal of Economic Growth National Bureau of Economic Research, Cambridge, MA. January.

Hanushek, Eric A., Woessmann, Ludger, 2009. Schooling, cognitive skills, and the Latin American growth puzzle. NBER Working Paper 15066. National Bureau of Economic Research, Cambridge, MA. June.

Hanushek, Eric A., Woessmann, Ludger, 2011a. The economics of international differences in educational achievement. In: Hanushek, Eric A., Machin, Stephen, Woessmann, Ludger (Eds.), Handbook of the Economics of Education, vol. 3. North Holland, Amsterdam, pp. 89–200.

Hanushek, Eric A., Woessmann, Ludger, 2011b. Sample selectivity and the validity of international student achievement tests in economic research. Economic Letters 110 (2), 79–82 (February).

Hanushek, Eric A., Zhang, Lei, 2009. Quality-consistent estimates of international schooling and skill gradients. Journal of Human Capital 3 (2), 107–143 (Summer).

Hanushek, Eric A., Peterson, Paul E., Woessmann, Ludger, 2012. Is the United States catching up? International and state trends in student achievement. Education Next 12 (4) (Fall).

Heckman, James J., Lochner, Lance J., Todd, Petra E., 2006. Earnings functions, rates of return and treatment effects: the Mincer equation and beyond. In: Hanushek, Eric A., Welch, Finis (Eds.), Handbook of the Economics of Education. North Holland, Amsterdam, pp. 307–458.

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

Hsieh, Chang-Tai, Klenow, Peter J., 2010. Development accounting. American Economic Journal: Macroeconomics 2 (1), 207–223.

Jamison, Eliot A., Jamison, Dean T., Hanushek, Eric A., 2007. The effects of education quality on mortality decline and income growth. Economics of Education Review 26 (6), 772–789 (December).

Kehoe, Timothy J., Ruhl, Kim J., 2010. Why have economic reforms in Mexico not generated growth? Journal of Economic Literature 48 (4), 1005–1027 (December).

Klenow, Peter J., Rodriquez-Clare, Andres, 1997. The neoclassical revival in growth economics: has it gone too far? In: Bernancke, Ben S., Rotemberg, Julio J. (Eds.), NBER Macroeconomics Annual 1997. MIT Press, Cambridge, MA, pp. 83–103.

Krueger, Alan B., Lindahl, Mikael, 2001. Education for growth: why and for whom? Journal of Economic Literature 39 (4), 1101–1136 (December).

Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 1998. First International Comparative Study of Language, Mathematics, and Associated Factors in Third and Fourth Grades. Latin American Educational Quality Assessment Laboratory, Santiago, Chile.

Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2001. Primer Estudio Internacional Comparativo sobre Lenguaje, Matemática y Factores Asociados, para Alumnos del Tercer y Cuarto Grado de la Educación Básica: Informe Técnico. UNESCO—Santiago, Oficina Regional de Educación para América Latina y el Caribe, Santiago, Chile.

Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2002. First International Comparative Study of Language, Mathematics, and Associated Factors in the Third and Fourth Grade of Primary School: Second Report. UNESCO—Santiago, Regional Office of Education for Latin America and the Caribbean, Santiago, Chile.

Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2005. Segundo Estudio Regional Comparativo Explicativo 2004–2007: Análisis Curricular. Oficina Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC/UNESCO), Santiago, Chile.

Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2008a. Los Aprendizajes de los Estudiantes de América Latina y el Caribe: Primer Reporte de los Resultados del Segundo Estudio Regional Comparativo y Explicativo. Oficina Regional de Educación de la UNESCO para América Latina y el Caribe (OREALC/UNESCO), Santiago, Chile.

Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación, 2008b. Student Achievement in Latin America and the Caribbean: Results of the Second Regional Comparative and Explanatory Study (SERCE) — Executive Summary. Regional Bureau for Education in Latin America and the Caribbean OREALC/UNESCO, Santiago, Chile.

Lazear, Edward P., 2003. Teacher incentives. Swedish Economic Policy Review 10 (3), 179–214.
Levine, Ross, Renelt, David, 1992. A sensitivity analysis of cross-country growth regressions. The American Economic Review 82 (4), 942–963 (September).
Mankiw, N. Gregory, Romer, David, Weil, David, 1992. A contribution to the empirics of economic growth. Quarterly Journal of Economics 107 (2), 407–437 (May).
Moreira, Marcelo J., 2003. A conditional likelihood ratio test for structural models. Econometrica 71 (4), 1027–1048.
Mulligan, Casey B., 1999. Galton versus the human capital approach to inheritance. Journal of Political Economy 107 (no. 6, pt. 2), S184–S224 (December).
Murnane, Richard J., Willett, John B., Duhaldeborde, Yves, Tyler, John H., 2000. How important are the cognitive skills of teenagers in predicting subsequent earnings? Journal of Policy Analysis and Management 19 (4), 547–568 (Fall).
Nelson, Richard R., Phelps, Edmund, 1966. Investment in humans, technology diffusion and economic growth. The American Economic Review 56 (2), 69–75 (May).
Pritchett, Lant, 2001. Where has all the education gone? World Bank Economic Review 15 (3), 367–391.
Pritchett, Lant, 2004. Access to education. In: Lomborg, Björn (Ed.), Global Crises, Global Solutions. Cambridge University Press, Cambridge, pp. 175–234.
Pritchett, Lant, 2006. Does learning to add up add up? The returns to schooling in aggregate data. In: Hanushek, Eric A., Welch, Finis (Eds.), Handbook of the Economics of Education. North Holland, Amsterdam, pp. 635–695.
Psacharopoulos, George, Patrinos, Harry A., 2004. Returns to investment in education: a further update. Education Economics 12 (2), 111–134 (August).
Rivkin, Steven G., Hanushek, Eric A., Kain, John F., 2005. Teachers, schools, and academic achievement. Econometrica 73 (2), 417–458 (March).
Rockoff, Jonah E., 2004. The impact of individual teachers on student achievement: evidence from panel data. The American Economic Review 94 (2), 247–252 (May).
Romer, Paul, 1990. Endogenous technological change. Journal of Political Economy 99 (no. 5, pt. II), S71–S102.
Topel, Robert, 1999. Labor markets and economic growth. In: Ashenfelter, Orley, Card, David (Eds.), Handbook of Labor Economics. North Holland, Amsterdam, pp. 2943–2984.
UNESCO, 2005. Education for All: the Quality Imperative, EFA Global Monitoring Report. UNESCO, Paris.
West, Martin R., Woessmann, Ludger, 2010. 'Every Catholic child in a Catholic school': historical resistance to state schooling, contemporary private competition and student achievement across countries. The Economic Journal 120 (546), F229–F255.