

GEOREFERENCING EDUCATIONAL QUALITY IN CALI, COLOMBIA IN 2018, WITH A FOCUS ON ACADEMIC AGGLOMERATION AND SEGREGATION

David Arango-Londoño
Valentina Farkas Sanchez
Maribel Castillo
Geovanny Castro-Aristizabal

D. Arango-Londoño

Pontificia Universidad Javeriana, Department of Natural Sciences and Mathematic, Cali, Colombia.
E-mail: david.arango@javerianacali.edu.co

V. Farkas Sanchez

Economics student. Universidad del Valle, Cali, Colombia. E-mail: valentina.farkas@correouni-valle.edu.co

M. Castillo

Ph. D. in Social Sciences Children and Youth, CINDE and University of Manizales. Professor at the Department of Economics and Researcher at the Research Group in Economics, Management and Health (ECGESA), Pontificia Universidad Javeriana Cali, Colombia. E-mail: mabelcas@javerianacali.edu.co

G. Castro-Aristizabal

Ph. D. in Economics University of Zaragoza, Spain. Professor at the Universidad Autónoma de Bucaramanga, Colombia. Researcher in the Sectoral Dynamics Research Group at the Universidad Autónoma de Bucaramanga. E-mail: gcastro793@unab.edu.co

Sugerencia de citación: Arango-Londoño, D., Farkas Sanchez, V., Castillo, M., & Castro-Aristizabal, G. (2021). Georeferencing educational quality in Cali, Colombia in 2018, with a focus on academic agglomeration and segregation. *Cuadernos de Economía*, 41(86), 109-128. <https://doi.org/10.15446/cuad.econ.v41n86.85723>

Este artículo fue recibido el 15 de marzo de 2020, ajustado el 12 de julio de 2020 y su publicación aprobada el 14 de septiembre de 2020.

Arango-Londoño, D., Farkas Sanchez, V., Castillo, M., & Castro-Aristizabal, G. (2022). Georeferencing educational quality in Cali, Colombia in 2018, with a focus on academic agglomeration and segregation. *Cuadernos de Economía*, 41(86), 109-128.

Our findings confirm the existence of educational quality agglomeration in Cali (Moran's Index = 0.311). High educational quality is concentrated in schools in the south of the city (commune 22), while low educational quality is found in schools located in the east (communes 13, 14 and 15). Our study also confirms the importance of the determinants for school performance identified in the previous literature. The main contribution of this study is the finding that crime negatively affects educational quality. When the number of homicides in the neighbourhood in which a school is located increases by one percentage point, the quality of education decreases, on average, by 0.63 points.

Keywords: Analysis of education, education and inequality, panel data models, spatio temporal models.

JEL: I21, I24.

Arango-Londoño, D., Farkas Sanchez, V., Castillo, M., & Castro-Aristizabal, G. (2022). Georreferenciación de la calidad educativa en Cali, Colombia, a partir del fenómeno de la aglomeración y segregación académica en el 2018. *Cuadernos de Economía*, 41(86), 109-128.

Este artículo confirma la existencia de aglomeración en la calidad educativa en Cali (índice de Moran = 0,311): la alta calidad se concentra en los colegios del sur de Cali (comuna 22), mientras la baja está ubicada al oriente de la ciudad (comunidades 13, 14 y 15). Como factores determinantes, se establece lo que la literatura en general concluye, además, y como aporte principal, se determinó que la criminalidad afecta de forma negativa la calidad educativa en Cali, esto es, cuando aumenta en un punto porcentual la cantidad de homicidios alrededor del colegio, la calidad educativa disminuye, en promedio, en 0,63 puntos.

Palabras clave: análisis de patrones puntuales, calidad educativa, segregación, aglomeración, índice de Moran.

JEL: I21, I24.

INTRODUCTION

This article seeks to measure the degree of agglomeration of educational quality in the city of Cali, Colombia, in order to determine its correlation with the phenomenon of academic segregation: the non-socialization (or non-mixture) between populations with low levels of educational achievement and their higher-performing peers. In this study, the educational quality of schools is measured using student scores achieved in SABER11 tests, the most important tests for secondary school students, administered by the Colombian Institute for Educational Assessment and Evaluation (ICFES, for its initials in Spanish).

Academic segregation is the term used by the Organisation for Economic Cooperation and Development (OECD) to designate the segregation of students by their level of school competencies. There are four different types of academic segregation. The first involves separating students into different groups by academic ability at an early age, with the result that they follow different academic tracks (early tracking). The second form of academic segregation results from the distinction between educational centres, based on the average skill levels of their students. The third modality derives from the ways educational centres group their students and the existence of options or sections reserved for the best students within a given centre. The final form of academic segregation occurs between state and private centres, as a result of the so-called right to freedom to choose educational centres (Rubia, 2013).

For its part, agglomeration of educational quality refers to the concentration of high quality educational institutions in a given space, a factor that creates a niche effect.

To analyse the relation between the agglomeration of educational quality and academic performance we use the most important ICFES tests, administered by each school in the city of Cali. The research focuses on Year 11 students. We examine their mathematics and critical reading skills using the results of the SABER11 tests for 2018, employing to the methodology proposed by ICFES (2016) for the purpose.

The quality of the education system is measured to determine the extent to which the goals and objectives defined for the teaching and learning processes have been achieved. Measurement using standardised tests or evaluations is a fundamental tool for obtaining the reliable information required to improve the education system. Based on these evaluations, it is possible to identify both the strengths and weaknesses of the system and to provide feedback for educational institutions and local authorities (Ministerio de Educación Nacional, Colombia, 2006).

Since 1968 Colombia has been applying what is currently known as the SABER tests to evaluate its education system, and participates in different international evaluations of school performance, which are applied at different educational levels and evaluate different competencies. The results of the tests have been used in various studies that have utilised multiple methodologies to identify factors associated with academic performance (Ayala *et al.*, 2011; Castro *et al.*, 2018; Sánchez,

2011), education gap estimates (Cárcamo & Mola, 2012; Castro et al., 2017), and efficiency studies (OECD, 2016).

The main objective of this research is to explain how the process of agglomeration of educational quality has generated academic segregation, by georeferencing educational quality in schools in Cali in 2018. First, we use classical and Bayesian methods to evaluate the different methodologies available for aggregating SABER11 test averages at school level. Next, we determine the degree of concentration of educational quality in Cali, using Kernel density estimation maps that illustrate the spatial distribution and patterns of the scores. Finally, the principal factors associated with the phenomenon of academic segregation are identified using econometric models to interpret point patterns.

Given the above, this article seeks to answer the following principal research question: What degree of agglomeration of educational quality existed in Cali in 2018? More specifically, how concentrated is educational quality in Cali? In particular, the authors seek to answer the following questions: What is the relationship between educational agglomeration and academic segregation in Cali?; what socioeconomic aspects are significant in explaining academic segregation? and, how is educational quality distributed in Cali?

The research problem is of interest because the results obtained in previous studies focus on the effectiveness or determinants of school performance, while ignoring a feature that has barely been addressed by studies of educational quality in Cali, namely the spatial location of schools. This aspect is associated with educational quality and might contribute to the creation of educational clusters (see Vélez, 2009). Likewise, there is a scarcity of research that explores whether educational agglomeration is a cause of academic segregation.

THE STATE OF THE ART

Research into education has increased since the second half of the 20th Century, and different links have been found between the personal characteristics of students, family and school factors, socioeconomic context and academic performance. Accordingly, this section provides a review of academic articles that have examined the issue at national (Colombian) and international level.

The literature review focuses on the economics of education and spatial relationships, with an emphasis on the concepts of educational agglomeration and segregation.

From an economic point of view, the notion of agglomeration is linked to three concepts: 1) economies of scale, 2) the size of the local market, and 3) transportation costs (Fujita et al., 1999). It is important to mention that in the case of services such as education positive externalities have a fundamental effect on the quality of life of the population. Thus, economic approaches must take spatiality into

account in order to avoid educational segregation; this is a topic that is addressed later on in the article.

Sabatini (2006) highlights three degrees of concentration: 1) the spatial concentration of social groups, 2) the social homogeneity found within different areas found in the cities including their ghetto areas, and 3) the social prestige (discredit) of neighbourhoods, associated with stigma. A study prepared by the Bogotá Dirección de Estratificación de Bogotá y Departamento de Estadística Universidad Nacional (2016), analyses three relevant aspects: socioeconomic and spatial dimensions and the characteristics of the agglomeration.

Kaztman (2011) states that children usually study close to their places of residence, and that the geographical location of their households is associated with situations that affect school performance. Accordingly, Vivas (2007) and Duarte *et al.* (2012) argue that students with lower socioeconomic status tend to be disadvantaged because of educational segregation, as they attend schools of lower educational quality. Seen from another angle, in a comparative analysis of Rio de Janeiro and Bogotá, Rubiano (2017) suggests that more egalitarian educational centres would contribute to creating societies where it would be easier for individuals to recognise the other and, therefore, to develop more peaceful and harmonious social scenarios.

Kaztman (2010), considers it important to distinguish between the concepts of segmentation and segregation, the former referring to differences in opportunities and reduced chances to interact with others, while the latter also includes a willingness to establish barriers that tend to exclude other groups or social categories. Kaztman argues that in Latin America it is more appropriate to speak of school segmentation than segregation. This article expands on this view.

Iregui *et al.* (2007) analyse the academic performance of students using a production function of the education system alongside stochastic frontier techniques. They conclude that the variables associated with school infrastructure and with the socio-economic environment of students have a positive and significant impact on academic performance. The results also suggest that private schools benefit from more favourable surroundings, given that, on average, they serve students with higher incomes. However, when equivalent contexts are assumed, they find no major differences in efficiency between state and private schools.

For their part, Jola (2011) seeks evidence of the determinants of education quality in Colombia by analysing the results of the OECD's 2006 PISA test. Employing an econometric analysis, they find that the physical resources of the campus, the curriculum or academic content and the competence of the teachers are positive determinants of test results. Similarly, the analysis reveals that a greater abundance of educational resources at home and the dedication and interest of individual students in science and mathematics are associated with higher performance.

In 1999 Aguerre (2002) conducted a study in Argentina and Uruguay, on the structure of social and institutional determinants of the performance of students in the sixth year of primary education. The results suggest that the strongest determinants of educational inequality are related to the theory of cultural reproduction, according to which education is a means by which existing unequal social relations are reproduced. He concludes, moreover, that the socio-cultural composition of schools has a greater marginal effect on academic performance than individual capital variables.

Murillo and Martínez (2017) find, for Colombia, that socio-economic, ethnic-cultural and immigrant/native school segregation is high in comparison with the rest of Latin America. They calculated the dissimilarity index, finding that segregation in Colombia is at 0.59 (high segregation is defined as between 0.5 and 0.6 in the study). Furthermore, using the 2015 PISA tests, Vazquez (2018) shows that Colombia is among the countries with the highest levels of school segregation in the world, alongside five other countries in the Latin American region (Argentina, Chile, Costa Rica, Mexico and Peru). The authors conclude that more studies should be carried out on the subject in the Latin American region.

At the municipal level, Correa (2011) uses econometric methods to analyse the educational achievement of Year 11 students in Cali. He finds that the gap in academic performance between students is due to factors that are specific both to schools and students.

It should be noted that studies have tried to demonstrate the relation between agglomeration and segregation and school performance and engagement with educational services and their effect on the quality of life. However, *the spatial relation of agglomeration and segregation by area has not been examined, let alone at the level of microeconomic units such as schools in Cali. For this reason, this study is novel in using such information and in the spatial treatment applied to it.* The objectives, data, and methods used to redress this situation are described below.

METHODOLOGY

This study uses cross-sectional data for the year 2018. It focuses on the city of Cali in Colombia. It is exploratory, descriptive, explanatory and correlational. The main hypothesis is that the educational quality of schools in the city is affected by spatial agglomeration. It is suggested that contextual socioeconomic factors associated with educational quality - such as the presence of homicides and thefts in the neighbourhood in which schools are located, the spatial location of the school, and the characteristics of students and their families - should be used as influencing variables.

We began by calculating the average scores in mathematics and critical reading of Year 11 students who attended schools in Cali in 2018, using the methodology proposed by ICFES (ICFES, 2016).

Subsequently, we aggregated students' scores for each school in Cali using Lajaunie's (1991) empirical Bayesian estimation. Inputs included information on individual student and family characteristics and school traits, extracted from the two ICFES databases for the year. We also included contextual variables for the schools, such as the number of homicides and thefts occurring within a 500 metre buffer zone around the institution.

This made it possible, on the one hand, to identify whether the agglomeration of educational quality generates academic segregation and, on the other, to establish the main determinants of the phenomenon. Once the school level information had been added, it was included in the cartography for each school using Geographic Information Systems (GIS) and R statistical software and the *raster*, *rgdal*, and *sp* packages.

Finally, we used methods such as *K-replay* functions to interpret point patterns and assess the degree of spatial agglomeration, followed by Kernel density estimation to generate heat maps enabling niches of schools containing high and low educational quality to be identified. Following the evaluation of agglomerations, different models for the quality of education were estimated and assessed according to the covariates context, school, household, and student. Among the models evaluated were classical regressions, models with spatial dependencies (SAR-Spatial Autoregressive Model) and models with spatial dependence on errors (SAE- Spatial Autoregressive Error Model). The metrics used to compare the models are: likelihood ratio tests and information criteria such as the Akaike and the Bayesian.

Data

We used information provided by ICFES, drawn from the two SABER11 tests it administered in 2018. We used the statistical software R and the packages *raster*, *rgdal*, and *sp*. First, we performed a match using the two 2018 databases. Since the information obtained comes from all over Colombia, the next step was to select the study area, using the municipality code to filter the data for Cali.

Next, using the cartography of the city's schools, the students who took the SABER11 tests were added to their respective schools using the establishment code, and cycles were created to contrast each variable of interest within the cartography.

The student variables added in this process were: scores achieved, hours in paid work, type of remuneration received, household stratum and gender. The variables for family characteristics were: household size, education achievement of parents and possession or not of internet, and television services and com-

puter. School variables were: whether the school is state or private and whether or not it is bilingual. Finally, we used the National Police database to record the number of homicides occurring within a 500 metre buffer zone around each educational institution.

Once the different variables of interest had been added to the cartography of the schools, an intersection was made between the schools and the database of Cali's constituent communes, providing the information that was subsequently inserted into the model ready for the different descriptive and spatial analyses that are presented in the following sections.

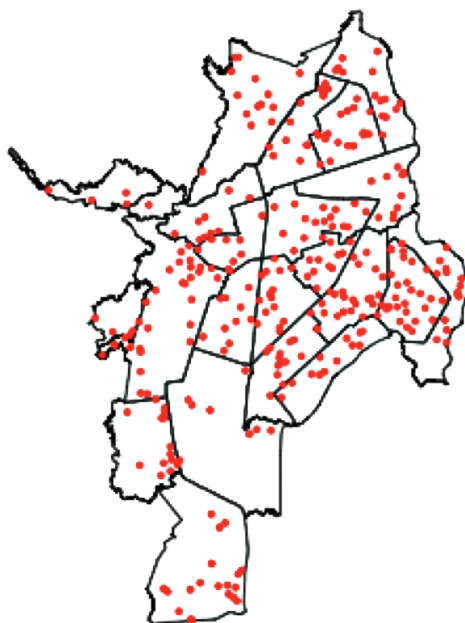
Descriptive Analysis

We used spatial statistics methods to develop the distribution maps of the schools. Specifically, we relied on the point pattern technique given the nature of the data involved.

Figure 1 is a map of Cali, showing its constituent communes. The points indicate the schools that participated in the SABER11 tests in 2018.

Figure 1.

Map of Cali with the Distribution of Schools



Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

Table 1 presents descriptive indicators of the characteristics of the students, families and schools. On average, the scores achieved by students of schools in Cali in the different subjects during 2018 were above 50. Critical reading scores and English scores were highest, with an average of 55.57 points and 54.89 points respectively. The average aggregate score obtained by students was 264.41 points.

79.2% of students were not in employment. 54.82% of those who presented the test were female. The commonest social strata - with 31.55% and 31.38% respectively - were 3 and 2 (in Colombia the population is divided into six social strata, 6 being the most affluent and 1 the poorest). The commonest level of parental educational achievement was completion of secondary school (30.37% of mothers and 28.67% of fathers).

Analysis of household size found that 55.19% of households were made up of 3 to 4 people. Most students (86.13%) had internet services, 51.66% of the schools were private and 96.71% were monolingual. Finally, in 2018 the average number of homicides per school within the 500 metre buffer zone was 9.22.

Table 1.
Descriptive Indicators

Variable	Percentage/ Average	[95% CI]	Spearman correlation coefficient
<i>Test scores</i>			
Overall Score	264.41	[260.49 -268.34] - 52,61]	
Mathematics Score	52.17	[51.32 - 53.02]	
Critical Reading Score	55.57	[54.91 - 56.23]	
Natural Science Score	52.01	[51.21 - 52.80]	
Social Science Score	51.09	[50.27 - 51.91]	
English Score	54.89	[53.86 - 55.93]	
<i>Working hours per week</i>			
Does not work	79.72	[79.10 – 80.32]	0.1756
Works less than 10 hours	14.85	[14.31 – 15.39]	-0.2211
Works between 11 and 20 hours	0	[0.00 – 0.00]	
Works between 21 and 30 hours	2.59	[2.36 – 2.85]	-0.2832
Works more than 30 hours	2.84	[2.59 – 3.11]	-0.3329

(Continued)

Table 1.
Descriptive Indicators

Variable	Percentage/ Average	[95% CI]	Spearman correlation coefficient
<i>Gender</i>			
Male	45.18	[44.47 - 45.89]	0.0113
Female	54.82	[54.10 - 55.53]	0.0294
<i>Stratum</i>			
Stratum 1	18.05	[17.48 - 18.63]	-0.4883
Stratum 2	31.38	[30.68 - 32.08]	-0.3838
Stratum 3	31.55	[30.85 - 32.25]	0.1332
Stratum 4	9.81	[9.37 - 10.27]	0.5075
Stratum 5	6.45	[6.09 - 6.83]	0.5111
Stratum 6	2.76	[2.53 - 3.02]	0.4067
<i>Education of the mother</i>			
None	1.27	1.11 - 1.46]	-0.3562
Incomplete primary schooling	9.12	[8.69 - 9.57]	-0.4884
Complete primary schooling	6.36	[5.99 - 6.74]	-0.4694
Incomplete secondary school	14.75	[14.22 - 15.29]	-0.3789
Complete secondary school	30.37	[29.68 - 31.07]	-0.1783
Incomplete technical or technological education	3.40	[3.14 - 3.69]	0.0871
Complete technical or technological education	12.08	[11.59 - 12.58]	0.2090
Incomplete undergraduate studies	3.32	[3.06 - 3.60]	0.3817
Complete undergraduate studies	15.55	[15.01-16.11]	0.6095
Graduate studies	3.78	[3.50 - 4.08]	0.5958
<i>Education of the father</i>			
None	2.44	[2.21 - 2 .69]	-0.4090
Incomplete primary schooling	12.42	[11.91-12.94]	-0.4393
Complete primary schooling	7.15	[6.59 - 7.57]	-0.4017
Incomplete secondary school	15.70	[15.14 -16.28]	-0.2458
Complete secondary school	28.67	[27.97 - 29.38]	-0.1094

(Continued)

Table 1.
Descriptive Indicators

Variable	Percentage/ Average	[95% CI]	Spearman correlation coefficient
Incomplete technical or technological education	2.67	[2.43 - 2.94]	0.1379
Complete technical or technological education	9.99	[9.53 - 10.47]	0.2030
Incomplete undergraduate studies	2.72	[2.48 - 2.99]	0.4117
Complete undergraduate studies	14.26	[13.72 - 14.82]	0.5949
Graduate studies	3.97	[3.68 - 4.29]	0.6380
<i>Characteristics of the household</i>			
<i>Household size</i>			
1 to 2	9.22	[8.81 - 9.66]	0.0688
3 to 4	55.19	[54.47 - 55.92]	0.1714
5 to 6	26.68	[26.04 - 27.33]	-0.0672
7 to 8	6.46	[6.11 - 6.83]	-0.3072
9 or more	2.44	[2.22 - 2.68]	-0.3178
Has access to the internet	86.13	[85.60 - 86.63]	0.1723
Does not have access to the internet	13.87	[13.36 - 14.39]	-0.4479
Has a TV	88.48	[87.99 - 88.94]	0.1306
Does not have a TV	11.52	[11.06 - 12.01]	-0.3081
Has a computer	78.46	[77.85 - 79.05]	0.1867
Does not have a computer	21.54	[20.94 - 22.15]	-0.4957
<i>Characteristics of the school</i>			
State School	48.34	[47.62 - 49.05]	-0.2339
Private School	51.66	[50.94 - 52.37]	0.3282
Bilingual School	3.29	[3.03 - 3.56]	0.2613
Non-bilingual School	96.71	[96.44 - 96.97]	-0.0494
Number of homicides	9.22	[8.73 - 9.72]	

Note: As this article does not aim to measure educational gaps and/or characterizations and differences between state and private schools, Table 1 shows aggregate statistics for educational facilities in Cali.

Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

Existence of Agglomeration

The K Function is a Complete Spatial Randomness (CSR) Test evaluated using distances. It is employed to analyse the spatial pattern of data and to summarise spatial dependence over a range of distances. Figures 2 and 3 present the results of the K function, both for the worst-performing schools and for those with better results.

Some patterns were discernible at the spatial level. The *poisson* pattern was random and, as may be seen in Figures 2 and 3, the curves observed are above the random pattern, indicating that the data confirms a spatial agglomeration. These results suggest the presence of an agglomeration of educational quality in Cali’s schools, a finding that is corroborated by Figure 4.

Figure 2.
K Function, worst-performing schools
(Low scores)

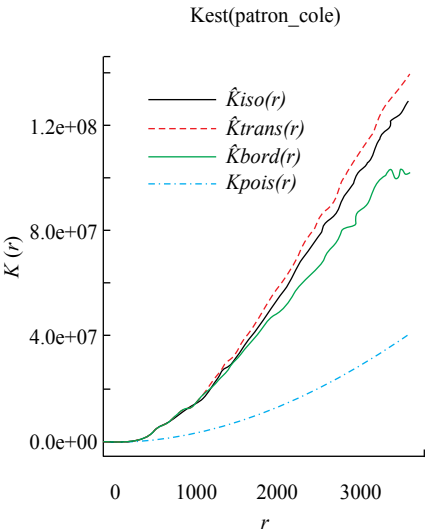
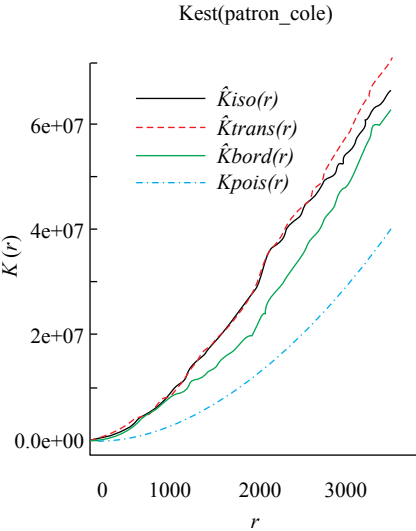


Figure 3.
K Function, Better-Performing Schools
(High scores)

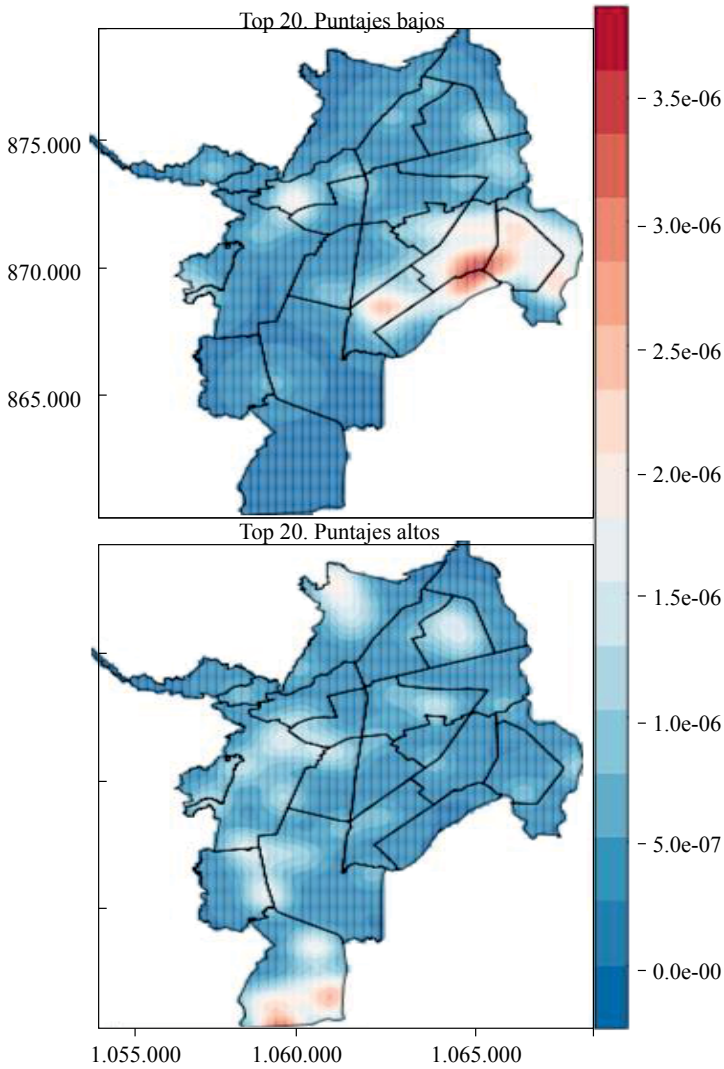


Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

Figure 4 shows the 20 schools that achieved the lowest scores in the test and the 20 that scored best. The former are located in the east of the city, specifically in communes 13, 14 and 15 while the latter are all located in the south, specifically, in commune 22.

This finding supports the initial hypothesis that educational quality in Cali’s schools presents a structure of agglomeration at the spatial level, meaning that high-quality educational institutions are spatially concentrated, creating a niche effect.

Figure 4.
Kernel Density Estimation



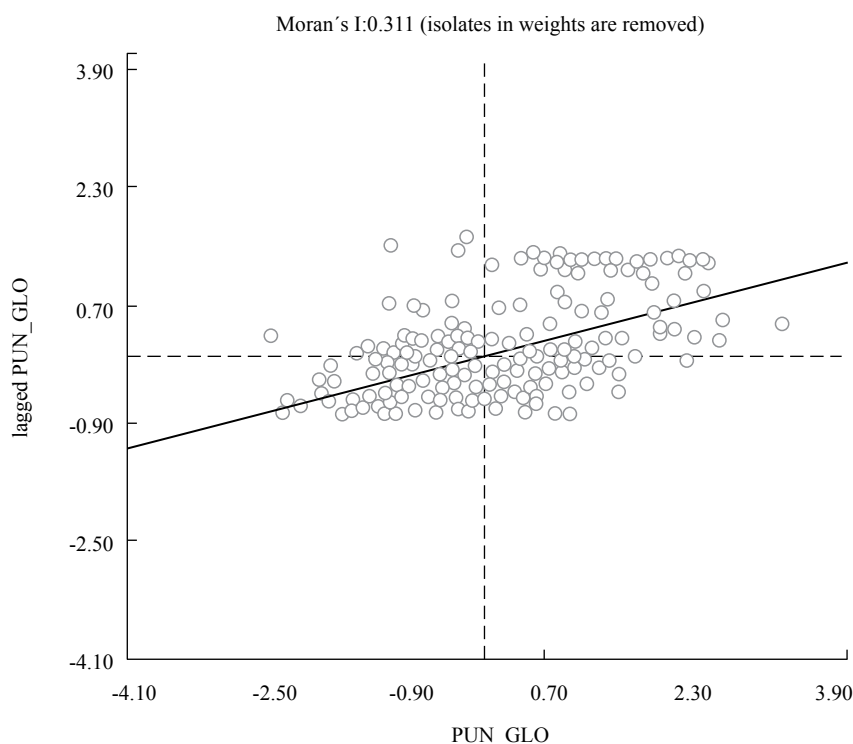
Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

To corroborate these findings, we calculated Moran's Index, which analyses spatial autocorrelation. This index can take values between -1 and 1. A value greater than zero indicates positive spatial autocorrelation (a cluster pattern) while a value below zero indicates a dispersion pattern.

We obtained a value of 0.311 for Moran's Index with a significance of 1%, indicating positive spatial autocorrelation confirming the existence of an agglomeration of educational quality in Cali. The educational quality of the schools is more related to the nearby schools than to those further away. Both low-quality and high-quality schools tend to be concentrated (lower left quadrant and upper right quadrant).

Figure 5.

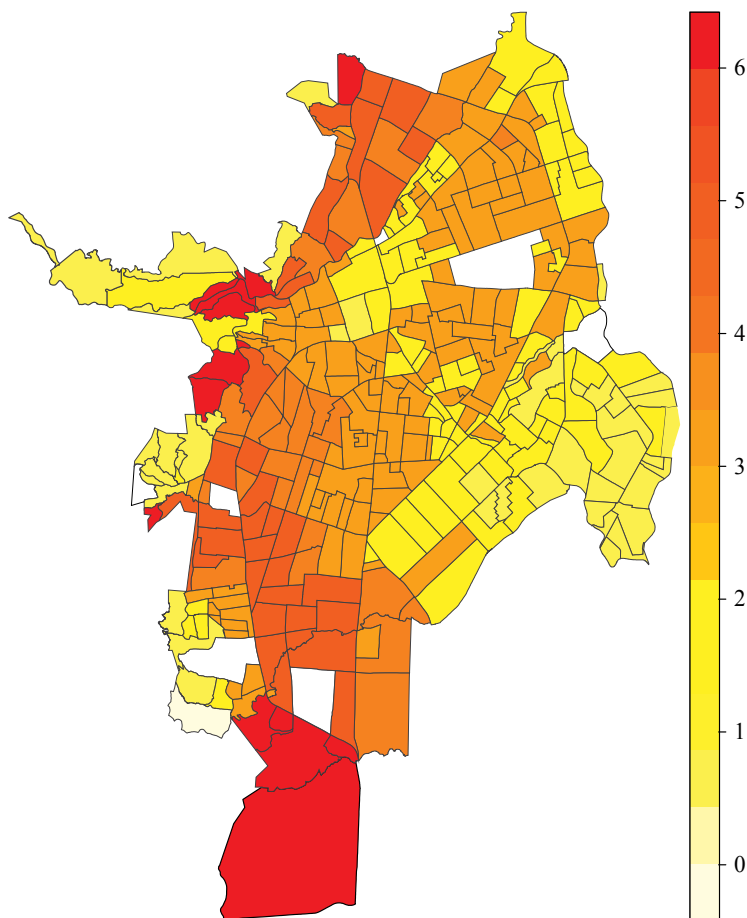
Moran's Index, Cali 2018



Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

Moreover, Figure 6 illustrates a clear trend at the spatial level. The schools with the highest test scores are located in the south of the city, specifically in neighbourhoods of high socioeconomic strata, while the schools with the lowest scores are located in the eastern part of the city where most of the neighbourhoods belong to lower socioeconomic strata.

Figure 6.
Stratum Map by Neighbourhood



Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

PROPOSED MODEL

The dependent variable in the model is the educational quality of Cali's schools. To determine the best combination of covariates that might affect educational quality, different personal and socio-economic variables were considered in the construction of the model.

We used the Spatial Lag Model, represented by:

$$y = \rho W y + X \beta + \varepsilon \quad (1)$$

Where y is the dependent variable. This is a vector that contains the scores obtained for each of the subjects evaluated by the SABER11 test. Its size is $(N \times 1)$. ρ is the spatial lag parameter to be estimated. W is the matrix of spatial weights; it is squared and its size is $N \times N$ and represents the intensity of interdependence that can exist between schools. For the construction of the W matrix, schools within a distance of 3,000 square metres of each other were considered to be neighbours. X represents the matrix of covariates and ε is a vector with uncorrelated errors.

The estimation of the model was carried out using GeoDa software. This modelling process enables the possible relationship between educational quality and explanatory variables in neighbouring schools to be analysed.

Including the possible covariates that might explain the quality of education in Cali, the statistical model is:

$$\begin{aligned} Score_{ij} = & \rho W_{ij} Score_{ij} + \beta_0 + \beta_1 Edu.M_{ij} + \beta_2 Work.hours_{ij} + \\ & \beta_3 House.S_{ij} + \beta_4 Women_{ij} + \beta_5 Men_{ij} + \beta_6 Bilingual_{ij} + \\ & \beta_7 Sch.typ_{ij} + \beta_8 Homicides_{ij} + \varepsilon_{ij} \end{aligned} \quad (2)$$

Where **Score** is the average score obtained by students in the subjects of Mathematics and Critical Reading, and the overall score achieved in the SABER11 test. Student scores were aggregated at school level.

Edu.M represents the level of educational achievement of the mother. For this variable, the most representative educational level was assigned to each school. **Work.hours** indicates the number of hours each student worked weekly. **House.S** defines the number of people with whom each student lived in their home.

Female and **Male** indicate the proportion of girls and boys who took SABER11 tests per school. **Bilingual** indicates whether the school is bilingual or not. We expected bilingual schools to increase student scores in the SABER11 tests.

Sch.Type indicates whether the school is state [$Sch.Type=1$] or private [$Sch.Type=0$]. We expect private schools to perform better in the tests. Finally, **Homicides** determines the number of homicides that occurred within a 500-meter buffer zone around each school. We expected a greater quantity of homicides to affect educational quality negatively.

EMPIRICAL RESULTS

Three simulations were performed using the spatial lag model in the GeoDa software, this model introduces a Lag spatial variable, which allows spatial dependence to be observed.

The results are as follows (see Table 1).

Table 2.

Estimates of the Educational Quality in Cali for the Year 2018

Variable	Overall	Mathematics	Critical Reading
(Spatial Lag)	0.247***	0.233**	0.214**
Intercept	168.383***	32.474***	39.766***
Edu.M	4.217***	0.933***	0.709***
Work.hours	-1.499***	-0.374***	-0.254***
House.S	-7.429***	-1.716***	-1.422***
Female	0.114	0.022	-0.000
Male	-0.005	0.004	0.016
Bilingual	25.196***	5.636***	3.757***
Sch.Typ	-13.952***	-2.526**	-2.311***
Homicides	-0.638***	-0.109**	-0.110***

*** Significant at 1 % ** Significant at 5 % * Significant at 10 %

Source: Data from the Colombian Institute for Educational Assessment and Evaluation (ICFES) for the SABER11 tests for 2018.

The dependent variables are the scores in mathematics, critical reading and overall, which were taken as proxy variables for educational quality. The regression shows that the spatial lag in all three cases is significant, as is the intercept.

In the model in which the overall score is the dependent variable, it was found that when the number of homicides around the school increases by one percentage point, educational quality decreases by 0.63 percentage points, a finding that is consistent with our expectations. Moreover, the education of the mother is significant at 1% and has a positive relationship with the dependent variable. That is, when the educational level of the mother increases by one percentage point, the student's overall score (educational quality) increases by 4.22 percentage points.

Another significant result in this model is the number of hours the student works, which has a value of -1.49 points and is consistent with our expectations. If students are obliged to distribute their time between working and studying, they may not be able to fulfil their school activities adequately.

In the three regressions, the relationships found were as expected. As the educational level of the mother increases, their children obtain higher test results, indicating better educational quality. Similarly, if the school is bilingual, test scores are higher. On the other hand, when the number of hours worked increases, test scores decline. Similarly, when the number of people in the household is higher, or the number of homicides around the school rises, test scores fall.

Additionally, in the three models it was found that educational quality is lower in state schools. This could be explained by the fact that there are approximately

three times more state schools than private and that consequently, the number of students who attend state schools and take the test is higher. A larger number of very low scores could lower the averages.

Another explanation might be that the educational quality of private schools tend to be higher as they are located in better geographical conditions. The variables related to gender were not significant in the models.

CONCLUSIONS

The point patterns methodology is an interesting alternative in the field of spatial statistics, as it allows the agglomeration of educational quality in Cali to be analysed in a novel way, as it provides geo-referenced information for each school. This makes it possible to observe the different socioeconomic factors that may affect agglomeration holistically.

It was found that high educational quality is concentrated in the south of the city, specifically in commune 22. This finding is consistent with the fact that most of the city's higher-income neighbourhoods are located there. The cluster of the worst-performing schools is found in the east of the city, in communes 13, 14 and 15.

In addition, school facilities are uneven. In commune 22 (high quality) the schools have superior educational materials and infrastructure to the schools in communes 13, 14 and 15 (low quality). Therefore, to minimise the negative effects of educational agglomeration and segregation in Cali, and to increase educational quality, efforts must be increased to provide better resources and facilities to schools located in disadvantaged environments and efforts should be made to attract better-trained teachers and managers.

One of the principal findings of this study is the negative effect of criminality on academic performance. This shows the necessity of improving social security and public order policies in the city.

Similar studies should be conducted to identify the large educational gaps in Colombia's major cities, such as Cali. The hillside areas that concentrate the largest number of state schools perform least well in the SABER11 tests. Additionally, the context of these schools (risk areas, high homicide rates, etc.), present challenges that might complicate further research into the situation.

Social division is reproduced in schools. For example, in commune 22 many schools operate a "club-type" model whereby parents pay for "shares" in return for high-quality quality bilingual education, while educational quality in many other communes is very low. This situation exacerbates differences that may generate problems, given the social gaps that characterise relations between inhabitants of the same city. It is a challenge for state policy to analyse educational segregation and build solutions so as not to keep aggravating the situation.

Finally, it is important to note that the investigations described in this article highlight the need to study educational segregation in Latin America. It is likewise necessary to deal with this issue in Colombia, where research on the problem is scarce. Simultaneously, it is crucial that these concerns are transmitted to policymakers.

REFERENCES

1. Aguerre, T. F. (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*, 7(16), 501-536.
2. Ayala, J., Marrugo, S., & Saray, B. (2011). Antecedentes familiares y rendimiento académico en los colegios oficiales de Cartagena. *Economía y Región*, 5(2), 43-85.
3. Cárcamo, C., & Mola, J. (2012). Diferencias por sexo en el desempeño académico en Colombia: un análisis regional. *Economía y Región*, 6(1), 133-169.
4. Castro, G., Giménez, G., & Pérez, D. (2017). Educational inequalities in Latin America, 2012 PISA: Causes of differences in school performance between public and private schools. *Revista de Educación*, 376, 33-61.
5. Castro, G., Giménez, G., & Pérez, D. (2018). Estimación de los factores condicionantes de la adquisición de competencias académicas en América Latina en presencia de endogeneidad. *Revista CEPAL, Chile*, 124, 35-59.
6. Correa, J. (2011). Determinantes del rendimiento educativo de los estudiantes de secundaria en Cali: un análisis multinivel. *Revista Sociedad y Economía*, 6, 81-105.
7. Dirección de Estratificación de Bogotá y Departamento de Estadística Universidad Nacional. (2016). *La estratificación en Bogotá: impacto social y alternativas para asignar subsidios*. Secretaría Distrital de Planeación.
8. Duarte, J., Bos, M., & Moreno, J. M. (2012). *Calidad, igualdad y equidad en la educación colombiana*. Banco Interamericano de Desarrollo. Notas Técnicas No. IDB-TN- 396. División de Educación, Banco Interamericano de Desarrollo.
9. Fujita, M., Krugman, P., & Venables, A. (1999). *The spatial economy cities, regions, and international trade*. MIT Press.
10. Instituto Colombiano para la Evaluación de la Educación (ICFES) (2016). *SABER en breve*. 5 ed. ICFES.
11. Iregui, A. M., Melo, L., & Ramos, J. (2007). Análisis de eficiencia de la educación en Colombia. *Revista de Economía del Rosario*, 10(1), 21-41.
12. Jola, A. F. (2011). Determinantes de la calidad de la educación media en Colombia: un análisis de los resultados PISA 2006 y del plan sectorial

- “Revolución Educativa”. *Revista Coyuntura Económica: Investigación Económica y Social*, 41(1), 25-61.
13. Kaztman, R. (2010). La dimensión espacial de la cohesión social en América Latina. In *Cohesión social en América Latina y el Caribe: una revisión de conceptos, marcos de referencia e indicadores* (pp. 39-84). Naciones Unidas, CEPAL.
 14. Kaztman, R. (2011). *Infancia en América Latina: privaciones habitacionales y desarrollo de capital humano*. Naciones Unidas, CEPAL.
 15. Lajaunie, C. (1991). *Local risk estimation for a rare noncontagious disease based on observed frequencies* (1a ed.). Center de Geostatistique, Ecole des mines de Paris.
 16. Ministerio de Educación Nacional, Colombia. (2006). *Al tablero*. <https://www.mineducacion.gov.co/1621/article-107321.html>
 17. Murillo, F. J., & Martínez-Garrido, C. (2017). Estimación de la magnitud de la segregación escolar en América Latina. *magis, Revista Internacional de Investigación en Educación*, 9(19), 11-30. <https://doi.org/10.11144/Javeriana.m9-19.emse>
 18. OCDE. (2016). *Education in Colombia. Reviews of National Policies for Education*. Organización para la Cooperación y el Desarrollo Económicos, OCDE.
 19. Rubia, F. A. (2013). La segregación escolar en nuestro sistema educativo. *Fórum Aragón*, 10, 47-52.
 20. Rubiano, M. (2017). *Segregación socioespacial. Miradas cruzadas desde Rio de Janeiro, Bogotá y sus áreas de influencia metropolitana*. Universidad Piloto de Colombia.
 21. Sabatini, F. (2006). *La segregación social del espacio en las ciudades de América*. Departamento de Desarrollo Sostenible, Banco Interamericano de Desarrollo.
 22. Sánchez, A. (2011). *Etnia y desempeño académico en Colombia* (Documentos de Trabajo sobre Economía Regional, 156). Banco de la República, Centro de Estudios Económicos Regionales (CEER).
 23. Vazquez, E. (mayo de 2018). Segregación socioeconómica en las aulas. ¿Es América Latina la región de mayor segregación del mundo? Blog del CEDLAS, <http://www.cedlas.econo.unlp.edu.ar/wp/segregacion-socioeconomica-en-las-aulas-es-america-latina-la-region-de-mayor-segregacion-del-mundo>
 24. Vélez, J. E. (2009). ¿Qué puede explicar los procesos de aglomeración en Colombia? Dos enfoques alternativos. *Revista de Economía & Administración*, 6(2), 59-91.
 25. Vivas, H. (2007). *Educación, background familiar y calidad de los entornos locales en Colombia*, Departamento de Economía Aplicada, Universidad Autónoma de Barcelona.