https://doi.org/10.46551/epp2024v12n0203

Economia e Políticas Públicas

Revista Eletrônica de Economia

Vol.12, n.2, jul./dez.2024
ISSN: 2318-647X (online)

https://www.periodicos.unimontes.br/index.php/economiaepoliticaspublicas

A influência da regularidade docente a partir de modelos econométricos espaciais para Minas Gerais

The influence of teaching regularity from spatial econometric models for Minas Gerais

La influencia de la regularidad docente a partir de modelos econométricos espaciales para Minas Gerais

Vanessa Ragone Azevedo[1]

Resumo: A importância da busca pelos determinantes educacionais é amplamente debatida na teoria econômica, e os resultados dos estudos nas últimas décadas do século passado indicavam apenas a relevância das características socioeconômicas do aluno e a escolaridade da mãe como fatores relevantes. Porém, a utilização de outros modelos e aplicações vêm demonstrando que características do professor, da escola e da turma são relevantes na modelagem do desempenho discente. A partir de dados do Saeb e do Censo da Educação Básica foi construída uma base de dados com o objetivo de analisar as características espaciais desta questão, assim como a influência da regularidade docente no desempenho dos estudantes do ensino fundamental (anos finais). Os resultados obtidos a partir da aplicação de métodos de georreferenciamento indicaram a forte influência das características socioeconômicas e da escolaridade materna, mas também reforçaram a importância da distorção idade-série e do parâmetro do multiplicador espacial, indicando que as notas dos alunos são influenciadas pela vizinhança. Outro resultado importante foi em relação ao impacto dos gastos governamentais em atividades de educação que indicou ser significativo em todos os modelos estimados.

[1]) Doctor in Economics from the Federal University of Juiz de Fora and Professor in the Department of Economics at the Federal University of Maranhão. ORCID: https://orcid.org/0000-0001-6998-7323; E-mail: vanessa.ragone@gmail.com.

Palavras-chave: Educação, Determinantes do Desempenho, Econometria Espacial.

Abstract: The importance of the search for educational determinants is widely debated in economic theory, the results of studies between the 60's and 90's indicated only the relevance of the student's socioeconomic characteristics and the mother's education as relevant factors, but the use of other models and Applications have shown that teacher, school and class characteristics are relevant in modeling student performance. From data from the Saeb and the Censo da Educação Básica, a database was built to analyze the spatial characteristics of this issue, as well as the influence of teacher regularity on the performance of elementary school students (final years). The results obtained from the application of georeferencing methods indicated the strong influence of socioeconomic characteristics and maternal education, but also indicated the importance of age-grade distortion and spatial multiplier parameter, indicating that students' grades are influenced by neighborhood. Another important result was in relation to the impact of government spending on education activities which indicated to be significant in all estimated models.

Keywords: Education, Educational Determinants, Spatial Econometric.

Resumen: La importancia de la búsqueda de determinantes educativos es ampliamente debatida en la teoría económica, los resultados de estudios en las últimas décadas del siglo pasado solo indicaron la relevancia de las características socioeconómicas del estudiante y la educación de la madre como factores relevantes. Sin embargo, el uso de otros modelos y aplicaciones ha venido demostrando que las características del docente, la escuela y la clase son relevantes para modelar el desempeño de los estudiantes. A partir de datos de la Saeb y del Censo de Educación Básica, se construyó una base de datos con el objetivo de analizar las características espaciales de este tema, así como la influencia de la regularidad docente en el desempeño de los alumnos de la enseñanza básica (últimos años). Los resultados indicaron la fuerte influencia de las características socioeconómicas y la educación materna, pero también reforzaron la importancia de la distorsión edad-grado y el parámetro multiplicador espacial, indicando que las calificaciones de los estudiantes están influenciadas por el barrio. Otro resultado importante estuvo relacionado con el impacto del gasto público en actividades educativas, que indicó ser significativo en todos los modelos estimados.

Palabras clave: Educación, Determinantes Educativos, Econometría Espacial.

Introduction

The economic development is linked to the improvement of local human capital. Therefore, a qualified educational system is among the main objectives of a government seeking economic growth and development. The search for improvements in the educational system is linked to the benefits that education provides in the individual and social spheres. Among the individual benefits, the ability to generate future income and increased life expectancy stand out (Murnane et al, 1995; Menezes- Filho, 2001; Murphy and Peltzman, 2004).

Based on the understanding of the importance of education and investments in this area from both an individual and economic development perspective, the literature has for years strived to identify the determinants of education, using school performance as a proxy to measure the quality of education. Coleman et al. (1966) found results indicating that approximately 80% of school performance can be explained by the family environment. Later studies, including those applied to Brazilian data, found results similar to those of Coleman et al. (1966), reinforcing the importance of family background in academic performance (Albernaz, Ferreira and Franco, 2002; Soares, 2005; Menezes-Filho, 2007; Machado et al., 2008).

With the emergence of new models and forms of analysis, several studies have indicated that, when controlling for student and family environment characteristics, school and school context characteristics have a significant influence on student performance. Several studies applying hierarchical models that allow for the control and analysis of different levels (school, class, teacher, student) have obtained results demonstrating that the influence of the school context, the performance of nearby students and the socioeconomic level of the class have an influence on educational results at the individual level (Felício and Fernandes, 2005; Alves and Soares, 2007; Riani and Rios-Neto, 2008).

Vernier, Bagolon and Jacinto (2015) use a more aggregated analysis, introducing, in addition to the socioeconomic characteristics of students, the characteristics of teachers, principals and the school, finding an impact on all dimensions, being significant for issues related to the region, so that location matters when analyzing school performance.

Using the theory developed by Fujita, Krugman and Venables (1999) as a basis, it is possible to understand how location plays an important role and the proximity effect becomes more intense as the exchange of information and knowledge is more fluid and effective.

Considering that the teacher in the classroom plays a significant role in the performance of his students, analyzing the relationship between the teacher regularity index (that is, an index that makes it possible to analyze places that have had a lower flow of teachers in the last 5 years) and its relationship with student proficiency is important to understand how the change of teachers, that is, changes in the teaching staff, can have an effect on student performance.

This article has the general objective of analyzing the impact of human capital at the municipal level for the state of Minas Gerais, analyzing school proficiency, a proxy variable for teaching quality, and the spatial aspect of this variable, as well as how teacher regularity affects school performance. This article proposes an investigation of human capital spillover based on the approach used by Ertur and Koch (2007), in which the stock of knowledge accumulated in a country depends on the stock of knowledge of neighboring countries. Based on a model of technological interdependence, the results obtained indicate that the stock of knowledge of a country generates spillovers that cross borders and affect neighboring countries, with this effect decreasing as the geographic distance increases.

From this perspective on geographic space and how it influences development and issues according to its limits, we intend to answer the following question: How do educational externalities measured by spending on education, teacher regularity and teaching effort impact school performance? The hypothesis initially adopted is that there is an impact of educational externalities on student performance.

Therefore, the general objective of this study is to investigate the impact of educational externalities using spatial models. To meet this demand, in terms of specific objectives, spatial econometric models will be used to assist in the study of externalities and to verify whether institutional, economic, cultural and social characteristics of a given region have a spillover effect to neighboring regions. Several studies that use spatial econometrics analyze how economic performance and growth are influenced by the neighborhood (Easterly and Levine, 1995; Moreno and Trehan, 1997; Silveira Neto, 2001; Hewings, Magalhães and Azzoni; 2005), while other studies analyze how spending and investments in the educational area affect neighbors (Case and Rosen, 1993).

The results found advance the literature in question by using the spatial autocorrelation model of the data to verify whether there is an influence on teacher regularity. Although the variable teacher regularity is not significant in the models analyzed, the other results demonstrate compatibility with the findings already consolidated in the literature on the subject. Therefore, it is relevant to investigate variables that capture the effect of the teacher-student relationship for the development of effective public policies to improve performance.

The study is structured in 6 sections. In addition to this introduction, the next section describes the data used. Sections 3 and 4 present the exploratory analysis of spatial data and the empirical strategy for estimating the models. The sixth section presents the results obtained from the model estimations, and finally, the final considerations of the study are presented.

Database and Empirical Strategy

The data used in this study come from two different sources. The first set of data comes from educational indicators available on the INEP website for the year 2017; in the specific case of this study, only public schools will be analyzed.

According to INEP (2020), social indicators allow us to assign statistical values to the quality of education, taking into account not only student proficiency, but also the socioeconomic context in which schools are located.

In addition, data on per capita spending on education activities and per capita income from the formal market were obtained from the website of the Minas Gerais Social Responsibility Index (IMRS) of the João Pinheiro Foundation, which has several indicators for all cities in Minas Gerais.

Table 1 describes the indicators used in this article, as well as a brief description and the source of the data in question; all data are at the municipal level.

Table 1 – Description of variables of interest

Variables	Ano	Description	Source
Standardized Grade	2017	Standardized SAEB score for the final years of elementary school	Saeb
Mother's Education	2017	Proportion of students whose mothers had completed high school or higher education	Saeb
Color	2017	Proportion of white students	Saeb
Sex	2017	Proportion of male students	Saeb
Education Spending	2018	Per capita expenditure on educational activities	IMRS
Per Capita Income	2018	Average per capita income from the formal market	IMRS
Teacher Regularity (> 3)	2018	Average number of years that teachers worked at the school in the last 5 years (Varies from 0 to 5)	Basic Education Census
Age-Grade Distortion	2018	Age-Grade Distortion Rate	Basic Education Census
Average Hour/Daily Class	2018	Average Hours of Daily Class	Basic Education Census
Teaching Effort (<=3)	2018	The lower the effort indicator, the lower the teacher's workload (Varies from 1 to 6)	Basic Education Census
Average Student/Class	2018	Average number of students per class	Basic Education Census

Source: Prepared by the author.

All data were collected for the 853 municipalities of Minas Gerais. The Basic Education Assessment System – SAEB – is a system for evaluating Brazilian schools conducted by INEP. The exam analyzes student performance in the subjects of mathematics and Portuguese. In addition to the tests, questionnaires are administered to principals, teachers and students, allowing for a broader assessment of schools.

The municipal school performance, obtained through standardized grades in Portuguese and mathematics, will be the dependent variable in this study. Data on race, sex and mother's education were introduced into the model to avoid problems of important omitted variables, since socioeconomic characteristics and mother's education are variables with significant importance in school performance.

Exploratory Analysis of Spatial Data

The application of a spatial regression model must be preceded by an investigation of the existence of spatial dependence. According to Gibbons and Overman (2012), it is important to pay attention to the need to apply spatial regressions given the research question and not only to treat endogeneity problems. Thus, this article focuses on an experimentalist approach that has a greater focus on identification and causality issues, as indicated by the authors.

Vernier, Bagolin and Fochezatto (2017) highlight that location plays an important role, and the proximity effect becomes more intense as the exchange of information and knowledge is more fluid and effective. Spatial models allow the study of externalities, thus making it possible to identify the institutional, economic, cultural and social variables of a given region that have spillover effects to neighboring regions.

Based on the idea of this spillover effect, it is necessary to perform exploratory spatial data analysis (ESDA) to assess whether there is any pattern of spatial association between regions. Since the variable of interest, the standard grade for elementary school in the final years, has regions without information (missing values), in order to perform the ESDA, the following municipalities were removed from the analysis: Fervedouro, São Geraldo do Baixo, Santana do Manhuaçu, Pedra do Anta and Lagoa Grande.

Determination of the Weight Matrix

The first step in implementing AEDE is to construct a spatial weight matrix, called the “W” matrix. Although there are several possible ways of constructing matrices (socioeconomic matrix, transactional matrix, etc.), Rêgo and Pena (2012) state that most studies use four types of matrices: rook, queen, distance/inverse distance and “k nearest neighbors”.

Since the objective of this article is to analyze the spillover of the educational effect from interactions between municipalities in Minas Gerais, it is necessary to specify how the municipalities are connected. The spatial weight matrix should be chosen according to the sample structure in order to ensure that there are no “islands” – regions without neighbors, which have non-negative and finite weights and exogenous weights.

The use of the distance matrix, despite allowing the identification of the proximity of municipalities, does not allow the verification of the existence of a border between them. And, since the sizes of Brazilian municipalities are not homogeneous, the use of a weighting matrix based on distance or contiguity can generate an unbalanced structure. A common solution to this problem consists of considering the weighting matrices based on the nearest neighbors, since, in this way, it would force each unit to have the same number of neighbors (Anselin, 2002; Dominicis et al., 2013).

Although the choice of the matrix is based on theoretical criteria, it is important to highlight an important result found by LeSage and Pace (2014) who concluded that the specification of the spatial weights matrix does not interfere with the factors of the spatial regression models.

In order to specify the order of the neighborhood matrix, the criterion of Almeida (2012) will be used. According to this criterion, after testing Moran's I for a set of matrices, the matrix that has generated the highest value and that is statistically significant is selected. The coefficient of Moran's I according to the number of neighbors is presented in the following table.

Table 2 – I of Moran for k nearest neighbor matrices

Variables	Moran's I coefficient		
	5 neighbors	10 neighbors	15 neighbors
Standardized Score (final years of elementary school)	0.387	0.364	0.362
Mother's Education	0.313	0.280	0.255
Color	0.770	0.747	0.738
Sex	0.008	0.004	0.005
Education Spending	0.112	0.094	0.093
Per Capita Income (formal market)	0.238	0.202	0.203
Teacher Regularity (Medium/High)	0.131	0.118	0.108
Age-Grade Distortion	0.338	0.290	0.270
Average Hour/Daily Class	0.111	0.096	0.086
Teaching Effort (1+2+3)	0.080	0.060	0.070
Average Student/Class	0.180	0.148	0.133

Source: Prepared by the author.

From the data presented in Table 2, it is possible to define the spatial weight matrix with the three nearest neighbors (k=5) as the matrix to be used for the AEDE and future estimations.

Method for analyzing spatial autocorrelation of data

Spatial dependence arises when the value of a variable in a given location depends on the value of that variable in neighboring regions, that is, whether or not the data are distributed randomly in space. The *Moran's I* is a method that allows us to verify and make inferences about spatial dependence.

In this study, statistics that allow the identification of spatial patterns will initially be used: global and local *Moran's I*. Global *Moran's I* provides a summary of the spatial distribution of the data, i.e., a single value (average) for all regions. Local *Moran's I* (LISA), by calculating a value for each observation unit, allows the identification of different spatial distribution patterns (clusters or outliers). According to Almeida (2012), the global *Moran's I* can be defined matrixically as:

$$I = \left(\frac{n}{S_0}\right)\left(\frac{z'Wz}{z'z}\right) \qquad (1)$$

In which n is the number of regions, z indicates the values of the standardized variable of interest, Wz represents the average values of the standardized variable of interest in the neighbors, defined according to a spatial weighting matrix W. S_0 is equal to the operation $\sum\sum w_{ij}$, means that all elements of the spatial weight matrix W must be added.

The null hypothesis being tested is spatial randomness. When the calculated value exceeds the expected value (within the limit of statistical significance) it indicates that there is positive spatial autocorrelation (similarity) and values below the expected indicate negative spatial autocorrelation (dissimilarity).

The verification of local patterns and determination of regions that contribute most to spatial autocorrelation can be done by using LISA. This indicator was initially suggested by Anselin (1995), and can be calculated as follows:

$$I_l = \frac{y_j \sum_{j=1}^{n} w_{ij} y_i}{\sum_{i=1}^{n} y_i^2} \qquad (2)$$

In which n indicates the number of regions, w_{ij} are the elements of the spatial weight matrix, y_i and y_j are the values of the variable used, while i and j refer to the different locations.

An additional tool for analyzing spatial autocorrelation is based on the Moran scatterplot, which shows the spatial lag of the variable of interest on the vertical axis and the value of the variable of interest on the horizontal axis. This type of diagram provides evidence of the type of linear autocorrelation (negative or positive). When the slope is negative, there is evidence of negative spatial autocorrelation (presence of outliers) and when the slope is positive, there is evidence of positive spatial autocorrelation (presence of clusters).

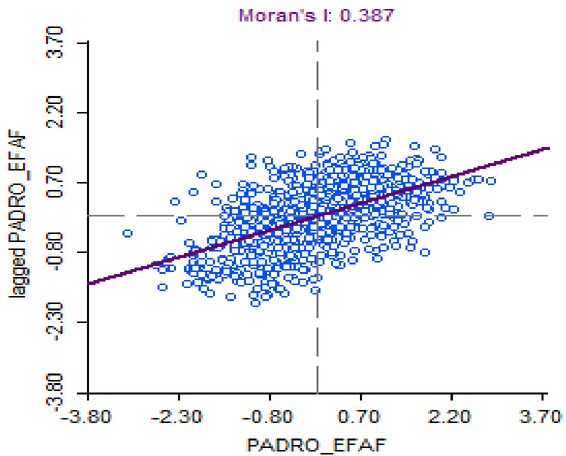
Furthermore, this diagram provides information on the types of spatial linear association, according to the distribution of data in the four quadrants, and there may be associations of the High-High (AA), Low-Low (BB), High-Low (AB) and Low-High (BA) types.

Using these methods, it is possible to perform exploratory analysis of spatial data, understanding how the data is distributed in space and the presence or absence of patterns, and studying the best model to be applied together with the theoretical framework.

Results of Exploratory Spatial Data Analysis

The results in Table 2 indicating Moran's I for the univariate spatial autocorrelation test indicate that all variables (except the gender variable) presented positive and significant results. From this, we reject the null hypothesis of spatial randomness, that is, there is spatial autocorrelation in the data for the variables investigated at the municipal level in Minas Gerais. Moran's scatter diagram, illustrated in Graph 1, reinforces the results indicated by the global Moran's I for the dependent variable.

Graph 1 – Moran scatter diagram for the standardized grade in elementary school – final years.



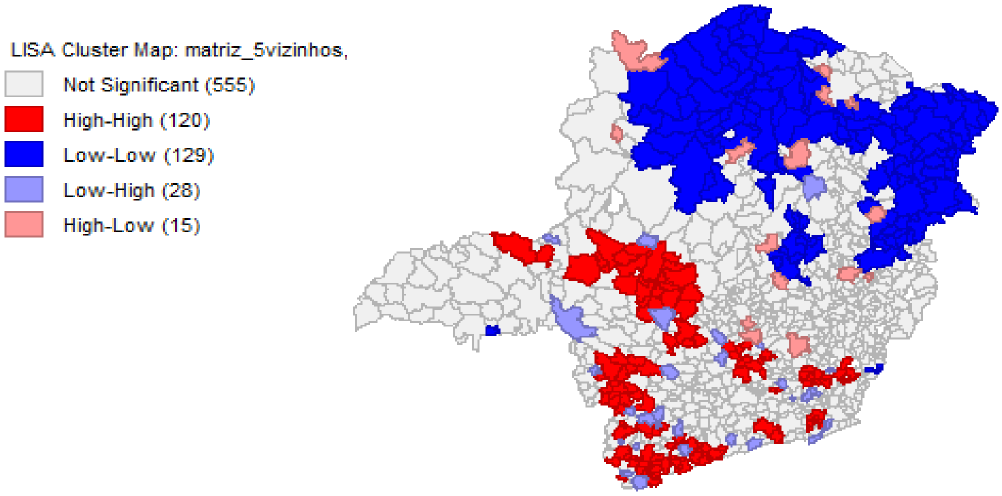
Source: Own elaboration.

Considering that Moran's I was positive for all variables, it is possible to conclude that there is a pattern of spatial concentration, indicating that municipalities that have indicators with high values have neighbors with high indicators and municipalities with low indicators have neighbors with low indicators.

With the help of local Moran's I, it is possible to better analyze local spatial patterns. Below in Figure 1 is the LISA cluster map based on the univariate local Moran's I at 5%

significance with 999999 permutations for the standardized grade of Elementary School – Final Years of Saeb.

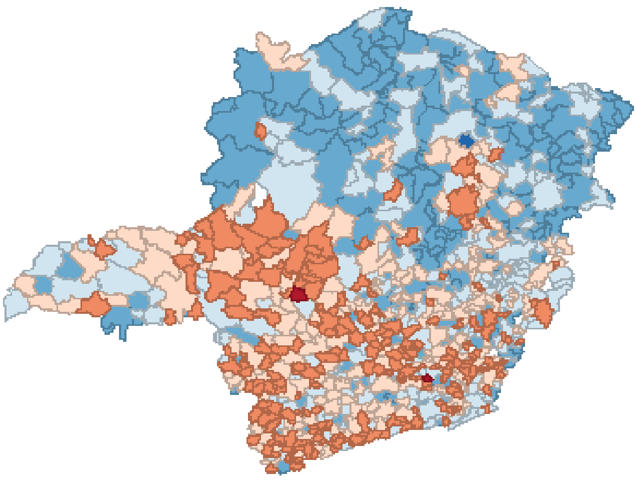
Figure 1 – LISA for Standardized Grades in Elementary Education – Final Years



Source: Own elaboration.

Although Moran's I was significant and positive, there are some regions that deviate from the pattern of high scores with neighbors with high scores and low scores with neighbors with low scores. The regions determined by the colors light blue and light red indicate, respectively, regions that have low scores and neighbors with high scores and regions with high scores with neighbors with low scores. Another noticeable pattern is that there is a north and south division on the map, with a greater concentration of high scores in the lower half of the map and a concentration of low scores in the upper part of the map.

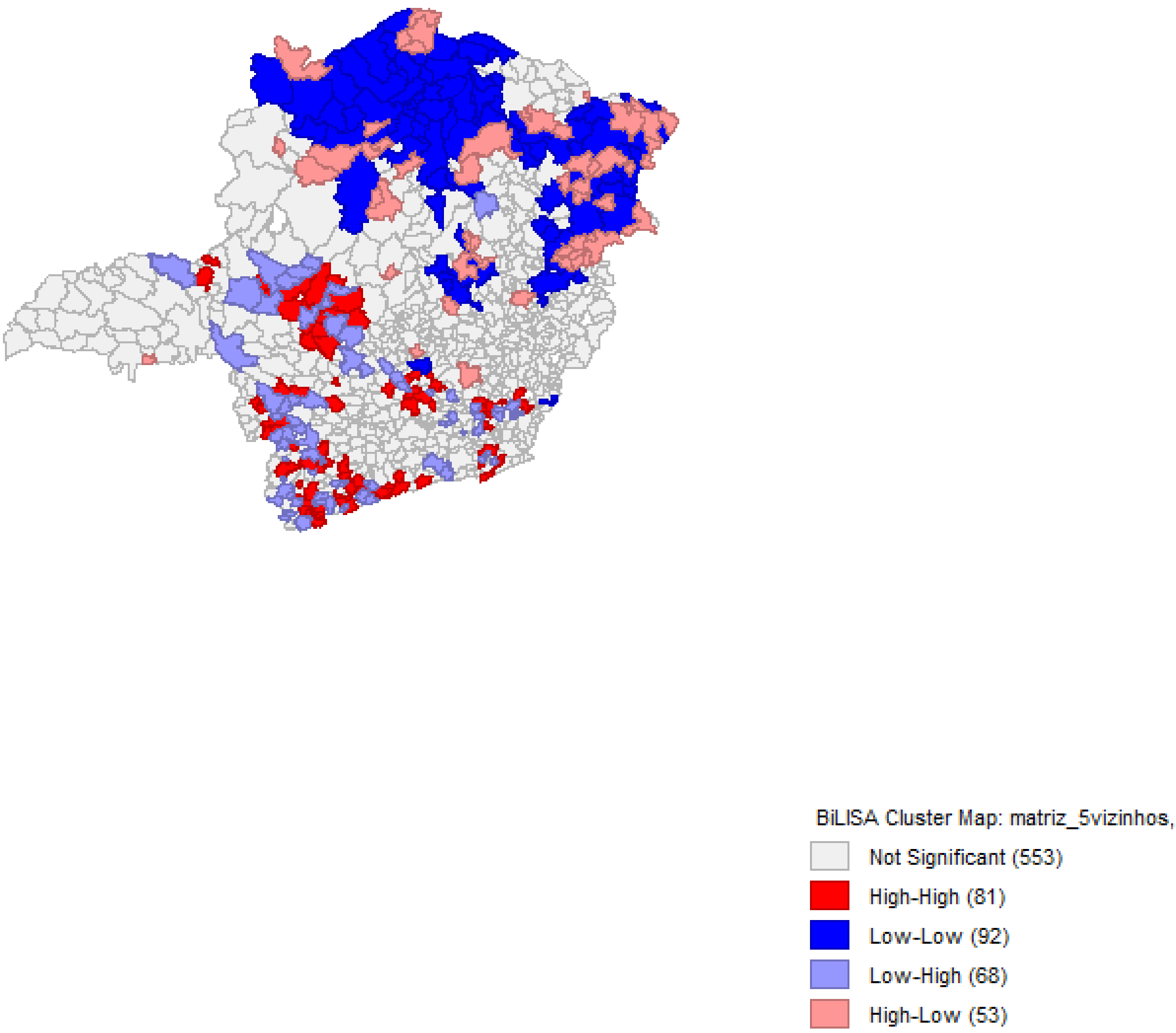
Figure 2 – Box Map using hinge 1.5 for Standardized Grade for Elementary School – Final Years



Source: Own elaboration.

Analyzing the outlier map from Figure 2, the north and south division becomes more discrepant. Using a hinge of 1.5, that is, multiplying the interquartile interval by one and a half times, only one municipality characterized as a lower outlier is identified (in the northern region of Minas Gerais) and 2 municipalities characterized as upper outliers (in the southern region of Minas Gerais). This result was obtained after treating the regions that did not have values for the variable of interest, standardized grade in elementary school final years. In order to verify the existence of spatial autocorrelation between the variable of teacher regularity and school performance, the bivariate cluster map will be analyzed based on the bivariate Moran's I of the variables teacher regularity and standardized grade.

Figure 3 – BLISA: Bivariate Cluster Map for Teacher Regularity and Standardized Grade in Elementary School – Final Years



Source: Own elaboration.

The bivariate cluster map makes it possible to analyze the association between teaching regularity and the standardized Saeb score of neighboring municipalities.

The predominance of clusters was of the low-low type, indicating that cities that have a low teacher regularity index are surrounded by neighbors with lower standardized grades; followed by high-high clusters that indicate that municipalities that have high regularity indexes are surrounded by neighbors with high grades and low-high clusters that indicate municipalities with a low regularity index surrounded by neighbors with high grades and high-low indicate municipalities with high regularity and low grades in their neighbors.

Empirical Strategy

From Exploratory Analysis of Spatial Data and measuring spatial heterogeneity, spatial dependence must be included in the model to be studied. Generally, the starting point of econometric analysis models is the classic linear regression model, estimated by ordinary least squares (OLS). However, Ertur et al. (2006) suggest that spatial dependence or heterogeneity can lead to unreliable OLS estimates, due to the possibility of heteroscedasticity generated by changes in the coefficients or in the error variance between observations.

Thus, to consider and explore the spatial nature of the problem in question we must consider the use of econometric-spatial models. Starting from equation 3:

$$\begin{aligned} y_i &= \alpha + \beta X_i + \rho W_{ij} y_j + \delta_i \\ \delta_i &= \lambda W_{ij} \delta_j + \varepsilon_i \\ \varepsilon &N(0, \sigma^2) \end{aligned} \tag{3}$$

in whichy_i (y_j) is the variable explained in the region *i* (*j*), α is the intercept, X is the matrix nxk of explanatory variables, β is the vector kx1 of coefficients, ρ is the parameter related to the spatial lag of the explained variable, λ is the noise variance parameter, W is the spatial weight matrix nxn, with $Wij > 0$, when the region *j* is a neighbor of the region *i*.

When the values of the parameters ρ and λ are changed, different models are obtained. In the case where there is no spatial dependence, neither in the dependent variable nor in the disturbances (ρ=0 e λ=0), the model would be the traditional MQO model, and could be represented as follows:

$$y_i = \alpha + \beta X_i + \delta_i \tag{4}$$

In cases where ρ≠0 and λ=0, the model to be estimated is the Spatial Autoregressive (SAR). Thus, spatial dependence is included in the model through the spatially lagged values of the dependent variable, as described in the following equation:

$$y_i = \alpha + \beta X_i + \rho W_{ij} y_j + \delta_i \tag{5}$$

Another possible model is the Spatial Error Model (SEM), which reflects the spatial dependence of the residuals (ρ=0 and λ ≠0). This specification indicates that a random shock introduced in a given region affects the others through the spatial structure. The model can be specified as follows:

$$\begin{aligned} y_i &= \alpha + \beta X_i + \delta_i \\ \delta_i &= \lambda W_{ij} \delta_j + \varepsilon_i \\ &\varepsilon N(0, \sigma^2) \end{aligned} \tag{6}$$

According to Anselin (1988), the most appropriate estimation for these models is maximum likelihood or instrumental variables, given that estimations via OLS generate biased and inconsistent results due to the simultaneity in the nature of autocorrelation caused by spatial lag.

In addition to the models usually used in spatial econometrics, there are also the SLX models or Spatial Cross Regressive Model, which only has the weight matrix in the explanatory variables, without a spatial multiplier. Another model is the SDM or Spatial Durbin Model, which can be described according to equation 7:

$$y_i = \alpha + \beta X + WX\tau + \rho W_{ij} y_j + \delta_i \tag{7}$$

The SDM Model is widely used in studies that analyze educational impact due to its characteristic of capturing local reach through WX and global reach through Wy. According to the literature, this would be the most suitable model for analyzing educational spillover.

Applied Model

In order to identify the impact of teacher regularity on student proficiency and the spillover effect of education, a model will be estimated that considers the spatial characteristics of the variables and the main determinants of proficiency to ensure that there will be no omitted variable bias, building reliable estimators. The base, theoretical model will be constructed by the following equation, which will also be analyzed by OLS:

$$edu(Score) = \beta_0 + \beta_1 RegTeacher + \beta_2 Race + \beta_3 Sex + \beta_4 EscMother + \beta_5 AgeGradDist + \beta_6 Income + \beta_7 Hclass + \beta_8 StudentClass + \beta_9 TeachingEffort + \beta_{10} GEduc + \rho i$$

where *edu(Score)* is the standardized Saeb grade in elementary school (final years) in each municipality; *RegTeacher* is the teacher regularity index per municipality; *EscMothe*r is the average education level of students' mothers per municipality; *AgeGradDist* is the age-grade distortion per municipality for the final years of elementary school; *HClass* is the average number of hours of class per municipality for the final years of elementary school; *StudentClass* is the average number of students per class in each municipality for the final years of elementary school; *TeachingEffort* is the average number of schools that have teachers with effort at levels 1, 2 and 3; and, finally, *GEduc* is the average expenditure on educational activities per municipality. The other models with spatial lags will be specified during the development of the results for better understanding and interpretation of the models.

Results

The first model to be tested was the Ordinary Least Squares Model. According to theory, when an OLS model is used to analyze models that should use spatial econometrics, unreliable estimates are generated due to the possibility of heteroscedasticity generated by changes in the coefficients or in the error variance between observations. Table 3 shows the results obtained from the application of the OLS model.

Table 3 – Model estimates using OLS.

Variables	Coefficient
Teacher Regularity	0.0005
Color	1.8194***
Sex	-0.4989**
Mother's Education	0.5091***
Age Grade Distortion	0.0040**
Incme	0.0001*
Average Class Hours	-0.0275
Average Students per Class	-0.0029
Teaching Effort	0.0002
Education Spending	0.0001*
Constant	4.8094***

Source: Own elaboration. Notes: *, **, *** denote p-value less than 10%, 5% and 1%.

The OLS model shows that only the variables denoting socioeconomic characteristics (race, sex, mother's education) and the age-grade distortion, income and education expenditure variables were significant in the model. In addition, the OLS model indicates that male sex is negatively associated with standardized SAEB grades, while the other significant variables in the model are positively associated with standardized elementary school grades. The following classic spatial analysis models will be analyzed: SAR, SEM and SAC models, which respectively consider: the spatial lag of the dependent variable, the spatial lag in the errors and the lag in the dependent variable and in the error, simultaneously.

Table 4 – Results of SAR, SEM and SAC Models

Variables	SAR	SEM	SAC
Teacher Regularity	0.0004	0.0003	0.0002
Color	1.6334***	1.5737***	1.3576***
Sex	-0.5088**	-0.4644*	-0.3890*
Mother's Education	0.7584***	0.7201***	0.7464***
Age Grade Distortion	0.0013	0.0023	0.0041**
Income	0.0000	0.00003	0.0000
Average Class Hours	00432	00483	0.06980
Average Students per Class	-0.0039	-0.00424	-0.0054
Teaching Effort	0.0009	0.0008	0.0008
Education Spending	0.0001**	0.0001**	0.0001**
Constant	4.1276***	4.2180***	3.9578***
P	0.0860***		01035***
Λ		1.0478***	3.0046***
Log likelihood (LIK)	-410.3891	-407.8721	-395.9727
Akaike info criterion (AIC)	421.93	474.29	419.38

Source: Own elaboration. Notes: *, **, *** denote p-value less than 10%, 5% and 1%.

The results in Table 4 indicate that the coefficients of the classical spatial econometric models SAR, SEM and SAC obtained coefficients very close to those of the Ordinary Least Squares model, with teacher regularity being non-significant in all of them. The main highlight is ρ and λ, which are statistically significant in the spatial models, indicating that the students' grades are correlated with the grades of their neighbors and that the autoregressive error parameter in the SEM and SAC models is also statistically significant. The analysis of the direct, indirect and total impacts of the SAR and SAC models is appropriate. In the case of the SEM model, since there is no spatial multiplier, there is no *Wy*, there is no direct impact, only indirect impact, being equal to the total.

Table 5 – Direct, indirect and total impacts of the SAR and SAC models

	SAR Model			SAC Model		
Variables	Direct	Indirect	Total	Direct	Indirect	Total
Teacher Regularity	0.0004	0.0000	0.0004	0.0002	0.0000	0.0002
Color	1.6338***	0.1407***	1.7745***	1.3576***	0.1434***	1.5011***
Sex	-0.5088**	-0,0439**	-0.5527**	-0.3890*	-0.0411	-0.4301*
Mother's Education	0.7585***	0.0653***	0.8638***	0.7464***	0.0789***	0.8253***

Age Grade Distortion	0.0013	0.0001	0.0015	0.0041**	0.0004*	0.0045**
Income	0.000	0.0000	0.0001	0.0000	0.0000	0.0000
Average Class Hours	0.0432	0.003725	0.4697	0.0698	0.0074	0.7718
Average Student/Class	-0.0039	-0.0003	-0.00428	-0.0054	-0.0006	-0.0059
Teaching Effort	0.0009	0.0001	0.0009	0.0008	0.0001	0.0008
Education Spending	0.0001**	0.00001**	0.0002**	0.0001**	0.0001*	0.0001**

Source: Prepared by the authors. Notes: *, **, *** denote p-values less than 10%, 5% and 1%.

The spillover effect is given by the indirect impact that indicates in both models that the variables that control the socioeconomic characteristics of the students have a significant direct, indirect and total impact (race, sex, mother's education), it is noted that the predominance of the male sex is associated with lower grades, which is already expected according to the literature.

Furthermore, only in the SAC model did the age-grade distortion variable present significant impacts, with only the direct and total ones being significant at 5%. The only variable introduced in the classical performance determinant model that presented significant impacts in both models described in Table 5 was spending on education, presenting significant direct, indirect and total impacts.

The following models will be analyzed, which are the most suitable for addressing this type of issue in the literature: SDM and SLX models. Both models introduce spatial lag in the explanatory variables, with the SLX model not having Wy, and therefore not having a spatial multiplier, being a simpler model, and the SDM model having a spatial multiplier and lag in the explanatory variables. In the case of the performance determinants model in question, the variables teaching effort and sex were not spatially lagged because they do not present spatial autocorrelation.

Table 6 – Results of the SDM and SLX Models

Variables	SDM	SLX
Teacher Regularity	0.0004	0.0004
Skin Color	1.3748***	1.3994***
Sex	-0.4239*	-0.4454*
Mother's Education	0.6984***	0.6973***
Age Grade Distortion	0.0040*	0.0036*
Income	0.0000	0.0000
Average Class Hours	0.0672	0.0501
Average Students per Class	-0.0054	-0.0055
Teaching Effort	0.0005	0.0004
Education Spending	0.0001**	0.0001**
Constant	4.0802***	4.1087***
WRegularity	0.0086	0.0107
WColor	-1.3630	-0.1505
WMother's Education	-7.7446*	-7.8138**
WAge Grade Distortion	-0.1093***	-0.1085***
WIncome	0.0018	0.0021
WAverage Class Hours	0.6308	1.4917**

WAverage Students per Class	-0.1274*	-0.1371*
WEducation Spending	0.0010	0.0013
ρ	0.8045***	
Log likelihood (LIK)	-397.0511	-400.8238
Akaike info criterion (AIC)	476.23	451.04

Source: Own elaboration. Notes: *, **, *** denote p-value less than 10%, 5% and 1%.

Analyzing the results of Table 6, SDM and SLX models, the coefficients found are very close to those of the other models, however, as these models consider the spatial lag of the explanatory variables, it is possible to note that the variable WAge-grade distortion is significant in both models at 1%, the variable WAverage Student per Class is also significant in both models, in addition, in the SDM model ρ remains significant at 1%. Below is the table with the impacts of the SDM and SLX models (the impact of the SLX model being composed of direct impact β, indirect impact τ and the total equal to β + τ).

Table 7 – Direct, indirect and total impacts of the SDM and SLX models

	SDM Model			SLX Model		
Variables	Direct	Indirect	Total	Direct	Indirect	Total
Teacher Regularity	0.0004	0.0411	0.0415	0.0003	0.0099	0.0100
Color	1.3732***	-1.1768	0.1964	1.3994***	-0.1380	1.2614*
Sex	-0.4261*	-1.5619	-1.9880	-0.4454*	0.0000	-0.4454*
Mother's Education	0.6526***	-32.8923	-32.2396	06973***	-7.1676*	-6.4703*
Age Grade Distortion	0.0033	-0.4857	-0.4823	0.0036*	-0.0995***	-0.0959***
Income	0.0000	0.0083	0.0083	0.0000	0.0019	0.0019
Average Class Hours	0.0716	3.1365	3.2080	0.0501	1.3684**	1.4185**
Average Student/Class	-0.0062*	-0.6032	-0.6094	-0.0055	-0.1258*	-0.1313*
Teaching Effort	0.0005	0.0018	0.0023	0.0004	0.0000	0.0004
Education Spending	0.0001**	0.0049	0.0050	0.0001**	0.0012	0.0014

Source: Own elaboration. Notes: *, **, *** denote p-value less than 10%, 5% and 1%.

The SDM and SLX models indicated significant direct impact of the variables that make up the socioeconomic characteristics and only direct impact of spending on education. In the SLX model, the variables: average class hour and average student class had a significant indirect and total impact, as well as age-grade distortion.

Using the AIC criterion to choose the model, we can indicate that the models that presented the best fit were respectively: SAC, SLX and SDM.

Conclusions

The educational determinants have been the focus of studies for many years in economic theory. School performance is used as a proxy for teaching quality. It is widely known in the literature on the economics of education that the socioeconomic characteristics of students have a great influence on performance, especially the mother's education, race and gender of the student.

However, when it comes to spatial econometric analysis, there are still few studies that address educational determinants. The big issue in observing the determinants of performance through traditional spatial econometric models is the problem of the ecological fallacy in which information from the whole is taken to the individual level, losing the nuances involved in microeconometric analysis.

With the development of models that use georeferencing, this problem can be solved, making it possible to analyze at an individual level and considering the space and spatial distribution of data in the analysis of individual data. However, Brazilian databases still do not allow this type of application, despite constant efforts and the reality that is increasingly closer with the development of georeferenced bases in large study centers.

Returning to the question of the article, the results indicated that teacher regularity was not significant in any of the models analyzed. Despite this, all models presented results compatible with those indicated in the literature, with a great influence of maternal education on school performance, sex and race. In addition, some other external variables (which are not in the traditional models of determinants of education) presented a significant effect when the analysis was based on the spatial model.

First, the standardized grade showed spatial autocorrelation, with ρ, a parameter that is linked to the spatial multiplier, being significant in all models that were estimated. In addition, spending on education activities was a significant variable in all models, indicating that government spending influences the academic performance of students in public schools in Minas Gerais. The age-grade distortion was also significant in the models that presented better adjustments based on the AIC criterion, with emphasis on the great significance of the variable WAge-grade distortion.

From the results it is possible to state that characteristics of the teaching staff such as effort and regularity in spatial models did not influence student performance, the use of hierarchical models may be a better suggestion for the analysis of different levels, enabling better analysis of how teachers, schools and classes impact student grades.

References

ALBERNAZ, Ângela; FERREIRA, Francisco HG; FRANCO, Creso. **Qualidade e**

equidade na educação fundamental brasileira. Texto para discussão, 2002.

ALMEIDA, Eduardo. **Econometria Espacial Aplicada**. Campinas: Alínea, 2012.

ALVES, M. T. G.; SOARES, J. F. Efeito-escola e estratificação escolar: o impacto da composição de turmas por nível de habilidade dos alunos. **Educação em Revista**, v. 45, p. 25-28, jun. 2007.

BARRO, R. J. e Xavier SALA-I-MARTIN. Technological Diffusion, Convergence, and Growth. **Journal of Economic Growth**, vol. 2, pp.1-26, 1997.

BARROS, R. P; MENDONÇA, R. Investimentos em Educação e Desenvolvimento Econômico. Rio de Janeiro, **IPEA**, 1997. (Texto para Discussão No. 525)

BIONDI, Roberta L.; FELÍCIO, Fabiana. Atributos escolares e o desempenho dos estudantes: uma análise de painel dos dados do SAEB. Brasília: **INEP**, 2007.

BIONDI, Roberta L.; VASCONCELLOS, Lígia; MENEZES-FILHO, Naércio A. Avaliando o Impacto da Olimpíada Brasileira de Matemática das Escolas Públicas no desempenho de matemática nas avaliações educacionais. In: 31º Encontro da Sociedade Brasileira de Econometria, 2009, Foz do Iguaçu. **Encontro Brasileiro de Econometria - SBE**, 2009.

BIONDI, Roberta L.; VASCONCELLOS, Lígia; MENEZES-FILHO, Naércio A. Avaliando o impacto do programa Escrevendo o Futuro no desempenho das escolas públicas na Prova Brasil e na renda futura dos estudantes. Anais do **XXXVII Encontro Nacional de Economia da ANPEC**, Foz do Iguaçu, Paraná, 2009.

BISHOP, J. Is the test score decline responsible for the productivity growth decline? **American Economics Review**. Vol. 79 (1), p. 178-97, 1989.

CÉSAR, C. C.; SOARES, J. Desigualdades acadêmicas induzidas pelo contexto escolar. **Revista Brasileira de Estudos de População**, v. 18, n. 1/2, p. 97-110, 2001.

COLEMAN, J. S. et al. Equality of educational Opportunity. U.S**. Government printing office**. Washington, DC, 1966.

ERTUR, C; KOCH, W. Growth, technological interdependence and spatial externalities: theory and evidence. **Jounal of Applied Econometrics**, v.22, issue 6, 2007.

FELÍCIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no Estado de São Paulo. Anais do **XXXIII Encontro Nacional de Economia**, Natal/RN, Brasil. ANPEC, 2005.

FUJITA, M.; KRUGMAN, P. & VENABLES, A. The spatial economy. Cambridge, **MIT**, 1999.

HANUSHEK, E. A.; KIMKO, D. D. Schooling, labor-force quality and the growth of nations? **The American Economic Review**, 90(5):1184–1208.2000

HOLLAND, P. W. Statistics and casual inference. **Journal of the American Statistical Association**, v. 81, n. 306, p. 945-960, 1986.

HEWINGS, G., MAGALHAES, A., AZZONI, C. R. Spatial dependence and regional inequality in Brazil. **Investigaciones Regionales**, Madrid, v. 6, p. 5-20, 2005.

MACHADO, Ana Flávia et al. Qualidade do ensino em matemática: determinantes do desempenho dos alunos em escolas públicas estaduais mineiras. **Revista Economia**, v. 9, n. 01, p. 23-45, jan/abr., 2008

MENEZES-FILHO, Naércio. A Evolução da Educação no Brasil e seu Impacto no Mercado de Trabalho. São Paulo, **FEA/USP**, 2001.

MENEZES-FILHO, Naércio. Os determinantes do desempenho escolar no Brasil. Instituto Futuro Brasil, **IBMEC São Paulo e Faculdade de Economia e Administração**

da Universidade de São Paulo. Sumário Executivo, 2007.

MORENO, R., TREHAN, B. Location and the Growth of Nations, **Journal of Economic Growth**, 2, pp.399-418, 1997.

MORETTI, E. Estimating the Social Return to Higher Education: Evidence from Longitudinal and Repeated Cross-sectional Data. **Journal of Econometrics**, n.121, p.175-212, 2004.

MURNANE, R. J., WILLETT, J. B. AND LEVY, F. The growing importance of cognitive skill in wage determination. **Review of Economics and Statistics**, vol. 77 (2) p. 251-66, Maio, 1995.

MURPHY, K. M. AND PELTZMAN, S. School Performance and the Youth Labor Market. **Journal of Labor Economics**, vol.22 (2), pp.299-325, The University of Chicago. 2004.

PALERMO, G. A.; DO NASCIMENTO SILVA, D. B.; NOVELLINO, M. S. F. Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro. **Revista Brasileira de Estudos de População**, v. 31, n. 2, p. 367-394, 2014.

RAMOS, R., SURIÑACH, J., ARTÍS, M. Human capital spillovers, productivity and regional convergence in Spain. **Papers in Regional Science**, Volume 89 n2 June 2010.

RIANI, J. L. R.; RIOS-NETO, E. L. G. Background familiar versus perfil escolar do município: qual possui maior impacto no resultado educacional dos alunos brasileiros? **Revista Brasileira de Estudos de População**, v. 25, n. 2, p. 251-269, jul./dez. 2008.

RIVKIN, S. G. Black/white differences in schooling and employment. **Journal of Human Resources**, vol.30 (4), p. 826-52, 1995.

RIVKIN, Steven G.; HANUSHEK, Eric A.; e KAIN, John F. Teachers, Schools, and Academic. **Achievement. Econometrica**. vol. 73, No.2, p417-58. 2005

ROSENTHAL S., STRANGE W. The attenuation of human capital externalities. **Journal of Urban Economics** 64: 373–389, 2008.

SILVEIRA NETO, R. M. Crescimento e spillovers: a localização importa? Evidências para os estados brasileiros. **Revista Econômica do Nordeste**, Fortaleza, v. 32, p. 524-545, 2001.

SOARES, Tufi M.Modelo de três níveis hierárquicos para a proficiência dos alunos de 4ª série avaliados no teste de língua portuguesa do SIMAVE/PROEB 2002. **Revista Brasileira de Educação**, n.29, p.73-87. 2005

SOUZA, M. R. P. Análise da variável escolaridade como fator determinante do crescimento econômico. **FAE**, Curitiba, Vol. 2, No. 3, p.47-56. 1999

VERNIER, Laura. D. S.; BAGOLIN, Izete. P.; JACINTO, Paulo. A. Fatores que Influenciam o Desempenho Escolar no Estado do Rio Grande do Sul: Uma Análise com Regressões Quantílicas. **Análise Econômica**, v. 33, n.64, 2015.

VERNIER, Laura Desirée Silva; BAGOLIN, Izete Pengo; FOCHEZATTO, Adelar. Distribuição e disseminação espacial da educação nos municípios brasileiros. **45º Encontro Nacional de Economia**, p. 18, 2017.