



What We Can Learn About Latin American Educational Systems from International Tests: A Brief Foray

Martin Carnoy
Stanford University

Tatiana Khavenson
HSE University

The Revista del Centro de Estudios Educativos, numero 3, 1971 included an early Carnoy article on the economics of education: “Un enfoque de sistemas para evaluar la educación, ilustrado con datos de Puerto Rico.” The article used a unique data set that had student test scores, students’ family background characteristics, and information about teachers and other school inputs for about one-third of all students in Puerto Rican schools to estimate relations between teacher characteristics and student test scores controlling for students’ social class, gender, and whether the school was urban or rural. Such data sets were rare in the late 1960s, and so were attempts to understand how education systems worked to produce student learning outcomes—that is, to improve the quality of education.

There is a lot to criticize in the empirical analysis in that early article, but it does show that there was considerable concern about the quality of education in Latin America even back in 1971. That concern has grown greatly in the past fifty years as countries in the region have expanded their educational systems to provide an increasing proportion of youth with secondary schooling and higher education. With that expansion, there has been a shift in focus from policies concerned with access to schooling to policies concerned with improving the quality of schooling (UNESCO, 2005).

VERSION: February 2023

Suggested citation: Carnoy, Martin, and Tatiana Khavenson. (2023). What We Can Learn About Latin American Educational Systems from International Tests: A Brief Foray. (EdWorkingPaper: 23-725). Retrieved from Annenberg Institute at Brown University: <https://doi.org/10.26300/8y90-6q29>

What We Can Learn About Latin American Educational Systems from International Tests: A Brief Foray

Martin Carnoy and Tatiana Khavenson

The *Revista del Centro de Estudios Educativos*, numero 3, 1971 included an early Carnoy article on the economics of education: “Un enfoque de sistemas para evaluar la educación, ilustrado con datos de Puerto Rico.” The article used a unique data set that had student test scores, students’ family background characteristics, and information about teachers and other school inputs for about one-third of all students in Puerto Rican schools to estimate relations between teacher characteristics and student test scores controlling for students’ social class, gender, and whether the school was urban or rural. Such data sets were rare in the late 1960s, and so were attempts to understand how education systems worked to produce student learning outcomes—that is, to improve the quality of education.

There is a lot to criticize in the empirical analysis in that early article, but it does show that there was considerable concern about the quality of education in Latin America even back in 1971. That concern has grown greatly in the past fifty years as countries in the region have expanded their educational systems to provide an increasing proportion of youth with secondary schooling and higher education. With that expansion, there has been a shift in focus from policies concerned with access to schooling to policies concerned with improving the quality of schooling (UNESCO, 2005).

Two factors have contributed to this shift. The first is research claiming that quality of education, as measured by international test scores, is a better predictor of economic growth than the number of years of schooling in the labor force (Hanushek and Kimko, 2000; Hanushek and Woessman, 2008). The second is the increase in testing itself, both at the national and international levels. Student test results are being used increasingly to pressure national and local educational systems, schools, and individual teachers to have their students do better on the tests (OECD, 2013). League tables comparing schools, local school districts, regions, and nations against others are now a regular feature of educational politics in many countries of the world. To some extent, international test scores are becoming important enough to affect government legitimacy.

In Latin America, the most prevalent of these international tests is the OECD’s PISA test, given every three years.¹ There has been a lot of research using Latin American PISA data, from overall assessments of the results (OECD, 2016) to estimating the “efficiency” of schooling

¹ The most extensive test in Latin America, however, is the LLECE test, applied by UNESCO Santiago, and now known as the PERCE (1997), SERCE (2006), and TERCE (2013) test, which cover a much larger number of countries than PISA, are grade-based, and focus on primary schooling. One advantage of the UNESCO test and survey is that they include Cuba, where students scored the highest in Latin America in 1997 and 2006. Unfortunately, Cuba did not participate in 2013. It would be possible to do a comparison across all years using these tests only for about seven countries. For an analysis of the 1997 LLECE test comparable to the analysis we undertake in this essay with the PISA test, see Carnoy and Marshall, 2005.

using PISA 2006 data (Deutsch et al, 2013) to estimating the effect of an additional year of schooling in seven countries of the PISA 2012 survey (Marchionni and Vazquez, 2019). An important takeaway from analyzing PISA results is that as for all measures of academic performance in school, individual students' family background (social class) makes a major difference in student academic outcomes, and the *average social class of students in a school* may make an even larger difference in how much students learn (student learning gains) as they go through school (Flores-Mendoza et al, 2021). Almost all the social class gap in learning occurs before students enter school at 6 or 7 years old (Jencks and Phillips, 1998), but we know surprisingly little about why it has been so difficult to close the gap in schools.

As the Carnoy article in *la Revista* illustrates, we already knew fifty years ago that high quality teachers are crucial to improving student outcomes in schools. That has been confirmed using more sophisticated statistical techniques (see, for example Rivkin et al, 2005). One major problem with the PISA survey is that until 2018, it—unlike the TIMSS and UNESCO tests—did not collect data from teachers, and since PISA a school based survey, even with the data available for the 2018 survey, it is not possible to link teachers with students' performance on the test and therefore not possible to estimate teacher effects accurately using the PISA data.

Thus, a serious problem in improving the analyses of educational production in Latin America is the paucity of longitudinal test and survey data with which we can follow the same students as they move from grade to grade in the same or different schools, identifying them with various teachers and schools, as well as with the characteristics of those teachers and schools. Given longitudinal data, we can learn a lot about which school inputs contribute to, say, test score gains and repetition rates. The existence of such data for individual students linked to schools in some Brazilian states and in Chile, for example, has allowed researchers to estimate the effects of various policies, such as full-day schools (Rosa et al, 2022; Bellei, 2009), literacy programs (Carnoy and Costa, 2015), or spending more on the schooling of low-income students (Carrasco, 2014; Murnane et al, 2017; Mizala and Torche, 2017).

Another serious problem is that we don't often take a hard comparative look at educational systems that work in similar social contexts yet produce quite different results. In this essay—and let us be clear, this is really more of an essay than an academic paper—we want to step back from the more sophisticated methods and longitudinal data needed to find answers to some of the more esoteric questions about educational production. Rather we want to conduct a simpler exploration. We want to see what, if anything, we can learn from international tests applied in various Latin American countries facing similar social conditions about what they have accomplished in twenty years of trying to improve their educational systems, and what this may imply about improving Latin American education in the future.

Ultimately our goal is to make a more thorough assessment of the effectiveness of public school systems in Latin America and how they may be changing over time than provided by simply comparing average international test scores for different countries. In order to do this, we will try to peel back two layers of variables that have important influence over how well students do on this test but probably don't reflect the quality of public schooling in these countries. The first of these layers of variables is student family resources and the second is mainly the average social class of the school attended by students (peer effect) and the grade attended when they

took the test, since this would reflect exposure to the mathematics and reading taught by schools and tested by the items on the test instrument.

Some countries in the region have taken the TIMSS test on occasion (Chile, for example), but six countries—Argentina, Brazil, Chile, Mexico, Peru, and Uruguay—have participated rather consistently in OECD’s PISA test since 2000 until the last test for which data are available—2018. Colombia joined in 2006, and Costa Rica and Panama joined in 2009. Furthermore, in Mexico, until 2012, PISA was applied in random samples in each of the Mexican states, which allows us to compare changes in PISA test scores across states within the same federal country.

We will focus on the six countries that have participated in the test since 2003 or earlier. One main contribution we make here is to “adjust” the comparisons of student performance for differences in social class and gender distributions of the students tested, the distribution of students across grades, the relative concentrations of students tested in lower and higher social class schools and in urban, rural, and private schools. We make these adjustments across the six countries for the PISA tests in 2000-2018 and we similarly adjust the Mexican state data for the period 2003-2012.

What do the results of many rounds of PISA tests applied to 15 year-olds in these countries reveal and hypothetical school public “quality” and changes in public school “quality” once we correct for the social class differences and length of time differences in school? What further lessons can we draw from results across states within Mexico?

The findings from these relatively simple “macroanalysis” comparative estimates are revealing. First, increases vary by students’ social class. Before adjustment for the average grade attended, the largest gains over time have generally come for lower-middle and low-family resources students. But one country—Peru—stands out in being able to increase student performance among advantaged students, albeit from initially very low levels.

Second, the analysis indicates that an important question for educational policy in the region should be why it has been so difficult for Latin American schools to increase the performance on these PISA tests of youth who come from resource advantaged families. A good comparison is that scores in 2015-2018 in the two Latin American countries—Chile and Uruguay—where this group of “privileged” students scored the highest, were about 470 points in math and 500 point in reading, whereas in Spain, a relatively low scoring European country, family resource-advantaged 15 year-olds scored about 520 in math and 530 in reading. Third, Mexico’s pattern of test scores stands out because very disadvantaged and disadvantaged students in Mexico score higher than in other countries throughout this period, even than students in higher average scoring countries, such as Uruguay and Chile.

Third, our results produce a surprising result that once we adjust the relatively much higher “published” Chilean student performance for the higher social class and educational attainment of the Chilean student sample, students in Mexico and Uruguay (and possibly Brazil) perform better than students in Chile, especially in mathematics.

Fourth, we find that although students in several of the six countries we analyze made substantial increases in average performance in the period 2000-2018—especially students in Peru and Uruguay—much of this increase was the result of increases in the average social class and average school attainment of 15 year-olds in the PISA samples rather than improvements in learning from the same amount of years in school for similar social class students.

The comparison of math scores within Mexico among states is also informative, although also descriptive rather than providing any insights into why in some states students made significant gains and students' relative math scores in a few states declined. There have been early attempts to estimate state education production functions using the PISA 2003 results (Alvarez et al, 2007). This shows that some institutional differences between Mexican states, such as the quality of the relations between the state administration and the teachers' union, help explain differences in PISA scores. However, to our knowledge no one has continued such research to uncover factors that may have influenced changes in test scores over time.

Our results show that students in only a handful of Mexican states made large gains in their mathematics test scores relative to the Mexican mean score over the nine year period 2003-2012 covered by the PISA surveys in the states. But the good news is that several of these states were initially low-scoring, and the other good news is that some states made rather large gains in this period. It will be important to understand why those states did so well and why students in other states, such as the Distrito Federal, made small gains.

Data and Method

The OECD's PISA is a test administered to a sample of 15 year-olds based on the distribution of 15 year-old students in various grades in each country. The PISA is a school sample, testing and apply a survey to about twenty-five randomly-drawn 15 year-olds in each selected school. In two Latin American countries—Brazil and Mexico—the sample have generally been much larger, as we will further explain below. The PISA assessment and survey data can be used to draw some inferences as to whether Latin America students are improving the amount of mathematics and language skills they have learned by the end of basic education into early high school.

Our methodological approach to approximating improvements in school system quality is to “net out” one major part of out-of-school influences by comparing students with similar family academic resources across countries. We argue that changes in test scores over time of students adjusting for student characteristics (such as gender) and family academic resources provide a better assessment of whether a country's educational system is improving than simply tracking average national scores. There are additional complexities concerning the PISA test because students sampled are in a given age group, not in a single grade, and, in some countries, such as Brazil, the test was applied at different dates in 2000, 2003/2006, 2009, and 2012, further biasing estimates of gains over time (Klein (2011)). We attempt to control for grade in estimating the test score changes in the PISA for advantaged and disadvantaged students.

Our empirical strategy is descriptive and comparative. We first estimate the level of and changes in each country's disadvantaged and advantaged students' PISA scores in mathematics

and compare them with their counterparts in other countries, focusing on Argentina, Brazil, Chile, and Mexico in 2000-2018. Secondly, we describe the test scores gains for 15-year-old in these countries, controlling for gender, student social class, school social class, whether the school attended was public or private, and the grade in which students in each year of the test were sampled. Third, we compare the trajectories of PISA mathematics scores in Mexican states in 2003-2015 using similar controls.

Results

Published PISA Reading and Mathematics scores

Before turning to our analysis by students' family resources and other "corrections" to make test scores among countries more comparable in terms of what might more closely reflect the effect that schooling and educational policy differences in various countries have on student performance on a test such as the PISA, let's take a look at the published results by country for reading and mathematics, which don't include any of these adjustments (Table 1).

From Table 1, we can see that the only countries in Latin America in which students have made significant gains in the PISA reading tests are Chile and Peru. Students in more countries have made significant gains in mathematics, with the largest gains coming in Peru, Brazil, Chile, and Mexico, in that order. The gains are in the order of more than one SD in Peru, about 0.5 SD in Brazil, 0.3 SD in Chile, and 0.2 SD in Mexico. Colombian students increased their math score by 0.2 SD as well, while score in Argentina, Uruguay, Costa Rica, and Panama have remained essentially unchanged.

Table 1. Latin America: Published PISA Reading and Mathematics Mean Scores, by Country, 2000-2018.

Mean Reading Score							
<i>Country</i>	<i>2000</i>	<i>2003</i>	<i>2006</i>	<i>2009</i>	<i>2012</i>	<i>2015</i>	<i>2018</i>
Argentina	418		374	398	396		402
Brazil	396	403	393	412	407	407	413
Chile	410		442	449	441	459	452
Colombia			385	413	403	425	412
Mexico	422	400	410	425	424	423	420
Peru	327			370	384	398	401
Uruguay		434	413	426	411	437	427
Costa Rica				443	441	427	426
Panama				371			377

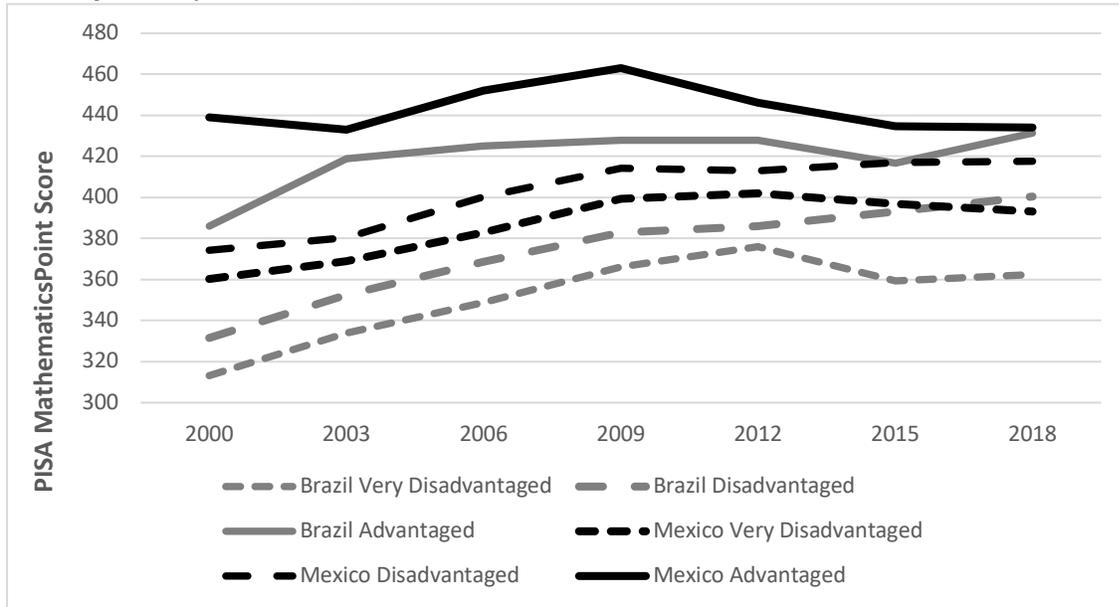
Mean Mathematics Score							
<i>Country</i>	<i>2000</i>	<i>2003</i>	<i>2006</i>	<i>2009</i>	<i>2012</i>	<i>2015</i>	<i>2018</i>
Argentina	388		381	388	388		379
Brazil	334	356	370	386	389	377	384
Chile	384		411	421	423	423	417
Colombia			370	381	376	390	391
Mexico	387	385	406	419	413	408	409
Peru	292			365	368	387	400
Uruguay		422	427	427	409	418	418
Costa Rica				409	407	400	402
Panama				360			353

Source: OECD PISA, published data in PISA reports, OECD PISA 2000 to OECD PISA 2018.
<https://www.oecd.org/pisa/publications/>

Comparing the relative performance of very disadvantaged, disadvantaged, and advantaged students

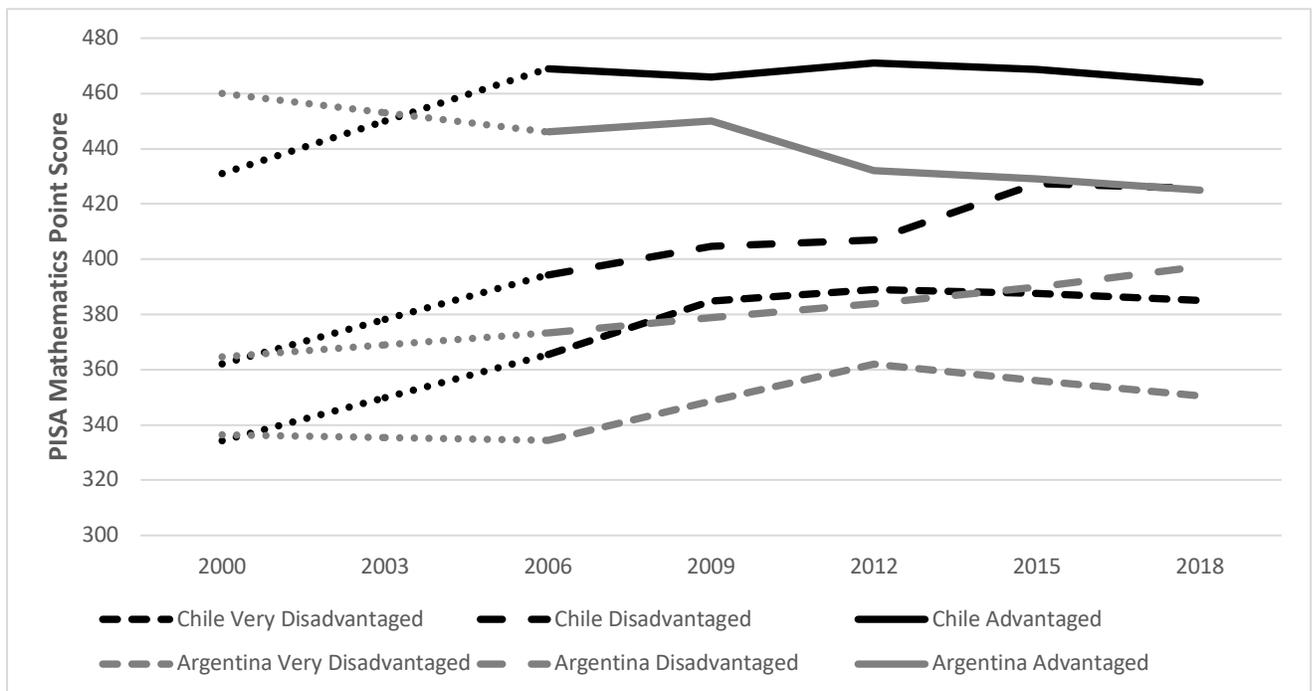
Figures 1, 2, 3, and 4, and 5 compare PISA test score trajectories for very disadvantaged (0-10 books in the home), disadvantaged students (11-99 books in the home), and advantaged students (>100 books in the home) in five pairs of countries, Brazil-Mexico, Chile-Argentina, Mexico-Chile, Peru-Mexico, and Chile-Uruguay.

Figure 1. PISA Average National Student Mathematics Scores for Brazil and Mexico, by Level of Family Resources, 2000-2018



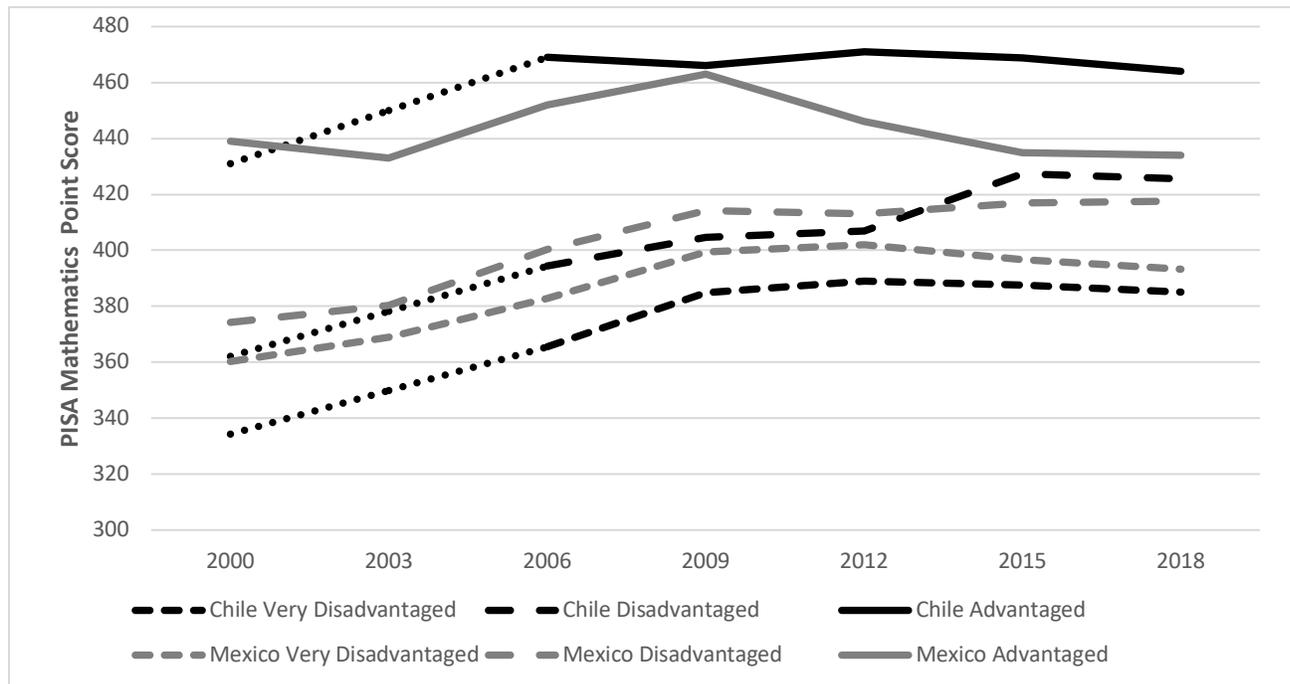
Source: OECD, PISA, publicly available microdata, 2000-2018. Authors' calculations.

Figure 2. PISA Average National Student Mathematics Scores for Chile and Argentina, by Level of Family Resources, 2000-2018



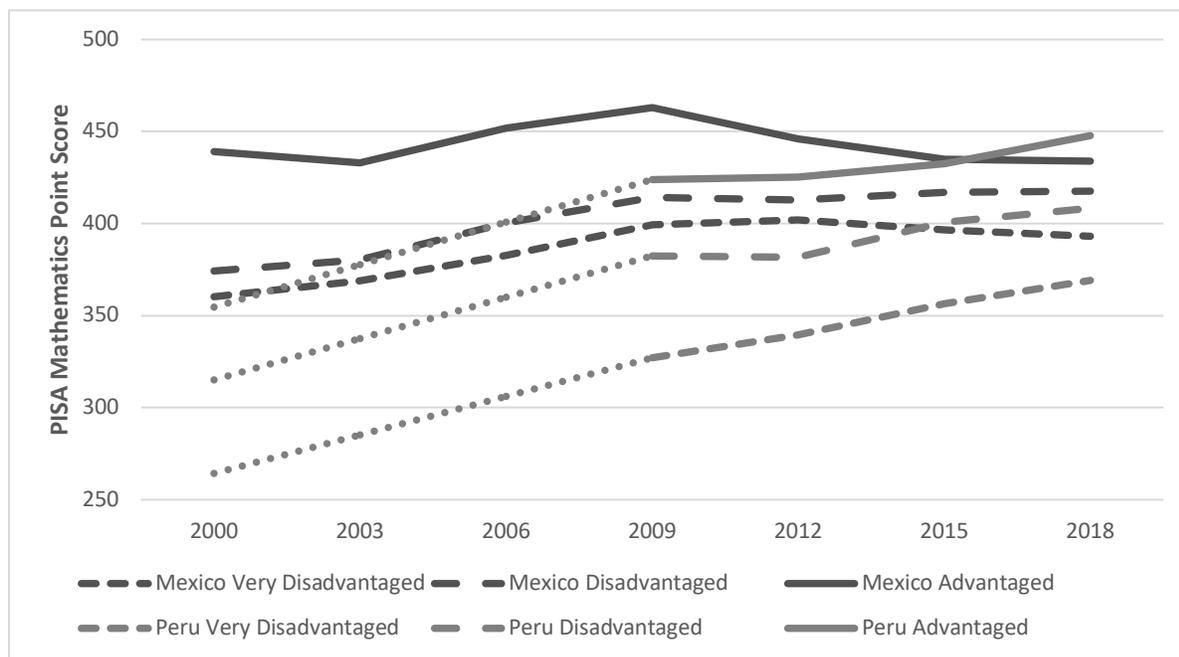
Source: OECD, PISA, publicly available microdata, 2000-2018. Authors' calculations.

Figure 3. PISA Average National Student Mathematics Scores for Mexico and Chile, by Level of Family Resources, 2000-2018.



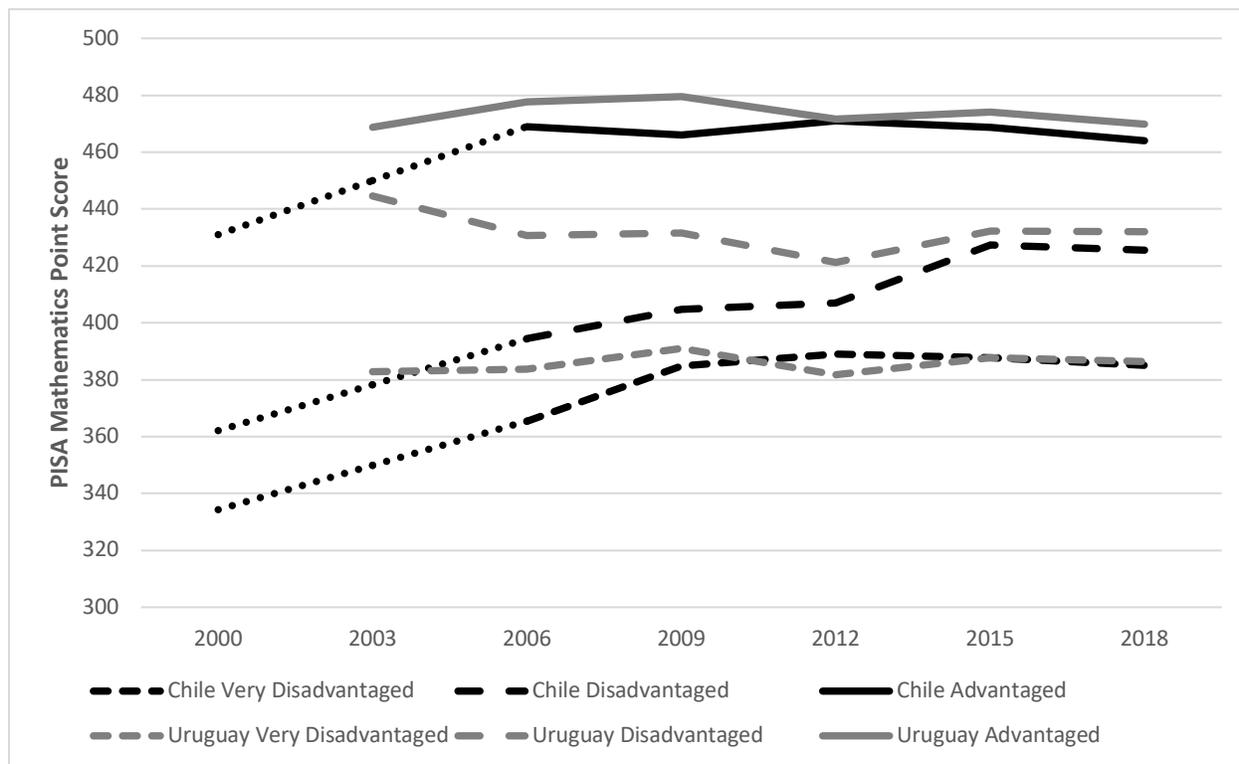
Source: OECD, PISA, publicly available microdata, 2003-2018. Authors' calculations.

Figure 4. PISA Average National Student Mathematics Scores for Mexico and Peru, by Level of Family Resources, 2000-2018



Source: OECD, PISA, publicly available microdata, 2003-2018. Authors' calculations.

Figure 5. PISA Average National Student Mathematics Scores for Chile and Uruguay, by Level of Family Resources, 2000-2018.



Source: OECD, PISA, publicly available microdata, 2003-2018. Authors' calculations.

We can draw several contrasts from these figures that raise questions about the kinds of education and other social policies that may have contributed to the patterns we observe. In Figure 1 we observe the two most populous Latin American countries—Brazil and Mexico. Students at all levels of family resources (FAR)—especially the very disadvantaged and the advantaged—scored higher in mathematics during this period in Mexico than in Brazil, yet in both countries student performance increased considerably on the PISA test in 2000-2018 for these lower family resource groups. During this period, the Brazil and Mexico PISA samples had similar FAR distributions—for example, in 2009, the very disadvantaged were 41% of the Brazil sample and 37% of the Mexican sample, whereas the advantaged were about 8% in Brazil and 10% in Mexico. Thus, the much higher test scores in Mexico for the very disadvantaged and disadvantaged suggest that Mexico has been able to get considerably better results in its schools for the vast majority of the student population.

In Figure 2 and Figure 5, we compare two pairs of relatively high income countries, Chile with Argentina and Chile with Uruguay. Chile and Argentina did not take the test in 2003, so the 2003 scores in these graphs for Chile and Argentina are interpolated. Students at all three levels of family resources in both Argentina and Uruguay have essentially flat math score trajectories in 2000-2018, but Uruguay has been able to maintain the highest PISA math test scores in Latin America (but not increase them) whereas Argentina has seen a distinct decline in student PISA math performance among its advantaged students. Chile, meanwhile has made significant gains

in PISA math at all levels of student family resources, so student performance has caught up with Uruguay for all three groups. We should remember that although these are all relatively high income per capita countries in the Latin American context, the PISA sample for Chile has a significantly higher family resource index than even Uruguay and Argentina. In addition, this has been a period of significant transition for Uruguay to a very low CO₂ emissions society, which includes a de-emphasis of rapid economic growth and a move to self-sustaining reduced “wants.” Has this intentionally reduced pressure to raise PISA scores? The reasons for the trajectory of Argentina’s PISA scores is likely very different, and Chile remains very focused on improving its educational system as measured by domestic (SIMCE) and international test performance, putting special emphasis on the performance of low-income students (Carrasco, 2014; Murnane et al, 2017) and on increasing spending on teachers (Bellei and Vanni, 2015). Has that made the difference in these trajectories of PISA scores?

Yet, interestingly enough, Mexico has had the most success with low-income students, at least in terms of PISA mathematics scores. In Figure 3, we observe that until 2015, the math scores of disadvantaged students in Mexico were higher than in Chile, and the scores of very disadvantaged students in Mexico continue to be higher than those in Chile. We have many studies of the reforms Chile has put in place in the 1990s and during the first decade of the 2000s to improve student test performance (for a summary, see Bellei and Vanni, 2015). Mexico has also attempted major reforms of the primary and secondary systems, especially in professionalizing teachers, but it is not clear what the overall effects of such reforms have been (Levinson et al, 2013; Ornelas, 2016). In any case, PISA scores have risen significantly in both countries for the vast majority of students, as represented by the very disadvantaged and disadvantaged groups in Figure 3.

Finally, in Figure 4, we compare Mexico, with its large Indigenous population, with Peru, which has proportionately and even larger proportion of Indigenous population. Among all the students in Latin American countries that take the PISA, students in Peru have made the largest mathematics score gains in this period. Further, unusual in Latin America, Peru has been able to increase scores for its advantaged group substantially from low levels to average scores higher than in Mexico. And students at lower levels of family resources in Peru have also made very large gains, almost catching up to students in Mexico from the same family resource background. What have been the reforms in Peru that have produced these enormous gains? We cannot answer this question in this essay, but it should be an important subject for research to uncover whether these gains are “real” and what their sources are.

Regression results and comparisons “corrected” for student grade level, gender, social class, and type of school.

In this section of the paper we compare individual student performance on PISA in six Latin American countries—Argentina, Brazil, Chile, Mexico, Peru, and Uruguay—in the period 2000-2018, using simple ordinary least squares (OLS) regression analysis to “adjust” student PISA scores in mathematics and reading for individual student socio-economic background

(ESCS),² the average socio-economic background of the school attended by the student (school ESCS), student gender, the grade in which the student was studying when he or she took the PISA, and whether the student attended private or public school. The coefficient of each country dummy (with Mexico as the reference country) show the student performance in that country relative to student performance in Mexico controlling for all the adjustment variables.

In making these adjustments using regression analysis coefficients and variable means, we have to deal with two problems. The first is the weighting of samples from the different countries. The PISA sample sizes vary from country to country. Mexico and Brazil sampled relatively large numbers of students from 2006 to 2012, after which Brazil continued to draw relatively large samples, but Mexico reduced its sample size from more than 30 thousand students to less than 8,000 students. Brazil and Mexico have large student enrollments, so it makes sense for them to have larger representation in the total sample of the six countries. Uruguay, on the other hand, samples about five thousand students every PISA year (about 20% of the Brazilian sample size), but has only about 2% of the number of students in secondary school as Brazil. To deal with this problem, we weighted the actual sample sizes in each country in each year by the relative size of the enrollment in secondary school in each country. Thus, the percentages of the overall sample for each country shown in Table 2 are the *weighted percentages*. A student in Uruguay has a relatively low weight in the regressions compared to a student in Brazil or Mexico.

The second major concern is that Brazil implemented a major reform in 2007 to change the admission age into first grade from 7 years-old to 6 years-old. A number of municipalities had already begun implementing this reform as early as 2003, and all had to comply by 2010 (Rosa et al, 2018). The effect of the reform was to gradually increase the total number of years of schooling that students were taking in primary education after 2007 from 4 to 5 years, and in total basic education, from 8 to 9 years. In the meantime, however, in the year that the reform was implemented, 7 year-olds who enrolled began school in 2nd grade, so that even in the 2015 PISA, many 15 year-olds had entered in 2008 in 2nd grade as 7 year-olds, were designated as in 10th grade, but only had had 9 years of schooling. Thus, it has taken many years for this increase in years attending school to work its way through the system, but by 2012, Brazil had effectively changed the numbers on the grades, so that the grade reported for students in the PISA sample had increased by one, yet most students in a designated grade had had one year less schooling than the designated grade because they had entered school at 7 years-old either in first or second grade. This is an issue for our analysis because we want to be able to compare the PISA scores of students in each country on the basis of being exposed to a hypothetically similar “dosage” of schooling (curriculum). If students in Brazil who entered the schools after 2003 or 2004 actually received more schooling than students in earlier years, we should allow the new grade numbers to reflect that extra education. However, if we think that the nominal grade assigned to students

² ESCS is an OECD constructed measure of student social class background, which includes, among other variables, books in the home, parents’ education, and articles in the home. It’s value is defined in terms of standard deviations from the OECD mean of ESCS, which is set as zero. Thus, the mean ESCS of the students sampled in these six Latin American countries is about one SD below the OECD mean ESCS, and actually declined relative to the OECD mean in the period 2003-2018.

who entered first grade at 6 years-old did not represent the number of years spent in school for most students in that grade, we should subtract a year from the new grade numbers beginning in 2012 and run the regressions with the adjusted grade classification for Brazil. Our solution to this issue was to present the results for both sets of regressions. The “true” score for Brazil is somewhere between the two values.

This adjusted PISA score for students goes a step farther in capturing school effectiveness than presenting the average scores over time divided into three family academic resource groups, as in the previous section. Since we also adjust here for the average *school social class* (also an indicator of social class segregation across schools and peer effect), which can influence test score gains (Willms, 2010; Treviño et al, 2016), and the grade in which students are studying when they took the PISA test, which influences their exposure to the material covered by the test, this “adjusted” student performance and its trajectory reflects how well students of similar socio-economic background and gender in the same grade of similar socio-economic school contexts in one Latin American country perform on the PISA test compared to students on other Latin American countries.

This comparison therefore somewhat better indicates how well a “randomly selected” student in one these countries would do in the middle/and secondary public schools of another of these countries. It therefore tells us more about the effectiveness of public schooling across these six countries and how it has changed over time than simply comparing raw PISA scores, which reflect the effects student performance of many factors that have little to do with the “quality” of a year of schooling in that country. Of course, there are other “unobservable” factors we are not controlling for, especially the “initial academic ability” of the students tested in each country. It would be far better to compare the “value-added” of schooling in each of these countries, say between 5th and 10th grades, than simply measuring a test score of students attending schools in 9th and 10th grades, but it is not possible to make these kind of estimates with PISA data.

Table 2 presents the mean values for the control variables in the overall PISA samples we use in our regression analyses. As indicated, for 2012-2018, we show two sets of distributions across grades. In the first set, we assume that the reported grade of students in the Brazilian sample represents additional schooling that resulted from entering schooling a year earlier. These are the actual reported grade for each student in the PISA survey and, as suggested by the means of each grade, show a rapid upward drift in the years of schooling being taken by students in Brazil by the time they are 15 years-old—gradually, those who started first grade are “taking an additional year of schooling” because they entered first grade at 6 years-old rather than 7 years-old. At the same time, this assumes that the additional year of schooling increases exposure to math and reading skills, and is not largely just a “renumbering upward” of the grades. In the second set of grade identifications, we have subtracted a year from the reported grade for each student in the Brazilian sample on the assumption that the reform was primarily a “renaming” of grades and did not, in reality represent more curriculum exposure in school.

Tables 3 (mathematics) and Table 4 (reading) show the estimated coefficients for the control variables and the country dummies (Mexico is the reference country) for all the countries in the group of six that took the PISA test in each year. There were only three years in which

students in all six countries we compare took the test—2009, 2012, and 2018. Only Mexico and Brazil, and Uruguay took the test in all seven PISA test years—2000-2018.

The coefficients of the country dummies represent a “residual effect” of being a student in that country controlling for student family resources, grades which the students were attending when they took the test, student gender, and whether they attended a public or private school. The regression estimates in these two tables use the “uncorrected” grades for Brazilian students—that is, these estimates assume that the reported grades in all years represent the nominal grade in which the sampled students were attending school at the time they took the test. We also estimated individual PISA scores including the Urban variable and interaction between urban and private school. The coefficients of the country dummies hardly changed in this alternative specification of the estimation model, so we do not report the results.

In Table 5, we show the coefficients of the country dummies when we used a grade variable for Brazilian students in the 2012, 2015, and 2018 test surveys which is constructed by subtracting a year from the reported grade. Thus, if a student is designated as attending 10th grade, we defined the student’s grade attended as 9th grade. Comparing the coefficients of the country dummies in Table 5 with the corresponding math and reading coefficients of country dummies in Tables 3 and 4, we observe very small differences in all cases except Brazil. The coefficients in 2012, 2015, and 2018 are much smaller for Brazil than the corresponding coefficients in Tables 3 and 4. This means that Brazilian score corrected for grade attended is higher when we reduce the grade by a year, since Brazilian students are not “penalized” for having studied more years than students in the other Latin American countries.

Table 2. Latin America: Means of Variables Used in Regression Estimates

Variables	2000 M	2000 R	2003	2006	2009	2012	2015	2018
Urban School ^a			0.861	0.903	0.906	0.909	0.892	0.891
Private School	0.183	0.183	0.148	0.193	0.179	0.202	0.181	0.195
School average ESCS ^b	43.347	43.280	-0.995	-0.985	-1.101	-1.102	-1.023	-1.093
Female	0.535	0.539	0.532	0.527	0.520	0.516	0.506	0.508
Individual ESCS ^b	43.363	43.281	-0.995	-0.985	-1.101	-1.103	-1.024	-1.094
Grade 7	0.109	0.108	0.098	0.070	0.044	0.009	0.023	0.025
Grade 8	0.197	0.199	0.199	0.148	0.124	0.075	0.050	0.064
Grade 9	0.407	0.407	0.424	0.374	0.319	0.206	0.199	0.161
Grade 10	0.280	0.279	0.272	0.374	0.469	0.468	0.488	0.546
Grade 11	0.007	0.007	0.007	0.027	0.044	0.228	0.227	0.197
Grade 12	0	0	0.000	0.007	0.000	0.014	0.013	0.007
Grade 6 (Δ Brazil grade)						0	0.01	0.018
Grade 7 (Δ Brazil grade)						0.047	0.034	0.043
Grade 8 (Δ Brazil grade)						0.106	0.078	0.087
Grade 9 (Δ Brazil grade)						0.300	0.319	0.246
Grade 10 (Δ Brazil grade)						0.501	0.515	0.570
Grade 11 (ΔBrazil grade)						0.045	0.043	0.036
Grade 12 (Δ Brazil grade)						0.001	0.001	0.000
Argentina	0.097	0.096		0.135	0.103	0.110		0.124
Brazil	0.617	0.620	0.648	0.489	0.451	0.480	0.458	0.428
Chile	0.045	0.045		0.059	0.050	0.045	0.051	0.043
Mexico	0.184	0.183	0.340	0.308	0.292	0.271	0.368	0.300
Peru	0.056	0.056			0.096	0.086	0.113	0.096
Uruguay			0.012	0.009	0.008	0.008	0.010	0.009

Observations 11,527 20,703 36,193 52,756 78,467 76,049 42,993 45,510

Source: OECD, PISA, publicly available microdata, 2003-2018. Authors' calculations. Note: a. The Urban variable is not available for Peru in 2000, so we did not estimate regressions using the variable for that one year. b. The social class variable is different in 2000 from ESCS, but we could correlate the SES variable in 2000 with the ESCS for later years, and the correlation coefficient is 0.8.

The results of taking these different adjustments to PISA scores into account for comparing the PISA scores for our six Latin American countries and the trajectories of those scores in 2003-2018 are shown in Figures 6, 7, and 8 for PISA mathematics and Figures 9, 10, and 11 for PISA reading. Figure 6 compares scores over time by country assuming that the reported grades for Brazilian students represent their years in school and adjusting the scores for the mean values of control variables *within each year*. This means that the comparison is not adjusted for the possibility that the control variables changed over time—that is, the likelihood that social class and grade attained increased between 2000 and 2018. This would not change the relative position of average test scores among countries in each year, but would change the trajectory of the test scores over time, raising earlier average test scores compared to later scores.

Table 3. Latin America: Estimated PISA Math Scale Scores Adjusted for Country Differences in Student and School SES and Grade Enrolled (“uncorrected” Brazilian scores), by Year, 2003-18.

Estimate of PISA Mathematics Scale Score							
Variables	2000	2003	2006	2009	2012	2015	2018
Private School	23.60*** (6.15)	-0.86 (9.64)	5.71 (5.03)	13.51*** (3.18)	12.32*** (2.93)	7.24* (4.26)	12.22*** (3.66)
School average ESCS	2.28*** (0.20)	48.37*** (5.12)	38.82*** (2.39)	34.30*** (1.69)	30.28*** (1.46)	29.41*** (2.22)	30.96*** (2.08)
Female	-27.53*** (2.85)	-22.79*** (2.18)	-22.70*** (1.48)	-21.38*** (0.95)	-22.23*** (0.96)	-16.93*** (1.44)	-18.14*** (1.30)
Individual ESCS	0.48*** (0.10)	4.14*** (1.15)	5.82*** (0.69)	4.86*** (0.40)	5.28*** (0.34)	7.80*** (0.71)	6.44*** (0.64)
Grade 7	-133.51*** (5.70)	-112.42*** (4.86)	-101.18*** (5.24)	-81.37*** (4.30)	-79.00*** (5.12)	-72.13*** (4.73)	-67.72*** (4.71)
Grade 8	-91.28*** (4.61)	-85.47*** (4.37)	-81.20*** (3.89)	-67.17*** (2.82)	-55.46*** (2.16)	-61.88*** (4.08)	-63.44*** (4.55)
Grade 9	-46.13*** (4.33)	-31.41*** (2.92)	-39.04*** (2.99)	-38.59*** (1.48)	-32.31*** (1.59)	-30.63*** (2.81)	-41.03*** (2.86)
Grade 11	12.29 (9.11)	11.55 (15.50)	-5.34 (3.78)	21.00*** (1.96)	32.85*** (1.48)	34.04*** (2.80)	22.84*** (2.45)
Grade 12			-12.85 (14.28)		48.34*** (4.46)	56.73*** (8.15)	31.04*** (9.22)
Argentina	-14.22*** (5.29)		-44.02*** (4.70)	-55.02*** (2.91)	-39.53*** (2.51)		-34.04*** (3.30)
Brazil	-17.27*** (3.69)	-13.80*** (3.67)	-3.82 (3.30)	-22.23*** (2.52)	-39.72*** (1.84)	-58.40*** (3.77)	-33.00*** (3.02)
Chile	-10.63** (4.36)		-22.44*** (4.44)	-39.22*** (2.82)	-20.55*** (2.55)	-19.65*** (3.67)	-18.58*** (4.03)
Peru	-85.38*** (3.56)			-60.67*** (3.07)	-51.00*** (2.20)	-40.43*** (3.30)	-17.88*** (3.23)
Uruguay		-10.55** (4.71)	-0.14 (3.12)	-11.79*** (2.34)	-6.06** (2.40)	-3.17 (2.99)	8.75*** (3.12)
Constant	312.82*** (8.39)	482.57*** (7.13)	483.37*** (2.97)	495.48*** (2.57)	475.87*** (1.90)	475.57*** (3.47)	471.10*** (3.34)

Observations	11,527	36,912	53,189	78,735	76,235	43,470	45,646
Adjusted R ²	0.43	0.44	0.43	0.42	0.39	0.31	0.36

Source: OECD, PISA, publicly available microdata, 2003-2018. Authors' calculations. Note: Statistical significance: *** p<0.01; ** p<.0.05; * p<0.10

We can estimate a set of adjusted mathematics test scores for all six countries using the means of our control variables in 2018 applied to the coefficients in each year from 2003 to 2018³—this effectively estimates what the level of Latin American PISA math scores would have been in, say, 2003 if the social class and grade attended composition of the PISA sample in 2003 had been the same as in 2018. These trajectories are shown in Figure 7. Finally, Figure 8 shows the trajectory of PISA math scores using the “corrected” Brazilian grades, assuming that the “real” grade is one year less than the reported grade in Brazil in 2012, 2015, and 2018. Figures 9, 10, and 11 present the same set of estimates for adjusted PISA reading scores.

Table 4. Latin America: Estimated PISA Reading Scale Scores Adjusted for Country Differences in Student and School SES and Grade Enrolled (“uncorrected” Brazilian scores), by Year, 2003-18.

Variables	Estimate of PISA Reading Scale Score						
	2000	2003	2006	2009	2012	2015	2018
Private School	14.05*** (4.58)	0.50 (8.06)	3.81 (5.81)	9.82*** (3.10)	10.44*** (3.17)	6.14 (4.60)	12.32*** (3.71)
School average ESCS	2.69*** (0.22)	44.09*** (4.42)	38.16*** (3.13)	38.60*** (1.69)	32.19*** (1.41)	35.73*** (2.28)	36.97*** (2.21)
Female	11.11*** (1.79)	22.39*** (2.22)	26.07*** (1.79)	19.50*** (0.77)	22.03*** (0.90)	12.14*** (1.37)	10.82*** (1.17)
Individual ESCS	0.36*** (0.07)	4.24*** (1.35)	6.10*** (0.79)	5.65*** (0.46)	5.13*** (0.40)	6.31*** (0.80)	6.10*** (0.57)
Grade 7	-121.55*** (5.10)	-114.73*** (6.64)	-117.39*** (5.44)	-104.62*** (4.97)	-96.41*** (7.22)	-83.84*** (5.05)	-80.26*** (4.10)
Grade 8	-85.03*** (3.97)	-84.14*** (4.83)	-92.70*** (4.93)	-82.66*** (3.10)	-67.72*** (2.52)	-68.35*** (4.25)	-70.05*** (3.39)
Grade 9	-46.34*** (3.16)	-33.25*** (3.40)	-38.80*** (2.88)	-42.88*** (1.81)	-37.81*** (1.88)	-33.49*** (2.75)	-42.29*** (3.00)
Grade 11	17.32*** (6.44)	10.00 (14.52)	-2.03 (3.32)	19.68*** (2.08)	33.12*** (1.53)	36.75*** (2.78)	26.53*** (2.32)
Grade 12			-12.36 (11.35)		47.82*** (4.67)	62.81*** (7.08)	39.42*** (8.51)
Argentina	-19.28*** (4.53)		-53.87*** (5.25)	-54.45*** (2.89)	-41.84*** (2.86)		-24.66*** (3.26)
Brazil	5.98* (3.27)	18.66*** (4.31)	17.13*** (3.94)	-1.55 (3.14)	-32.66*** (2.19)	-45.76*** (3.31)	-16.60*** (2.95)
Chile	-16.04*** (3.83)		7.19 (4.95)	-19.89*** (3.17)	-13.13*** (2.59)	-2.24 (3.50)	2.02 (4.08)
Peru	-86.26*** (3.58)			-61.15*** (2.78)	-45.19*** (2.78)	-45.91*** (3.19)	-29.71*** (3.22)

³ It is not possible to apply the 2018 means of covariates to the 2000 coefficients because the measure of social class in 2000 is different from the ESCS, which is not available for that year—it only began being estimated in 2003. However, it is highly likely that the 2000 results for all countries would increase substantially if we were able to apply the 2018 means to the 2000 coefficients of the control variables.

Uruguay		-9.94** (4.64)	-17.75*** (3.89)	-21.59*** (2.48)	-13.99*** (2.86)	-2.09 (2.99)	6.46* (3.52)
Constant	316.32*** (8.44)	469.55*** (6.05)	463.73*** (3.81)	491.27*** (2.27)	468.41*** (2.15)	483.78*** (3.62)	475.58*** (3.55)
Observations	20703	36,910	53,188	78,734	76,235	43,470	45,619
Adjusted R ²	0.44	0.34	0.38	0.44	0.38	0.32	0.36

Source: OECD, PISA, publicly available microdata, 2003-2018. Authors' calculations. Note: Statistical significance: ***: p<0.01; **: p<.0.05; *: p<0.10

One big surprise in these estimated adjusted trajectories is that Chilean student PISA mathematics test scores—once adjusted for the higher social class of the Chilean PISA sample (both individual social class and average school social class)—are lower than test scores in both Mexico and Uruguay and higher in Brazil only when no correction is made for the possible “renaming” increase of grades in Brazil. Chile is seen as the “star” of Latin American education (along with Uruguay) in the PISA literature (see OECD, 2016). These results suggest that the “high” (in Latin America) Chilean scores are largely the result of higher average levels of family resources and higher levels of student attainment in the Chilean PISA sample rather than the result of a more effective school system.

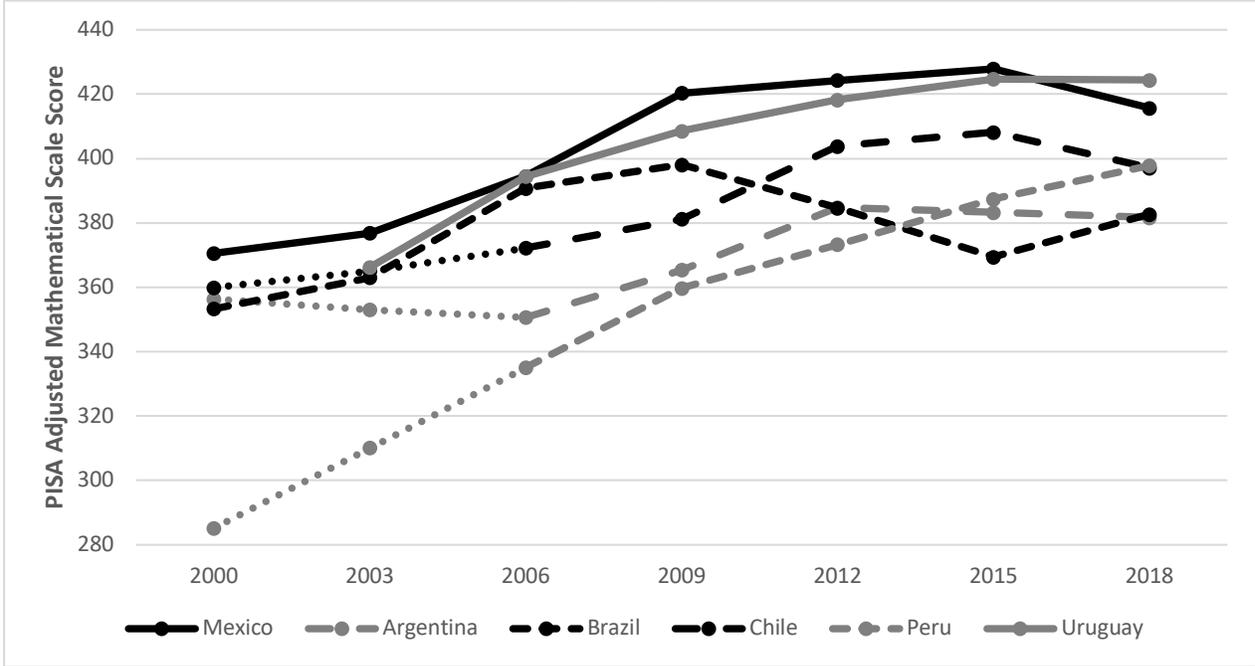
The second big surprise is the relatively low gains in PISA mathematics scores in several of these countries over the first two decades of the 21st century despite a major focus on improving mathematics test scores in the region during this period. Taking the results of Figure 7 as controlling for changes in social class and grade attended in the PISA samples over time, we observe that in Argentina, Chile, and Mexico, students increased their PISA math scores about 10-15 points in 12-15 years—0.10 to 0.15 SD. The math test scores in Brazil increased much more((about 30 points) when we assume that Brazilian students attended more years of schooling in later PISA years in “name only,” and students in Uruguay and Peru made substantial gains even when we correct for changes in control variables over the period 2003-2018.

Table 5. Latin America: Estimated PISA Mathematics and Reading Scale Scores Adjusted for Country Differences in Student and School SES and Grade Enrolled (“corrected” Brazilian scores), Country Coefficients, by Year 2012-2018.

Country	Mathematics			Reading		
	2012	2015	2018	2012	2015	2018
Argentina	-37.94*** (2.47)		-33.87*** (3.24)	-40.19*** (2.83)		-24.64*** (3.24)
Brazil	-8.04*** (1.88)	-26.44*** (3.30)	-5.55* (3.06)	1.92 (1.99)	-10.59*** (3.24)	14.29*** (3.04)
Chile	-20.00*** (2.54)	-19.87*** (3.64)	-18.47*** (3.96)	-12.13*** (2.57)	-2.32 (3.47)	2.33 (4.01)
Peru	-47.34*** (2.16)	-37.95*** (3.27)	-15.64*** (3.24)	-40.63*** (2.85)	-42.45*** (3.18)	-26.74*** (3.24)
Uruguay	-5.33** (2.41)	-2.20 (2.94)	9.18*** (3.13)	-13.29*** (2.80)	-1.04 (2.96)	6.72* (3.45)

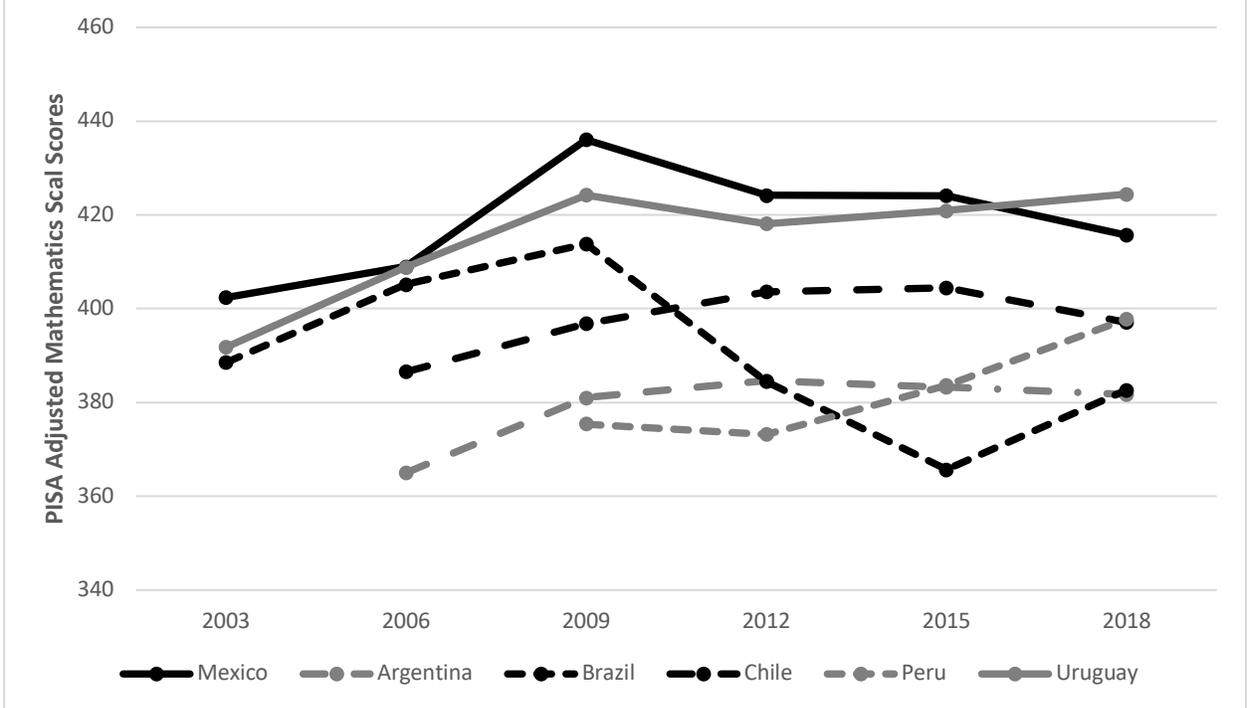
Source: OECD, PISA, publicly available microdata, 2012-2018. Authors' calculations. Note: Statistical significance: *** p<0.01; ** p<.0.05; *:p<0.10

Figure 6. Latin America: PISA Mathematics Scale Scores Adjusted for Means of Control Variables within Each Year, No Correction for Reported Brazilian Grades, 2000-2018



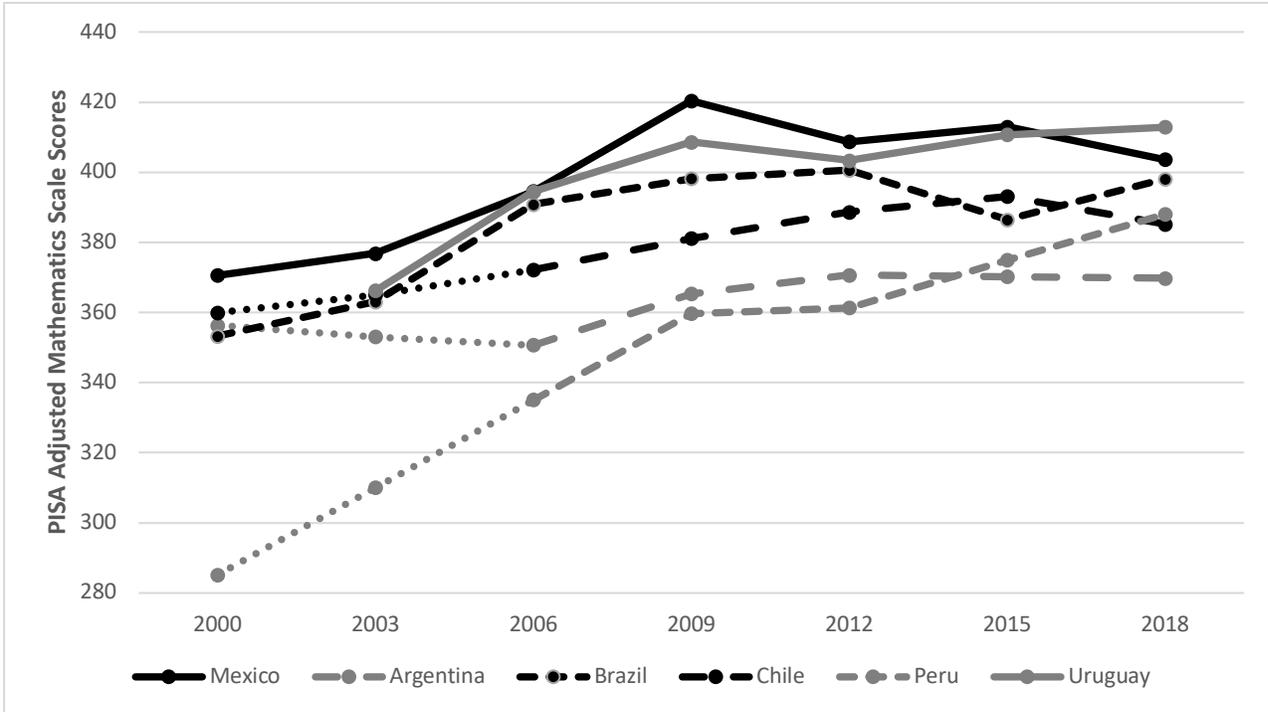
Source: Table 3 and authors' calculations.

Figure 7. Latin America: PISA Mathematics Scale Scores Adjusted for Means of Control Variables in 2018, No Correction for Reported Brazilian Grades, 2003-2018



Source: Table 3 and authors' calculations.

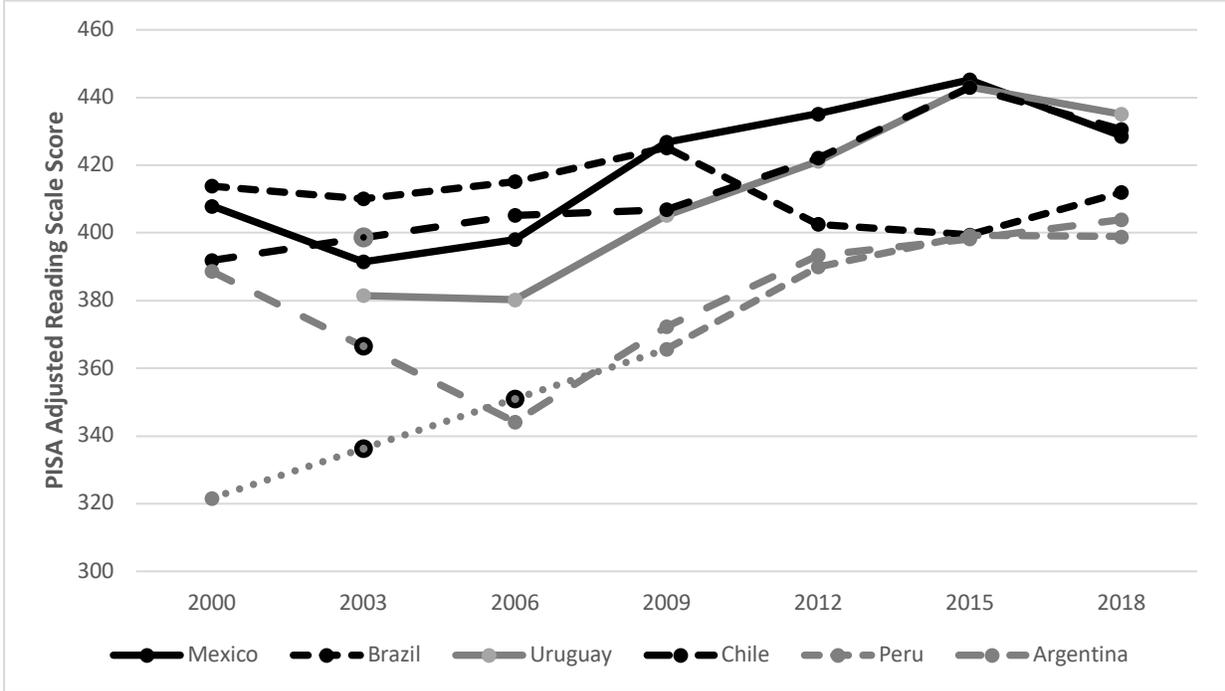
Figure 8. Latin America: PISA Mathematics Scale Scores Adjusted for Means of Control Variables within Each Year, Correcting Reported Brazilian Grades, 2000-2018.



Source: Table 3, Table 5, and authors’ calculations.

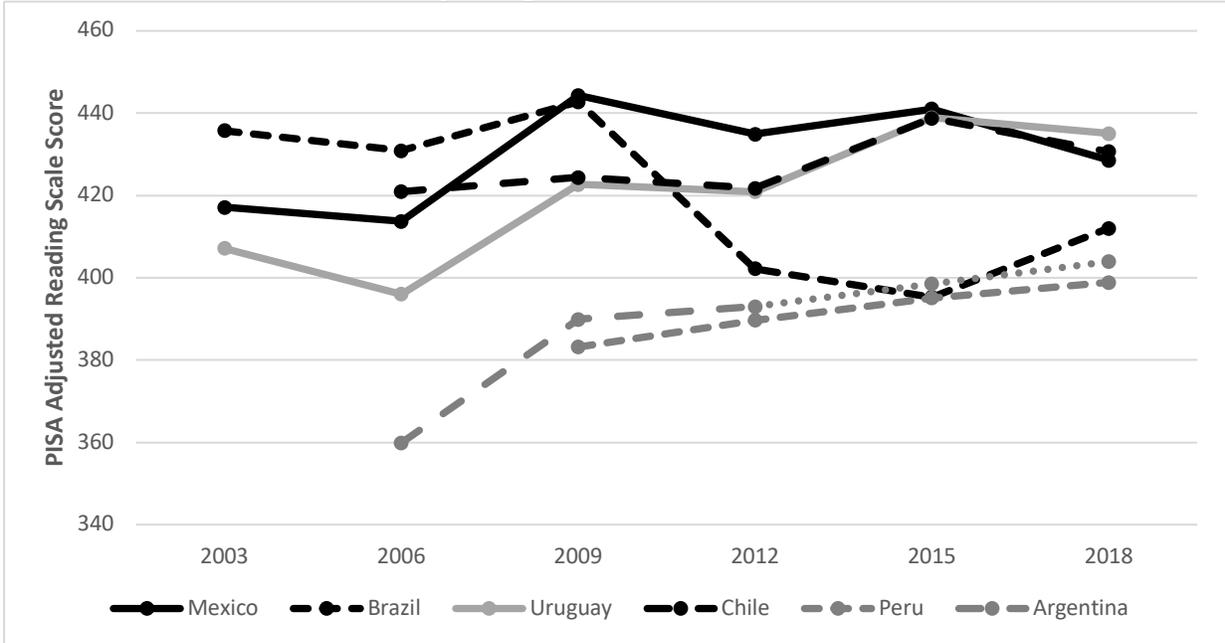
The results for PISA reading scores are somewhat different, but many of the conclusions are the same. We can observe one group of countries—Mexico, Uruguay, and Chile—that have converged in equal higher average adjusted scores in 2015-2018. The gains are highest in Uruguay among this group (Figure 9) even when we adjust the gains for the change in social class and grade distribution in the samples by using the 2018 means for our adjustments across years (Figure 10). Again Chile’s advantage in raw scores is eliminated by our adjustments—mainly because of the higher social class of students in the Chile sample. If we correct Brazilian scores for the upward move in grade distribution in 2012-2018, Brazil joins the high scoring group. If the increase in grade attended reflects possible greater real time in school, adjusted Brazilian scores drop in 2012-2018. Students in Peru and Argentina have lower PISA reading scores, but Peru and Uruguay have made considerable reading gains during this period even when we adjust for changes in student social class and grade attended during this period (Figure 10).

Figure 9. Latin America: PISA Reading Scale Scores Adjusted for Means of Control Variables within Each Year, No Correction for Reported Brazilian Grades, 2000-2018



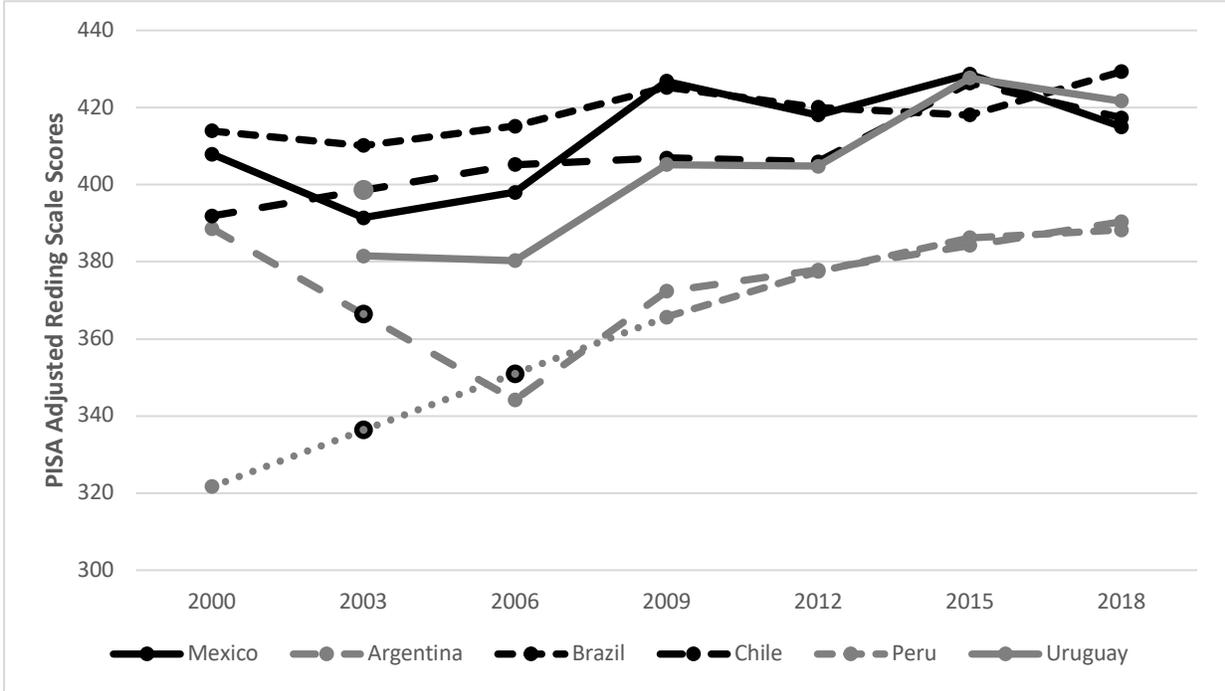
Source: Table 4 and authors' calculations.

Figure 10. Latin America: PISA Reading Scale Scores Adjusted for Means of Control Variables in 2018, No Correction for Reported Brazilian Grades, 2003-2018



Source: Table 4 and authors' calculations.

Figure 11. Latin America: PISA Reading Scale Scores Adjusted for Means of Control Variables within Each Year, Correcting Reported Brazilian Grades, 2000-2018.



Source: Table 4, Table 5, and authors’ calculations.

Comparisons among Mexican states

One country in Latin America, Mexico, took equal random PISA samples in each state from 2003 to 2012. This allows us to analyze the results across states to estimate whether there exist significant differences in student performance on PISA among states and whether some states made larger gains than others, accounting for demographic and family academic resources differences among states, cross-sectionally and at different points in time. These state comparisons are similar to the comparisons we made among Latin American countries in the previous section.

In the case of Mexican states, we estimated the individual student’s scale score in the PISA mathematics test as a function of a set of dummy variables for the OECD PISA social class variable (ESCS), student age, gender, language spoken at home other than Spanish, and the grade enrolled, whether the student had attended pre-school, the average ESCS in the school the student attended, and whether that school was private or public, and in a rural area, village, or large city. Finally, the estimated PISA scale score estimate included a set of dummy variables for each state, with Nuevo Leon as the reference state. The standard errors of estimate of the coefficients are corrected for the clustering of the variables at the state level by including a cluster correction.

Our coefficients of interest are the coefficients for each state, which show how much of a difference in mathematics scale score attending school in each state (relative to Nuevo Leon) made in each year after controlling for all those other variables. It could be argued that these state coefficients provide a more accurate estimate of educational quality differences among the states' public education than the differences in raw score, since we have controlled for observable differences in parents' academic resources, the degree to which students have been exposed to mathematics curriculum (grade attended), peer effects (average school ESCS) the possibly greater difficulty of delivering schooling in small villages, and whether a student attended pre-school and public or private school. However, these state coefficients just represent a *closer* approximation than the raw mean to the actual impact of public schools on PISA mathematics score. We did similar estimates for the PISA reading score, which generally showed smaller differences among states than on mathematics scores, but only show the results here for mathematics (for those readers interested in the regression results, please contact the authors).

The mean mathematics score for the reference state, Nuevo Leon, is equal to the sum of the coefficients of all the control variables weighted by the means of the variables in the sample and added to the intercept term. *We do this weighting separately for each year.* In Table 6, we present the mean mathematics scores in each state adjusted for the control variables in each of the four years, 2003, 2006, 2009, and 2012. Thus, for example, in 2003, the 15 year-old students in the Nuevo Leon sample scored 389.5 on the PISA mathematics test and 414.4 in 2012. However, this gain of 24 points may in part be the result of changes in the social class and grade attended at the time of the test in the 2003 and 2012 sample. If we adjust the 2003 results for such changes, students in Nuevo Leon in 2003 with similar family background and grade attended would have scored 404.1 on the 2003 PISA mathematics test. Thus the adjusted net improvement in math score resulting from factors associated with "better" schooling in Nuevo Leon rather than having been in school longer at age 15 or having greater family resources is 10.3 points, and since all other state scores are estimated relative to Nuevo Leon, this reduces the net gain in every state possibly due to better education by 14.6 points.

Students in most states in the country, who in 2003 performed on average at the highest and lowest ends of relative PISA mathematics scores (adjusted for students' family resources and grade of school), also performed at the highest and lowest ends in 2012. The consistently higher end performing state school systems were Chihuahua, Queretaro, Jalisco, and Guanajuato. The consistently low-end performing state school systems were Tabasco, Guerrero, Chiapas, and Baja California Sur. More interesting were the states where students made big gains or small gains in PISA math score relative to the Mexican average gain during this period. We can divide such states into five different groups and show the gains under no adjustment for the social class and grade attended composition of the PISA sample and under an adjustment for an increase in social class and grade attainment (see Table 7a and 7b).

- Group 1 is initially high-scoring states—Queretaro and Jalisco—that made high gains in the 9-year period.
- Group 2 is initially middle-scoring states that made high gains—Puebla (the largest gain) and Veracruz.
- Group 3 is initially low-scoring states that made high gains—Baja California Sur, Yucatan, Tlaxcala, and Campeche

- Group 4 is initially high-scoring state that made low gains—Quintana Roo, Distrito Federal, and Hidalgo.
- Group 5 is initially low-scoring states that made low gains—Tabasco.

Table 6. Mexico: Adjusted Standardized Student PISA Mathematics Achievement Scale Scores, by State, 2003-2012 (no adjustment for change in covariate means between 2003 and 2012.

State	2003	State	2006	State	2009	State	2012
Tabasco	358.7	Guerrero	372.4	Tabasco	385.4	Tabasco	379.1
Guerrero	362.8	Tabasco	380.0	Coahuila	401.1	Guerrero	395.3
Campeche	364.3	Baja California	386.0	Baja California Sur	402.0	Chiapas	398.2
Tlaxcala	364.6	Tamaulipas	388.1	Tamaulipas	402.9	Sinaloa	402.7
Baja California Sur	366.0	Baja California Sur	389.1	Chiapas	403.8	Baja California Sur	402.9
Chiapas	367.6	Coahuila	390.5	Guerrero	406.3	Tamaulipas	403.5
Sinaloa	370.2	Nayarit	391.5	Michoacan	409.9	Baja California	404.7
Coahuila	373.7	Sonora	393.0	Sinaloa	410.0	Michoacan	405.7
Baja California	375.5	Michoacan	393.9	Durango	410.1	Coahuila	406.8
Yucatan	376.8	Sinaloa	395.6	San Luis Potosi	411.0	Nayarit	408.8
Mexico	378.6	Chihuahua	395.7	Sonora	412.7	Tlaxcala	409.3
Sonora	378.7	Chiapas	396.2	Tlaxcala	414.5	Mexico	409.6
Nayarit	378.7	Zacatecas	397.9	Baja California	415.0	Campeche	409.7
Morelos	380.0	Puebla	397.9	Campeche	416.6	Distrito Federal	412.1
Tamaulipas	380.3	Campeche	400.5	Morelos	417.9	Zacatecas	412.4
Puebla	380.5	Tlaxcala	400.8	Aguascalientes	419.5	Quintana Roo	412.8
Durango	382.2	Durango	401.5	Quintana Roo	419.9	Sonora	414.4
Colima	384.5	Quintana Roo	403.3	Yucatan	420.1	Nuevo Leon	414.4
Veracruz	384.9	Colima	406.2	Puebla	422.0	Hidalgo	414.8
Zacatecas	387.9	Nuevo Leon	407.3	Chihuahua	422.3	Morelos	415.1
Michoacan*	NA	Distrito Federal	408.1	Mexico	422.4	Colima	415.9
Nuevo Leon	389.5	Aguascalientes	408.2	Nuevo Leon	423.2	Durango	417.6
Guanajuato	390.0	San Luis Potosi	411.3	Jalisco	423.5	Yucatan	420.0
San Luis Potosi	390.6	Morelos	411.9	Nayarit	424.8	Aguascalientes	421.1
Distrito Federal	392.2	Jalisco	412.7	Hidalgo	424.8	Veracruz	421.7
Chihuahua	392.4	Guanajuato	413.0	Zacatecas	425.3	Chihuahua	421.9
Aguascalientes	392.8	Yucatan	415.3	Distrito Federal	427.7	Guanajuato	423.0
Hidalgo	394.0	Mexico	415.9	Guanajuato	430.2	San Luis Potosi	424.1
Jalisco	394.7	Hidalgo	416.4	Veracruz	432.6	Puebla	430.8
Queretaro	394.8	Veracruz	416.4	Colima	434.2	Queretaro	434.9
Quintana Roo	395.0	Queretaro	435.0	Queretaro	435.7	Jalisco	436.1

Source: OECD, PISA, publicly-available microdata—authors’ calculations. Notes: *Michoacan did not participate in the 2003 PISA. Numbers in red indicate that residual is significantly different from Nuevo Leon, the reference state. Reference variables: state—Nuevo Leon; grade attended—grade 10; preschool—more than one year; language at home—language of test; community size—town.

Thus, an important policy implication of these results is whether we can identify strategies that states in Groups 1, 2, and 3 employed in the period before or after 2003 that might have contributed to these relatively higher math gains and whether we can contrast them with conditions in state in Groups 4 and 5 that contributed to relatively low gains in those states—Group 4 is especially interesting because those three states started out relatively high in the state rankings and fell considerably. Another consideration is the finding of about a third of the states in Mexico with low gains in mathematics during this nine year period but a sizable group with rather high gains (more than 0.2 SD) even when we adjust for increases in social class and grade

attended at 15 years-old. It is important to understand what may have happened differently in these two groups of states.

Table 7a. Mexico: Gains, 2003-2012, in Adjusted Standardized Student PISA Mathematics Achievement Scores, by State (scale points), Not Adjusting for Changes in Mean Social Class, Grade Attended, and Other Covariates Between 2003 and 2012

State	Gain 2003-2012	State	Gain 2003-2012
Michoacan	NA	Sinaloa	32.54
Quintana Roo	17.72	Guanajuato	32.93
Distrito Federal	19.86	Coahuila	33.17
Tabasco	20.33	San Luis Potosi	33.57
Hidalgo	20.78	Morelos	35.15
Tamaulipas	23.16	Durango	35.45
Zacatecas	24.55	Sonora	35.68
Nuevo Leon	24.9	Baja California Sur	36.85
Aguascalientes	28.28	Veracruz	36.87
Baja California	29.16	Queretaro	40.12
Chihuahua	29.58	Jalisco	41.46
Nayarit	30.04	Yucatan	43.19
Chiapas	30.61	Tlaxcala	44.67
Mexico	30.99	Campeche	45.42
Colima	31.46	Puebla	50.34
Guerrero	32.5		

Source: OECD, PISA, publicly available microdata for Mexico, 2003-2012. Authors' calculations.

Table 7b. Mexico: Gains, 2003-2012, in Adjusted Standardized Student PISA Mathematics Achievement Scores, by State (scale points), Adjusting for Changes in Mean Social Class, Grade Attended, and Other Covariates Between 2003 and 2012

State	Gain 2003-2012	State	Gain 2003-2012
Michoacan	NA	Sinaloa	17.99
Quintana Roo	3.17	Guanajuato	18.38
Distrito Federal	5.31	Coahuila	18.62
Tabasco	5.78	San Luis Potosi	19.02
Hidalgo	6.23	Morelos	20.6
Tamaulipas	8.61	Durango	20.9
Zacatecas	10	Sonora	21.13
Nuevo Leon	10.35	Baja California Sur	22.3
Aguascalientes	13.73	Veracruz	22.32
Baja California	14.61	Queretaro	25.57
Chihuahua	15.03	Jalisco	26.91
Nayarit	15.49	Yucatan	28.64
Chiapas	16.06	Tlaxcala	30.12
Mexico	16.44	Campeche	30.87
Colima	16.91	Puebla	35.79
Guerrero	17.95		

Source: OECD, PISA, publicly available microdata for Mexico, 2003-2012. Authors' calculations.

Brief Conclusions

We have presented considerable data on PISA test scores in Latin America with the intent of drawing attention to whether educational quality is improving or not, and if so, which countries (and in Mexico, which states) are leading the way in such improvement. We have argued that to get at estimates of changing and comparative educational quality through these PISA scores we need to account for the fact that the social class composition of students and schools in the PISA samples varies across countries and over time. We know that students' family background influences PISA scores both at the individual and school levels, and that much of this influence is independent of the quality of classroom teaching or other measures of what could be called educational quality. We also argued that because the PISA tests 15 year-olds, not students in a given grade, we also need to account for the fact that students in PISA samples in Latin America vary across countries and across PISA test years in the grade they were attending when they took the PISA test.

Of course, quantity of schooling (time in school) affects how much students learn and could be considered as a measure of school quality, but the most prevalent way that school quality or effectiveness is defined is in terms of how much students are learning for the same amount of time in school. Thus, to compare quality of schooling among Latin American countries and to compare how this quality may have changed over time, we argue, means comparing how well students of similar social class background attending the same number of grades and attending schools with similar social class students perform on the PISA test in the six countries we study and how their performance may have changed over time controlling for any changes in social class and grade attended that may have occurred over the past 15 to 18 years.

Having done this, we find that students in two countries—Peru and Uruguay—made important gains in mathematics and reading in 2003-2018, even adjusting for changes in the social class and grade composition of the PISA student samples in this period. Students in Argentina also made quite large gains in reading in these years. The adjusted mathematics and reading scores suggest that the quality of mathematics taught in Mexican and Uruguayan schools is higher than in other countries and that the quality of Mexican, Uruguayan, and Chilean teaching of reading is higher than in Argentina and Peru.

In many ways, the results showing Mexican education to be of relatively high quality in the Latin American context may be surprising, given the politicization of the Mexican teachers' union and the resulting difficulties of reforming the system. Further, according to our results, PISA mathematics and reading scores have improved little in Mexico in 2003-2018 once we adjusted for changes in the social class and grade composition of the PISA samples. This makes our comparison of states within Mexico even more relevant, since the results of those estimates suggest considerable variation in adjusted students' gains across states, with students in some making quite large gains and students in other states, rather negligible gains. Elsewhere, we have found a similar variation across Brazilian states on the Brazilian national test (Carnoy et al, 2017). We suspect that were test score data available for provinces in another federal country, Argentina, we might also find considerable variation there.

REFERENCES

- Álvarez, J., Moreno, V. G., & Patrinos, H. A. (2007). *Institutional effects as determinants of learning outcomes: Exploring state variations in Mexico* (Vol. 4286). World Bank Publications.
- Bellei, C. (2009). Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile. *Economics of Education Review*, 28(5), 629-640.
- Bellei, C., & Vanni, X. (2015). Chile: The evolution of educational policy, 1980-2014. In Schwartzman, S. (ed.) *Education in South America*, London: Bloomsbury Publishing, 179-200.
- Carnoy, M., & Marshall, J. (2005). Cuba's academic performance in comparative perspective. *Comparative Education Review*, 49(2), 230-261.
- Carnoy, M., Marotta, L., Louzano, P., Khavenson, T., Guimarães, F. R. F., & Carnauba, F. (2017). Intranational Comparative Education: What State Differences in Student Achievement Can Teach Us about Improving Education—the Case of Brazil. *Comparative Education Review*, 61(4), 726-759.
- Carrasco, R. (2014). Leveling the playing field: How can we address educational inequalities? Unpublished PhD dissertation, Stanford University Graduate School of Education.
- Deutsch, J., Dumas, A., & Silber, J. (2013). Estimating an educational production function for five countries of Latin America on the basis of the PISA data. *Economics of Education Review*, 36, 245-262.
- Flores-Mendoza, C., Ardila, R., Gallegos, M., & Reategui-Colareta, N. (2021, March). General intelligence and socioeconomic status as strong predictors of student performance in latin american schools: Evidence from pisa items. In *Frontiers in Education* (Vol. 6, p. 632289). Frontiers Media SA.
- Hanushek, E. A., & Kimko, D. D. (2000). Schooling, labor-force quality, and the growth of nations. *American economic review*, 90(5), 1184-1208.
- Hanushek, E. & Woessmann, L. (2008). The role of cognitive skills in economic development. *Journal of Economic Literature*. 46 (3), 607-668.
- Jencks, C. & Phillips, M. (1998). *The Black-White Test Score Gap*. Washington, DC: The Brookings Institution.
- Klein, R. (2011). Uma re-análise dos resultados do Pisa: Problemas de comparabilidade. *Ensaio: Avaliação e Políticas Públicas em Educação*, 19(73), 1-20..

- Levinson, B. A., Blackwood, J., & Cross, V. (2013). Recipients, agents, or partners?: The contradictions of teacher participation in Mexican secondary education reform. *Journal of Educational Change*, 14(1), 1-27.
- Marchionni, M., & Vazquez, E. (2019). The causal effect of an extra year of schooling on skills and knowledge in Latin America. Evidence from PISA. *Assessment in Education: principles, policy & practice*, 26(4), 489-515.
- Mizala, A., & Torche, F. (2017). Means-tested school vouchers and educational achievement: Evidence from Chile's universal voucher system. *The Annals of the American Academy of Political and Social Science*, 674(1), 163-183.
- Murnane, R. J., Waldman, M. R., Willett, J. B., Bos, M. S., & Vegas, E. (2017). *The consequences of educational voucher reform in Chile* (No. w23550). National Bureau of Economic Research.
- OECD, PISA. (2013). *Lessons from the PISA 2012 for the United States*. Paris: OECD.
- OECD (2016). *Skills in Ibero-America. Insights from PISA 2012*. Paris: OECD.
- Ornelas, C. (2016). *El sistema educativo mexicano: la transición de fin de siglo*. Fondo de Cultura Económica.
- Rosa, L., Bettinger, E., Carnoy, M., & Dantas, P. (2022). The effects of public high school subsidies on student test scores: The case of a full-day high school in Pernambuco, Brazil. *Economics of Education Review*, 87, Article 102201.
- Treviño, E., Valenzuela, J. P., & Villalobos, C. (2016). Within-school segregation in the Chilean school system: What factors explain it? How efficient is this practice for fostering student achievement and equity? *Learning and individual differences*, 51, 367-375.
- UNESCO (2005). *Education for all: The quality imperative*. EFA Global Monitoring Report. Paris: UNESCO
- Willms, J. D. (2010). School composition and contextual effects on student outcomes. *Teachers College Record*, 112(4), 1008-1037.