Estimation of factors conditioning the acquisition of academic skills in Latin America in the presence of endogeneity

Geovanny Castro Aristizabal, Gregorio Giménez, Domingo Pérez Ximénez-de-Embún[1]

Abstract

This article identifies the main determinants of skill acquisition in Latin America. Not having repeated a grade, sex, the number of books in the home and the mother's education are defined as individual and family characteristics. In the case of school characteristics, the results are more heterogeneous between countries. The key factors seem to be attending a private school, the number of students per classroom, the quality of the educational materials available, and larger school size and autonomy. The characteristics of the schools explain most of the variability of the results, followed by family characteristics and then individual ones. School-based factors play a particularly crucial role in Argentina, Brazil and Costa Rica; family characteristics are very important in Chile, Colombia and Peru; and individual ones are important in Colombia and Mexico.

Keywords

Capacity-building, academic achievement, education, educational quality, evaluation, educational indicators, educational research, Latin America

JEL classification

C29, I21, I24, I28, I29

Authors

Geovanny Castro Aristizabal is a professor and research fellow in the Department of Economics at Pontificia Universidad Javeriana de Cali (Colombia). Email: gcastro@javerianacali.edu.co.

Gregorio Giménez is a professor and research fellow in the Department of Applied Economics at the University of Zaragoza (Spain). Email: gregim@unizar.es.

Domingo Pérez Ximénez-de-Embún is a professor and research fellow in the Department of Economic Analysis at the University of Zaragoza (Spain). Email: dpxe@unizar.es.

[1] The authors are grateful for financial assistance received from Santander Bank, the University of Zaragoza and Pontificia Universidad Javeriana de Cali (Colombia). They also appreciate the comments made by Macarena Iranzo and anonymous referees, which helped improve the final version of the article. Any remaining errors are the authors' exclusive responsibility.

I. Introduction

The Latin American countries participating in the Programme for International Student Assessment (PISA) 2012 have accepted the challenge of improving educational quality. Public expenditure per student has been increased, and laws have been amended to guarantee the right to education for the least privileged population groups. Although not all countries have progressed at the same pace, they have generally improved and expanded education supply by building schools and creating new teaching posts.

This has led to an increase in the average number of years' schooling and better results achieved on the various international tests in which they have participated (CIPPEC, 2011). A comparison of the results of the Second and Third Regional Comparative and Explanatory Studies (SERCE 2006 and TERCE 2013, respectively) generally reveals a significant increase in scores obtained by 3rd and 6th grade students in the three subject areas evaluated, especially mathematics (Rivas, 2015).[2] Nonetheless, at the secondary school level, the performance of the Latin American region still lags behind global standards. In terms of the three skills evaluated, the eight Latin American countries that took part in the 2012 PISA obtained results that ranked them among the last 20 of the 65 participants.[3]

Although there are many empirical studies on the economics of education generally, which shed light on the factors that condition school performance or skill acquisition, there are few focused on Latin America.

These research projects are based on estimating an education production function (EPF) through a variety of methodologies. In particular, the use of multilevel techniques (or hierarchical models) has gained increasingly broad acceptance as one of the best ways of studying and analysing educational data. This is because the characteristics of the student and the school (the inputs into EPF) are nested through their school performance; in other words, they are hierarchically structured (Gaviria and Castro, 2005).

Nonetheless, these hierarchical models assume the absence of correlation between the independent variables and the model's errors. Yet not all of the EPF inputs are statistically exogenous (zero correlation with the error term), since the model omits variables that cannot be measured directly. For this reason, they may be correlated with the error and cause an endogeneity problem (Hanushek and Woessmann, 2011). In the presence of this problem, multilevel techniques yield inconsistent and biased coefficients, which is why the literature suggests using instrumental variables (Greene, 2012). Nonetheless, this method has seldom been used in the economics of education (see Hanushek and Woessmann, 2011) and never to analyse school performance in Latin America.

In view of the above, this article aims to identify the factors that are decisive for acquiring skills in Latin America, by analysing data from countries participating in the 2012 PISA in the three subject areas evaluated. The education production function is also estimated using the generalized method of moments (GMM), firstly because a problem of endogeneity has been detected and secondly because aggregation at the country level is not possible owing to the lack of regional relative weights in the selected sample.

The article is structured as follows. Following this introduction, section II reviews the literature on factors determining the acquisition of skills in developing countries. A methodological section III describes the model used and its variables, along with the econometric techniques applied to analyse

[2] The scores represent evaluations made by the Latin American Laboratory for the Evaluation of Educational Quality (LLECE) of the United Nations Educational, Scientific and Cultural Organization (UNESCO) for Latin American countries, with the aim of measuring the learning outcomes of the region's students in reading, mathematics and science.

[3] The following Latin American countries participated: Argentina, Brazil, Chile, Colombia, Costa Rica, Mexico, Peru and Uruguay.

it. Section IV reports the results of the empirical analysis, in which the EPF estimation reveals that the key determinants of a student's school success are not having repeated a grade, sex, the number of books in the home and the educational level of his/her mother. Lastly, section V sets out the conclusions.

II. The literature on academic performance in Latin America

The earliest studies on educational quality and school performance (Alexander and Simmons, 1975; Jencks, 1972; and Coleman and others, 1966) found that family background largely determines students' academic performance. More recently, in the same vein, Woessmann (2010) argued that school performance is strongly related to family background and weakly associated with the characteristics of the school. In Latin America, Cervini (2012) analysed the school effect on academic performance, by applying bivariate multilevel models (at three levels: country, student and school), drawing on results of the 2006 SERCE in mathematics and reading. When extracurricular factors were controlled for, these tests found that the effect of the school on educational performance was very weak.

In recent decades, the number of studies on education and academic performance has increased considerably, with varying results. Some research finds that the characteristics of the student or those of his/her family environment are the key factors. In contrast, there is other empirical evidence that factors associated with schools and institutions have the greatest influence on school performance.

The results found for Latin American countries reflect this ambiguity. Some studies find a strong association between academic performance and the students' individual and family characteristics; but others conclude that features of the school are the most important factors. Nonetheless, relatively few research projects arrive at this second conclusion (see table 1).

Table 1
Latin America: determinants of school performance

Factor		Author(s)	Country or region	Data
Individual	Sex	Cárcamo and Mola (2012); Woessmann (2010); Vegas and Petrow (2007)	Colombia; Argentina; Latin countries in PISA	SABER11 2009; PISA 2000-2003
	Grade repetition (-)	Oreiro and Valenzuela (2013); Méndez and Zerpa (2011); Post (2011)	Uruguay; Uruguay and Chile; Chile, Colombia, Ecuador and Peru	PISA 2003; PISA 2006; SERCE 2006 and Regional employment surveys
	Race (-)	Marteleto (2012); Viáfara and Urrea (2006)	Brazil, Colombia	National survey 1982 and 2007; ENH 2000
Family	Group[a] (+)	Thieme, Prior and Tortosa-Ausina (2013); Donoso and Hawes (2002); Vivas, Correa and Domínguez (2011)	Chile; Chile; Colombia	SIMCE 2008; SIMCE 2000; ECV2003
	Parents' education (+)	Ayala, Marrugo and Saray (2011); Sánchez (2011)	Colombia	SABER11 2010
School	Teachers (+)	World Bank (2005); León, Manzi and Paredes (2004)	Mexico; Chile	PISA 2003; Teacher appraisal system
	School day (+)	Bonilla (2011)	Colombia	SABER11 2009

Source: Prepared by the authors.
Note: SABER11 refers to the results obtained in the tests applied to students in their last year of high school in the National Education Quality Assessment System. PISA refers to the Programme of International Student Assessment. SERCE means Second Regional Comparative and Explanatory Study. ENH is the National Household Survey. SIMCE is the Education Quality Measurement System. ECV stands for the Quality of Life Survey. The symbol in brackets that appears next to each of the factors refers to the relationship between performance and the factor. When sex is measured with a dummy variable that takes the value a 1 if the student is a girl and 0 if a boy, the effect on the score in reading is (+) and in mathematics (-).
[a] Refers to socioeconomic characteristics.

In the case of individual characteristics, sex, grade repetition and race are the factors that most influence school performance; and the following observations can be made (i) on average, girls obtain better results in reading, while boys do better in mathematics and science; (ii) grade repetition negatively affects the average score obtained in each of the areas evaluated; and (iii) black students have more disadvantages than those of other races and, thus, achieve worse results.

In the case of family factors, previous empirical studies of school performance in Latin American countries indicate that family characteristics and the environment jointly benefit student performance. The same is true of the parents' education, in particular, the mother's schooling level.

Lastly, in terms of school characteristics, the quality of the teachers and the type of school day (morning or afternoon) seem to be decisive for students' academic performance (see table 1). Studies that estimate education production functions in Latin America are still few and far between; and those that exist seldom use advanced econometric techniques. This study takes account of these limitations and tries to adopt a new perspective. To this end, techniques are used to impute data for which records are missing, together with estimators that are robust and unbiased in the presence of endogeneity stemming from the dual causation between endogenous and exogenous variables. The use of these techniques shows the greater weight of family characteristics in explaining the variability of performance, with respect to the sample worked with. It also highlights the existence of performance disparities between public and private schools.

III. Methodology

1. Model and description of the inputs of the education production function

The education production function (EPF) is similar to a typical production function. It relates the product (output), in other words the academic results or the marks scored by the students, with the inputs, namely, a set of variables associated with the student and the school. This article follows Hanushek, Link and Woessmann (2013), and Hanushek and Woessmann (2012 and 2011), working with the following EPF:

$$PM_{ij}^{p} = \beta_0 + \sum\nolimits_{i=1}^{m} \beta_i CE_i^p + \sum\nolimits_{i=m+1}^{h} \beta_i CE_i^p + \sum\nolimits_{i=h+1}^{w} \beta_i FE_i^p + \varepsilon_i \quad (1)$$

where PM_{ij}^{p} represents the average value of the five possible results that student i from country p can obtain in skill j (see OECD, 2014); ε_i denotes the model's random error term, which encompasses variables that are not directly observable or measurable, such as learning capacity, the student's innate abilities, or the "peer effect".[4] Lastly, the three summations include the inputs as follows:

- Student characteristics (CE) are encompassed in the first of the summations contained in equation (1). The following variables have been created:

  (i) The variable *sex* which takes the value 1, if the student is female, or 0 if male. It measures inequalities between the sexes by specific skills (see Woessmann, 2010, in relation to reading; García, Hidalgo and Robles, 2010, and Mullis and others, 2007, on mathematics; and Vegas and Petrow, 2007, with respect to science).

[4] The "peer effect" is located in this residual part of the function because it raises certain theoretical and empirical issues that make it difficult to measure. Murnane (1981) demonstrates the difficulty of identifying a student's relevant peer group (same class, same course or same school?). This author also highlights the difficulty of identifying which features of this peer group can really affect each student's school performance (the students' socioeconomic level, racial composition, academic level, sex?). In contrast, other authors, such as Brunello and Rocco (2008) or Angrist (2014), point to empirical problems that limit these results —mainly selection biases, spurious correlations and endogeneity problems.

(ii) The variable *non-repeater*, which takes the value 1, if the student has repeated at least one grade, and 0 otherwise. It is used to measure the impact of what is defined in the literature as "grade retention" (Méndez and Zerpa, 2011; Hong and Yu, 2007).

(iii) The variable *effort*, which takes the value 1, if the student seeks additional information to clarify a topic that she is studying but does not understand, and 0 if she does not perform this search.

(iv) The variable *discipline*, which is given the value 1, if the students pay attention to the teacher in most classes, and 0 otherwise (Post, 2011; Cervini, 2003).

- The second summation collects family characteristics, which make it possible to measure the impact of the household's socioeconomic and cultural status on skill acquisition. For this purpose, the following variables are specifically generated:

  (i) *books*, which takes the value 1, if there are more than 200 books in, and 0 otherwise (Woessmann and others., 2007).

  (ii) *educamother* and *educafather*, which refer to the minimum schooling level of the mother and father, respectively. Following Hanushek and Luque (2003) a bachelor degree gives this variable a value of 1 otherwise it is 0.

  (iii) *empfather*, which takes the value 1, if the father's is employed is part-time or full-time, and 0 otherwise (Hanushek and Woessmann, 2011; Woessmann and others, 2007).

- The last summation includes the school factors, which are the following:

  (i) The ownership of the school, which makes it possible to detect educational disparities between public and private schools. Thus, the variable *public* is created, which takes the value of 1 if the school is publicly owned, and 0 if it is private (Gamboa and Waltenberg, 2012).

  (ii) The student-teacher ratio (STRATIO), the number of students enrolled in the school (SCHSIZE) and the quality of educational materials (SCMATEDU), which are incorporated as proxies of the school's educational expenditure (Hanushek, 2011; Vignoles and others, 2000).

  (iii) The dummy variable autonomy, which is based the proposal by Hindrinks and others (2010), and takes the value 1 if the school director and teachers have decision-making autonomy on important matters related to the institution, and 0 otherwise.

2. Sources of information and treatment of data without registration

The information with which the EPF inputs are constructed and used to estimate model (1) is taken from the 2012 PISA, available from OECD (2015). This takes into account information on both students and schools in Latin American countries that participated in the 2012 PISA tests. In total, 90,799 observations of students are obtained from 3,722 schools, distributed as follows: 9,073 students and 352 schools in Colombia; 5,908 and 226 in Argentina; 19,204 and 839 in Brazil; 6,856 and 221 in Chile; 4,602 and 193 in Costa Rica; 33,806 and 1,471 in Mexico; 6,035 and 240 in Peru and 5,315 and 180 in Uruguay. The observations in question are statistically representative of the population of each of the countries studied (OECD, 2015).

Nonetheless, this database contains numerous missing values, in other words data corresponding to the information that are unrecorded because of a failure to reply to the questionnaires that students and directors of schools are required to complete. This can lead to biases in statistical inference, which is why the absent data must be imputed (Medina and Galván, 2007). Thus, in this study the proposal of the aforementioned authors is followed and the hot-deck method is applied to variables for which more than 10% of the values are missing in the database. According to Durrant (2009), this nonparametric method maintains the probability distribution of the imputed variables, so it is more efficient than multiple imputation methods.

Table 2 shows the mean and standard deviation of the variables used in EPF. It should be noted that the average scores obtained by Latin American students in reading, mathematics and science in the 2012 PISA (around 500 points in each of the subject areas) are unsatisfactory compared to those of the Organization for Economic Cooperation and Development (OECD). This is reflected in the low positions that these countries occupy in the ranking of the 65 participating economies (34 from the OECD and 31 associates), which are shown to the right of average columns in the table.

Table 2

Latin America (8 countries): means and standard deviations of the variables included in the education production function, 2012[a]

	Mean score: Reading	Rank	Mean score: Mathematics	Rank	Mean score: Science	Rank	*non-repeater*	*sex*	*effort*	*discipline*
ARG	395.98	60	388.43	59	405.63	58	0.62	0.51	0.24	0.09
	90.81		73.28		81.45		0.48	0.50	0.43	0.28
BRA	406.53	55	388.51	58	401.62	59	0.61	0.52	0.21	0.16
	81.48		74.78		74.62		0.49	0.50	0.41	0.37
CHL	441.40	47	422.63	51	444.93	46	0.74	0.52	0.30	0.15
	73.91		77.69		76.03		0.44	0.50	0.46	0.35
COL	403.40	57	376.49	62	398.68	60	0.58	0.53	0.32	0.17
	79.40		70.77		71.77		0.49	0.50	0.47	0.37
CRI	440.55	47	407.00	56	429.35	51	0.65	0.53	0.34	0.19
	69.21		64.72		65.04		0.48	0.50	0.47	0.39
MEX	423.55	52	413.28	53	414.92	55	0.80	0.51	0.26	0.19
	75.42		70.59		65.91		0.40	0.50	0.44	0.39
PER	384.15	65	368.10	65	373.11	65	0.69	0.51	0.27	0.14
	88.31		80.67		73.04		0.46	0.50	0.44	0.34
URY	411.35	54	409.29	55	415.84	54	0.61	0.53	0.30	0.16
	90.83		85.29		90.05		0.49	0.50	0.46	0.37
Total	409.29		393.38		404.85		0.67	0.52	0.25	0.16
	82.12		75.35		74.28		0.47	0.50	0.43	0.37

	books	*educamother*	*educafather*	*empfather*	*public*	*STRATIO*	*SCMATEDU*	*autonomy*	*SCHSIZE*
ARG	0.15	0.64	0.60	0.89	0.66	10.63	-0.54	0.96	519.31
	0.35	0.48	0.49	0.31	0.47	13.60	1.07	0.19	372.21
BRA	0.07	0.45	0.41	0.80	0.83	28.41	-0.58	0.95	979.25
	0.26	0.50	0.49	0.40	0.38	16.47	1.05	0.22	611.51
CHL	0.16	0.67	0.66	0.90	0.36	21.92	-0.38	0.92	902.16
	0.37	0.47	0.48	0.30	0.48	7.29	1.00	0.27	576.66
COL	0.08	0.55	0.57	0.85	0.84	27.01	-1.38	0.89	1455.14
	0.27	0.50	0.49	0.36	0.36	9.08	1.17	0.31	1106.35
CRI	0.10	0.57	0.58	0.88	0.85	20.22	-1.08	0.99	823.54
	0.30	0.50	0.49	0.33	0.36	23.85	1.24	0.11	631.80
MEX	0.09	0.39	0.42	0.85	0.88	30.69	-0.86	0.92	856.37
	0.29	0.49	0.49	0.36	0.33	31.62	1.14	0.27	946.47
PER	0.10	0.53	0.66	0.84	0.76	18.45	-1.16	0.99	672.27
	0.29	0.50	0.47	0.37	0.42	7.62	1.24	0.11	561.77
URY	0.14	0.47	0.45	0.89	0.83	15.48	0.12	0.86	905.73
	0.35	0.50	0.50	0.31	0.37	7.54	1.03	0.35	685.64
Total	0.09	0.48	0.48	0.83	0.80	25.87	-0.76	0.94	921.47
	0.29	0.50	0.50	0.37	0.40	20.84	1.14	0.24	776.33

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2012 test.

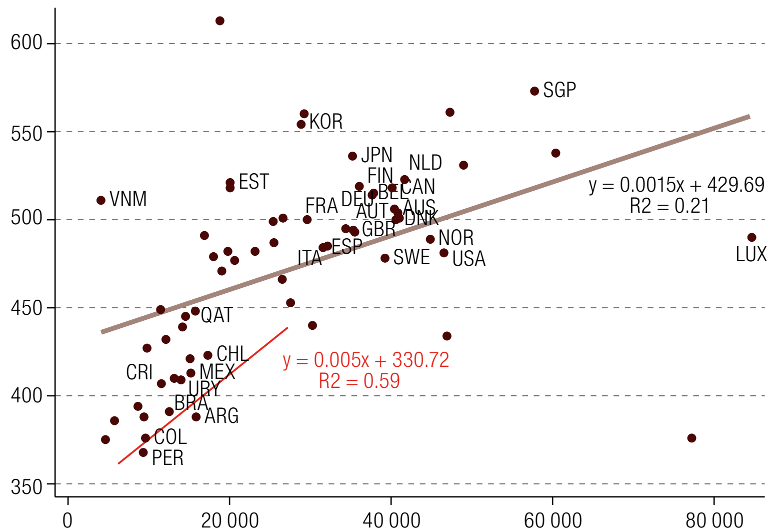
Note: The total has been calculated using the average data for all students in Latin America. The position refers to the country's ranking among those participating in the PISA test.

[a] The mean is shown on the upper line, with the standard deviation below.

The heterogeneity observed in the results may be a reflection of the inputs considered in EPF and shown in the table; but additional factors may also be in play. For example, if school performance is conditioned by socioeconomic circumstances, the circumstances prevailing in Latin American countries are very different from the OECD average. Figure 1 shows the relationship between income and performance in Mathematics (black dotted line), for the 65 economies participating in PISA. This relationship suggests that income differences can explain 21% of the variation among school results in the participating economies. The richest countries thus have a clear advantage.

Moreover, this relationship is stronger when exclusively considering the eight Latin American countries that are the focus of this research (the dashed red line), since 59% of the variability in the countries' results can be explained by their income differences. Average income in the region (US$ 13,175) lies between the extremes of Chile (US$ 17,312) and Peru (US$ 9,350).

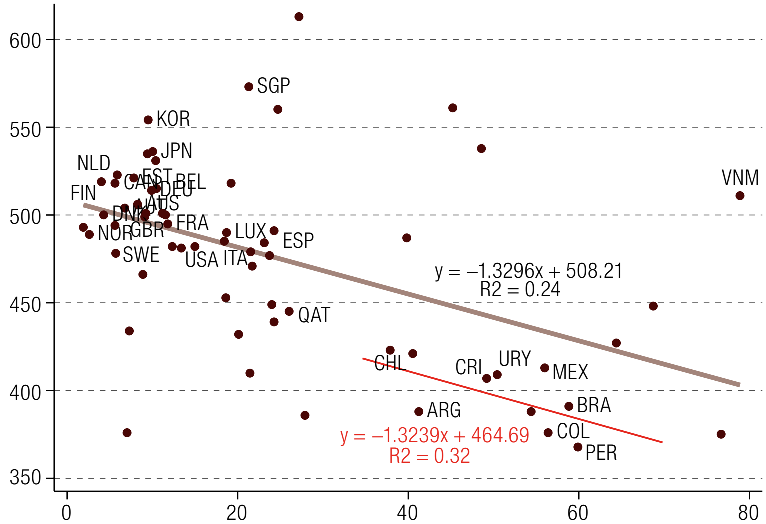
Figure 1
Results in mathematics and GDP per capita of the 65 economies that participated in the PISA tests, 2012
(Average scores and current dollars corrected for purchasing power parity (PPP))



Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) relating to the PISA 2012 test.

Cross-country differences in income level are compounded as a limiting factor by its distribution, since income in Latin America is distributed very unequally; and inequality seriously affects children and adolescents. Countries that have a larger proportion of students with socioeconomic problems face more severe educational challenges. This situation is reflected in figure 2, which shows the relationship between the results achieved in mathematics and the percentage of students who are in adverse socioeconomic circumstances. These appear to have a negative effect on school performance (dashed black line), since the index is associated with 24% of the variability in the results achieved in that subject. The relationship is more intense in the context of Latin American countries (dashed red line): 32% of the variability in the results obtained by students in these countries is associated with differences related to adverse circumstances, since 51% of Latin American students live in unfavourable environments. Peru had the largest proportion of students in this situation (59.9%) and Chile the smallest (37.9%).

Figure 2
Results in mathematics and proportion of students in adverse socioeconomic circumstances in the 65 economies that participated in the PISA test, 2012
(Average scores and percentages)



Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2012 test.

Nonetheless, while income level and the percentage of students in adverse socioeconomic circumstances are important limiting factors, they do not justify the differences between the results obtained in Latin America and those of other countries that participated in the PISA test. In figures 1 and 2, the Latin American countries are below the least-squares regression line, which means that their average result in the mathematics test is lower than expected, even after controlling for their level of income and the percentage of students in unfavourable environments. According to the regression line shown in figure 1, the expected result for Latin American countries would be 449 points, or 58 above the score actually obtained. Similarly, according to the regression line in figure 2, the expected result would be 440 points: 47 higher than actually achieved.

3. The endogeneity problem in the education production function

The EPF inputs are likely to be correlated with the model's error or disturbance term (1), so the correlation between them differ from zero because of omitted variables. Since some factors that determine performance are not directly measurable, not all inputs are statistically exogenous (Hanushek and Woessmann, 2011). For example, non-repeater status (an input corresponding to individual characteristics) may be associated with the student's capacity to learn, innate abilities or motivation. These factors cannot be observed directly, so they are encompassed by the error term, which results in a non-zero correlation between the fact of being a grade-repeater and the disturbance.[5]

[5] Similarly, school characteristics —such as the number of students to a class, and the school's size, ownership and autonomy— may depend on the educational policy and decisions made in administrative institutions, such as education ministries, or else may be decided on by the school's directors and teachers themselves. Consequently, there may be a non-zero correlation between these factors (which are unobservable and included in the error term) and school inputs.

When this correlation exists in at least one of the EPF inputs , the problem of endogeneity arises, which means estimates obtained through ordinary least squares (OLS) or hierarchical (multilevel) models are not appropriate. Nonetheless, multilevel models are still widely used in education, because they offer the advantage of making it possible to avoid possible selection biases in schools. While classical models assume fixed effects, that is effects common to all individuals, multilevel models are composed of two differentiated parts: one that is common and fixed for all contexts, and a second that varies and is estimated according to level. Thus, by simultaneously modelling multiple units of analysis, it is possible to accurately estimate the contribution of the variables of each of the levels (schools in PISA) to the student's academic performance.

As noted above however, the coefficients estimated by multilevel models in the presence of endogeneity will be biased and inconsistent (Wooldridge, 2010). If endogeneity exists, the literature suggests applying propensity score matching (PSM) methodologies or instrumental variables (IV). These methods would be consistent and would also make it possible to deal with the problem of selection biases. The main difference between the PSM and IV methods is that PSM is normally used to compare groups: one of them receives treatment and the other does not. In addition, PSM uses observable factors to construct the weights in the estimates, while the VI method is based on the use of instruments from unmeasured or unobserved factors. Thus, the advantage of using VIs is that the existence of these unobserved factors that are correlated with school results is taken into account. This is of vital importance when working with EPF, since it is inevitable that not all elements that influence the results will be included in the inputs used.

The problem with the VI method is that it does not deliver efficient estimators if endogeneity does not really exist, so the presence of the latter needs to be verified. It can also be difficult to find valid instruments that satisfy the necessary conditions, in other words instruments that correlate with the inputs of EPF but not directly with the schools' results. To detect the problem of endogeneity, this study uses the generalized method of moments (MGM) test statistic. The instruments are identified and analysed using the statistic developed by Hansen (1982) (see Hall, 2005; Baum, Schaffer and Stillman, 2003; and Hayashi, 2000). The following section describes the VI methodology as used in this study to estimate EPF.

IV. Results

1. Analysis of endogeneity

The null hypothesis used to detect this problem is $H_0{:}cov(X, \varepsilon_i) = 0$ (exogenous EPF inputs). If the p-value associated with the GMM statistic is lower than the significance level, then there is not enough statistical evidence to accept the null hypothesis. This would imply the presence of endogeneity. When applying the statistic to each input, it is found that, at 1% significance, non-repetition status is the only factor correlated with the error term. The other variables do not present problems of endogeneity; this is true for each of the skills evaluated in PISA 2012.

For this reason, the model is instrumentalized using the following instruments: first, the motivation of the student (*motivation*) as measured through the reply given to the following question: "In the last two weeks of classes, how many times did you skip school for a whole day?" (the variable takes the value 1 if the answer is "none"; otherwise, it is 0); second, the average duration of classes in minutes (*minuesp*, *minumat* and *minusci*); and third, the number of hours of reinforcement classes that the student takes outside school (*clasesp, clasmat* and *classci*). In the case of reading in particular, the instruments used are *motivation*, *minuesp* and *clasesp*; for mathematics, *motivation*, *minumat* and *clasmat*; and for science, *motivation*, *minusci* and *classci*.[6]

To validate the instruments, the hypothesis of their lack of correlation (and that of the other inputs) with the error term is formulated, and then the Hansen test (1982) is applied. The results show that the p-value associated with this statistic is greater than the one-percent significance level in each skill area, so there is not enough statistical evidence to reject the hypothesis. It is therefore concluded that the inputs and instruments are exogenous (see table 3).

Table 3
Latin America (8 countries): endogeneity and overidentification of the education production function, 2012

Endogeneity test	Argentina	Brazil	Chile	Colombia	Costa Rica	Mexico	Peru	Uruguay
H_0: exogenous inputs								
GMM - χ^2 statistic	72.175	82.762	60.005	8.721	27.398	290.462	56.215	9.721
[Reading]	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.002)
MGM - χ^2 statistic	71.061	5.586	26.886	8.311	7.973	353.245	90.597	11.175
[Mathematics]	(0.000)	(0.018)	(0.000)	(0.004)	(0.005)	(0.000)	(0.000)	(0.001)
GM - χ^2 statistic	104.694	89.051	40.532	22.027	26.249	294.152	54.181	3.218
[Science]	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.073)
Instruments test								
Hansen J - χ^2	3.915	0.357	3.361	1.419	0.993	0.504	1.220	0.022
[Reading]	(0.141)	(0.550)	(0.186)	(0.492)	(0.609)	(0.478)	(0.269)	(0.882)
Hansen J - $\chi^2$2	5.337	3.484	2.000	0.590	7.554	0.708	0.374	0.016
[Mathematics]	(0.069)	(0.062)	(0.157)	(0.745)	(0.023)	(0.400)	(0.541)	(0.899)
Hansen J - χ^2	2.311	2.291	3.753	1.188	4.312	0.545	1.058	1.667
[Science]	(0.317)	(0.130)	(0.053)	(0.552)	(0.116)	(0.460)	(0.304)	(0.197)

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2012 test.

Note: Instrumentalized variable: *non-repeater*. Instruments included: *sex, effort, discipline, books, educamother, educafather, empfather, public, STRATIO, SCMATEDU, autonomy* and *SCHSIZE*. Instruments excluded: *motivation, minuesp, clasesp* and *pre-school*, which takes the value 1 if the student attended pre-school, but 0 otherwise.] Imputation of omitted data through using the hot-deck methodology according to Medina and Galván (2007). p-value in parentheses.

Thus, once the problem of endogeneity has been corrected for (finding instruments correlated with non-repeater status, but not correlated with the model error of the Latin American countries participating in the 2012 PISA. The following sections present and interpret the results. As a measure of robustness, the results of the analysis using PISA 2009 data are reported at the end of the article. It should be noted that the results were similar to those obtained for 2012 (see annex A1).

[6] *clasesp*, *clasmat* and *classci* are dichotomous variables that take the value of 1 if the student attends 2-4 hours of reading, mathematics or science classes, respectively, outside the school, and 0 otherwise.

2. Estimating the education production function

(a) Reading comprehension

Table 4 displays the results obtained in reading.

Table 4
Latin America (8 countries): factors conditioning the acquisition of reading skills, PISA 2012

	Independent variable ↓	Argentina	Brazil	Chile	Colombia	Costa Rica	Mexico	Peru	Uruguay
Individual characteristics	*constant*	187.79	289.97	266.02	320.72	344.40	222.56	256.58	331.23
		(24.995)*	(17.412)*	(19.678)*	(30.813)*	(15.251)*	(13.521)*	(15.222)*	(10.379)*
	non-repeater	268.44	230.79	210.21	155.43	129.51	227.03	180.24	134.17
		(39.553)*	(30.952)*	(28.176)*	(43.896)*	(18.069)*	(15.489)*	(21.903)*	(20.337)*
	sex	7.38	6.61	9.11	1.36	14.17	5.57	3.87	22.01
		(5.533)	(4.087)	(3.158)*	(5.419)	(3.007)*	(1.949)*	(2.877)	(2.758)*
	effort	10.45	-1.17	6.26	7.35	5.35	10.15	5.86	7.21
		(4.493)**	(3.096)	(2.805)**	(2.987)**	(2.389)**	(1.759)*	(3.046)***	(2.428)*
	discipline	-11.10	0.29	5.48	8.59	3.99	4.44	-1.05	6.66
		(6.948)	(3.166)	(3.437)	(3.810)**	(2.927)	(2.021)**	(3.631)	(3.108)**
Family characteristics	*books*	5.21	2.45	28.77	19.12	19.51	13.65	21.94	17.15
		(6.090)	(4.130)	(3.140)*	(5.039)*	(4.089)*	(2.719)*	(4.185)*	(3.379)*
	educamother	13.97	6.45	6.93	11.03	1.21	0.07	9.64	0.97
		(4.408)*	(2.805)**	(4.120)	(4.168)*	(0.321)*	(1.833)	(3.384)*	(0.531)***
	educafather	-5.20	1.77	19.12	10.89	-2.11	9.31	6.26	-1.28
		(4.231)	(2.665)	(3.377)*	(3.157)*	(2.468)	(1.772)*	(3.546)***	(2.387)
	empfather	10.84	1.94	-13.87	2.10	-3.41	6.84	2.31	5.17
		(6.636)	(2.897)	(4.691)*	(4.406)	(3.493)	(2.222)*	(3.478)	(3.583)
School characteristics	*public*	-16.04	-31.47	-0.84	-27.18	-25.48	-2.55	-25.17	-35.41
		(9.854)	(5.903)*	(4.736)	(5.543)*	(4.718)*	(3.160)	(5.036)*	(7.505)*
	STRATIO	-0.08	-0.41	-0.53	-0.37	0.11	-0.11	-0.09	-0.16
		(0.135)	(0.068)*	(0.196)*	(0.152)**	(0.043)**	(0.016)*	(0.194)	(0.1669)
	SCMATEDU	-9.58	-1.64	0.39	3.43	3.54	3.25	10.59	9.15
		(3.224)*	(1.447)	(1.523)	(1.782)***	(1.397)**	(0.893)*	(1.179)*	(1.418)*
	autonomy	35.45	1.94	9.08	0.72	7.56	7.05	3.30	-2.91
		(8.998)*	(5.818)	(5.418)***	(4.265)	(11.100)	(3.185)**	(10.519)	(4.157)
	SCHSIZE	-0.01	0.00	0.01	0.01	0.01	0.00	0.02	0.00
		(0.009)	(0.002)	(0.003)*	(0.001)*	(0.002)*	(0.001)**	(0.002)*	(0.003)
	Observations	5 632	16 573	5 898	8 059	4 281	29 614	5 442	4 754
	Instruments	*motivation minuesp clasesp*	*pre-school minuesp*	*motivation minuesp clasesp*	*motivation minuesp clasesp*	*motivation minuesp clasesp*	*motivation pre-school*	*motivation minuesp*	*motivation minuesp*

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2012 test.

Note: * significant at 1%, ** significant at 5%, *** significant at 10%. Robust standard deviations in parentheses. Imputation of the omitted data through the hot-deck methodology according to Medina and Galván (2007).

In terms of individual characteristics, the results obtained in reading show that there are gender gaps in favour of women, which concurs with the findings obtained by Woessmann (2010). Thus, in Uruguay and Costa Rica, the countries where the gap is largest, women are 22.1 and 14.1 points above men, respectively. In contrast, in Argentina, Brazil, Colombia and Peru, the gaps are not significant. The non-grade repeater condition is also a decisive factor in the acquisition of this skill in all countries, since it has a positive effect on the average score according to Méndez and Zerpa (2011) and Martin (2011). It should be noted that the effect of the *non-repeater* variable is stronger in Argentina and Brazil. The same applies to students who seek additional information to clarify a topic they study and do not understand (*effort* variable). The effect of this variable is highest in Argentina and Mexico. Lastly, discipline is a positive and significant factor in Colombia, Mexico and Uruguay, but it is not significant in the other countries.

In terms of the socioeconomic and cultural status of the household, the results show that the number of books positively influences the average reading score, as indicated by Crespo, Díaz and Pérez (2012), and Woessmann and others (2007). The greatest impact of this indicator is observed in Chile, with 28.7 points, and in Peru, with 21.9. Positive effects of the mother's education are also observed, such as reported by Hanushek and Luque (2003). Thus, students with mothers whose educational level is high school, at least, obtain a higher average score than those whose mothers lack these studies. In the case of Chile and Mexico, the effect of the mother's education is not significant, which coincides with the result obtained by Meunier (2011) for Switzerland. In the case of the father's education level, the estimates differ between countries. In Chile, Colombia, Mexico and Peru, the relationship is positive and significant, while in the other the countries it is not. Lastly, the effect of the time that the father spends working is very heterogeneous: negative in Chile, positive in Mexico, and not significant in the other countries.

In the case of school characteristics, the relationship between the ownership of the school and the acquisition of reading skills is negative in all countries studied. Students who attend private schools obtain a higher average score than those who attend public schools. Nonetheless, in Argentina, Chile and Mexico, this relationship is not significant. Formichella (2011) obtains the same results for Argentina. Among countries in which this relationship is significant, Uruguay and Brazil display the greatest differences between public and private schools, while Costa Rica and Peru have the smallest (Giménez and Castro, 2017; Fernández and Del Valle, 2013; Gamboa and Waltenberg, 2012; Montero and others, 2012).

The coefficient of the variable that represents the student/teacher ratio is negative, coinciding with the findings reported by Krueger (2003) and Krueger and Whitmore (2001). In the cases of Argentina, Peru and Uruguay, this variable is not significant. The values obtained with respect to school size were not significant in Argentina, Brazil and Uruguay.

The quality of educational materials is also a condition for school performance, except in Brazil and Chile, where the estimated ratio is not significant. Peru and Uruguay are the countries where this type of expenditure has greatest effect.

Lastly, in terms of autonomy, it is seen that the decisions taken by school directors and teachers on all major institutional issues have a positive effect in Argentina, Chile and Mexico; but this effect is not significant in the other countries (Benton, 2014).

(b) Mathematics

Table 5 shows the factors that contribute to the results in mathematics.

Table 5

Latin America (8 countries): factors conditioning the acquisition of mathematics skills, 2012

	Independent variable ↓	Argentina	Brazil	Chile	Colombia	Costa Rica	Mexico	Peru	Uruguay
Individual characteristics	*constant*	243.70	375.54	294.16	349.90	374.86	229.23	273.57	357.42
		(20.223)*	(11.971)*	(17.557)*	(22.527)*	(12.773)*	(14.124)*	(16.109)*	(9.347)*
	non-repeater	221.31	101.05	161.01	121.02	88.22	239.09	193.38	130.63
		(32.461)*	(21.946)*	(24.873)*	(31.947)*	(15.222)*	(16.106)*	(23.282)*	(17.966)*
	sex	-38.68	-27.01	-34.91	-38.63	-30.27	-33.31	-36.06	-23.21
		(4.589)*	(2.801)*	(2.793)*	(4.111)*	(2.441)*	(1.993)*	(2.972)*	(2.489)*
	effort	10.84	6.12	7.75	6.52	5.37	9.33	3.98	9.45
		(3.665)*	(2.009)*	(2.439)*	(2.525)*	(1.947)*	(1.789)*	(3.155)	(2.284)*
	discipline	-6.95	3.78	2.28	7.53	2.75	6.24	-2.31	6.68
		(5.640)	(1.983)***	(2.936)	(3.028)**	(2.489)	(2.051)*	(3.668)	(2.839)**
Family characteristics	*books*	9.70	10.84	34.63	21.61	24.72	16.12	27.43	23.12
		(5.078)***	(2.832)*	(2.851)*	(4.342)*	(3.833)*	(2.740)*	(4.405)*	(3.062)*
	educamother	8.06	10.74	11.30	10.09	0.96	-1.17	4.83	1.41
		(3.614)**	(1.774)*	(3.441)*	(3.404)*	(0.261)*	(1.876)	(3.511)	(0.479)*
	educafather	1.95	6.66	20.49	9.00	0.42	6.69	3.04	-2.13
		(3.475)	(1.666)*	(2.833)*	(2.576)*	(1.972)	(1.814)*	(3.711)	(2.192)
	empfather	8.91	-0.04	-6.03	0.54	-1.37	0.65	-0.98	5.34
		(5.418)	(1.754)	(3.930)	(3.538)	(2.797)	(2.307)	(3.607)	(3.145)***
School characteristics	*public*	-8.66	-51.66	-8.80	-24.03	-28.04	-0.35	-25.88	-30.01
		(8.046)	(3.956)*	(4.092)**	(4.557)*	(4.200)*	(3.257)	(5.435)*	(6.686)*
	STRATIO	-0.10	-0.45	-0.76	-0.59	0.17	-0.10	-0.11	-0.64
		(0.107)	(0.046)*	(0.163)*	(0.127)*	(0.044)*	(0.017)*	(0.205)	(0.170)*
	SCMATEDU	-7.88	3.61	1.37	3.30	6.64	1.47	9.15	7.13
		(2.628)*	(0.956)*	(1.300)	(1.403)**	(1.142)*	(0.910)	(1.229)*	(1.287)*
	autonomy	22.47	2.70	14.19	-4.29	-1.36	5.55	-6.39	-5.18
		(7.158)*	(3.432)	(4.897)*	(3.456)	(9.592)	(3.227)***	(11.135)	(3.778)
	SCHSIZE	-0.02	0.01	0.01	0.01	0.01	0.00	0.01	0.00
		(0.008)**	(0.002)*	(0.003)*	(0.001)*	(0.002)*	(0.001)	(0.003)*	(0.003)
	Observations	5 632	16 968	5 898	8 059	4 281	29 614	5 442	4 754
	Instruments	*motivation minumat*	*motivation minumat clasmat*	*motivation minumat clasmat*	*motivation minumat clasmat*	*motivation minumat clasmat*	*motivation pre-school*	*motivation minumat*	*motivation minumat*

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2012 test.

Note: * significant at 1%, ** significant at 5%, *** significant at 10%. Robust standard deviations in parentheses. Imputation of the omitted data through the hot-deck methodology according to Medina and Galván (2007).

Considering gender differences, in all of the countries included in this research, boys obtained a higher average score than girls in mathematics, similar to the results obtained by Vegas and Petrow (2007). The largest gap occurs in Argentina, Colombia and Peru. Moreover, the effect of the *non-repeater* variable is positive and significant. This estimate coincides with that found by Oreiro and Valenzuela (2013), and Méndez and Zerpa (2011). The students with greatest academic capacity (greater effect of the non-repeater variable) are from Mexico and Argentina. Effort, on the other hand, has a positive and significant differential effect, except among Peruvian students. In terms of disciplinery climate, a positive and significant relationship is noted in Brazil, Colombia, Mexico and Uruguay. Thus, students who pay attention to the teacher achieve average scores that are 3.7, 7.5, 6.2 and 6.6 points higher than those who do not, respectively. The effect is not significant in Argentina, Chile, Costa Rica and Peru.

In terms of family characteristics, the number of books is a key factor in skill acquisition, with the greatest effect in Chile and Peru. The gap by educational level of the mother in general is positive and significant, as in Hanushek and Luque (2003). This determines a higher average score for students whose mothers completed high school at least. The countries where the mother's education has the greatest impact are Brazil and Chile. In contrast, this is not significant in Peru or in Mexico.

The effect of the father's educational level is positive and significant only in Brazil, Chile, Colombia and Mexico. In these countries, students whose parents have at least completed high school achieve, respectively, 20.4, 9.0, 6.6 and 6.6 points more than those whose father has not attained this educational level. In the cases of Argentina, Costa Rica and Peru, a positive but not significant relationship is calculated. In the case of Uruguay, the relationship is negative and not significant. Lastly, type of employment contract is not a condition for skill acquisition, since the estimated coefficients are not significant, except in Uruguay.

In school-related variables, there are significant gaps between the performance of public and private schools, except in Argentina and Mexico. Brazil and Uruguay display the greatest divergences, as with reading comprehension. Meanwhile, in Chile and Colombia these differences are minor.

Class size has a negative and significant effect (except in Costa Rica). The size of the school has positive and significant effects in all countries, except Argentina (where it is negative and significant) and in Mexico and Uruguay (not significant).

The results for autonomy are ambiguous. In the case of Argentina, Chile and Mexico, the fact that decisions on major institutional issues are taken by the director and teachers of the school favours the performance of students in mathematics. On average, in these countries, schools where there is autonomy score 22.4, 14.1 and 5.5 points more than that in other schools. In Brazil, Colombia, Costa Rica, Peru and Uruguay, the relationship is not significant.

Lastly, the quality of educational materials is positively related to performance in mathematics. The countries where there is a greatest effect are Peru, Uruguay and Costa Rica. It should be noted that Argentina is a special case, since the impact of educational spending on performance is negative.

(c) Sciences

The key factors determining the acquisition of science skills include sex and non-repeater status among the individual characteristics (see table 6). On gender, significant gaps were estimated, as in Vegas and Petrow (2007), with girls scoring worse. In terms of the second factor, the relationship between non-repeater status and academic performance is positive and significant. Students who do not repeat any grade perform better than those who have done so at least once (see Oreiro and Valenzuela, 2013, and Méndez and Zerpa, 2011).

Table 6
Latin America: factors conditioning the acquisition of science skills, 2012

	Independent variable ↓	Argentina	Brazil	Chile	Colombia	Costa Rica	Mexico	Peru	Uruguay
Individual characteristics	*constant*	193.34	403.52	309.39	298.60	365.93	249.65	286.18	374.63
		(27.767)*	(11.480)*	(15.814)*	(35.901)*	(15.541)*	(12.699)*	(12.842)*	(10.428)*
	non-repeater	299.20	51.37	158.77	214.08	123.92	203.04	145.94	109.86
		(44.942)*	(21.133)**	(21.786)*	(51.400)*	(18.022)*	(14.451)*	(18.886)*	(19.980)*
	sex	-26.43	-5.17	-16.58	-40.42	-22.28	-22.71	-19.28	-9.80
		(6.118)*	(2.658)***	(2.731)*	(6.354)*	(2.870)*	(1.767)*	(2.473)*	(2.673)*
	effort	8.06	8.76	7.21	3.61	3.23	8.34	4.51	10.27
		(4.855)***	(1.914)*	(2.520)*	(3.583)	(2.236)	(1.592)*	(2.599)***	(2.407)*
	discipline	-11.50	6.89	4.67	5.40	-0.58	3.15	-1.63	4.12
		(7.742)	(1.907)*	(3.120)	(4.402)	(2.804)	(1.835)***	(3.079)	(3.080)
Family characteristics	*books*	3.51	10.39	28.47	18.05	17.48	15.52	27.31	21.59
		(6.687)	(2.896)*	(2.967)*	(5.916)*	(3.712)*	(2.469)*	(3.678)*	(3.375)*
	educamother	11.16	12.46	8.10	0.87	0.99	0.44	8.73	1.92
		(4.834)**	(1.738)*	(3.450)**	(4.952)	(0.315)*	(1.680)	(2.886)*	(0.507)*
	educafather	2.21	8.71	19.30	10.89	-2.94	8.37	3.66	0.41
		(4.633)	(1.626)*	(2.998)*	(3.731)*	(2.352)	(1.630)*	(3.050)	(2.321)
	empfather	9.99	1.02	-7.76	2.69	-1.06	2.27	-2.24	-0.70
		(7.234)	(1.705)	(4.042)***	(5.279)	(3.276)	(2.012)	(3.056)	(3.401)
School characteristics	*public*	-3.60	-58.60	-9.37	-8.96	-27.70	-1.32	-19.00	-41.90
		(11.119)	(3.786)*	(3.903)**	(6.617)	(4.583)*	(2.911)	(4.456)*	(7.454)*
	STRATIO	-0.07	-0.49	-0.48	-0.51	0.05	-0.10	0.06	-0.20
		(0.150)	(0.045)*	(0.167)*	(0.183)*	(0.046)	(0.015)*	(0.169)	(0.171)
	SCMATEDU	-16.04	2.91	2.30	0.99	3.41	1.86	8.37	6.25
		(3.548)*	(0.945)*	(1.351)***	(2.126)	(1.314)*	(0.822)**	(1.015)*	(1.379)*
	autonomy	29.06	10.31	14.77	-0.24	3.86	8.86	0.78	-11.87
		(9.171)*	(3.219)*	(4.826)*	(5.143)	(11.766)	(2.917)*	(8.681)	(4.036)*
	SCHSIZE	-0.02	0.01	0.01	0.01	0.01	0.00	0.01	0.01
		(0.011)***	(0.002)*	(0.003)*	(0.002)*	(0.002)*	(0.001)	(0.003)*	(0.004)
	Observations	5 632	16 968	5 898	8 059	4 281	29 614	5 442	4 754
	Instruments	*motivation minucie*	*motivation minucie clascie*	*motivation minucie clascie*	*motivation minucie clascie*	*motivation minucie clascie*	*motivation pre-school*	*motivation minucie*	*motivation minucie*

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2012 test.

Note: * significant at 1%, ** significant at 5%, *** significant at 10%. Robust standard deviations in parentheses. Imputation of the omitted data through the hot-deck methodology according to Medina and Galván (2007).

Among the socioeconomic and cultural characteristics of the student, the key factors are the number of books in the home (Woessmann and others, 2007) and the mother's education (Hanushek and Luque, 2003). Argentina is an exception in the case of books, and Colombia and Mexico are exceptions in relation to the mother's education, since the coefficients associated with these variables were not significant. In terms of the father's education, although the estimated effect on the acquisition of science skills is positive (except in Costa Rica), it is only significant in Brazil, Chile, Colombia and Mexico. In contrast, the results of the father's type of employment contract are more varied. In some cases, the effect is negative and not significant (Costa Rica, Peru and Uruguay), in others, negative and significant (Chile) and, in others, positive and not significant (Argentina, Brazil, Colombia and Mexico).

In the case of school characteristics, average gaps in school performance were also estimated between public and private schools, favouring the latter; but the gaps are not significant for Argentina, Colombia and Mexico. Brazil and Uruguay are the countries with the greatest differences, as was also the case in reading and mathematics. With regard to the student-teacher ratio, the results are similar to those found in reading and mathematics. The effect on school performance is generally negative and significant, although not significant in the cases of Argentina, Costa Rica, Peru and Uruguay. School size has a positive and significant impact, except in the cases of Mexico and Uruguay.

Positive and significant effects are estimated for school autonomy in Argentina, Brazil, Chile and Mexico. In contrast, the effect is negative and significant in Uruguay, and it is not significant in Colombia, Costa Rica and Peru.

Lastly, the estimated coefficients of the quality of educational materials are positive and statistically significant in all countries, except Argentina (negative and significant) and Colombia (not significant).

(d) Decomposition of the causes of performance differences

The unequal scores obtained by the students may reflect differences in the effort they make, which corresponds to the residual part of EPF, or factors beyond their control, in other words the inputs defined in that function: individual, family and school factors.

If r is the result obtained in the PISA tests and C is a matrix of non-controllable factors, the expected test result, conditional on these factors, will be:

$$\hat{r} = \mathrm{E}[r|C] \tag{2}$$

From (2), the variance of the results in the different areas, $\frac{1}{N}\sum_{i=1}^{N}\left(r_i - \bar{r}\right)^2$, can be decomposed as a function of the relative contribution of each explanatory factor included in EPF. This is obtained using the Shapley-Shorrocks methodology,[7] which is based on the calculation of the variance, considering all possible permutations of the explanatory variables encompassed in EPF. As school PISA results are continuous variables of arbitrary mean and variance, the best option for estimating model (2) will be to use a linear function (Ferreira and Gignoux, 2014).

Table 7 displays the Shapley-Shorrocks decomposition by student, home and school characteristics. The variances of the results in mathematics, reading and science in relation to all observations in Latin America are estimated at 0.186, 0.185 and 0.181, respectively. By country, the minimum variance corresponds to Mexico, 0.141 in mathematics and science, and the maximum variance to Chile, 0.373 in mathematics. The validity of these estimates is supported by the low value of the standard errors of the bootstrap resampling method, which is significant at the standard levels with 100 replicas.

The rest of the table displays the variance decomposition percentages, in other words the extent to which school success is affected by individual, family and school characteristics. On average, individual characteristics appear to explain 12% of the variation in school results, family characteristics 28%, and school factors 60%. Individual characteristics play a particularly important role in Colombia and Mexico, family factors stand out in Chile, Colombia and Peru, and school characteristics are key in Argentina, Brazil and Costa Rica.

[7] For a detailed explanation, see Shorrocks (1982).

Table 7

Latin America (8 countries): Shapley-Shorrocks decomposition of the variance in school results, 2012

(Percentages of the variance)

	Latin America			Argentina			Brazil			Chile			Colombia		
	Read	Mathematics	Science	Read	Mathematics	Science	Read	Mathematics	Science	Read	Mathematics	Science	Read	Mathematics	Science
Individual characteristics	18.32	12.75	5.64	14.31	5.36	1.76	21.52	8.05	4.00	7.60	7.75	2.48	11.91	22.52	14.58
Family characteristics	24.77	28.32	30.84	17.22	25.27	25.90	18.93	21.24	22.85	49.14	47.24	47.00	37.53	34.39	39.36
School characteristics	56.91	58.93	63.52	68.47	69.37	72.34	59.55	70.71	73.15	43.26	45.01	50.52	50.56	43.09	46.06
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Var. results[a]	0.186	0.185	0.181	0.277	0.273	0.262	0.222	0.245	0.207	0.343	0.373	0.313	0.159	0.200	0.162
Bootstrap standard error	0.003	0.002	0.002	0.013	0.009	0.009	0.005	0.006	0.005	0.009	0.010	0.009	0.008	0.009	0.008
Observations		81 062			5 632			16 968			5 898			8 059	
R.R.[b]								100							

	Costa Rica			Mexico			Peru			Uruguay		
	Read	Mathematics	Science	Read	Mathematics	Science	Read	Mathematics	Science	Read	Mathematics	Science
Individual characteristics	17.47	15.72	5.11	22.63	15.39	9.17	5.15	8.36	3.87	17.47	7.19	3.67
Family characteristics	16.41	18.26	18.42	18.37	21.84	25.60	34.62	34.59	39.66	23.03	32.65	31.13
School characteristics	66.12	66.02	76.47	59.00	62.77	65.23	60.23	57.05	56.47	59.50	60.16	65.20
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Var. results[a]	0.220	0.249	0.201	0.169	0.141	0.141	0.289	0.304	0.259	0.252	0.251	0.222
Bootstrap standard error	0.009	0.013	0.009	0.004	0.004	0.004	0.010	0.010	0.010	0.009	0.010	0.010
Observations		4 340			29 760			5 442			4 963	
R.R.[b]							100					

Source: Prepared by the authors.

[a] Variance, Ferreira-Gignoux methodology (unscaled).

[b] Bootstrap resampling replicas.

V. Conclusions

This paper identifies the determinants of academic performance in the Latin American countries that participated in the 2012 PISA. When working with education production functions, the presence of endogeneity problems means that the estimated coefficients are inconsistent and biased. For this reason, estimates obtained through multilevel models are suboptimal. This study's methodological contribution stems from its use of the instrumental variables technique, which makes it possible to correct endogeneity problems.

A decomposition of the variance of school results reveals that school characteristics explain most of the variability of results between students (and about 60% of this variability across Latin America as a whole). The next most important factors are family characteristics (28%) and individual ones (12%). School factors play a particularly important role in Argentina, Brazil and Costa Rica; family characteristics are most important in Chile, Colombia and Peru; and individual ones in Colombia and Mexico. A decomposition to identify the contribution of each factor and verify the importance of school-based factors are an innovation in studies on Latin America as a whole.

In the case of individual and family factors, significant gender disparities were estimated, with girls obtaining a higher average score in reading and boys scoring higher in mathematics and science. Grade retention proved to be a determinant of school performance, along with the number of books in the home and the educational level of the mother.

In the case of school factors, the effects are more heterogeneous. Attending a private school is found to have a positive and significant effect in Brazil, Chile, Colombia, Costa Rica, Mexico, Peru and Uruguay. Smaller class size has a positive effect in the schools of Brazil, Chile, Colombia and Mexico. Higher spending on educational materials appears to produce better results, although apparently not in Argentina; and school size has a positive effect, except in Argentina and Uruguay. Lastly, greater school management autonomy benefits students' results in Argentina, Chile and Mexico.

The presence of endogeneity within the Latin American countries' education production function makes it advisable to use consistent and robust techniques that address these problems. In contrast to other techniques, which are common in earlier studies, but where endogeneity among the model's variables is not considered, the use of instrumental variables makes it possible to approximate the true effect of educational inputs on school results.

Bibliography

Alexander, L. and J. Simmons (1975), "The determinants of school achievement in developing countries: the education production function", *Working Paper*, No. 201, Washington, D.C., World Bank.

Angrist, J. D. (2014), "The perils of peer effects", *Labour Economics*, vol. 30, Amsterdam, Elsevier.

Ayala, J., S. Marrugo and B. Saray (2011), "Antecedentes familiares y rendimiento académico en los colegios oficiales de Cartagena", *Economía y Región*, vol. 5, No. 2, Cartagena de Indias, Technological University of Bolivar.

Baum, C. F., M. E. Schaffer and S. Stillman (2003), "Instrumental variables and GMM: estimation and testing", *The Stata Journal*, vol. 3, No. 1, Texas, Stata Corporation.

Benton, T. (2014), *A Re-evaluation of the Link Between Autonomy, Accountability and Achievement in PISA 2009*, Cambridge, Cambridge Assessment.

Bonilla, L. (2011), "Doble jornada escolar y calidad de la educación en Colombia", *Documentos de Trabajo sobre Economía Regional y Urbana*, No. 143, Bogota, Bank of the Republic.

Brunello, G. and L. Rocco (2008), "Educational standards in private and public schools", *The Economic Journal*, vol. 118, No. 533, Wiley.

Cárcamo, C. and J. Mola (2012), "Diferencias por sexo en el desempeño académico en Colombia: un análisis regional", *Economía y Región*, vol. 6, No. 1, Cartagena de Indias, Technological University of Bolivar.

Cervini, R. (2012), "El 'efecto escuela' en países de América Latina: reanalizando los datos del SERCE", *Education Policy Analysis Archives*, vol. 20, No. 39.
___(2003), "Diferencias de resultados cognitivos y no-cognitivos entre estudiantes de escuelas públicas y privadas en la educación secundaria de Argentina: un análisis multinivel", *Education Policy Analysis Archives*, vol. 11, No. 6.
CIPPEC (Public Policies for Equity and Growth Implementation Centre) (2011) [online] http://www.cippec.org/publicaciones-educacion.
Coleman, J. and others (1966), *Equality of Educational Opportunity*, Washington, D.C., Department of Education.
Crespo, E., C. Díaz and J. Pérez (2012), "Factores condicionantes de la desigualdad educativa: un análisis para el caso español", *Investigaciones de economía de la educación*, E. Pacheco Vieira (ed.), vol. 7, Oporto, Economics of Education Association.
Donoso, S. and G. Hawes (2002), "Eficiencia escolar y diferencias socioeconómicas: a propósito de los resultados de las pruebas de medición de la calidad de la educación en Chile", *Educação e Pesquisa*, vol. 28, No. 2, São Paulo, University of São Paulo.
Durrant, G. B. (2009), "Imputation methods for handling item-nonresponse in practice: methodological issues and recent debates", *International Journal of Social Research Methods*, vol. 12, No. 4, Taylor & Francis.
Fernández, A. and R. del Valle (2013), "Inequality in education in Costa Rica: the gap between students in public and private schools. An analysis of the results of the Programme for International Student Assessment (PISA)", *CEPAL Review*, vol. 111 (LC/G.2597-P), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).
Ferreira, F. and J. Gignoux (2014), "The measurement of educational inequality: achievement and opportunity", *World Bank Economic Review*, vol. 28, No. 2, Washington, D.C., World Bank.
Formichella, M. M. (2011), "Do private schools in Argentina perform better because they are private?", *CEPAL Review*, No. 105 (LC/G.2508-P), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).
Gamboa, L. F. and F. D. Waltenberg (2012), "Inequality of opportunity for educational achievement in Latin America: evidence from PISA 2006–2009", *Economics of Education Review*, vol. 31, No. 5, Amsterdam, Elsevier.
García, I., M. Hidalgo and A. Robles (2010), "Diferencias regionales en rendimiento educativo en España: ¿la familia lo explica todo?", *Working Paper*, No. 10.12, Sevilla, Pablo de Olavide University.
Gaviria, J. L. and M. Castro (2005), *Modelos jerárquicos lineales*, Madrid, La Muralla.
Giménez, G. and G. Castro (2017), "¿Por qué los estudiantes de colegios públicos y privados de Costa Rica obtienen diferentes resultados académicos?", *Perfiles Latinoamericanos,* vol. 25, No. 49, Mexico City, Latin American Faculty of Social Sciences (FLACSO).
Greene, W. H. (2012), *Econometric Analysis*, New Jersey, Prentice Hall.
Hall, A. R. (2005), *Generalized Method of Moments*, Oxford, Oxford University Press.
Hansen, L. P. (1982), "Large sample properties of generalized method of moments estimators", *Econometrica*, vol. 50, No. 4, New York, The Econometric Society.
Hanushek, E. A. (2011), "The economic value of higher teacher quality", *Economics of Education Review*, vol. 30, No. 3, Amsterdam, Elsevier.
Hanushek, E. A., S. Link and L. Woessmann (2013), "Does school autonomy make sense everywhere? Panel estimates from PISA", *Journal of Development Economics*, vol. 104, Amsterdam, Elsevier.
Hanushek, E. A. and J. Luque (2003), "Efficiency and equity in schools around the world", *Economics of Education Review*, vol. 22, No. 5, Amsterdam, Elsevier.
Hanushek, E. A. and L. Woessmann (2012), "Schooling, educational achievement, and the Latin American growth puzzle", *Journal of Development Economics*, vol. 99, No. 2, Amsterdam, Elsevier.
___(2011), "The economics of international diferences in educational achievement", *Handbook of the Economics of Education*, E. A. Hanushek, S. Machin and L. Woessmann (eds.), vol. 3. Amsterdam, North-Holland.
Hayashi, F. (2000), *Econometrics*, New Jersey, Princeton University.
Hindrinks, J. and others (2010), "School autonomy and educational performance: within-country evidence", *CORE Discussion Papers*, No. 2010082, Louvain, Catholic University of Louvain.
Hong, G. and B. Yu (2007), "Early-grade retention and children's reading and math learning in elementary years", *Educational Evaluation and Policy Analysis*, vol. 29, No. 4, SAGE.
Jencks, C. (1972), *Inequality: A Reassessment of the Effect of Family and Schooling in America*, New York, Basic Books.
Krueger, A. B. (2003), "Economic considerations and class size", *The Economic Journal*, vol. 113, No. 485, Wiley.

Krueger, A. B. and D. M. Whitmore (2001), "The effect of attending a small class in the early grades on college-test taking and middle school test results: evidence from project STAR", *The Economic Journal*, vol. 111, No. 468, Wiley.

León, G., J. Manzi and R. Paredes (2004), *Calidad docente y rendimiento escolar en Chile: evaluando la evaluación,* Santiago, Catholic University of Chile.

Marteleto, L. (2012), "Educational inequality by race in Brazil, 1982-2007: structural changes and shifts in racial classification", *Demography*, vol. 49, No. 1, Springer.

Martin, A. (2011), "Holding back and holding behind: grade retention and students' non academic and academic outcomes", *British Educational Research Journal*, vol. 37, No. 5, Taylor & Francis.

Medina, F. and M. Galván (2007), "Imputación de datos: teoría y práctica", *Statistical and Prospective Studies*, No. 54 (LC/L.2772-P), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).

Méndez, N. and M. Zerpa (2011), "Desigualdad en las capacidades educativas. Los casos de Uruguay y Chile", *Revista de Economía*, vol. 18, No. 1, Montevideo, Central Bank of Uruguay.

Meunier, M. (2011), "Immigration and student achievement: evidence from Switzerland", *Economics of Education Review*, vol. 30, No. 1, Amsterdam, Elsevier.

Montero, E. and others (2012), *Costa Rica en la pruebas PISA 2009 de competencia lectora y alfabetización matemática*, San José, National Council of Rectors.

Mullis, I. V. S. and others (2007), *PIRLS 2006 International Report. IEA's Progress in International Reading Literacy Study in Primary Schools in 40 Countries*, Chesnut Hill, Massachusetts, Boston College.

Murnane, R. J. (1981), "Interpreting the evidence on school effectiveness", *Teachers College Record,* vol. 83, No. 1.

OECD (Organization for Economic Cooperation and Development) (2015), "Programme for International Student Assessment" [online] http://pisa2012.acer.edu.au/.

___(2014), *PISA 2012 Technical Report*, Paris.

Oreiro, C. and J. P. Valenzuela (2013), "Determinants of educational performance in Uruguay, 2003-2006", *CEPAL Review*, No. 107 (LC/G.2536-P), Santiago, Economic Commission for Latin America and the Caribbean (ECLAC).

Post, D. (2011), "Primary school student employment and academic achievement in Chile, Colombia, Ecuador and Peru", *International Labour Review*, vol. 150, No. 34, Geneva, International Labour Organization (ILO).

Rivas, A. (2015), *América Latina depués de PISA. Lecciones aprendidas de la educación en siete países (2000-2015)*, Buenos Aires, Public Policies for Equity and Growth Implementation Centre (CIPPEC)/ Natura/Natura Institute.

Sánchez, A. (2011), "Etnia y desempeño académico en Colombia", *Documentos de Trabajo sobre Economía Regional*, No. 156, Cartagena, Bank of the Republic.

Shorrocks, A. F. (1982), "Inequality decomposition by factor components", *Econometrica*, vol. 50, No. 1, New York, The Econometric Society.

Thieme, C., D. Prior and E. Tortosa-Ausina (2013), "A multilevel decomposition of school performance using robust nonparametric frontier techniques", *Economics of Education Review*, No. 32, Amsterdam, Elsevier.

Vegas, E. and J. Petrow (2007), *Raising Student Achievement in Latin America: The Challenge for the 21st Century*, Washington, D.C., World Bank.

Viáfara, C. and F. Urrea (2006), "Efectos de la raza y el género en el logro educativo y estatus socio-ocupacional para tres ciudades colombianas", *Desarrollo y Sociedad*, No. 58, Bogota, University of the Andes.

Vignoles, A. and others (2000), "The relationship between resource allocation and pupil attainment: a review", *Working Paper*, No. 2, London, London School of Economics and Political Science.

Vivas, H., J. B. Correa and J. A. Domínguez (2011), "Potencial de logro educativo, entorno socioeconómico y familiar: una aplicación empírica con variables latentes para Colombia", *Sociedad y Economía*, No. 21, Cali, Del Valle University.

Woessmann, L. (2010), "Families, schools and primary-school learning: evidence for Argentina and Colombia in an international perspective", *Applied Economics*, vol. 42, No. 21, Taylor & Francis.

Woessmann, L. and others (2007), "School accountability, autonomy, choice, and the level of student achievement. International evidence from PISA 2003", *OECD Working Papers*, No. 13, Paris, Organization for Economic Cooperation and Development (OECD).

Wooldridge, J. M. (2010), *Introducción a la econometría. Un enfoque moderno*, Madrid, Cengage Learning.

World Bank (2005), "Mexico: determinants of learning policy note", *Report*, No. 31842-MX, Washington, D.C.

Annex A1

Estimation of the model with PISA 2009 data

As a robustness analysis, the estimates in tables 4 to 6 were replicated with PISA 2009 data. As can be seen, the results obtained do not differ significantly from those found for 2012.

Table A1.1

Latin America (10 countries): factors conditioning the acquisition of reading skills, PISA 2009

	ARG	BRA	CHL	COL	CRI	MEX	PAN	PER	URY	VEN
constant	315.57	348.34	285.51	328.80	335.89	308.30	223.94	231.66	337.17	169.93
	(23.610)*	(11.447)*	(28.735)*	(17.931)*	(27.350)*	(7.893)*	(88.753)**	(21.815)*	(11.573)*	(118.951)
non-repeater	143.21	206.51	235.64	196.66	262.58	204.91	265.79	305.64	147.03	346.66
	(30.962)*	(16.371)*	(44.379)*	(24.569)*	(41.803)*	(11.486)*	(120.424)**	(41.728)*	(14.192)*	(172.670)**
sex	22.59	4.59	0.85	3.49	6.67	8.80	3.90	-4.87	18.95	-16.55
	(4.145)*	(3.266)	(5.030)	(3.866)	(4.687)	(1.930)*	(7.975)	(5.282)	(2.857)*	(17.429)
effort	2.08	6.78	4.48	8.02	-3.04	3.62	-2.11	-4.04	2.69	9.06
	(3.926)	(2.889)**	(4.039)	(3.659)**	(4.721)	(1.787)**	(7.868)	(4.503)	(2.576)	(9.491)
discipline	10.61	5.74	8.30	10.58	-6.65	4.93	-5.23	8.86	8.23	-6.38
	(4.243)**	(3.499)	(4.769)***	(5.350)**	(6.194)	(2.395)**	(10.806)	(6.709)	(3.344)**	(11.131)
books	14.83	3.76	26.47	19.54	13.62	18.87	1.72	9.30	17.53	-1.56
	(7.619)***	(7.428)	(6.115)*	(9.011)**	(11.833)	(4.626)*	(14.197)	(10.965)	(4.432)*	(15.731)
educamother	21.48	2.22	17.02	12.93	-13.34	14.41	3.74	-10.33	14.77	3.43
	(5.091)*	(3.351)	(5.342)*	(4.303)*	(6.748)**	(2.052)*	(10.114)	(9.071)	(3.805)*	(15.708)
empfather	9.52	-0.89	-9.14	3.63	-0.78	7.33	-5.39	-5.65	-3.52	21.92
	(6.767)	(3.374)	(6.431)	(5.068)	(7.134)	(2.728)*	(14.243)	(5.960)	(4.314)	(23.214)
public	-28.86	-49.88	-0.56	-21.31	-20.64	-12.15	22.35	-6.34	-30.29	44.24
	(6.415)*	(6.088)*	(5.300)	(6.471)*	(10.110)**	(4.276)*	(54.604)	(8.586)	(6.365)*	(61.015)
STRATIO	0.03	-0.34	-1.02	-0.95	1.11	-0.29	0.60	-0.99	0.79	-0.54
	(0.091)	(0.099)*	(0.233)*	(0.1999)*	(0.479)**	(0.029)*	(0.562)	(0.307)*	(0.212)*	(0.224)**
SCMATEDU	7.38	4.17	4.46	6.50	3.80	2.01	16.20	8.89	1.36	16.29
	(2.170)*	(1.663)**	(1.701)*	(2.040)*	(2.166)***	(1.118)***	(4.956)*	(2.285)*	(1.269)	(4.329)*
autonomy	-4.96	4.96	-3.03	6.39	-19.53	5.76	5.09	13.54	0.36	-4.89
	(4.252)	(2.837)***	(3.872)	(3.629)***	(5.486)*	(2.094)*	(8.141)	(4.604)*	(2.608)	(12.838)
SCHSIZE	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.01	0.00	0.02
	(0.004)	(0.002)	(0.002)*	(0.001)*	(0.006)	(0.001)	(0.016)	(0.004)*	(0.003)	(0.010)***
Observations	2 485	10 976	3 194	5 866	3 403	27 172	1 950	4 686	3 988	1 660

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2009 test.

Note: * significant at 1%, ** significant at 5%, *** significant at 10%. Robust standard deviations in parentheses.

Table A1.2
Latin America (10 countries): factors conditioning the acquisition of mathematics skills, PISA 2009

	ARG	BRA	CHL	COL	CRI	MEX	PAN	PER	URY	VEN
constant	344.07	385.70	294.99	329.84	359.21	334.95	267.97	293.87	373.24	243.21
	(16.278)*	(9.064)*	(26.479)*	(15.719)*	(21.377)*	(7.177)*	(78.365)*	(18.268)*	(9.751)*	(85.576)*
non-repeater	107.51	149.90	217.52	176.96	212.80	179.02	220.13	240.25	131.49	211.27
	(23.083)*	(12.893)*	(41.318)*	(21.674)*	(32.489)*	(10.385)*	(111.873)**	(34.484)*	(11.980)*	(124.853)***
sex	-18.72	-33.39	-38.69	-39.02	-31.79	-28.19	-23.10	-39.87	-33.52	-37.60
	(3.862)*	(2.579)*	(4.646)*	(3.349)*	(3.727)*	(1.764)*	(6.926)*	(4.259)*	(2.478)*	(12.108)*
effort	-1.19	3.05	1.94	8.00	-4.74	4.16	4.83	-2.77	-0.74	2.95
	(3.427)*	(2.267)	(3.789)	(3.198)**	(3.733)	(1.610)*	(6.616)	(3.657)	(2.259)	(5.994)
discipline	8.58	4.24	8.55	10.10	-5.29	3.95	-1.50	6.62	5.66	0.01
	(3.894)**	(2.795)	(4.459)***	(4.585)**	(4.974)	(2.159)***	(8.707)	(5.343)	(2.873)**	(7.102)
books	17.87	7.72	31.09	26.57	8.22	20.94	22.08	11.12	24.99	8.50
	(6.487)*	(5.869)	(5.883)*	(7.468)*	(9.007)	(4.208)*	(14.779)	(8.998)	(4.011)*	(10.598)
educamother	15.68	2.62	15.39	13.13	-6.21	13.27	-6.30	-0.62	14.92	20.59
	(4.314)*	(2.665)	(4.996)*	(3.793)*	(5.367)	(1.858)*	(9.299)	(7.291)	(3.368)*	(10.800)***
empfather	9.93	0.21	-4.50	-5.64	-9.46	4.19	-16.69	-10.26	-2.03	9.75
	(5.885)***	(2.612)	(5.924)	(4.254)	(5.733)***	(2.448)***	(13.003)	(4.875)**	(3.709)	(15.496)
public	-19.86	-65.18	0.40	-18.26	-17.75	-10.21	9.92	-18.28	-24.62	41.30
	(5.087)*	(4.861)*	(4.918)	(5.810)*	(8.043)**	(4.170)**	(49.118)	(7.035)*	(5.493)*	(42.045)
STRATIO	-0.01	-0.22	-1.54	-0.85	0.89	-0.24	0.65	-1.33	0.82	-0.60
	(0.065)	(0.076)*	(0.223)*	(0.172)*	(0.377)**	(0.026)*	(0.538)	(0.254)*	(0.187)*	(0.142)*
SCMATEDU	10.05	2.94	4.76	6.27	5.41	3.34	14.45	7.99	3.56	14.62
	(2.013)*	(1.322)**	(1.585)*	(1.802)*	(1.751)*	(1.098)*	(4.004)*	(1.868)*	(1.103)*	(2.790)*
autonomy	-4.40	2.37	1.88	7.09	-16.71	5.62	3.62	8.77	3.39	-4.34
	(3.710)	(2.235)	(3.600)	(3.192)**	(4.283)*	(1.865)*	(7.572)	(3.751)**	(2.281)	(8.751)
SCHSIZE	0.01	0.00	0.01	0.01	-0.01	0.00	0.00	0.01	-0.01	0.03
	(0.004)***	(0.002)	(0.002)*	(0.001)*	(0.004)	(0.000)	(0.016)	(0.003)*	(0.002)**	(0.007)*
Observations	1 994	10 976	3 194	5 866	3 403	27 172	1 903	4 686	3 981	1 660

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2009 test.

Note: * significant at 1%, ** significant at 5%, *** significant at 10%. Robust standard deviations in parentheses.

Table A1.3
Latin America (10 countries): factors conditioning the acquisition of science skills, PISA 2009

	ARG	BRA	CHL	COL	CRI	MEX	PAN	PER	URY	VEN
constant	317.37	390.64	318.85	367.09	374.51	338.22	149.89	274.51	347.51	242.73
	(30.387)*	(9.809)*	(27.390)*	(15.844)*	(22.095)*	(6.718)*	(169.262)	(18.803)*	(11.098)*	(88.688)*
non-repeater	169.84	167.24	217.65	172.59	207.84	167.99	384.86	253.86	155.80	239.19
	(48.917)*	(13.936)*	(42.761)*	(21.940)*	(33.925)*	(9.701)*	(237.546)	(35.974)*	(13.325)*	(129.713)***
sex	-8.61	-22.94	-26.52	-27.50	-23.94	-20.31	-30.00	-26.64	-22.88	-32.67
	(5.978)	(2.810)*	(4.768)*	(3.390)*	(3.850)*	(1.708)*	(13.680)**	(4.485)*	(2.811)*	(13.220)**
effort	-2.45	4.53	5.12	8.00	-0.13	4.62	-9.72	0.17	1.97	5.72
	(4.308)	(2.483)***	(3.908)	(3.220)**	(3.885)	(1.575)*	(11.934)	(3.847)	(2.551)	(6.877)
discipline	-1.77	4.32	8.85	6.86	-3.25	3.37	-6.73	8.21	10.76	0.21
	(4.826)	(3.021)	(4.569)***	(4.696)	(5.102)	(2.118)	(16.402)	(5.671)	(3.317)*	(7.953)
books	11.39	6.42	31.78	29.05	13.42	21.22	-1.08	8.36	17.62	8.91
	(7.942)	(6.949)	(5.736)*	(7.525)*	(9.424)	(4.092)*	(23.054)	(9.275)	(4.510)*	(11.909)
educamother	20.35	2.93	9.76	10.94	-11.85	16.89	-5.46	-11.13	14.74	15.69
	(7.559)*	(2.918)	(5.091)***	(3.783)*	(5.653)**	(1.823)*	(17.478)	(7.726)	(3.723)*	(12.015)
empfather	2.82	-0.30	-9.71	-4.81	0.68	3.40	-15.58	-7.55	-5.15	26.67
	(8.170)	(2.899)	(6.002)	(4.347)	(5.814)	(2.362)	(24.547)	(5.065)	(4.133)	(16.381)
public	-17.95	-59.28	-2.15	-26.48	-32.77	-15.08	122.05	-14.25	-18.52	15.03
	(9.101)**	(5.150)*	(5.015)	(5.556)*	(8.312)*	(3.769)*	(114.428)	(7.347)***	(6.090)*	(44.987)
STRATIO	0.23	-0.27	-1.25	-1.14	0.74	-0.27	1.22	-0.81	1.00	-0.30
	(0.111)**	(0.087)*	(0.226)*	(0.180)*	(0.391)***	(0.026)*	(1.198)	(0.261)*	(0.205)*	(0.163)***
SCMATEDU	6.39	2.64	3.54	4.50	2.01	2.99	30.17	7.23	1.53	14.88
	(2.825)**	(1.392)***	(1.635)**	(1.765)**	(1.792)	(0.967)*	(9.193)*	(1.959)*	(1.263)	(3.193)*
autonomy	-1.06	1.57	-0.61	2.57	-13.52	7.02	8.93	10.78	0.22	-6.03
	(5.154)	(2.455)	(3.680)	(3.197)	(4.470)*	(1.851)*	(12.664)	(3.918)*	(2.581)	(9.452)
SCHSIZE	0.00	-0.01	0.01	0.00	0.00	0.00	-0.03	0.01	-0.01	0.02
	(0.006)	(0.002)**	(0.002)*	(0.001)***	(0.005)	(0.000)	(0.033)	(0.003)*	(0.003)	(0.008)***
Observations	2 011	10 976	3 194	5 866	3 403	27 172	1 699	4 686	3 929	1 660

Source: Prepared by the authors, on the basis of data from the Organization for Economic Cooperation and Development (OECD) on PISA 2009 test.

Note: * significant at 1%, ** significant at 5%, *** significant at 10%. Robust standard deviations in parentheses.