Revista EIA

UNIVERSIDAD EIA

What elements impact academic achievement in students Colombians? A multilevel approach

Revista EIA
ISSN 1794-1237
e-ISSN 2463-0950
Año XIX/ Volumen 21/ Edición N.42
Julio - diciembre de 2024
Reia4221 pp. 1-33

Publicación científica semestral
Universidad EIA, Envigado, Colombia

Galvis-Gonzalez Jaime Eduardo[1]
Candelo-Viáfara, Juan Manuel[1]
Rivera-Díaz, María del Pilar[1]

1. Universidad del Valle, Colombia

Para citar este artículo / To reference this article /
Galvis-Gonzalez, J. E.; Candelo-Viáfara, J. M.; Rivera-Díaz, M, P. What elements impact academic achievement in students Colombians? A multilevel approach
Revista EIA, 21(42), Reia4221. pp. 1-33.
https://doi.org/10.24050/reia.v21i42.1754

Autor de correspondencia:

Rivera-Díaz, María del Pilar
Maestría en prevención de riesgos
Universidad del Valle, Colombia
maría.delpilar.rivera@correounivalle.edu.co

Recibido: 04-12-2023
Aceptado: 27-05-2024
Disponible online: 01-07-2024

Abstract

objective: to analyze the factors that influence the academic performance of students in Colombia, considering both individual characteristics and factors related to the educational institution. **Methodology**: a hierarchical econometric model was used to nest the data at different levels, and the individual results of the Saber 11° tests of the Colombian Institute for the Evaluation of Education were used for both semesters of 2021. **Results**: indicate that the nature of the school (public or private) is a key factor in student performance, explaining almost 11% of the differences between schools, despite the fact that students attending public schools have lower socioeconomic, cultural and social conditions than those attending private schools, there is still a significant gap in academic performance that cannot be fully explained. **Conclusion**: private schools were found to have more variability in their effect on student performance, indicating that the unique characteristics of private schools, such as admission fees and incentives granted to teachers, may influence student academic performance more randomly.

Keywords: academic performance, econometric model, education evaluation, student performance, higher education, academic skills, school level

¿Qué elementos inciden en el rendimiento académico de los estudiantes colombianos? Un enfoque multinivel

Resumen

Objetivo: analizar los factores que influyen en el rendimiento académico de los estudiantes en Colombia, considerando tanto las características individuales como los factores relacionados con la institución educativa. **Metodología:** se utilizó un modelo econométrico jerárquico para anidar los datos en diferentes niveles, y se emplearon los resultados individuales de las pruebas Saber 11° del Instituto Colombiano para la Evaluación de la Educación para ambos semestres de 2021. **Resultados**: indican que la naturaleza del colegio (público o privado) es un factor clave en el desempeño de los estudiantes, explicando casi el 11% de las diferencias entre colegios, a pesar de que los estudiantes que asisten a escuelas públicas tienen condiciones socioeconómicas, culturales y sociales más bajas que los que asisten a escuelas privadas, aún existe una brecha significativa en el rendimiento académico que no puede ser explicada en su totalidad. **Conclusión:** se observó que las escuelas privadas tienen más variabilidad en su efecto sobre el rendimiento de los alumnos, lo que indica que las características únicas de las escuelas privadas, como los derechos de admisión y los incentivos concedidos a los profesores, pueden influir en el rendimiento académico de los alumnos de forma más aleatoria.

Palabras clave: rendimiento académico, modelo econométrico, evaluación educativa, rendimiento de los estudiantes, educación superior, competencias académicas, nivel escolar

1. Introduction

Education is an element that allows for increasing a country’s productivity, and in Colombia, it is a fundamental right. Likewise, it is one of the most effective mechanisms for reducing regional inequalities. Investing in education reduces poverty, promotes social mobility, and increases wages, among other factors that make it one of the pillars of both economic and social growth and development, given its implications at the individual and global levels (Sánchez et al., 2014).

Therefore, quality education directly influences students’ performance, which, in turn, affects their access to higher education and opportunities in the labor market. Higher education levels lead to increased productivity and, consequently, higher income and well-being for individuals (Prada, 2006). In Colombia, students must take the "Saber 11 tests" to assess their academic competencies and qualify for higher education, these tests, administered by ICFES, also serve to monitor the quality of secondary education, the results of these tests significantly impact students' access to higher education, as they are used in the admission criteria, which vary across institutions and programs (Icfes, 2020).

In this order of ideas, it is essential to identify which factors influence a student's ability to obtain results that allow them to meet the minimum access requirements demanded by the different higher education institutions. A vast literature suggests that various characteristics of students (personal, family, and socioeconomic), as well as characteristics specific to schools, influence students' performance in various academic performance evaluation tests. Some of these are: gender, socioeconomic stratum, education of parents, nature of the school, level of education of the teacher, among others (Gaviria and Barrientos, 2001, Correa 2004; López, 2012; Zambrano, 2013).

The article aims to identify the individual and school factors that affect students' performance. Specifically, the study seeks to evaluate the influence of socioeconomic factors that are considered determinants of gaps in academic performance according to the nature of students' schools, based on the Saber 11 tests. From the above, the research question to be answered is, "How do household socioeconomic factors influence students' academic performance in the Saber 11 tests, according to the nature of the institution, in the municipalities of interest?" This research question leads us to suppose that, with similar household socioeconomic conditions, it is possible to expect the same academic performance from the student, regardless of the school's nature to which he or she belongs.

This research is divided into an introduction, a reference framework, an empirical application, where the data and variables to be used are presented, as well as a descriptive analysis and a hierarchical model of two levels. It also presents the model's results that will allow us to see which student and institutional variables influence the results of the Saber 11 tests, the discussion, and finally, the conclusions.

2. Literature review

Leckie and Prior (2022) found that simple approaches are effective in measuring the impact of schools on student learning in England, while Rhiannon (2022) suggests that value-added models are preferable to multilevel models for measuring academic achievement in low-income countries such as India.

Lozano (2003) identified personal and school factors, such as gender, parental education and classroom interactions, as significant influences on school failure among secondary school students in Almeria, Spain, on the other hand, Lizasoain et al. (2007) found that students from low socioeconomic backgrounds who attend high level schools obtain the best academic results in the Basque Autonomous Community, closely followed by students from higher socioeconomic backgrounds who attend schools of the same level.

For its part, García, Rodríguez, and Torres (2020) studied contextual characteristics and their relationship with performance in mathematics and language using data from the 2017 ESCALA tests conducted in Andalusia, Spain. Through the segmentation tree data mining methodology, they identified a positive influence of social and cultural status and family expectations on test performance, this suggests a relationship between contextual characteristics and student performance, as well as the identification of students at risk of not reaching the required minimum levels.

In the Latin American context, Ortega, Malmberg, and Sammons (2018) investigated the effects of schools on the academic trajectory of primary students in language and math. Their study involved over 19,000 students in 156 schools, the results indicated that school centers significantly impact students' academic performance, but no positive effect was found regarding the composition of the school center.

In Mexico, Hoyos, Espino, and García (2012) conducted a study that considered various factors, including home characteristics, individual attributes, institutional factors, and school resources, to explain differences in math grades among students. The study found that individual characteristics, particularly academic background, had the greatest influence on cognitive achievement, followed by school resources, institutional environment, and family background.

Revista EIA https://doi.org/10.24050/reia.v21i42.1754

Using data from PISA 2012, Moreno and Cortez (2020) found that socioeconomic status, the presence and educational level of the mother and, to a lesser extent, the father's education, positively influenced student performance in public and private schools in Mexico, they also observed a reduction in the achievement gap between these institutions, similarly León and Collahua (2016) conducted a study in Peru and highlighted the significant role of socioeconomic status on student performance at both the individual and institutional level and Taber (2018), found a strong correlation between investment in infrastructure and educational quality.

In Peru, Jordán et al. (2019) found that directing public investment towards educational inputs improved outcomes, while investing in infrastructure had no positive effect. Arcidiácono et al. (2014) documented increasing school segregation based on socioeconomic backgrounds in Latin America, with low-income students having limited interaction with other socioeconomic groups.

In Colombia, Duarte and Bos (2012) explored educational disparities and found significant variation in academic achievement related to socioeconomic status. School-level disparities were more pronounced than within-school disparities, highlighting the challenges faced by economically disadvantaged, urban, and rural schools in terms of students' academic performance.

Sarmiento et al. (2000) found that, after accounting for socioeconomic status, public schools outperformed private schools in terms of academic achievement. Jiménez and Pinzón (1998), Gaviria and Barrientos (2001a, 2001b), Sánchez and Otero (2012), and Correa and Orejuela (2017) also emphasize the impact of socioeconomic factors on student performance, including parental education and school type, these studies highlight the multifaceted influences on academic achievement, Chica, Galvis and Ramírez (2011) in their study find similar results to those mentioned above, and also detail that having a higher socioeconomic level allows the student to have access to benefits and technological tools (computer, internet), tutors, good food, among others, which facilitate the academic process and lead to a better performance in the Saber 11° tests. For this reason, they recommend directing some policies to improve the unfavorable socioeconomic conditions of Colombians and thus contribute to better academic performance of students.

The studies consulted provide evidence of a relationship between academic achievement and individual-level variables (such as parental education, socioeconomic status, among others) and school-level variables (such as the nature of the school, the number of teachers, among others) that are key to school achievement and its effect on individual well-being (Mantilla and Cortés, 2016). For this reason, this work presents a novelty in the use of information and its results, by using a multilevel model and its nesting (interaction) of data.

3. Methodology

3.1 Multilevel Model

To determine the academic performance gap according to the nature of the school, associated with the socioeconomic conditions of students belonging to the municipalities of interest for Valle, based on the Saber 11° tests of the year 2021, the hierarchical econometric model will be estimated, which achieves the nesting of data in levels, where students belong to level 1 and schools to level 2. With this model, better accuracy of the effects is achieved due to the interaction of each of these levels, as it models the mean and variance simultaneously (Rasbash, 2008; Bernal-Ruiz et al., 2018).

Gaviria and Castro (2005) serve as a fundamental reference for the examination of multilevel models. Based on their work, this study largely adopts their notation and argumentative approach. The authors acknowledge that hierarchical models account for the nested structure and inherent complexity of social science data. These models offer a suitable means to address the variation arising from different levels of aggregation, thereby providing a statistical solution for simultaneously analyzing individual differences and contextual influences.

Furthermore, Hox (1995) highlights the necessity of employing multilevel models when working with grouped data, as observations within the same group tend to exhibit greater similarity compared to observations across different groups, (Boado, 2013), this violation of the assumption of independence

Revista EIA https://doi.org/10.24050/reia.v21i42.1754

among all observations can be effectively captured by expressing it as an intra-class correlation coefficient (ρ).

The Saber 11° tests applied by the State to evaluate the academic performance of high school students require nested structures, with students grouped within schools, this makes students belonging to one type of school have different characteristics than those belonging to another, the above makes the techniques traditionally used limited; therefore, the most suitable methodology for the study is that of multilevel models (hierarchical regression models) (Chel and Omar, 2015). This technique seeks to solve the estimation problems (distortion of the error term, standard error, and significance levels of the estimates) generated by the nested structure of the data.

To measure the academic performance gap of students, different factors of the home (socioeconomic) and schools (nature, character, etc.) that determine it must be taken into account, thus seeking to measure the proportion of the differences (variation) in the results attributed to the student and schools, (Albor, Dau, & Ruíz, 2014; Rodriguez et al., 2020).

Therefore, estimating a multilevel model requires a systematic analysis that starts with an empty model, which estimates the global mean score (dependent variable), the variance between schools and students, and does not include predictors. Then explanatory variables of students and schools are added to observe if adding a variable generates an effect on the total variance of the model, in addition to analyzing what proportion of the variation is explained by the difference between schools.

3.2 Definition of a two-level multilevel mode

Drawing from Bryk et al., (1996), work the suggested multilevel model comprises two distinct sub-models: one at level 1 and another at level 2, the focus of the study revolves around data that involves students nested within schools, at the level 1, the model examines the interrelationships among variables specific to individual students, while at the level 2, it analyzes the impact of school-level factors.

There are $i=1,\ldots,nj$ units at level 1 (students) nested within $j=1,\ldots,J$ units at level 2 (schools).

In the level 1 model, the dependent variable for case *i* (student) within unit *j* is represented as:

$$Y_{ij} = \beta_{0j} + \beta_{1j} X_{1ij} + \beta_{2j} X_{2ij} + \ldots + \beta_{Qj} X_{Qij} + e_{ij}$$
$$Y_{ij} = \beta_{0j} + \Sigma^{Q}{}_{q=1} \beta_{qj} X_{qij} + e_{ij}$$

Where:

The coefficients(q=0, 1,…,Q) correspond to the level 1 of the model;The variable X_{qij} repsetenst the predictor at level 1 for case *i* in unit *j;t*he term represent the random efecto at level 1; and σ^2 is the variance of *eij,* the varianceat level 1. The random efecto $\sim$ N(0, σ2) fowllows a normally distributed.In the level 2 model, each of the coefficients β_{qj} (q=0, 1,…,Q) defined in the level 1 model becomes a dependent variable in the level 2 model.

$$B_{qj} = \gamma_{q0} + \gamma_{q1} W_{1j} + \gamma_{q2} W_{2j} + \ldots + \gamma_{qS} W_{-Sqj} + \mu_{qj}$$
$$B_{qj} = \gamma_{q0} + \Sigma^{Sq}{}_{S=1} \gamma_{qs} W_{sj} + \mu_{qj}$$

Where: coefficient correspond to (q=0, 1,…, $_{Sq}$) Level 2 coefficient; Level 2 predictor; and level 2 random effect.It is assumed that, for each unit *j* of Level 2, the vector ($\mu_{0j}, \mu_{1j}, \ldots, \mu_{Qj}$) is distributed as a multivariate normal, and each element of μ_{qj} has mean zero and variance:Var() = For each pair of random effects *q* and *q'*, it holds that: Cov (, $_j$) = .The variance and covariance components are grouped into a dispersion matrix, *T*, whose dimensions are The level 1 coefficients can be modeled in level 2 in three different ways:

Fixed level 1 coefficient, = , level 1 coefficient with non-random variation across level 2 units.

$$B_{qj} = \gamma_{q0} + \Sigma^{Sq}{}_{S=1} \gamma_{qs} W_{sj}$$

Coefficient of level 1 with random variation in level 2 units

$$B_{qj} = \gamma_{q0} + \mu_{qj}$$

https://doi.org/10.24050/reia.v21i42.1754

Or with level 2 variables.

$$B_{qj} = \gamma_{q0} + \Sigma^{Sq}{}_{S=1}\, \gamma_{qs}\, W_{sj} + \mu_{qj}$$

The size of T depends on the number of level 1 coefficients designated as random. For model estimation, the Iterative Generalized Least Squares (IGLS) method was utilized, which is a series of the Generalized Least Squares (GLS) procedure. However, this approach yields estimators of the random parameters that are biased since it does not take into account the sample variance of the fixed part of the model (Millán and Hoyo, 2005), as stated by Correa (2004), this technique is suitable for examining variations in student performance within a school setting, as it allows for the decomposition of a variable (performance) into its intra-school and inter-school components. Moreover, it facilitates the analysis of the relationship between variables at different levels of aggregation, such as student or school characteristics.

3.3. Empty Model

This is the starting point for every multilevel model, in which there are no explanatory variables. If the variance of this model is not statistically different from zero (significant), there will be nothing to explain, that is, it would not make sense to include explanatory variables in the multilevel model in either of its two levels (Gaviria and Castro, 2005).

Level 1 would be represented by

$$Y_{ij} = \beta_{0j} + e_{IJ}$$

The level 2 by:

$$\beta_{0j} = \beta_0 + \mu_{0j}$$

The complete model by:

$$Y_{ij} = \beta_{0j} + (e_{ij} + \mu_{0j})$$

e_{ij} is how much the performance of student *i* in school *j* deviates from the school mean.

μ_{0j} is how much the mean of school *j* deviates from the overall mean.

$(e_{ij} + \mu_{0j})$ is the total variance.

$e_{ij} \sim N(0,\sigma^2_{e})\ y\mu_{0j} \sim N(0, \sigma^2_{\mu})$.

σ^2_{e} is the variance within students in each school.

$\sigma\mu^2$ is the variance between schools

Data

First, a descriptive analysis of the data is performed with some personal and socioeconomic characteristics of the student as well as the institution to which they belong, considered relevant in the global score obtained by the students in the Saber 11° tests for the year 2021, in the municipalities jointly studied.

The data used for the analysis were taken from the Colombian Institute for the Evaluation of Education and correspond to the individual results of the Saber 11° State tests for both semesters of 2021 these incorporate general information of students and schools from the municipalities together. The academic performance of each student is measured from the global score obtained in the Saber 11° tests, evaluated on a scale of 0 to 500, where the highest value is considered better performance.

The database contains a total of 3341, of which 3048 records were taken for the study, taking into account students from 96 schools that provide complete information for all variables of interest, where 2,496 students belong to public schools and 552 to their private peers. Table 1 presents the variables to be used.

Revista EIA https://doi.org/10.24050/reia.v21i42.1754

Table 1. Dictionary of Variables

Variable	Value	Variable Description
Sex	0	Male
	1	Female
Stratum	1	Very low stratum
	2	Low stratum
	3	Medium stratum
	4	Upper middle stratum
	5	High stratum
	6	Very high stratum
Household size	1	1 a 2
	2	3 a 4
	3	5 a 6
	4	7 a 8
	5	9 o más
Mother's education	1	None
	2	Incomplete primary school
	3	Primary school complete
	4	Secondary school incomplete
	5	High school complete
	6	Technical or technological incomplete
	7	Technical or technological complete
	8	Incomplete professional education
	9	Professional education complete
	10	Graduate degree
Internet	0	If you have Internet
	1	No Internet
Computer	0	If you have a computer
	1	No Computer
Nature College	0	Official
	1	Non-official
Character College	1	Academic
	2	Technical
	3	Technical/Academic
	4	Not applicable

Table 1. Dictionary of Variables

Variable	Value	Variable Description
School Area	**0**	Urban
	1	Rural
School Day	**1**	Morning
	2	Afternoon
	3	Evening
	4	Full
	5	Single
	6	Saturday

Source: Own elaboration based on ICFES Saber 11-2021 test results.

In the descriptive analysis, Table 2 shows some descriptive statistics, where the mean of the global score of the students included in the study is 252 more than half of the analyzed sample corresponds to female students; 92% of the learners belong to low socioeconomic strata[1] (1, 2, and 3) when observing household size, 67.56% are households with 1 to 4 people, in addition, 24% of the mothers of the students have completed higher education (complete technical or technological, complete professional, and postgraduate), being one of the factors that have been given greater relevance in the literature regarding a student's performance in high school.

The study included socioeconomic and technological variables that help the student in their learning, noting that the majority of students (64.76%) have a computer, and 84.22% have access to the internet at home, on the school characteristics side, it is observed that most learners belong to the public and urban sectors, and students belonging to the morning shift (57.15%) and technical/academic character (41.83%) are the most relevant.

1 The socioeconomic stratification is a classification into strata of residential properties that are entitled to receive public services. It is mainly done to charge differential rates for public utilities based on the strata, allowing for the allocation of subsidies and contributions in this area

https://doi.org/10.24050/reia.v21i42.1754

Table 2. Descriptive Indicators of the Student

Variables	Descripción de la variable	Porcentaje/Media
Puntaje Global		252
Sex	Male	54,40%
	Female	45,60%
Stratum	without stratum	1,71%
	Stratum 1	27,23%
	Stratum 2	50,92%
	Stratum 3	13,88%
	Stratum 4	4,40%
	Stratum 5	1,54%
	Stratum 6	0,33%
Household size	1 a 2	11,29%
	3 a 4	56,27%
	5 a 6	24,80%
	7 a 8	5,94%
	9 o más	1,71%
Mother's education	None	3,48%
	Incomplete primary school	10,43%
	Primary school complete	6,86%
	Secondary school incomplete	15,16%
	High school complete	33,79%
	Technical or technological incomplete	3,38%
	Technical or technological complete	12,34%
	Incomplete professional education	2,92%
	Professional education complete	9,97%
	Graduate degree	1,67%
Internet	If you have Internet	84,22%
	No Internet	15,78%
Computer	If you have a computer	64,76%
	No Computer	35,24%
Nature College	Official	81,89%
	Non-official	18,11%
Character College	Academic	1,51%
	Technical	25,26%
	Technical/Academic	31,40%
	Not applicable	41,83%

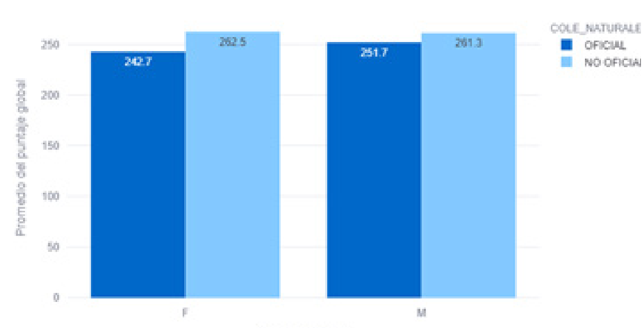
Table 2. Descriptive Indicators of the Student

Variables	Descripción de la variable	Porcentaje/Media
School Area	Urban	21,29%
	Rural	78,71%
School Day	Morning	20,51%
	Afternoon	4,99%
	Evening	57,15%
	Full	10,96%
	Single	3,74%
	Saturday	2,66%

Source: Own elaboration based on ICFES Saber 11-2021 test results.

Graph 1, shows that students from non-official institutions had a higher average score than their peers from official institutions. In addition, male students perform better in official schools, unlike non-official ones where females obtain higher scores, Gaviria and Barrientos (2001b) support this finding in their study, based on the hypothesis that men and women use different strategies to answer tests, for example, young men tend to look at the answers before reading the question, while young women tend to be more reflective, additionally, they mention that psychologists and pedagogues argue that male strategies yield better results in multiple-choice tests, such as the Saber 11 tests, while female strategies are more effective in university evaluations.

Graph 1. Average student performance by gender and type of school discriminated

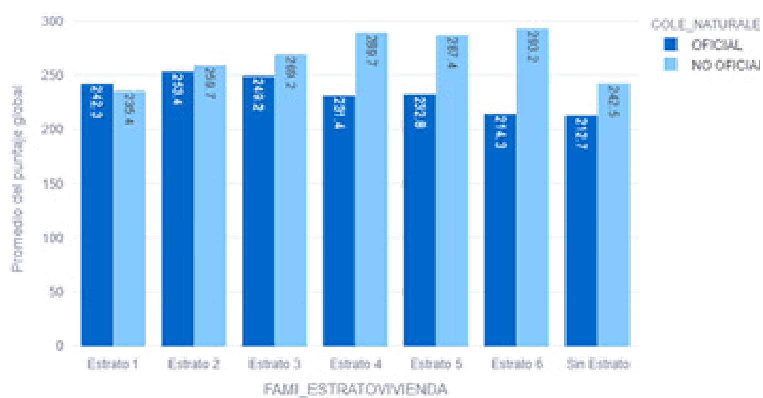


Source: Own elaboration based on ICFES Saber 11-2021 test results.

Revista EIA https://doi.org/10.24050/reia.v21i42.1754

On the other hand, in graph 2, it can be observed that students belonging to socioeconomic strata 4, 5, and 6 and attending non-official schools obtained a higher average performance, given that official schools are relatively more efficient for students from low socioeconomic strata, where these differences originate from the way public schools operate and the incentives, compared to their private counterparts (Núñez et al., 2002; Fernández et al., 2013), as it is noted that as the stratum increases from stratum 4 in official schools, the average tends to decrease.

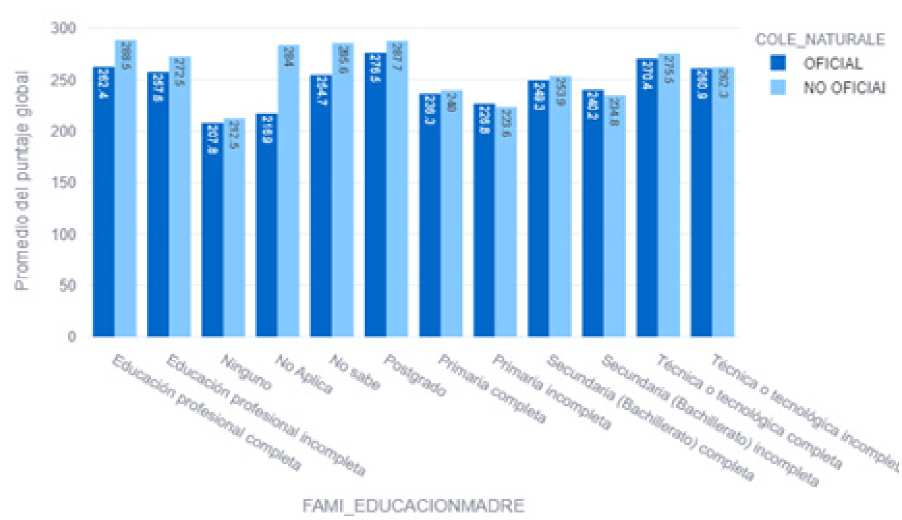
Graph 2. Average student performance by stratum and type of school



Source: Own elaboration based on ICFES Saber 11-2021 test results.

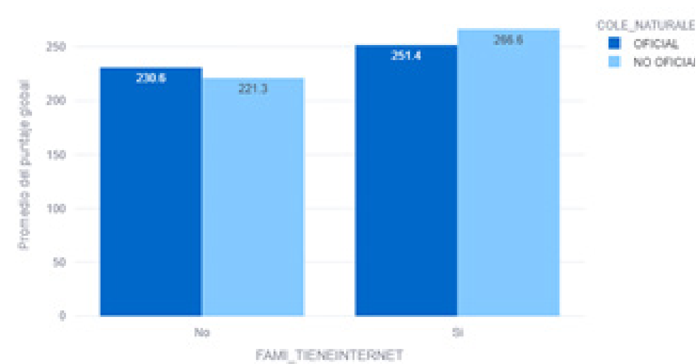
Graph 3 shows a direct relationship between mother's education and student performance, meaning that the average achievement score of the student increases as the mother's education level increases; the best results were obtained by students whose mothers have completed higher education. Other factors that qualitative analysis allows us to relate to good student performance are access to a computer and internet service at home, which is true for both types of institutions graphs 4 and 5, overall, it can be seen that students who belong to a non-official school and a high socio-economic level obtained better average results, as they may have easy access to technological tools.

Graph 3. Average student performance, discriminated by mother's education and nature of school.



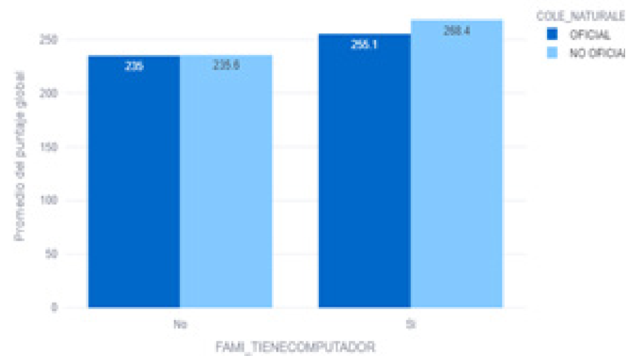
Source: Own elaboration based on ICFES Saber 11-2021 test results.

Graph 4. Average student performance discriminated by Internet access and type of school.



Source: Own elaboration based on ICFES Saber 11-2021 test results

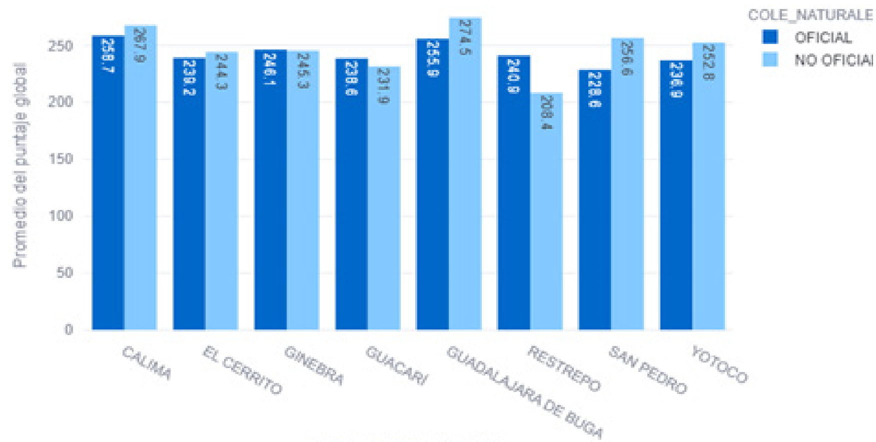
Graph 5. Average student performance by computer and type of school



Source: Own elaboration based on ICFES Saber 11-2021 test results

In graph 6, it can be observed that students who belong to a private school obtained better results, except for the municipalities of Guacari, Ginebra, and Restrepo, where the average performance of students was higher in official schools. Students from Guadalajara de Buga and Calima el Darien achieved a higher-than-average performance in both types of institutions, while the municipalities with the worst average scores in both types were El Cerrito and Restrepo. This may be due to a lack of administrative capacity in terms of quality and coverage of education (Saavedra and Forero, 2017; Macas, 2016), a problem that is reflected in the results of the Saber 11 tests.

Graph 6. Average student performance by municipality and type of school



Source: Own elaboration based on ICFES Saber 11-2021 test results

4. Results

4.1 Empty model and inter-school and intra-school variation

This model does not include any explanatory variables at either the student or school level, therefore β_{0j} represents the intercept or mean performance of students, i.e., the average score obtained by students on the Saber 11° tests. Additionally, In the table 3, the random effects show the existence of variance at the first level (intra-school variance), indicating that students differ from each other within schools, and the existence of variance between schools (inter-school variance), indicating that schools also differ in their average performance. This implies that student characteristics are more relevant when it comes to explaining inequalities in Saber 11° test performance (higher proportion of variance), although these factors are not able to fully explain the variation in student achievement, so it is necessary to observe the impact of institutional variables on overall scores.

Furthermore, it is worth noting that the variances for both levels are significant (when the ratio between their estimator and their standard error is greater than two ($p<0.05$), indicating unexplained variation between student performance and the average performance of schools.

Table 3. Results of the Null Model

Fixed Effect	
Intercept	***Estimate***
β_{0ij}	249.743 (3.139109)
Random effect	
Variance	***Estimate***
Level 1 variance	
σ^2_e = var (e_{ij})	1800.179 (46.75079)
Level 2 variance	
σ^2_μ= var (μ_{ij})	677.3221 (126.7044)

Source: Own elaboration based on ICFES Saber 11-2021 test results.

Standard error in parentheses.

As the variance is significant for both levels, the assumption of independence of all observations is not met; this can be explained by the intra-class correlation (ρ), which is equal to the proportion of total variance explained by one level, in this case level 2.

The intra-school correlation coefficient (ICC) is equal to:

$$\rho = \frac{\sigma_\mu^2}{\sigma_e^2+\sigma_\mu^2} = \frac{677.3221}{1800.179 + 677.3221} = 0.2733\%$$

This indicates that 27.33% of the total variance in student performance is due to variance between schools, while the remainder is explained by level 1 (intra-student), meaning that differentiation in student achievement on the Saber 11° test is explained in this proportion by differences between schools, therefore, institutional factors are important in the study of student performance on the 2021 Saber 11° tests in the municipalities of interest that is, attending a specific school implies a higher/lower probability of achieving a high/low educational performance (overall score on the Saber 11° test) than expected in another school.

It should be noted that Correa (2004) mentions in his article that, for OECD countries, the Intra-School Correlation Coefficient (ρ) does not represent more than 10% or 15% of the total variance of student performance, however, in developing countries (such as Colombia), these differences are greater, where the Intra-School Correlation Coefficient (CCI) is between 30% and 40%.

4.2 Model with separate student-level variables

In Table 4, the impact of individual student variables on average academic performance variance is shown. Student characteristics have a significant effect on performance, but little impact on inter-school variance. The variable "computer ownership" has the most influence on inter-school variance, reducing it from 677.3 to 586.9. However, adding additional variables does not significantly reduce intra-school variance. "Computer ownership" explains 3% of the unexplained variance in inter-school performance.

Table 4. Results of the student variables separately

Variables and levels	Model 0 estimation	Model 1 estimation	Model 2 estimation	Model 3 estimation	Model 4 estimation	Model 5 estimation	Model 6 estimation
Gender		-10.87***					
Stratum			9.89**				
Mother's education				3.80***			
Persons per household					-1.66*		
Internet						-12.94***	
Computer							-11.00***
Levels							
School	677.3221	692.1904	714.9902	500.7361	672.7125	607.2603	586.9642
Student	1800.179	1770.52	1792.886	1760.851	1798.669	1785.094	1782.024
CCI=ρ	**0.2733**	**0.2810**	**0.2850**	**0.2214**	**0.2722**	**0.2538**	**0.2477**

Source: Own elaboration based on ICFES Saber 11-2021 test results

Probabilities p: * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Revista EIA | https://doi.org/10.24050/reia.v21i42.1754

4.3 Model with separate school-level variables

In Table 5, the impact of school variables on average academic performance variance is explained. These variables have a significant effect when estimated separately, except for school character. However, they have little impact on inter-school variance in the Saber 11th grade tests. The school area and nature have the most influence on inter-school variance, reducing it from 677.3 to 536.2 and 560.5 respectively. The school schedule is most influential in student variation, decreasing it from 1800.1 to 1788.9. The percentage of school variation does not change significantly with additional variables, with the area and nature of the school being the only ones with higher variability, they explain 4% of the unexplained variance in average performance between schools.

Table 5. Results of the school variables separately

Variables and levels		Model 0 estimation	Model 1 estimation	Model 2 estimation	Model 3 estimation	Model 4 estimation
School area			-23.10***			
School day	Single			***		
	Full			-19.21**		
	Morning			-31.25***		
	Afternoon			-2.41		
	Evening			-7.71		
	Saturday			-41.02***		
School Character	Not applicable				1.74	
	Academic				-.026	
	Technical				-1.05	
	Technical/ Academic				23.2	
Nature College						22.19***
Levels						
School		677.3221	536.2788	686.6314	675.3772	560.5735
Student		1800.179	1798.195	1788.992	1800.146	1800.446
CCI=ρ		**0.2733**	**0.2297**	**0.2773**	**0.4102**	**0.2374**

Source: Own elaboration based on ICFES Saber 11-2021 test results.

Probabilities p: * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

4.4 Total model with joint student-level and school-level variables

In Table 6, it is observed whether the variables associated with the student or school as a whole have any effect on the achievement of the Saber 11th grade test. Given the above, a model is calculated with all the student variables, one with all the school variables, a total model with all the variables without a random part, and finally a total model, which includes all the variables in the empty model and a random part formed by the variable of interest, the nature of the school, in order to observe the impact of this variable on the students' results in the Saber 11th grade test.

Table 6. Results of variables in combination at the student and school level

Levels	Empty Model	Students Model	Schools Model	Total model without nature in the random part	Total Model with nature in the random part
School	677.3221	497.741	487.5014	417.3868	167.3273
Student	1800.179	1706.975	1790.139	1699.864	1698.768
CCI=ρ	**0.2733**	**0.2257**	**0.2140**	**0.1971**	**0.0896**

Source: Own elaboration based on ICFES Saber 11-2021 test results

By combining student variables, the school effect is reduced, providing a better explanation of average student performance in different schools, the unexplained inter-school variance significantly decreases from 677.3 to 497.7, and the inter-student variance decreases from 1800.1 to 1706.9, indicating a relationship between level 1 variables and Saber 11° test performance. The unexplained intra-school variance decreases from 27.33% to 22.57%, explaining at least 4.7% of the differences in average academic performance between schools.

When adding school variables together, the unexplained intra-school variance decreases from 27.33% to 19.71%, explaining at least 7.6% of the differences in average achievement between schools.

Revista EIA https://doi.org/10.24050/reia.v21i42.1754

In the total model, with both student and school variables, the unexplained intra-school variance decreases from 27.33% to 8.9%, explaining at least 18.4% of the differences in average performance between schools, the school nature variable shows strong explanatory power, reducing intra-school variance from 19.7% to 8.9% and explaining 11 percentage points of the differences between schools. This model provides the best explanation of Saber 11° test performance.

Table 7 shows the total model with the random part, where the student socioeconomic variables are statistically significant, unlike the number of people per household. It is observed that the mother's educational level and socioeconomic status are the variables that have the greatest relevance in the student's academic performance on the global score of the Saber 11° tests for 2021 in the municipalities together. Moreover, the characteristics of the school are statistically significant, unlike the school's character, where there are no differences between the performances of students who belong to an academic, technical, and/or technical/academic school; the variable nature of the school has the greatest relevance in the academic achievement of students in the Saber 11° tests, where belonging to an unofficial institution results in better scores, meaning that the individual achieves approximately 11 additional points in the overall average score, showing that the gap reaches points of difference in academic performance in the Saber 11° tests, in favor of private institutions.

Table 7. Total model results with randomized part

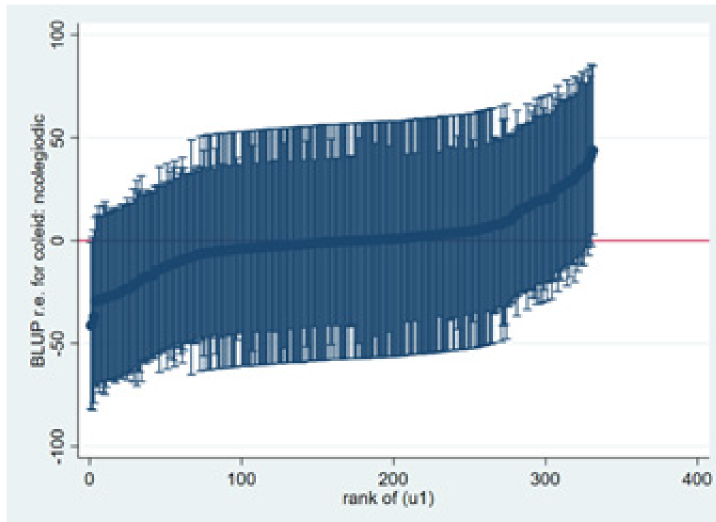
Gender	-9,95***
Stratum	13,86***
Mother's education	3,09***
Persons per household	-1,14
Internet	-8,15***
Computer	-5,96***
School area	-14,66***
School day	-1,96**
School character	-0,47
Nature of school	1,077*
Constate	239,04
Number of observations	
Number of groups	
Wall chi2(10)	217,4

Source: Own elaboration based on ICFES Saber 11-2021 test results.

Probabilities p: * Significant at 10%; ** Significant at 5%; *** Significant at 1%.

Finally, the box plot is presented to observe the variability of the residuals for the school nature and their respective confidence intervals. It can be seen that some schools are situated below the cutoff line (0,0), others are above the line, and others are on the line, indicating marked differences between groups. The distinctive characteristics of students belonging to different schools stand out. Additionally, it is shown that the mean of the residuals is approximately zero.

Graph 7. Variability of residuals for nature of school



Source: Own elaboration based on ICFES Saber 11-2021 test results.

The model validation is performed, and the residuals and standardized residuals are estimated (which yielded better results). First, regarding normality, it can be observed that these residuals have a relatively normal behavior according to the Q-Q[2] plots, with a good trend of the standardized residuals towards the line that marks normality. Also, with the density[3] function, an approximation is observed, although there is a deviation in the peak of the function. Finally, the histogram[4] illustrates that the data distribution is approximately normal without problems of kurtosis.

Finally, another graphical test is used to define the homoscedasticity[5] of the model or whether it presents problems with the variance of the residuals. It is pertinent to suspect problems of heteroscedasticity in the model, as a modest portion of the residuals surpass the confidence bands, although this proportion is not so representative. It should be noted that one of the solutions to this

2 See appendix 1

3 See appendix 2

4 See appendix 3

5. See appendix 4

problem is these multilevel models since the hierarchy is essential, where it is possible to specify or analyze a population taking into account different contexts. For this, it is also necessary to include other types of variables that explain the performance of students.

5. Discussion

Numerous studies have explored the determinants of academic achievement, recognizing that it is influenced by personal, family, institutional and socioeconomic factors. This study aims to unveil these determinants for the students of the municipalities attached to the Buga Chamber of Commerce, analyzing the results of the Saber tests for grade 11 of 2021, which reveals that 27.3% of the total variance in student performance can be attributed to differences between schools as pointed out by Álvarez-Sotomayor and Martínez-Cousinou, (2020).

When examined individually, the student and school variables have limited explanatory power, with the nature of the school and students' access to computers exerting the greatest influence. However, the combination of the student and school variables provides a better explanation of the unexplained variance. Overall, the inclusion of all variables decreases the intra-school correlation coefficient (ICC) by 7.6%. In addition, the incorporation of the nature of school variable in the random part decreases the ICC by approximately 11 percentage points, previous studies by López (2012) and Govorova et al., (2020), support the notion that non-official schools have distinct characteristics that can positively influence academic outcomes.

The study reveals a difference in performance between students in different types of schools, with unofficial schools obtaining better mean scores. The decrease in intra-school variance from 27.3% to 8.9% confirms this disparity, as highlighted by Boado (2013), the literature review supports the notion that students in public schools often face inferior economic, cultural and social conditions, resulting in lower test preparation compared to their counterparts in private schools as mentioned by Villar-Aldonza, (2023).

Revista EIA | https://doi.org/10.24050/reia.v21i42.1754

The hypothesis that students from centers of different nature obtain similar academic performance in similar socioeconomic conditions is not corroborated. The variable nature of the center is the most relevant variable for academic performance, while the variable character of the center is not significant in accordance with the findings reported by Mantilla and Cortés (2016).

Findings from studies by López (2012), Cox and Jiménez (1991), and Núñez et al. (2002) support the idea that non-official schools have distinct characteristics that can positively influence academic outcomes. These studies highlight factors such as the right of admission reserved by private schools, which affects student performance even after controlling for various individual and school characteristics.

Among the socioeconomic variables of the students, the educational level of the mother and the socioeconomic stratum are the most relevant for the overall results. The study recognizes the limitations of the available information and suggests further research with additional variables, such as municipality or neighborhood, to inform policies for improving education, reducing the gap between private and public schools, alleviating poverty, promoting economic development and improving well-being, as Moreno and Cortez (2020) state.

6. Conclusions

this research sheds light on the factors influencing student performance on the Saber 11° state tests, the study confirms that individual and school-level variables have a significant impact on academic achievement. Non-official institutions show better performances, but it is important to consider the broader context and conditions of students and schools.

The nature of the school variable emerges as a powerful determinant, explaining a considerable portion of the variation between schools by incorporating this variable into a random structure, the effect of school type on student performance becomes more random, suggesting the presence of unexplained gaps between different types of schools.

The observed randomness in private schools can be attributed to their distinct characteristics and policies, such as selective admissions and curriculum management, these factors contribute to the variability in student outcomes. It is crucial to note that this study is limited by available information, timing, municipalities, and variables. Further research is recommended, including the inclusion of additional variables and a more detailed analysis of students and schools. A broader analysis, potentially incorporating a third level such as municipality or neighborhood, can provide a more comprehensive understanding of the factors influencing education in the region.

The empirical evidence generated by this research can guide the development of policies aimed at improving education, reducing the gap between private and public schools, promoting economic development, and enhancing overall well-being.

CONFLICTS OF INTEREST

The authors declare that they have no financial, professional or personal conflicts of interest that could inappropriately influence the results obtained or the interpretations proposed.

CONTRIBUTION OF AUTHORS

For the development of this project all authors have made a significant contribution specified below:

Galvis, Gonzalez Jaime Eduardo: conceptualization, research development, methodology and data analysis.

Candelo, Viáfara, Juan Manuel: conceptualization, methodology and analysis of results.

Rivera Díaz, María del Pilar: conceptualization, design, data analysis and final revision of the manuscript.

All authors participated in the drafting of the manuscript.

7. REFERENCES

Albor, G. R., Dau, M. A., & Ruíz, J. L. R. (2014). Calidad institucional y rendimiento académico: El caso de las universidades del Caribe colombiano. Perfiles educativos, 36(143), 10-29. https://www.scielo.org.mx/pdf/peredu/v36n143/v36n143a2.pdf

Álvarez-Sotomayor, A., & Martínez-Cousinou, G. (2020). Inmigración, lengua y rendimiento académico en España. Una revisión sistemática de la literatura. Revista internacional de sociología, 78(3), e160-e160. https://doi.org/10.3989/ris.2020.78.3.19.083

Arcidiácono, Malena; Cruces, Guillermo Antonio; Gasparini, Leonardo; Jaume, David; Serio, Montserrat (2014). La segregación escolar público-privada en América Latina, Santiago de Chile: Serie Políticas Sociales N° 195, Naciones Unidas – CEPAL. https://hdl.handle.net/11362/36757

Bernal-Ruiz, F., Rodríguez-Vera, M., González-Campos, J., & Torres-Álvarez, A. (2018). Competencias parentales que favorecen el desarrollo de funciones ejecutivas en escolares. Revista Latinoamericana de Ciencias Sociales, Niñez y Juventud, 16(1), 163-176. https://doi.org/10.11600/1692715x.16109

Bryk, A., Raudenbush, S., & Congdon, R. (1996). HLM: Hierarchical linear and nonlinear modeling with the HLM/2L and HLM/3L programs. Scientific Software International

Boado, H. C. (2013). Introducción al análisis multinivel (Vol. 49). CIS-Centro de Investigaciones Sociológicas.

Chel, G., & Omar, R. (2015). Evaluación de modelos empíricos, matemáticos y redes neuronales para estimar datos faltantes en estaciones meteorológicas en México (Master's thesis). http://hdl.handle.net/10521/2621

Chica, S., Galvis, D., & Ramírez, A. (2011). Determinantes Del Rendimiento Académico En Colombia: Pruebas ICFES Saber 11, 2009 (Academic Performance Determinants in Colombia: ICFES Saber 11, 2009 Exam). Center for Research in Economics and Finance (CIEF), Working Papers, (11-5).ES https://dx.doi.org/10.2139/ssrn.2419792

Correa, J, and Orejuela, C. (2017). Incidencia de los factores del contexto en el desempeño en matemáticas de los estudiantes de básica secundaria en Colombia. Semestre Económico, 20(44), 159-188. https://doi.org/10.22395//seec.v20n44a8

Correa, J. (2004). Determinantes del rendimiento educativo de los estudiantes de secundaria en Cali: un análisis multinivel. Sociedad y economía, (6),81-105. ISSN: 1657-6357. https://www.redalyc.org/pdf/996/99617648003.pdf

Duarte, J., Bos, M. (2012). Desigualdad en los aprendizajes de los alumnos colombianos. BID, (17), 1-8. http://mapeal.cippec.org/wp-content/uploads/2014/05/DUARTE-y-BOS-Desigualdad-en-los-azjes-de-alumnos-colombianos.pdf

Escardíbul, J. (2008). Los determinantes del rendimiento educativo en España. Un análisis a partir de la evaluación de PISA-2006. Investigaciones de Economía de la Educación, 3, 153-162. DOI: 10.4438/1988-592X-RE-2011-362-161

Fernández, E., Bernardo, A., Suárez, N., Cerezo, R., Núñez, J. C., & Rosario, P. (2013). Predicción del uso de estrategias de autorregulación en la educación superior: Un análisis a nivel individual y de contexto. Anales de Psicología/Annals of Psychology, 29(3), 865-875. https://doi.org/10.6018/analesps.29.3.139341

Fertig, M., & Schmidt, C. (2002). The role of background factors for reading literacy: Straight national scores in the PISA 2000 study. https://dx.doi.org/10.2139/ssrn.323599

García-Jiménez, Rodríguez-Santero & Torres-Gordillo, (2020). "Influence of Contextual Variables on Educational Performance: A Study Using Hierarchical Segmentation Trees," Sustainability, MDPI, vol. 12(23), pages 1-10, November https://doi.org/10.3390/su12239933

Gaviria, A.; & Barrientos, J. (2001a). Determinantes de la calidad de la educación en Colombia, Archivos de Economía, núm. 159, Departamento Nacional de Planeación. http://hdl.handle.net/11445/1249

Gaviria, A.; & Barrientos, J. (2001b). Calidad de la educación y rendimiento académico en Bogotá, Coyuntura Social, vol. 24, pp. 111-127. https://hdl.handle.net/10495/7183

Gaviria, J y Castro, M, (2005) Modelos Jerárquicos lineales. Serie Cuadernos de Estadística, No. 29. Madrid: Editorial La Muralla S.A.

Govorova, E., Benítez, I., & Muñiz, J. (2020). Predicting student well-being: Network analysis based on PISA 2018. International Journal of Environmental Research and Public Health, 17(11), 4014. https://doi.org/10.3390/ijerph17114014

Hoyos, R., Espino, J., & García, V. (2012). Determinantes del logro escolar en México. Primeros resultados utilizando la prueba ENLACE media superior. El trimestre económico, 79(316), 783-811.

Instituto Colombiano para la Evaluación de la Educación Superior (ICFES). (2020). Guía de orientación Saber 11° 2020-1. [En línea] Available at: https://www.icfes.gov.co/documents/20143/1628228/Guia+de+orientacion+saber+11+2020-1.pdf/ec534dff-b171-d51b-5ee8-c05139100635

Jiménez, P., & Pinzón, A. (1998). Los insumos escolares en la educación secundaria y su efecto sobre el rendimiento académico de los estudiantes: un estudio en Colombia. Washington, DC: The World Bank/Latin America and the Caribbean Region/Department of Human Development.

Jordán, O. R. C., González-Martí, I., & Gil-Madrona, P. (2019). La dificultad de la implementación de una enseñanza por competencias en España. Education Policy Analysis Archives, 27, 121-121. https://doi.org/10.14507/epaa.27.4053

León, J., & Collahua, Y. (2016). El efecto del nivel socioeconómico en el rendimiento de los estudiantes peruanos: un balance de los últimos 15 años. MISC.

Leckie, G. y Prior, L. (2022). Una comparación de modelos de valor agregado para la responsabilidad escolar. Efectividad Escolar y Mejoramiento Escolar, 33 (3), 431-455. https://doi.org/10.6018/rie.417881

Lizasoain, L., Joaristi, L., Lukas, J. F., & Santiago, K. (2007). Efectos contextuales del nivel socioeconómico sobre el rendimiento académico en la educación secundaria obligatoria en la Comunidad Autónoma. https://www.redalyc.org/pdf/2750/275020546008.pdf

López, S. F. (2012). ESTIMACIÓN DEL EFECTO COLEGIO EN COLOMBIA: 1980–2009. Estudios Gerenciales, 28(122), 49-68. https://doi.org/10.1016/S0123-5923(12)70193-6

Lozano, A. (2003). Factores personales, familiares y académicos que afectan al fracaso escolar en la educación secundaria. Electronic journal of research in educational psychology. 1(1) PP. 43 - 66. https://www.redalyc.org/pdf/2931/293152876002.pdf

Macas Chulde, C. M. (2016). La planificación estratégica como factor que impulsa el desarrollo de la educación. Estudio de caso del circuito 3-4 Colegio Cumbayá, periodo 2014-2015 (Master's thesis, IAEN).

Mantilla Sánchez, I. L., & Cortés Velásquez, H. H. (2016). Relación entre la autoestima de los docentes, el logro académico y la convivencia escolar de los estudiantes de grado tercero en los Colegios Colsubsidio San Vicente y Torquigua: un estudio exploratorio de tipo descriptivo. http://hdl.handle.net/11396/418

Millán, D. S. S. I., and del Hoyo Bernat, D. J. (2005). Análisis de sistemas cointegrados mediante métodos de subespacios (Doctoral dissertation, Universidad de Valladolid).

Moreno Treviño, Jorge O., & Cortez Soto, Sara N.. (2020). Rendimiento académico y habilidades de estudiantes en escuelas públicas y privadas: evidencia de los determinantes de las brechas en aprendizaje para México. Revista de economía, 37(95), 73-106. Epub 22 de enero de 2021.https://doi.org/10.33937/reveco.2020.148

Núñez, J., Steiner, R., Cadena, X., & Pardo, R. (2002). ¿Cuáles colegios ofrecen mejor educación en Colombia? Archivos de Economía, 193, 1-56.

OCDE, Oecd Publishing, Organisation for Economic Cooperation, Development, Paris (France). (2001). PISA Knowledge and Skills for Life: First Results from PISA 2000. París: OCDE, 2001

Ortega, L., Malmberg, L. E., & Sammons, P. (2018). School effects on Chilean children's achievement growth in language and mathematics: An accelerated growth curve model. School Effectiveness and School Improvement, 29(2), 308-337. http://dx.doi.org/10.1080/09243453.2018.1443945

Prada, C. F. (2006). ¿Es rentable la decisión de estudiar en Colombia? Revista Ensayos Sobre Política Económica; Vol. 24. No. 51, edición especial Educación. Junio, 2006. Pág.: 226-323. https://doi.org/10.32468/Espe.510

Rasbash, J. (2008). Module 4: Multilevel structures and classifications. LEMMA (Learning Environment for Multilevel Methodology and Applications), Centre for Multilevel Modelling, University of Bristol.

Rhiannon Moore (2022) Variation, context, and inequality: comparing models of school effectiveness in two states in india, School Effectiveness and School Improvement, 33:4, 588-609, https://doi.org/10.1080/09243453.2022.2089169

Rodriguez-Hernandez, C. F., Cascallar, E., & Kyndt, E. (2020). Socio-economic status and academic performance in higher education: A systematic review. Educational Research Review, 29, 100305. http://dx.doi.org/10.1016/j.edurev.2019.100305

Saavedra, V., & Forero, D. (2017). Diagnóstico y plan de acción educativo para los municipios certificados del Valle del Cauca: Buenaventura, Buga, Cali, Cartago, Jamundí, Palmira, Tuluá y Yumbo. Bogotá: FDI Pacífico y Fedesarrollo, 389 p. http://hdl.handle.net/11445/3699

Sánchez, A. M., and Otero, A. (2012). RE No. 154 marzo de 2012--Educación y reproducción de la desigualdad en Colombia. Reportes del emisor.

Sánchez, A., Otero, A., Bonilla, L., Galvis, L. A., Cepeda, L., & Meisel, A. (2014). Educación y desarrollo regional en Colombia. Libros Banco de la República.

Sarmiento, A., Becerra, L., y González, J. (2000). La incidencia del plantel en el logro educativo del alumno y su relación con el nivel socioeconómico (No. 012965). Fedesarrollo. p. 53-62. http://hdl.handle.net/11445/1767

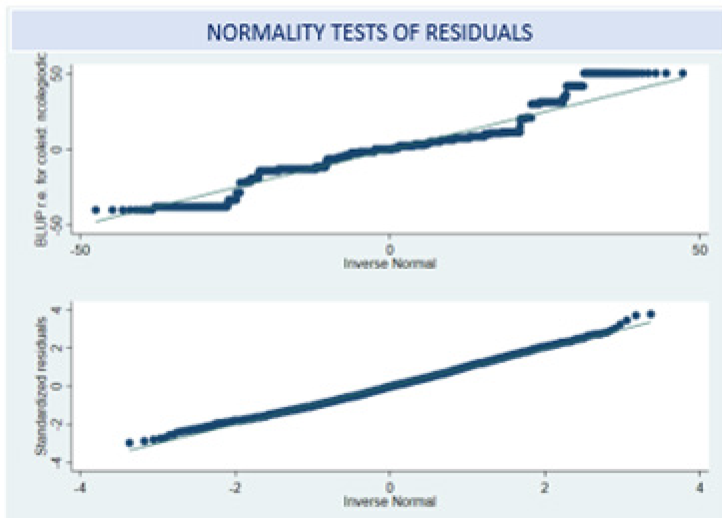
Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. Research in science education, 48, 1273-1296. https://doi.org/10.1007/s11165-016-9602

Villar-Aldonza, A. (2023). To what extent a modern teaching style benefits students? Why do teachers act the way they do?. Journal of Computer Assisted Learning, 39(2), 578-590. https://doi.org/10.1111/jcal.12765

Zambrano, J. (2013). Análisis multinivel del rendimiento escolar en matemáticas para cuarto grado de Educación Básica Primaria en Colombia. Sociedad y economía, (25)

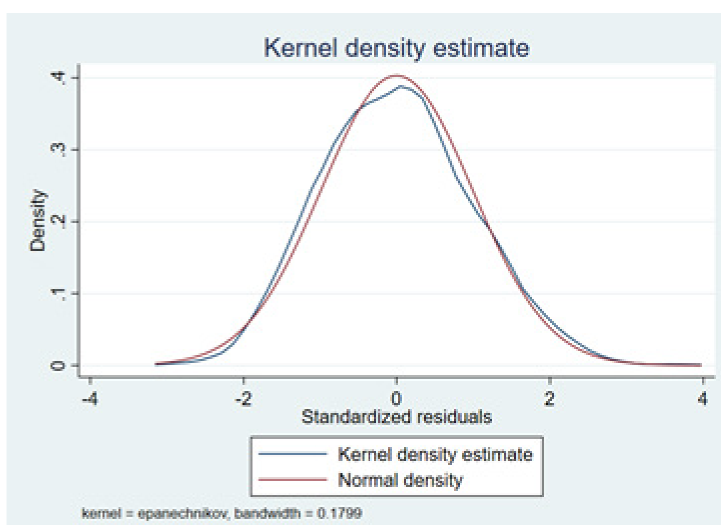
Anexos

Anexo 1



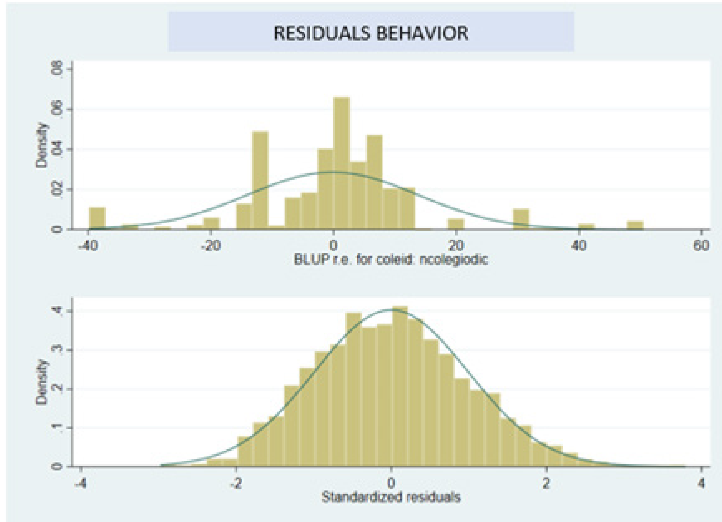
SOURCE: Own elaboration based on ICFES Saber 11-2021 test results.

Anexo 2



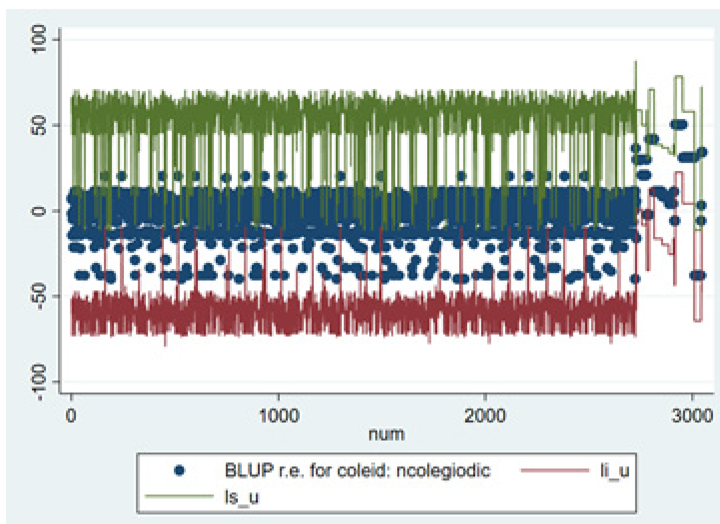
SOURCE: Own elaboration based on ICFES Saber 11-2021 test results..

Anexo 3



SOURCE: Own elaboration based on ICFES Saber 11-2021 test results.

Anexo 4



SOURCE: Own elaboration based on ICFES Saber 11-2021 test results.