

**KEY WORDS**

Quality of education  
Academic performance  
Measurement  
Private schools  
Public schools  
Social aspects  
Argentina

# Do private schools in Argentina perform better because they are private?

*María Marta Formichella*

**T**he objective of this study is to analyse the determinants of the quality of education in Argentina and, in particular, to look at what influence a school's ownership structure has. A multilevel regression model and 2006 data from the Programme for International Student Assessment (PISA) were used for this purpose. One of the main findings is that the correlation between a school's administrative structure (public or private) and its students' scholastic performance fades when the socioeconomic school environment is taken into consideration.

María Marta Formichella  
Assistant Professor for the  
Seminar on Economic Accounting  
and Fundamental Principles  
of Economics, Department of  
Economics, Universidad Nacional  
del Sur (UNS)

✉ [mformichella@uns.edu.ar](mailto:mformichella@uns.edu.ar)

# I

## Introduction

Countless research projects have shown how important education is as a determinant of socioeconomic development. The beneficial effects of education have been shown to range from employment-related advantages to less tangible ones having to do with the fact that educated persons may have a better quality of life than uneducated ones (Formichella, 2010).

One of the relevant considerations in this connection is the differing levels of scholastic achievement, in terms of both quantity and quality, that a given individual can attain. In quantitative terms, 51% of the people between the ages of 25 and 65 do not have a high school diploma, and the educational Gini coefficient, calculated on the basis of completed years of schooling, is 0.21. In qualitative terms, the rating on the Theil index –calculated on the basis of scores on the 2006 PISA science test–is 0.059, with 41% of this difference being accounted for by differences between schools (Formichella, 2010).

The existing literature on determinants of scholastic performance in Argentina indicates that qualitative differences in achievement can be accounted for by the traits of individuals and their households, characteristics of the schools that they attend, and, according to many of these papers, the socioeconomic background of a student's classmates.

While these studies agree on many points, there is one issue about which they differ, and that is how the type of administrative structure influences academic

achievement. This divergence is also seen in the findings of empirical studies conducted in other countries.<sup>1</sup>

As we will see, some empirical studies focusing on education in Argentina have found that students' performance is strongly influenced by whether the school is privately or publicly run, while others have found that this correlation is eclipsed by the socioeconomic environment of the students enrolled in each type of school.

The aim of this study is therefore to find evidence to settle this difference of opinion. Its working hypothesis tends to favour the second of these positions (i.e. that the influence that the form of school administration appears to have on students' performance is actually attributable to the effect that the socioeconomic school environment has on scholastic achievement).

The sections that follow this introduction will explore different aspects of this question. Section II presents a review of the major studies that have focused on the determinants of scholastic performance in Argentina. Section III examines the data sources used for this study. Section IV covers the methodology, while section V outlines the proposed models and their results. The conclusions are presented in section VI.

---

<sup>1</sup> For a review of these studies, see Calero and Escardíbul (2007).

# II

## Background information about the education system in Argentina

Various empirical studies have explored the determinants of scholastic performance in Argentina. Cervini (1999) analyses the performance (as measured by scores on mathematics tests) of students in the seventh year of the basic general education track using qualitative information from the National Evaluation Campaign (ONE) of 1997. Using a hierarchical methodology applied at two different levels (schools and students), he finds that the socioeconomic level of the household and of the school

environment have a significant impact. In addition, the following factors were identified as having a negative impact on achievement: (i) if a student is over-age;<sup>2</sup> (ii) if a student has repeated a grade; (iii) if a student

---

<sup>2</sup> "Over-age students" are defined as students who are two or more years behind. This factor is measured by the number of students in different age groups who are attending grade levels below those that would usually correspond to their age group.

has changed schools; (iv) if a student lives far away from the school; and (v) if the student is female. This author also finds that self-motivation, personal effort and the student's view of how much help he or she receives from family members have a positive impact. Finally, he finds evidence that better school infrastructure and better-quality educational resources have positive effects on performance, but he does not conclude that the fact that a school is privately or publicly run has any significant explanatory value.

In 2002, Cervini used 1997 ONE data to study scholastic achievement at the primary level. For this study, he used a three-level hierarchical methodology based on provinces, schools and students. He found that differences among students and, after that, schools were the most powerful explanations for the differences in achievement, while differences at the provincial level were of virtually no importance. His main conclusion is that the theory of cultural reproduction holds true for Argentina, with scholastic performance being largely accounted for by households' cultural and economic capital. He also finds that the influence of socioeconomic factors varies a great deal across schools. He concludes that this variance is attributable to different schools' varying abilities to offset underlying inequalities. His ultimate conclusion is that, of all the various factors that have an influence on performance, the one that has the greatest impact of all is the socioeconomic status of students' classmates.

Cervini (2002a) analyses the determinants of scholastic performance at the middle-school level using data from the national secondary-school completion survey conducted in 1998. He uses the same tri-level (provinces, schools and students) analytical methodology as in the preceding study and, here again, finds that inequalities in outcomes are primarily found at the individual student level, although a significant percentage of the existing inequalities can also be accounted for by inter-school differences. As in the preceding case, he found that inter-province differences had very little explanatory power. He also found that there was a negative correlation between being female (male) and performance in mathematics (reading) and that the economic assets of a household were a determining factor for access to middle school, while a household's cultural assets had a positive effect on performance. Here again, he found that schools differ in their ability to offset inequalities in the students' backgrounds.

Fernández Aguerre (2002) uses ONE data from the year 2000 and a logistic regression methodology to study the determinants of success in school for sixth-grade students in Argentina. Like Cervini, he finds evidence

of a positive correlation between the socioeconomic status of students' households and their performance. He also finds that a school's sociocultural environment has more of an impact on student outcomes than their households' economic status does. His results indicate that students who attend private schools outperform those who attend public schools and that men perform better than women.

Wößmann and Fuchs (2005) also study scholastic performance at the primary-school level in Argentina using clustering-robust linear regressions (CRLRS) and data from the Progress in International Reading Literacy Study (PIRLS) of 2001. These authors identify household characteristics and family background as the most influential factors in determining a student's scholastic performance. They do not find convincing evidence of a correlation between academic achievement and how well-funded a school is or other school characteristics. They do find, however, that a school that uses a centralized curriculum or that uses a skills-building approach has a positive effect on performance.

Cervini (2005a and 2006) then took his 2002 middle-school study on math performance to another level. He found that being a female, having repeated a grade or working longer hours all have an adverse effect on scholastic performance and that the degree of institutional selectivity based on students' socioeconomic status was quite high. He also found evidence of a positive correlation between the school environment and performance. In addition, his findings indicate that schools' capacity to offset household-based differences varies and that inequities related to inter-school inequalities of effect diminish when schools of a similar make-up are compared to one another.

Cervini (2005b) has also studied non-cognitive (as well as cognitive) educational outcomes linked to general socio-educational attitudes, which he measures using other variables gleaned from the student questionnaire that he then uses to capture related subjective factors. He uses a tripartite (students, schools, provinces) hierarchical model for this purpose together with data from the national secondary-school completion survey conducted in 1998. His comparison of the determinants of these two types of results leads him to the following conclusions: (i) individuals' traits have a greater impact on cognitive than on non-cognitive outcomes; (ii) schools' ability to offset inequalities of origin relative to cognitive and non-cognitive outcomes differ; and (iii) schools' performance in terms of their students' outcomes are inconsistent (i.e. there is no strong correlation between cognitive and non-cognitive outcomes).

Gertel and others (2006) have also used ONE data (mathematics and reading test results at the primary school level for 2000), together with a dual-level (students and classrooms) hierarchical model, to study the determinants of academic performance in Argentina. Their findings underscore the importance of grade-repetition and gender variables: having repeated a grade lowers performance and being a male improves it (lowers it) in mathematics (reading). Their results in terms of household-related variables indicate that having siblings who have dropped out of school or who never attended in the first place depresses student performance, while there is a positive, although fairly weak, correlation between socioeconomic status and educational outcomes. They also find that a number of school-related variables, including the quality of infrastructure, teacher experience and training, and the type of school administration, have a significant impact (with students attending private schools scoring higher).

In view of their results concerning the impact of the private/public school variable, Gertel and others (2007a) followed up with a tripartite model (students, grades and schools) in which the third level (schools) is represented by all the different variables linked to the public/private differential. Their findings corroborate the conclusion that the school's ownership structure has a greater influence on actual learning outcomes.

These same researchers (Gertel and others, 2007b) later looked at scholastic performance using not only ONE data, but also 1997 scores of primary-school students on the international Latin American Laboratory for Assessment of the Quality of Education (LLECE). They use a hierarchical data methodology and the same three levels as before. One of their main findings is that the public/private school variable has a strong explanatory value in terms of scholastic performance in Argentina.

Santos (2007) has also studied the determinants of educational quality in Argentina at the middle-school level, but uses PISA scores from 2000 on mathematics and reading tests and quantile regressions to do so. She found that gender is an influential variable, with females scoring higher on reading tests and males higher on mathematics tests. She also finds that class size should be limited to a maximum of 32 students in order to achieve positive learning outcomes and that the quality of a school's educational resources has a major impact on performance. She also finds that, after controlling for other features, private schools actually do not have better results than public schools, and she posits that the differences in scores may be attributable to the fact

that the private-school students come from a higher socioeconomic stratum. She also concludes that the presence or absence of educational resources in students' homes has a highly significant effect.

Abdul-Hamid (2007) also analyses the determinants of educational quality in Argentina using quantile regressions and 2000 PISA scores. He finds that the factors that have a positive influence on scholastic achievement include the following: (i) the school learning environment; (ii) the geographic location of the school (schools in larger cities perform more successfully); (iii) the academic guidance and encouragement provided to students; (iv) the type of school administration (private-school students outperform public-school students); (v) a majority of women in the student body; (vi) whether or not the student has education resources in the home; and (vii) whether or not the student's mother has an above-average level of education.

Finally, Cervini conducted two studies in which he compares the educational levels of primary and secondary students in Argentina using ONE (2000) data, the results of the 1998 national secondary-school completion survey and a multilevel methodology.

In the first of these studies (Cervini, 2009), he uses three levels of analysis: students (level 1), schools (level 2) and provinces (level 3). His objective is to compare the pattern of (in)equity in the distribution of scores on reading and mathematics tests. He finds that schools differ markedly in terms of their students' performance in the evaluations conducted at the two different grade levels and that the intra-school variance was greater at the primary education level. He therefore concludes that the characteristics of a student's household are more influential during the primary education cycle. He also finds that, in general, all the variables associated with the students' socioeconomic status influence their test scores at both grade levels and that the variables associated with the school environment are also significant factors at both levels, although the economic status of classmates exerts a greater influence in the primary education cycle, whereas the students' educational environment, proxied by the level of education of their parents, has a greater impact in the secondary cycle.

In the second study (2010), Cervini uses various models to disaggregate information at different levels (in which he considers, alternatively, three of the following variables: students, classrooms, schools, municipalities and provinces) and focuses on an analysis of a comparison of the "school effect" across primary and secondary grade levels. He concludes that the results are determined by the methodology that is employed, that the school effect

in Argentina is close to what it is in developed countries, and that no significant differences can be detected between the primary and secondary levels of education.

In drawing this section to a close, we can synthesize some of the results as follows: the studies discussed above appear to represent a consensus view as to the impact of the household's socioeconomic environment on academic performance, as well as the school environment. As for the variable that, based on our stated objective, is of the greatest interest to us here, some of the authors use the

form of school administration (public/private) as an explanatory value, whereas others do not, and even among those who do use this variable, a consensus is lacking. Some reach the conclusion that it is an important variable in explaining scholastic achievement (Fernández Aguerre, 2002; Gertel and others, 2007a and 2007b; Abdul-Hamid, 2007), whereas others conclude that it loses explanatory value and actually becomes irrelevant when the model is controlled for school-related socioeconomic variables (Cervini, 2010; Santos, 2007).

### III

## The data

The regression analysis employed in this study<sup>3</sup> was performed on data from the 2006 PISA tests administered by the Organization of Economic Cooperation and Development (OECD), which are designed to measure the knowledge and skills of 15-year-olds in different countries in order to determine how well prepared they are to participate fully in society as adults. The assessment is composed of a series of mathematics, science and reading tests. It is conducted once every three years, with a particular area being given priority each time. The 2006 PISA emphasized knowledge and skills in the field of science (OECD, 2006a).

The PISA assessments are supplemented with information on the characteristics of each student's household and school. The household data are collected by means of a questionnaire that the students fill out at the time that they take the tests. In some countries, parents also answer a questionnaire. Information on the schools is gathered by interviewing their directors.

Scores on the PISA test go from 0 to 1,000 and are grouped into six brackets. The cut-off scores for each bracket vary from one skill set to the next (OECD, 2006a).

The results of these assessments are presented using "plausible values" (PVs), which represent the sum total of a student's proficiencies. Arriving at PVs based on this type of evaluation entails extrapolating a continuum from a group of discontinuous variables (the test scores) (OECD, 2003).

Martínez Arias (2006) explains that the problem which the PISA programme has to deal with is that each individual answers only a limited number of questions, so it becomes necessary to estimate how he or she would have answered all the others. These responses are predicted on the basis of the answers to the questions to which the individual did respond and of what are known as "conditioning" variables, which are gleaned from the background questionnaire. Thus, rather than making just one prediction, a distribution of values is generated a posteriori for each student, along with the probabilities associated with each of those values. Five values are taken from this distribution at random; these are the PVs. This approach avoids the bias that would be generated if a (non-observable) proficiency were to be predicted on the basis of a limited number of observable data. The PISA team uses tailor-made software to estimate these values.

The population statistics and the parameters for the regression models have to be estimated using each of the PVs separately. The value of each population statistic will be equal to the mean of the estimates for that population statistic using each of the five PVs generated from the PISA test. The same is true of the model parameters: the value of each parameter will be equal to the mean of the estimates for that parameter arrived at using each of the five PVs (OECD, 2003).

Although the countries that take part in the PISA programme are not always exactly the same ones from one test to the next, they are always in one of two groups: OECD member countries and OECD partner countries. Argentina, which falls into the second group, participated in the PISA assessments for 2000, 2006 and 2009.

<sup>3</sup> This econometric regression methodology is explained in section IV.

At the time of writing, the latest available data are from the 2006 PISA assessment, and this is therefore the database used here. In Argentina, the 2006 PISA test was administered between 14 and 16 August 2006. A total of 4,339 students of 15 years of age in 179 different schools took part. The sample used for this study was reduced to 3,860 observations owing to data losses. According to OECD (2003), this does not distort the results, however, since the reduction in data amounts to somewhat less than 10%.<sup>4</sup>

<sup>4</sup> In order to avoid having to dispense with a larger number of observations, some variables have not been used in this analysis. To this end, complex imputation techniques have not been used, since they would make it necessary to estimate the value of the missing variables using regression analyses that would make use of the rest of the explanatory variables. In the case of a number of the observations,

It should also be pointed out that the sample for the PISA programme in each country is constructed using a two-stage procedure: first the schools are selected and, then, the students within those schools are chosen. This is why, as we will see in the following section, it is best to use multilevel models for analysing scholastic performance.

data for more than one variable were missing. Santos (2007) explains that this can cause more problems than it solves. In these cases, a simple imputation technique (whereby the missing datum would be replaced with the mean value for the sample) was not used either because, since a very large percentage of data were missing, failing to use an imputation technique based on the characteristics of the observation may distort the results. However, as we will see later on, in some cases (the CULTPOSS, SCMATEDU and CLSIZE variables), the simple imputation technique is used because the percentage of missing data was small.

## IV Methodology

This study uses multilevel regression techniques to arrive at its findings. Cervini (1999) characterizes this as an appropriate correlation technique for examining variations in the characteristics of individual members of a group (e.g., in the case, schools).

In multilevel regression analyses, sample units are nested within larger units. Rather than calculating a regression equation for the entire dataset, a regression equation is calculated for each of these larger units (OECD, 2003). Therefore, these models are the most suitable ones to use when the data are grouped together, since they incorporate information about the nesting structure of the data.

In addition, when the population sample is selected in stages, we are dealing with a hierarchical multilevel analysis, and the observations within each group are therefore not independent (Hox, 1995). In other words, when we first select a larger structure and, then, select specific cases within it, we have created a multilevel structure. One advantage of using a multilevel analysis in this case is that it makes it possible to study the effects that both the group variables and the individual variables have on the individual results at the same time. It also incorporates the intra-group dependence of the observations (Diez Roux, 2002).

Thus, the use of multilevel models allows us to obtain better estimators of the regression coefficients and

their variance than we could achieve using traditional models such as ordinary least squares regressions (De la Cruz, 2008). Moreover, the standard statistical tests rely heavily on the assumption that the observations are independent, and if this assumption is violated (as occurs in multilevel structures), the estimates of standard errors on conventional statistical tests are much smaller and their “significant” results are spurious (Hox, 1995).

One example of such a multilevel structure is the case of schools composed of different classes which are in turn made up of different groups of students. This is why this type of analysis is suitable for studying the determinants of scholastic performance (Hox, 1995; Bryk and Raudenbusch, 1988, cited in Calero, Choi and Waisgrais, 2009). When information on all three of these levels is available, then we have a tri-level hierarchical model, and when information is available only on the schools and the students involved, then we have a dual-level hierarchical model.

Multilevel models may incorporate fixed and random effects. In its simplest form, this occurs when inter-group variance (random effects) appears only in the intercept; more complex forms emerge when this variance is present in the coefficients of explanatory variables. In the first case, there is a straight regression line for each group and all the lines are parallel. In the second case, the lines will also have different slopes.

The fact that the intercept is defined by random effects is crucial; if it were to be defined by fixed effects, there would be no point in using a multilevel model of analysis (OECD, 2003).

Since hierarchical models involve different categories, there are variables to be considered within each one of those categories. For example, in the case of an assessment of scholastic performance in which two levels (centres<sup>5</sup> and students) are examined, some variables will refer to the centre (level 2) and others to the students (level 1).

The level 2 variables will be identical for all the students that are in the same school, and these variables will therefore influence only the intercepts for the schools. The level 1 variables can be incorporated via fixed or random effects depending on the theoretical aspects of the subject under study and the research objectives (OECD, 2003). If a level 1 variable is incorporated via fixed effects, the assumption is that there is no difference between schools with regard to that variable's effect on the dependent variable. On the other hand, if a level 1 variable is included by assigning it random effects, the assumption is that its effect on the dependent variable differs across schools.

In the case of school performance, if the effects are fixed, then the regression coefficients provide information about what happens in terms of scholastic achievement —within a given school— when the value of an explanatory value is altered (*ceteris paribus*). Since the effects are fixed, this coefficient will be the same for all schools (OECD, 2003).

By the same token, if the effects are random, then the coefficients are interpreted in the same way, but their values will be different in each school. These coefficients can be divided into two parts: a fixed aspect (which is shared by all the schools), and a random one (which represents the residual distance between the coefficient for each school and the shared aspect) (OECD, 2003).

Formally, assuming that we are dealing with a model of scholastic performance made up of two levels with three explanatory variables (one at level 2 (P) and two at level 1 (X for fixed effects and Z for random ones)), then the above explanation can be expressed as follows:

— Level 1:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + \beta_{2j} Z_{ij} + r_{ij} \quad (1)$$

<sup>5</sup> The PISA programme uses the word “centre” as a synonym for “school”.

— Level 2 :

$$\beta_{0j} = \gamma_{00} + \gamma_{01} P_j + \mu_{0j} \quad (2)$$

$$\beta_{1j} = \gamma_{10} \quad (3)$$

$$\beta_{2j} = \gamma_{20} + \mu_{2j} \quad (4)$$

— Complete model:

$$Y_{ij} = \gamma_{00} + \gamma_{01} P_j + \gamma_{10} X_{ij} + \gamma_{20} Z_{ij} + \mu_{0j} + \mu_{2j} Z_{ij} + r_{ij} \quad (5)$$

where:

$Y_{ij}$  is the (expected) educational outcome for student “*i*” in school “*j*”.

$\beta_{0j}$  is the intercept of the straight regression line for school “*j*”.

$\beta_{1j}$  is the coefficient associated with explanatory variable *X*, which is incorporated into the model via fixed effects; consequently,  $\beta_{1j}$  is the same for all centres and is represented by  $\gamma_{10}$ .

$\beta_{2j}$  is the coefficient associated with explanatory variable *Z*, which is incorporated into the model via random effects; consequently, it varies across centres. It is composed of a fixed part ( $\gamma_{20}$ ) and a random part ( $\mu_{2j}$ ). The latter represents the residual distance between the regression coefficient for the centres and  $\gamma_{20}$ .

$\gamma_{01}$  is the coefficient that is associated with explanatory variable *P* (the one and only variable at level 2, which is why its subscript is “01”). Since it is a level 2 variable, there is a different *P* value for each centre “*j*”. As can be seen from the equation, this variable influences the value of the intercept  $\beta_{0j}$ .

$r_{ij}$  is the residual variance within each centre.

$\mu_{0j}$  is the residual variance between centres.

In multilevel analyses, it is helpful to estimate a model without including any explanatory variables. This “null model” provides information about how much of the inequality seen in scholastic achievement scores is attributable to inter-centre differences and how much is attributable to intra-centre differences. Formally, this model can be depicted as:

$$Y_{ij} = \gamma_{00} + \mu_{0j} + r_{ij} \quad (6)$$

where:

$\gamma_{00}$  represents the fixed or deterministic effects (global intercept).

$\mu_{0j}$  and  $r_{ij}$  represent the random or stochastic effects.

In this case, the intercepts for each centre ( $\beta_{0j}$ ) are equal to or very close to the means for the centres. As can

be seen,  $\beta_{0j}$  is made up of a fixed component, which is the same for all schools ( $\gamma_{00}$ ), and a random component ( $\mu_{0j}$ ), which represents the deviation of school “j” from  $\gamma_{00}$ , with  $\gamma_{00}$  being the result for the schools as a group (the mean), i.e. the “global intercept”. Thus,  $\mu_{0j}$  is the deviation of school “j” from the global intercept and represents the variance across the different centres.

For its part,  $r_{ij}$  is the deviation of the result for individual student “i” from the mean for school “j” (the school that the student attends). Since, in this model, each student is assigned the mean value for his or her school as a predicted score,  $r_{ij}$  is equal to the variance within each school.

Therefore, when there are no explanatory variables in the model, the residual and estimated intra- and inter-centre variances are equal to each other (OECD, 2003).

In sum, there are generally two types of relevant indices in multilevel analyses: regression coefficients and the decomposition of the variance between the different levels. A frequently used indicator, known as the “intra-class correlation coefficient” ( $\rho$ ), can be derived from the second of these indices. This coefficient represents the proportion of the residual variance that is explained by inter-school differences ( $\rho = \mu_{0j} / (\mu_{0j} + r_{ij})$ ). In the null model, it represents the proportion of the variance in educational outcomes between schools. If this value were zero (0), then there would be no point in using a multilevel model.

When the models include explanatory variables, multilevel analyses usually provide information about the variance in the residual, i.e. the variance in the results that is not explained by the predictor variables that have been included in the model. However, it is interesting to look at the proportion of the variance that is explained. This can be obtained by comparing the proposed model with the null model at each level and overall. The way that this variance can be calculated is detailed in table 1:

TABLE 1

**Percentage of variance accounted for by variables over the null model**

Total	$1 - (\mu_{0j} + r_{ij})$ Proposed model / $(\mu_{0j} + r_{ij})$ Null model
Level 1 (students)	$1 - (r_{ij})$ Proposed model / $(r_{ij})$ Null model
Level 2 (schools)	$1 - (\mu_{0j})$ Proposed model / $(\mu_{0j})$ Null model

Source: Author’s original calculations based on Organization for Economic Cooperation and Development (OECD), *El programa PISA de la OCDE. ¿Qué es y para qué sirve?*, Paris, 2003.

It will be helpful to provide a general description of the econometric models used in this study. As noted in the preceding section, data at the classroom level are not available, and a two-level hierarchical model (students and schools) is therefore used. In addition, variance between schools is permitted in the constant term and the slope coefficients. In other words, there can be variance not only in the intercepts of the straight regression lines for each school, but also in the coefficients associated with some of the level 1 explanatory variables. The formal expression of these models is as follows:

Level 1

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^K \gamma_{k0} X_{kij} + \sum_{m=1}^M \beta_{mj} Z_{mij} + r_{ij} \quad (7)$$

Level 2

$$\beta_{0j} = \gamma_{00} + \sum_{f=1}^F \gamma_{0f} P_f + u_{0j} \quad (8)$$

$$\beta_{mj} = \gamma_{m0} + u_{mj} \quad (9)$$

where:  $K$  is the number of level 1 explanatory variables which are incorporated into the model via fixed effects ( $X$ );  $M$  is the number of level 1 explanatory variables which are incorporated into the model via random effects ( $Z$ ); and  $F$  is the number of level 2 explanatory values ( $P$ ). (The rest of the nomenclature has already been defined above.)

The Hierarchical Linear and Nonlinear Modeling for Windows (vHLM) Programme was used to calculate the parameters for the above equations, which are estimated using simultaneous iteration methods that maximize the probability function.

In all of the models discussed here, observations are weighted using the  $w\_fstuwt$ <sup>6</sup> weighting variable provided in the PISA database, and the weightings are standardized using the specific option offered by the vHLM software. The five plausible values for the “score on science test” variable (since this was the area emphasized in the 2006 PISA assessment) are used as the dependent variable, and the programme generates the final mean result for the population estimators. Finally, in order to preclude any problems in the significance tests in the event of heteroskedasticity, the parameters calculated for the equations and their corresponding standard errors are robust to heteroskedasticity.

<sup>6</sup> “Weights are therefore inversely proportional to the probability of selection.” (OECD, 2003).



## V

### The models and their results

Multilevel models are proposed in this section as a means of attaining the research objective. As mentioned earlier, in order for a hierarchical analytical model to be appropriate and to produce meaningful results, there must be differences between the nesting units (in this case, the schools). In order to make sure that multilevel models are suitable for this purpose, we seek to determine how much of the divergence in performance is attributable to inter-school disparities and how much to intra-school differences. To this end, we estimate the null model, as explained in section IV (methodology). This model will be called model 1.

As shown in table 2, over half of the differences in performance are due to differences between schools, which indicates that a considerable degree of inter-school inequality exists. This confirms that it makes sense to use multilevel models to study the determinants of scholastic achievement. Cervini (2002b) explains that this is important in order to separate out the variance of each group and thus improve the quality of the estimates.

In analysing the different factors that may influence academic achievement, the effect of the type of school administration (public/private) is of particular interest here. It is observed that, on average, students in private schools score higher in science than those in public schools do (442.34 versus 377.02).<sup>7</sup> The next question

to be answered is therefore whether these differences are significantly correlated with the type of school administration or if there are other factors that account for this disparity. Earlier writings (Calero and Escardíbul, 2007; Santos, 2007; Cervini, 2002a, 2005a, 2005b and 2009; among others) indicate that it may be useful to include and analyse the school environment in which individual students find themselves. This factor is also known as the “peer effect”.

The concept of the peer effect is based on the idea that students do not learn only from their teachers but also by interacting with their schoolmates; this interaction provides learning experiences that lead to the development of both skills and knowledge. In addition, a series of forms of stimulation are also transmitted (motivation, standards of behaviour) that are often an indirect reflection of parents’ habits (Mediavilla, 2010).

In this study we are looking at whether the school environment makes an important difference in terms of educational outcomes and whether this difference influences the variable that indicates whether the school is publicly or privately run. Since there are a number of possible models, the possible variables that could be used as control variables for those options will be described below. This will be followed by a detailed discussion of the two variables that appear to be the best options for achieving the proposed objective. Table 3 provides a characterization of the sample as a function of the different explanatory variables.

- (a) *Variables at the student level (their own characteristics and those of their households):*
- Sex (gender): dichotomous; male = 1.
  - Age: continuous; a student’s age is calculated as the difference between the year and month that the test is administered and the year and month that the student was born. A special formula is used to obtain the inter-observation variance.<sup>8</sup>
  - Upper secondary school (secondary\_level): dichotomous; students at the “upper secondary” level (i.e. in either of the last two grades of secondary school) = 1; students at the “lower secondary level”

<sup>7</sup> The percentage of independent (i.e. unsubsidized) private schools is quite small (7.9%), and these institutions are therefore grouped together with subsidized private schools.

TABLE 2

#### Null model results

Dependent variable	Plausible values in science
Value of the intercept	366.51
Intra-centre variance ( $\sigma^2$ )	4 143.37
Inter-centre variance ( $\tau^2$ )	5 725.91
Total variance	9 869.28
Percentage of residual variance accounted for by differences between students	42%
Percentage of residual variance accounted for by differences between centres ( $\rho$ )	58%

Source: Author’s original calculations based on the Programme for International Student Assessment (PISA) database, 2006.

<sup>8</sup> Age = (100 + year of test – year of birth) + (month of test – month of birth)/12.

TABLE 3

## Categorization of observations as a function of the different explanatory variables

Quantitative variables	Mean	Standard deviation
Dependent variable		
Plausible values in science	398.28	98.78
Explanatory variables: level 1 (students)		
Age	15.69	0.28
Occupational status of parents (HISEI)	45.81	16.88
Educational level of parents (PARED)	12.25	4.35
Cultural assets of the household (CULTPOSS)	-0.21	0.87
Educational resources of the household (HEDRES)	-0.67	1.05
Economic status of the household (WEALTH)	-1.29	0.89
Socioeconomic status of the household (ESCS)	-0.61	1.14
Household possessions: wealth and educational and cultural resources (HOMEPOS)	-1.05	1.01
Independent variables: level 2 (schools)		
Class size (CLSIZE)	30.81	9.78
Proportion of female students (Proportion of girls)	0.53	0.16
Faculty responsibility in connection with curriculum (RESPRES)	-0.39	0.51
Faculty responsibility in connection with resource allocation (RESCURR)	0.29	0.86
Quality of educational resources (SCMATEDU)	-0.69	1.34
Shortage of teachers (TCSHORT)	-0.19	0.97
Mean socioeconomic level (PROESCS)	-0.70	0.78
Qualitative variables	Category	Frequency
Explanatory variables: level 1 (students)		
Sex (gender)	Male	46.2%
	Female	53.8%
Upper secondary school (secondary level)	Upper	71.2%
	Lower	28.8%
Place of birth (native)	Argentine	97.5%
	Not Argentine	2.5%
Explanatory variables: level 2 (schools)		
Form of administration (public)	Public	67.4%
	Private	32.6%

Source: Author's original calculations based on the Programme for International Student Assessment (PISA) database, 2006.

(i.e. in one of the first three grades of secondary school) = 0.<sup>9</sup>

- Occupational status of parents (HISEI): continuous; this OECD index reflects the status of the father or mother, whichever is higher.

<sup>9</sup> All the students who participate in the assessment are 15 years old, but they are divided into three groups: those who have repeated a grade and are attending a lower grade than other students of their age; those attending the grade that corresponds to their age and, given the beginning and ending dates of the academic year, are in the second year of secondary school; and those who are attending the grade that corresponds to their age but that, given the beginning and ending dates of the academic year, are in the third year of secondary school. The students in the first two groups are in lower secondary school (i.e. one of the first two years of secondary school) while those in the third group are in upper secondary school (i.e. one of the last three years of secondary school).

- Educational level of parents (PARED): continuous; this OECD index measures the number of years of schooling completed (of either the father or mother, whichever is higher).
- Place of birth (native): dichotomous; native = 1; first- or second-generation immigrant = 0.
- Cultural assets of household (CULTPOSS):<sup>10</sup> continuous; this OECD index reflects the amount of such assets by measuring such variables as the

<sup>10</sup> Some of the data for this variable are missing. Since the percentage of missing data is small (3% of the observations), however, the mean value is used in those cases. In addition, a flag variable is added to indicate how many observations include imputed data. This variable is designated as CULTPOSS\_O; it is dichotomous. Original CULTPOSS value = 1; presence of imputed value = 0.

availability of books (classic literature and poetry) in the home and the presence or absence of works of art. The higher the ranking on this index, the higher the household's cultural level.

- Educational resources of the household (HEDRES): continuous; this OECD index reflects the amount of such resources (e.g., whether there is a desk and a quiet place to study, if a computer is available for schoolwork, if it is loaded with an educational computer programme, and if reference books and a dictionary are present). The higher the ranking on this index, the higher the level of educational resources in the household.
- Economic status of household (WEALTH): continuous; this OECD index measures the possessions present in the student's household (e.g., whether or not the student has his or her own room and whether the household has an Internet connection, washing machine, DVD player, refrigerator with freezer, landline telephone and cable television). It also registers the number of mobile phones, televisions sets, computers and motor vehicles owned by the household. The higher the value of the index, the higher the household's economic status.

(b) *Variables at the school level:*

- Class size (CLSIZE):<sup>11</sup> indicates the average number of students per class.
- Proportion of female students: indicates the percentage of female students in the school.
- Faculty responsibility in connection with the curriculum (RESPRES): this OECD index reflects the degree of responsibility that the teaching staff has in regard to curricular issues. Higher rankings indicate greater levels of responsibility.
- Faculty responsibility in connection with resources (RESCURR): this OECD index reflects the degree of responsibility that the teaching staff has in regard to resource allocation. Higher rankings indicate greater levels of responsibility.
- Quality of educational resources (SCMATEDU): this OECD index represents the quantity and quality of the school's educational resources. (It includes measurements of laboratory equipment, books, computers, Internet connections, audiovisuals, etc.)

Positive values indicate that the school provides good conditions in terms of educational resources; negative values indicate the opposite. Higher positive values signal a greater stock of educational resources.

- Shortage of teachers (TCSHORT): this OECD index measures any teacher shortages. Higher values indicate that a more serious problem exists owing to a lack of qualified teachers.

Variables relating to the prior individual performance of students are not included here. This is because academic performance is usually recurrent, and using past individual performance to account for current performance may not be appropriate (Viego, 2006).

Among the main variables considered in this study, the information available on the type of ownership structure is represented by a dichotomous variable that has been labeled "public"; thus public schools = 1 and private schools = 0. The characteristics of the student body at the school level are used to gauge the school environment, since the 2006 PISA does not report information at the classroom level. This environment is represented by the variable "Socioeconomic level of the school". In order to quantify it, the mean ESCS indicator for the school is used.

The ESCS indicator is constructed by the PISA OECD team and sums up the information provided by the HISEI, PARED and HOMEPOS indices. HISEI and PARED have been explained above. HOMEPOS amalgamates the information supplied by the WEALTH, HEDRES and CULTPOSS indices, together with the number of books in the home. In short, ESCS represents the household's socioeconomic status. This index is designed in such a way that a positive value indicates that the household's socioeconomic status exceeds the mean for OECD countries, while a negative value indicates the opposite. The higher the value of the index, the higher the household's socioeconomic status.

The results of the estimated models will now be explored. As noted earlier, model 1 includes only the constant term. In contrast, explanatory variables have gradually been incorporated into the other models (first, variables relating to the students and, later, variables relating to the schools).

Consequently, we now propose that all the level 1 variables be incorporated as independent variables. These are first incorporated with fixed effects (model 2). Later, another model is proposed (model 3)<sup>12</sup> that includes

<sup>11</sup> There are some (a few) data missing for this variable and for SCMATEDU. Where information is not available, the mean value is imputed. There are also two flag variables (CLZISE\_O and SCMATEDU\_O) which, as in the case of CULTPOSS\_O, show whether the observation includes an imputed datum or not.

<sup>12</sup> No problems of multicollinearity arise in models 2 or 3; the mean result for the variance inflation factor (VIF) is 1.2 and does not surpass 5 for any of the variables. The value of 5 is taken as a benchmark in line with Calero and Escardibul (2007).

the same explanatory variables as model 2, together with random effects for those variables in model 2 that proved to be statistically significant. Finally, we seek to determine whether these random effects are significant. Table 4 shows the results for these models.

In the case of model 2, the variables that turned out to be significant (with a confidence level of 95%) are “SecLevel”, HISEI, PARED and HEDRES.<sup>13</sup> The results can be summed up as indicating that students who live in households with a better educational environment (represented by the occupational status and educational level of the parents) and more educational resources (represented by HEDRES) score higher. Model 2 thus indicates that education is not equitable, since individual students obtain differing educational outcomes as a function of the characteristics displayed by their households, and schools do not succeed in offsetting these differences of origin or evening out the outcomes. In addition, students who are already in upper secondary school also outperform the others.

The random effects included in model 3 are statistically significant at less than 5%.<sup>14</sup> This implies that the extent to which a scarcity of educational resources in the home or an unfavourable environment for studying depresses performance differs depending on the school which the students attend. In other words, students’ limitations of origin may be counterbalanced to a greater or lesser degree by the school or may not be compensated for at all.

Next, in order to improve the explanation provided by the preceding model, the model can be expanded by including variables at the school level (model 4),<sup>15</sup> but a variable that quantifies or measures the “environment” has yet to be included.

As shown in table 4, the CLSIZE, RESCURR, SCMATEDU and “Pública”<sup>16</sup> variables are now included. All of them

TABLE 4

**Estimated coefficients: models 2 through 5**  
(Final estimates of fixed effects with standard errors that are robust to heteroskedasticity)

Variables	Coefficients: fixed portion			
	Model 2	Model 3	Model 4	Model 5
Constant	173.64 (0.041)	166.90 (0.047)	191.84 (0.024)	206.4 (0.015)
Gender	3.06 (0.448)	2.21 (0.578)	2.06 (0.601)	1.96 (0.619)
Age	4.91 (0.330)	5.39 (0.281)	5.65 (0.260)	5.32 (0.293)
SecLevel	67.65 (0.000)	67.64 (0.000)	65.02 (0.000)	63.56 (0.000)
HISEI	0.35 (0.003)	0.32 (0.009)	0.33 (0.006)	0.29 (0.016)
PARED	1.92 (0.000)	2.02 (0.000)	1.91 (0.000)	1.81 (0.000)
NATIVO	-3.58 (0.652)	-3.07 (0.701)	-2.42 (0.764)	-1.67 (0.837)
CULPOSS	2.64 (0.221)	2.56 (0.252)	2.42 (0.276)	2.09 (0.348)
HEDRES	7.64 (0.001)	7.74 (0.000)	7.73 (0.000)	7.86 (0.000)
WEALTH	-1.86 (0.375)	-1.50 (0.462)	-1.84 (0.364)	-2.35 (0.244)
CULTPOSS_O	51.56 (0.000)	51.16 (0.000)	50.24 (0.000)	50.02 (0.000)
CLSIZE			1.37 (0.014)	1.28 (0.009)
RESCURR			8.39 (0.120)	7.0 (0.192)
SCMATEDU			8.09 (0.019)	7.15 (0.030)
PUBLIC			-38.39 (0.001)	-15.07 (0.138)
CLSIZE_O			-18.87 (0.099)	-27.97 (0.055)
SCMATEDU_O			-17.49 (0.383)	-11.02 (0.602)
PROESCS				25.90 (0.001)

Source: Author’s original calculations based on the Programme for International Student Assessment (PISA) database, 2006.

Note: The *p* value is shown in parentheses.

except RESCURR prove to be significant at 5%. It can also be seen that the significance and effect of the level 1 variables have not been altered. Calero and Escardíbul (2007) explain that this result is to be expected in multilevel models.

The extra information provided by model 4 that is not present in model 3 is that individuals’ performance is better in schools that: (i) have a larger class size; (ii) have better educational resources; and (iii) are privately run.

RESPRES variable is not included either because it is highly correlated with RESCURR (a positive Pearson coefficient that is significant at 1%), nor is TCSHORT, since it is highly correlated with SCMATEDU (a negative Pearson coefficient that is significant at 1%).

<sup>13</sup> The programme reports the result of running an individual hypothesis test for each variable using the *t* statistic and based on a null hypothesis that the coefficient is equal to 0.

<sup>14</sup> The VHLM programme displays a table showing the result of a chi-squared test run to analyse the random effect of each variable. This test encompasses the coefficients for all the centres, taken together, for each variable. The null hypothesis is that the coefficients of all the schools are equal (i.e. that the “random part” of the coefficient for each centre is 0 (the distance between the coefficient for each centre and the mean coefficient is null). If, as occurs here, the three variables in question fall into the rejection region, then this hypothesis is not accepted and there are legitimate grounds for stating that there are differences in the slopes of the regressions for the schools.

<sup>15</sup> There are no problems of multicollinearity; the mean value for the variance inflation factor (VIF) is 1.21, and it does not exceed a value of 5 for any individual variable.

<sup>16</sup> The “proportion of female students” variable is not included because models 2 and 3 indicate that gender is not a significant factor. The

In fact, students attending public schools, *ceteris paribus*, score 38 points less than those who go to private schools. It would, however, be somewhat rash to attribute a great deal of importance to the effect of the private/public variable on academic performance without first analysing what is happening with this variable by including information on the “environment” in model 4, since –as noted previously– there is empirical evidence that the incorporation of the peer effect reduces or eliminates the significance of the form of school administration.

In order to determine whether or not effects linked to the school environment are at work, we will now incorporate a variable into model 4 that is intended to quantify the peer effect. This variable is the “mean ESCS”, and it itself proves to be significant (value  $P = 0.001$ ). This new model will be designated as model 5.<sup>17</sup>

Thus, there is evidence that points to the presence of a peer effect. In addition, the inclusion of this variable reduces the significance of the “public” variable, with the  $P$  value of that variable shifting from 0.001 (model 4) to 0.138, which means that it ceases to be significant. This result signals that what is important is not so much whether a school is public or private but rather the characteristics of the households that the students attending each type of school come from.

Finally, the proposed models can be compared in order to determine which is the most suitable one to use as a basis for drawing definitive conclusions. As may

be seen from table 5, model 5 is the most appropriate choice for the following reasons:

- (i) Overall, model 5 has the greatest explanatory power of the three (36% over the null model). In other words, it fails to explain a smaller percentage of residual variance relative to the null model.
- (ii) Given our research objective, school-level effects on performance are of particular interest. This is another reason why model 5 is the most appropriate, since it provides more information in that respect. On the one hand, the variables used at the school level (level 2) in this model explain a greater percentage of the inter-school variance in test scores (52%) than the other models do. On the other hand, if we look at the percentages of residual variance that are not explained in each case, we see that the smallest percentage for level 2 is that of model 5 (28%).

Consequently, model 5 has been used to analyse the role of what have been designated as the major variables for this study. Table 5 shows the estimated coefficients calculated using this model.

The results yielded by the chosen model (model 5) indicate that the fact that a school is public or private is not a significant determinant of scholastic performance. By contrast, the socioeconomic environment of each school is an important explanatory variable for academic achievement and represents the peer effect. Consequently, the findings support the hypothesis that the determining factor for high performance is not whether the school is public or private but rather the kinds of students who attend each type (public or private) of school.

<sup>17</sup> There are no problems of multicollinearity; the mean FIV value is 1.27, and it does not exceed a value of 5 for any individual variable.

TABLE 5

## Main results for the non-null models

	Model 2	Model 3	Model 4	Model 5
Residual variance (level 1)	3 621.43	3 500.41	3 520.14	3 517.26
Residual variance (level 2)	2 922.06	4 242.93	3 307.76	2 795.64
Total residual variance	6 543.49	7 743.34	6 827.90	6 312.90
Percentage of residual variance accounted for by the variables over the null model: reduction in total residual variance	34	22	31	36
Comparison with the null model				
Percentage of residual variance accounted for by variables over the null model: level 1 (students)	12	15	15	15
Percentage of residual variance accounted for by the variables over the null model: level 2 (schools)	50	27	43	52
Percentage of total residual variance not accounted for by the model	66	78	69	64
Structure of percentages of total residual variance not accounted for by the model				
Percentage of residual variance attributable to differences between students (intra-school)	36	35	35	35
Percentage of residual variance attributable to differences between schools	29	43	33	28

Source: Author's original calculations based on the Programme for International Student Assessment (PISA) database, 2006.

## VI

### Conclusions

Throughout this study we have been examining the determinants of secondary-school students' scholastic achievement in Argentina based on data from the 2006 PISA test. The aim of this exercise is to answer the following question: Is the fact that a school is public or private a significant factor in accounting for educational outcomes?

The evidence presented here supports the hypothesis put forward in the introduction, according to which the correlation between the type of school administration (public/private) and scholastic performance fades when the socioeconomic environment of the school is brought into the picture. This appears to indicate that private schools provide a better education than public schools do, but not because of their ownership structure. Instead, this is due to the fact that their students are better-placed, because of their background, to study and attain a high level of performance.

This raises an interesting question for further research: although the way in which private schools are organized does not appear to be a decisive factor in terms of performance, it is nonetheless true that such schools provide some sort of signal that causes individuals of a certain socioeconomic status to choose to attend them. In other words, the reasons why families choose this type of school (and, frequently, why these schools choose their students) may account for the favourable learning environment that is provided for students.

The considerations that we have explored here furnish grounds for stating that education policies need to go beyond the scope of the educational environment as such

because socioeconomic status appears to be an extremely influential determinant of scholastic performance. The results also show, however, that schools differ in their ability to offset inequalities of origin. This shows that there is indeed scope for education policy. In other words, since some schools perform better than others, there is scope for improvements that will, at the very least, bring the lower-performing schools up to the level of the better-performing ones.

As to the question of what kinds of education policies should be put in place, the information presented here would seem to indicate that one aspect to be taken into account is the quality of each school's educational resources. A more detailed study will be needed in order to pinpoint exactly which aspects should be given priority, however. Future research efforts should incorporate more level 2 (school-level) variables into the analysis in order to fine-tune the related policy considerations.

Finally, it should be noted that many governments may be seeking to achieve a basic minimum of educational services for their students. In order to assess the extent to which this goal is achieved and the conditions that lead to success or failure in school, the specific factors that determine the likelihood that a given student will fail need to be explored.<sup>18</sup> Further research will also need to be done in this area.

*(Original: Spanish)*

---

<sup>18</sup> This can be done thanks to the development of software that allows logistic regression models to be used in multilevel analyses.

## Bibliography

- Abdul-Hamid, H. (2007), "Evaluación de lo preparada que está Argentina para la economía del conocimiento: Medición del conocimiento y las habilidades de los alumnos en matemáticas y ciencias con resultados de PISA 2000", *Bienestar y política social*, vol. 3, No. 2, Mexico City, Inter-American Conference on Social Security/Ibero-American University.
- Calero, J., A. Choi and S. Waisgrais (2010), "¿Qué determina el fracaso escolar en España? Un estudio a través de PISA-2006" [online] <http://www.um.es/dp-hacienda/eep2010/comunicaciones/eep2010-6.pdf>.
- \_\_\_\_\_ (2009), "Determinantes del rendimiento educativo del alumnado de origen nacional e inmigrante en PISA-2006", *Cuadernos económicos*, No. 77, Madrid, Información Comercial Española, Ministry of Industry, Tourism and Trade.
- Calero, J. and J.O. Escardibul (2007), "Evaluación de servicios educativos: El rendimiento en los centros públicos y privados medido en PISA-2003", *Hacienda pública española*, vol. 183, No. 4, Madrid, Institute of Fiscal Studies.
- Cervini, R. (2010), "El 'efecto escuela' en la educación primaria y secundaria: El caso de Argentina", *Revista electrónica iberoamericana sobre calidad, eficiencia y cambio en educación (REICE)*, vol. 8, No. 1, Madrid, Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar.
- \_\_\_\_\_ (2009), "Comparando la inequidad en los logros escolares de la educación primaria y secundaria en Argentina: Un estudio multinivel", *Revista electrónica iberoamericana sobre calidad, eficiencia y cambio en la educación (REICE)*, vol. 7, No. 1, Madrid, Red Iberoamericana de Investigación sobre Cambio y Eficacia Escolar.
- \_\_\_\_\_ (2006), "Los efectos de la escuela y del aula sobre el logro en matemáticas y en lengua de la educación secundaria. Un modelo multinivel", *Perfiles educativos*, vol. 28, No. 112, Mexico City, National Autonomous University of Mexico.
- \_\_\_\_\_ (2005a), "Nivel y variación de la equidad en la educación media de Argentina", *Revista iberoamericana de educación*, Madrid, Organization of Ibero-American States for Education, Science and Culture.
- \_\_\_\_\_ (2005b), "Variación de la equidad en resultados cognitivos y no cognitivos de la educación media de Argentina", *Revista electrónica de investigación educativa*, vol. 7, No. 1 [online] <http://redie.uabc.mx/vol7no1/contenido-cervini3.html>.
- \_\_\_\_\_ (2002a), "Desigualdades socioculturales en el aprendizaje de matemática y lengua de la educación secundaria en Argentina: un modelo de tres niveles", *Revista electrónica de investigación y evaluación educativa (RELIEVE)*, vol. 8, No. 2, Valencia, University of Valencia.
- \_\_\_\_\_ (2002b), "Desigualdades en el logro académico y reproducción cultural en Argentina. Un modelo de tres niveles", *Revista mexicana de investigación educativa*, vol. 7, No. 16, Mexico City, Consejo Mexicano de Investigación Educativa.
- \_\_\_\_\_ (1999), *Calidad y equidad en la educación básica en la Argentina*, Buenos Aires, Ministerio de Cultura y Educación de la Nación.
- De La Cruz, F. (2008), "Modelos multinivel" [online] [http://rpe.epiredperu.net/rpe\\_ediciones/v12\\_n03\\_2008/AR1.pdf](http://rpe.epiredperu.net/rpe_ediciones/v12_n03_2008/AR1.pdf)
- Diez Roux, A. (2002), "A glossary for multilevel analysis", *Journal of Epidemiology and Community Health*, vol. 56, No. 8, BMJ Publishing Group.
- Fernández Aguerre, T. (2002), "Determinantes sociales e institucionales de la desigualdad educativa en sexto año de educación primaria de Argentina y Uruguay, 1999. Una aproximación mediante un modelo de regresión logística", *Revista mexicana de investigación educativa*, vol. 7, No. 16, Mexico City, Consejo Mexicano de Investigación Educativa.
- Formichella, M. (2010), "Educación y desarrollo: análisis desde la perspectiva de la equidad educativa interna y del mercado laboral", thesis, Buenos Aires, Universidad Nacional del Sur.
- Gertel, H. and others (2007a), "El rendimiento escolar de la población de estudiantes de la educación básica en Argentina: ¿cómo contribuye la gestión de la escuela?", *IX Jornadas Argentinas de Estudios de Población*, Cordoba, Argentine Association of Population Studies.
- \_\_\_\_\_ (2007b), "Incidencia de la gestión sobre el rendimiento escolar en la escuela argentina. El mensaje de las pruebas internacionales y nacionales" [online] <http://www.aep.org.ar>.
- \_\_\_\_\_ (2007c), "Los factores determinantes del rendimiento escolar al término de la educación básica en Argentina. Una aplicación de técnicas de análisis jerárquico de datos", *Anales de las XVI Jornadas de la Asociación de Economía de la Educación*.
- \_\_\_\_\_ (2006), "Análisis multinivel del rendimiento escolar al término de la educación básica en Argentina", document presented at the Annual meeting of the Argentine Association of Political Economy [online] [http://www.aep.org.ar/esp/anales/works06/Gertel\\_Giuliodori\\_Herrero.pdf](http://www.aep.org.ar/esp/anales/works06/Gertel_Giuliodori_Herrero.pdf)
- Hox, J.J (1995), *Applied Multilevel Analysis*, Amsterdam, TT-Publikaties.
- Martínez Arias, R. (2006), "La metodología de los estudios PISA", *Revista de educación*, Madrid, Complutense University of Madrid.
- Mediavilla, M. (2010), "Las becas y ayudas al estudio como elemento determinante de la continuidad escolar en el nivel secundario post-obligatorio. Un análisis de sensibilidad a partir de la aplicación del Propensity Score Matching", *Investigaciones de economía de la educación*, Asociación de Economía de la Educación.
- OECD (Organization for Economic Cooperation and Development) (2006a), *El programa PISA de la OCDE. ¿Qué es y para qué sirve?*, Paris.
- \_\_\_\_\_ (2006b), *PISA 2006: Science Competencies for Tomorrow's World*, Paris.
- \_\_\_\_\_ (2006c), *PISA Data Analysis Manual. SPSS, Second Edition*, Paris.
- \_\_\_\_\_ (2003), *PISA 2003. Data Analysis Manual*, Paris.
- Santos, M. (2007), "Quality of education in Argentina: determinants and distribution using PISA 2000 test scores", *Well-being and Social Policy*, vol. 3, No. 1, Mexico City, Inter-American Conference on Social Security/Ibero-American University.
- Viego, V. (2006), "Comentario al trabajo 'Análisis multinivel del rendimiento escolar al término de la educación básica en Argentina'" [online] [www.aep.org.ar](http://www.aep.org.ar).
- Wößmann, L. and T. Fuchs (2005), "Families, schools, and primary-school learning: evidence for Argentina and Colombia in an international perspective", *Policy Research Working Paper Series*, No. 3537, Washington, D.C., World Bank.