EXPLORING THE GAP DIFFERENCE IN 2000-2009 PISA TEST SCORES BETWEEN ARGENTINA, CHILE AND MEXICO[1]

GIGENA, Manuel
VERA, María L.
GIULIODORI, Roberto F.
GERTEL, Héctor R[*]

Abstract

Student reading proficiency scores in the PISA 2000 study were, on average, equally poor in Argentina, Chile and Mexico. By contrast, important differences were observed in PISA 2009. In this paper, the 2000-2009 difference is decomposed into coefficient and covariate effects by applying Oaxaca-Blinder decomposition techniques. Decomposing the total gap into characteristic and coefficient effects showed striking country differences. In Argentina, both effects were moderate, had similar weight and negative sign. In Mexico these effects were moderate, similar in absolute value, and opposed in sign; and in Chile these effects were high, had similar value and positive sign. This paper seeks to examine potential factors of policy interest explaining divergences in trajectories between these countries.

JEL Code: I21, H52
Keywords: PISA Study, Latin America, Oaxaca-Blinder decomposition

1 - Introduction

Considerable effort has been exercised in the construction of tests that made students' performance in the major areas of cognitive abilities internationally comparable. These types of programs rapidly attracted the attention of growth-economists intrigued by the fact that education, when it was measured as school attainment of the population, resulted to be a poor predictor of cross-country growth differences. In this vein, recent empirical work done by economists such as Hanushek and others[2] has contributed evidence pertaining to the powerful effects of educational quality, measured by what people know, on individual earnings, on the distribution of income, and on economic growth. What students achieve in High School has produced in the Latin American countries immediate concerns about the future of their economies because the countries in this region have consistently performed at the bottom of any international test of cognitive abilities despite their long standing achievements in extending school coverage of the population[3].

[1] This research was supported by a grant from the FONCyTProgramme (Ministry of Science and Technology Innovation, Argentina) under PICT 2007 Grant #803

[*] Facultad de Ciencias Económicas Universidad Nacional de Córdoba – Argentina, e-mail: hgertel@gmail.com

[2] The idea is developed in Hanushek and Kimko (2000), Hanushek and Luque (2003), and Hanushek and Woessmann, 2011: 124 - 142. Measures of quantity of education were earlier introduced in growth accounting by Barro and Lee (1993) while Breton (2011) identified complementarity of school quality and school attainment on the rate of economic growth.

[3] The growth accounting literature provides evidences pointing at the fact that low quality education that fails to render citizens functionally literate has the potential to curb long-term

Notwithstanding this general observation, recent data collected through the Program for International Student Assessment (PISA) shows that, within Latin America, the performances of individual countries have moved between 2000 and 2009 in different directions: i.e. while some countries that participated in both rounds of PISA exhibited an important improvement (Chile), in others the average score has remained quite stable (Mexico) or even decreased (Argentina). Thus, a relevant question of immediate research interest that captured our attention is: what made a difference for student achievement in Latin America countries, between 2000 and 2009?

In this study, we use the well-established Oaxaca-Blinder technique (Blinder 1973, Oaxaca 1973, Elder et.al. 2010), of decomposing inter-group differences in mean levels of an outcome into those due to different observable characteristics or "endowments" across groups and those due to different effects of characteristics or "coefficients" of groups to identifying the underlying causes of time-period disparities in performance of individual countries in the reading test of PISA.

This study is structured as follows: major features for the three countries under analysis are presented in Section 2. Section 3 outlines the educational production function that regress PISA scores on a reduced number of explanatory factors, indicates the decomposition method included for studying time-period disparities in mean estimates of PISA results by country, and discusses aspects of the PISA databases relevant to this study. Section 4 presents the results of the estimated model and section 5 concludes.

2–Countries Characteristics

Table 1 presents the basic GDP, and PISA statistics by country. The first column shows the GDP per capita for the three countries in 2009. Corresponding long-term annual rates of growth are shown in the next 2 columns. It can be appreciated that GDP per capita of Chile grew at a 2.4 % per annum between 1960 and 2009 while Argentina grew slower, at a modest 1.3% between 1960 and 2009 with 2.7% in 2000-2009.The Mexican economy grew at about 2% annually. At these rates, they can expect to double their GDP in about 30, 54 and 37 years, respectively. A slightly more optimistic result is obtained, however, when using the 1999-2009 rates.

Table 1. GDP, School life expectancy and reading score in PISA by country

Country	GDP per cápita 2009	Annual growth rate		School life expectancy		Reading score in PISA		Students below level 2 (in %)	
		1960-2009	1999-2009	1999	2008	2000	2009	2000	2009
Argentina	11961.3	1.3	2.7	14.4	15.6	418	398	43.9	52.0
Chile	11998.8	2.3	3.2	12.8	14.5	409	449	48.2	31.0
Mexico	11629.6	1.9	2.0	11.9	13.7	422	425	44.1	40.0
Latin America	10117.2	1.8	2.5	12.6	13.7	394	413	...	...

Source: Penn World Table, UNESCO report 2011 (table 4 and 7) and PISA 2009 executive summary.

economic growth.Hanushek and Woessmann (2009) made explicit that whereas high schooling levels in Latin America do not explain why the area has trailed other regions in terms of economic development for the last four decades, the low cognitive capabilities of its population do.

Argentina exhibits the highest school life expectancy value, being the only country of the 3 where an average student could have expected to be at school by age 15 in 1999. Yet, the higher selectivity found in the Chilean and the Mexican systems decreased between 1999 and 2009. Next, the results in PISA (reading test) are presented. They cast a shadow on the growth perspectives of Argentina, by reporting a fallback of about 5% in the reading capabilities of its students. The last two columns indicate the proportion of students in each country performing at the bottom of the scores distribution. In fact, last year's release of PISA 2009 results brought different conclusions among education experts in Argentina. While some considered how reading performance improved in the country (compared to PISA 2006) or worsened (against PISA 2000), others focused in the backlash suffered by Argentina compared to the rest of Latin America. As one of the aspects of PISA is international comparability, we found it more interesting to explore this second issue because mean score decreased from 418 to 398 points in Argentina, increased from 409 to 449 points in Chile and was almost the same in Mexico (422 and 425 points), and the interesting question that arises is what makes their paths different.

Possible sources of variation, to be explored in this study, are: changes in quantity of factors affecting outcome, changes in the productivity of these assets and changes in the overall context of the country.

A framework based on Blinder (1973) and Oaxaca (1973), according to which the mean difference in outcome between two groups can be decomposed into a characteristics and a coefficient effect[4] is adopted to analyze the relative importance of each of these sources of variation. In our model "characteristics" would mean the set endowments that contribute to the learning capability of students and "coefficients" measure their returns[5] or productivity in production. We then proceed to partition the mean total difference between 2000 and 2009 outcomes in the three countries into endowments and coefficient, or productivity effects. The coefficient effect can, in turn, be broken-down into marginal effects associated with productivity levels of endowments and fixed effects that capture the overall contribution of the institutional environment.

3– Database and Theoretical Model

In this paper we use reading test scores from the PISA databases as measure of academic attainment of Argentinean, Chilean and Mexican students in 2000 and 2009[6]. Additionally, the Student and the School principal databases provided information about

[4]The methodology is presented in section 3.

[5] In education production functions, the estimated coefficients represent the returns to student's characteristics (the change in the mean expected score of a student with average characteristics derived from a marginal change in a covariate) plus a fixed effect caught by the intercept, accounting for the state of technology and overall contribution of institutions unrelated to the covariates. Because technology evolves quite slowly in education, the fixed-effect captured in our model would be mostly reflecting the overall contribution of the institutional environment.

[6]PISA 2009 is the fourth edition of the assessment developed by UNESCO and OECD to evaluate 15 year olds' preparation to cope with real-life challenges related to three subjects: reading, mathematics and science. PISA was first implemented in the year 2000 and has been continued every three years, each time focusing on a different subject. In total, 43 countries participated in the first assessment and 74 in the 2009 edition. Argentina and Chile participated in all but the 2003 edition, while México participated in all four editions.

main student and family characteristics and about the public/private characteristic of the school that affect results.

An important issue throughout the discussion of differences in results between two populations (i.e. students in PISA 2000 and 2009) has been the relationship between test scores, the student, family and school endowments and the effectiveness with which these endowments are transformed into results. Our model for the analysis of differences includes a linear education production function and the Oaxaca-Blinder decomposition for the linear model.

Consider firstly, a linear regression model[7], estimated separately for the groups "a" (2009) and "b" (2000) of PISA students in Argentina, Chile and Mexico, respectively.

$$R_{ij} = \beta_0 + \beta_F F_{ij} + \beta_S S_{ij} + \beta_P P_j + \varepsilon_{ij}$$

Where R_{ij} is the reading plausible value[8] of student i in school j, F_{ij} is the vector of family background variables, S_{ij} is the vector of student individual variables, P_j is the school variable used as control and ε_{ij} is the error term (Wooldridge, 2002).

The constant $\beta_{0.}$ captures the fixed effect of institutions not accounted for by the covariates while β_F, β_S and β_P measure the impact of the covariates included in the regression. Improved OLS was applied to gain efficiency. More specifically, due to the two-way stratified sample design in PISA a survey regression and balanced repeated replication (BRR) replicate weights are used. This improved technique provides us with estimates of the effect of each of the explanatory variables at the conditional mean of the scores' distribution (PISA 2009).

This paper uses a reduced set of independent variables that capture the principal features of schools, family background and student endowments affecting results. Student variables are: repeated course[9], most likely to negatively affect results, and enjoy reading, likely to be positively associated with scores —interpreted as summary of previous effective coaching and home incentives. Family background endowment is taken into account by these variables: Highest Parental Socio-Economic Index of Occupational Status (Hisei), Cultural Possessions (Cultposs) and Home Educational Resources (Hedres); these three variables have been reformulated as indexes on a 0/100 scale and a positive correlation with test scores is expected to result. Gender and school management

[7]The linear functional form is adopted following Hanushek (1979), in that "the issue of functional form appears to be a second order problem […]. Within a limited range of variation, many functional forms look very similar".

[8]"A plausible value is a random number drawn from the distribution of scores that could be reasonably assigned to each individual—that is, the marginal posterior distribution." (Adams and Wu 2002). Five plausible values were calculated for each student in each of the subjects. The regression and decomposition are done five times using each value at a time as dependent variable and then the mean of the coefficient values of each explanatory variable is calculated.

[9] Although PISA 2009 database provides a construct of grade repetition, this research uses information about the grade currently being attended by students to decide whether they have repeated (grade 9th or below) or not (grade 10th or above). Correlation between both measures is 84 %. This formulation is used in order to make comparisons between PISA 2000 and PISA 2009, as the former variables related to grade repetition were not included in PISA 2000.

(public/private) fixed effects are also explored by using appropriate dummy variables. Appendix I provides us a description of the variables included in our model.

School variables such as teacher characteristics, or the student/teacher ratio were not included at this stage because the quantity of missing values was high in the three countries and they are not randomly distributed (i.e. the share of school-level missing data is greater among schools attended by low-performing students).

To control for the presence of endogeneity, the Durbin-Wu-Hausman test (augmented regression test) was applied. The null hypothesis of exogeneity was rejected for the variable joyread, and for this reason two stage least squares were applied, with PISA's index of diversity in reading material (Divread) as instrumental variable. Correlation between joyread and divread was found to be strong, whereas between divread and reading scores was almost negligible[10].

Once we have the estimates, we proceed to apply the Oaxaca-Blinder decomposition technique for the linear model (Blinder 1973, Oaxaca 1973). This procedure originally proposed to study unexplained gaps in labor market outcomes by groups by allowing coefficients to vary by sub-groups and decomposing the difference into a part that is "explained" by group differences in productivity characteristics such as education or work experience and a residual part that cannot be accounted for by such differences in endowments and is supposed to come from some type of market discrimination (Jann 2008) was later extended to other fields. A recent application in education is Ammermüller (2007). This author applies Oaxaca-Blinder to decompose country differences in PISA test scores between Finland and Germany. In our model, unexplained time-differences in school achievement may arise from individual, family and school characteristics being more/less favorable in 2009 than in 2000 and changes in productivity. This would resemble the explained and unexplained parts in traditional wage-gaps models. An additional feature of the model is the capacity it has to identify changes in productivity of individual, family and school assets or endowments included in the production function as marginal effects and changes in the overall effectiveness of the institutional environment as fixed effects.

The decomposition methodology applied for studying disparities of mean outcomes between two years (SG_{a-b}) is explained as:

$$SG_{a-b} = E(R^a) - E(R^b)$$

Where SGa-b stands for the estimated mean gap, E(R) denotes the expected value of the outcome variable and a and b represent two different points in time[11].

[10] More instruments were used in each edition of PISA in order to test robustness. The final model is exactly identified in order to use the same instruments in both editions (2000 and 2009) and divread was the only variable that fulfilled that requirement. Post estimation tests (following Cameron and Trivedi Stata manual) were performed in order to confirm that divread is not a weak instrument of joyread.

[11] For further details see Elder, Goddeeris, Haider (2010)

$$E(A^a) = E(\hat{\beta}_0^a + \hat{\beta}_F^a F_i^a + \hat{\beta}_S^a S_i^a + \hat{\beta}_P^a P_i^a + \varepsilon_i^a)$$
$$= E(\hat{\beta}_0^a + \hat{\beta}_F^a F_i^a + \hat{\beta}_S^a S_i^a + \hat{\beta}_P^a P_i^a) + E(\varepsilon_i^a)$$
$$= \hat{\beta}_0^a + \hat{\beta}_F^a E(F_i^a) + \hat{\beta}_S^a E(S_i^a) + \hat{\beta}_P^a E(P_i^a)$$

With F, S, and P representing family, school and student variables,

And, $E(\hat{\beta}^a) = \hat{\beta}^a$ and $E(\varepsilon^a) = 0$ by assumption.

$$SG_{a-b} = \hat{\beta}_0^a + \hat{\beta}_F^a E(F_i^a) + \hat{\beta}_S^a E(S_i^a) + \hat{\beta}_P^a E(P_i^a) - \hat{\beta}_0^b - \hat{\beta}_F^b E(F_i^b) - \hat{\beta}_S^b E(S_i^b) - \hat{\beta}_P^b E(P_i^b)$$

Rearranging,

$$SG_{a-b} = [(E(F_i^a) - E(F_i^b))\hat{\beta}_F^a + (E(S_i^a) - E(S_i^b))\hat{\beta}_S^a + (E(P_i^a) - E(P_i^b))\hat{\beta}_P^a] +$$
$$[(\hat{\beta}_0^a - \hat{\beta}_0^b) + (\hat{\beta}_F^a - \hat{\beta}_F^b)E(F_i^b) + (\hat{\beta}_S^a - \hat{\beta}_S^b)E(S_i^b) + (\hat{\beta}_P^a - \hat{\beta}_P^b)E(P_i^b)] +$$
$$[(\hat{\beta}_F^a - \hat{\beta}_F^b)(E(F_i^a) - E(F_i^b)) + (\hat{\beta}_S^a - \hat{\beta}_S^b)(E(S_i^a) - E(S_i^b)) + (\hat{\beta}_P^a - \hat{\beta}_P^b)(E(P_i^a) - E(P_i^b))]$$

This is a "three-fold" decomposition: the first 3 terms in the right-hand side of the equation refer to the endowment effects, that is, the differences in family, individual and school characteristics between 2009 and 2000; a second component is formed by the following four summands that together account for the coefficient effect (fixed and marginal) and the last three terms summarize interactions. The previous decomposition is formulated from the point of view of moment b. Therefore, the allocation component measures the expected change in the mean outcome of moment b, if the country had moment a's predictor levels. Moreover, the coefficient component measures the expected change in point b's mean outcome, if it had moment a's coefficients. Our analysis is formulated with the year 2000 as moment b since we want to address the evolution of each country between 2000 and 2009. This procedure is done with the five plausible values available for each student and the effect is calculated at the mean of the five coefficient values of each variable. The model was estimated using Stata 11(Jann 2008, Kreuter and Valliant 2007, Oyeyemi et. al. 2010).

4– Results

Contributing factors to PISA results in 2000 and 2009

In this study, contributing factors to explain disparities between 2000 and 2009 PISA results in the three countries are the regression coefficients reported in Table 2 (the crude measure of productivity used in this study) and the mean value of covariates, the endowments, shown in Appendix II.

In terms of the covariates component the most significant changes between 2000 and 2009 are associated in Argentina and Chile with the repetition rate (reported in Appendix II). Whereas the latter country managed to reduce in this period the rate of repetition from 39% to 25%, in the former the change operated with equal magnitude but in the opposite direction. Repeaters in Argentina by 2009 represented 38% of 15 year old students. In Mexico, about one every two 15 years old students were identified in both years as a repeater.

Enrolment of 15 years old students in private schools increased in Chile between 2000 and 2009 by about 27%, with non-publicly run institutions accounting for nearly 60% of

15 years old enrolments. These numbers double the Argentinean figures and are three to five times as large as the Mexicans (see Appendix II).

The pattern of changes between 2000 and 2009 in the index of socioeconomic status of the families (Hisei, Cultposs, Hedres) present differences between the three countries. In Argentina, only the index of occupational status of parents shows an improvement; the index of cultural possessions and educational resources at home deteriorated. Whilst Chilean students benefit from a substantial improvement in their parents' job status between 2000 and 2009, Mexican ones experienced in the same period a retreat of their home educational resources of great magnitude. In fact, they end up being by 2009 the less endowed of the three populations in terms of possessions available for didactic use.

Table 2 reports the coefficients obtained from estimation of the IV linear regression model. All the coefficients had the expected sign and those marked with an asterisk were statistically significant at the 5% level.

	Argentina		Chile		México	
Variable	2000	2009	2000	2009	2000	2009
Intercept	223.32*	307.21*	259.32*	307.04*	200.88*	374.92*
Gender (Male=1)	-8.95*	-20.55*	-0.42	-5.20*	4.89*	-16.11*
Repetition (Yes=1)	-86.75*	-63.03*	-60.65*	-58.91*	-66.47*	-55.79*
Management (Private=1)	28.27*	36.37*	21.88*	21.31*	44.18*	18.14*
Hisei	0.84*	1.00*	1.07*	0.86*	0.68*	0.68*
Cultposs	-0.09	0.19*	0.05*	0.17*	0.04	0.04*
Hedres	0.69*	0.62*	0.39*	0.43*	0.57*	0.57*
Joyread (instrumented)	3.07*	0.59*	2.11*	1.64*	3.32*	0.51*
R^2	0.38	0.36	0.37	0.37	0.30	0.29
N	3,190	4,195	4,600	4,876	3,946	35,730
F	293.542	268.352	672.376	518.014	387.636	799.932

Dependent variable: Reading score
Instrumented: joyread
Instruments: gender, repetition, management, hisei, cultposs, hedres and divread
Note: * indicates significance at 5% level

On gender issues, the three countries exhibited a similar pattern, with boys performing in 2000 and 2009 below girls (aggravating significantly by 2009). In fact, in 2000, in Chile, the mean expected score of a student with average characteristics was similar for boys and girls, whereas in 2009 boys are expected to obtain a score 5.2 points lower. Both the Argentinean and Mexican boys had also a lower performance relative to girls and the distance increased by 2009. Students attending a privately-run school tend to perform in the reading test better than their peers[12] attending government administered schools. In Mexico the size of the effect of attending a privately-run school was smaller in 2009 (18 points against 44 in 2000), attending a private school in Chile represented a handicap of about 22 points in both rounds of PISA and attending a private school in Argentina produced in 2009 a strong (positive) effect on reading scores. (It increases from 22 points in 2000 to 36 points in 2009). Repeating a grade negatively affects the mean expected score of the reading test. In the three countries, the level of the resulting effect of repeating grade was smaller in 2009. Argentina exhibits the better results in reducing the test score gap between repeaters and non-repeaters (from -86 points in 2000 to -66 points

[12]This effect can be biased due to auto-selection, but that effect is not be considered in this study.

in 2009). The coefficients of the family background and socioeconomic covariates have the expected sign in all three countries. The level of the marginal effect for the socioeconomic index of parent occupation increased in Argentina, but decreased in Chile and was stable in Mexico between 2000 and 2009. The coefficient of the index of cultural possessions turned from not being significant in Argentina to being so in 2009, and the same happened in Mexico. The value of the coefficient for the Home Educational Resources Index decreased in Argentina, increased in Chile and was stable in Mexico between 2000 and 2009. The coefficient of the "enjoy reading" index, after controlling for endogeneity effects, develops negatively in the three countries, as it contracted through the decade.

Oaxaca-Blinder decomposition

First, the estimated mean score gap was obtained for the three countries by using Table 2 coefficients and mean value of variables (Appendix II). This unexplained mean gap was then broken down using the Oaxaca-Blinder technique for linear models developed in section 3, so as to identify the contribution made to it by changes in coefficients (returns and fixed effects) and characteristics (endowment). Results are presented in Table 3 below.

In Argentina and Chile the estimated total gaps are statistically significant at the 5% level, and they account for nearly all of the observed difference. Quite the contrary, the estimated total gap was found in Mexico to be non-significant. However, being both the real and the estimated gap of less than 4 points, the time-difference become negligible.

In Argentina, the 5% retreat of reading competences between 2000 and 2009 (about 20 points or one fifth of the international standard deviation) owes itself to a weighted contraction of characteristics (about 2/3) and coefficients (about 1/3) -after the netting of the combined effect. In this respect, the decomposition technique helped to identify that the "quantity of capital" in terms of endowments that intervene in the production of education shrank in 2009 relative to 2000 and that, their (negative) effect on PISA results is reinforced by a lower level of effectiveness by which endowments are transformed in knowledge in 2009. The exact opposite happens in Chile.

In turn, the average reading score of Mexican students are in 2000 and 2009 quite similar. However, Oaxaca-Blinder decomposition

Table 3. Oaxaca-Blinder Decomposition of Reading Score Gap

	Argentina		Chile		Mexico	
Real Total Gap	-19.99		39.81		3.30	
Estimated Total Gap	-19.98	*	38.26	*	1.16	
Characteristicscs	-13.82	*	13.67	*	-22.92	*
Joyread (Instrumented)	-2.89	*	-1.45	*	-9.77	*
Gender (Male=1)	-0.18		-0.01		-0.02	
Repetition (Yes=1)	-7.99	*	8.05	*	1.68	*
Management (Private=1)	-0.33		2.74	*	-2.46	*
Hisei	1.45	**	4.65	*	-1.33	*
Cultposs	0.71	**	0.07	**	0.24	
Hedres	-4.59	*	-0.37	*	-11.27	*

Coefficients	-8.95	*	25.55	*	14.51	*
Joyread (Instrumented)	-115.29	*	-22.76	*	-149.49	*
Gender (Male=1)	-5.02	*	-2.23		-10.26	*
Repetition (Yes=1)	6.10	*	0.65		4.74	*
Management (Private=1)	3.21		-0.27		-4.48	*
Hisei	6.05	**	-6.99	*	0.01	
Cultposs	17.79	*	6.03	*	0.09	
Hedres	-5.69		3.41		-0.13	
Intercept	83.89	*	47.71	*	174.04	*
Combined	2.79	*	-0.96		9.58	*
Joyread (Instrumented)	2.33	*	0.33	**	8.28	*
Gender (Male=1)	-0.23		-0.17		0.07	
Repetition (Yes=1)	2.18	*	-0.23		-0.27	**
Management (Private=1)	-0.09		-0.07		1.45	*
Hisei	0.28		-0.93	*	0.00	
Cultposs	-2.17	*	0.15	*	0.01	
Hedres	0.49		-0.04		0.03	
Note: * indicates significance at 5% ** indicates significance at 10%						

shows they suffer from a fallback in characteristics between 2000 and 2009 that was responsible for a decrease of about 22 points in the mean score of the reading test However, this lost is balanced against a positive coefficients effect of about 14 points, which according to our model reflects gains in productivity and another 10 extra points associated with combined effects. Further examination of fixed and marginal coefficients effect would suggest that what prevented Mexico from falling any further in PISA is that, despite the lower level of inputs and input-productivity of 2009 relative to 2000, overall institutional conditions improved (as suggested by the important increase between 2000 and 2009 in the value of the intercept). In Argentina, on the other hand, the improvement in overall institutional conditions between 2000 and 2009 was smaller. In Chile, the improvement in overall institutional environmental conditions that took place between 2000 and 2009 was not as important as the contribution made by the increase in the marginal effect of covariates to explain the rise in PISA test scores between 2000 and 2009. In short, a variety of outcomes may be expected. The combination of endowments and coefficient effects helped to identify relevant sources of the improvement of PISA results in Chile, the stability found in Mexico and the decrease that took place in Argentina.

5 - Conclusions

The 2000-2009 evolution of PISA reading scores in three South American countries having analogous initial situations though different final states was analyzed in this work and the Oaxaca-Blinder decomposition of means gap framework was adapted to study time-disparities in educational outcome. The Oaxaca-Blinder technique, applied to the mean expected time-gap in reading scores allowed us to explore the disparities in students reading ability as by a combination of characteristics, coefficients and combined effects. The results reported in Section 4 illustrate that more than one outcome may be obtained, because characteristics and coefficients effects can either be complementary (reinforcing

reciprocal impact) or opposite. Argentina's disappointing PISA performance in the last decade is due to a reduced characteristics stock and decreased coefficients. In contrast, Mexico's scores were similar in 2000 and 2009. The exercise unveiled that, in this country, a decrease in the stock of characteristics was roughly compensated with an increase in effectiveness between 2000 and 2009. A major contributing source to this effect was an improvement of overall institutional environment (fixed) effect, rather than better use of existing endowments. Chile, instead, displays a quite different mix of effects: total coefficient effects are large (nearly twice as much as the estimated for the other countries under analysis), positive and reinforcing positive characteristics effect. Interestingly, estimates of improvement of overall institutional efficiency in Chile were small in this period, in comparison to improvements in the productivity of specific endowments applied to educational production. This result is different from the findings for Argentina and Mexico where the productivity level of intervening endowments remained stable or even decreased but the overall institutional environment effect increased markedly along the period. The contribution of improvements in the overall institutional environment to PISA results was important in all three countries: approximately 47, 83 and 174 points in Chile, Argentina and Mexico respectively.

Finally, this study highlights the usefulness of exploring how changes in the relative contribution of endowments, fixed and marginal coefficient effects help to understand the sources of time-gaps in PISA results within a country as well as how these contribute to explain differences in PISA results trajectories among various countries. Educational leaders and practitioners in general, and people that participate in the design of more effective treatments for the improvement of the students' present learning performance, may find information provided by the type of time-gap decomposition methods useful for the analysis of the sources of time-change in PISA learning outcomes.

References

Adams, R. and Margaret Wu (2002), "PISA 2000 Technical Report". Technical Report, Paris: OECD.

Ammermüller, A. (2007), "PISA: What makes the difference? Explaining the Gap in PISA Test Scores Between Finland and Germany", Centre for European Economic Research.

Barro R. and Jong-wha Lee (1993), "International comparisons of educational attainment" Journal of Monetary Economics 32, No 3, pp 363-394.

Blinder, A. S. (1973), "Wage Discrimination: Reduced Form and Structural Estimates", Journal of Human Resources, 8, pp. 436 - 455.

Breton T. R. (2011) "The quality vs. the quantity of schooling: What drives economic growth?" Economics of Education Review, forthcoming.

Elder T. E., John H. Goddeeris and Steven J. Haider (2010), "Unexplained gaps and Oaxaca-Blinder decompositions" Labour Economics, 17, pp 284-290.

Hanushek, E. A. (1979), "Conceptual and empirical issues in the estimation of educational production functions" The Journal of Human Resources, Vol. 14, No 3, pp. 351-388.

Hanushek E. A. and Dennis D. Kimko (2000), "Schooling, labor force quality, and the growth of nations" American Economic Review 90 No 5, pp 1184 – 1208.
Hanushek E. A.and Javier A. Luque (2003), "Efficiency and equity in schools around the world" Economics of Education Review 22 No 5, pp 481 – 502.
Hanushek, E.A. (2007), "Education Production Functions: Developed Country Evidence" International Encyclopedia of Education, 3rd ed. Elsevier Ltd, 2010.
Hanushek E.A. and Ludger Woessmann (2009) "Schooling, cognitive skills, and the Latin American growth puzzle" CESifo Working Paper No 2667.
Hanushek E.A. and Ludger Woessmann (2011) "The economics of International differences in educational achievement" in Hanushek E.A., Stephen Machin and Ludger Woessman, Handbook on the Economics of Education, vol. 3, pp 89 – 200.
Jann, B. (2008), "A Stata implementation of the Blinder-Oaxaca decomposition". ETH Zurich Sociology Working Paper, 5.
Kreuter, F., and Richard Valliant (2007), "A survey on survey statistics: What is done and can be done in Stata", The Stata Journal, Volume 7, Number 1, pp. 1–21.
Oaxaca, R. L.(1973), "Male-Female Wage Differentials in Urban Labor Markets", International Economic Review, 14, 3, pp 693-709.
Oyeyemi, G. M., A. A. Adewara, and R. A. Adeyemi (2010), "Complex Survey Data Analysis: A comparison of SAS, SPSS and STATA", Asian Journal of Mathematics and Statistics, 3, 1, pp 33-39
PISA, (2001) PISA 2000 Database, OECD.
_____ (2009), Data Analysis Manual, SPSS 2nd Edition, OECD.
_____ (2010), PISA 2009 Database, OECD.
Willms, J. D. and Thomas Smith (2005) "A manual for conducting analyses with data from TIMMS and PISA".
Wooldridge, J. M. (2002) "Econometric Analysis of Cross Section and Panel Data" The MIT Press, Cambridge, Mass., London.

Appendix I. Description of variables

Variable	Reference	Expected sign	Values	Definition
Plausible value for combined reading literacy	pvread			Random numbers, drawn from the distribution of combined reading literacy scores, that could be reasonably assigned to each individual.
Student attributes				
Repeated course	repetition	(-)	Dummy	The variable takes the value of 0 if the student is in either the tenth or eleventh grade, and 1 if he is in seventh, eighth or ninth grade.
Engagement in reading	joyread	(+)	0-100 (continuous)	The index measures student´s engagement in reading. It derives from students' level of agreement with eight statements about reading habits, enjoyment and attitude towards this activity. The variable was rescaled to the 0/100 range for the three countries together.
Diversity of reading materials	divread		0-100 (continuous)	The index was derived from the frequency with which students read the following materials because they want to: magazines, comic books, fiction, non-fiction books and newspapers. The higher values on this index indicate higher diversity in reading."
Family background				

Highest Parental Socio-Economic Index of Occupational Status	hisei	(+)	0-100 (continuous)	Occupational data for student´s parents was obtained by asking open ended questions. The response were coded to four-digit ISCO codes and mapped to the international socio-economic index of occupational status (ISEI). This index captures the highest ISEI attained by either the mother or father of the student. The variable was rescaled to the 0/100 range for the three countries together.
Cultural activities	cultposs	(+)	0-100 (continuous)	The index measures the frequency with which students engage in activities related to classical culture. The variable was rescaled to the 0/100 range for the three countries together.
Home Educational Resources	hedres	(+)	0-100 (continuous)	The index builds on the availability and number of certain educational items at home, namely a quiet place to study, a desk, text books and calculators. The variable was rescaled to the 0/100 range for the three countries together.
Control variables				
Gender (male=1)	gender	(+)	Dummy	0 was assigned to females and 1 to males.
School type (private=1)	Management	(+)	Dummy	Schools were classified as either public or private according to whether a public agency or a private entity had the ultimate decision-making power concerning its affairs.

AppendixII. Descriptive statistics

	Argentina				Chile				Mexico			
	2000		2009		2000		2009		2000		2009	
	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.	Mean	s.d.
Gender (Male=1)	0.44	0.01	0.46	0.01	0.47	0.01	0.51	0.01	0.50	0.01	0.49	0.00
Management (Private=1)	0.38	0.01	0.36	0.01	0.46	0.01	0.58	0.01	0.15	0.01	0.12	0.00
Repetition (Yes=1)	0.29	0.01	0.38	0.01	0.39	0.01	0.25	0.01	0.46	0.01	0.44	0.00
Hisei	34.21	0.51	38.46	0.39	29.54	0.30	36.56	0.31	32.71	0.34	34.12	0.19
Cultposs	60.78	0.83	53.54	0.53	51.48	0.51	53.11	0.44	38.07	0.58	44.78	0.26
Hedres	75.82	0.43	69.02	0.29	74.14	0.29	73.33	0.27	77.81	0.31	58.46	0.15
Joyread	46.42	0.36	45.62	0.18	47.65	0.23	47.05	0.17	53.04	0.19	50.11	0.09

Note: Standard Errors (S.E.) were obtained using balanced-repeated replicate weights (BRR).