

UNIVERSIDADE FEDERAL DE SANTA MARIA  
CENTRO DE CIÊNCIAS SOCIAIS E HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA E  
DESENVOLVIMENTO

**Wallace Lobato Siqueira**

**DETERMINANTES DA QUALIDADE DA PROFICIÊNCIA DOS  
ALUNOS DAS ESCOLAS PÚBLICAS BRASILEIRAS DE 2011 A 2017:  
UMA ABORDAGEM COM DADOS EM PAINEL ESPACIAL**

Santa Maria, RS  
2020

**Wallace Lobato Siqueira**

**DETERMINANTES DA QUALIDADE DA PROFICIÊNCIA DOS ALUNOS DAS  
ESCOLAS PÚBLICAS BRASILEIRAS DE 2011 A 2017: UMA ABORDAGEM COM  
DADOS EM PAINEL ESPACIAL**

Dissertação apresentada ao Curso de Pós-graduação em Economia e Desenvolvimento, da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do título de **Mestre em Economia e Desenvolvimento**.

Orientador: Dr. Clailton Ataídes de Freitas  
Coorientadora: Dra. Kalinca Léia Becker

Santa Maria, RS  
2020

Siqueira, Wallace Lobato

Determinantes da qualidade da proficiência dos alunos das escolas públicas brasileiras de 2011 a 2017: uma abordagem com dados em painel espacial / Wallace Lobato Siqueira.- 2020.

86 p.; 30 cm

Orientador: Clailton Ataídes de Freitas

Coorientadora: Kalinca Léia Becker

Dissertação (mestrado) - Universidade Federal de Santa Maria, Centro de Ciências Sociais e Humanas, Programa de Pós-Graduação em Economia e Desenvolvimento, RS, 2020

1. Educação 2. Desempenho escolar 3. Sistema de Avaliação da Educação Básica 4. Dados em painel espacial  
I. de Freitas, Clailton Ataídes II. Becker, Kalinca Léia  
III. Título.

Sistema de geração automática de ficha catalográfica da UFSM. Dados fornecidos pelo autor(a). Sob supervisão da Direção da Divisão de Processos Técnicos da Biblioteca Central. Bibliotecária responsável Paula Schoenfeldt Patta CRB 10/1728.

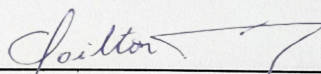
Declaro, WALLACE LOBATO SIQUEIRA, para os devidos fins e sob as penas da lei, que a pesquisa constante neste trabalho de conclusão de curso (Dissertação) foi por mim elaborada e que as informações necessárias objeto de consulta em literatura e outras fontes estão devidamente referenciadas. Declaro, ainda, que este trabalho ou parte dele não foi apresentado anteriormente para obtenção de qualquer outro grau acadêmico, estando ciente de que a inveracidade da presente declaração poderá resultar na anulação da titulação pela Universidade, entre outras consequências legais.

Wallace Lobato Siqueira

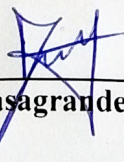
**DETERMINANTES DA QUALIDADE DA PROFICIÊNCIA DOS ALUNOS DAS  
ESCOLAS PÚBLICAS BRASILEIRAS DE 2011 A 2017: UMA ABORDAGEM COM  
DADOS EM PAINEL ESPACIAL**

Dissertação de mestrado apresentada ao Programa de Pós-graduação em Economia e Desenvolvimento da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para a obtenção do título de **Mestre em Economia e Desenvolvimento**.

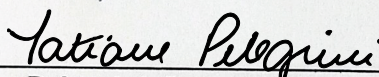
Aprovado em 18 de maio de 2020:



\_\_\_\_\_  
**Clailton Ataídes de Freitas, Dr. (UFSM) – Parecer**  
(Presidente/Orientador)



\_\_\_\_\_  
**Dieison Lenon Casagrande, Dr. (UFGD) – Parecer**



\_\_\_\_\_  
**Tatiane Pelegrini, Dra. (PUC-RS) – Parecer**



*Dedico a todos que contribuíram de alguma forma para que eu chegasse até aqui.*

*Educação nunca foi despesa.  
Sempre foi investimento com retorno garantido.*

(William Arthur Lewis)

## RESUMO

### **DETERMINANTES DA QUALIDADE DA PROFICIÊNCIA DOS ALUNOS DAS ESCOLAS PÚBLICAS BRASILEIRAS DE 2011 A 2017: UMA ABORDAGEM COM DADOS EM PAINEL ESPACIAL**

AUTOR: Wallace Lobato Siqueira  
ORIENTADOR: Dr. Clailton Ataídes de Freitas  
COORIENTADORA: Dra. Kalinca Léia Becker

Esta dissertação apresenta um estudo sobre os condicionantes da proficiência escolar dos alunos em escolas públicas nos municípios brasileiros. Verificaram-se os efeitos que a infraestrutura física do ambiente escolar, a qualificação dos professores, as características dos alunos e de suas famílias exerceram no desempenho em matemática dos alunos do 9º ano do ensino fundamental das escolas públicas agregadas em municípios. Para isso, foram utilizados os dados do Sistema de Avaliação da Educação Básica (SAEB) e do Censo Escolar entre 2011 e 2017, sob a metodologia de dados em painel espacial. Os resultados apontaram que as condições familiares foram os fatores que mais influenciaram o desempenho médio dos alunos no SAEB em nível municipal no período estudado, principalmente no que se refere à escolaridade da mãe, de modo que a proporção de mães mais instruídas teve impactos tanto diretos quanto indiretos para os municípios vizinhos, demonstrando que o *background* familiar de uma região pode afetar positivamente tanto o desempenho escolar local quanto o de outras regiões. Em relação aos demais parâmetros, a proporção de alunos que leem com frequência e fazem os exercícios para casa, em geral, tiveram efeitos positivos e significativos sobre o desempenho nas avaliações do SAEB, assim como a proporção de alunos que tiveram suas tarefas corrigidas pelos professores. Para além desses, verificou-se forte dependência espacial da nota municipal em matemática do SAEB, constatando que a estrutura espacial é relevante para determinação do desempenho escolar dos alunos na esfera municipal no Brasil, em que a performance do município nesse quesito está associada de forma positiva à dos demais municípios, ou seja, há transbordamentos positivos das variáveis internas para a sua vizinhança. Em contrapartida, os resíduos possuíram efeito inverso, em que choques podem diminuir o desempenho médio do município em questão, transbordando para seus vizinhos.

**Palavras-chave:** Educação. Desempenho escolar. SAEB.

## ABSTRACT

### **DETERMINANTS THE QUALITY OF BRAZILIAN PUBLIC SCHOOL STUDENT PROFICIENCY FROM 2011 TO 2017: A SPACE PANEL DATA APPROACH**

AUTHOR: Wallace Lobato Siqueira  
ADVISOR: Clailton Ataídes de Freitas  
COADVISOR: Kalinca Léia Becker

This dissertation presents a study on the determinants of quality of Brazilian education. It was verified the effects that the school infrastructure, the teacher qualifications, the characteristics of students and their families had on the mathematics performance of students in the 9th grade of elementary education in public schools in municipalities of Brazil. For this, were used data from SAEB and School Census between 2011 and 2017, using spatial panel methodology. The results showed that family conditions were the factors that most influenced the average performance of students in SAEB at the municipal level in the period studied, especially about the mother's education. Because the proportion of mothers with more education had direct and indirect impacts, for neighbouring municipalities, demonstrating that the family background of a region can positively affect the local school performance and that of its neighbours. Regarding the parameters, the proportion of students who read often and do the exercises at home had positive effects and significant performed in the SAEB assessments, as well as the proportion of students who had their homework corrected by the teachers. In addition to these, there was a strong spatial dependence on the municipal grade in mathematics from the SAEB, noting that the spatial structure is relevant for determining the school performance of students at the municipal level in Brazil, where the performance of the municipality is positively associated with from the other municipalities, that is, there are positive overflows of internal variables for their neighbourhood. On the other hand, the residues had the opposite effect, in which shocks can decrease the average performance of the municipality in question, overflowing to its neighbours.

**Keywords:** Education. School performance. SAEB.



## LISTA DE FIGURAS

Figura 1 - Representação das matrizes de ponderação espacial. ....	40
Figura 2 - Índice de Moran local univariado da nota média de matemática municipal para o nono ano do ensino fundamental das escolas públicas .....	54
Figura 3 - Distribuição das notas do SAEB por intervalos de nota.....	57

## LISTA DE GRÁFICOS

Gráfico 1 - IDEB do Ensino Fundamental (anos iniciais) por região .....	34
Gráfico 2 - Proporção de professores com ensino superior completo por região.....	35
Gráfico 3 - Distribuição do investimento público em educação e por nível de ensino em relação ao PIB.....	36
Gráfico 4 - Distribuição do investimento público em educação e por nível de ensino em relação ao PIB <i>per capita</i> , sem o ensino superior .....	37
Gráfico 5 - Distribuição do investimento público em educação e por nível de ensino em relação ao PIB <i>per capita</i> , com o ensino superior .....	38

## LISTA DE TABELAS

Tabela 1 - Escolha da matriz de contiguidade para o tratamento da variável dependente (Índice de Moran) .....	53
Tabela 2 - Número de cidades nos <i>clusters</i> do <i>I</i> de Moran local.....	55
Tabela 3 - Quantidade de cidades por intervalo de notas do SAEB.....	56
Tabela 4 - Resumo das estatísticas dos testes realizados .....	58
Tabela 5 - Escolha da matriz de contiguidade para o tratamento dos resíduos segundo o critério de Critério de Baumont (Índice de Moran) .....	60
Tabela 6 - Estatísticas do critério de Akaike (AIC) para seleção do modelo econométrico espacial .....	60
Tabela 7 - Resumo das estimativas dos modelos econométricos espaciais (*) .....	60
Tabela 8 - Impacto direto, indireto e total das variáveis independentes.....	65

## LISTA DE QUADROS

Quadro 1 - Sinais dos parâmetros à guisa da teoria.....	48
Quadro 2 - Testes e suas hipóteses .....	49
Quadro 3 - Nomenclatura e descrição das variáveis utilizadas .....	51



## LISTA DE ABREVIATURAS E SIGLAS

AA	Alto-alto
AB	Alto-baixo
AEDE	Análise exploratória de dados espaciais
AIC	Critério de informação de Akaike
AMC	Áreas mínimas comparáveis
ANA	Avaliação Nacional de Alfabetização
ANEB	Avaliação Nacional da Educação Básica
ANRESC	Avaliação Nacional do Rendimento Escolar
BA	Baixo-alto
BB	Baixo-baixo
ENCCEJA	Exame Nacional para Certificação de Competências de Jovens e Adultos
ENEM	Exame Nacional do Ensino Médio
FE	Efeitos fixos
FINBRA	Finanças do Brasil
FUNDEB	Fundo de Manutenção e Desenvolvimento da Educação Básica
GWR	<i>Geographically Weighted Regression</i>
$I$	Índice de Moran global univariado
$I_i$	Índice de Moran local univariado
ICMS	Imposto sobre Circulação de Mercadorias e Serviços
IDEB	Índice de Desenvolvimento da Educação Básica
IED	Investimentos externos diretos
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
IQIM	Índice de Qualidade Institucional Municipal
ISCED	<i>International Standard Classification of Education</i>
LISA	<i>Local indicators of spatial association</i>
MEC	Ministério da Educação
MQO	Mínimos quadrados ordinários
OECD	<i>Organisation for Economic Co-operation and Development</i>
PIB	Produto Interno Bruto
PISA	<i>Programme for International Student Assessment</i>
PNAD	Pesquisa Nacional por Amostra de Domicílios
PNE	Portadores de necessidades especiais
POLS	<i>Pooled ordinary least square</i>
PPV	Pesquisa sobre Padrões de Vida
RCE	Retornos constantes de escala
RE	Efeitos aleatórios
SAC	<i>Spatial Autoregressive</i>
SAEB	Sistema de Avaliação da Educação Básica
SAR	<i>Spatial Autogressive Model</i>
SARMA	<i>Spatial Autogressive model with Moving Average</i>
SC	Critério de informação de Schwarz
SDEM	<i>Spatial Durbin and Error Model</i>

SDM	<i>Spatial Durbin Model</i>
SEM	<i>Spatial Error Model</i>
SLX	<i>Spatial model with Lag in X</i>
SLXMA	<i>Spatial model with Lag in X and Moving Average</i>
SMA	<i>Spatial Moving Average Model</i>
STN	Secretaria do Tesouro Nacional
TALIS	<i>Teaching and Learning International Survey</i>
UNESCO	<i>United Nations Educational, Scientific and Cultural Organization</i>
VI	Variáveis instrumentais

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>14</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> .....	<b>19</b>
<b>3</b>	<b>REVISÃO DE LITERATURA</b> .....	<b>22</b>
<b>4</b>	<b>FATOS ESTILIZADOS DO SISTEMA DE EDUCAÇÃO NO BRASIL</b> .....	<b>33</b>
<b>5</b>	<b>METODOLOGIA</b> .....	<b>39</b>
<b>5.1</b>	<b>ASPECTOS DOS PROCEDIMENTOS ECONOMÉTRICOS ESPACIAIS</b> .....	<b>39</b>
<b>5.1.1</b>	<b>Matriz de pesos espaciais</b> .....	<b>39</b>
<b>5.1.2</b>	<b>Autocorrelação espacial</b> .....	<b>40</b>
<b>5.2</b>	<b>DEFINIÇÃO E ESPECIFICAÇÃO DOS MODELOS ECONOMÉTRICOS</b> .....	<b>41</b>
<b>5.2.1</b>	<b>Modelagem clássica de dados em painel</b> .....	<b>42</b>
<b>5.2.2</b>	<b>Modelagem de dados em painel espacial</b> .....	<b>44</b>
<b>5.2.3</b>	<b>Relacionando os testes econométricos</b> .....	<b>48</b>
<b>5.3</b>	<b>FONTE E BASE DE DADOS</b> .....	<b>50</b>
<b>6</b>	<b>ANÁLISE E DISCUSSÃO DOS RESULTADOS</b> .....	<b>53</b>
<b>7</b>	<b>CONCLUSÃO</b> .....	<b>68</b>
	<b>BIBLIOGRAFIA</b> .....	<b>69</b>
	<b>APÊNDICE A – TESTE DE WHITE DESCONSIDERANDO AS DEFASAGENS ESPACIAIS</b> .....	<b>76</b>
	<b>APÊNDICE B – TESTE DE CHOW DESCONSIDERANDO AS DEFASAGENS ESPACIAIS</b> .....	<b>77</b>
	<b>APÊNDICE C – TESTE DE BREUSCH-PAGAN DESCONSIDERANDO AS DEFASAGENS ESPACIAIS</b> .....	<b>78</b>
	<b>APÊNDICE D – TESTE DE HAUSMAN</b> .....	<b>79</b>
	<b>APÊNDICE E – TESTE DE HAUSMAN ROBUSTO</b> .....	<b>80</b>
	<b>APÊNDICE F – TESTE CD DE PESARAN DE DEPENDÊNCIA TRANSVERSAL DOS MODELOS DE DADOS EM PAINEL</b> .....	<b>81</b>
	<b>APÊNDICE G – TESTE CD DE PESARAN DE FRACA DEPENDÊNCIA DA SEÇÃO TRANSVERSAL</b> .....	<b>82</b>
	<b>APÊNDICE H – TESTE CD DE PESARAN PARA AS VARIÁVEIS</b> .....	<b>83</b>
	<b>APÊNDICE I – TESTE CD DE PESARAN PARA OS RESÍDUOS DA REGRESSÃO POR EFEITOS FIXOS</b> .....	<b>84</b>
	<b>APÊNDICE J – TESTE DE WOOLDRIDGE PARA CONSTATAR AUTOCORRELAÇÃO</b> .....	<b>85</b>
	<b>APÊNDICE K – IMPACTO DIRETO, INDIRETO E TOTAL</b> .....	<b>86</b>

## 1 INTRODUÇÃO

A educação é um dos alicerces da economia, sendo responsável por formar e aprimorar capacidades das quais os indivíduos necessitam para compor a força de trabalho e gerar progresso tecnológico. Além disso, possui um papel civilizatório, pois permite desenvolver pensamento crítico e reflexivo que permeia todos os aspectos da vida (CUNHA *et al.*, 2006).

Grandes têm sido os esforços de vários pesquisadores, tanto teóricos quanto empíricos, para o desenvolvimento de parâmetros e explicações para os determinantes do desempenho escolar dos alunos das escolas públicas brasileiras, principalmente, com o intuito de avaliar as políticas públicas e analisar os fatores que geram a qualidade da educação e a eficiência dos gastos (BIONDI; FELÍCIO, 2007; CURI; MENEZES-FILHO, 2009; CADAVAL; MONTEIRO, 2011)

Por uma perspectiva sociológica, a educação é autojustificável, uma vez que promove a dignidade humana, visto que cria condições de inserção no mercado de trabalho, permitindo o fomento a novas oportunidades, para que os indivíduos sejam o que almejam ser, além de apresentar-se como fator imprescindível para também promover alívio da pobreza e suas mazelas (SEN, 2000; BUENO; FIGUEIREDO, 2012).

Sob o ângulo econômico, a educação é essencial para qualidade da força de trabalho, para os avanços tecnológicos e para o crescimento e o desenvolvimento econômico de uma nação, quer seja pela equalização das oportunidades, pela sua influência na mobilidade social, ou mesmo na distribuição de renda. Assim, sua ação se torna importante, tanto para o recrudescimento da produtividade do trabalho quanto para as fontes das inovações, juntamente com o aumento da segurança e a melhora na saúde (HANUSHEK; KIMKO, 2000; BARBOSA FILHO; PESSÔA, 2008; 2010). A educação também pode exercer influência sobre a diminuição da criminalidade e da violência dentro e fora das escolas (BECKER; KASSOUF, 2016; HANUSHEK, 2019). Para além desses, a educação, como objeto de análise, é importante porque envolve recursos públicos que poderiam ser alocados para outros fins, logo devem ser designados de forma eficiente (RAMOS, 2015).

Em virtude da importância dessa temática, pode-se perceber que a educação é capaz de afetar positivamente todos os âmbitos da vida dos indivíduos, demonstrando que o processo de ensino é um interessante determinante para o crescimento e para o desenvolvimento econômico de um país. Todavia, são necessários investimentos na qualidade da educação,



como, por exemplo, na infraestrutura escolar, no aprimoramento e treinamentos de professores, principalmente, na escola pública (CUNHA *et al.*, 2006).

Assim, os governos nacionais e as organizações internacionais têm implementado um sistema de avaliação da qualidade da educação ofertada nas escolas públicas e privadas por meio da coleta de informações nas escolas e aplicação de testes padronizados para mensurar o nível de aprendizado dos alunos. A proficiência escolar, medida pelos testes, pode ser entendida como o domínio das habilidades que o aluno atingiu nas disciplinas que a prova afere (KLEIN, 2009). No Brasil, o Ministério da Educação (MEC), através do INEP (Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira), mensura bianualmente a evolução do desempenho dos alunos por meio do Sistema de Avaliação da Educação Básica (SAEB). Nessa avaliação, são aplicados testes padronizados das disciplinas de língua portuguesa e de matemática na população de escolas públicas e, facultativamente, particulares em todo o território nacional. Além da proficiência dos alunos, a base de dados do SAEB fornece informações acerca dos professores, dos diretores e da infraestrutura escolar, em conjunto com o Censo Escolar.

Porém, tanto nas avaliações internas quanto nas comparações internacionais, a proficiência dos alunos brasileiros tem se mostrado precária. Nas avaliações do PISA (*Programme for International Student Assessment*), o Brasil tem se posicionado entre as últimas colocações em todas as áreas do conhecimento no *ranking* entre os países amostrados. A performance dos alunos brasileiros, em 2018, ficou abaixo da média dos alunos dos países da OECD (*Organisation for Economic Co-operation and Development*). Nesse ano, o Brasil atingiu 404 pontos em ciências; em leitura, 413 pontos e em matemática, 384 pontos, enquanto os países membros da OECD, em média, atingiram 489 pontos, 487 pontos e 499 pontos, respectivamente (OECD, 2020).

Nas avaliações internas, o IDEB (Índice de Desenvolvimento da Educação Básica) calcula, por meio do desempenho no SAEB, na Prova Brasil e no Censo Escolar<sup>1</sup>, um índice que mede a evolução e a qualidade da educação recebida pelos estudantes brasileiros do ensino fundamental e, recentemente, do ensino médio e tem por finalidade auxiliar a criação de metas e o desenho de políticas específicas para que esse índice seja melhorado. Entretanto, ao analisar sua trajetória, esse tem crescido vagarosamente desde o início de sua mensuração e os melhores resultados concentram-se nas regiões com alta dinâmica econômica, como Sul e

---

1 A Prova Brasil e o Censo Escolar são outras formas e de mensurar as condições da educação no país.

Sudeste. Em 2017, apenas São Paulo, Minas Gerais, Paraná, Santa Catarina, Ceará e o Distrito Federal atingiram a média desejável (IDEB, 2020).

Ao serem observados os resultados do IDEB por regiões, identificam-se disparidades que permanecem ao longo do período de 2005 a 2017. Enquanto as regiões Sul e Sudeste atingem o índice de 3,6, em 2005, para o ensino fundamental (anos finais), nas demais regiões verifica-se um IDEB menor, sendo 3,2 na região Centro-Oeste; 3 na região Norte e, no Nordeste, 2,6. Já em 2017, notou-se recrudescimento do IDEB de todas as regiões. Contudo, a região Nordeste situou-se meio ponto abaixo do verificado nas regiões Sul e Sudeste (IDEB, 2020).

Nessa perspectiva, com o propósito de avaliar o desempenho dos alunos e as disparidades entre as regiões, estudos têm se dedicado a identificar os fatores que exercem influências sobre os conhecimentos e o desempenho dos alunos (VERNIER, 2016; BARROS; *et al.* 2001). A literatura indica a importância da infraestrutura física do ambiente escolar como relevante para caracterização da qualidade da educação (SOARES NETO *et al.* 2013; HANUSHEK, 2006; BEZERRA; KASSOUF, 2006). Nesses estudos, é apontado que a atmosfera onde o ensino é efetivado exerce influência no entendimento e no desenvolvimento das habilidades dos alunos. Soares Neto *et al.* (2013) averiguaram que, aproximadamente, 85% das escolas públicas detiveram condições elementares ou básicas de qualidade do ambiente de ensino escolar<sup>2</sup>.

No que se refere à qualificação dos professores, Hanushek e Rivkin (2006) verificaram que, em países em desenvolvimento, como o caso do Brasil, a titulação dos professores influencia na qualidade da educação. No Censo Escolar de 2017, verificou-se que, aproximadamente, 66% dos professores da educação infantil brasileira possuíam educação superior, já no ensino fundamental e médio esse percentual subiu para 78% e 87%, respectivamente (INEP, 2020).

Deve-se considerar que uma boa qualidade da infraestrutura física do ambiente escolar é capaz de influenciar os escores dos alunos, assim como maiores patamares de qualificação dos professores podem aumentar o nível de habilidades dos alunos (DUARTE, 2009; HANUSHEK; RIVKIN, 2006). Porém, conforme a literatura, os aspectos idiossincráticos

---

<sup>2</sup> Define-se elementar como ambiente escolar que tem pouca ou nenhuma condição de oferta de infraestrutura, possuindo apenas água e energia elétrica. E básica, quando há o mínimo de recursos empregados na oferta de infraestrutura, como poucos aparelhos, como computadores, impressora e televisores, segundo Soares Neto *et al.* (2013).

(como, por exemplo, as condições econômicas, a educação dos pais, a origem do estudante) são os que mais influenciam na proficiência em testes padronizados (MENEZES-FILHO, 2011; ALBERNAZ; FERREIRA; FRANCO, 2002). Mesmo assim, esses autores têm inserido variáveis sobre condições físicas da escola e titulação dos professores, a fim de verificar se esses fatores são capazes de influenciar as capacidades e habilidades dos estudantes, visto que a escola e os professores são afetados pelos efeitos das políticas públicas.

Assim, diante da problemática que permeia o presente estudo, surgem as seguintes indagações: existem disparidades geográficas na proficiência municipal do SAEB em matemática? A infraestrutura física do ambiente escolar e a formação dos professores têm contribuído para o aumento dos escores dos alunos a nível municipal na disciplina de matemática no SAEB?

Para responder a esse problema de pesquisa, tem-se por objetivo determinar os efeitos que o *background* familiar, a infraestrutura física escolar e a qualificação dos professores exerceram na proficiência em matemática dos alunos do 9º ano do ensino fundamental das escolas públicas nos municípios brasileiros, na avaliação do SAEB, no período de 2011 a 2017.

A disciplina de matemática foi escolhida por ser um conhecimento universal, que permite comparações entre testes, tanto nacionais quanto internacionais. No que se refere ao período, esse foi escolhido pela disponibilidade dos dados georreferenciados, ou seja, informações que permitem verificar a localização geográfica da origem do dado aferido, fato esse que passou a ser de acesso público partir de 2011.

O presente estudo tem por pretensão contribuir com o debate acadêmico no sentido de refletir acerca dos determinantes do desempenho dos alunos brasileiros, utilizando o modelo de dados em painel espacial, no período de 2011 a 2017, visto que Vernier (2016) evidenciou o comportamento espacial dos dados do SAEB entre 2011 e 2013, e também discutir a qualidade da educação, uma vez que os recursos públicos são escassos. Dessa forma, a consecução do presente estudo poderá contribuir com a discussão acerca da qualidade de alocação desses recursos, ou seja, se sua aplicação está ou não fazendo a educação brasileira aumentar a sua performance no que se refere às metas traçadas. Cabe ressaltar que, para medir a eficiência, estuda-se a relação entre a qualidade da educação alcançada pelos alunos em âmbito municipal e as condições de oferta da educação por parte do ente público. Assim, busca-se analisar a educação brasileira dentro de seu aspecto espacial, pois os resultados demonstraram que a proficiência atingida no SAEB é afetada pela localização geográfica, verificando um efeito de transbordamento do desempenho para as cidades próximas.

Com relação aos aspectos metodológicos, recorre-se ao modelo econométrico de dados em painel espacial, o qual consegue captar, concomitantemente, as características das séries temporais e *cross-section*, bem como os efeitos de transbordamentos do desempenho dos alunos em matemática entre os municípios.

A dissertação encontra-se estruturada em seis capítulos, sendo o primeiro composto por esta introdução; o Capítulo 2 contempla o referencial teórico; no Capítulo 3, aborda-se a revisão bibliográfica; no Capítulo 4, tem-se a metodologia adotada, em que se caracterizam a amostra, as variáveis e os métodos utilizados para conduzir a pesquisa; no quinto capítulo, realizam-se a apresentação e a análise dos resultados obtidos. E, por fim, no sexto capítulo, tem-se a conclusão a respeito dos resultados da pesquisa.

## 2 REFERENCIAL TEÓRICO

A educação, notório fator para economia de uma nação, é importante desde as primeiras formalizações do que se vinha delineando como ciência econômica. Smith (2017), que se dedicou a analisar a sociedade no século XVIII, percebeu que educação tinha um duplo papel, tanto no seu aspecto social, quanto no econômico, pelo qual defendeu que deveria haver ao menos uma instrução básica para as pessoas comuns, sendo ler, escrever e calcular quesitos essenciais para que a sociedade fosse minimamente civilizada, assim como para que esses indivíduos conseguissem exercer alguma profissão. Essa percepção revela o quanto algum nível de instrução básica, principalmente para as pessoas mais pobres, apresentava-se relevante já nos primórdios da ciência econômica.

De forma similar, Thomas Robert Malthus, contemporâneo a Smith, também entendia a importância da instrução para toda a população, defendendo que a educação poderia manter a sociedade em harmonia, vislumbrando-a como fonte para o autocontrole demográfico (MALTHUS, 1798).

Em meados do século XX, com a economia já consolidada como ciência, Mincer (1958) deu os primeiros passos na análise dos impactos que a instrução proporciona na distribuição de renda dos indivíduos, denominando-a de capital humano. O autor apontou que fatores não econômicos podem exercer papéis na distribuição de renda, como a escolha da profissão, por exemplo. Contudo, verificou que parte significativa dessas diferenças era proporcionada pelos treinamentos. Assim, além de habilidades naturais intrínsecas a cada indivíduo, quando esses recebiam algum treinamento, obtinham rendas maiores.

Friedman (1955) compartilhou das ideias neoclássicas da importância da educação básica para o aumento do bem-estar dos indivíduos. Para o autor, a educação e a instrução constituíam-se como o caminho de que a sociedade dispunha para que houvesse efeito *spillover*, promovendo maior capacidade da mão de obra. Ademais, sua análise apregoou que os conhecimentos escolares contribuem para o aumento do bem-estar, não somente individual, mas também promovem um efeito transbordamento para todas as áreas em torno da vida desses indivíduos. Ou seja, quando se investe em educação, as pessoas que convivem com o indivíduo (que adquiriu conhecimentos e habilidades através da educação) e a sua adjacência também beneficiam-se, seja por transmissão de conhecimentos dados como tácitos ou mesmo por melhor possibilidade de representação social e política.

Schultz (1961) introduziu o entendimento do que era o capital humano e suas nuances. O autor entendeu os indivíduos como capital, resgatando as ideias de Smith, que também

percebia as pessoas como a grande riqueza das nações. Contudo, advertiu que os indivíduos são ativos, que não podem ser obtidos por domínio ou mesmo por comércio. Entretanto, ao investirem em si, os indivíduos podiam aumentar a possibilidade de auferir maior renda e, conseqüentemente, melhorias no seu bem-estar. Para além disto, Schultz constatou falhas na noção clássica de os indivíduos serem dotados igualmente das mesmas capacidades, sendo justamente as diferenças as responsáveis por fazerem com que países fossem mais ou menos desenvolvidos, de modo que, países mais avançados investiam mais em capital humano, o que lhes conferia maior competitividade no processo de produção. Ou seja, as diferenças de desenvolvimento e de rendimentos eram produto dos investimentos em educação, atuando de maneira direta na dissemelhança étnica e espacial.

Schultz (1961) ainda arguia que um dos fatores que intensificava a diferença entre países desenvolvidos e não desenvolvidos consistia nos investimentos externos diretos (IED). Nações em desenvolvimento recebiam esses investimentos em infraestrutura, máquinas e equipamentos, não provendo investimentos em capital humano. Isso intensificava ainda mais as disparidades regionais e entre a capacidade instalada e a capacidade de utilização da infraestrutura disponível, demonstrando a importância não só do capital humano geral (leitura, escrita, raciocínio lógico, operações matemáticas etc.), mas do capital humano específico, como, por exemplo, treinamentos. O autor argumenta que a educação era a principal encarregada pela melhoria de bem-estar na qualidade de vida da população (política, institucional etc.), sobretudo quando se tratavam de países pobres e em desenvolvimento.

Becker (1962) aprofundou-se também na compreensão de capital humano e determinou que as habilidades e a educação são complementares, de forma que indivíduos mais habilidosos investiam mais em educação, o que lhes proporcionava um aumento significativo do salário durante a vida profissional, e isso fazia aumentar ainda mais a disparidade de renda em relação às pessoas menos habilidosas. Verificou também que os indivíduos com maior escolaridade, em média, recebiam salários maiores e quando deslocavam-se para países em desenvolvimento esses ganhos eram ainda mais expressivos (BECKER, 1994)

Mincer contribuiu de formas mais incisivas para a teoria do capital humano. Para entender a diferenciação de salários entre os indivíduos, formulou um modelo econométrico, usando variáveis instrumentais (VI), para estimar os impactos que a educação formal recebida pelos indivíduos teria sobre as suas próprias rendas auferidas. Os resultados alcançados apontaram que a assimetria dos salários recebidos pela força de trabalho era provocada pelas diferentes proporções de aquisição de escolarização (MINCER, 1970; 1974).

Lucas (1988) articulou importantes argumentos para o debate acerca das fontes do crescimento econômico, em especial, sobre a adoção de novas tecnologias, entendendo que a utilização de novas técnicas produtivas permite retornos constantes de escala (RCE) e, conforme o efeito transbordamento, gera externalidades positivas com o investimento no insumo capital humano (formação, capacitação e treinamento), que é um insumo não rival. Assim, diagnosticou que políticas que estimulavam o aumento da poupança e da qualificação profissional apresentavam impactos positivos diretos no crescimento da economia.

Mankiw, Romer e Weil (1992) apresentaram reforços teóricos ao modelo de crescimento de Solow (1956), ao verificarem que o crescimento da economia norte-americana era significativamente maior que o crescimento conferido pelo capital e pelo trabalho, atribuindo esse excedente a fatores exógenos ao modelo proposto, como o progresso tecnológico e outros fatores externos. Os autores concluíram, assim como Solow, que as diferenças de poupança, educação e crescimento eram as principais responsáveis pelas diferenças entre os países, indicando que políticas educacionais, por exemplo, poderiam diminuir essas disparidades.

Mankiw, Romer e Weil (1992) incluíram no modelo de Solow a tecnologia e o capital humano, que eram considerados fatores exógenos. Esses elementos, somados aos avanços dos conhecimentos dos métodos de estimação, trouxeram a seguinte reflexão: quando a taxa de poupança de um país está elevada, deve-se ao fato de a renda estar mais elevada, o que, conseqüentemente, é efeito direto do capital humano utilizado. Assim, um maior rendimento estimula os indivíduos a investirem mais na sua capacitação, constituindo-se de um processo que se retroalimenta. Logo, simplificando a realidade, pode-se afirmar que o crescimento econômico está intimamente ligado ao capital humano, sob a perspectiva dos autores.

No próximo capítulo, parte-se para revisão bibliográfica, que traz as principais reflexões sobre o tema contemporaneamente e que foram essenciais para determinação das variáveis que foram utilizadas no presente estudo.

### 3 REVISÃO DE LITERATURA

Mais recentemente, a ampla disponibilidade de banco de dados e os avanços dos métodos quantitativos têm permitido aos pesquisadores, cada vez mais, melhores condições para avaliar vários aspectos de uma grande variedade de fenômenos econômicos, como a relação entre as melhorias nas condições educacionais, o desempenho escolar por parte dos alunos e o crescimento econômico. Dessa forma, neste capítulo são apresentados estudos que analisaram os efeitos proporcionados pelo *background* familiar, pela infraestrutura física do ambiente escolar e pelas características dos docentes sobre o desempenho dos alunos.

A função produção expressa, por si, relações de entradas (matéria prima, tecnologia etc.) e saídas (produtos e serviços), de maneira que os *inputs* sejam combinados de formas diversas, a fim de verificar/obter os *outputs* possíveis. Como resultado, através do 1º relatório Coleman (1966), considerado um marco para avaliação de políticas públicas da educação, verificou-se a adoção do conceito da função produção educacional, que expressou o resultado atingido pelos alunos em testes padronizados em função das características do aluno, do seu *background* familiar e da escola.

Hanusheck (2019) analisou os impactos que as metodologias de avaliação educacional causaram na literatura especializada. Como principais resultados: mostrou que, nos Estados Unidos, a responsabilização dos agentes (formadores de políticas públicas, professores e diretores) tem relação fundamental com o resultado em testes padronizados da educação básica e que houve uma elevação do bem-estar econômico individual desde que foram adotados sistemas de responsabilização dos professores quanto ao desempenho dos alunos em testes padronizados.

Belmonte *et al.* (2019) exploraram os efeitos que os investimentos na infraestrutura física escolar promoveram no desempenho dos alunos das escolas públicas italianas. Para isso, utilizaram, através do método econométrico de diferenças em diferenças, os dados das escolas pouco afetadas pelo terremoto em 2012 e que receberam recursos para reformas, direcionando-os para melhoras na infraestrutura física. Os autores verificaram que, a cada 100 euros extras por aluno, aumentaram os resultados em testes padronizados de matemática em 4,6% e 1,2% em italiano. Esses resultados demonstraram que investimentos adicionais foram responsáveis por 10% das diferenças dos resultados dos testes na disciplina de matemática.

Oliveira, Menezes-Filho e Komatsu (2018), recorrendo ao Índice de Qualidade Institucional Municipal (IQIM) e ao IDEB dos municípios brasileiros entre 2009 e 2012, estudaram a qualidade da gestão municipal, usando o modelo de dados em painel por efeitos



fixos, como fator determinante no desempenho da educação básica. Demonstraram que a gestão municipal tem correlação positiva com o IDEB e com as notas da Prova Brasil, havendo mais influência nos primeiros anos da escolarização, visto que essa etapa do ensino é municipalizada. Concluíram, então, que a maneira como é conduzida a gestão municipal tem relação positiva com a qualidade da educação básica.

Scriptore, Azzoni e Menezes-Filho (2018) analisaram os efeitos do saneamento em relação ao desempenho atingido pelos alunos em provas padronizadas sob metodologia de dados em painel e em painel espacial, entre 2000 e 2010. Os resultados apontaram que a variação de 1% de saneamento provocou um aumento de 0,11% no desempenho médio desses alunos. Além disso, verificou-se que o saneamento básico provocou efeitos para muito além da diminuição da morbidez, de modo que a coleta e o tratamento de esgoto e o abastecimento de água potável melhoraram a frequência escolar, diminuindo as chances de evasão e repetência, assim como a distorção idade-série no ensino fundamental.

Coenen *et al.* (2018) analisam a literatura empírica internacional sobre as características dos professores que afetam o desempenho dos alunos em testes padronizados. O resgate apontou que a escolaridade adicional dos professores, principalmente os que lecionam matemática e ciências exatas, foi significativa para o resultado em testes, observando que o nível de certificação estava correlacionado com o desempenho, de modo que professores com mestrado tinham alunos com resultados mais altos que professores com especializações. Em relação à experiência, observaram que esta é relevante, porém que perde efeitos com o tempo, dependendo do método utilizado.

Tavares, Camelo e Paciência (2018), utilizando-se dos dados do Censo Escolar e da Prova Brasil de 2011 das escolas públicas urbanas, determinaram os fatores que geram a desigualdade de oportunidades de aprendizado, distinguindo entre circunstâncias e esforços. Os resultados apontam que mais da metade dos fatores que colaboram para equalização do desempenho escolar não depende das decisões sob controle da escola e, sim, da rede escolar (50% a 60%), de modo que de 10% a 15% dessa assimetria de notas foi sensível aos efeitos das políticas educacionais executadas pelo MEC; que a comunidade, o tamanho da escola e as etapas do ensino oferecidas pelo estabelecimento ajudaram a diminuir essas dessemelhanças; que entre 10% e 23% do *background* familiar exerceu influência nas notas dos alunos, de modo que as escolas com maior percentual de alunos brancos e com maior nível socioeconômico obtiveram notas mais altas, assim como escolas com maior percentual de alunos (sexo masculino) e com atraso escolar receberam notas mais baixas. E, finalmente, que as notas dos alunos tiveram correlação positiva com a experiência dos professores.

Franco e Menezes-Filho (2017) analisaram a evolução da proficiência das escolas públicas e privadas brasileiras nas provas do SAEB entre 1997 e 2005. Para isso, utilizaram a metodologia de dados em painel, pela qual foram construídos dois painéis: um de 1997 a 2005 e outro de 1999 a 2005. O modelo utilizou características das famílias, professores, diretores e escolas, controlando os efeitos não observáveis como fixos. Como principais resultados, verificaram que as características socioeconômicas dos alunos foram responsáveis por explicar grande parcela do desempenho na prova; observaram ainda que a infraestrutura escolar não foi significativa estatisticamente, mas ressaltaram que esse resultado pode ter se dado devido ao fato de o espaço temporal com o qual trabalharam ter sido limitado a uma amostra de poucos anos. Para além desses resultados, inferiram que o desempenho dos alunos é maior nas escolas que possuem um turno superior a cinco horas-aula e que quanto maior o número de matrículas na 4ª série, menor o desempenho médio dos alunos. O estudo ressalta, ainda, que as escolas privadas diferem em até meio ponto na proficiência *vis-à-vis* às escolas públicas.

Silva (2017) teve como propósito analisar os efeitos que a arrecadação de ICMS (Imposto sobre Circulação de Mercadorias e Serviços) exerceu no IDEB em 32 municípios de São Paulo entre 2005 e 2015. Para isso, utilizou a metodologia de dados em painel controlado por efeitos fixos. Os resultados apontam que as despesas com saneamento e a remuneração média dos professores possuem efeitos estatisticamente significativos e positivos no IDEB, assim como a arrecadação da cota-parte do ICMS, de modo que quanto maior a arrecadação, maior o índice. Esse fato é importante, visto que o cálculo do IDEB utiliza o SAEB em sua construção. Dessa maneira, para a amostra de municípios que possuem maior índice de atividades econômicas, maior a possibilidade de obter melhores resultados em testes padronizados.

Vernier, Bagolin e Fochezatto (2017) tiveram como propósito estudar a existência de efeitos espaciais na educação através dos dados do SAEB de 2013 dos municípios brasileiros. Para isso, utilizaram, respectivamente, a técnica de análise exploratória de dados espaciais e a estimação de um *cross-section* espacial. Os autores averiguaram que a proficiência atingida nos testes sofria influência da região e das cidades vizinhas, ou seja, há transbordamentos da proficiência de uma região para seu entorno, sendo tais transbordamentos positivos ou negativos. Analisou-se, ainda, que professores mais instruídos (que possuíam especialização, pós-graduação, mestrado e/ou doutorado) conferiam um resultado bastante robusto ao desempenho dos estudantes.

Hong e Zimmer (2016) analisaram as relações entre os gastos de capital no desempenho dos alunos e nos gastos escolares entre 1996 e 2009, considerados os alunos da 4ª e da 7ª séries no Estado do Michigan (Estados Unidos da América). Os autores verificaram fracos resultados no curto prazo dos investimentos por aluno, observando que os efeitos positivos foram observáveis apenas a partir de 4 anos após os investimentos iniciais, havendo impactos positivos ao longo de até 10 anos. Os autores concluíram que o método mais eficiente para verificar os benefícios na sociedade é aguardar que esses alunos cheguem à fase adulta, visto que existem relações positivas entre as despesas e o desempenho.

Vernier (2016) com objetivo de verificar a existência de efeitos espaciais na educação brasileira utilizou a metodologia de painel espacial nos dados dos SAEB de 2011 e 2013. Os resultados apontaram que o desempenho escolar está correlacionado de forma significativa com a etnia e a educação dos pais. Para além desses, as cidades que detiveram maior percentual de alunos homens apresentaram pior proficiência. Já a qualificação do professor foi significativa estatisticamente nas estimações, com exceção da regressão por efeitos fixos.

Amâncio-Vieira *et al.* (2015), utilizando os dados do SAEB e da Secretaria Municipal de Educação de Londrina para o ano de 2011, analisaram a qualidade da educação municipal de Londrina – PR, enfocando, especialmente, as influências dos custos operacionais diretos no desempenho dos alunos do ensino fundamental. Para tanto, recorrendo à análise fatorial e à regressão por MQO, apontaram que os custos administrativos, a quantidade de professores graduados e pós-graduados e a experiência do professor influenciaram o desempenho escolar dos alunos em cerca de 65%, estando aproximadamente 34% desse desempenho atrelado à infraestrutura escolar.

Dos Santos e De Tolentino-Neto (2015) estudaram o desempenho médio das escolas gaúchas na prova de matemática do SAEB, para os estudantes do 4º ano e 9º ano do ensino fundamental e da 3ª série do ensino médio, entre 2005 e 2013. Como principais resultados verificaram: que houve tímidas oscilações entre a performance dos alunos no período estudado; que, em geral, houve aumento das médias dos alunos que já haviam realizado o teste em edições anteriores; que o estado obteve rendimento maior que a média nacional; que, de maneira geral, os alunos do Rio Grande do Sul obtiveram notas médias baixas, quando comparadas com as edições anteriores, em especial, nos anos finais do ensino fundamental e no ensino médio, apesar de o ensino fundamental, anos iniciais, apresentar uma leve tendência de crescimento da média.

Luz (2014), com objetivo de verificar os determinantes das desigualdades de desempenho dos alunos das escolas públicas municipais entre 2007 e 2009, através de um

painel hierárquico por efeitos fixos dos dados da Prova Brasil, refuta, parcialmente, os efeitos da qualidade e da equidade educacionais verificados por Franco *et al.* (2007). Desse modo, Luz (2014) verificou que os mesmos fatores que geraram qualidade da educação promoveram equidade educacional. Identificou também que existe alta correlação entre o déficit escolar e a renda familiar e que esses efeitos provocaram transbordamento para seus vizinhos, observados sob modelos hierárquicos. Contudo, a análise sobre dados em painel simples observou-se que um melhor desempenho e igualdade são inversamente correlacionados.

Chevalier *et al.* (2013) analisaram o impacto da renda e da educação dos pais sobre a evasão escolar, utilizando os dados do *Labour Force Survey* (Reino Unido) por MQO. Os resultados apresentaram que a educação dos pais, sobretudo a materna, é significativa para a permanência das crianças na escola. Além desses, a renda permaneceu relevante em todas as estimações. Isso aponta que, em famílias mais vulneráveis, os jovens têm maior probabilidade de abandonar a escola antes dos 16 anos. Os autores concluíram que há um efeito multiplicador de renda e de escolaridade final entre as gerações (de pais para filhos).

Gonçalves e França (2013), com os dados da Prova Brasil, do Censo Escolar, da FINBRA (Finanças do Brasil) e da STN (Secretaria do Tesouro Nacional), estudaram as características municipais das capitais brasileiras que afetaram a gestão educacional no ano de 2005. Para atingir essa finalidade, utilizaram regressão em três estágios. Os resultados apontaram que a rotatividade partidária é favorável à eficiência na gestão municipal da educação e que esse processo de descentralização da educação (do Governo Federal ao municipal) foi benéfico, apesar de não ter sido observado em todos os municípios. Os autores verificam, também, que a eficiência não se traduz diretamente em IDEB.

Soares Neto *et al.* (2013) avaliaram, em 2011, a qualidade da infraestrutura das escolas brasileiras de pequeno porte (escolas com até 10 turmas e 200 alunos), usando os dados do Censo Escolar. Os resultados apontaram que 84,5% dessas escolas apresentaram condições elementar ou básica, ou seja, escolas com pouca ou nenhuma condição de oferta de infraestrutura de ambiente escolar. Vale ressaltar que essa categoria de escolas atenderam a 11% das matrículas brasileiras na educação básica em 2011. Os autores também verificaram que as condições físicas das escolas afetaram os resultados da Prova Brasil, principalmente, nas regiões Norte e Nordeste, especialmente, nas suas áreas rurais. Além disso, verificaram que quase a metade dos alunos dessas escolas recebem benefício do Programa Bolsa Família, indicativo de que os alunos desses estabelecimentos enfrentam baixas condições socioeconômicas.

Franco e Menezes-Filho (2012), com objetivo de classificar as escolas através do desempenho médio de seus estudantes da 4ª série do ensino fundamental, utilizaram-se de regressões por MQO com os dados do SAEB, entre 1999 e 2003. O estudo concluiu, como principal resultado, que a proficiência média dos alunos, nas avaliações realizadas, estava correlacionada com o ambiente socioeconômico em que a instituição estava inserida, assim como o dos alunos que a frequentavam.

Oliveira e Soares (2012) apuraram os determinantes da repetência escolar por meio do Censo Escolar, no período entre 2007 e 2010, utilizando a metodologia de dados em painel por efeitos fixos. Os resultados apontam que uma boa infraestrutura física do ambiente escolar proporciona uma maior chance de não repetência e, pelos mesmos dados, verificam correlação entre as condições do ambiente escolar e uma melhor atuação em testes padronizados. Os autores verificaram, ainda, que os alunos foram mais suscetíveis à reprovação que as alunas; que os PNEs (portadores de necessidades especiais) são mais vulneráveis à reprovação, visto que têm necessidade de um transporte escolar adequado e que escolas municipais têm piores condições de prover ensino de qualidade a esse público.

Cadaval e Monteiro (2011) avaliaram o desempenho dos alunos da 8ª série do ensino fundamental das escolas brasileiras na prova do SAEB entre 2001 e 2005. Para isso, usaram o modelo de dados em painel por efeito fixo. Como resultados principais, identificaram que o fator mais significativo para explicar a proficiência dos alunos foi a educação dos pais, sobretudo, a educação da mãe. O estudo caracteriza que essa repercussão está associada ao fato de que há aumento da escolaridade média dos pais ao longo do tempo. Ressaltaram, também, que houve crescimento e quase universalização da educação, através da proporção de matrículas. Contudo, o aumento do acesso não promoveu repercussão escolar em termos de melhora na qualidade, verificando que, apesar das evoluções percebidas, os resultados não refletiram em melhores notas nos anos subsequentes (2003 e 2005).

Duarte (2009) analisou os fatores que influenciaram na rotatividade dos professores brasileiros, usando uma regressão *cross-section* para cada nível escolar avaliado pelo SAEB em 2003. Os principais resultados foram: que, na disciplina de matemática, na 4ª série do ensino fundamental, a proporção de alunos brancos afetou negativamente a permanência dos professores durante o ano letivo, divergindo dos resultados de Oliveira e Soares (2012) e Tavares, Camelo e Paciência (2018), bem como a falta de computadores para o desenvolvimento das atividades pedagógicas. Em contrapartida, possuir quadra esportiva afetou positivamente o desempenho dos alunos. Para a 8ª série do ensino fundamental, a proporção de alunos homens, diretores brancos, bibliotecas e cidades que possuíam menos de

200 mil habitantes tiveram efeitos negativos e estatisticamente significativos. Já para o 3º ano do ensino médio, cidades com até 200 mil habitantes tiveram efeitos negativos no desempenho dos alunos; todavia, sendo o diretor branco, o efeito foi considerado positivo no aprendizado escolar – ambas variáveis com coeficientes estatisticamente significativos. O estudo ainda apontou que, aproximadamente, 15% dos professores de matemática tinham outros empregos, além da docência, demonstrando a pouca atratividade da profissão de professor.

Alves (2008) analisou como políticas educacionais afetaram o desempenho escolar dos alunos da 4ª série do ensino fundamental e de crianças de dez anos das capitais brasileiras, de 1999 a 2003, com base no SAEB. Utilizando dados em painel hierárquico, estimou-se que o nível socioeconômico é determinante do desempenho médio dos alunos, de modo que quanto maior esse nível, maior o desempenho. Porém, a autora percebeu que a maior autonomia dos municípios em relação à distribuição das responsabilidades educacionais foi positiva para a permanência das crianças na escola. Contudo, essas políticas têm pouco ou nenhum efeito no desempenho verificado nesse período.

De Andrade e Soares (2008), usando os dados do SAEB de 1997 a 2003, analisaram o efeito escola sobre o desempenho cognitivo dos alunos. Para isso, recorreram ao modelo de regressão hierárquico em dois níveis: alunos e escola. Os resultados apontaram uma diminuição nas heterogeneidades no 5º e 9º anos do ensino fundamental e, quanto à 3ª série do ensino médio, observaram que de 9,5% a 31,2% das diferenças de desempenho são intraescolares. Isso levou os autores a concluir que políticas focalizadas em melhoras na infraestrutura física do estabelecimento escolar têm efeitos positivos no desempenho dos alunos.

Biondi e Felício (2007) analisaram os fatores escolares que exerceram efeitos no desempenho dos alunos da 4ª série do ensino fundamental da rede pública brasileira entre 1999 e 2003. Usaram o modelo de regressão múltipla e de dados em painel por efeitos fixos das características dos alunos, dos professores, dos diretores e das condições escolares em relação ao desempenho no SAEB. De modo geral, verificaram que a infraestrutura escolar e a qualificação dos professores foram significativas na performance dos alunos nas avaliações do SAEB. As escolas que possuíam laboratório e bibliotecas obtiveram uma pontuação média mais elevada, assim como quando os professores detinham o ensino superior completo, utilizavam computador como recurso pedagógico e possuíam mais de dois anos de experiência profissional. Contraintuitivamente, professores que se submetiam a treinamentos

na área de atuação de ensino e escolas que usufruíam de acesso à internet mostraram-se como fatores que tiveram efeitos negativos e significativos.

Franco *et al.* (2007) investigaram os fatores que conferiam a qualidade e a equidade da educação aos alunos da 4ª série do ensino fundamental, em 2001, por meio dos dados do SAEB. Os resultados indicaram alta correlação entre o aumento do desempenho escolar e o aumento da desigualdade dentro da escola, de maneira que enfatiza a complexidade e a importância de estudar o desempenho escolar. Jales (2010) observou resultados similares ao analisar o efeito da heterogeneidade da turma (diferentes etnias, condições socioeconômicas e gênero) no desempenho escolar para amostra do SAEB 2005. Assim, ambos os estudos verificam a necessidade de serem adotadas políticas públicas com enfoque na melhora da qualidade educacional, controlando aspectos que possibilitem maior equidade. E mais: Franco *et al.* (2007) apontaram que a existência de biblioteca melhora em até 4,4 pontos a performance dos alunos nos testes padronizados.

Hanushek (2006) teve como propósito estudar quais recursos escolares impactavam no desempenho escolar através de uma revisão bibliográfica dos países desenvolvidos e em desenvolvimento. Como principais resultados, percebeu-se que os fatores que mais afetaram o desempenho em avaliações padronizadas foram as idiosincrasias, constatando que a renda individual e o desempenho geral da economia nacional influenciaram diretamente nas habilidades adquiridas. O autor entendeu, por um lado, que a performance dos estudantes nos testes está diretamente correlacionada com o crescimento da economia, o que pode gerar impactos econômicos de longo prazo, como produtividade da mão de obra. Por outro lado, ressalta a importância que a atmosfera escolar, a infraestrutura física e a qualificação dos professores exercem para a melhora, tanto da qualidade do conhecimento adquirido pelos alunos, quanto do desempenho dos alunos nos resultados dos testes a que esses são submetidos, principalmente, em países em desenvolvimento.

Nesse sentido, Bezerra e Kassouf (2006) tiveram por objetivo analisar, usando o SAEB 2003, os fatores que afetam o desempenho escolar dos estudantes nas disciplinas de português e matemática da 4ª série do ensino fundamental nas escolas rurais e urbanas brasileiras. Para isso, usaram o método de regressão por mínimos quadrados. Como resultados, verificaram que as condições do ambiente escolar são significativas para a qualidade, determinando que é preciso ter, à disposição dos alunos, computadores, bibliotecas e laboratórios, pois esses relacionam-se, diretamente, com um melhor desempenho dos alunos em testes padronizados e que fatores distintos afetaram o desempenho das localidades rurais e urbanas.

Klein (2006) avaliou como estava ocorrendo o processo de evolução da educação brasileira comparando os dados do SAEB e da PNAD (Pesquisa Nacional por Amostra de Domicílios), entre 1992 e 2003. Entre os resultados, destacou que houve poucas mudanças relevantes para o ensino fundamental, anos finais, e para o ensino médio; que foi dada uma atenção maciça à universalização da educação, porém, não foram providas medidas necessárias para que o aluno concluísse as etapas do ensino; que a conclusão do ensino depende não só da capacidade de sua oferta e matrícula, depende de condições de qualidade para que o indivíduo consiga adquirir as habilidades necessárias de forma adequada; e que há diferenças de proficiência entre os alunos do ensino público e privado e, quando público, diferenças ligadas à rede de ensino, sendo municipal ou estadual. O autor ainda destacou que, mesmo controlando fatores como o nível socioeconômico e a idade, os alunos autodeclarados negros e pardos atingiram performances mais baixas nas avaliações, evocando a necessidade de atenção das políticas públicas. Mostrou-se, também, que professores que passam exercícios de casa exercem impacto no desempenho do SAEB, aumentando-se ainda mais esses resultados quando também os corrigem. Além desses, observou-se que quanto mais cedo os alunos entram para a escola, melhores são seus resultados, verificando-se a importância da pré-escola e das creches para aquisição de habilidades medidas no teste.

Lange e Topel (2006) verificaram que os lugares que apresentavam maior expansão da oferta de escolarização e crescimento da renda produziam melhor qualidade da educação, concluindo que o efeito social da educação era muito mais efetivo que os retornos individuais, de modo que a expansão da instrução provocava efeitos transbordamentos nos salários e em outros aspectos da vida humana.

Felício e Fernandes (2005) verificaram formas de melhorar o desempenho escolar no Estado de São Paulo. Para isso, avaliaram as desigualdades de desempenho escolar dos estudantes, que poderiam ser explicadas pelos insumos escolares, utilizando a decomposição pelo índice de Theil; em seguida, foi realizada uma regressão por MQO, utilizando os dados do SAEB para o ano de 2001. Os autores explanaram que de 65% a 71% das desigualdades do desempenho dos alunos foram atribuídas às idiossincrasias (características dos alunos e suas famílias) e que cerca de 30% da performance pode ser atribuída ao efeito escola. Para além disso, caso o aluno da 4ª série do ensino fundamental das escolas públicas estudasse nas melhores instituições do município, sua performance seria igual à atingida pela 7ª série do ensino fundamental ou coincidiria com os resultados apurados pelo ensino privado. Nas regressões, identificaram que a educação do pai e da mãe tem efeitos positivos e estatisticamente significativos e crescentes conforme maior escolaridade dos pais.



Castro e Sadeck (2003) analisaram a capacidade de financiamento público em educação dos estados e municípios brasileiros no ano de 2000. Com os resultados, destacaram a prioridade estrutural na política brasileira vinculada à educação, demonstrando que houve altos investimentos na educação básica. Foram transferidos, pela União e estados, até 314% do valor do orçamento de arrecadação municipal de impostos para o financiamento da educação municipal, em 2000. Cabe ressaltar, que esses valores representam o mínimo previsto constitucionalmente, entretanto, denunciam que a maior parte do valor dos recursos do MEC é direcionada ao financiamento do ensino superior, encurtando a possibilidade de equalização das desigualdades e inovações na educação básica. Ressaltam, ainda, que alguns estados têm capacidade muito maior de financiamento da sua educação, evidenciando disparidades aviltantes entre as regiões brasileiras. O estudo salientou que São Paulo investiu, por matrícula, cerca de R\$ 1.434,00, valor aproximadamente 4,5 vezes maior do que o investido pelo Maranhão. Destaca-se, ainda, que apenas São Paulo, Brasília, Rio Grande do Sul, Rio de Janeiro e Roraima atingiram valor acima de mil reais por matrícula. Em contrapartida, Maranhão, Pará e Bahia não atingiram nem a metade da média nacional de gastos por matrícula, que é equivalente acerca de R\$ 830,00.

Behrman e Rosenzweig (2002), com o objetivo de avaliarem o impacto da variação da escolaridade dos pais na educação das crianças, utilizaram os dados de gêmeos idênticos do *Minnesota Twin Registry* (Estados Unidos da América) entre 1983 e 1990. Os resultados apontam que a educação dos pais (sexo masculino) tem efeitos significativos, porém de baixo impacto, no aumento da probabilidade de a próxima geração dos filhos ter maior escolaridade e habilidades. Em contrapartida, a educação das mães teve efeito contrário, principalmente quanto às mães que possuíam alta instrução, na amostra, pois diminuíram o seu tempo em casa.

Ferrão *et al.* (2001) descrevem os objetivos e a importância das características computadas pelo SAEB para investigar as escolas eficazes e a qualidade da educação, de modo que as boas condições da infraestrutura escolar refletem-se em maior propensão das escolas em obter resultados melhores. Porém, observaram que as escolas que ofertam os ambientes escolares superiores foram alocadas onde as condições socioeconômicas da região também são mais altas, além de perceberem que houve diferenças nos resultados entre as unidades da federação.

Barros *et al.* (2001) investigaram os determinantes do desempenho educacional dos adolescentes e jovens das áreas urbanas do Nordeste e do Sudeste do Brasil, a partir do conjunto de dados da PNAD de 1996 e da PPV (Pesquisa sobre Padrões de Vida) de 1996/97.

Para isso, utilizaram-se de regressões por MQO, conseguindo apurar que as escolaridades dos pais, sobretudo da mãe, e dos professores são determinantes para o desempenho dos alunos, em especial, para o ensino fundamental, visto que no ensino médio essa importância cai significativamente. O estudo ainda sugeriu que, tão importante quanto a qualificação dos professores, são os recursos escolares, como livros, computadores disponíveis e funcionamento noturno, entre outros. Enfim, os autores demonstraram que os insumos escolares exercem efeitos diretos no desempenho educacional dos indivíduos.

Hanushek e Kimko (2000) exploraram as interações entre o crescimento econômico e a qualidade da educação e perceberam que essa última foi refletida de forma mais adequada quando se utilizava a proficiência escolar, visto que reduz os impactos da quantidade de educação. Outra verificação foi que a elevação dos gastos se mostrou altamente significativa para a qualidade da proficiência dos alunos.

Em síntese, os autores demonstraram que diversos aspectos exercem efeito no desempenho escolar nos mais diversos métodos de análise, de modo que os efeitos mais contundentes foram relacionados ao *background* familiar dos alunos. Contudo, observou-se que o efeito escola tem se destacado, e que as condições e as características do ambiente escolar têm se evidenciado eficientes para explicar a performance dos estudantes em testes padronizados. Além desses, destacou-se que a qualificação dos professores foi determinante da proficiência dos alunos, principalmente quando esses (professores) detêm atualização, pós-graduação e mestrado. Para além desses, mostrou-se: que o processo de responsabilização dos educadores e gestores tem sido cada vez mais positivo; que a qualidade da gestão municipal e a alocação eficiente dos gastos inferem diretamente no IDEB; que o saneamento tem relação tanto com a melhora das condições de saúde dos alunos, quanto com a diminuição da evasão e repetência escolar; e que há transbordamento espacial dos resultados do SAEB.

No próximo capítulo, parte-se para uma análise da estrutura do sistema educacional brasileiro e a evolução dos resultados do sistema de avaliação ao longo dos anos.

#### 4 FATOS ESTILIZADOS DO SISTEMA DE EDUCAÇÃO NO BRASIL

A Constituição Federal de 1988 garante a oferta da educação de forma pública e gratuita, de modo que todo cidadão, em qualquer idade, deve ter acesso facilitado ao sistema educacional, sendo imperativo dos 4 aos 17 anos de idade. Para além disso, o sistema de ensino nacional segue o padrão ISCED (*International Standard Classification of Education*). Isso permite que a UNESCO (*United Nations Educational, Scientific and Cultural Organization*) compare o desempenho educacional entre nações (BRASIL, 2020; UNESCO, 2012).

A educação básica brasileira consiste em três etapas: educação infantil, ensino fundamental e ensino médio. A educação infantil consiste no primeiro ciclo da educação básica, que comporta crianças de 0 a 5 anos de idade, sendo separadas em dois seguimentos: creche (0 a 3 anos) e pré-escola (4 a 5 anos). Esta última é essencial, pois pode auxiliar e nortear o desempenho educacional do aluno durante toda sua vida escolar (BRASIL, 2020). Klein (2009) verificou que o desempenho em provas padronizadas é maior quando o aluno começa a vida escolar na pré-escola, pois os alunos desenvolvem, por exemplo, habilidades sociais, cognitivas, físicas, linguísticas e espaciais.

O ensino fundamental consiste no segundo ciclo da educação básica, etapa esta separada em anos iniciais (6 a 10 anos) e finais (11 a 14 anos). A primeira etapa é responsável pela alfabetização, que consiste em apreender a leitura, a escrita, a história e os cálculos simples. A etapa seguinte é a responsável pelo alicerçamento do saber construído até então, mas com conteúdo mais sofisticado, com a inclusão das ciências naturais e humanas, além de outras habilidades importantes para formação do cidadão (BRASIL, 2020).

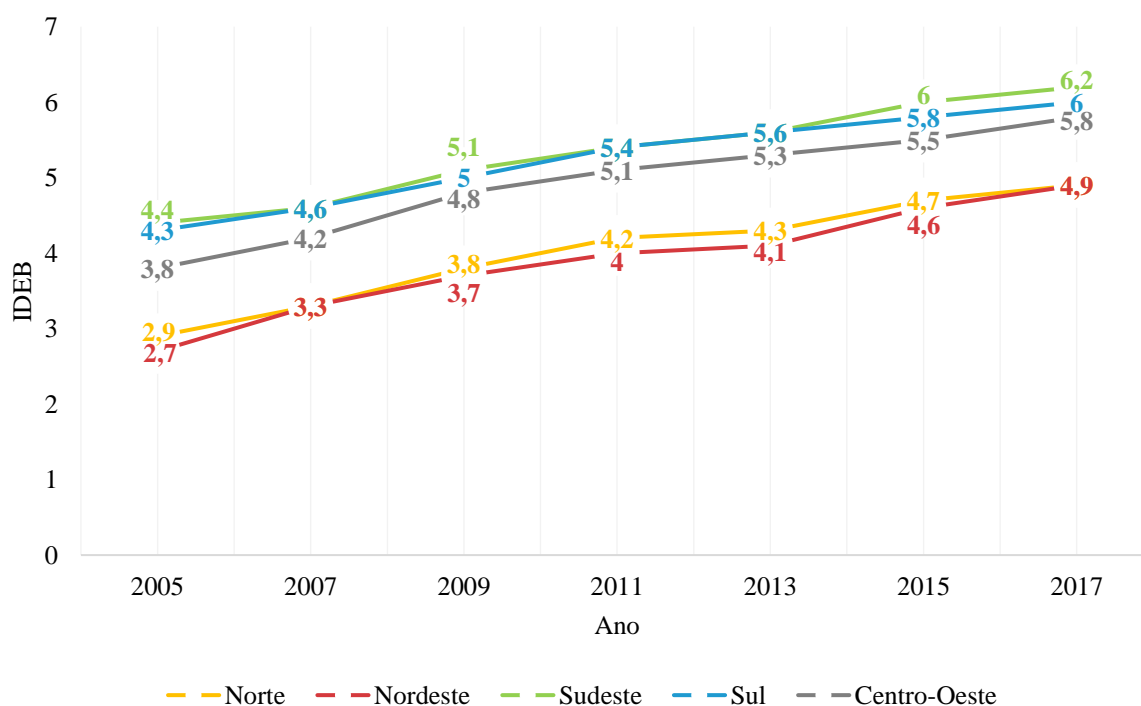
O ensino médio (15 a 17 anos), terceiro e último ciclo da educação básica, consiste em preparar os discentes para qualificação profissional (curso técnico e ensino superior), para reflexão crítica, além de prepará-los para o mercado de trabalho. Essa etapa é responsável pela consolidação e pelo aprimoramento dos métodos e técnicas construídas durante a vida escolar e pela profundidade em técnicas matemáticas/estatísticas, ciências da natureza, sociais e humanas, idiomas de língua inglesa e espanhol e métodos eficientes de leitura e interpretação de textos (BRASIL, 2020).

O sistema de educação brasileiro é periodicamente avaliado, a fim de verificar como ele tem se caracterizado nesses ciclos e como pode ser melhorado. Várias são as avaliações, como: Avaliação Nacional de Alfabetização (ANA), Avaliação Nacional da Educação Básica (ANEBC), Avaliação Nacional do Rendimento Escolar (ANRESC) - ou Prova Brasil, ENEM

(Exame Nacional do Ensino Médio), ENCCEJA (Exame Nacional para Certificação de Competências de Jovens e Adultos), PISA, TALIS (*Teaching and Learning International Survey*) e o SAEB.

Atualmente, o SAEB avalia 5º e 9º anos do ensino fundamental e o último ano do ensino médio. Essa avaliação é importante por contribuir para o cálculo do IDEB, em conjunto com o Censo Escolar e a Prova Brasil. Diante desse índice, surge o primeiro sinal de diferenças educacionais entre as regiões no Brasil no desempenho do IDEB. Essa verificação é importante, visto que é um dos sinais dos efeitos espaciais, conforme está testado no capítulo 6. Observa-se, conforme Gráfico 1, que no período de 2005 a 2017 a região Sudeste do Brasil teve um crescimento de, aproximadamente, 40%; o Nordeste, no mesmo período, apresentou um crescimento, ainda mais robusto, de 81,5%, mas situou-se 1,3 ponto abaixo do Sudeste; já o Centro-Oeste situou-se próximo ao verificado pelas regiões Sul e Sudeste, para ensino fundamental (anos iniciais).

Gráfico 1 - IDEB do Ensino Fundamental (anos iniciais) por região



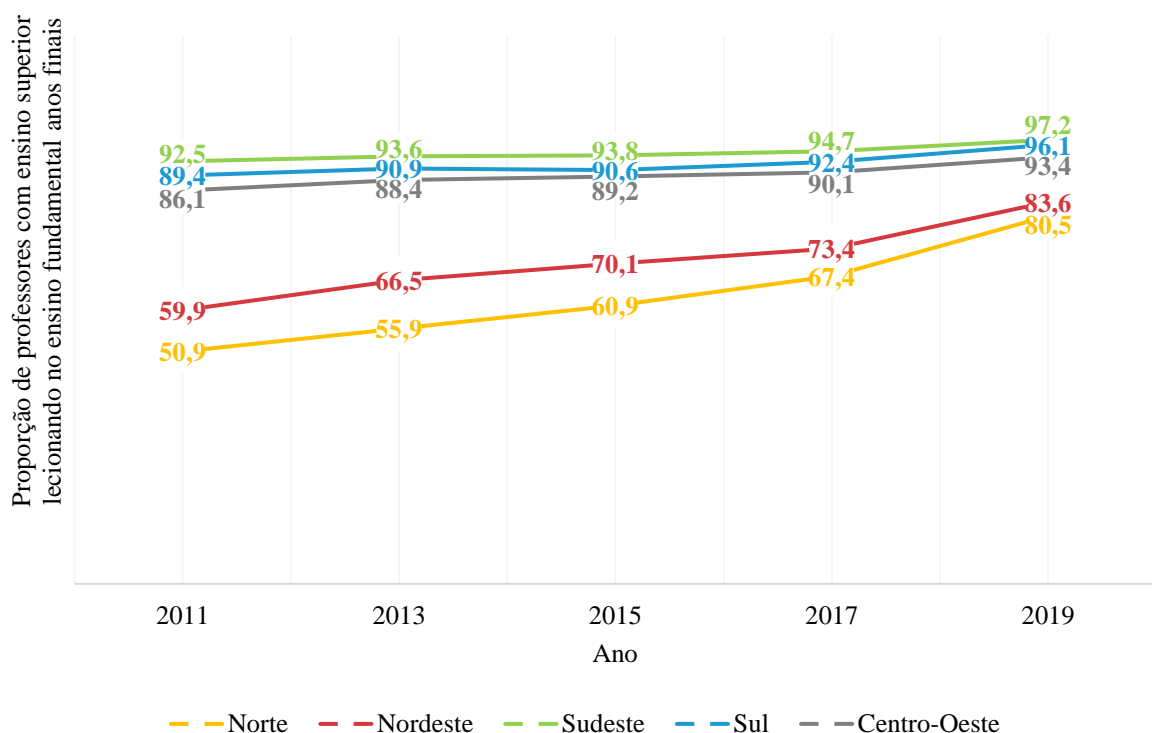
Fonte: IDEB (2020)

As regiões Sul e Sudeste atingiram o índice de 4,3 e 4,4 em 2005, respectivamente, para o ensino fundamental (anos iniciais), nas demais regiões verificou-se um IDEB menor, sendo 3,8 na região Centro-Oeste; 2,9 na região Norte e, no Nordeste, 2,7. Em 2017, notou-se

recrudescimento do IDEB de todas as regiões. Contudo, apenas as regiões Sul e Sudeste atingiram a média almejavável (IDEB, 2020).

Outro fato relevante a se observar é que a proporção de professores com ensino superior completo que se dedicam ao ensino fundamental mostra diferenças preocupantes entre as regiões. Essa diferença vem diminuindo de forma contínua, porém as regiões Norte e Nordeste ainda são dotadas de uma capacidade menor de oferta de professores qualificados com ensino superior em exercício, sendo esse índice significativamente menor do que nas demais regiões, haja vista que, em 2011, o Nordeste possuía apenas 50% do seu corpo docente com ensino superior; o Norte, com 60%; já o Sul, 89,4% e o Sudeste, 92,5%, encontrando-se, essas duas últimas regiões, já próximas da universalização de professores qualificados, desde 2011. Em 2017, houve expressiva melhora desse indicador nas regiões Norte (80,5%) e Nordeste (83,6%), porém, permanecendo, ainda, abaixo do conferido nas demais regiões, conforme retratado no Gráfico 2.

Gráfico 2 - Proporção de professores com ensino superior completo por região



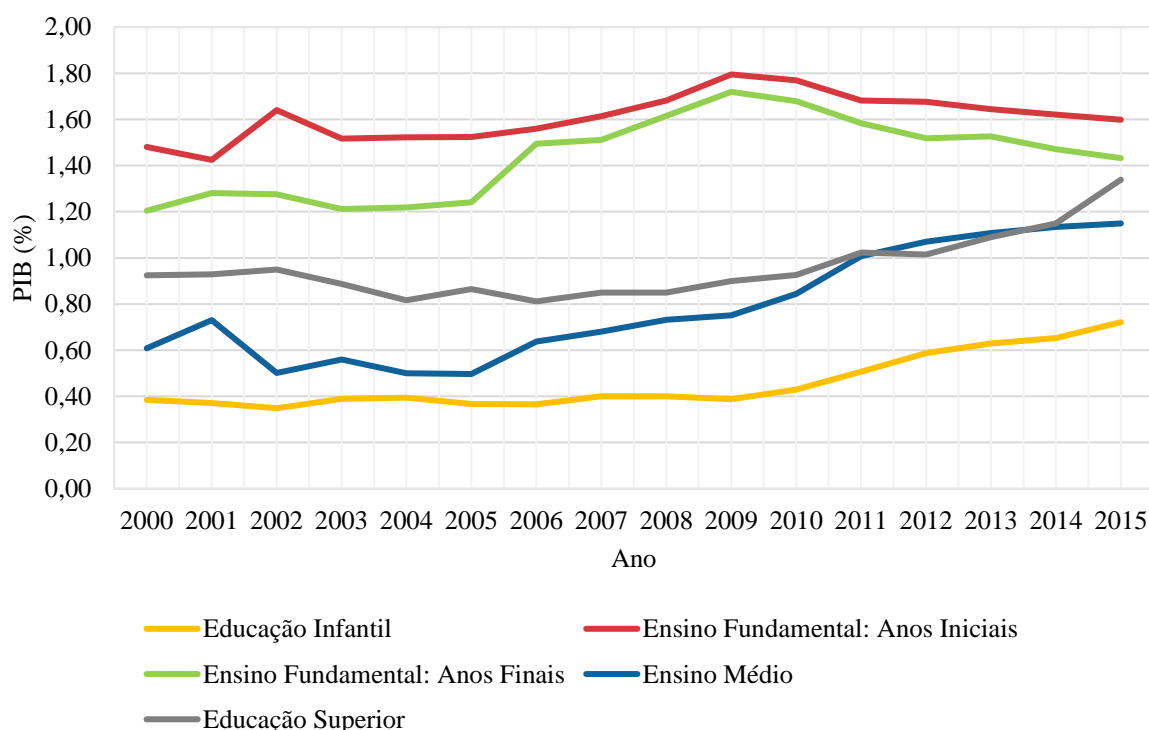
Fonte: INEP (2020)

Cabe ressaltar que a literatura abordada indica que os professores com ensino superior completo, principalmente os que possuem pós-graduação, são determinantes na qualidade da educação e no desempenho dos alunos (HANUSHEK; RIVKIN, 2006; VERNIER, 2016).

Assim, para além das disparidades regionais, quando se analisam os investimentos em educação, constatam-se, também, diferenças quanto aos valores empenhados em cada nível de ensino. Ou seja, quando avaliados pela perspectiva dos recursos públicos destinados à educação em relação ao Produto Interno Bruto (PIB), por nível de ensino, percebe-se o foco das políticas nacionais na promoção da educação básica e na sua universalização.

O Gráfico 3 apresenta a relação de gastos do PIB em educação por nível de ensino. Nota-se que grande parte dos recursos são direcionados para a educação básica, expressivamente no ensino fundamental, demonstrando o comprometimento das políticas públicas com o foco na educação básica. Entre 2000 e 2015, houve pouca variação do percentual de gastos por nível de ensino.

Gráfico 3 - Distribuição do investimento público em educação e por nível de ensino em relação ao PIB

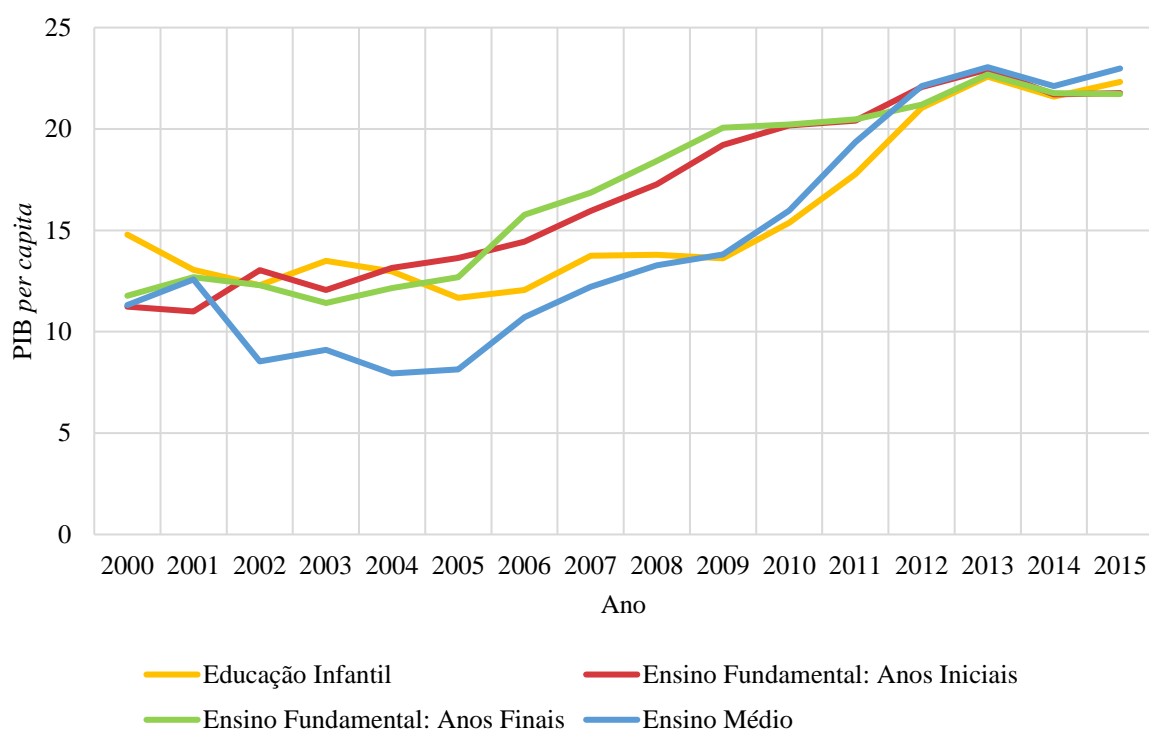


Fonte: INEP (2020)

Em contrapartida, sob a perspectiva do investimento público em relação ao PIB *per capita*, revela-se um foco altamente direcionado à educação superior. Em 2011, os

investimentos no ensino superior superavam em até 12 vezes o montante investido na educação básica em termos *per capita*. Esse resultado tem diminuído gradualmente, contudo, em 2017, constatou-se, ainda, uma diferença de 3,6 vezes mais investimentos na educação superior em PIB *per capita*. Para ilustrar isso, os Gráficos 4 e 5 apresentam os investimentos públicos em relação ao PIB *per capita* por nível de ensino e o Gráfico 6 mostra a relação PIB *per capita* entre a educação básica e superior.

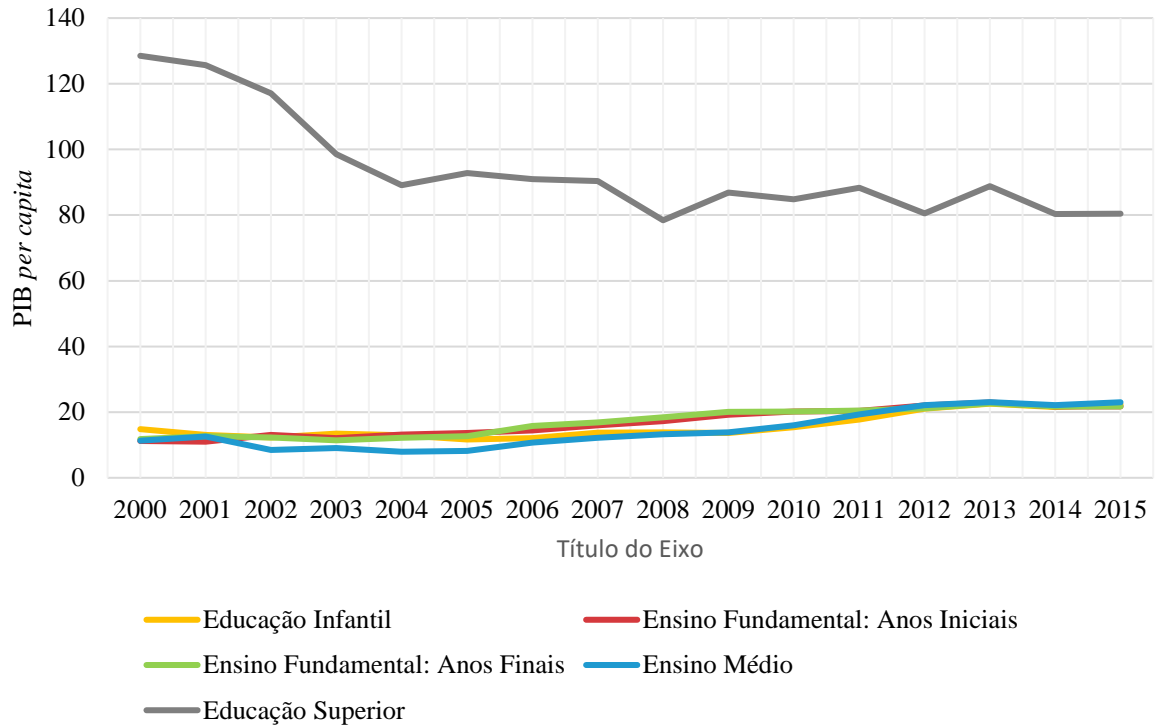
Gráfico 4 - Distribuição do investimento público em educação e por nível de ensino em relação ao PIB *per capita*, sem o ensino superior



Fonte: INEP (2020)

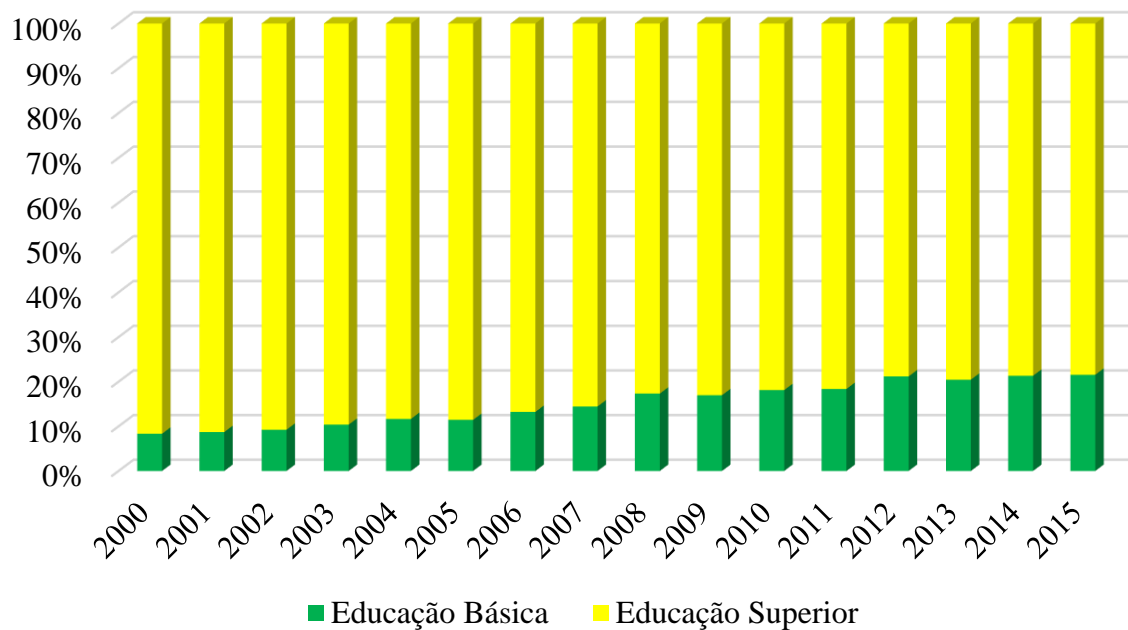
Neste capítulo, demonstrou-se que existem disparidades tanto regionais, quanto em relação aos investimentos por nível de ensino, de modo que os resultados apontam que as dessemelhanças entre as regiões são estruturais e verificadas ao longo do tempo (2000 a 2017). Isso indica um comportamento espacial dos dados, sendo necessário levar esse fato em consideração. No próximo capítulo parte-se para metodologia utilizada, em que se identificam as técnicas espaciais, o modelo e os testes utilizados.

Gráfico 5 - Distribuição do investimento público em educação e por nível de ensino em relação ao PIB *per capita*, com o ensino superior



Fonte: INEP (2020)

Gráfico 6 - Distribuição do investimento público em educação e por nível de ensino percentual



Fonte: INEP (2020)



## 5 METODOLOGIA

Neste capítulo, são apresentados os princípios e as noções da econometria espacial, de dados em painel clássico e da construção de dados em painel espacial. Inicia-se com os métodos espaciais, que consistem na criação de matrizes de pesos espaciais, e o método de verificação da autocorrelação, a saber  $I$  de Moran global e local. Em seguida, são apresentados os procedimentos econométricos que contemplam modelos de dados em painel, os modelos de regressão espacial, modelos de dados em painel espacial e o modelo utilizado no presente estudo e, por fim, a fonte e a base de dados.

### 5.1 ASPECTOS DOS PROCEDIMENTOS ECONOMÉTRICOS ESPACIAIS

A econometria espacial traz *insights* que os modelos econométricos de regressão linear clássicos não detiveram. Os modelos espaciais levam em consideração a autocorrelação e a heterogeneidade espaciais, ou seja, diferenças provocadas pelo posicionamento geográfico de uma dada observação, no caso aqui estudado, os municípios, em um determinado momento do tempo. Com isso, pode-se estabelecer as influências que as características de um município conferem a si e aos demais (ALMEIDA, 2012), de modo que, caso seja ignorada a dependência espacial, pode-se obter estimativas ineficientes e tendenciosas dos parâmetros do modelo, além da invalidação das inferências estatísticas (CLIFF; ORD, 1981).

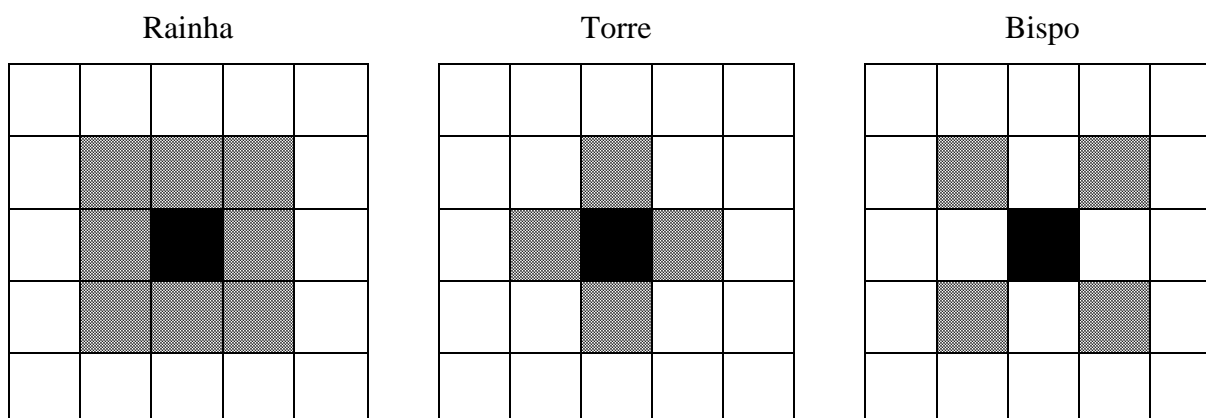
#### 5.1.1 Matriz de pesos espaciais

A dependência espacial pode ser percebida por contiguidades (vizinhança), distância inversa (quanto mais longe, menor são suas conectividades), distância social, distância econômica e  $k$  vizinhos (sendo  $k = 1, 2, 3, \dots, n$ ). Contudo, conforme considera a Primeira Lei da Geografia<sup>3</sup>, regiões mais próximas e limítrofes têm maior efeito de dependência que as demais mais distantes. Assim, a análise espacial perpassa pela representação dos dados espaciais na forma matricial, em que é considerada a noção de proximidade. Para tanto, as matrizes de contiguidades, conforme a Figura 1, podem ser construídas considerando as noções de rainha, torre e bispo, como definidas no jogo de xadrez (ALMEIDA, 2012).

---

<sup>3</sup>Tobler (1970) inferiu que todas as coisas estão relacionadas com todo restante, porém coisas mais próximas estão mais relacionadas do que coisas mais distantes.

Figura 1 - Representação das matrizes de ponderação espacial.



Fonte: Elaboração própria, com base em Almeida (2012).

No presente estudo, a escolha da matriz de ponderação espacial foi baseada no critério similar a Baumont (2004), que sugere regredir o modelo proposto por MQO – sob todas as possibilidades de matriz, obter-se os resíduos e realizar o teste do  $I$  de Moran e, então, escolher a configuração de maior valor do  $I$  de Moran estatisticamente significativo. Foram consideradas, neste estudo, as matrizes rainha, torre e distância inversa.

Definida a matriz de pesos espaciais (contiguidade), o passo a seguir é verificar se os dados seguem distribuição espacial. Se não há autocorrelação espacial, recorre-se, então, à metodologia econométrica linear clássica, caso contrário, a análise requer o uso dos recursos da econometria espacial. Dessa forma, inicia-se pela a AEDE (análise exploratória de dados espaciais), em que se averigua a existência de autocorrelação espacial e *clusters* de transbordamento espaciais. Assim, a seção a seguir apresenta a forma de mensurar a autocorrelação espacial usando as estatísticas de  $I$  de Moran local e global, em que, a partir do diagnóstico de autocorrelação espacial, justifica-se a estimação dos modelos econométricos espaciais (ALMEIDA, 2012; STAKHOVYCH; BIJMOLT (2009)).

### 5.1.2 Autocorrelação espacial

O Índice de Moran global ( $I$ ) univariado determina a dependência espacial linear entre os municípios. Essa estatística permite medir a proporção de desvios em torno da média da autocorrelação espacial e varia de  $-1 > I < 1$ . Valores próximos a zero indicam ausência de dependência linear, enquanto valores próximos à unidade indicam evidente autocorrelação espacial. O índice é dado, formalmente, por:

$$I = \left( \frac{n}{\sum_{i=1}^n W_{ij}} \right) \left( \frac{z'Wz}{z'z} \right) \quad (1)$$

em que  $n$  é o número de unidades espaciais;  $\sum_{i=1}^n W_{ij}$  é o somatório dos elementos da matriz de contiguidade, pelo qual  $j$  denota os valores dos vizinhos próximos da região  $i$ ;  $z$  é o vetor de  $n$  observações e  $W$  é a própria matriz de contiguidade (ALMEIDA, 2012).

Apesar de sua eficiência, a autocorrelação global pode negligenciar alguns padrões locais de autocorrelação espacial. Por essa razão, as pesquisas empíricas também utilizam-se do  $I$  de Moran local univariado ( $I_i$ ), também conhecido por LISA (*local indicators of spatial association*), o qual estabelece a autocorrelação para cada unidade analisada. Esse índice é dividido em quatro quadrantes: alto-alto (AA), baixo-baixo (BB), alto-baixo (AB) e baixo-alto (BA), sendo utilizado para determinar como o  $n$ -ésimo município se relaciona com os demais (ALMEIDA, 2012). A título de referência, emprega-se o  $I_i$  maior que zero, indicando que há semelhanças (quando maior que zero) ou dissemelhanças (quando menor que zero) entre si e sua vizinhança. Quando igual a zero, determina-se ausência de *clusters*. Essa estatística é dada por:

$$I_i = z_i \sum_{j=1}^J W_{ij} z_j \quad (2)$$

em que  $I_i$  é o índice de Moran local univariado;  $z_i$  é o valor da observação na região  $i$ ;  $\sum_{j=1}^J W_{ij} z_j$  é o somatório dos elementos da matriz de contiguidade, pelo qual  $j$  denota os valores dos vizinhos próximos da região  $i$ , multiplicado pelo vetor de observações próximas a região  $i$  e  $W$  como definido na última equação (ALMEIDA, 2012).

Além do índice de Moran local univariado, pode-se obter o  $I$  de Moran bivariado, o qual segue lógica similar. Contudo, a estatística bivariada indica o grau de associação linear entre o valor de uma variável no município  $i$  em relação a outra variável nos municípios vizinhos. A seguir serão descritas as noções de dados em painel, dados em painel espacial e o modelo proposto.

## 5.2 DEFINIÇÃO E ESPECIFICAÇÃO DOS MODELOS ECONÔMICOS

Neste estudo, conforme já sinalizado no capítulo 1, utiliza-se a modelagem de dados em painel, que engloba o universo de *cross-section* e séries temporais. Esse recurso econométrico permite aumentar os graus de liberdade, o que, conseqüentemente, aumenta a eficiência dos estimadores. Além disso, como o modelo foi expandido com a inclusão de

variáveis apontadas na literatura como possíveis efeitos estatísticos sobre o SAEB, tem-se um modelo empírico estruturado teoricamente, excluindo a possibilidade de omissão de variáveis relevantes, o que gera estimativas consistentes dos parâmetros. Também, como o teste de autocorrelação espacial aponta para este fenômeno, contempla-se, nos dados de painel clássico, a dependência espacial. Essa dependência não observável se manifesta no intercepto, nos parâmetros de inclinação e no vetor de resíduos do modelo.

Ao fazer uso de dados em painel espacial é importante se ter painéis balanceados, pois esses estão sujeitos à sensibilidade de dados faltantes. Para além disso, faz-se necessário interpolar as unidades por meio de recursos como áreas mínimas comparáveis (AMC), o que permite elaborar comparações intertemporais de áreas, como municípios que mudaram de área por divisão ou aglomeração ao longo do tempo (ELHORST, 2014).

### 5.2.1 Modelagem clássica de dados em painel

Os modelos de dados em painel clássico sem dependência espacial podem ser organizados em: POLS (*pooled ordinary least square*) e efeitos não observados. O modelo de estimação *pooled* contém uma constante comum, ou seja, não há heterogeneidades, possuindo estimativas consistentes e eficientes por MQO. Quanto aos modelos de efeitos não observáveis, esses são acondicionados em dois tipos: os efeitos fixos (FE) e os efeitos aleatórios (RE). A modelagem dos FE capta as heterogeneidades existentes nos dados por meio dos interceptos, permitindo que essas heterogeneidades se correlacionem com os regressores do modelo. Greene (2012) aborda que a omissão dessa variável captada pelo intercepto torna a estimação *pooled* inconsistente e viesada. Já a modelagem de RE considera os efeitos não observados como estocásticos e invariantes no tempo, de modo que a heterogeneidade indica diferentes termos de intercepto entre os indivíduos que não se correlacionam com os regressores, pelos quais seu termo constante é aleatório e incluído no termo de erro (resíduos), o que torna a estimação pelo método *pooled* consistente, porém ineficiente (GREENE, 2012).

O modelo geral de dados em painel *pooled*, já trazendo para a realidade do presente estudo, é especificado como:

$$SAEB_{it} = Z_i\alpha + X'_{it}\beta + \varepsilon_{it} \quad (3)$$

em que  $Z_i$  é o vetor das heterogeneidades individuais consideradas, neste caso, constantes ao longo do tempo;  $\alpha$  é o vetor de parâmetro da constante estimada;  $X'$  é o vetor transposto das variáveis independentes, conforme descrito no quadro 3;  $\beta$  é o vetor de parâmetros estimados

e  $\varepsilon$  é o vetor de resíduos não autocorrelacionados; o subscrito  $i$  indica os municípios e variam  $i = 1, \dots, 5561$  e o subscrito  $t$  indica a série temporal bianual, ou seja,  $t = 2011, 2013, 2015$  e  $2017$ . Ressalta-se que o modelo, representado pela equação (3), respeita os pressupostos do modelo clássico de regressão (média condicional zero do  $\varepsilon_{it}$ ; homoscedasticidade, independência entre as observações  $i$  e exogeneidade estrita de  $X_{it}$  (GREENE, 2012; ELHORST, 2003).

O modelo geral de dados em painel com FE é especificado como:

$$SAEB_{it} = \text{Dum}'_i \alpha + X'_{it} \beta + u_{it} \quad (4)$$

em que  $\text{Dum}'_i$  refere-se ao vetor transposto de variáveis *dummies* que reporta os FE;  $\alpha$  é o vetor de parâmetro dos FE estimado;  $u$ , o vetor de resíduos não autocorrelacionados (GREENE, 2012).

Ademais, pressupõe-se que os efeitos não observados estão correlacionados com as variáveis independentes, que os efeitos fixos não estejam autocorrelacionados no espaço e não estejam correlacionados com o termo de erro, respectivamente, isto é:

$$E[\text{Dum}X] \neq 0 \quad (5)$$

$$E[\alpha_i \alpha_j] = 0 \quad \text{caso } i \neq j \quad (6)$$

$$E[\text{Dum}\varepsilon] = 0 \quad (7)$$

Finalmente, o modelo geral de dados em painel com RE é especificado como:

$$SAEB_{it} = \alpha + X'_{it} \beta + u_i + \varepsilon_{it} \quad (8)$$

em que  $\alpha$  é o intercepto;  $u_i$  refere-se ao vetor de variável que reporta os RE; e  $\varepsilon_{it}$  é o vetor do termo de erro. Cabe ressaltar que:  $u \sim \sigma_u^2$  caracteriza o termo de erro aleatório que varia através dos agrupamentos, mas é invariante ao longo do tempo; já  $\varepsilon_{it}$  é o vetor de resíduos não autocorrelacionados (GREENE, 2012).

Pressupõe-se, ainda, a ortogonalidade, ou seja, que os efeitos não observados não podem se correlacionar com qualquer variável independente exógena; que o termo aleatório possua média zero e variância constante; que os efeitos aleatórios não estejam autocorrelacionados no espaço e variância constante dos efeitos aleatórios, respectivamente, representados por:

$$E[\alpha X] = 0 \quad (9)$$

$$E[\varepsilon] = 0 \quad E[\varepsilon \varepsilon'] = \sigma_\varepsilon^2 I_{nT} \quad (10)$$

$$E[\alpha_i \alpha_j] = 0 \quad \text{caso } i \neq j \quad (11)$$

$$E[\alpha_i \alpha_j] = \sigma^2 \quad \text{caso } i = j \quad (12)$$

Entretanto, quando há violação da hipótese de ausência de efeitos espaciais, faz-se necessário avançar para a modelagem de dados em painel espacial, tema da próxima seção, sob pena de reduzir a eficiência das estimativas. Contudo, é importante haver painéis balanceados, pois as estimativas estão sujeitas à sensibilidade dos dados faltantes (ELHORST, 2014; GREENE, 2012; ALMEIDA, 2012).

## 5.2.2 Modelagem de dados em painel espacial

Portanto, ao violar a hipótese de ausência de autocorrelação espacial  $\left( \begin{matrix} H_0: [\alpha_i \alpha_j] = 0 \\ H_a: [\alpha_i \alpha_j] \neq 0 \end{matrix} \right)$ , os clássicos modelos FE e RE não mais se sustentam. Para testar essa hipótese, usa-se a estatística do  $I$  de Moran (ELHORST, 2014; ALMEIDA, 2012).

A dependência espacial exige a remodelagem dos modelos de dados em painéis com a inclusão da defasagem espacial. Há várias especificações de painéis espaciais, sendo a mais geral representada por:

$$SAEB_{it} = \rho W_i SAEB_{it} + X_{it} \beta + W_i X_{it} \tau + \xi \quad (13)$$

$$\xi = \lambda W_j \xi_{it} + \varepsilon_{it} \quad (14)$$

em que,  $\rho$ ,  $\tau$  e  $\lambda$  são parâmetros espaciais da variável dependente, das variáveis independentes e do erro espacialmente correlacionado, respectivamente, situados no intervalo aberto entre -1 e 1;  $W_i SAEB_{it}$  é a defasagem espacial da variável explicada;  $W_i X_{it}$  representa a defasagem espacial das variáveis explicativas; e  $W_j \xi_{it}$  são os erros defasados espacialmente. Cabe ressaltar que  $W$  representa a matriz de ponderação espacial<sup>4</sup> (ELHORST, 2014; ALMEIDA, 2012).

Pressupõe-se que as matrizes de pesos espaciais podem ser diferentes, qualquer que seja a defasagem. Contudo, deve se ter cuidado, pois, caso seja especificada uma matriz de contiguidades de ordem diferente, por exemplo, segunda ordem, deve-se incluir apenas a segunda ordem, visto que, se especificada uma matriz de ordem superior, que inclua a ordem inferior, incorre-se nos erros de redundância e circularidade. Assim, se esse for o caso, faz-se necessária a diferenciação entre  $W_i$  e  $W_j$  (ALMEIDA, 2012; STAKHOVYCH; BIJMOLT, 2009).

---

<sup>4</sup> Detalhes dessa matriz estão na seção 5.1.1.

Dependendo das restrições impostas na equação (13), esta degenera-se nos seguintes tipos: modelo espacial não autorregressivo, modelos espaciais globais, modelos espaciais locais e modelos de dependência espacial de alcance global e local (ELHORST, 2014; ALMEIDA, 2012).

O modelo espacial não autoregressivo propõe defasagem espacial apenas nas variáveis independentes, sendo especificado como o modelo SLX (*Spatial model with Lag in X*) (ALMEIDA, 2012).

Em relação aos modelos espaciais globais, estes são os que utilizam um único parâmetro defasado espacialmente para captar a estrutura de correlação espacial, podendo ser segmentados em: modelo autorregressivo espacial – SAR (*Spatial Autoregressive Model*), modelo com erro autorregressivo – SEM (*Spatial Error Model*) e modelo autorregressivo com erro – SAC (*Spatial Autoregressive*) (ALMEIDA, 2012).

Quanto aos modelos espaciais locais, esses utilizam parâmetros defasados que variam continuamente no espaço, definidos como: GWR (*Geographically Weighted Regression*), SMA (*Spatial Moving Average Model*) e SLXMA (*Spatial model with Lag in X and Moving Average*) (ALMEIDA, 2012).

Já os modelos espaciais de alcance global e local podem ser determinados da seguinte maneira: SDM (*Spatial Durbin Model*), SDEM (*Spatial Durbin and Error Model*) e SARMA (*Spatial Autoregressive model with Moving Average*) (ALMEIDA, 2012).

Apesar dessa grande variedade de modelos espaciais, o foco do presente estudo se limita aos mais conhecidos, a saber: o SAR, o SEM e o SAC, visto que esses modelos englobam as principais defasagens globais que se almeja empiricamente encontrar diante da Política de Educação Infantil no nível municipal.

O modelo SAR surge ao se impor, na equação (13), primeiramente, as seguintes restrições:  $\rho \neq 0$ ;  $\tau = 0$  e  $\lambda = 0$ , de modo que:

$$SAEB_{it} = \rho W_i SAEB_{it} + X_{it} \beta + \varepsilon_{it} \quad (15)$$

em que  $\rho$  permite captar o efeito de interação acerca do quanto uma variação da média do SAEB de uma cidade pode afetar as regiões vizinhas (ELHORST, 2014; ALMEIDA, 2012).

Segundo, se  $\rho$  for positivo, aponta para autocorrelação espacial global positiva, ou que um valor alto da nota do SAEB nas localidades vizinhas aumenta o valor da nota no município  $i$  e vice-versa. Mas, se  $\rho$  apresenta-se negativo, isso indica autocorrelação espacial global negativa, pelo qual se tem efeito inverso ao anterior. Se a nota é alta na cidade  $i$ , as regiões vizinhas possuem notas baixas e vice-versa (ELHORST, 2014; ALMEIDA, 2012).

Trazendo para a realidade deste estudo, o modelo SAR é válido, visto que os agentes que promovem as políticas públicas tomam decisões no presente baseados no passado e na influência da tomada de decisões que outros agentes tomaram em períodos anteriores.

Da mesma forma, ao restringir a equação (13), fazendo  $\rho = 0$ ;  $\tau = 0$  e  $\lambda \neq 0$  na equação (16), define-se o modelo SEM, representado por:

$$\begin{aligned} SAEB_{it} &= X_t\beta + \xi \\ \xi &= \lambda W_i \xi_{it} + \varepsilon_{it} \end{aligned} \quad (16)$$

em que  $\lambda$  possibilita captar a dependência espacial dos resíduos caracterizados por efeitos não modelados ou por ausência de medidas apropriadas (ELHORST, 2014; ALMEIDA, 2012).

Cabe ressaltar que o padrão espacial não pode estar autocorrelacionado com as variáveis independentes. Assim, pode-se entender que o termo do erro do município  $i$  está correlacionado com a média  $j$  das regiões vizinhas, mais um resíduo aleatório (ELHORST, 2014; ALMEIDA, 2012).

Esse modelo surge da dificuldade em modelar todas as fontes da dependência espacial, manifestando-se, assim, no erro aleatório. Dessa maneira, entende-se que, além das influências já expostas que afetam a nota do SAEB, podem existir choques externos – por exemplo, o pleno emprego em uma determinada região –, que provoquem um multiplicador espacial, acarretando um efeito de dependência global, atingindo todas as regiões com diferentes intensidades, a depender da sua conectividade com núcleo de um acontecimento (ELHORST, 2014; ALMEIDA, 2012).

Finalmente, com as restrições  $\rho \neq 0$ ;  $\tau = 0$  e  $\lambda \neq 0$ , na equação (17), define-se o modelo SAC, representado por:

$$\begin{aligned} SAEB_{it} &= \rho W_i SAEB_{it} + X_{it}\beta + \xi \\ \xi &= \lambda W_i \xi_{it} + \varepsilon_{it} \end{aligned} \quad (17)$$

Conhecidas as especificações dos modelos anteriores, a saber equações (15) e (16), procura-se utilizá-los concomitantemente, de modo que tanto a nota do SAEB quanto um componente de erro aleatório possam ser defasados espacialmente. Isso, torna a modelagem mais próxima da realidade, pela qual tanto os fatores enfáticos da literatura quanto ocorrências não modeladas confrontam o empirismo.

Sabe-se que a equação 13 está representada de forma geral, com o indicador SAEB respondendo a um conjunto de regressores. A partir de então, especifica-se o modelo de dados em painel espacial no que concerne ao município utilizado no presente estudo para avaliar a



performance educacional em matemática dos alunos das escolas públicas do 9º ano do ensino fundamental. Verificam-se as condições da oferta do ensino público conforme a especificação a seguir:

$$\begin{aligned} \text{SAEB}_{it} &= \text{Dum}'_i\alpha + \text{CF}_{it}\beta + \text{CD}_{it}\beta + \text{CE}_{it}\beta + \rho W_i \text{SAEB}_{it} + \xi \\ \xi &= \lambda W_i \xi_t + \varepsilon_t \end{aligned} \quad (18)$$

$$\begin{aligned} \text{SAEB}_{it} &= \alpha + u_{it} + \text{CF}_{it}\gamma + \text{CD}_{it}\gamma + \text{CE}_{it}\gamma + \rho W_i \text{SAEB}_{it} + \xi \\ \xi &= \lambda W_i \xi_{it} + \varepsilon_{it} \end{aligned} \quad (19)$$

em que as equações (18) e (19) representam o modelo de FE e RE, respectivamente. Nota-se  $\text{SAEB}_{it}$ , conforme definido nas equações (13) e (14);  $\text{CF}_{mt}$  é o vetor linha de variáveis com características das condições dos alunos e de suas famílias no município  $i$  no ano  $t$ ;  $\text{CD}_{mt}$  é o vetor linha de variáveis com características das condições docentes do município  $i$  no ano  $t$ ;  $\text{CE}_{mt}$ , condições de infraestrutura escolar do município  $i$  no ano  $t$ , conforme descrito no quadro 3;  $\rho$  é o parâmetro espacial da variável explicada;  $\lambda$  é o parâmetro espacial do distúrbio aleatório;  $W_i$  é a matriz de pesos espaciais, sendo  $W_{ij} > 0$ , quando a região  $j$  tem alguma contiguidade com a região  $i$ ;  $\alpha$  é a constante;  $\text{Dum}'_i\alpha$  é o vetor de efeitos fixos;  $u_{it}$  é o vetor de efeitos aleatórios;  $\varepsilon$  são os resíduos; e  $\xi_{it}$  surge quando os distúrbios estão espacialmente autocorrelacionados (ELHORST, 2014; ALMEIDA, 2012).

Com a inclusão da dependência espacial no modelo de dados em painel, faz-se necessário entender três conceitos derivados da metodologia espacial, sendo estes o efeito direto, indireto e total. O efeito direto expressa os conhecidos efeitos estimados na econometria clássica, ou seja, é o efeito exercido por uma determinada variável sobre a variável dependente. O efeito indireto é o efeito transbordamento das variáveis dependentes e independentes sobre a variável dependente. Por exemplo, a variação de um ponto percentual na proporção de professores com mestrado na média das cidades  $j$  converte-se em uma variação na nota média da nota do SAEB na cidade  $i$ . O efeito total é o resultado do impacto direto e indireto (LESAGE, 2008; ELHORST, 2014).

Cabe ressaltar que, segundo Almeida (2012), caso o modelo apropriado seja por RE e a estimação é feita por FE, os parâmetros estimados permanecem consistentes, porém não são eficientes. Contudo, caso a metodologia adequada seja por FE, mas seja realizada estimação por RE, os coeficientes são todos ineficientes, viesados e inconsistentes.

Destaca-se que os modelos anteriores, equações (18) e (19), podem ser utilizados tanto para regressões com dados em painel com FE quanto para RE, respectivamente, sendo o método de estimação utilizado o método de máxima verossimilhança. Esse método

proporciona estimativas que possibilitam maximizar a probabilidade de se observarem os dados populacionais, de modo que assintoticamente os parâmetros são consistentes, eficientes e seguem distribuição normal (LEE, 2004; ELHORST, 2014).

Para obter-se a regressão, recorreu-se ao método de máxima verossimilhança. Primeiramente, estimam-se por MQO a variável dependente e as variáveis defasadas espacialmente em relação às variáveis independentes e, em seguida, obtém-se os resíduos de ambas as estimações. De posse das estimativas dos resíduos e da constante, esses são utilizados na função de log-verossimilhança, pela qual se detém os parâmetros espaciais e, em seguida, utilizam-se os parâmetros espaciais para calcular os demais parâmetros dos modelos estimados (LEE, 2004).

Definidos os métodos de estimação, o Quadro a seguir apresenta os sinais esperados pelos parâmetros das variáveis.

Quadro 1 - Sinais dos parâmetros à guisa da teoria

Variável	Sinal Esperado	Variável	Sinal Esperado	Variável	Sinal Esperado	Variável	Sinal Esperado	Variável	Sinal Esperado
CF <sub>1</sub>	+	CF <sub>8</sub>	+	CF <sub>15</sub>	+	CD <sub>5</sub>	+	CE <sub>7</sub>	+
CF <sub>2</sub>	-	CF <sub>9</sub>	-	CF <sub>16</sub>	+	CE <sub>1</sub>	+	CE <sub>8</sub>	+
CF <sub>3</sub>	-	CF <sub>10</sub>	+	CF <sub>17</sub>	+	CE <sub>2</sub>	+	CE <sub>9</sub>	+
CF <sub>4</sub>	+	CF <sub>11</sub>	+	CD <sub>1</sub>	+	CE <sub>3</sub>	+	CE <sub>10</sub>	+
CF <sub>5</sub>	+	CF <sub>12</sub>	+	CD <sub>2</sub>	+	CE <sub>4</sub>	+	CE <sub>11</sub>	+
CF <sub>6</sub>	-	CF <sub>13</sub>	+	CD <sub>3</sub>	+	CE <sub>5</sub>	+	CE <sub>12</sub>	+
CF <sub>7</sub>	-	CF <sub>14</sub>	+	CD <sub>4</sub>	+	CE <sub>6</sub>	+	CE <sub>13</sub>	+

Fonte: Elaboração própria

Na seção a seguir, são descritos os principais testes realizados para especificação do modelo econométrico de dados em painel e dados em painel espacial.

### 5.2.3 Relacionando os testes econométricos

Para correta especificação e verificação da robustez do modelo proposto, fazem-se necessários alguns testes, sendo estes: o teste de White, de Chow, de Breusch-Pagan, de Hausman e Hausman robusto e os testes CD de Pesaran.

O teste de White tem por objetivo verificar a existência de heterocedasticidade na distribuição de erros, de modo que se constitui em fazer uma regressão auxiliar dos resíduos ao quadrado sob todos os regressores, de modo a verificar se a variância dos resíduos é

constante. Assim sendo, o teste utiliza a estatística do multiplicador de Lagrange com distribuição  $\chi^2$ , sob a hipótese nula de homoscedasticidade e alternativa de heterocedasticidade, ao passo que o teste de Wooldridge faz o diagnóstico da presença de autocorrelação dos resíduos nos modelos RE ou FE. Sob a hipótese nula, verifica-se a não autocorrelação, e a não rejeição dessa hipótese identifica autocorrelação (GREENE, 2012).

Já o teste de Chow, a tradicional estatística  $F$ , destina-se a comparar o modelo *pooled* contra o modelo de FE. A não rejeição da hipótese nula indica o modelo *pooled* e vice-versa. Assim como o teste de Chow, o teste de Breusch-Pagan empenha-se em comparar os modelos *pooled* e RE, utilizando a estatística do multiplicador de Lagrange. A não rejeição da hipótese nula aponta para o modelo *pooled* e vice-versa. Já o teste de Hausman (1978) compara as estimativas dos modelos RE e FE. A hipótese nula tem como apropriado o modelo RE e a hipótese alternativa o FE. Assim, ao rejeitar hipótese nula, pela estatística  $\chi^2$  calculada, indica que o modelo de FE é mais apropriado. Já o teste de Hausman robusto consiste das mesmas hipóteses, contudo é calculado a partir da distribuição conjunta de *bootstrap*. O teste, em sua versão robusta é necessário caso seja violada a pressuposição de homoscedasticidade e não autocorrelação e correlação nos modelos estimados (GREENE, 2012).

Por fim, os testes CD de Pesaran (2004; 2015) visam aferir a dependência transversal nos modelos de dados em painel e a fraca dependência transversal (hipótese mais fraca). O teste consiste em verificar a correlação média entre as unidades de *cross-section* do painel. Sob a hipótese nula do teste há correlação entre as unidades de corte transversais. No quadro a seguir são resumidos os testes realizados:

Quadro 2 - Testes e suas hipóteses

Testes	Hipóteses dos Testes
Chow	$H_0$ : Modelagem <i>pooled</i> $H_A$ : Modelagem de FE
Breusch-Pagan	$H_0$ : Modelagem <i>pooled</i> $H_A$ : Modelagem de RE
White	$H_0$ : Homoscedasticidade $H_A$ : Heterocedasticidade
Wooldridge	$H_0$ : Não autocorrelação $H_A$ : Autocorrelação
Hausman robusto	$H_0$ : RE $H_A$ : FE
Pesaran CD	$H_0$ : Correlação entre os <i>cross-section</i> $H_A$ : Independência dos <i>cross-section</i>

Fonte: Elaboração própria

Definidos os testes, Almeida (2012) sugere os seguintes procedimentos para especificação de modelo: realizar o teste Breush-Pagan, com objetivo de verificar se no modelo especificado há heterogeneidades não observadas (FE ou RE); proceder o teste de Hausman (1978), ou do Hausman robusto, a fim de esclarecer qual é o modelo adequado para tratar os efeitos não observados, se FE ou RE; ajustar o modelo indicado pelo teste de Hausman, desconsiderando a dependência espacial; fazer o teste de  $I$  de Moran nos resíduos, com intuito de verificar a existência da dependência espacial. Caso não haja evidências dessa dependência, utilizar o modelo de regressão como painéis clássicos, caso contrário estimar o modelo de efeitos (fixo ou aleatório) com dependência espacial; finalmente, utilizar o critério de informação de Akaike (AIC) e Schwarz (SC) pelo menor valor, de modo a selecionar o melhor modelo.

### 5.3 FONTE E BASE DE DADOS

Foram selecionadas trinta e cinco variáveis independentes relacionadas aos resultados do SAEB. O presente estudo utiliza-se dos microdados do SAEB dos anos de 2011 a 2017 (dados bianuais), em conjunto com os dados do Censo Escolar dos respectivos anos nas escolas públicas brasileiras agregadas em municípios. Ambas as bases constituem-se de informações a nível de aluno, classe e escola, apresentando características que podem ser importantes para determinar o desempenho dos estudantes. Assim, o banco de dados é constituído por cerca de 1,5 milhão de observações ou, mais precisamente, 1.403.978 observações. Na constituição da base de dados, buscou-se contemplar todas as possíveis características relacionadas ao desempenho escolar dos alunos em matemática do 9º ano do ensino fundamental das escolas públicas brasileiras, em consonância com a literatura abordada no decorrer do presente trabalho e pela disponibilidade de dados relacionados à infraestrutura escolar, à titulação dos professores e aos aspectos socioeconômicos, detalhados no quadro 3.

Cabe ressaltar, que o método de dados em painel espacial, no software Stata 15, não permite a utilização de painéis desbalanceados, pois os dados faltantes podem gerar viés nas análises. Assim, os dados faltantes foram substituídos pela sua média temporal disponível. Em 2011, foram substituídas 133 observações da variável independente; em 2013, foram 123; em 2015, foram 286; e 45 em 2017. Contudo, nove cidades, sendo oito gaúchas e uma goiana, não detiveram observações para nenhum dos anos do estudo, dessa forma foram excluídas.

Quadro 3 - Nomenclatura e descrição das variáveis utilizadas

(continua)

Variável	Descrição
<b>Score municipal no SAEB (variável dependente)</b>	
<b>SAEB<sub>i</sub></b>	Média municipal da nota de matemática na avaliação do SAEB para o 9º ano do ensino fundamental
<b>Características/condições dos estudantes e do seu <i>background</i> familiar</b>	
<b>CF<sub>1</sub></b>	Proporção de estudantes do sexo masculino
<b>CF<sub>2</sub></b>	Proporção de estudantes negros, pardos e indígenas
<b>CF<sub>3</sub></b>	Proporção de estudantes que possuem um ou mais anos de distorção idade-série
<b>CF<sub>4</sub></b>	Proporção de estudantes que moram com a mãe
<b>CF<sub>5</sub></b>	Proporção de estudantes que moram com o pai
<b>CF<sub>6</sub></b>	Proporção de estudantes que dedicam uma hora ou mais aos afazeres domésticos
<b>CF<sub>7</sub></b>	Proporção de estudantes que trabalham (formalmente ou informalmente)
<b>CF<sub>8</sub></b>	Proporção de estudantes que ingressaram na pré-escola
<b>CF<sub>9</sub></b>	Proporção de estudantes que foram reprovados uma ou mais vezes
<b>CF<sub>10</sub></b>	Proporção de estudantes que fazem o dever de casa de matemática com frequência
<b>CF<sub>11</sub></b>	Proporção de estudantes que afirmaram que o professor corrige o dever de casa de matemática com frequência
<b>CF<sub>12</sub></b>	Proporção de mães que possuem o ensino médio completo
<b>CF<sub>13</sub></b>	Proporção de mães que possuem o ensino superior completo
<b>CF<sub>14</sub></b>	Proporção de pais que possuem o ensino médio completo
<b>CF<sub>15</sub></b>	Proporção de pais que possuem o ensino superior completo
<b>CF<sub>16</sub></b>	Proporção de responsáveis que frequentam reunião de pais e mestres
<b>CF<sub>17</sub></b>	Proporção de estudantes que praticam leitura de livros, jornais e revistas <sup>5</sup>
<b>Condições docentes</b>	
<b>CD<sub>1</sub></b>	Proporção de professores formados em matemática
<b>CD<sub>2</sub></b>	Proporção de professores com atualização em matemática
<b>CD<sub>3</sub></b>	Proporção de professores com mestrado
<b>CD<sub>4</sub></b>	Proporção de professores com doutorado
<b>CD<sub>5</sub></b>	Experiência média dos professores ( <i>gaps</i> de 5 anos)
<b>Condições de infraestrutura escolar</b>	
<b>CE<sub>1</sub></b>	Proporção de escolas que possuem água filtrada
<b>CE<sub>2</sub></b>	Proporção de escolas que possuem abastecimento de água da rede pública
<b>CE<sub>3</sub></b>	Proporção de escolas que possuem abastecimento de energia elétrica da rede pública
<b>CE<sub>4</sub></b>	Proporção de escolas que possuem coleta de esgoto da rede pública
<b>CE<sub>5</sub></b>	Proporção de escolas que possuem coleta periódica de lixo
<b>CE<sub>6</sub></b>	Proporção de escolas que possuem laboratório de informática

<sup>5</sup> Exceto revistas de comportamento, celebridades, esportes ou TV.

Quadro 3 - Nomenclatura e descrição das variáveis utilizadas

(conclusão)

<b>CE<sub>7</sub></b>	Proporção de escolas que possuem laboratório de ciências
<b>CE<sub>8</sub></b>	Proporção de escolas que possuem quadra de esportes (coberta ou descoberta)
<b>CE<sub>9</sub></b>	Proporção de escolas que possuem biblioteca
<b>CE<sub>10</sub></b>	Proporção de escolas que possuem sala de leitura
<b>CE<sub>11</sub></b>	Média municipal de computadores disponíveis na escola para uso dos alunos
<b>CE<sub>12</sub></b>	Proporção de escolas que possuem acesso à internet
<b>CE<sub>13</sub></b>	Proporção de escolas que oferecem alguma atividade complementar

Fonte: Elaboração própria.

No próximo capítulo, apresentam-se os resultados obtidos e suas respectivas análises.

## 6 ANÁLISE E DISCUSSÃO DOS RESULTADOS

É sabido que a educação é um dos alicerces da Economia, de modo que estudar os determinantes do desempenho dos alunos das escolas públicas pode promover fatores que geram qualidade da educação, eficiência dos gastos e desenvolvimento econômico. Assim, o presente estudo tem por objetivo analisar os determinantes do desempenho médio dos alunos das escolas públicas brasileiras em seu aspecto espacial. Com isso, neste capítulo, são apresentados os passos realizados através da análise espacial dos modelos econométricos espaciais e seus respectivos resultados.

Inicia-se esta análise com a verificação *I* de Moran global do desempenho municipal dos alunos das escolas públicas do 9º ano do ensino fundamental na prova de matemática do SAEB. Essa etapa é fundamental, visto que, se a interação espacial dos dados for verificada, o modelo de dados em painel clássico não pode ser utilizado, conforme detalhado no capítulo 5.

Para a análise de verificação da dependência espacial, primeiramente, é necessário escolher a matriz de pesos espaciais, sendo as mais comuns as matrizes rainha (*queen*), torre (*rook*) e distância inversa. Assim, o *I* de Moran foi calculado para cada ano e o critério de escolha da matriz de pesos foi definido a partir da maior estatística calculada estatisticamente significativa (valor em negrito). Essas informações estão resumidas na Tabela 1.

Tabela 1 - Escolha da matriz de contiguidade para o tratamento da variável dependente (Índice de Moran)

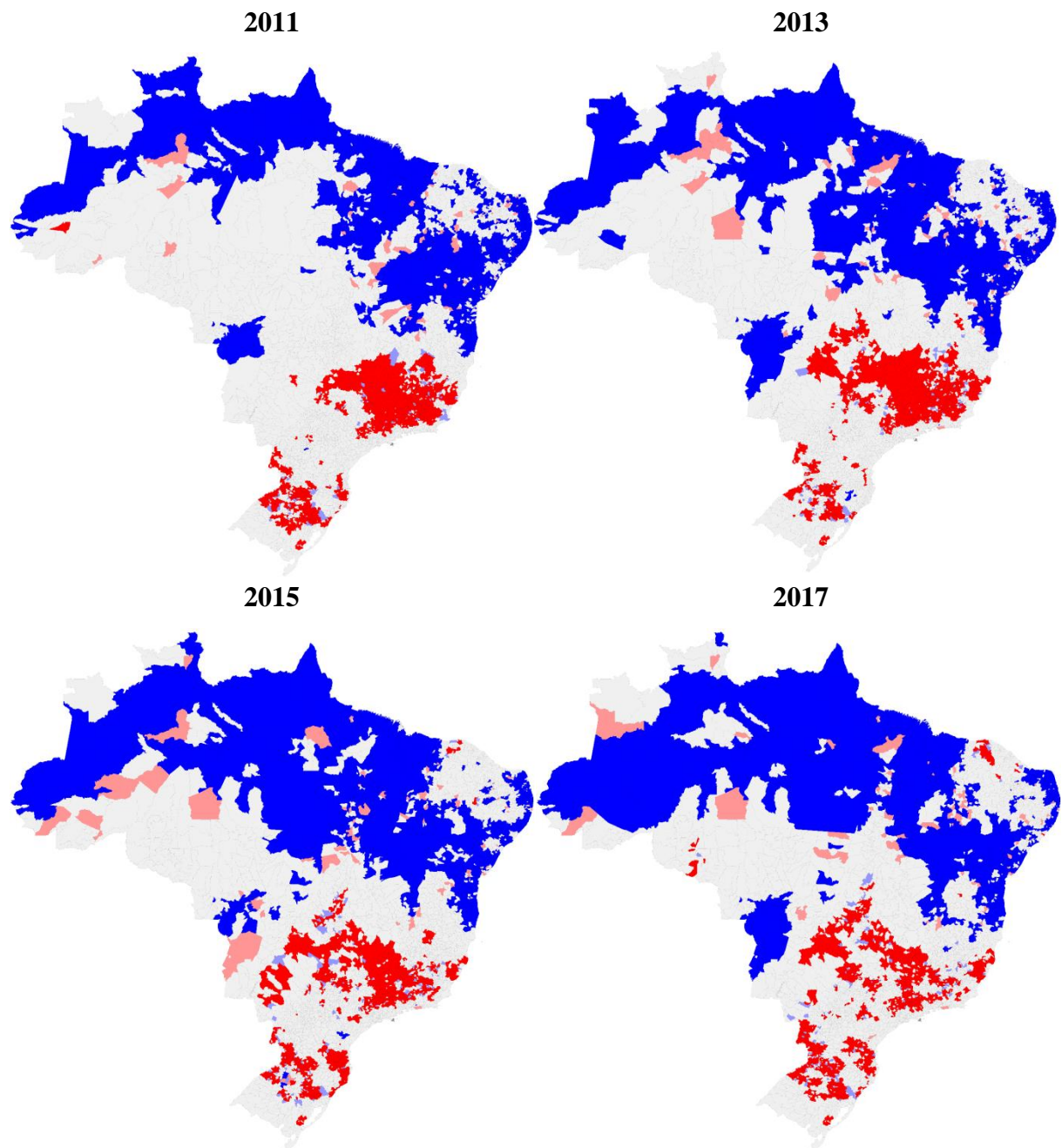
Ano	<i>I</i> de Moran (variável dependente)		
	Rainha <sup>(*)</sup>	Torre <sup>(*)</sup>	Distância Inversa <sup>(*)</sup>
<b>2011</b>	0,6653	<b>0,6665</b>	0,5117
<b>2013</b>	0,6393	<b>0,6399</b>	0,4827
<b>2015</b>	0,6090	<b>0,6097</b>	0,4398
<b>2017</b>	0,6100	<b>0,6111</b>	0,4416

Fonte: INEP (2020)

(\*) Nota: o nível de significância das estatísticas calculadas é dado por:  $-\left(\frac{1}{1-n}\right) = -0,0002$ , sendo *n* o número de municípios.

Observa-se que as estatísticas de *I* de Moran para as três matrizes de pesos foram significativas estatisticamente a 1%, tendo os seus valores bem próximos. A contiguidade do tipo torre apresentou o maior valor para os quatro anos da série. Logo, foi selecionada para interação espacial da variável dependente. A seguir, é apresentada a distribuição do *I* de Moran Local para os anos de 2011, 2013, 2015 e 2017.

Figura 2 - Índice de Moran local univariado da nota média de matemática municipal para o nono ano do ensino fundamental das escolas públicas



**Legenda**

Não significativo
  Alto-Alto
  Baixo-Baixo
  Baixo-Alto
  Alto-Baixo
  Ilhas

Fonte: INEP (2020)

Nota: Idealizado a partir do *software* GeoDa 1.14.0

As cores mais vibrantes na figura 2 indicam clusters significativos e a cor cinza claro indica ausência de clusters, ou seja, não significativos. Assim, a cor vermelha revela concentração Alto-Alto (AA); em azul, conglomerados Baixo-Baixo (BB); lilás e rosa representam Baixo-Alto (BA) e Alto-Baixo (AB), respectivamente. De modo geral, nos



quatro anos analisados, houve *clusters* de municípios significativos entre 42% e 44% da amostra. A tabela (2), a seguir, indica o número de cidades em cada agrupamento por ano:

Tabela 2 - Número de cidades nos *clusters* do *I* de Moran local

<b>Ano</b>	<b>AA</b>	<b>BB</b>	<b>BA</b>	<b>AB</b>	<b>Não Significativo</b>
<b>2011</b>	1103	1249	64	57	3086
<b>2013</b>	1117	1271	72	68	3031
<b>2015</b>	1045	1271	97	68	3078
<b>2017</b>	1005	1182	104	62	3206

Fonte: Elaboração própria

Nota: Idealizado a partir do software GeoDa 1.14.0

Na Região Norte, houve concentrações BB nos quatro anos analisados, em que os *clusters* se acumularam nas divisas nacionais e em seus limites com o Nordeste do país. Em 2011, identificaram-se aglomerados BB em algumas regiões limítrofes com cidades que se destacaram com notas mais altas. Contudo, a partir de 2013, os *clusters* BB foram se tornando mais comuns, passando a cobrir quase todo o seu território entre 2015 e 2017.

No Nordeste brasileiro, identificou-se uma grande congloeração BB nos quatro anos analisados, principalmente nas regiões limítrofes e da costa oceânica. Destaca-se, a partir de 2015, um *cluster* AA localizado nas proximidades de Sobral, no Ceará.

Na Região Centro-Oeste, em 2011, houve uma aglomeração BB nos limites entre Brasil e Bolívia, que perdurou até 2017. Contudo, em 2015, essa aglomeração dispersou-se, retornando, em 2017, à condição observada entre 2011-2013. Cabe ressaltar que, a partir de 2013, houve transbordamentos de notas altas nas cidades próximas à Região Sudeste, o que permaneceu nos anos subsequentes e espalhou-se, adentrando em direção ao centro daquela região.

Na Região Sudeste, houve um grande *cluster* AA, identificado na parte central, entre 2011 e 2013. Todavia a partir de 2015, esse aglomerado foi diminuindo e dispersando-se, de modo que, em 2017, verificou-se uma proporção visivelmente menor que os anos anteriores. Para além disso, na fronteira com a Região Nordeste, constatam-se *clusters* BB, que perduraram durante os quatro anos da análise.

Na Região Sul, observam-se quase exclusivamente *clusters* AA, concentrando-se, principalmente, na divisa com Paraguai e Argentina, nas regiões metropolitanas e costa oceânica.

De modo geral, observa-se grande concentração de *clusters* do tipo AA, ou seja, municípios com notas altas circundados por municípios com, também, notas altas no SAEB, principalmente nos conglomerados urbanos das Regiões Sul e Sudeste. Esse agrupamento pode ter sido causado pela maior dinâmica econômica da região, que proporciona tanto maior acesso a escolas, quanto famílias que objetivam qualificar-se para melhorar as condições de acesso ao mercado de trabalho (CASTRO; SADECK, 2003).

Em contrapartida, as Regiões Norte e Nordeste concentram, em quase todo o seu território, *clusters* BB. Esse perfil de *cluster* foi também identificado na Região Centro-Oeste, nas proximidades com a Bolívia. Cabe ressaltar que esses conglomerados BB convergem com o PIB *per capita* dessas regiões, que também é baixo. Esses resultados são importantes por verificarem correlação e dependência entre uma variável econômica e a educação. Esse resultado pode ser devido ao fato de que a média de pontos dos alunos, através dos anos, vem diminuindo, proporcionando esses conglomerados.

Mesmo quando se observa a educação em intervalos de notas, percebe-se forte concentração das notas altas no Sul e Sudeste brasileiro. Em 2011, quase que exclusivamente as notas acima de 275 pontos se concentraram nessas Regiões. Enquanto as notas média-baixas e abaixo de 200 pontos concentraram-se nas Regiões Norte e Nordeste.

Em 2013, observou-se uma concentração de notas baixas mais acentuada nas Regiões Norte e Nordeste, sem alterações para o período seguinte, em 2015. Já para as Regiões Sul e Sudeste, percebe-se uma maior dispersão das performances dessas notas. Contudo, em 2015, essa aglomeração retornou e provocou transbordamento para parte da Região Centro-Oeste. Por fim, em 2017, observa-se a formação de um grande conglomerado de notas altas e média-altas nas Regiões Centro-Oeste, Sul e Sudeste. Já no Norte e no Nordeste, os conglomerados foram de notas média-baixas e baixas. A tabela a seguir apresenta a quantidade de cidades em cada intervalo e, em seguida, a figura 3 ilustra essa distribuição.

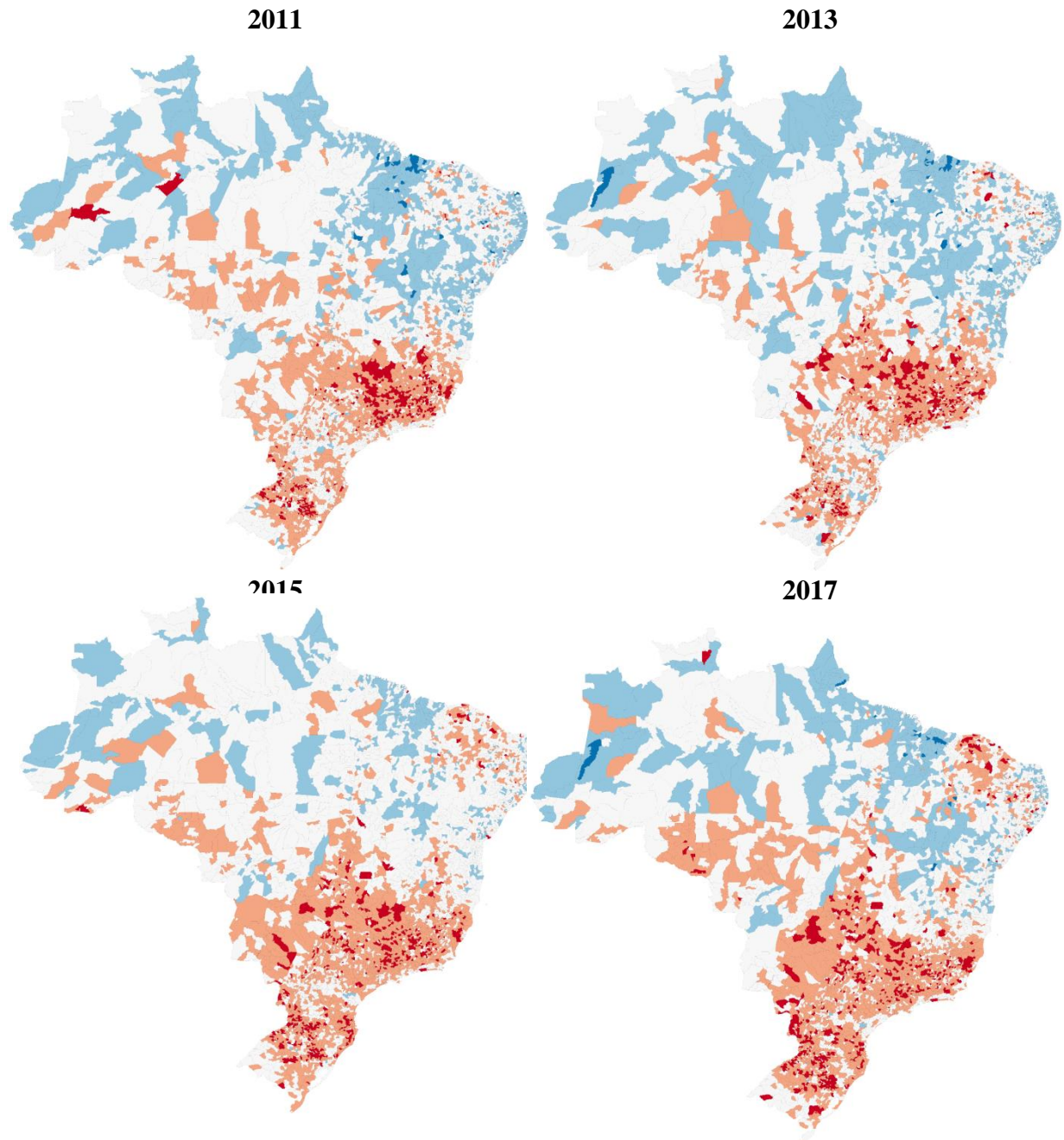
Tabela 3 - Quantidade de cidades por intervalo de notas do SAEB

Ano	<200	[200, 225]	[225, 250]	[250, 275]	>275
<b>2011</b>	40	987	2352	1749	433
<b>2013</b>	32	1032	2347	1768	382
<b>2015</b>	0	422	2428	2200	511
<b>2017</b>	19	612	2027	2285	618

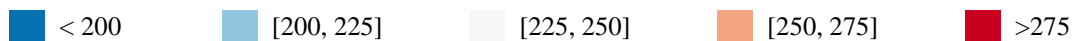
Fonte: Elaboração própria

Nota: Idealizado a partir do software GeoDa 1.14.0

Figura 3 - Distribuição das notas do SAEB por intervalos de nota



**Legenda**



Fonte: INEP (2020)

Nota: Idealizado a partir do software GeoDa 1.14.0

Verificada a existência da interação espacial dos dados em cada ano, seguem-se os procedimentos de especificação dos modelos de dados em painel e dados em painel espacial. Assim, primeiramente, realizou-se o teste de White para verificar a heterocedasticidade no modelo *pooled*. Conforme tabela 4, o teste apontou para a rejeição da hipótese nula com significância estatística de 1%. Logo, há presença de heterocedasticidade. Já o teste de

Wooldridge, para o diagnóstico de autocorrelação, apresentou uma estatística significativa a 1%, apontando para a violação da hipótese de ausência de autocorrelação.

Tabela 4 - Resumo das estatísticas dos testes realizados

<b>Teste</b>	<b>Estatística</b>
<b>White</b>	<b>1404.52</b> 0.0000 <sup>(*)</sup>
<b>Chow</b>	<b>6.98</b> 0.0000 <sup>(*)</sup>
<b>Breusch-Pagan</b>	<b>7232.65</b> 0.0000 <sup>(*)</sup>
<b>Hausman</b>	<b>4720.77</b> 0.0000 <sup>(*)</sup>
<b>Hausman robusto</b>	<b>1562.40</b> 0.0000 <sup>(*)</sup>
<b>CD de Pesaran (2004)</b>	<b>28.701</b> 0.0000 <sup>(*)</sup>
<b>CD de Pesaran (2015)</b>	<b>57.603</b> 0.0000 <sup>(*)</sup>
<b>Wooldridge</b>	<b>24.219</b> 0.0000 <sup>(*)</sup>

Fonte: Elaboração própria

<sup>(\*)</sup> Nota: indica p-valor.

Na sequência, foi verificado se as especificidades não observadas de cada município exercem efeito no desempenho dos alunos nas provas do SAEB. Para isso, primeiramente, foi realizado o teste de Chow, o qual apresentou uma estatística significativa a 1% de significância, apontando que a regressão por FE é preferível à regressão *pooled*. Em seguida, foi realizado o teste de Breusch-Pagan, conforme tabela 4, permitindo também rejeitar a hipótese nula a 1% de significância, o que significa que a estimação do modelo RE comporta melhor a regressão que *pooled*. Assim, como ambos os testes foram significativos estatisticamente, segue-se para o teste de especificação de qual modelo de efeitos não observados (efeito aleatório - RE ou efeitos fixos - FE) ajusta melhor o modelo proposto. Para tanto, foi realizado o teste de Hausman, que permitiu, conforme tabela 4, rejeitar a hipótese nula ao nível de 1% de significância estatisticamente, indicando o modelo de FE. Como complemento, foi realizado o teste de Hausman robusto, visto que a presença de heterocedasticidade pode invalidar os resultados do teste de Hausman tradicional. Assim, o teste robusto foi efetuado com a técnica de *bootstrap* de reamostragem mil vezes consecutivas. A estatística do teste determinou que o modelo de FE permaneceu sendo o mais apropriado a 1% de significância.

Além desses testes, foram implementados os testes CD de Pesaran (2004; 2015) para verificar a dependência transversal dos modelos de dados em painel e a fraca dependência transversal. O teste CD de Pesaran (2004) de independência transversal rejeitou a hipótese nula, de modo que foi verificada dependência entre as unidades de *cross-section*. O teste CD de Pesaran (2015), para fraca dependência da seção transversal, também rejeitou a hipótese nula, de modo que é reafirmada a dependência entre as unidades de *cross-section*, conforme tabela 4. Esse resultado sinaliza para a existência de dependência espacial na base dos dados utilizados.

Para efeito de robustez foi realizado o teste de CD de Pesaran (2004; 2015), igualmente para cada variável do modelo proposto e seus resíduos da regressão por efeitos fixos. Os resultados retratados demonstram que apenas a proporção de professores com mestrado ( $CD_3$ ) e doutorado ( $CD_4$ ) não apresenta dependência da seção transversal (detalhados no apêndice H e I).

Dessa forma, os testes realizados apontam para o modelo de efeito fixo, com a presença de heterocedasticidade, autocorrelação e correlação contemporânea. Após essa fase, o próximo passo é definir a matriz de pesos espaciais, que é utilizada na estimação dos modelos com dados em painel espacial. Para isso, realiza-se a regressão clássica por FE e se obtém os resíduos. De posse do vetor de resíduos, verifica-se qual matriz de ponderação espacial representa melhor a distribuição dos efeitos não observados e se os seus resíduos seguem distribuição espacial. Esse critério foi sugerido por Baumont (2004). De forma similar à realizada para a nota de matemática do SAEB, testaram-se as matrizes de contiguidades do tipo rainha, torre e distância inversa, dada a limitação do *software* utilizado, a saber Stata versão 15.1 – que permite a criação apenas das matrizes apresentadas neste trabalho. Contudo, o critério desenvolvido por Baumont sugere que teste seja realizado também para  $k$  vizinhos próximos (sendo  $k = 1, 2, 3, \dots, K$ ). Dessa maneira, a matriz de contiguidade que melhor estima a dependência é a contiguidade torre, conforme a Tabela 3, uma vez que as estatísticas de  $I$  de Moran dessa contiguidade são maiores (valores em negritos), *vis-à-vis* às demais retratadas na tabela em toda a série analisada, embora, nos três casos, essas estatísticas sejam significativas.

Definidos todos os critérios necessários para a regressão de dados em painel espacial, parte-se para os resultados da estimação. A tabela 6 apresenta os principais resultados das estimativas dos modelos *pooled* e dos modelos de FE nas especificações SAR, SEM e SAC.

Tabela 5 - Escolha da matriz de contiguidade para o tratamento dos resíduos segundo o critério de Critério de Baumont (Índice de Moran)

Ano	I de Moran (resíduos)		
	Rainha (*)	Torre (*)	Distância Inversa (*)
<b>2011</b>	0,2811	<b>0,2830</b>	0.1281
<b>2013</b>	0.2046	<b>0.2060</b>	0.0633
<b>2015</b>	0.1845	<b>0.1849</b>	0.0487
<b>2017</b>	0.2300	<b>0.2309</b>	0.0676

Fonte: INEP (2020)

(\*) Nota: o nível de significância das estatísticas calculadas é dado por:  $-\left(\frac{1}{1-n}\right) = -0,0002$ , sendo  $n$  o número de municípios.

Em seguida, utilizou-se o critério de Akaike (AIC) para seleção do modelo mais adequado para a estimação do modelo proposto. A Tabela 6 revela que o modelo SAC é o que apresentou a menor estatística calculada do AIC (valor em negrito). Portanto, o modelo de FE com a especificação SAC é o estatisticamente selecionado para o ajuste do modelo proposto e sua interpretação no presente estudo.

Tabela 6 - Estatísticas do critério de Akaike (AIC) para seleção do modelo econométrico espacial

Valor da estatística AIC		
SAR	SEM	SAC
117899.5	117977.1	<b>117734.6</b>

Fonte: estimado pelo autor a partir do *software* Stata 15.1

O modelo proposto associa as condições do aluno e de sua família, a qualificação dos professores e as características da infraestrutura escolar ao desempenho na prova de matemática do SAEB nos municípios brasileiros. Assim, os sinais dos parâmetros estimados, de modo geral, estão de acordo com a teoria.

Tabela 7 - Resumo das estimativas dos modelos econométricos espaciais (\*)

Variáveis	(continua)			
	Pooled p-valor	SAR p-valor	SEM p-valor	SAC p-valor
<b>CF<sub>1</sub></b>	-1.6477 <b>0.2450</b>	4.1087 <b>0.0000</b>	5.4645 <b>0.0000</b>	2.8713 <b>0.0050</b>
<b>CF<sub>2</sub></b>	-24.4526 <b>0.0000</b>	0.1316 <b>0.8930</b>	0.5258 0.5950	-0.5083 0.5710
<b>CF<sub>3</sub></b>	-15.6899 <b>0.0000</b>	-8.7201 <b>0.0000</b>	-10.5952 <b>0.0000</b>	-5.9549 <b>0.0000</b>
<b>CF<sub>4</sub></b>	39.0549 <b>0.0000</b>	1.8603 <b>0.2650</b>	1.5021 <b>0.3660</b>	2.2272 <b>0.1540</b>

Tabela 7 - Resumo das estimativas dos modelos econométricos espaciais (\*)

Variáveis	(continuação)			
	Pooled p-valor	SAR p-valor	SEM p-valor	SAC p-valor
CF <sub>5</sub>	-3.0700 <b>0.0250</b>	-5.7497 <b>0.0000</b>	-6.9560 <b>0.0000</b>	-4.1846 <b>0.0000</b>
CF <sub>6</sub>	6.9321 <b>0.0000</b>	-5.3007 <b>0.0000</b>	-5.0566 <b>0.0000</b>	-3.6123 <b>0.0000</b>
CF <sub>7</sub>	12.5549 <b>0.0000</b>	-12.3940 <b>0.0000</b>	-13.3869 <b>0.0000</b>	-9.8926 <b>0.0000</b>
CF <sub>8</sub>	11.0359 <b>0.0000</b>	8.5518 <b>0.0000</b>	8.8202 <b>0.0000</b>	7.5225 <b>0.0000</b>
CF <sub>9</sub>	-25.7496 <b>0.0000</b>	-20.3887 <b>0.0000</b>	-23.6604 <b>0.0000</b>	-15.3798 <b>0.0000</b>
CF <sub>10</sub>	11.8611 <b>0.0000</b>	11.5322 <b>0.0000</b>	11.7062 <b>0.0000</b>	10.1275 <b>0.0000</b>
CF <sub>11</sub>	31.3714 <b>0.0000</b>	12.8634 <b>0.0000</b>	11.6466 <b>0.0000</b>	13.1827 <b>0.0000</b>
CF <sub>12</sub>	1.7217 <b>0.4150</b>	20.4421 <b>0.0000</b>	21.0863 <b>0.0000</b>	18.3966 <b>0.0000</b>
CF <sub>13</sub>	32.4696 <b>0.0000</b>	32.2241 <b>0.0000</b>	32.7837 <b>0.0000</b>	28.1048 <b>0.0000</b>
CF <sub>14</sub>	24.0012 <b>0.0000</b>	15.7773 <b>0.0000</b>	16.0652 <b>0.0000</b>	13.6733 <b>0.0000</b>
CF <sub>15</sub>	29.5956 <b>0.0000</b>	7.1187 <b>0.0450</b>	7.7201 <b>0.0280</b>	5.2948 <b>0.1120</b>
CF <sub>16</sub>	5.9887 <b>0.0000</b>	2.2332 <b>0.0500</b>	1.8654 <b>0.1010</b>	2.4287 <b>0.0220</b>
CF <sub>17</sub>	39.1520 <b>0.0000</b>	8.6587 <b>0.0000</b>	7.6056 <b>0.0000</b>	8.0493 <b>0.0000</b>
CD <sub>1</sub>	1.6716 <b>0.0000</b>	-0.4995 <b>0.0140</b>	-0.3895 <b>0.0630</b>	-0.5575 <b>0.0020</b>
CD <sub>2</sub>	-1.9341 <b>0.0230</b>	1.2891 <b>0.0410</b>	0.9922 <b>0.1160</b>	1.3876 <b>0.0190</b>
CD <sub>3</sub>	0.8844 <b>0.5790</b>	1.3379 <b>0.3370</b>	1.3825 <b>0.3180</b>	1.4084 <b>0.2830</b>
CD <sub>4</sub>	-1.0799 <b>0.8730</b>	5.8638 <b>0.2160</b>	5.2601 <b>0.2630</b>	6.4317 <b>0.1500</b>
CD <sub>5</sub>	0.4371 <b>0.0000</b>	0.3331 <b>0.0000</b>	0.3200 <b>0.0000</b>	0.2985 <b>0.0000</b>
CE <sub>1</sub>	1.9953 <b>0.0000</b>	0.0322 <b>0.8650</b>	0.3205 <b>0.1130</b>	-0.1357 <b>0.4030</b>
CE <sub>2</sub>	-2.7820 <b>0.0000</b>	-1.2042 <b>0.0320</b>	-1.2603 <b>0.0270</b>	-0.8587 <b>0.0950</b>
CE <sub>3</sub>	-2.8929 <b>0.0020</b>	1.3086 <b>0.2680</b>	2.3998 <b>0.0610</b>	-0.2834 <b>0.7740</b>
CE <sub>4</sub>	3.0707 <b>0.0000</b>	-1.9317 <b>0.0000</b>	-1.4666 <b>0.0000</b>	-1.8603 <b>0.0000</b>
CE <sub>5</sub>	6.2998 <b>0.0000</b>	1.4847 <b>0.0140</b>	1.9675 <b>0.0020</b>	0.7560 <b>0.1560</b>
CE <sub>6</sub>	2.4506 <b>0.0000</b>	0.6386 <b>0.0960</b>	0.4760 <b>0.2270</b>	0.7659 <b>0.0290</b>

Tabela 7 - Resumo das estimativas dos modelos econométricos espaciais (\*)

Variáveis	(conclusão)			
	Pooled p-valor	SAR p-valor	SEM p-valor	SAC p-valor
CE <sub>7</sub>	0.5517 <b>0.1980</b>	-0.2163 <b>0.6410</b>	-0.2157 <b>0.6510</b>	-0.1461 <b>0.7260</b>
CE <sub>8</sub>	5.4002 <b>0.0000</b>	-0.2594 <b>0.5240</b>	0.0951 <b>0.8180</b>	-0.5075 <b>0.1700</b>
CE <sub>9</sub>	3.2646 <b>0.0000</b>	-1.3606 <b>0.0010</b>	-0.9846 <b>0.0160</b>	-1.4440 <b>0.0000</b>
CE <sub>10</sub>	-0.5398 <b>0.1880</b>	1.9899 <b>0.0000</b>	1.6454 <b>0.0000</b>	1.9537 <b>0.0000</b>
CE <sub>11</sub>	0.1005 <b>0.0000</b>	0.0141 <b>0.1720</b>	0.0330 <b>0.0030</b>	-0.0047 <b>0.5910</b>
CE <sub>12</sub>	3.7597 <b>0.0000</b>	-1.0933 <b>0.0050</b>	-0.4321 <b>0.2960</b>	-1.2400 <b>0.0000</b>
CE <sub>13</sub>	0.1586 <b>0.5740</b>	1.8116 <b>0.0000</b>	1.7053 <b>0.0000</b>	1.5166 <b>0.0000</b>
constante	136.5263 <b>0.0000</b>	-	-	-
$\rho$	-	0.3670 <b>0.0000</b>	-	0.65556 <b>0.0000</b>
$\lambda$	-	-	0.3986 <b>0.0000</b>	-0.5149 <b>0.0000</b>
$\sigma_e$	-	8.1910	8.1953	7.8299
R <sup>2</sup>	0.5540	0.1888	0.2462	0.0819
Teste de Wald	-	1096.21 <b>0.0000</b>	1060.37 <b>0.0000</b>	3017.23 <b>0.0000</b>

Fonte: Estimado pelo autor a partir do *software* Stata 15.1

(\*) Estatísticas em negrito indicam o p-valor.

A proporção municipal de alunos do sexo masculino (CF<sub>1</sub>) tem um efeito positivo na nota média do SAEB municipal. Assim como já evidenciado na literatura, os resultados apontaram que as características da família são determinantes importantes do desempenho médio dos alunos na disciplina de matemática. Resultados controversos foram observados apenas para a presença dos pais no domicílio. A estimativa do parâmetro da variável proporção de alunos que residem com a mãe (CF<sub>4</sub>) foi positiva e significativa apenas no modelo *pooled*. Já a proporção de alunos que residem com pai (CF<sub>5</sub>) teve efeito negativo em 4,18 pontos nas notas do SAEB, sendo estatisticamente significativo a 1%. Ressalta-se que essa variável foi negativa em todas as regressões.

Em relação à educação dos pais, quando há um aumento percentual de mães que possuem o ensino médio completo (CF<sub>12</sub>), isso traz um reflexo de 18,39 pontos no SAEB no modelo SAC. Quanto ao percentual de mães que possuem ensino superior completo (CF<sub>13</sub>) o efeito é ainda mais expressivo, cerca de 28,11 pontos na avaliação do SAEB. Assim como o



percentual das mães, o percentual de educação dos pais também influencia o desempenho médio municipal dos alunos, de modo que a média cresce 13,67 e 5,29 pontos no SAEB, caso os pais possuam ensino médio completo (CF<sub>14</sub>) e ensino superior completo (CF<sub>15</sub>), respectivamente. Porém, o pai deter ensino superior não foi estatisticamente significativo. Ressalta-se que a educação da mãe foi altamente significativa e houve pouca variação, independentemente do tipo de regressão. Esse resultado está de acordo com as evidências dos estudos de Barros *et al.* (2001) e Vernier (2016).

A proporção de responsáveis que frequentaram reuniões escolares (CF<sub>16</sub>) aumenta 2,43 pontos a nota do SAEB. Esse resultado vai ao encontro das evidências do estudo De Araújo, Souza e Andrade (2019), que demonstraram que os responsáveis que se preocupam com o desempenho dos alunos e se envolvem com suas atividades na escola proporcionam uma melhora no desempenho escolar médio desses alunos.

A proporção de alunos que possuíam um ano ou mais de defasagem em relação à série que cursavam (CF<sub>3</sub>) e/ou que foram reprovados em algum momento (CF<sub>9</sub>) diminuiu -5,96 e -15,38 pontos em média, respectivamente, a nota do SAEB do município. Esses resultados estão de acordo com as evidências de Oliveira e Soares (2012) e Scriptore, Azzoni e Menezes-Filho (2018). As demais regressões superestimaram os efeitos dessas variáveis. Alguns autores, como Scriptore, Azzoni e Menezes-Filho (2018), Ortigão e Aguiar (2013) e Oliveira e Soares (2012), têm discutido os efeitos proporcionados pela reprovação e, de modo geral, não se verificou consenso nacional ou internacional de como proceder sob esse caso. Porém, esses estudos apontaram que reunir alunos com capacidades diferentes (alunos com baixo desempenho ou reprovados com os demais alunos) pode promover aos alunos de menor desempenho, com atrasos ou reprovações, mais *gaps* de conhecimento do que repetir a série que estavam.

Em relação à proporção de alunos que se dedicaram a afazeres domésticos (CF<sub>6</sub>) e/ou trabalham (CF<sub>7</sub>), os parâmetros estimados indicaram um efeito negativo de 3,61 e 9,89 pontos no SAEB, respectivamente. No entanto, sob a regressão *pooled*, percebeu-se um efeito positivo de os alunos trabalharem, o que é contraintuitivo, demonstrando a importância dos efeitos espaciais quando são desconsiderados.

Além das características citadas, foram verificadas variáveis de envolvimento do aluno com os afazeres escolares. Assim, os resultados indicaram que a proporção de alunos que fazem com frequência os exercícios para casa (CF<sub>10</sub>) aumenta, em média, 10,14 pontos o desempenho municipal do SAEB. E mais: quando os professores corrigem esses exercícios (CF<sub>11</sub>), promovem um efeito positivo de 13,18 pontos no SAEB. Esse fato demonstra a

importância do papel do professor e sua responsabilização no processo de promoção de qualidade da educação pública, conforme ressalta Klein (2006). Assim, os formuladores de políticas para a educação podem atuar por meio de algum tipo de incentivo que estimule os professores a aplicar e corrigir os exercícios para casa, uma vez que isso pode auxiliar na promoção da qualidade e no melhor desempenho do aluno no SAEB.

Como verificado por Hanushek e Kimko (2000) e Klein (2006), ter iniciado a vida escolar na pré-escola tem um efeito positivo e significativo, que acompanha o aluno durante toda sua vida acadêmica e, possivelmente, em todos os outros aspectos de sua vida. A proporção de alunos que frequentaram a pré-escola ( $CF_8$ ) proporcionou um efeito médio de cerca de 7,52 pontos no SAEB. Esse efeito é similar ao conferido pela proporção de alunos que têm hábito de leitura ( $CF_{17}$ ), que trouxe incremento de 8,05 pontos. Cabe ressaltar que a leitura, seja de jornais, livros (gerais ou literatura infanto-juvenil), revistas (com exceção de revistas de comportamento, celebridades, esportes e TV) e histórias em quadrinhos (gibis) é capaz de desenvolver tanto aspectos linguísticos quanto desempenho cognitivo matemático.

Avaliados os aspectos relacionados aos docentes, verificou-se que a proporção de professores formados em matemática ( $CD_1$ ) teve um efeito de meio ponto negativo. Em relação à atualização ( $CD_2$ ), mestrado ( $CD_3$ ) ou doutorado ( $CD_4$ ), todos tiveram efeito positivo. Contudo, somente a atualização teve efeito estatisticamente significativo de 1,39 ponto no SAEB. Esses resultados, também, foram encontrados por Amâncio *et al.* (2015).

Já os resultados da variável de experiência média dos professores ( $CD_5$ ) foram significativos, embora com magnitude de apenas 0,63 ponto no SAEB, o que pode ser devido à forma de mensuração da variável, que é computada em intervalos de cinco anos.

Quando se avaliam os efeitos da infraestrutura do ambiente escolar, verifica-se que fatores como água filtrada ( $CE_1 = -0,14$  ponto), abastecimento de água ( $CE_2 = -0,86$ ), de eletricidade ( $CE_3 = -0,28$ ), coleta de lixo periódica ( $CE_5 = 0,75$ ), laboratório de ciências ( $CE_7 = -0,15$ ), quadra de esportes ( $CE_8 = -0,51$ ) e computador para uso dos alunos ( $CE_{11} = 0,005$ ) não foram significativos estatisticamente.

Por sua vez, a coleta de esgoto ( $CE_4$ ) teve um efeito negativo de 1,86 ponto, assim como a proporção de bibliotecas ( $CE_9$ ) e acesso à internet ( $CE_{12}$ ), com efeitos de -1,44 e -1,24 ponto no SAEB, respectivamente. Esses resultados vão contraintuitivamente ao esperado e de encontro à literatura, visto que se aponta que a coleta de rejeitos por rede de esgoto tem efeito positivo no desempenho médio dos alunos, assim como o acesso a bibliotecas (BARROS *et al.* 2001; BEZERRA; KASSOUF, 2006; BIONDI; FELÍCIO, 2007; DUARTE, 2009).

Contudo, a proporção de salas de leitura ( $CE_{10}$ ) teve um efeito positivo de 1,96 ponto no SAEB. A proporção laboratórios de informática ( $CE_6$ ) e a proporção municípios que oferecem alguma atividade complementar ( $CE_{13}$ ) foram significativas, exercendo um efeito de, respectivamente, 0,77 e 1,51 ponto no SAEB. A variável atividade complementar refere-se ao percentual de escolas que oferecem atividades extracurriculares, como idiomas, esportes, xadrez e afins, demonstrando que os municípios, ao oferecerem essas atividades extraclasse para seus alunos, podem atingir maiores notas no SAEB.

Os parâmetros espaciais foram significativos em todas as estimações, principalmente, no modelo SAC, pelo qual tiveram um efeito de 0,66. Isso significa que há uma associação positiva de transbordamento dos efeitos internos da proporção das variáveis em seus vizinhos no perfil de torre. Diferentemente do último parâmetro, a defasagem dos resíduos exerceu um efeito negativo, em que os efeitos não modelados (choques) proporcionam um transbordamento que prejudica o desempenho médio municipal das notas dos alunos no SAEB.

A tabela 8 apresenta os impactos diretos, indiretos e totais das variáveis independentes. Pode-se observar que as condições do aluno e dos familiares (CF) têm um efeito transbordamento do município observado para os municípios vizinhos, de modo que os coeficientes dessas variáveis apresentaram o mesmo sinal e impactos indiretos significativos a 1% estatisticamente. Isso aponta que melhores condições dos alunos e de seus familiares no município  $i$  podem ter efeito transbordamento positivo para o município  $j$ .

Os efeitos do mestrado ( $CD_3$ ) e do doutorado ( $CD_4$ ) dos professores não foram estatisticamente relevantes. Contudo, os *gaps* de experiência dos docentes ( $CD_5$ ) provocaram efeito *spillover* estatisticamente significativo a 1%.

Tabela 8 - Impacto direto, indireto e total das variáveis independentes

Variáveis	Impacto Direto	Impacto Indireto	Impacto Total
	p-valor	p-valor	p-valor
$CF_1$	3.1008 <b>0.0050</b>	3.9771 <b>0.0060</b>	7.0779 <b>0.0050</b>
$CF_2$	-0.5489 <b>0.5710</b>	-0.7040 <b>0.5720</b>	-1.2529 <b>0.5710</b>
$CF_3$	-6.4308 <b>0.0000</b>	-8.2483 <b>0.0000</b>	-14.6791 <b>0.0000</b>
$CF_4$	2.4052 <b>0.1540</b>	3.0849 <b>0.1550</b>	5.4901 <b>0.1540</b>
$CF_5$	-4.5190 <b>0.0000</b>	-5.7962 <b>0.0000</b>	-10.3152 <b>0.0000</b>

(continua)

Tabela 8 - Impacto direto, indireto e total das variáveis independentes

Variáveis	Impacto Direto	Impacto Indireto	Impacto Total
	p-valor	p-valor	p-valor
CF <sub>6</sub>	-3.9010 <b>0.0000</b>	-5.0035 <b>0.0000</b>	-8.9044 <b>0.0000</b>
CF <sub>7</sub>	-10.6831 <b>0.0000</b>	-13.7024 <b>0.0000</b>	-24.3855 <b>0.0000</b>
CF <sub>8</sub>	8.1236 <b>0.0000</b>	10.4195 <b>0.0000</b>	18.5432 <b>0.0000</b>
CF <sub>9</sub>	-16.6088 <b>0.0000</b>	-21.3027 <b>0.0000</b>	-37.9115 <b>0.0000</b>
CF <sub>10</sub>	10.9369 <b>0.0000</b>	14.0278 <b>0.0000</b>	24.9647 <b>0.0000</b>
CF <sub>11</sub>	14.2362 <b>0.0000</b>	18.2595 <b>0.0000</b>	32.4957 <b>0.0000</b>
CF <sub>12</sub>	19.8668 <b>0.0000</b>	25.4815 <b>0.0000</b>	45.3482 <b>0.0000</b>
CF <sub>13</sub>	30.3508 <b>0.0000</b>	38.9284 <b>0.0000</b>	69.2792 <b>0.0000</b>
CF <sub>14</sub>	14.7660 <b>0.0000</b>	18.9391 <b>0.0000</b>	33.7051 <b>0.0000</b>
CF <sub>15</sub>	5.7179 <b>0.1120</b>	7.3339 <b>0.1130</b>	13.0518 <b>0.1120</b>
CF <sub>16</sub>	2.6228 <b>0.0220</b>	3.3641 <b>0.0240</b>	5.9869 <b>0.0230</b>
CF <sub>17</sub>	8.6926 <b>0.0000</b>	11.1492 <b>0.0000</b>	19.8418 <b>0.0000</b>
CD <sub>1</sub>	-0.6021 <b>0.0020</b>	-0.7722 <b>0.0030</b>	-1.3743 <b>0.0020</b>
CD <sub>2</sub>	1.4984 <b>0.0190</b>	1.9219 <b>0.0200</b>	3.4204 <b>0.0190</b>
CD <sub>3</sub>	1.5210 <b>0.2830</b>	1.9508 <b>0.2850</b>	3.4718 <b>0.2840</b>
CD <sub>4</sub>	6.9457 <b>0.1500</b>	8.9087 <b>0.1520</b>	15.8544 <b>0.1500</b>
CD <sub>5</sub>	0.3223 <b>0.0000</b>	0.4134 <b>0.0000</b>	0.7358 <b>0.0000</b>
CE <sub>1</sub>	-0.1466 <b>0.4030</b>	-0.1880 <b>0.4050</b>	-0.3346 <b>0.4040</b>
CE <sub>2</sub>	-0.9274 <b>0.0940</b>	-1.1894 <b>0.0950</b>	-2.1168 <b>0.0940</b>
CE <sub>3</sub>	-0.3060 <b>0.7740</b>	-0.3925 <b>0.7740</b>	-0.6985 <b>0.7740</b>
CE <sub>4</sub>	-2.0090 <b>0.0000</b>	-2.5767 <b>0.0000</b>	-4.5857 <b>0.0000</b>
CE <sub>5</sub>	0.8164 <b>0.1550</b>	1.0471 <b>0.1540</b>	1.8635 <b>0.1540</b>
CE <sub>6</sub>	0.8271 <b>0.0290</b>	1.0608 <b>0.0310</b>	1.8879 <b>0.0300</b>
CE <sub>7</sub>	-0.1578 <b>0.7260</b>	-0.2024 <b>0.7260</b>	-0.3602 <b>0.7260</b>

(continuação)

Tabela 8 - Impacto direto, indireto e total das variáveis independentes

Variáveis	Impacto Direto	Impacto Indireto	(conclusão) Impacto Total
	p-valor	p-valor	p-valor
<b>CE<sub>8</sub></b>	-0.5480 <b>0.1700</b>	-0.7029 <b>0.1730</b>	-1.2510 <b>0.1710</b>
<b>CE<sub>9</sub></b>	-1.5594 <b>0.0000</b>	-2.0001 <b>0.0000</b>	-3.5595 <b>0.0000</b>
<b>CE<sub>10</sub></b>	2.1098 <b>0.0000</b>	2.7060 <b>0.0000</b>	4.8158 <b>0.0000</b>
<b>CE<sub>11</sub></b>	-0.0051 <b>0.5910</b>	-0.0066 <b>0.5920</b>	-0.0117 <b>0.5920</b>
<b>CE<sub>12</sub></b>	-1.3391 <b>0.0000</b>	-1.7175 <b>0.0000</b>	-3.0566 <b>0.0000</b>
<b>CE<sub>13</sub></b>	1.6378 <b>0.0000</b>	2.1007 <b>0.0000</b>	3.7385 <b>0.0000</b>

Fonte: estimado pelo autor a partir do *software* Stata 15.1

Logo, os resultados apontam que fatores como a infraestrutura do ambiente escolar e a capacitação docente são importantes, contudo, fatores associados ao *background* familiar do aluno têm um efeito mais efetivo quando analisados em dados em painel espacial. Dessa maneira, políticas públicas voltadas para a família, principalmente para mães com baixa escolaridade, podem contribuir para melhorar o desempenho acadêmico das crianças. Políticas educacionais que incentivem os professores a aplicar e a corrigir os exercícios dos alunos também podem contribuir nesse sentido.

## 7 CONCLUSÃO

O trabalho buscou analisar os fatores que afetam o desempenho médio em matemática dos alunos das escolas públicas dos municípios brasileiros no 9º ano do ensino fundamental, com variáveis relacionadas, por município, às características médias dos alunos, das famílias, dos professores e da infraestrutura física escolar. Para isso, foram utilizados dados do SAEB e do Censo Escolar agregados em municípios entre 2011 e 2017.

Os resultados apontaram que as condições familiares foram os fatores que mais influenciaram o desempenho médio no SAEB dos alunos nos municípios no período estudado, em acordo com as evidências já apontadas na literatura, principalmente no que se refere à variável associada à escolaridade da mãe. Quando se observaram os efeitos indiretos dessas variáveis, constatou-se que a proporção de mães mais instruídas apresentou impactos que transbordaram para os municípios vizinhos positivamente, demonstrando que o *background* familiar de uma região pode afetar o desempenho escolar em outras regiões pelo efeito *spillover*.

O grau de comprometimento dos alunos com a leitura e ao fazerem os exercícios para casa com frequência também foram determinantes para o desempenho médio dos alunos nos municípios. Contudo, as variáveis de condição docente e de infraestrutura escolar tiveram efeitos menores no exame do SAEB. Ademais, alunos que tiveram suas tarefas corrigidas pelos professores apresentaram melhor desempenho nas avaliações do SAEB.

Para além desses, verificou-se forte dependência espacial, demonstrando que a estrutura espacial é relevante para determinação do desempenho escolar dos alunos no plano municipal no Brasil, de modo que a performance do município nesse quesito está associada de forma positiva à essa performance nos demais municípios, ou seja, há transbordamentos positivos das variáveis internas para a sua vizinhança. Em contrapartida, os resíduos possuíram efeito inverso, o que demonstra que choques ocorridos podem diminuir o desempenho médio do município em questão e seus vizinhos.

Assim, o estudo demonstrou a importância de considerar-se os aspectos espaciais ao analisar a educação em seu contexto nacional, visto que há disparidades entre as regiões quanto às notas do SAEB. Além disso, as características dos professores e da infraestrutura do ambiente escolar exercem efeitos; contudo, as características familiares exercem efeitos ainda mais expressivos.

Sugere-se que, para estudos futuros, possam-se utilizar demais modelos de interação espacial, assim como analisar seus aspectos em níveis hierárquicos.

**BIBLIOGRAFIA**

ALBERNAZ, Â; FERREIRA, F. H. G.; FRANCO, C. Qualidade e equidade no ensino fundamental brasileiro. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, RJ, v. 33, n. 3, p. 453-476, dez. 2002.

ALMEIDA, E. **Econometria espacial aplicada**. Campinas, SP: Ed. Alínea, 2012. 498 p.

ALVES, F. Políticas educacionais e desempenho escolar nas capitais brasileiras. **Cadernos de Pesquisa**. v. 38, n. 134, p. 413-440, 2008. Disponível em: <https://doi.org/10.1590/S0100-15742008000200008>. Acesso em: 15 set. 2019.

AMÂNCIO-VIEIRA, S. F.; BORINELLI, B.; NEGREIROS, L. F.; DALMAS, J. C. A relação entre custo direto e desempenho escolar: uma análise multivariada nas escolas de Ensino Fundamental de Londrina/PR. **Educação em Revista**, Belo Horizonte, MG, v. 31, n. 1, p. 169-194, Jan./Mar. 2015.

BARBOSA FILHO, F. D. H.; PESSÔA, S. A. Educação e crescimento: o que a evidência empírica e teórica mostra? **Revista Economia**. Brasília, DF, v. 11, n. 2, p. 265-303, maio/ago. 2010.

BARBOSA FILHO, F. D. H.; PESSÔA, S. Retorno da educação no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, RJ, v. 38, n. 1, p. 97-125, Abr. 2008.

BARROS, R. P. D.; MENDONÇA, R.; SANTOS, D. D. D.; QUINTAES, G. Determinantes do desempenho educacional no Brasil. **Pesquisa e Planejamento Econômico**, Rio de Janeiro, RJ, v. 31, n. 1, p. 1-42, abr. 2001.

BAUMONT, C. Spatial effects in housing price models: do housing prices capitalize urban development policies in the agglomeration of Dijon (1999)? **Working Papers**. Laboratoire d'économie et de gestion, 26 p. 2004. Disponível em: <https://hal.archives-ouvertes.fr/hal-01525664>. Acesso em: 24 nov. 2019.

BECKER, G. S. Human capital revisited. In: **Human capital: A theoretical and empirical analysis, with special reference to education**. University of Chicago press, 1994. cap. 2, p. 15-26.

BECKER, G. S. Investment in human capital: A theoretical analysis. **Journal of political economy**, University of Chicago press, v. 70, n. 5, Part 2, p. 9-49, oct. 1962.

BECKER, K. L.; KASSOUF, A. L. Violência nas escolas públicas brasileiras: uma análise da relação entre o comportamento agressivo dos alunos e o ambiente escolar. **Nova Economia**, Belo Horizonte, MG, v. 26, n. 2, p. 653-677, maio/ago. 2016.

BEHRMAN, J. R.; ROSENZWEIG, M. R. Aumentar a escolaridade das mulheres eleva a escolaridade da próxima geração? **American Economic Review**. v. 92, n. 1, pág. 323-334, 2002.

BELMONTE, A. *et al.* School infrastructure spending and educational outcomes: Evidence from the 2012 earthquake in Northern Italy. **Economics of Education Review**. v. 75, p. 101951, 2019.

BEZERRA, M. G.; KASSOUF, A. L. Análise dos fatores que afetam o desempenho escolar nas escolas das áreas urbanas e rurais do Brasil. In: **XLIV Congresso da Sociedade Brasileira de Economia e Sociologia Rural**. Fortaleza, CE, jul. 2006. 17 p. Disponível em: <http://ageconsearch.umn.edu/record/144811/files/169.pdf>. Acesso em: 10 maio 2019.

BIONDI, R. L.; FELÍCIO, F. D. Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados do SAEB. **Série Documental. Textos para Discussão**. Brasília, DF: MEC/INEP, v. 28, 2007. 19 p. Disponível em: <https://bit.ly/2PeuuPl>. Acesso em: 18 maio 2019.

BRASIL. Constituição (1988). **Constituição da República Federativa do Brasil**: promulgada em 5 de outubro de 1988: íntegra das emendas constitucionais. Brasília, DF: Senado Federal, 2020. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: 1 jan. 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Censo Escolar**. Brasília, DF, 2020. Disponível em: <http://portal.inep.gov.br/web/guest/microdados>. Acesso em: 1 jan. 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Índice de Desenvolvimento da Educação Básica**. Brasília, DF, 2020. Disponível em: <http://portal.inep.gov.br/web/guest/educacao-basica/ideb/resultados>. Acesso em: 1 jan. 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Indicadores Educacionais**. Brasília, DF, 2018. Disponível em: <http://portal.inep.gov.br/indicadores-educacionais>. Acesso em: 10 fev. 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Investimentos Públicos em Educação**. Brasília, DF, 2018. Disponível em: <http://portal.inep.gov.br/web/guest/microdados>. Acesso em: 10 fev. 2020.

BRASIL. Ministério da Educação. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. **Sistema de Avaliação da Educação Básica**. Brasília, DF, 2018. Disponível em: <https://bit.ly/2J9HfY2>. Acesso em: 10 fev. 2020.

BUENO, C. A.; FIGUEIREDO, I. M. Z. A relação entre educação e desenvolvimento para o Banco Mundial: a ênfase na “satisfação das necessidades básicas” para o alívio da pobreza e sua relação com as políticas de Educação Infantil. In: **IX Associação Nacional de Pós-Graduação e Pesquisa em Educação Sul**: Seminário de Pesquisa em Educação da Região Sul, Universidade Caxias do Sul, Caxias do Sul, RS, 2012. Disponível em: <http://www.ucs.br/etc/conferencias/index.php/anpedsul/9anpedsul/paper/viewFile/1024/128>. Acesso em: 1 maio 2019.



CADAVAL, A. F.; MONTEIRO, S. M. M. Determinantes da qualidade da educação fundamental no Brasil: uma análise com dados do SAEB. In: **39º Encontro Nacional de Economia**. Foz do Iguaçu, PR, 2011. Disponível em: <https://bit.ly/3c2H6ms>. Acesso em: 20 maio 2019.

CASTRO, J. A. D.; SADECK, F. Financiamento do gasto em educação das três esferas de governo em 2000. **Texto para discussão**. Brasília, DF: MP/IPEA, v. 955, jun. 2003. 31 p. Disponível em: [http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td\\_0955.pdf](http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/td_0955.pdf). Acesso em: 20 maio 2019.

CHEVALIER, A. *et al.* O impacto da renda e da educação dos pais na escolaridade dos filhos. **IZA Journal of Labor Economics**. v. 2, n. 1, pág. 8, 2013.

CLIFF, A. D.; ORD, J. K. **Spatial processes: models & applications**. London: Taylor & Francis, 266 p, 1981.

COENEN, J. *et al.* Teacher characteristics and their effects on student test scores: A systematic review. **Journal of economic surveys**. v. 32, n. 3, p. 848-877, 2018.

COLEMAN, J. S. **Equality of educational opportunity**. Washington, D.C.: US Department of Health, Education, and Welfare. Office of Education. 1966.

CUNHA, F. *et al.* Interpreting the evidence on life cycle skill formation. **Handbook of the Economics of Education**. v. 1, p. 697-812, 2006.

CURI, A. Z.; MENEZES-FILHO, N. A. A relação entre educação pré-primária, salários, escolaridade e proficiência escolar no Brasil. **Estudos Econômicos (São Paulo)**. v. 39, n. 4, p. 811-850, 2009.

DE ANDRADE, R. J.; SOARES, J. F. O efeito da escola básica brasileira. **Estudos em avaliação educacional**. v. 19, n. 41, p. 379-406, 2008. Disponível em: <https://bit.ly/2VrM7iF>. Acesso em: 10 set. 2019.

DOS SANTOS, J. B. P.; DE TOLENTINO-NETO, L. C. B. O que os dados do SAEB nos dizem sobre o desempenho dos estudantes em Matemática? = What SAEB data tell us about student performance in mathematics? **Educação Matemática Pesquisa**. São Paulo, SP, v. 17, n. 2, p. 309-333, 2015. Disponível em: <https://revistas.pucsp.br/emp/article/download/22442/pdf>. Acesso em: 18 maio 2019.

DUARTE, R. G. **Os determinantes da rotatividade dos professores no Brasil: uma análise com base nos dados do SAEB 2003**. 34 p. Dissertação (Mestrado em Economia) - Universidade de São Paulo, Ribeirão Preto, SP, 18 dez. 2009.

ELHORST, J. P. **Spatial econometrics: from cross-sectional data to spatial panels**. Heidelberg: Springer, v. 479, p. 480, 2014.

ELHORST, J. P. Specification and estimation of spatial panel data models. **International regional science review**. v. 26, n. 3, p. 244-268, 2003.

FELÍCIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de São Paulo. In: **33º Encontro Nacional de Economia**. Salvador, BA, 2005. Disponível em: <http://www.anpec.org.br/encontro2005/artigos/A05A157.pdf>. Acesso em: 27 jan. 2019.

FERRÃO, M. E.; BELTRÃO, K. I.; FERNANDES, C.; SANTOS, D.; SUÁREZ, M.; ANDRADE, O. D. C. SAEB - Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. **Revista Brasileira de Estudos de População**, v. 18, n. 1/2, p. 111-130, 2001.

FRANCO, A. M. P.; MENEZES-FILHO, N. A. Os determinantes do aprendizado com dados de um painel de escolas do SAEB. **Economia Aplicada**. v. 21, n. 3, p. 525-548, 2017. Disponível em: <https://bit.ly/39Xnw9d>. Acesso em: 18 maio 2019.

FRANCO, A. M. P.; MENEZES-FILHO, N. A. Uma análise de rankings de escolas brasileiras com dados do SAEB. **Estudos Econômicos (São Paulo)**. São Paulo, SP, v. 42, n. 2, p. 263-283, abr./jun. 2012. Disponível em: <http://www.scielo.br/pdf/ee/v42n2/02.pdf>. Acesso em: 19 maio 2019.

FRANCO, C.; ORTIGÃO, I.; ALBERNAZ, Â.; BONAMINO, A.; AGUIAR, G.; ALVES, F.; SÁTYRO, N. Qualidade e equidade em educação: reconsiderando o significado de "fatores intraescolares" = Calidad y equidad en educación: reconsiderando el significado de "factores intra-escolares" = Quality and equality in education: reconsidering the meaning of "within-school factors". **Ensaio: avaliação e políticas públicas em educação**. Rio de Janeiro, RJ, v. 15, n. 55, p. 277-298, abr./jun. 2007.

FRIEDMAN, M. **The role of government in education**. New Brunswick, N. J.: Rutgers University Press, 1955. Disponível em: <https://bit.ly/2SPQm5Q>. Acesso em: 29 abr. 2019.

GONÇALVES, F. D. O.; FRANÇA, M. T. A. Eficiência na provisão de educação pública municipal: uma análise em três estágios dos municípios brasileiros. **Estudos Econômicos**. São Paulo, SP, v. 43, n. 2, p. 271-299, 2013. Disponível em: <https://doi.org/10.1590/S0101-41612013000200003>. Acesso em: 10 set. 2019.

GREENE, W. H. **Econometric analysis**. 7 ed. Harlow, UK: Pearson Education, 2012. 1240 p.

HANUSHEK, E. A. School Resources. In: HANUSHEK, E. (Ed.). **Handbook of the Economics of Education**. Elsevier, v. 2, p. 865-908, 2006.

HANUSHEK, E. A.; KIMKO, D. D. Schooling, labor-force quality, and the growth of nations. **American economic review**. v. 90, n. 5, p. 1184-1208, 2000.

HANUSHEK, E. A.; RIVKIN, S. G. Teacher quality. In: HANUSHEK, E. (Ed.). **Handbook of the Economics of Education**. Elsevier, v. 2, p. 1051-1078, 2006.

HANUSHEK, E. Addressing Cross-National Generalizability in Educational Impact Evaluation. **National Bureau of Economic Research**. Cambridge, M. A. v. 25460, 2019. 25 p. Disponível em: <https://www.nber.org/papers/w25460>. Acesso em: 25 abr. 2019.

HAUSMAN, J. A. Specification tests in econometrics. **Econometrica**: Journal of the econometric society. v. 46, n. 6, p. 1251-1271, nov. 1978.

HONG, K.; ZIMMER, R. Does investing in school capital infrastructure improve student achievement? **Economics of Education Review**. v. 53, p. 143-158, 2016.

JALES, H. B. **Peer effects na educação no Brasil**: evidência a partir dos dados do SAEB. 2010. 81 p. Dissertação (Mestrado em Economia) - Escola de Economia de São Paulo. Fundação Getúlio Vargas, São Paulo, SP, 9 ago. 2010.

KLEIN, R. Como está a educação no Brasil? O que fazer? **Ensaio**: avaliação e políticas públicas em educação. v. 14, n. 51, p. 139-172, 2006.

KLEIN, R. Utilização da teoria de resposta ao item no Sistema Nacional de Avaliação da Educação Básica (SAEB). **Meta**: Avaliação. v. 1, n. 2, p. 125-140, maio/ago. 2009. Disponível em: <http://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/38/17>. Acesso em: 19 maio 2019.

LANGE, F.; TOPEL, R. The social value of education and human capital. In: HANUSHEK, E. (Ed.). **Handbook of the Economics of Education**. Elsevier, v. 1, p. 459-509, 2006.

LEE, L. F. Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. **Econometrica**. v. 72, n. 6, p. 1899-1925, 2004.

LESAGE, J. P. An introduction to spatial econometrics. **Revue d'économie industrielle**. n. 123, p. 19-44, 2008.

LUCAS, R. E. On the mechanics of economic development. **Journal of monetary economics**. v. 22, n. 1, p. 3-42, 1988.

LUZ, L. N. **Dois ensaios sobre eficácia e equidade na educação**. 2014. 131 p. Dissertação (Mestrado em Economia) - Universidade Federal de Juiz de Fora, Juiz de Fora, MG, 4 fev. 2014. Disponível em: <https://repositorio.ufjf.br/jspui/handle/ufjf/714>. Acesso em: 1 maio 2019.

MALTHUS, T. R. **Essai sur le principe de population**. Paris: Flammarion, v. 2, 1798.

MANKIW, N. G.; ROMER, D.; WEIL, D. N. A contribution to the empirics of economic growth. **The quarterly journal of economics**. v. 107, n. 2, p. 407-437, May, 1992.

MENEZES-FILHO, N. A. **Os determinantes do desempenho escolar do Brasil**. In: DELFIM NETTO, A. (Coord.). **O estado da arte em economia**: O Brasil e a ciência econômica em debate. Saraiva, 2011.

MINCER, J. Investment in human capital and personal income distribution. **Journal of political economy**. The University of Chicago Press, v. 66, n. 4, p. 281-302, 1958.

MINCER, J. **Schooling, Experience, and Earnings**. Human Behavior & Social Institutions No. 2. National Bureau of Economic Research, Inc. 1974. 167 p.

MINCER, J. The distribution of labor incomes: a survey with special reference to the human capital approach. **Journal of economic literature**. v. 8, n. 1, p. 1-26, 1970.

OLIVEIRA, L. F. B. D.; SOARES, S. S. Determinantes da repetência escolar no Brasil: uma análise de painel dos censos escolares entre 2007 e 2010. **Texto para discussão**. Brasília, DF: SAE/IPEA, v. 1706, fev. 2012. 34 p. Disponível em: <https://bit.ly/32mM4WQ>. Acesso em: 19 maio 2019.

OLIVEIRA, V.; MENEZES-FILHO, N.; KOMATSU, B. A Relação entre a Qualidade da Gestão Municipal e o Desempenho Educacional no Brasil. **Policy Paper**. São Paulo, SP, n. 34, ago. 2018. Disponível em: <https://bit.ly/2VgJz6A>. Acesso em: 1 de maio de 2019.

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **Programme for International Student Assessment**. OECD Publishing, 2020. Disponível em: <https://pisadataexplorer.oecd.org/ide/idepisa/>. Acesso em: 23 fev. 2020.

ORTIGÃO, M. I. R.; AGUIAR, G. S. Repetência escolar nos anos iniciais do ensino fundamental: evidências a partir dos dados da Prova Brasil 2009. **Revista Brasileira de Estudos Pedagógicos**. Brasília, DF, v. 94, n. 237, p. 364-389, ago. 2013. Disponível em: <https://bit.ly/2VivaXE>. Acesso em: 19 maio 2019.

PESARAN, M. H. General Diagnostic Tests for Cross-Section Dependence in Panels. **CESifo Working Paper Series**. No. 1229, aug. 2004. Disponível em: <https://ssrn.com/abstract=572504>. Acesso em: 24 nov. 2019.

PESARAN, M. H. Testing Weak Cross-Sectional Dependence in Large Panels. **Econometric Reviews**. v. 34, n. 6-10, p. 1089-1117, 2015.

RAMOS, C. A. **Introdução à economia da educação**. Rio de Janeiro, RJ: Alta Books Editora, 2015.

SCHULTZ, T. W. Investment in human capital. **The American Economic Review**. v. 51, n. 1, p. 1-17, 1961. Disponível em: <https://www.jstor.org/stable/1818907?seq=1>. Acesso em: 12 jun. 2018.

SCRIPTORE, J. S.; AZZONI, C. R.; MENEZES-FILHO, N. A. Os impactos do saneamento básico sobre a educação: usando a privatização como variável instrumental. In: **46º Encontro Nacional de Economia**. Niterói, RJ, 2018. Disponível em: <https://bit.ly/2vYiWZQ>. Acesso em: 19 maio 2019.

SEN, A. **Development as Freedom**. New York: Anchor, 2000. 382 p.

SILVA, C. D. D. **Fatores econômicos e sociais: impacto no resultado do Índice de Desenvolvimento da Educação Básica (IDEB) de municípios do Estado de São Paulo de 2005 a 2015**. 2017. Dissertação (Mestrado em Economia) - Universidade Católica de Brasília, Brasília, DF, 4 dez. 2017. Disponível em: <https://bdtd.ucb.br:8443/jspui/handle/tede/2372>. Acesso em: 1 de maio 2019.

SMITH, A. **A riqueza das nações: uma investigação sobre a natureza e as causas da riqueza das nações.** 4ª ed. Rio de Janeiro, RJ: Nova Fronteira, 1-591 p. 2017.

SOARES NETO, J. J.; KARINO, C. A.; JESUS, G. R.; ANDRADE, D. F. D. A infraestrutura das escolas públicas brasileiras de pequeno porte. **Estudos em Avaliação Educacional.** v. 24, n. 54, p. 78-99, jan./abr. 2013. Disponível em: <https://repositorio.enap.gov.br/handle/1/1915>. Acesso em: 1 de maio 2019.

SOLOW, R. M. A contribution to the theory of economic growth. **The quarterly journal of economics.** v. 70, n. 1, p. 65-94, Fev. 1956.

STAKHOVYCH, S.; BIJMOLT, T. H. Specification of spatial models: A simulation study on weights matrices. **Papers in Regional Science.** v. 88, n. 2, p. 389-408, 2009.

TAVARES, P.; CAMELO, R.; PACIÊNCIA, L. Uma análise do papel das escolas e das redes de ensino sobre as desigualdades de oportunidades educacionais. **Economia Aplicada.** v. 22, n. 2, p. 47-80, 2018. Disponível em: <https://bit.ly/2IkDgHsc>. Acesso em: 4 set. 2019.

TOBLER, W. R. A Computer Movie Simulating Urban Growth in the Detroit Region. **Economic Geography.** vol. 46, p. 234-240, jun. 1970. Disponível em: <https://bit.ly/2wvJLVk>. Acesso em: 24 nov. 2019.

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION. **International standard classification of education: ISCED 2011.** Montreal: UNESCO Institute for Statistics, 2012, 88 p. Disponível em: <https://bit.ly/2IoVGqQ>. Acesso em: 1 jan. 2020.

VERNIER, L. D. S. **Crescimento educacional brasileiro: uma análise da distribuição e disseminação dos efeitos espaciais.** 2016. 82 p. Tese (Doutorado em Economia) - Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, RS, 2016.

VERNIER, L. D. S.; BAGOLIN, I. P.; FOCHEZATTO, A. Distribuição e disseminação espacial da educação nos municípios brasileiros. In: **I Congress Latin American and Caribbean Regional Science Association International.** FEA/USP, São Paulo, SP, 2017. 20 p. Disponível em: <https://bit.ly/2HPyrWt>. Acesso em: 1 maio 2019.

## **APÊNDICE A – TESTE DE WHITE DESCONSIDERANDO AS DEFASAGENS ESPACIAIS**

White's general test statistic : 1404.518 Chi-sq(633) P-value = 1.9e-60

## APÊNDICE B – TESTE DE CHOW DESCONSIDERANDO AS DEFASAGENS ESPACIAIS

Fixed-effects (within) regression  
 Group variable: CIDADE

Number of obs = 22,244  
 Number of groups = 5,561

R-sq:  
 within = 0.2469  
 between = 0.1913  
 overall = 0.1854

Obs per group:  
 min = 4  
 avg = 4.0  
 max = 4

F(35,16648) = 155.98  
 Prob > F = 0.0000

corr(u\_i, Xb) = 0.1552

MATEMÁTICA	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
SEXO	3.894808	1.144418	3.40	0.001	1.651626	6.13799
NPI	.7606289	1.020162	0.75	0.456	-1.238998	2.760256
DIS	-12.21223	.5809132	-21.02	0.000	-13.35088	-11.07358
MORA_MÃE	2.254226	1.73729	1.30	0.194	-1.151048	5.6595
MORA_PAI	-6.771295	1.260568	-5.37	0.000	-9.242143	-4.300446
TRABALHO_DOMÉSTICO	-8.579458	1.132322	-7.58	0.000	-10.79893	-6.359986
TRABALHA	-15.39278	1.226179	-12.55	0.000	-17.79622	-12.98934
PRÉ_ESCOLA	7.503678	.982338	7.64	0.000	5.578191	9.429165
REPROVADO	-22.23562	.8807852	-25.25	0.000	-23.96205	-20.50918
FAZ_EXERCÍCIOS	11.90219	.8885218	13.40	0.000	10.16059	13.64379
PROFESSOR_CORRIGE	12.87496	2.1538	5.98	0.000	8.653279	17.09663
MÃE_EM	21.66715	1.92419	11.26	0.000	17.89554	25.43877
MÃE_ES	35.90056	2.63424	13.63	0.000	30.73717	41.06395
PAI_EM	17.03502	2.362919	7.21	0.000	12.40345	21.6666
PAIS_ES	7.749763	3.692084	2.10	0.036	.5128855	14.98664
REUNIÃO_PAIS_E_MESTRES	2.268605	1.188735	1.91	0.056	-.0614417	4.598652
LEITURA	10.42874	1.178451	8.85	0.000	8.118847	12.73863
FORMADO_EM_MAT_PROF	-.9192096	.2118262	-4.34	0.000	-1.334412	-.5040076
ATUALIZAÇÃO_MAT_PROF	1.269869	.6584728	1.93	0.054	-.0208082	2.560545
MESTRADO_PROF	1.417043	1.451957	0.98	0.329	-1.428946	4.263033
DOCTORADO_PROF	6.00089	4.930074	1.22	0.224	-3.662581	15.66436
EXPERIÊNCIA_PROF	.3333891	.0662325	5.03	0.000	.2035663	.4632118
ÁGUA_FILTRADA	.180382	.1968504	0.92	0.360	-.2054658	.5662298
ABASTECIMENTO_ÁGUA_RP	-1.161438	.5861239	-1.98	0.048	-2.310303	-.0125723
ABASTECIMENTO_ELETRICIDADE_RP	2.406854	1.228607	1.96	0.050	-.0013459	4.815054
ABASTECIMENTO_REDE_DE_ESGOTO_RP	-3.018725	.3566623	-8.46	0.000	-3.717821	-2.319629
LIXO_COLETA_PERIÓDICA	2.440969	.6269139	3.89	0.000	1.212151	3.669787
LABORATÓRIO_DE_INFORMÁTICA	.3803804	.3998287	0.95	0.341	-.4033265	1.164087
LABORATÓRIO_DE_CIÊNCIAS	-.3589604	.4832864	-0.74	0.458	-1.306253	.5883325
QUADRA	.1018145	.4240646	0.24	0.810	-.7293974	.9330264
BIBLIOTECA	-1.420235	.419153	-3.39	0.001	-2.241819	-.5986501
SALA_DE_LEITURA	2.768044	.4495403	6.16	0.000	1.886897	3.649191
COMPUTADOR_PARA_USO_DOS_ALUNOS	.0310995	.0107268	2.90	0.004	.0100739	.0521252
INTERNET	-1.056489	.4073298	-2.59	0.010	-1.854899	-.2580792
ATIVIDADE_COMPLEMENTAR	2.649202	.2443674	10.84	0.000	2.170215	3.128188
_cons	226.2351	2.989326	75.68	0.000	220.3757	232.0945
sigma_u	16.846723					
sigma_e	8.5292008					
rho	.79597431	(fraction of variance due to u_i)				

F test that all u\_i=0: F(5560, 16648) = 6.98

Prob > F = 0.0000

## APÊNDICE C – TESTE DE BREUSCH-PAGAN DESCONSIDERANDO AS DEFASAGENS ESPACIAIS

Breusch and Pagan Lagrangian multiplier test for random effects

MATEMÁTICA[CIDADE,t] = Xb + u[CIDADE] + e[CIDADE,t]

Estimated results:

	Var	sd = sqrt(Var)
MATEMÁT~A	406.8154	20.16967
e	72.74727	8.529201
u	82.50899	9.083446

Test: Var(u) = 0

[chibar2\(01\)](#) = 7232.65  
 Prob > chibar2 = 0.0000



## APÊNDICE D – TESTE DE HAUSMAN

	Coefficients		(b-B) Difference	sqrt (diag (V_b-V_B)) S.E.
	(b) HFE	(B) HRE		
SEXO	3.894808	3.511463	.3833449	.
NPI	.7606289	-22.28575	23.04637	.6919956
DIS	-12.21223	-11.46558	-.7466517	.162592
MORA_MÃE	2.254226	16.94939	-14.69517	.
MORA_PAI	-6.771295	-.9200878	-5.851207	.274627
TRABALHO_D~O	-8.579458	1.061218	-9.640675	.
TRABALHA	-15.39278	2.573179	-17.96596	.5040833
PRÉ_ESCOLA	7.503678	12.66118	-5.1575	.5929113
REPROVADO	-22.23562	-27.0354	4.799779	.3040144
FAZ_EXERCÍ~S	11.90219	11.55519	.347002	.1998819
PROFESSOR_~E	12.87496	18.68207	-5.807118	.
MÃE_EM	21.66715	11.78372	9.883438	.3900827
MÃE_ES	35.90056	33.8485	2.052063	.6232276
PAI_EM	17.03502	24.26429	-7.229264	.632708
PAIS_ES	7.749763	13.26461	-5.514845	.
REUNIÃO_PA~S	2.268605	2.464606	-.1960002	.5629612
LEITURA	10.42874	21.81814	-11.38941	.3053095
FORMADO_EM~F	-.9192096	.1391429	-1.058353	.
ATUALIZAÇÃ~F	1.269869	.2436359	1.026233	.
MESTRADO_P~F	1.417043	1.844918	-.427875	.1818798
DOCTORADO_~F	6.00089	3.085342	2.915548	.
EXPERIÊNCI~F	.3333891	.3907454	-.0573564	.
ÁGUA_FILTR~A	.180382	.8265199	-.6461379	.0447377
ABASTEC~A_RP	-1.161438	-2.373796	1.212358	.2853668
ABASTEC~E_RP	2.406854	7.016404	-4.60955	.6731556
ABASTEC~O_RP	-3.018725	.1792058	-3.197931	.1799827
LIXO_COLET~A	2.440969	4.881357	-2.440388	.2485956
LABORATÓRI~A	.3803804	1.285348	-.9049673	.0885589
LABORATÓRI~S	-.3589604	1.745327	-2.104287	.2003361
QUADRA	.1018145	3.883937	-3.782122	.1841237
BIBLIOTECA	-1.420235	1.010377	-2.430611	.1860413
SALA_DE_LE~A	2.768044	.9647432	1.803301	.1840392
COMPUTADOR~S	.0310995	.0661848	-.0350853	.
INTERNET	-1.056489	.7952291	-1.851718	.
ATIVIDADE_~R	2.649202	1.632404	1.016797	.

b = consistent under Ho and Ha; obtained from xtreg

B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

chi2(35) = (b-B)' [(V\_b-V\_B)^(-1)] (b-B)  
 = 4720.77  
 Prob>chi2 = 0.0000  
 (V\_b-V\_B is not positive definite)

## APÊNDICE E – TESTE DE HAUSMAN ROBUSTO

Cluster-Robust Hausman Test  
(based on 1000 bootstrap repetitions)

```
b1: obtained from xtreg MATEMÁTICA SEXO NPI DIS MORA_MÃE MORA_PAI TRABALHO_DOMÉSTICO TRABALHA
> PRÉ_ESCOLA REPROVADO FAZ_EXERCÍCIOS PROFESSOR_CORRIGE MÃE_EM MÃE_ES PAI_EM PAIS_ES REUNIÃO_P
> AIS_E_MESTRES LEITURA FORMADO_EM_MAT_PROF ATUALIZAÇÃO_MAT_PROF MESTRADO_PROF DOUTORADO_PROF
> EXPERIÊNCIA_PROF ÁGUA_FILTRADA ABASTECIMENTO_ÁGUA_RP ABASTECIMENTO_ELETRICIDADE_RP REDE_DE_E
> SGOTO_RP LIXO_COLETA_PERIÓDICA LABORATÓRIO_DE_INFORMÁTICA LABORATÓRIO_DE_CIÊNCIAS QUADRA BIB
> LIOTECA SALA_DE_LEITURA COMPUTADOR_ALUNOS INTERNET ATIVIDADE_COMPLEMENTAR, fe
b2: obtained from xtreg MATEMÁTICA SEXO NPI DIS MORA_MÃE MORA_PAI TRABALHO_DOMÉSTICO TRABALHA
> PRÉ_ESCOLA REPROVADO FAZ_EXERCÍCIOS PROFESSOR_CORRIGE MÃE_EM MÃE_ES PAI_EM PAIS_ES REUNIÃO_P
> AIS_E_MESTRES LEITURA FORMADO_EM_MAT_PROF ATUALIZAÇÃO_MAT_PROF MESTRADO_PROF DOUTORADO_PROF
> EXPERIÊNCIA_PROF ÁGUA_FILTRADA ABASTECIMENTO_ÁGUA_RP ABASTECIMENTO_ELETRICIDADE_RP REDE_DE_E
> SGOTO_RP LIXO_COLETA_PERIÓDICA LABORATÓRIO_DE_INFORMÁTICA LABORATÓRIO_DE_CIÊNCIAS QUADRA BIB
> LIOTECA SALA_DE_LEITURA COMPUTADOR_ALUNOS INTERNET ATIVIDADE_COMPLEMENTAR, re
```

Test: Ho: difference in coefficients not systematic

```
chi2(35) = (b1-b2)' * [V_bootstrapped(b1-b2)]^(-1) * (b1-b2)
          =      1562.40
Prob>chi2 =      0.0000
```

**APÊNDICE F – TESTE CD DE PESARAN DE DEPENDÊNCIA TRANSVERSAL  
DOS MODELOS DE DADOS EM PAINEL**

Pesaran's test of cross sectional independence = 57.603, Pr = 0.0000

## APÊNDICE G – TESTE CD DE PESARAN DE FRACA DEPENDÊNCIA DA SEÇÃO TRANSVERSAL

Pesaran (2015) test for weak cross-sectional dependence.  
Residuals calculated using *predict, e* from *xtreg*.

H0: errors are weakly cross-sectional dependent.  
CD = 57.603  
p-value = 0.000

## APÊNDICE H – TESTE CD DE PESARAN PARA AS VARIÁVEIS

Panelvar: CIDADE

Timevar: ANO

Variable	CD-test	p-value	average joint T	mean $\rho$	mean abs( $\rho$ )
MATEMÁTICA	985.352	0.000	4.00	0.13	0.52
SEXO	337.031	0.000	4.00	0.04	0.49
NPI	404.465	0.000	4.00	0.05	0.50
DIS	1320.463	0.000	4.00	0.17	0.54
MORA_MÃE	721.363	0.000	4.00	0.09	0.50
MORA_PAI	2123.618	0.000	4.00	0.27	0.53
TRABALHO_DOM~O	3716.37	0.000	4.00	0.47	0.63
TRABALHA	2147.239	0.000	4.00	0.27	0.54
PRÉ_ESCOLA	296.802	0.000	4.00	0.04	0.51
REPROVADO	1201.572	0.000	4.00	0.15	0.52
FAZ_EXERCÍCIOS	149.513	0.000	4.00	0.02	0.49
PROFESSOR_CO~E	1444.777	0.000	4.00	0.18	0.47
MÃE_EM	1251.008	0.000	4.00	0.16	0.51
MÃE_ES	691.215	0.000	4.00	0.09	0.50
PAI_EM	910.859	0.000	4.00	0.12	0.50
PAIS_ES	664.557	0.000	4.00	0.08	0.49
REUNIÃO_PAIS~S	181.7	0.000	4.00	0.02	0.51
LEITURA	562.23	0.000	4.00	0.07	0.51
FORMADO_EM_M~F	1159.338	0.000	4.00	0.15	0.53
ATUALIZAÇÃO~F	571.478	0.000	4.00	0.07	0.08
MESTRADO_PROF	.015	0.988	4.00	0.00	0.00
DOUTORADO_PROF	.001	0.999	4.00	0.00	0.00
EXPERIÊNCIA~F	1604.231	0.000	4.00	0.20	0.51
ÁGUA_FILTRADA	2997.167	0.000	4.00	0.38	0.40
ABASTECIM~A_RP	1470.833	0.000	4.00	0.19	0.41
ABASTECIM~E_RP	300.226	0.000	4.00	0.04	0.05
ABASTECIM~O_RP	107.926	0.000	4.00	0.01	0.01
LIXO_COLETA~A	1800.924	0.000	4.00	0.23	0.33
LABORATÓRIO~A	1638.212	0.000	4.00	0.21	0.47
LABORATÓRIO~S	588.056	0.000	4.00	0.07	0.24
QUADRA	2825.933	0.000	4.00	0.36	0.51
BIBLIOTECA	1745.123	0.000	4.00	0.22	0.50
SALA_DE_LEIT~A	249.059	0.000	4.00	0.03	0.33
COMPUTADOR_P~S	4803.893	0.000	4.00	0.61	0.70
INTERNET	2654.398	0.000	4.00	0.34	0.43
ATIVIDADE_CO~R	1700.563	0.000	4.00	0.22	0.49

Notes: Under the null hypothesis of cross-section independence,  $CD \sim N(0,1)$ 

P-values close to zero indicate data are correlated across panel groups.

## APÊNDICE I – TESTE CD DE PESARAN PARA OS RESÍDUOS DA REGRESSÃO POR EFEITOS FIXOS

Panelvar: CIDADE

Timevar: ANO

Variable	CD-test	p-value	average joint T	mean $\rho$	mean abs( $\rho$ )
RESÍDUOS_FE	57.603	0.000	4.00	0.01	0.51

Notes: Under the null hypothesis of cross-section independence,  $CD \sim N(0,1)$   
P-values close to zero indicate data are correlated across panel groups.

## APÊNDICE J – TESTE DE WOOLDRIDGE PARA CONSTATAR AUTOCORRELAÇÃO

```
Wooldridge test for autocorrelation in panel data
H0: no first-order autocorrelation
    F( 1, 5560) = 24.219
    Prob > F = 0.0000
```

