

# TRANSBORDAMENTOS ESPACIAIS DA EDUCAÇÃO NOS MUNICÍPIOS BRASILEIROS

Gabriella da Silva Cavalcanti<sup>1</sup>  
Diego de Maria André<sup>2</sup>  
Júlia Rocha Araújo<sup>3</sup>

**RESUMO:** O objetivo do estudo consiste em investigar o transbordamento da qualidade da educação entre os municípios via proficiência média em português e matemática para alunos dos anos finais do ensino fundamental por meio dos modelos econométrico-espaciais. Foram utilizados microdados do SAEB, dados do Observatório da Criança e do Adolescente, do IBGE e do portal FINBRA, sendo o período de análise os anos de 2011 e 2015 para os 5570 municípios brasileiros. Inicialmente, foi realizada a Análise Exploratória de Dados Espaciais e observada a predominância de *clusters* espaciais, indicando a presença de autocorrelação espacial positiva no desempenho escolar. Os resultados obtidos a partir do modelo SDM indicam que o desempenho escolar de um município está condicionado ao desempenho escolar prévio do próprio município, ao desempenho escolar dos municípios vizinhos e às características socioeconômicas dos municípios e dos seus vizinhos.

**Palavras-Chave:** Externalidades Educacionais, Econometria Espacial, Desempenho Escolar.

**ABSTRACT:** The objective of this study is to investigate the overflow of the quality of education among municipalities through average proficiency in Portuguese and mathematics for students of the final years of elementary school by of econometric spatial models. Were are used microdata of SAEB and data from the Child and Adolescent Observatory, of the IBGE and of the FINBRA Portal, the period of analysis is the years 2011 and 2015 for the 5570 Brazilian municipalities. Initially, the Exploratory Analysis of Spatial Data was performed and the predominance of spatial clusters was observed, indicating the presence of positive spatial autocorrelation in school performance. The results obtained from the SDM model indicate that the school performance of a municipality is conditioned to previous school performance of the municipality itself, on the school performance of neighboring municipalities and the socioeconomic characteristics of the municipalities and their neighbors.

**Keywords:** Educational Externalities, Spatial Econometrics, School Performance.

**Classificação JEL:** C21, I21.

**Área 2:** Econometria espacial.

---

<sup>1</sup>Mestre em Economia Regional, Programa de Pós-Graduação em Economia, Universidade Federal do Rio Grande do Norte (PPECO/UFRN). E-mail: <gabriella.s.cavalcanti@hotmail.com>

<sup>2</sup>Professor Doutor do Programa de Pós-Graduação em Economia e do Departamento de Economia da Universidade Federal do Rio Grande do Norte (UFRN). E-mail: <diegomandre@gmail.com>

<sup>3</sup>Professora Doutora do Departamento de Economia da Universidade Federal do Rio Grande do Norte (UFRN). E-mail: <araujorjulia@gmail.com>

# 1 INTRODUÇÃO

Após a fase de ampliação do acesso à educação pública vivenciada no Brasil a partir da década de 1990, a qual praticamente universalizou o ensino fundamental e promoveu o crescimento da escolaridade média da mão de obra, o desafio passou a ser a melhoria da qualidade dessa educação (MENEZES-FILHO, 2007).

A promoção da melhoria da qualidade da educação requer a compreensão sobre o verdadeiro impacto das variáveis extra e intraescolares sobre a proficiência dos alunos para o estabelecimento de regras que tornem a alocação de recursos eficientes nessa área. Nesse sentido, uma agenda cujo objetivo seja a melhoria da qualidade da educação pública deve buscar entender o papel das variáveis que circundam a vida do aluno, que vão desde os aspectos individuais e da sua família, passando pelos insumos escolares até a influência do espaço.

Sabe-se que na condição de bem-público não-exclusivo e não-rival, a educação gera externalidades positivas advindas do seu retorno individual e social. Estudos têm encontrado evidências da existência de externalidades positivas do aumento do capital humano nos municípios brasileiros (ARAÚJO e SILVEIRA NETO, 2004; FALCÃO e SILVEIRA NETO, 2007; QUEIROZ, 2003; MORETTI, 2002). Tais externalidades tenderiam a afetar em maior grau os indivíduos com maiores níveis educacionais e as diferenças de produtividade entre os agentes estariam vinculadas a vantagens produtivas geradas pela presença de um maior estoque de capital humano nestas localidades (FALCÃO E SILVEIRA NETO, 2007).

Queiroz (2003) encontrou evidências dos efeitos sociais positivos do capital humano sobre o diferencial regional de salários em Minas Gerais. Nessa mesma linha, Araújo e Silveira Neto (2004) apresentam evidências de que as diferenças na disponibilidade do estoque de capital humano nas regiões geram uma relação positiva entre a concentração de capital humano e a produtividade dos trabalhadores nas grandes cidades, sendo importantes para explicar os diferenciais regionais de renda no Brasil. Ainda segundo Moretti (2002), altos níveis locais de capital humano tem a possibilidade de elevar a produtividade agregada por meio do efeito direto que o capital humano individual pode causar na redução da criminalidade, dos gastos com saúde e tornar a população mais consciente e participativa politicamente.

Na ciência regional, ficou a cargo da Nova Geografia Econômica atribuir a localização um papel importante antes renegado e deixado em segundo plano pelas demais áreas de pesquisa, dando-a importância crucial a medida que, quanto mais próximas as regiões, maiores as externalidades advindas das trocas de informações e conhecimentos (ALMEIDA, 2012; FUJITA, KRUGMAN e VENABLES, 1999; VERNIER, BAGOLIN e FOCHEZATTO, 2017).

O Brasil é caracterizado por possuir significativas diferenças territoriais e regionais, sendo que o Norte e Nordeste são as regiões que apresentam os piores indicadores socioeconômicos do país. Estudos empíricos sugerem que parte dessa desigualdade é explicada pelas disparidades educacionais (MENEZES-FILHO, 2001; MENEZES-FILHO et al., 2006; RODRIGUES et al., 2017; SOUZA et al., 2013). Portanto, é importante entender como a qualidade da educação está espacialmente distribuída, já que disso dependerá o nível e a distribuição dos retornos educacionais de um país.

A maioria dos estudos nessa área abordam os transbordamentos dos gastos em educação e a influência exercida pela vizinhança sobre tais variáveis (CARNEIRO, 2014; FOWLES e TANDBERG, 2018). Todavia, poucos pesquisadores se empenharam em entender os transbordamentos da qualidade da educação (LOBÃO e DA SILVA, 2018; RODRIGUES et al., 2017; VERNIER, 2016; VERNIER, BAGOLIN e FOCHEZATTO, 2017). Nesse sentido, este estudo pretende contribuir com o debate acerca da importância da consideração do espaço na determinação da qualidade da educação ofertada por meio do reconhecimento dos limites ge-

ográficos no estudo das externalidades educacionais.

O transbordamento da qualidade da educação será analisado a partir da proficiência média em português e matemática para alunos dos anos finais do ensino fundamental por meio dos modelos econométrico-espaciais.

Para execução do exercício econométrico utilizou-se os dados do SAEB referentes ao ano de 2015 para os municípios brasileiros. Para controlar por fatores não observáveis, a nota do SAEB de 2011 do 5º ano do ensino fundamental foi utilizada como controle.

Além desta introdução, este trabalho está dividido em mais três seções. A seção 2, na sequência, descreve os procedimentos metodológicos adotados e a base de dados utilizada. Na seção 3 são apresentados e discutidos os resultados encontrados e, finalmente, na seção 4, são feitas as considerações finais.

## 2 PROCEDIMENTOS METODOLÓGICOS

### 2.1 Análise Exploratória de Dados Espaciais

Similarmente à econometria convencional, na econometria espacial é preciso realizar previamente uma análise exploratória dos dados antes de estimar os modelos de regressão. Segundo Anselin (1988), a Análise Exploratória de Dados Espaciais (AEDE) consiste em um conjunto de procedimentos técnicos que servem para descrever e observar distribuições espaciais, identificar pontos extremantes e padrões de associação no espaço denominados *outliers* e *clusters* espaciais, respectivamente. De acordo com Almeida (2012), a AEDE auxilia na etapa posterior de especificação dos modelos espaciais.

O primeiro passo de uma AEDE é testar a existência de autocorrelação espacial, isto é, se os dados são aleatoriamente distribuídos ou não no espaço. No entanto, para isso, é necessário, anteriormente, construir uma matriz de ponderação espacial ( $W$ ) que retrate o arranjo espacial das interações geradas a partir do fenômeno analisado. Os pesos espaciais ( $w_{ij}$ ) representam o grau de influência de uma região  $i$  sobre a região  $j$  segundo algum critério de proximidade.

Na literatura, o critério de proximidade geográfica definido de acordo com a contiguidade e a distância são os mais adotados (RÊGO e PENA, 2012). A matriz binária de contiguidade segundo as convenções rainha ou torre possuem a desvantagem de não assegurarem uma conexão balanceada entre as regiões, pois pode haver regiões com muitos vizinhos e outras com poucos. Já a vantagem delas consiste em permitir a clara definição de contiguidades de ordens superiores. Já em relação às matrizes baseadas na distância geográfica, a grande vantagem da adoção da convenção da matriz de  $K$ -vizinhos mais próximos é que ela garante a inexistência de “ilhas”, isto é, regiões que não possuam vizinhos, dado que haverá equilíbrio no número de vizinhos para todas as regiões (ALMEIDA, 2012).

Após a definição da matriz de pesos espaciais a ser adotada, o próximo passo consiste na AEDE por meio da estatística de teste do  $I$  de Moran. O  $I$  de Moran consiste em um coeficiente de autocorrelação espacial do tipo produto cruzado que testa a hipótese nula de distribuição aleatória dos dados espaciais permitindo verificar e fazer inferência acerca da presença de dependência espacial. O  $I$  de Moran global pode ser definido como na Equação 1 a seguir:

$$I = \frac{n}{S_0} \frac{z'Wz}{z'z'} \quad (1)$$

em que  $n$  é o número de regiões,  $z$  representa os valores da variável padronizada,  $W$  é a matriz de ponderação espacial e  $S_0$  é um escalar resultante do somatório dos elementos da matriz de

pesos espaciais  $W$ . O valor do coeficiente varia entre -1 e 1, quanto mais próximo de -1 maior a dispersão dos dados sinalizando presença de autocorrelação negativa e, quanto mais próximo de 1, maior a concentração dos dados através das regiões indicando a presença de autocorrelação positiva (ALMEIDA, 2012).

Além da medida global de associação linear global, Anselin (1988) propôs o *Local Indicator of Spatial Association* (LISA) que permite identificar padrões de autocorrelação espacial local, podendo ser calculado segundo a Equação 2 a seguir:

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

onde  $n$  trata-se do número de regiões,  $w_{ij}$  consistem nos elementos da matriz de ponderação espacial,  $y_i$  e  $y_j$  são os valores da variável de interesse e os subscritos  $i$  e  $j$  são as diferentes regiões.

Combinando as informações do diagrama de dispersão de Moran e do mapa de significância LISA, pode-se extrair informação mais simples e interessante do mapa de *clusters* LISA por meio da análise dos seus quadrantes que representam quatro tipos de associação espacial linear, sendo eles: Alto-Alto (AA), Baixo-Baixo (BB), Alto-Baixo (AB), Baixo-Alto (BA). Os quadrantes AA e BB denotam autocorrelação positiva (*clusters*), significando que regiões com valores altos (baixos) da variável considerada estão próximos de regiões que também apresentam valores altos (baixos). Em oposição a esse padrão, os quadrantes AB e BA sugerem autocorrelação negativa (*outliers*), indicando que regiões com valores altos (baixos) da variável considerada estão próximos de regiões que apresentam valores baixos (altos).

## 2.2 Modelos Espaciais

Após detectada e mensurada a autocorrelação espacial, cabe incorporar a dependência espacial no modelo econométrico convencional. Parte-se do processo a-espacial representado pelo modelo de regressão linear clássico e incorpora-se as defasagens espaciais  $Wy$  (na variável dependente),  $WX$  (nas variáveis independentes),  $W\xi$  e/ou  $W\varepsilon$  (nos termos de erro) visando controlar a dependência espacial (ALMEIDA, 2012).

Parte-se do modelo de dependência espacial geral (SGM) apresentado na Equação 3 a seguir:

$$y = \rho W_1 y + X\beta + W_1 X\tau + \varepsilon \quad (3)$$

$$\varepsilon = \lambda W_2 \varepsilon + \epsilon \quad (4)$$

ou

$$\varepsilon = \gamma W_2 \epsilon + \epsilon \quad (5)$$

$$\epsilon \sim Normal(0, \sigma^2 I_n) \quad (6)$$

em que  $y$  é a variável explicada na região  $i$ ,  $X$  é a matriz  $n \times k$  de variáveis explicativas,  $\beta$  é o vetor  $k \times 1$  de coeficientes,  $\rho$  é o parâmetro relacionado com a defasagem espacial da variável dependente,  $\tau$  é o parâmetro da defasagem espacial das variáveis explicativas,  $\lambda$  o parâmetro de variância do termo de erro,  $\gamma$  o coeficiente de média móvel espacial,  $W$  é a matriz de pesos espaciais  $n \times n$ , com  $W_{ij} > 0$ , quando a região  $j$  é vizinha da região  $i$ .

Ao impor restrições sobre os parâmetros de da Equação 3, pode-se derivar todos os modelos espaciais conhecidos. Os modelos espaciais estimados foram o SLX, SDEM e SDM, caracterizados por serem capazes de controlar a influência das características das observações

vizinhas sobre as demais observações, isto é, os transbordamentos espaciais entre as regiões contíguas. Conforme Almeida (2012) e Le Sage e Pace (2009), no caso em que  $\rho=\tau=\lambda=\gamma=0$ , isto é, não existindo dependência espacial na variável dependente, nas explicativas nem no termo de erro, o modelo seria o convencional Modelo Clássico de Regressão Linear (MCRL), como representado na Equação 7 abaixo:

$$y = X\beta + \epsilon \quad (7)$$

Se impostas as restrições  $\rho=\lambda=\gamma=0$  e  $\tau \neq 0$ , obtêm-se o Modelo Regressivo Cruzado Espacial (SLX). O modelo SLX consiste num modelo de defasagem espacial de alcance local nas variáveis explicativas, isso quer dizer que algumas ou todas as variáveis explicativas da região  $i$  influenciam a variável dependente da região vizinha  $j$  e vice-versa. A motivação da modelagem deriva das externalidades positivas ou negativas geradas por fatores ou características das regiões vizinhas (LESAGE e PACE, 2009). O modelo SLX pode ser representado como na Equação 8 abaixo:

$$y = X\beta + WX\tau + \epsilon \quad (8)$$

Se considerarmos  $\lambda=\gamma = 0$  e  $\rho \neq 0$  e  $\tau \neq 0$ , têm-se o *Spatial Durbin Model* (SDM). Segundo LeSage e Pace (2009), além do mesmo intuito do modelo SAR de comportar a dependência espacial na variável dependente, a motivação adicional para consideração do modelo SDM é a omissão de variáveis relevantes correlacionadas com alguma(s) variável(is) explicativa(s), sendo essa(s) autocorrelacionada(s) espacialmente, uma situação muito recorrente em econometria, principalmente quando se trata de estudos empíricos relacionados à variáveis de educação. Podendo este ser representado como na Equação 9 a seguir:

$$y = \rho Wy + X\beta + WX\tau + \epsilon \quad (9)$$

No caso de  $\rho = \gamma = 0$  e  $\lambda \neq 0$  e  $\tau \neq 0$ , têm-se o *Spatial Durbin Error Model* (SDEM). No modelo SDEM a dependência espacial é de alcance local nas variáveis explicativas (X) e de alcance global no termo de erro ( $\epsilon$ ). Os transbordamentos locais das variáveis explicativas exógenas nas regiões vizinhas e a existência de efeitos não-modelados de alcance global implícitos no termo de erro seriam as motivações principais do modelo. A representação formal do modelo SDEM é apresentada na Equação 10 a seguir:

$$y = X\beta + WX\tau + \epsilon \quad (10)$$

$$\epsilon = \lambda W\epsilon + \epsilon \quad (11)$$

no qual  $\lambda$  é o coeficiente do erro autorregressivo da defasagem espacial  $W\epsilon$ .

Especificado corretamente o modelo econométrico espacial, cabe encontrar o melhor método de estimação. Segundo Almeida (2012), na econometria espacial, a maioria dos trabalhos empíricos estimam os modelos por meio do Método da Máxima Verossimilhança (MV), do Método de Variáveis Instrumentais (VI) e do Método de Momentos Generalizados (GMM) visto que, a estimação por MQO produz estimadores viesados e/ou ineficientes, levando a interpretações enganosas. É importante salientar que o emprego da MV requer a hipótese de normalidade do termo de erro, enquanto os métodos de VI e GMM prescindem desse pressuposto.

Partindo de uma função geral de produção da educação para estimar o desempenho educacional por Mínimos Quadrados Ordinários (MQO) como na Equação 12 a seguir:

$$SAEB_{2015m} = \beta_0 + \beta_1 SAEB_{2011m} + \beta_2 Alu_m + \beta_3 Fam_m + \beta_4 Doc_m + \beta_5 Eco_m + \varepsilon \quad (12)$$

onde  $SAEB_{2015}$  é o desempenho escolar medido pela proficiência em português e matemática nos anos finais do ensino fundamental de cada município  $m$ ;  $SAEB_{2011}$  é o desempenho escolar prévio de cada município  $m$  medido pela proficiência nas mesmas disciplinas dos alunos dos anos iniciais do ensino fundamental em 2011, tal como visto em Hanushek (2002);  $Alu$  é o vetor de características dos alunos agregadas por município;  $Fam$  é o vetor de características do *background* familiar agregadas por município;  $Doc$  é o vetor de características dos docentes do município;  $Eco$  é o vetor de variáveis que caracterizam a economia e as despesas com educação dos municípios;  $\beta$  é o vetor coluna dos coeficientes dos estimadores de MQO e  $\varepsilon$  é o vetor coluna dos resíduos da regressão.

Com o intuito de especificar corretamente o modelo econométrico-espacial, foram realizados testes difusos ou focados. Os testes difusos são aqueles em que testa-se apenas a hipótese nula de independência espacial dos resíduos e não se conclue nada a respeito do tipo de autocorrelação espacial presente no modelo. No presente trabalho, foi calculado o teste difuso do  $I$  de Moran sugerido por Cliff e Ord (1973). No que se refere aos testes focados, eles possuem a propriedade de indicar a especificação da autocorrelação espacial testada na hipótese alternativa, pois esta trata-se de um modelo espacial específico. Existe um conjunto de testes focados, a maioria do tipo multiplicador de Lagrange, em que a hipótese nula é o modelo restrito (MQO) e a hipótese alternativa algum modelo espacial irrestrito.

Visando definir o modelo espacial mais robusto dentre os estimados, são utilizados o indicador do valor da Função de Verossimilhança (LIK), os Critérios de Informação de Akaike (AIC), Schwarz (BIC) e o teste da Razão de Verossimilhança (*LR test*). Segundo os testes AIC e BIC deve ser escolhido o modelo que apresentar o menor valor da estatística para os referidos testes, o que ocorre quando o valor da estatística LIK é a maior dentre as analisadas. Já o LR test é calculado a partir da comparação do modelo mais restrito, que consiste na hipótese nula, com modelos irrestritos, que constituem a hipótese alternativa.

O último passo da análise empreendida nesse estudo consiste nos testes de especificação da dependência espacial para escolha do melhor modelo econométrico-espacial e, finalmente, a estimação do efeito marginal total (efeitos diretos e indiretos) de uma mudança na variável explicativa de uma região sobre as regiões vizinhas. De acordo com LeSage e Pace (2009), nos modelos espaciais o efeito marginal de  $Y$  em relação à  $X$  não é igual ao  $\beta$  do MQO, visto que, uma mudança em  $X$  de uma região tem potencial de afetar não apenas a própria variável  $Y$  da região diretamente, mas também a variável  $Y$  das regiões vizinhas indiretamente devido ao multiplicador espacial.

### 2.3 Base de dados e estatísticas descritivas

Os dados utilizados neste estudo foram provenientes de fonte secundária. Foram utilizados os microdados do Sistema de Avaliação da Educação Básica (SAEB) e dados do Observatório da Criança e do Adolescente divulgados pelo Instituto Nacional de Estudos e Pesquisas Anísio Teixeira (INEP), do Instituto Brasileiro de Geografia e Estatística (IBGE) e do portal Finanças Brasil (FINBRA), sendo o período de análise os anos de 2011 e 2015 para os 5570 municípios brasileiros. A descrição, a fonte e a periodicidade das variáveis são apresentadas no Quadro 1 a seguir:

Quadro 1: Fonte dos dados e variáveis utilizadas

Variável	Descrição	Fonte	Ano
Desempenho escolar atual	Proficiência média em português e matemática dos alunos do 9º ano do EF	SAEB	2015
Desempenho escolar prévio	Proficiência média em português e matemática dos alunos do 5º ano do EF	SAEB	2011
Características socioeconômicas	Prop. de indivíduos da cor branca	SAEB	2015
	Prop. de indivíduos do sexo masculino	SAEB	2015
Características dos docentes	Prop. de professores pós-graduados	SAEB	2015
	Prop. de professores que corrigem deveres de casa	SAEB	2015
	Prop. de docentes com formação superior de licenciatura (ou bacharelado com complementação pedagógica) na mesma área da disciplina que leciona	Observatório da Criança e do Adolescente	2015
<i>Background</i> familiar	Prop. de mães com pelo menos ensino médio completo	SAEB	2015
	Prop. de pais com pelo menos ensino médio completo	SAEB	2015
	Prop. de filhos que moram com o pai	SAEB	2015
	Prof. de alunos que possuem computador doméstico	SAEB	2015
	Prop. de alunos que trabalham fora de casa	SAEB	2015
	Prop. de alunos que residem com mais de 4 pessoas (incluindo ele próprio)	SAEB	2015
	Prop. de pais que incentivam o estudo dos filhos	SAEB	2015
Renda Municipal	PIB municipal per capita	IBGE	2014
Gastos com educação	Despesas educacionais municipais per capita no ensino fundamental em valores correntes (R\$)	FINBRA	2015

Fonte: Elaboração própria.

O SAEB avalia o desempenho do alunado nas disciplinas básicas de português e matemática e consiste em uma avaliação triparte, visto que além dos alunos, professores e diretores respondem os questionários aplicados.

O desempenho escolar dos municípios, considerado a variável dependente deste estudo, foi obtido através das proficiências médias nas disciplinas de português e matemática. A consideração dessas duas disciplinas se justifica pelo fato de constituírem as disciplinas básicas do currículo escolar.

Estudos empíricos sugerem que a influência do *background* familiar no desempenho escolar tem uma relação inversa com o nível de escolaridade do indivíduo. Segundo Silva e Hasenbalg (2002), esse efeito diminui nas fases finais do ensino fundamental. Sendo assim, achou-se mais conveniente utilizar as notas de proficiência dos alunos de 8ª série/9º ano do ensino fundamental.

A nota conceito do SAEB varia em uma escala de 0 a 500, sendo que a média municipal nas disciplinas de português e matemática em 2015 foram, respectivamente, de 244,91 e 249,81. As notas máximas nas disciplinas de português e matemática foram de 304,23 e 322,74, respectivamente. Na Figura 1 abaixo, apresenta-se a distribuição da educação escolar nos municípios brasileiros segundo as proficiências de português e matemática.

Como percebe-se da visualização da Figura 1, os municípios com as maiores notas localizam-se predominantemente nas regiões Sul e Sudeste, apesar de já percebida uma maior disseminação da qualidade da educação escolar no espaço territorial brasileiro, em contraste com a concentração observada nos trabalhos de Vernier (2016) e Vernier, Bagolin e Fochezatto (2017).

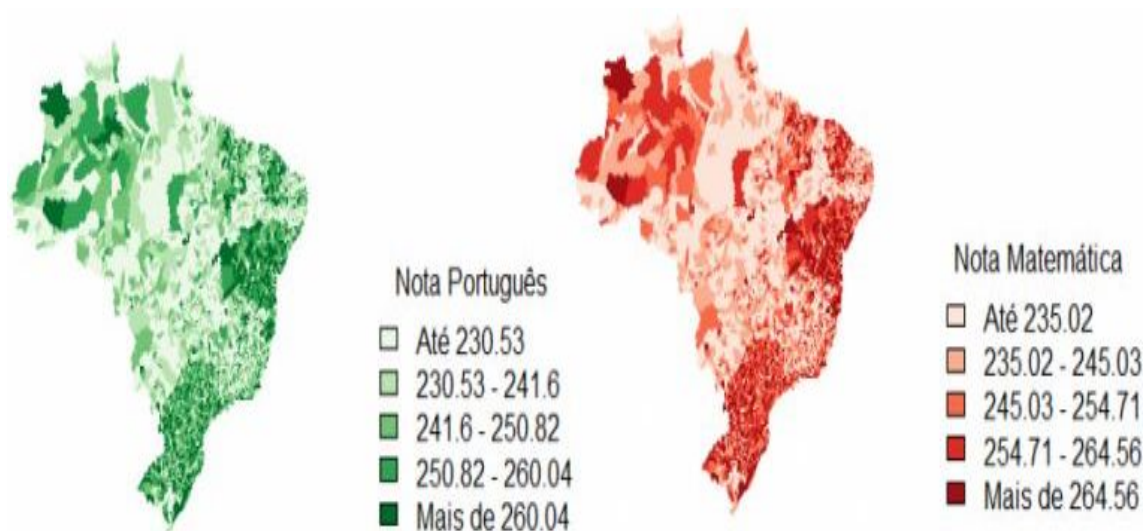


Figura 1: Distribuição da Educação Escolar segundo a proficiência em Português e Matemática

Fonte: Elaboração própria a partir de dados do INEP (2015) e IBGE (2015).

Visando controlar por fatores não observáveis, as proficiências médias em português e matemática dos alunos do 5º ano obtidas em 2011 foram utilizadas como controle<sup>4</sup>. A idéia da inclusão de tais variáveis também se baseia na hipótese da existência de um processo de aprendizagem cumulativa por parte do aluno, no sentido de haver uma associação positiva entre o seu desempenho escolar prévio e o seu desempenho escolar atual, como previsto em Hanushek (2002) e Ding e Lehrer (2006).

De acordo com a literatura da educação, diversas variáveis afetam o desempenho escolar como o sexo do indivíduo, cor/raça, *background* familiar (escolaridade dos pais, renda familiar, incentivo dos pais, etc) e características dos docentes. Dessa forma, para controlar o efeito do espaço no desempenho escolar, procede-se pela inclusão de tais variáveis. Para construir a base de dados a nível de município, as variáveis relativas aos alunos e aos docentes foram agregados por município a partir do cálculo da proporção de indivíduos com determinada característica.

A Tabela 1 mostra as estatísticas descritivas das variáveis utilizadas no presente estudo.

<sup>4</sup>Atenta-se para o fato de que apenas essas variáveis serão referentes ao ano de 2011, permanecendo todas as demais referentes ao ano de 2015, mantendo a estrutura dos dados *cross-section*.



Tabela 1: Estatística descritiva das variáveis utilizadas

Variável	Média	Mediana	Desvio-padrão	Mínimo	Máximo
Nota português	244,9	245,7	17,27	177,7	304,2
Nota matemática	249,8	249,3	17,40	201,3	322,7
Prop. de alunos do sexo masculino	0,48	0,48	0,07	0,19	0,88
Prop. de alunos da cor branca	0,30	0,25	0,17	0,00	1,00
Prop. de mães com ao menos ensino médio completo	0,30	0,30	0,10	0,00	0,80
Prop. de pais com ao menos ensino médio completo	0,20	0,19	0,09	0,00	0,65
Prop. de alunos que moram com o pai	0,65	0,64	0,09	0,30	1,00
Prop. de alunos com computador em casa	0,55	0,55	0,20	0,00	1,00
Prop. de alunos que trabalham fora de casa	0,18	0,17	0,08	0,00	0,65
Prop. de alunos que moram em casas com mais de 4 pessoas	0,40	0,39	0,11	0,00	0,86
Prop. de pais que incentivam o estudo dos filhos	0,52	0,52	0,12	0,07	0,95
Prop. de professores pós-graduados	0,72	0,75	0,25	0,00	1,00
Prop. de professores esforçados de português	0,84	0,85	0,10	0,23	1,00
Prop. de professores esforçados de matemática	0,85	0,87	0,09	0,22	1,00
Prop. da adequação formação docente-ensino	0,46	0,46	0,21	0,00	1,00
PIB per capita	18.648	13.880	21.188,75	3.086	815.094
Despesas educação (Bilhões)	15,2	6,2	2,6274e-7	49,3	3.120

Fonte: Elaboração própria a partir de dados do INEP (2015), IBGE (2015), do Observatório da Criança e do Adolescente (2015) e do Portal FINBRA (2015).

Como pode-se observar na Tabela 1 acima, as médias das proficiências das disciplinas de português e matemática se aproximam, sendo a última superior à primeira. Quanto às variáveis socioeconômicas dos alunos, tem-se uma amostra equilibrada de alunos do sexo feminino e masculino, com mães com nível de escolaridade maior que o do pai, pais que moram com seus filhos e os incentivam razoavelmente ao estudo. Ainda em relação às características familiares, os alunos possuem em sua maioria computador em casa, em geral, dedicam seu tempo exclusivamente ao estudo e moram em famílias de tamanhos medianos (com no máximo 4 pessoas, incluindo o aluno).

Além disso, tais alunos estudam, em sua maioria, em escolas com professores pós-graduados e considerados esforçados, porém, ainda não adequadamente alocados em disciplinas correspondentes à sua formação. Acrescido a isso, estudam em escolas localizadas em municípios com um moderado PIB per capita e que gastam significativamente com a educação no nível fundamental.

## 3 RESULTADOS

### 3.1 Análise Exploratória de Dados Espaciais

De acordo com Baumont (2004), existe uma técnica de escolha da especificação mais adequada no sentido da matriz de pesos espaciais a ser escolhida. Segundo tal procedimento testa-se os resíduos do modelo de regressão linear estimado usando o coeficiente  $I$  de Moran para um conjunto de matrizes  $W$  e define-se  $K$  que gere o valor mais alto para tal estatística de teste e que seja estatisticamente significativa. Dessa forma, o valor da estatística do  $I$  de Moran para as matrizes de  $K$ -vizinhos testadas são apresentados a seguir na Tabela 2:

Tabela 2: I de Moran segundo Matriz de Ponderação Espacial

Variável	Convenção					
	Rainha	Torre	K = 5	K = 10	K = 15	K = 20
portugues2015	0,3519***	0,3514***	0,3526***	0,3434***	0,3366***	0,3307***
matematica2015	0,3688***	0,3686***	0,3751***	0,3628***	0,3542***	0,3468***

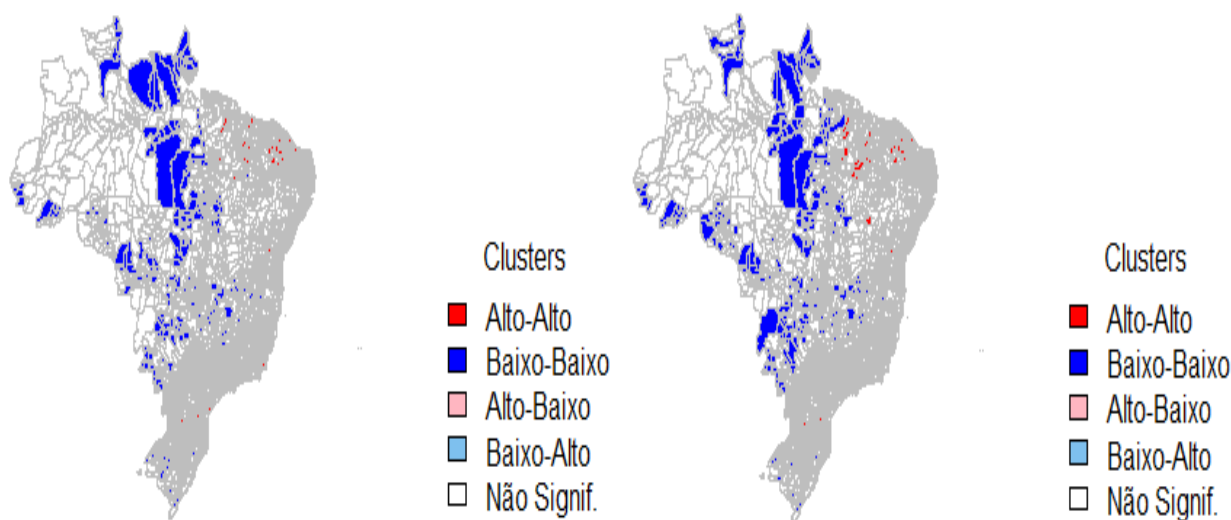
Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

Segundo esse critério, foi escolhida a matriz de ponderação espacial com os cinco vizinhos mais próximos ( $k=5$ ). Ademais, adotou-se a matriz de  $K$ -vizinhos tendo em vista que, nos estudos empíricos de educação que utilizam a metodologia de econométrico-espacial citados anteriormente, a mesma é utilizada com maior frequência, já que essa convenção garante a inexistência de municípios sem vizinhos (LOBÃO e DA SILVA, 2018; RODRIGUES et al., 2017; VERNIER, BAGOLIN e FOCHEZATTO, 2017).

Após a definição da matriz de pesos espaciais a ser adotada, o próximo passo consiste na AEDE por meio da estatística de teste do  $I$  de Moran. O valor do coeficiente da variável dependente do desempenho em português e em matemática significativo e positivo acima do valor esperado sugere autocorrelação espacial positiva, isto é, concentração dos dados, dando um prévio indício que municípios com alto desempenho em português e matemática estão próximos a municípios também com altas notas nessas disciplinas.

Adicionalmente, são apresentados nas Figuras 2(a) e 2(b) os mapas de *clusters* LISA referentes às mesmas variáveis para auxiliar na identificação de padrões espaciais. Como pode-se ver, há a predominância de *clusters* Alto-Alto nas regiões Sudeste, Sul e Nordeste e *clusters* Baixo-Baixo nas regiões Norte e Centro-Oeste, indicando a presença de autocorrelação positiva nas notas de português e matemática, significando que municípios com notas altas (baixas) em português e em matemática estão próximas de municípios que também apresentam notas altas (baixas), isto é, melhores (piores) desempenhos nas avaliações do SAEB.



(a) Figura 2: LISA para Nota de Português

(b) Figura 2: LISA para Nota de Matemática

Os *clusters* de alto desempenho identificados na região Nordeste podem ser explicados pela melhora recente das notas de desempenho de alguns municípios em 2015. Exemplo disso são os municípios dos Estados do Piauí e Ceará que vêm se destacando frente aos demais estados da região que, apesar disso, ainda apresenta grande desigualdade dentro dos próprios estados.

Já os *clusters* de baixo desempenho possivelmente retratam uma piora acentuada observada nos indicadores educacionais em 2015 dos municípios da região Norte e uma estagnação na região Centro-Oeste.

### 3.2 Modelos Espaciais

O primeiro passo na avaliação da especificação e teste da dependência espacial consiste na estimação do modelo por Mínimos Quadrados Ordinários (MQO). As Tabelas 3 e 4 a seguir mostram os resultados das estimações dos modelos por MQO:

Tabela 3: Resultado do modelo MQO para nota de português e dos Testes para Autocorrelação Espacial

Variáveis	Coefficientes	E.P	p-valor
Constante	138.065	7.952	0.000
portugues2011	0.372	0.009	0.000
sexomasc	-25.672	2.431	0.000
corbranca	9.104	1.397	0.000
escolmae	1.377	1.900	0.469
moracpai	-2.780	2.033	0.172
incpais	5.211	1.391	0.000
internet	17.249	1.697	0.000
trabfora	-7.584	2.034	0.000
qtdeconj	-12.133	1.845	0.000
profposgrad	-0.252	0.616	0.682
profesfport	32.156	1.707	0.000
adformacao	-0.016	0.010	0.110
lnpibpcpt	0.663	0.333	0.046
lndespesapcpt	1.026	2.281	0.653
lndespesapcpt2	-0.094	0.198	0.634
R <sup>2</sup> ajustado	0.625		0.000
F	532.95		0.000
Testes para Autocorrelação Espacial			
I de Moran	0.0843		0.000
ML <sub>-ρ</sub>	60.6843		0.000
ML* <sub>-ρ</sub>	2.7367		0.098
ML <sub>-λ</sub>	94.3947		0.000
ML* <sub>-λ</sub>	36.4471		0.000
ML <sub>-ρ-</sub>	97.1314		0.000

Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

Tabela 4: Resultado do modelo MQO para nota de matemática e dos Testes para Autocorrelação Espacial

Variáveis	Coefficientes	E.P	p-valor
Constante	123.017	7.562	0.000
matematica2011	0.501	0.011	0.000
sexomasc	2.970	2.303	0.197
corbranca	10.451	1.319	0.000
escolmae	-1.232	1.796	0.493
moracpai	10.326	1.920	0.000
incpais	5.997	1.311	0.000
internet	15.358	1.611	0.000
trabfora	-1.544	1.921	0.422
qtdeconj	-11.982	1.755	0.000
profposgrad	0.574	0.583	0.325
profesfmat	25.506	1.668	0.000
adformacao	-0.010	0.009	0.271
lnpibpcpt	-0.916	0.315	0.004
lndespesapcpt	0.208	2.161	0.923
lndespesapcpt2	0.053	0.188	0.779
R <sup>2</sup> ajustado	0.671		0.000
F	654.29		0.000
Testes para Autocorrelação Espacial			
I de Moran	0.0678		0.000
ML_ρ	11.0562		0.0009
ML*_ρ	4.5258		0.0334
ML_λ	61.1041		0.000
ML*_λ	54.5737		0.000
ML_ρ-	65.6299		0.000

Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

A partir dos resultados deste, utiliza-se de teste para detectar a autocorrelação espacial e especificá-la de acordo com o modelo espacial mais adequado. Nesse sentido, foram realizados testes difusos e focados. Como pode-se ver dos resultados do teste difuso de  $I$  de Moran sugerido por Cliff e Ord (1973) das Tabelas 3 e 4, rejeita-se a hipótese nula de independência espacial para ambos os modelos estimados ao nível de 1% de significância, sugerindo autocorrelação positiva nos resíduos das regressões.

No que se refere aos resultados dos testes focados, pode-se observar que foram encontradas evidências de dependência espacial na variável dependente e/ou no termo de erro para os modelos estimados segundo a significância estatística dos testes de ML. Um resultado interessante encontrado e já esperado foi que, após a inclusão da variável controle de desempenho escolar prévio nos modelos, houve uma queda significativa das estatísticas dos testes ML, o que pode ser explicado pelo fato de tal variável ser uma *proxy* de efeito fixo do município, isolando assim o efeito espacial.

Em face de tais evidências encontradas, cabe a seguir analisar a melhor especificação econométrico-espacial para os referidos modelos.

Primeiramente, da análise dos resíduos do modelo estimado por MQO observou-se a presença de heterocedasticidade, autocorrelação e ausência de normalidade nos termos de erro. O modelo espacial local SLX e os modelos espaciais globais e locais SDEM e SDM foram estimados utilizando o método de Máxima Verossimilhança.

Visando definir o modelo espacial mais robusto dentre os especificados anteriormente, foram utilizados o indicador do valor da Função de Verossimilhança (LIK), os Critérios de Informação de *Akaike* (AIC), *Schwarz* (BIC) e o teste da Razão de Verossimilhança (*LR test*). Sendo este último calculado a partir da comparação do modelo SLX mais restrito com os modelos SDEM e SDM irrestritos, respectivamente.

Segundo os testes de robustez realizados, a especificação com melhor ajuste para todos os modelos estimados é via SDM, sendo os resultados das estimações apresentados nas Tabelas 5 e 6 a seguir:

Tabela 5: Resultados das estimações do modelo da nota de português

Variáveis	SLX	SDEM	SDM
Constante	141.078*** (18.441)	145.792*** (20.360)	119.481*** (18.475)
portugues2011	0.365*** (0.010)	0.365*** (0.010)	0.362*** (0.010)
demais controles	⋮	⋮	⋮
$\rho$			0.170*** (0.022)
$\lambda$		0.163*** (0.022)	
lag.portugues2011	0.092*** (0.018)	0.084*** (0.019)	0.019 (0.020)
lag.p_sexomasc	-7.732 (5.296)	-8.424 (5.584)	-3.224 (5.272)
lag.p_corbranca	-4.522* (2.654)	-5.133* (2.797)	-6.088** (2.628)
lag.p_escolmae	-9.238** (3.756)	-8.375** (3.977)	-8.262** (3.712)
lag.p_moracpai	-5.996 (4.206)	-5.210 (4.454)	-4.739 (4.158)
lag.p_incpais	-6.978** (2.905)	-5.829* (3.066)	-6.825** (2.871)
lag.p_internet	9.531*** (3.568)	8.923** (3.785)	5.572 (3.558)
lag.p_trabfora	4.306 (4.132)	3.752 (4.355)	5.158 (4.085)
lag.p_qtdeconj	4.840 (3.506)	4.286 (3.736)	6.387* (3.470)
lag.p_profposgrad	-0.606 (1.258)	-0.645 (1.333)	-0.576 (1.243)
lag.p_profesfport	-3.744 (3.584)	-3.169 (3.799)	-8.600** (3.605)

lag.p_adformacao	-0.155*** (0.019)	-0.143*** (0.020)	-0.138*** (0.019)
lag.lnpibpcpt	0.028 (0.727)	0.171 (0.765)	-0.007 (0.718)
lag.lndespapcpt	-1.514 (4.850)	-3.341 (5.158)	-2.268 (4.793)
lag.lndespapcpt2	0.206 (0.420)	0.350 (0.447)	0.258 (0.415)
Log Likelihood		-17,960.94	-17,957.40
AIC	36039.05	35987.89	35980.81
BIC	36246.29	36201.6	36194.52
LR Test (df = 1)		53.17***	60.24***

Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

Tabela 6: Resultados das estimações do modelo da nota de matemática

Variáveis	SLX	SDEM	SDM
Constante	133.099*** (17.251)	134.823*** (18.805)	114.941*** (17.336)
matematica2011	0.486*** (0.012)	0.486*** (0.011)	0.484*** (0.012)
demais controles	⋮	⋮	⋮
$\rho$			0.143*** (0.022)
$\lambda$		0.138*** (0.023)	
lag.matematica2011	0.076*** (0.021)	0.072*** (0.022)	-0.002 (0.024)
lag.p_sexomasc	-5.570 (5.031)	-5.493 (5.264)	-5.752 (4.985)
lag.p_corbranca	-12.387*** (2.504)	-12.664*** (2.616)	-13.528*** (2.486)
lag.p_escolmae	-15.322*** (3.549)	-14.242*** (3.724)	-13.521*** (3.521)
lag.p_moracpai	-5.102 (3.987)	-3.943 (4.184)	-5.655 (3.955)
lag.p_incpais	-2.709 (2.734)	-2.126 (2.861)	-3.106 (2.711)
lag.p_internet	13.117*** (3.376)	12.295*** (3.551)	9.567*** (3.380)
lag.p_trabfora	-1.146 (3.906)	-2.041 (4.083)	-1.405 (3.871)
lag.p_qtdeconj	5.000 (3.350)	4.926 (3.533)	6.538** (3.326)

lag.p_profposgrad	-1.322 (1.191)	-1.140 (1.251)	-1.251 (1.180)
lag.p_profesmat	-5.492 (3.468)	-4.580 (3.650)	-8.376** (3.475)
lag.p_adformacao	-0.100*** (0.017)	-0.093*** (0.018)	-0.091*** (0.017)
lag.lnpibpcpt	-0.555 (0.682)	-0.352 (0.713)	-0.290 (0.676)
lag.lndespapcpt	-3.650 (4.586)	-5.066 (4.832)	-4.035 (4.544)
lag.lndespapcpt2	0.450 (0.398)	0.553 (0.419)	0.455 (0.394)
Log Likelihood		-17706.41	-17704.33
AIC	35513.87	35478.82	35474.66
BIC	35721.11	35692.54	35688.37
LR Test (df = 1)		37.05***	41.21***

Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

Ao analisar a relação entre o desempenho escolar dos alunos do 5º ano do ensino fundamental em 2011 nas disciplinas de português e matemática e o desempenho escolar dos alunos da 9º ano da mesma fase de ensino nessas mesmas disciplinas em 2015 foram observados os efeitos significativo e positivo confirmando a hipótese inicial da existência de fatores não-observáveis e invariantes no tempo do município que influenciam o desempenho escolar e também de que o processo de aprendizagem é cumulativo, ou seja, há uma correlação positiva entre o conhecimento prévio e o atual, de forma que os melhores alunos do passado são também os do presente, estando de acordo com o previsto por Hanushek (2002) e Ding e Lehrer (2006).

Após a inclusão da dependência espacial no modelo de regressão clássico, verifica-se a manutenção da significância estatística dos parâmetros em todos os modelos espaciais testados para os modelos da variável dependente da proficiência em português e em matemática.

E, corroborando com os indícios encontrados na AEDE, o coeficiente  $\rho$  foi significativo ao nível de 5% e positivo para o modelo SDM tanto para os modelos da variável dependente da nota de português quanto de matemática, assim como nos resultados encontrados no estudo de Vernier, Bagolin e Fochezatto (2017), confirmando a existência de dependência espacial no desempenho escolar entre os municípios. Os coeficientes  $\tau$  das defasagens espaciais das variáveis explicativas da cor do aluno, da escolaridade das mães, do incentivo dos pais, do tamanho da família, do esforço dos docentes e da adequação formação docente-ensino apresentaram-se estatisticamente significantes ao nível de 5% de significância para os modelos da nota de português e matemática, enfatizando a importância de especificar o transbordamento das variáveis explicativas para evitar o viés de variável omitida que ocasionaria o problema da endogeneidade. Sendo assim, a significância estatística dos coeficientes  $\rho$  e  $\tau$  do modelo SDM sugere a presença de uma dependência espacial mais intrincada, manifestada pelo multiplicador espacial global do processo SAR de defasagem espacial impulsionado pelos multiplicadores adicionais de alcance local nas variáveis explicativas do modelo SLX.

### 3.3 Efeitos Marginais

Dessa forma, foram estimados os efeitos marginais diretos, indiretos e totais para o modelo SDM visando a análise do efeito de uma variação marginal e sua retroalimentação através das regiões vizinhas. Os resultados são apresentados nas Tabelas 7 e 8 a seguir:

Tabela 7: Efeito marginal direto, indireto e total para o modelo da nota de português

	Direto	Indireto	Total
portugues2011	0.36***	0.09***	0.46***
p_sexomasc	-25.00***	-8.74	-33.73***
p_corbranca	10.68***	-5.00*	5.68*
p_escolmae	-0.03	-9.67**	-9.69**
p_moracpai	-2.93	-6.12	-9.05*
p_incpais	6.00***	-6.79**	-0.79
p_internet	16.37***	9.77**	26.14***
p_trabfora	-9.88***	4.07	-5.82
p_tamfamilia	-12.04***	5.07	-6.97
p_profposgrad	0.16	-0.64	-0.48
p_profesfport	30.47***	-4.00	26.47***
p_adformacao	0.02*	-0.16***	-0.14***
lnpibpcpt	0.58*	0.11	0.69
lnspesapcpt	0.51	-2.55	-2.04
lnspesapcpt2	-0.04	0.29	0.25

Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

De acordo com os resultados obtidos e expostos na Tabela 7, uma variação de 1 ponto na nota de português do 5º ano do ensino fundamental de um município gera, em média, um aumento em 0,36 pontos na nota de português do 9º ano do próprio município pelo efeito marginal direto. Já o aumento de 1 ponto na nota de português do 5º ano do ensino fundamental dos municípios vizinhos aumenta, em média, em 0,09 a nota na disciplina do 9º ano do município pelo efeito indireto, resultando em um efeito marginal total ou impacto médio para o município de um aumento de 0,46 pontos na nota de desempenho de português.

Tabela 8: Efeito marginal direto, indireto e total para o modelo da nota de matemática

	Direto	Indireto	Total
portugues2011	0.49***	0.08***	0.56***
p_sexomasc	3.83*	-5.92	-2.09
p_corbranca	15.08***	-12.94***	2.13
p_escolmae	-1.40	-15.62***	-17.02***
p_moracpai	10.66***	-4.70	5.96
p_incpais	5.40***	-2.66	2.74
p_internet	15.33***	13.39***	28.72***
p_trabfora	-0.76	-1.72	-2.49



p_tamfamilia	-12.58***	5.39	-7.19*
p_profposgrad	0.93	-1.27	-0.34
p_profesfmat	25.11***	-5.44	19.67***
p_adformacao	0.02	-0.10***	-0.09***
lnpibpcpt	-0.62*	-0.43	-1.05
lndespapcpt	-0.34	-4.54	-4.20
lndespapcpt2	0.04	0.52	0.56

Nota: \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.10$

Fonte: Elaboração própria.

De acordo com os resultados obtidos e expostos na Tabela 8, uma variação de 1 ponto na nota de matemática do 5º ano do ensino fundamental do município gera, em média, um aumento em 0,49 pontos na nota de matemática do 9º ano do próprio município pelo efeito marginal direto. Ao aumentar em 1 ponto da nota na disciplina do 5º ano dos municípios vizinhos, aumenta-se, em média, em 0,08 a nota do 9º ano do município devido ao efeito indireto. Isto resulta em um efeito marginal total ou impacto médio para o município de um aumento de 0,56 pontos na nota de desempenho de matemática.

## 4 CONSIDERAÇÕES FINAIS

Este estudo teve como objetivo central investigar, por meio de modelos econométrico-espaciais, os transbordamentos espaciais da educação nos municípios brasileiros. Buscou-se também controlar por efeitos não observáveis, sendo a nota do SAEB de 2011 do 5º ano do ensino fundamental utilizada como controle.

Ademais, foi observado o efeito significativo e positivo da variável de controle dos fatores não observáveis, a saber, o desempenho escolar prévio dos anos iniciais do ensino fundamental em português e matemática sobre o desempenho escolar atual do município nos anos finais do ensino fundamental nessas mesmas disciplinas. Esse resultado confirma a hipótese inicial do estudo da existência de fatores não observáveis do município que influenciam as notas de proficiência, em que se pese que tal variável de controle comporta diversos componentes fixos como esforço do aluno, habilidades, gosto pelos estudos, além das próprias características do município invariantes no tempo como insumos escolares, gastos com educação, renda municipal, dentre outros que afetam o desempenho escolar.

Tal evidência também confirma a natureza cumulativa do processo de aprendizagem já que a proficiência atual dos alunos depende da proficiência passada, o que sugere que a função de produção da educação deva incorporar na sua especificação um componente de fatores não-observáveis do desempenho escolar do município. Tal resultado chama atenção da necessidade da oferta de um ensino básico de qualidade, que proporcione aos alunos uma formação continuada que promova o acúmulo de conhecimento e que permita, conseqüentemente, o alcance de melhores indicadores educacionais.

Além disso, um resultado que merece ser evidenciado é que mesmo após a inclusão do controle não-observável do desempenho escolar municipal prévio, o efeito espacial se mantém, sugerindo, assim, o alto poder de explicação do espaço em relação ao desempenho escolar dos municípios. Essa evidência encontrada no estudo abre espaço para uma vasta discussão da importância da consideração do espaço na determinação da qualidade da educação ofertada por meio do reconhecimento dos limites geográficos no estudo das externalidades educacionais, que

foi durante muito tempo negligenciado em grande parte dos estudos empíricos.

Corroborando com os indícios encontrados na AEDE, o coeficiente  $\rho$  apresentou-se estatisticamente significativo e positivo para o desempenho em português e matemática, sugerindo que há interdependência espacial no desempenho escolar dos municípios. Acrescenta-se a isso, a significância estatística encontrada do coeficiente  $\tau$  de defasagem espacial nas variáveis explicativas comumente utilizadas na literatura empírica indicando a existência de transbordamentos espaciais locais das características dos vizinhos influenciando o desempenho escolar dos municípios e, acima de tudo, a importância de considerar o transbordamento das variáveis explicativas visando mitigar o viés da variável omitida.

Sendo assim, os resultados obtidos sustentam a hipótese inicial do estudo de existência de *spillovers* educacionais entre os municípios, nos quais foram encontradas evidências de uma dependência espacial mais intrincada, visto que há dependência espacial tanto na variável dependente quanto nas variáveis explicativas do modelo, sugerindo que a melhor especificação é dada pelo *Spatial Durbin Model* (SDM). Os efeitos marginais diretos, indiretos e totais estimados indicam que o desempenho escolar de um município está condicionado ao desempenho escolar prévio do próprio município, ao desempenho escolar dos municípios vizinhos e às demais características escolares e econômicas dos municípios e dos seus vizinhos.

Apesar de obtidos os resultados esperados no estudo, tem-se que reconhecer o ainda muito recente uso da econometria espacial no estudo das externalidades educacionais. Entretanto, a incipiência do tema consiste em uma motivação a mais para futuros trabalhos. Por fim, uma sugestão para trabalhos futuros seria a incorporação na análise de variáveis mais passíveis de intervenção de políticas públicas, como infraestrutura escolar, gestão escolar, entre outras.

## REFERÊNCIAS

ALMEIDA, E. **Econometria espacial aplicada**. 1 ed. Campinas: Alínea, 2012.

ANSELIN, L. **Spatial Econometrics: Methods and Models**. 1988. 151 f.

ARAÚJO, I.; SILVEIRA NETO, R.M. Concentração Geográfica de Capital Humano, Ganhos de Produtividade e Disparidades Regionais de Renda: Evidências para o Brasil Metropolitano. **Revista Econômica do Nordeste**, Fortaleza, v.35, n.3, jul-set, 2004.

BARRO, R. J.; LEE, J. W. A new data set of educational attainment in the world, 1950-2010. **Journal of Development Economics**, v. 104, p. 184-198, 2010.

BAUMONT, C. **Spatial effects in housing price models: do house prices capitalize urban development policies in the agglomeration of Dijon (1999)?** Mimeo. Université de Bourgogne, 2004.

BRASIL. Secretaria do Tesouro Nacional. **Sistema de Informações Contábeis e Fiscais do Setor Público Brasileiro (Siconfi)**. Brasília: STN, 2019. Disponível em: <[https://siconfi.tesouro.gov.br/siconfi/pages/public/consulta\\_finbra](https://siconfi.tesouro.gov.br/siconfi/pages/public/consulta_finbra)>. Acesso em: 15 abr. 2019.

BIONDI, R. L.; FELÍCIO, F. **Atributos escolares eo desempenho dos estudantes: uma análise em painel dos dados do Saeb**. Ministério da Educação, Instituto Nacional de Estudos e Pes-

quisas, 2007.

CARNEIRO, V. L. **Análise de spillovers nos gastos municipais**. Dissertação de Mestrado - Universidade de Brasília. Brasília: DF, jul. 2014.

CLIFF, A.; ORD, J. K. **Spatial autocorrelation**. London: Pion, 1973.

DING, W.; LEHRER, S. F. Do peers affect student achievement in China's secondary schools?. **National Bureau of Economic Research**, n. 12305, p. 1-42, june 2006.

FALCÃO, N. de A.; SILVEIRA NETO, R. da M. Os efeitos das economias de aglomeração sobre os salários industriais: uma aplicação ao caso brasileiro. In: ENCONTRO DA ASSOCIAÇÃO NACIONAL DOS CENTROS DE PÓS-GRADUAÇÃO EM ECONOMIA, XXXV, 2007, **Anais eletrônicos...** Recife: UFPE, 2007. Disponível em: <<http://www.anpec.org.br/encontro2007/artigos/A07A159.pdf>>. Acesso em: 26 nov. 2018.

FOWLES, J. T.; TANDBERG, D. A. State Higher Education Spending: A Spatial Econometric Perspective. **American Behavioral Scientist**, v. 61, n. 14, p. 1773-1798, 2018.

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

FUNDAÇÃO Abrinq. **Observatório da Criança e do Adolescente**. São Paulo: Abrinq, 2015. Disponível em: <<https://observatoriocrianca.org.br/cenario-infancia/temas/ensino-fundamental/1077-adequacao-da-formacao-docente-no-ensino-fundamental?filters=1,1679>>. Acesso em: 15 abr. 2019.

HANUSHEK, E. A. **National Bureau of Economic Research**, n. 8799, p. 1-95, feb. 2002.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Produto Interno Bruto dos Municípios 2010-2014**. Brasília: IBGE, 2018. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9088-produto-interno-bruto-dos-municipios.html?edicao=18760&t=downloads>>. Acesso em: 10 nov. 2018.

INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA. **Microdados da Aneb e da Anesc 2015**. Brasília: Inep, 2017. Disponível em: <<http://portal.inep.gov.br/basica-levantamentos-acessar>>. Acesso em: 10 nov. 2018.

LESAGE, J.; PACE, R. K. **Introduction to spatial econometrics**. Chapman & Hall/CRC, 2009.

LOBÃO, M. S. P.; DA SILVA, R. G. Spillovers espaciais no desenvolvimento educacional dos municípios da Amazônia Clássica Brasileira. **REDES: Revista do Desenvolvimento Regional**, v. 23, n. 2, 290-315, 2018.

MACHADO, A. F. et al. Qualidade do Ensino em Matemática: Determinantes do Desempenho de Alunos em Escolas Públicas Estaduais Mineiras. **Revista Economia**, v. 9, n.1, 23-45, jan. 2008.

MENEZES-FILHO, N. Educação e desigualdade. In: MENEZES-FILHO, N. A.; LISBOA, M. (Eds.). **Microeconomia e sociedade**. Rio de Janeiro: Contracapa, 2001.

MENEZES-FILHO, N.; FERNANDES, R.; PICCHETTI, P. Rising human capital, but constant inequality: the education composition effect in Brazil. **Revista Brasileira de Economia**, v. 60, p. 200-250, 2006.

MENEZES-FILHO, N. **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.

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, 2002.

QUEIROZ, B. L. Diferencial Regional de Salários e Retornos Sociais à Educação: Uma Abordagem Hierárquica. In: **Mercado de Trabalho: Uma Análise a Partir das Pesquisas Domésticas no Brasil**. WAJNMAN, S.; MACHADO, A. F. (Org.). Belo Horizonte: Editora UFMG, 2003.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2019. Disponível em: <<https://www.R-project.org/>>.

RÊGO, C. V.; PENA, M. G. **Análise dos modelos de regressão espacial SAR, SEM e SAC**. Trabalho de Conclusão de Curso (Bacharelado em Estatística) - Universidade de Brasília, Brasília, 2012. 88 f.

RODRIGUES, L. et al. Mensuração da desigualdade educacional entre os municípios nordestinos. **Revista de Economia Contemporânea**, v. 21, p. 1-31, 2017.

SILVA, N. do V.; HASENBALG, C. Recursos familiares e transições educacionais. **Cadernos de Saúde Pública**, v. 18, p. S67-S76, 2002.

SOUZA, M. I. de. A. et al. **Textos de Economia**, Florianópolis, v.16, n.2, p.111-142, jul./dez. 2013.

VERNIER, L. D. S. **Crescimento educacional brasileiro: uma análise da distribuição e disseminação dos efeitos espaciais**. Tese (Doutorado em Economia do Desenvolvimento da Escola de Negócios) - Pontifícia Universidade Católica do Rio Grande do Sul, 2016. 81 f.

VERNIER, L. D. S.; BAGOLIN, I. P.; FOCHEZATTO, A. Distribuição e disseminação espacial da educação nos municípios brasileiros. In: ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS, XV, 2017.