



UNIVERSIDADE FEDERAL DO CEARÁ
CENTRO DE CIÊNCIA AGRÁRIAS
DEPARTAMENTO DE ECONOMIA AGRÍCOLA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA RURAL

JAYANE FREIRES FERREIRA

**ENSAIOS SOBRE OS DETERMINANTES DA QUALIDADE EDUCACIONAL NO
BRASIL**

FORTALEZA

2024

JAYANE FREIRES FERREIRA

ENSAIOS SOBRE OS DETERMINANTES DA QUALIDADE EDUCACIONAL NO
BRASIL

Dissertação apresentada ao Programa de Pós-Graduação em Economia Rural da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Economia Rural. Área de concentração: Políticas Públicas e Desenvolvimento Rural.

Orientador: Prof. Dr. Jair Andrade de Araújo

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação
Universidade Federal do Ceará
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

F441c Ferreira, Jayane Freires.

Ensaaios sobre os determinantes da qualidade educacional no Brasil / Jayane Freires Ferreira. – 2024.
98 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências Agrárias, Programa de Pós-Graduação em Economia Rural, Fortaleza, 2024.

Orientação: Prof. Dr. Jair Andrade de Araújo.

1. Desempenho escolar . 2. Qualidade da educação. 3. Função de produção educacional. 4. Gastos com educação. 5. Rural e urbano . I. Título.

CDD 338.1

JAYANE FREIRES FERREIRA

ENSAIOS SOBRE OS DETERMINANTES DA QUALIDADE EDUCACIONAL NO
BRASIL

Dissertação apresentada ao Programa de Pós-Graduação em Economia Rural da Universidade Federal do Ceará, como requisito parcial à obtenção do título de Mestre em Economia Rural. Área de concentração: Políticas Públicas e Desenvolvimento Rural.

Aprovada em: 07/02/2024.

BANCA EXAMINADORA

Prof. Dr. Jair Andrade de Araújo (Orientador)
Universidade Federal do Ceará (UFC)

Prof. Dr. Francisco José da Silva Tabosa
Universidade Federal do Ceará (UFC)

Prof^a. Dra. Andréa Ferreira da Silva
Universidade Regional do Cariri (URCA)

Dr. Gerrio dos Santos Barbosa
Núcleo de Estudos Racias (NERI) – INSPER

A Deus.

A minha família pelo apoio incondicional.

AGRADECIMENTOS

A Deus, antes de tudo, por me proporcionar determinação e saúde mental para atingir os meus objetivos.

A minha família, especialmente, a meus pais Elizângela e Cícero, meus irmãos Brenda e Bruno, e minha avó Lourdes pelo apoio incondicional. A Elias e Eva pelo suporte contínuo. Ao meu avô Francisco e aos meus tios Júnior e Wedson.

Ao meu orientador, Prof. Dr. Jair Andrade de Araújo, pelos ensinamentos oferecidos para a realização da pesquisa. Aos membros da banca examinadora, Prof. Dr. Francisco José da Silva Tabosa, Prof^a. Dra. Andréa Ferreira da Silva e o Dr. Gerrio dos Santos Barbosa pelas valiosas contribuições e sugestões para o desenvolvimento final da dissertação. A todos os docentes do Programa de Pós-Graduação em Economia Rural.

Às amigas construídas durante esse período e que me apoiaram imensamente, em especial, a Gescilene, Tiago, Nicole, Trícia, Ivan, Moisés, Rubens, Aline e a Ingrid.

À Luciana Rodrigues, por ter compartilhado os seus conhecimentos e me ensinado os primeiros passos para a construção de uma base de dados.

À Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (Funcap), pelo apoio financeiro com a manutenção da bolsa de auxílio.

À Universidade Federal do Ceará, excelente instituição de aprendizado.

Por fim, para todas as pessoas que contribuíram de forma direta e indiretamente durante esse decurso. Muito obrigada!

“Educação nunca foi despesa. Sempre foi investimento com retorno garantido.”

(Sir Arthur Lewis)

RESUMO

Esta dissertação é composta por três ensaios que relacionam os principais determinantes para a qualidade da educação no Brasil. O primeiro ensaio analisa os determinantes do desempenho educacional para o 5º ano do ensino fundamental nas escolas públicas, com ênfase nas características dos alunos, da família, dos professores, das escolas e da localização, com a aplicação de um painel de efeitos fixos entre os anos de 2011 e 2017. Os resultados centrais do estudo sugeriram que os fatores que mais explicam o desempenho escolar são as características individuais dos alunos e educação da mãe, trazendo ilação do que é discutido na literatura. Por sua vez, o segundo ensaio examina a relação entre gastos públicos com educação por aluno e qualidade educacional nos municípios brasileiros entre o hiato de 2013 e 2019. Metodologicamente, foram utilizados os microdados do Sistema de Avaliação da Educação Básica (Saeb) no 5º e 9º ano do ensino fundamental nas disciplinas de Língua Portuguesa e Matemática, adotando uma abordagem de dados em painel a nível município com efeitos fixos que controla as características não observáveis e constantes no tempo. Os resultados apresentaram que todos os coeficientes dos gastos educacionais foram positivos e significantes, indicando que os aumentos nas despesas elevam a qualidade educacional nos municípios, porém, é válido reconhecer que essa relação ainda é consideravelmente pequena. Por fim, o terceiro ensaio busca analisar a dimensão da lacuna de desempenho educacional entre alunos que frequentam escolas rurais e urbanas no ensino médio entre os anos de 2017 a 2021 no Brasil. Como estratégia de identificação, foram utilizados os microdados do Saeb para o 3º do ensino médio, aplicando o modelo de Mínimos Quadrados Ordinários (MQO) com efeitos fixos de estado e município que controlam as características não observáveis. Os resultados em torno da média apresentaram que a inserção gradual das covariadas e dos efeitos fixos reduzem o diferencial de desempenho entre os alunos rurais e urbanos, mas a magnitude das diferenças de aprendizado ainda é consideravelmente elevada. Logo, aumentar os investimentos ligados à distribuição de renda para a melhoria das condições socioeconômicas das famílias dos alunos que residem na área rural é importante para impulsionar um melhor bem-estar e, conseqüentemente, melhorar os resultados educacionais.

Palavras-chave: desempenho escolar; qualidade da educação; função de produção educacional; gastos com educação; rural e urbano.

ABSTRACT

This dissertation consists of three essays that examine the main determinants of education quality in Brazil. The first essay analyzes the determinants of educational performance for the 5th grade in public schools, focusing on student, family, teacher, school, and location characteristics, using a fixed-effects panel from 2011 to 2017. The main findings suggest that the factors most explaining school performance are the individual characteristics of students and the mother's education, aligning with the discussions in the literature. The second essay, in turn, examines the relationship between public education spending per student and educational quality in Brazilian municipalities from 2013 to 2019. Methodologically, microdata from the Basic Education Assessment System (Saeb) were used for the 5th and 9th grades in Portuguese and Mathematics, adopting a panel data approach at the municipal level with fixed effects that control for unobservable and time-invariant characteristics. The results showed that all coefficients of educational expenditures were positive and significant, indicating that increased spending raises educational quality in municipalities, though it is worth acknowledging that this relationship remains relatively small. Finally, the third essay seeks to analyze the performance gap between students attending rural and urban high schools from 2017 to 2021 in Brazil. As an identification strategy, Saeb microdata for the 3rd year of high school were used, applying the Ordinary Least Squares (OLS) model with state and municipal fixed effects that control for unobservable characteristics. The average results showed that the gradual inclusion of covariates and fixed effects reduces the performance differential between rural and urban students, but the magnitude of learning differences remains considerably high. Therefore, increasing investments related to income distribution to improve the socioeconomic conditions of students' families living in rural areas is important to enhance their well-being and, consequently, improve educational outcomes.

Keywords: school performance; education quality; educational production function; education spending; rural and urban.

LISTA DE GRÁFICOS

Gráfico 1 –	Histograma das notas do 5º ano do ensino fundamental em língua portuguesa entre 2011 e 2017.....	35
Gráfico 2 –	Histograma das notas do 5º ano do ensino fundamental em língua portuguesa entre 2011 e 2017.....	36
Gráfico 3 –	Desempenho médio das escolas em língua portuguesa e matemática por localização entre 2011 e 2017.....	36
Gráfico 4 –	Histograma da distribuição dos gastos por aluno entre 2013 e 2019.....	63
Gráfico 5 –	Média total das notas padronizadas em língua portuguesa e matemática por localidade	80
Gráfico 6 –	Média total das características individuais e socioeconômicas dos alunos por localidade	80

LISTA DE QUADROS

Quadro 1 – Descrição das variáveis	29
Quadro 2 – Descrição das variáveis	57
Quadro 3 – Descrição das variáveis	77

LISTA DE TABELAS

Tabela 1 –	Média e desvio padrão da proficiência em língua portuguesa e matemática para o 5º ano entre 2011 e 2017.....	35
Tabela 2 –	Média e desvio padrão das referente as características individuais dos alunos e seu <i>background</i> familiar, professores, diretores e escolas entre os anos de 2011 e 2017.....	37
Tabela 3 –	Testes de especificação	39
Tabela 4 –	Resultados da estimação da função de produção educacional para língua portuguesa e matemática no 5º ano do ensino fundamental entre 2011 e 2017.....	41
Tabela 5 –	Estatística descritiva da proficiência em Língua Portuguesa e Matemática para o 5º e 9º ano entre 2013 e 2019.....	62
Tabela 6 –	Quantidade de <i>missing</i> para cada variável entre 2013 e 2019.....	62
Tabela 7 –	Testes de especificação.....	64
Tabela 8 –	Resultados do modelo para a proficiência em língua portuguesa e matemática no 5º ano do ensino fundamental.....	65
Tabela 9 –	Resultados do modelo para a proficiência em língua portuguesa e matemática no 9º ano do ensino fundamental.....	66
Tabela 10 –	Resultados do modelo para a proficiência em língua portuguesa e matemática no 5º ano do ensino fundamental com gastos em educação defasados.....	68
Tabela 11 –	Resultados do modelo para a proficiência em língua portuguesa e matemática no 9º ano do ensino fundamental com gastos em educação defasados.....	68
Tabela 12 –	Resultados das estimações em língua portuguesa e matemática em 2017.....	82
Tabela 13 –	Resultados das estimações em língua portuguesa e matemática em 2019.....	83
Tabela 14 –	Resultados das estimações em língua portuguesa e matemática em 2021.....	85

LISTA DE ABREVIATURAS E SIGLAS

EA	Efeitos Fixos
EF	Efeitos Aleatórios
IBGE	Instituto Brasileiro de Geografia e Estatística
Ideb	Índice de Desenvolvimento da Educação Básica
Inep	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
IPCA	Índice de Preços ao Consumidor Amplo
MQO	Mínimos Quadrados Ordinários
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
Pisa	<i>Programme for International Student Assessment</i>
POLS	<i>Pooled Ordinary Least Squares</i>
PIB	Produto Interno Bruto
Saeb	Sistema de Avaliação da Educação Básica
Siconfi	Sistema de Informações Contábeis e Fiscais do Setor Público Brasileiro

SUMÁRIO

1	INTRODUÇÃO	16
2	DETERMINANTES DO DESEMPENHO EDUCACIONAL NAS ESCOLAS PÚBLICAS DO BRASIL	19
2.1	Introdução	19
2.2	Os determinantes do desempenho escolar: um passeio pela literatura econômica	21
2.2.1	<i>Evidências empíricas internacionais</i>	21
2.2.2	<i>Evidências empíricas no Brasil.....</i>	25
2.3	Metodologia.....	27
2.3.1	<i>Base de dados e descrição das variáveis</i>	28
2.3.2	<i>Função de Produção Educacional (FPE).....</i>	30
2.3.3	<i>Modelo de dados em painel.....</i>	32
2.3.4	<i>Testes de especificação</i>	33
2.3.4.1	<i>Teste Breush-Pagan: efeitos aleatórios versus pooled.....</i>	33
2.3.4.2	<i>Teste F: pooled versus efeitos fixos</i>	33
2.3.4.3	<i>Teste de Hausman: efeitos fixos versus efeitos aleatórios</i>	34
2.4	Resultados e discussão.....	34
2.4.1	<i>Estatística descritiva</i>	34
2.4.2	<i>Análise Econométrica.....</i>	39
2.5	Considerações finais	44
3	A RELAÇÃO DOS GASTOS PÚBLICOS POR ALUNO E QUALIDADE EDUCACIONAL DOS MUNICÍPIOS BRASILEIROS.....	47
3.1	Considerações iniciais	47
3.2	A discussão acerca da qualidade educacional: Consensos e dissensos na literatura existente.....	49
3.2.1	<i>A relação entre recursos educacionais e desempenho escolar</i>	49
3.2.2	<i>Políticas públicas e performance estudantil.....</i>	53
3.3	Metodologia.....	55
3.3.1	<i>Base de dados e descrição das variáveis</i>	55
3.3.2	<i>Função de Produção Educacional.....</i>	58
3.3.3	<i>Modelo de dados em painel.....</i>	58

3.3.4	<i>Testes de especificação do painel</i>	59
3.3.4.1	<i>Teste Breush-Pagan: efeitos aleatórios versus pooled</i>	59
3.3.4.2	<i>Teste F: pooled versus efeitos fixos</i>	60
3.3.4.3	<i>Teste de Hausman: efeitos fixos versus efeitos aleatórios</i>	60
3.4	Resultados e discussão	61
3.4.1	<i>Estatística descritiva</i>	61
3.4.2	<i>Estimação do modelo</i>	63
3.5	Considerações finais	69
4	DESIGUALDADES DE DESEMPENHO EDUCACIONAL ENTRE ALUNOS DE ESCOLAS RURAIS E URBANAS NO BRASIL	71
4.1	Introdução	71
4.2	Diferenças de desempenho educacional no meio urbano-rural	73
4.3	Metodologia	76
4.3.1	<i>Base de dados e descrição das variáveis</i>	76
4.3.2	<i>Mínimos Quadrados Ordinários (MQO) com efeitos fixos</i>	78
4.4	Resultados e discussão	79
4.4.1	<i>Estatística descritiva</i>	79
4.4.2	<i>Análise Econométrica</i>	81
4.5	Considerações finais	86
5	CONCLUSÃO	88
	REFERÊNCIAS	90
	APÊNDICE A – RESULTADOS DO MODELO COM BALANCEAMENTO	97
	APÊNDICE B – ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS ENTRE 2017 E 2021	98

1 INTRODUÇÃO

Desde a década de 1960, quando houve o início dos primeiros estudos empíricos em Economia da Educação com a teoria do capital humano (Mincer, 1958; Schultz, 1961; Becker, 1962), surgiram trabalhos que tiveram a finalidade de responder o seguinte questionamento: os recursos destinados à educação são um fator determinante para o desempenho educacional? Pode parecer uma questão simples, porém existe consensos e dissensos na literatura sobre essa relação, o que pode ser determinante para os percursos de possíveis políticas públicas em educação.

Nessa perspectiva, a intensa procura no ambiente acadêmico internacional em determinar essa relação teve início quando o governo americano publicou um famoso relatório que, posteriormente, ficou conhecido como *Coleman Report* (Relatório Coleman). Publicado em 1966, esse relatório trouxe a ideia de que existiria uma relação positiva e significativa entre os insumos educacionais e aprendizado dos alunos (Coleman *et al.*, 1968; Hanushek, Rivkin, Taylor, 1996; Card, Krueger, 1996, Hanushek, 2020). Com isso, essa análise é fundamentada por um modelo de produção simples, em que os insumos incluem atributos familiares, qualidade dos professores, recursos escolares, e o resultado final é o desempenho do aluno.

Nesse cenário, os estudos econômicos que se esforçaram em determinar os fatores que influenciam o sucesso escolar, medido pelos testes de proficiência, baseiam as suas investigações no conceito de função de produção educacional. Em essência, tal conceito postula que o sucesso do aluno em exames padronizados é o resultado do processo de produção, em que as escolas são tratadas como “fábricas” que produzem conhecimento, professores e funcionários são os trabalhadores e o recursos físicos são o capital (Hanushek, 1979, 2002; Todd; Wolpin, 2003; Glewwe, 2006). Portanto, o produto dessa investigação entra diretamente no processo de formulação de políticas públicas.

Segundo Hanushek (2020), embora o desempenho possa ser mensurado por pontos discretos no tempo, o processo educacional é cumulativo, ou seja, os insumos empregados em um determinado momento no tempo podem influenciar o desempenho dos alunos no presente. Nesse caso, o histórico familiar é caracterizado pela educação dos pais, rendimento e tamanho da família; os insumos escolares geralmente são medidos pelas particularidades dos professores, como escolaridade, experiência, sexo, cor/raça, dentre outros; a entidade escolar normalmente inclui tamanho da sala de aula, instalações e despesas administrativas;

e os fatores do distrito ou comunidade, são os níveis médios de gastos.

A qualidade da educação é algo difícil de se medir precisamente. Então, a literatura refere-se a esse termo como sendo a base do conhecimento e das habilidades analíticas, pois são o foco das escolas. Desse modo, os testes de desempenho padronizados são os melhores indicadores de qualidade que estão disponíveis (Hanushek; Luque, 2003; Hanushek; Woessmann, 2010). No Brasil, a constante decadência da qualidade de ensino é percebida por meio dos resultados do *Programme for International Student Assessment (Pisa)*¹. Uma avaliação internacional realizada pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), que investiga a qualidade de ensino pelo mundo, e a performance do Brasil nas áreas de Matemática, Leitura e Ciências são consideradas inferiores a outros países.

Concernente a isso, este trabalho está filtrado em três ensaios correlatos sob a perspectiva de explicar os determinantes da qualidade educacional, tendo como base os microdados do Sistema de Avaliação da Educação Básica (Saeb), pois é uma avaliação que mensura a qualidade de ensino no país, oferecendo um panorama amplo dos resultados educacionais. Inicialmente, o escopo do primeiro ensaio intitula-se “Determinantes do desempenho educacional nas escolas públicas do Brasil”, pois faz-se necessário uma revisitação para melhor entendimento dos fatores que influenciam o aprendizado dos estudantes brasileiros. Neste, são inseridos atributos dos alunos, da família, dos professores, das escolas e de localização com um painel de dados a nível escola para o 5º do ensino fundamental nos anos de 2011, 2013, 2015 e 2017.

O segundo ensaio é denominado de “A relação entre os gastos públicos por aluno e qualidade educacional dos municípios brasileiros”, pois torna-se fundamental para identificar se alocação dos recursos públicos está sendo eficiente, de modo que reflita na melhora da qualidade da educação. Para tanto, adota-se uma abordagem por meio de um painel de dados em nível municipal com o controle de efeitos fixos aplicando, desse modo, os microdados do Saeb em português e matemática para o 5º e 9º ano do ensino fundamental nos anos de 2013, 2015, 2017 e 2019.

Finalmente, o terceiro ensaio denomina-se “Desigualdades de desempenho educacional entre alunos de escolas rurais e urbanas no Brasil”. A motivação para tal

¹ O Programa Internacional de Avaliação de Alunos (Pisa) é uma avaliação internacional que mede o conhecimento e as habilidades dos alunos nas áreas de leitura, matemática e ciências, permitindo, dessa forma, a comparação dos resultados entre países.

temática é a escassez na literatura em estudar o componente geográfico como determinante no aprendizado escolar, especialmente com uma análise a nível aluno em um cenário de pós-pandemia. Dito isso, busca examinar a dimensão da lacuna de desempenho educacional entre alunos que frequentam escolas rurais e urbanas no ensino médio entre 2017 e 2021. Como estratégia de identificação, foi empregado o modelo de Mínimos Quadrados Ordinários (MQO) com efeitos fixos de estado e município que controlam as características não observáveis.

Esta dissertação contempla cinco capítulos. Além desta introdução geral, o Capítulo 2 apresenta o primeiro ensaio sobre os principais determinantes do desempenho escolar. O Capítulo 3 aborda o segundo ensaio acerca da influência dos gastos públicos sobre a qualidade da educação. Em seguida, o Capítulo 4 enfatiza o terceiro ensaio, elucidando as disparidades de desempenho entre alunos localizados em áreas rurais e urbanas. Por fim, o Capítulo 5 é destinado para realizar as considerações finais.

2 DETERMINANTES DO DESEMPENHO EDUCACIONAL NAS ESCOLAS PÚBLICAS DO BRASIL

2.1 Introdução

Desde muito tempo, é consenso na literatura de que a educação é um instrumento essencial para o aumento da renda individual, do crescimento e desenvolvimento dos países. Em tal caso, investimentos nesta área estão ligados a um conjunto de benefícios econômicos, como o aumento salarial e da produtividade, e também não econômicos, dentre os quais é possível destacar a redução da criminalidade, maior expectativa de vida e melhoria nos índices de saúde (Mincer, 1958; Schultz, 1961; Becker, 1962; Mincer, 1974; Heckman; Lochner; Todd, 2006; Hanushek; Woessmann, 2008; Becker, 2009).

Embora a teoria da produção padrão concentre os seus estudos em quantidades variáveis de uma produção homogênea, isso não é simplesmente reproduzido num equivalente educacional. Mensurar o desempenho educacional e entender os seus determinantes são importantes para a concepção de políticas públicas relacionadas a questões como integração escolar, os sistemas financeiros educativos e responsabilização dos professores (Hanushek, 1979). Nesse sentido, o entendimento sobre a qualidade educacional, e não exclusivamente quantidade da educação, estimulou o número de pesquisas da literatura de capital humano (Hanushek; Woessmann, 2010).

No Brasil, as políticas educacionais ganharam destaque pelo baixo desempenho que o país vem apresentando em avaliações nacionais e internacionais. De acordo com os dados do Pisa para o ano de 2022, quando ao todo foram avaliados 81 países, as médias do país foram praticamente as mesmas de 2018 em matemática, leitura e ciências. Essa avaliação mostrou que 73% dos estudantes obtiveram baixo desempenho na disciplina de matemática², 50% em leitura³ e 50% em ciências⁴. Ademais, ao equiparar à média dos países da Organização para a Cooperação e Desenvolvimento Econômico (OCDE), o Brasil tem resultados muito inferiores⁵ nas três avaliações (MEC, 2023).

² Quando compara o desempenho aos países da América do Sul, o Brasil fica atrás de países como Chile (412), Uruguai (409) e Peru (391), apresentando um desempenho médio de 379 pontos.

³ O desempenho médio do Brasil foi de 410 pontos, sendo essa pontuação estatisticamente inferior à média do Chile (448) e Uruguai (430).

⁴ O desempenho médio do Brasil foi de 403 pontos, rendimento inferior às médias do Chile (444), do Uruguai (435) e da Colômbia (411), ficando em último lugar (empatado com Argentina e Peru) dos países que compõem a América do Sul.

⁵ Os países da OCDE apresentaram resultados em leitura, matemática e ciências de 476, 472 e 485,

Nessa perspectiva, o rol de pesquisas sobre a performance educacional abrange várias medidas diferentes de recursos destinados às escolas, visto que as mais comumente aplicadas incluem: características individuais e socioeconômicas do aluno e da família; os recursos reais da sala de aula, como a formação e experiência dos professores e razão professor-aluno; agregados financeiros são as despesas por aluno e salários dos professores; e os recursos nas escolas, como insumos administrativos, instalações, características dos diretores e das turmas (Hanushek, 1997; Woessmann, 2003; Rivkin; Hanushek; Kain, 2005; Hanushek; Woessmann, 2010; Curi; Souza, 2015; Franco; Menezes Filho, 2017; Kroth; Gonçalves, 2019; Hanushek, 2020).

Sob a luz dessas ideias, e considerando que a temática não está consolidada na literatura, o presente capítulo tem o objetivo de analisar os determinantes do desempenho educacional nas escolas públicas de ensino no Brasil. Ante o exposto, buscar-se responder algumas indagações: os atributos individuais e socioeconômico dos alunos e o *status* familiar são de fato importantes para explicar o aprendizado escolar? Particularidades dos professores, como maior grau de escolaridade, experiência e salários agregam para a qualidade do ensino? Em geral, a infraestrutura escolar e a quantidade de turmas presente na escola adiciona algum valor na performance nas avaliações padronizadas? São questionamentos importantes que serão respondidos no decorrer deste capítulo.

Procurou-se aplicar um modelo que melhor descrevesse a influência de atributos observáveis sobre o desempenho escolar em língua portuguesa e matemática, na qual foi construído um modelo de função de produção educacional com dados em painel, atribuindo o modelo de efeitos fixos que controla as características não observáveis e constantes no tempo. Contudo, utilizando os microdados do Sistema de Avaliação da Educação Básica (Saeb) no 5º ano do ensino fundamental para os períodos de 2011, 2013, 2015 e 2017, foi incorporada as variáveis relacionadas aos alunos, *background* familiar, professores, diretor, turma e infraestrutura escolar.

Além dessa seção introdutória, o artigo contempla na seção 2.2 uma discussão das principais literaturas acerca dos determinantes da qualidade de ensino. Em seguida, a seção 2.3 delinea os dados do Saeb e do Censo Escolar, bem como apresenta o método de análise empregado. A seção 2.4 apresenta os principais resultados da estatística descritiva e análise econométrica. Por fim, a seção 2.5 faz as considerações finais extraídas a partir do estudo

realizado.

2.2 Os determinantes do desempenho escolar: um passeio pela literatura econômica

Quando surge a questão do desempenho educacional e os fatores que os influenciam, a literatura dominante internacional sobre a Economia da Educação não é consensual em relação aos insumos escolares e à melhoria nos testes padronizados, assim também como na literatura nacional. Tendo por embasamento os estudos econométricos, alguns autores costumam chegar a conclusões divergentes sobre a influência do perfil dos professores, atributos da escola e do volume de gastos com educação. Não obstante, existe uma tendência majoritária da importância das características socioeconômicas dos indivíduos e da família em explicar fortemente o aprendizado educacional.

2.2.1 Evidências empíricas internacionais

Após o início dos trabalhos em Economia da Educação na década de 1960, surgiram alguns trabalhos que buscaram estabelecer uma relação entre recursos escolares e aprendizado dos alunos, trazendo uma analogia da teoria da firma para o processo educacional, na qual se intitulou de função de produção educacional. De acordo com um dos primeiros estudos, Coleman *et al.* (1968) analisou a relação entre recursos escolares e desempenho dos estudantes depois de um estudo publicado pelo governo dos Estados Unidos denominado de “*Equality of Educational Opportunity*”, que abordava a importância dos gastos em educação. Os principais resultados demonstraram uma alta correlação da performance escolar com as características familiares, e uma pouca relação com características das escolas e professores.

Diante disso, um dos trabalhos pioneiros em analisar a eficiência do sistema educacional foi o estudo seminal de Hanushek (1971), ao qual avaliou se as características dos professores e da sala de aula são eficazes para elevar o desempenho dos estudantes. Desse modo, utilizou uma amostra de dados extraída do sistema escolar da Califórnia durante o período de 1962 a 1969, em que nela detinha as informações dos alunos, *background* familiar e pontuações nos testes de desempenho de Stanford, além da escolaridade dos professores. Em síntese, o autor chegou em algumas conclusões principais: primeiro, a experiência docente e a formação de pós-graduação não contribuem para o aumento nas

notas dos estudantes; e segundo diferentes professores e diversas composições de sala de aula afeta o desempenho dos alunos.

Este estudo, como mencionou Hanushek (1971), é melhor visto como sendo algo sugestivo e não definitivo, em outras palavras, como sendo um protótipo em vez de uma análise final. Além disso, Hanushek (1979) aborda que a importância política das análises de *input-output*, proporcionada pelo *Coleman Report*, levou a um crescimento substancial do número de estudos e esforços para interpretar muitos resultados divergentes e contraditórios. De todo modo, o autor explana que, à medida que foram iniciadas outras pesquisas nessa área, a relação estimada passou a ser chamada de “Função de Produção Educacional”, ao invés de, simplesmente, análise insumo-produto.

Em suma, Hanushek (1979) faz algumas conclusões gerais em sua análise: as diferenças do *status* socioeconômico da família ocasionam, de fato, diferenças significativas no desempenho educacional; as escolas não têm efeitos homogêneos sobre os alunos; existem evidências de que as escolas são economicamente ineficientes. Ou seja, não aplicam as combinações mais adequadas de insumos; e por fim, há diferenças significantes entre as funções de produção por cor/raça e, eventualmente, antecedentes familiares. Os insumos escolares interagem de maneira mais relevante com as particularidades base do indivíduo.

No mesmo período, a pesquisa de Boardman e Murnane (1979) ganhou notoriedade ao estabelecer a educação como um processo cumulativo. Isto é, o desempenho dos alunos corresponderá a uma função de investimentos realizados em períodos passados, em que os autores propuseram a construção de um modelo de dados em painel. Os autores afirmam que o *Coleman Report* foi uma realização importante, mas que vários críticos enfatizaram que a natureza transversal dos dados trouxe limitação sobre a viabilidade dos resultados. Nesse caso, destacaram que os dados em painel são mais eficazes para derivar inferências causais relacionadas aos determinantes do desempenho dos alunos.

Posteriormente, Hanushek (1986) realizou revisão de 187 equações contidas em 38 artigos e livros diferentes concentrando-se em fatores escolares como: razão professor/aluno, formação, experiência e salário do professor, fatores administrativos e instalações. Utilizando a análise dos coeficientes estimados por meio de um sistema de “*vote countin*”, o autor concluiu que não há relação sistemática entre os insumos escolares e desempenho dos alunos. Ademais, essas evidências aconteceram tanto para os gastos agregados, quanto para seus principais determinantes.

Em uma noção contraintuitiva, Hedges, Laine e Greenwald (1994) defendem uma

meta-análise da literatura, em que fazem uma reanálise dos achados examinados por Hanushek usando métodos de síntese mais sofisticados. De acordo com os autores, o método analítico “*vote countin*” utilizado por Hanushek para sintetizar os resultados dos seus estudos tem baixo poder estatístico. Utilizando métodos analíticos mais poderosos, os autores encontraram fortes resultados da existência de algum impacto positivo dos recursos e pouca existência de efeitos negativos. Mais precisamente, os efeitos medianos são positivos em grande parte das variáveis de recursos, com exceção da formação de professores que apresentou efeito negativo.

Com mais de três décadas de análise, Hanushek (1997) realiza atualização e ampliação de estudos passados (Hanushek, 1986, 1989) sobre o debate acerca do efeito dos recursos escolares na performance dos estudantes. Na sua concepção, em um determinado momento, o desempenho dos alunos está correlacionado aos insumos primários, como as características familiares, ambiente escolar e colegas. Além disso, enfatiza que o processo educacional é cumulativo, de forma que os fatores históricos e contemporâneos impactam no desempenho atual. Nessa perspectiva, o autor reforça que as escolas demonstram ineficiência em suas atividades, uma vez que não existe uma relação forte entre variações dos recursos escolares e desempenho dos estudantes.

Com o intuito de investigar os principais determinantes da qualidade educacional, mensurado pelas pontuações em testes internacionais, taxas de repetência e taxas de abandono, Lee e Barro (2001) utilizam um painel com efeitos fixos de escola que inclui medidas *output* e *input* voltados para uma ampla variedade de países. Os autores encontraram que as características familiares, como renda e educação dos pais, estão associadas com o desempenho dos alunos. Outras descobertas apontam que a razão professor-aluno está forte e positivamente relacionado com os resultados escolares, já os efeitos mais fracos, mas positivos, emergem do salário médio dos professores e período escolar. Ademais, o PIB per capita correlaciona-se insignificamente com desempenho nas avaliações de matemática e ciências e de forma positiva nos testes de leitura.

Por sua vez, Woessmann (2003) buscou estimar o efeito do *background* familiar, recursos e instituições escolares sobre o desempenho em matemática e ciências em 39 países entre anos de 1994 e 1995. Utilizando dados em nível aluno do *Trends in International Mathematics and Science Study* (TIMSS) com uma amostra representativa de 260.000 estudantes, o autor aplicou uma função de produção educacional estimando por Mínimos Quadrados Ponderados. Os resultados apresentaram que as diferenças de desempenho entre

países, no que condiz com a qualidade educacional, não estão correlacionados com as diferenças de gastos com educação, mas com a discrepância dos sistemas de ensino, por exemplo: testes centrais; instrumentos de controle em assuntos curriculares e orçamentais; e autonomia escolar em processos e decisões de pessoal.

Já Häkkinen, Kirjavainen e Uusitalo (2003) analisaram alterações nas despesas escolares conforme o desempenho dos alunos na Finlândia entre os anos de 1990 a 1998. Diante disso, como uma forma de identificação da variação, os autores estimaram modelos de dados em painel de efeitos fixos que usam alterações dramáticas nas despesas escolares promovidas pela recessão de 1990. Em consonância com os resultados, não houve efeito significativo de mudanças nas despesas docentes sobre as pontuações em testes, porém, uma das variáveis que mais explicam o desempenho dos alunos é a escolaridade dos pais.

Os pesquisadores Rivkin, Hanushek e Kain (2005) procuraram avaliar o impacto de características dos professores e das escolas sobre o desempenho educacional, com atenção especial aos problemas potenciais relativos à omissão ou medição incorreta de variáveis de alunos e escolas. Por meio da aplicabilidade da função de produção educacional com um painel de escolas do *UTD Texas Schools Project*, os resultados apresentaram que a instrução de alta qualidade pode compensar substancialmente as desvantagens relativas ao baixo nível socioeconômico. Ademais, encontrou que a experiência não é estatisticamente significativa associada com o desempenho após os anos iniciais de profissão, e que a redução de dez alunos no tamanho da turma é menor do que o benefício e aumentar a qualidade do professor.

Em outra perspectiva, Glewwe e Kremer (2006) desenvolveram um modelo de função de produção educacional para analisar o efeito das políticas educacionais sobre os anos de escolaridades e habilidades adquiridas dos indivíduos. Em todas as análises, grande parte das variáveis de professores e escola não apresentaram coeficientes significativamente diferentes de zero. Porém, os autores explicam que isso pode decorrer do tamanho pequeno da amostra e uma alta correlação entre muitas variáveis. Em suma, eles finalizaram seu estudo enfatizando que as escolas dos países em desenvolvimento defrontam-se com alguns problemas institucionais, tais como: distorções no orçamento da educação, que podem resultar em gastos ineficientes de fundos; e o baixo incentivo aos professores podem vir a ocasionar problemas de aumento nas taxas de desocupação.

Partindo para um panorama que aborda as diferenças de desempenho entre escolas urbanas e rurais, Ramos, Duque e Nieto (2012) examinaram essas disparidades na Colômbia utilizando os microdados do Pisa entre os anos de 2006 e 2009 com os resultados das

avaliações em matemática, ciências e leitura. Os autores aplicaram uma função de produção educacional por meio da ampliação da metodologia desenvolvida em Juhn-Murphy-Pierce (1993) para verificar os efeitos da variação no tempo. Os achados mostraram que a maior parte da diferença de desempenho educacional entre escolas rurais e urbanas está ligada as características familiares e pouco associadas com os insumos da escola.

Na mesma linha, Amini e Nivorozhkin (2015) analisam os determinantes do desempenho escolar rural-urbano para os alunos da Rússia. Com isso, estimaram uma função de produção de educação empregando técnica de variável instrumental com o intento de controlar a endogeneidade dos insumos escolares por intermédio de um modelo multinível de dados hierárquicos. Os resultados para o método de decomposição apontaram que as disparidades das escolas rurais e urbanas no desempenho escolar está relacionada com particularidades socioeconômicas da família que influencia incentivos educacionais.

A partir de uma ótica distinta, Lavy (2020) estima empiricamente o efeito dos recursos escolares e aumento na quantidade de horas que os alunos passam na escola sobre os resultados cognitivos dos estudantes da 5ª série de todas as escolas de Israel. Utilizando dados longitudinais a nível escola, o autor desenvolveu um painel especificando a função de produção educacional que vincula o desempenho educacional e seus determinantes relevantes. Os resultados indicaram que os recursos escolares, totais ou por turma, e o aumento do tempo de instrução impactam positiva e significativamente nas pontuações de matemática, ciências e inglês.

2.2.2 Evidências empíricas no Brasil

Dando o início das linhas de pesquisa, Barros *et al.* (2001) buscaram analisar os determinantes do baixo desempenho escolar centrado no universo de pessoas entre 11 e 25 anos de idade localizados nas áreas urbanas da região Nordeste e Sudeste. Os autores examinaram quatro tipos de determinantes: a qualidade e disponibilidade dos serviços educacionais; a disponibilidade de recursos familiares, sejam eles financeiros ou não; a atratividade do mercado de trabalho na área; e o agregado de recursos da comunidade em que o indivíduo vive. Concisamente, as características familiares demonstraram maior importância para explicar o desempenho educacional.

Com o intuito de compreender os determinantes do desempenho educacional brasileiro, Albernaz, Ferreira e Franco (2002) procuram quantificar o impacto da

escolaridade e experiência dos professores e infraestrutura escolar sobre o aprendizado dos estudantes da 8ª série em 1999. Utilizando dados do Saeb com um tamanho médio amostral de aproximadamente 35 alunos por escola, os autores aplicaram a função de produção educacional com modelos hierárquicos lineares. Os resultados revelaram que a escolaridade dos professores é um fator que contribui para a melhoria do desempenho dos alunos, mas de maneira crescente de acordo com o nível socioeconômico dos mesmos.

O estudo de Amaral e Menezes Filho (2008) procura identificar se os gastos com educação elevam o aprendizado dos estudantes das 4ª e 8ª séries do ensino fundamental. Para mensurar esse efeito, os autores utilizaram dados das notas da Prova Brasil em língua portuguesa e matemática e das despesas com educação fundamental dos municípios em 2005. Ao controlar pela escolaridade média da população, proporção de docentes com ensino superior, quantidade de horas-aula e *dummies* de unidade da federação, os autores encontraram efeitos pequenos e estatisticamente insignificantes dos gastos sobre o desempenho nos testes. Ademais, por meio de regressões quantílicas, encontraram que o impacto dos gastos sobre o desempenho escolar na quarta série ocorre somente nos municípios situados nos quantis mais altos.

Projetando um debate na literatura acerca da relação entre investimento em educação e desempenho escolar, Menezes Filho e Oliveira (2014) conduziram uma pesquisa correlacionando gastos por aluno e a qualidade educacional nos municípios brasileiros. Como forma de mensurar a qualidade da educação, os autores utilizaram os dados do Saeb nos anos de 2005, 2007, 2009 e 2011, aplicando a função de produção educacional com diversos métodos de regressão para dados em painel. Logo, os resultados apresentaram que há uma correlação positiva entre gastos por aluno e a performance dos estudantes em língua portuguesa e matemática.

Sob essa ótica, Franco e Menezes Filho (2017) verificaram o impacto das características dos alunos, professores, turmas, diretores e escolas sobre o a proficiência em matemática da 4ª série/5º ano, utilizando dados de dois painéis de escolas entre os anos de 1997 e 2005. Como efeito, o *background familiar* foi determinante para explicar desempenho, enquanto poucas variáveis de relacionados a escolas demonstraram algum efeito. Outras descobertas apontaram efeitos distintos das variáveis de controle sobre desempenho dependendo se a escola for da rede pública ou privada.

Em outro enfoque, a pesquisa de Silva Filho (2019) realizou uma estimação do efeito da formação dos professores sobre o desempenho dos alunos em português e matemática no

5º ano do EF, além de adicionar uma medida de defasagem para o 3º ano do EF no período de 2015. Por intermédio de um modelo de valor adicionado com efeitos fixos de escola, foi estimado também o efeito do nível superior específica do professor. Os achados do estudo mostraram que não existe efeito significativo da formação superior dos professores sobre a performance dos estudantes nas duas avaliações. Ou seja, o fato de o estudante ter um professor com ensino superior não ocasionou em maiores ganhos, em termos de desempenho, em comparação aqueles que tiveram professores sem nível superior.

Mais recentemente, Souza, Sampaio e Sampaio (2022) avaliaram o impacto da modalidade da formação dos professores de matemática nas pontuações do 9º ano do ensino fundamental nas escolas públicas entre o hiato de 2011 a 2017. Empiricamente, os autores delimitaram alguns grupos de tratamento, a saber: docentes com curso de licenciatura em português, matemática, outras áreas e até mesmo professores sem formação superior; aplicando Regressão Generalizada Impulsionada. Os principais resultados apresentaram que os docentes sem formação superior impactam de maneira negativa nas pontuações dos exames de proficiência quando equiparados a outros grupos de professores que possuem ensino superior. Adicionalmente, esse efeito negativo é mais acentuado para os alunos em situação de vulnerabilidade social, nível socioeconômico inferior ou escolas localizadas na área rural.

Mediante a literatura exposta, torna-se evidente a heterogeneidade dos resultados dos diversos determinantes à educação sobre as pontuações nos testes de proficiência, seja em âmbito, nacional ou internacional. É importante destacar a utilização da função de produção educação como forma de examinar a melhor combinação de insumos voltados à educação, de forma que o desempenho final seja maximizado. Posto isso, a seção subsequente descreverá os procedimentos adotados na pesquisa para a obtenção dos principais resultados.

2.3 Metodologia

Esta seção apresenta os principais procedimentos metodológicos empregados na pesquisa. Primeiramente, tem-se a demonstração da base de dados com a descrição das variáveis que compõem a função de produção educacional. Em seguida, há a estratégia empírica a partir do modelo teórico exposto, estimada por dados em painel. Logo após, os testes de especificação para destacar qual o melhor painel que se adequa a base de dados construída.

2.3.1 Base de dados e descrição das variáveis

Como forma de mensurar a qualidade do ensino, a base de dados é proveniente do Saeb e do Censo Escolar, ambos conduzidos pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). O Saeb é um conjunto de avaliações externas em larga escala que tem a finalidade de fazer um diagnóstico da educação básica brasileira e também fatores que podem interferir no desempenho dos alunos. Nesse sentido, o presente estudo vai recorrer aos microdados da Prova Brasil nos anos de 2011, 2013, 2015 e 2017 para o 5º do ensino fundamental em língua portuguesa e matemática das escolas da rede pública.

A Prova Brasil é uma avaliação censitária bianual que envolve os alunos do 5º e 9º ano do ensino fundamental regular das escolas públicas que possuem, no mínimo, 20 alunos matriculados nos anos avaliados. O objetivo principal é realizar avaliação da qualidade do ensino praticados nas escolas públicas, atribuindo resultados para cada escola participante, além de apresentar indicadores sobre as condições de infraestrutura da escola. Ademais, a partir de 2017 as avaliações se tornaram censitárias para o ensino médio. Logo, os testes fazem avaliação do desempenho escolar nas áreas do conhecimento em língua portuguesa (com foco em leitura) e matemática (com ênfase na resolução de problemas) (INEP, 2023).

Além desses instrumentos que mensuram o desempenho, são aplicados quatro tipos de questionários: o de alunos, que respondem perguntas sobre o ambiente e nível socioeconômico familiar, hábitos de estudo e de leitura, motivação, trajetória escolar, entre outros aspectos; o dos professores e o do diretor, que fornecem informações sobre sua formação profissional, nível socioeconômico e cultural, formas de gestão, práticas pedagógicas, dentre outros. Enfim, informações sobre a escola, em que possuem perguntas relacionadas ao ambiente físico, infraestrutura, recursos, entre outros aspectos.

Complementarmente, os dados do Censo Escolar⁶ são utilizados para aferir as variáveis de controle relacionadas às características das escolas, visto que é o principal instrumento de coleta de informações da educação básica e mais importante pesquisa estatística educacional brasileira conduzida anualmente pelo Inep. Por meio dela, os atores educacionais analisam como está a situação educacional nas escolas e, a partir disso, acompanham a efetividade das políticas públicas. Com isso, o Quadro 1 apresenta a descrição de todas as variáveis, a nível escola, extraídas das fontes de dados já supracitadas.

⁶ São dados administrativos referentes as escolas.

Uma das maiores dificuldades para a construção de um painel de escolas com dados do Saeb é harmonizar as respostas de todos os questionários sem perder uma certa quantidade de observações, desse modo, foram excluídas todas as observações com dados *missing*. O fato de escolher apenas os anos de 2011 a 2017 se dá pela inexistência de algumas variáveis no ano de 2019⁷ que seriam importantes para função de produção educacional. No caso de anos anteriores ao intervalo em estudo, optou-se por não utilizar, para obter um maior número de escolas que se repetem ao longo do tempo.

Contudo, as variáveis relacionadas a experiência dos professores foram divididas em três níveis em razão da pergunta do questionário ser de caráter categórico, onde utilizou-se a variável “leciona até 5 anos” como base, da mesma forma para as variáveis de diretor. No que tange ao salário dos professores, também foram realizados três níveis usando como base “possui até 5 salários mínimos”. Em detrimento da forma como o Saeb aplica esse questionamento, isto é, em intervalos de rendimentos, foi necessário equiparar essa variável em termos de salários mínimos para os quatros anos.

Quadro 1 – Descrição das variáveis

(Continua)

Dimensão	Variável	Descrição
Aluno e background familiar	Masculino	Proporção de alunos do sexo masculino
	Trabalha	Proporção de alunos que trabalham fora de casa
	Já reprovou	Proporção de alunos que já reprovaram uma ou mais vezes
	Já abandonou	Proporção de alunos que já abandonaram uma ou mais vezes
	Preta	Proporção de alunos da cor/raça preta
	Computador	Proporção de alunos que possuem computador em casa
	Mãe que nunca estudou	Proporção de mães que nunca estudou
	Mãe com ensino médio	Proporção de mães que tem ensino médio
	Mãe com ensino superior	Proporção de mães que tem ensino superior
Professor	Ensino superior	Proporção de professores que tem ensino superior
	Leciona entre 6 a 15 anos	Proporção de professores que leciona entre 6 a 15 anos
	Leciona há mais de 15 anos	Proporção de professores que leciona há mais de 15 anos
	Recebe entre 5 e 10 SM	Proporção de professores que possui entre 5 e 10 salários mínimos
	Recebe mais de 10 SM	Proporção de professores ganham mais de 10 salários mínimos

⁷ Não tem variáveis sobre o grau de escolaridade e margem salarial dos professores.

(Conclusão)

Dimensão	Variável	Descrição
Diretor	Ensino superior	<i>Dummy</i> : 1 sim, 0 caso contrário
	Exerce a função entre 5 a 10 anos	<i>Dummy</i> : 1 sim, 0 caso contrário
	Exerce a função há mais de 10 anos	<i>Dummy</i> : 1 sim, 0 caso contrário
Escola	Biblioteca	Possui biblioteca? <i>Dummy</i> : 1 sim, 0 caso contrário
	Laboratório de ciências	Possui laboratório de ciências? <i>Dummy</i> : 1 sim, 0 caso contrário
	Laboratório de informática	Possui laboratório de informática? <i>Dummy</i> : 1 sim, 0 caso contrário
	Quadra de esportes	Possui quadra de esportes? <i>Dummy</i> : 1 sim, 0 caso contrário
	Docentes no 5º do EF	Quantidade de docentes do ensino fundamental nos anos iniciais
	Turmas no 5º do EF	Número de turmas do ensino fundamental nos anos iniciais
Localização	Rural	<i>Dummy</i> : 1 rural, 0 urbana

Fonte: Elaboração própria a partir dos dados do Saeb e Censo Escolar de 2011 a 2017.

Logo, mesmo que o Saeb seja de caráter censitário para as turmas de 5º e 9º do ensino fundamental, nem todas as escolas públicas participam da avaliação, pois é necessário que elas se enquadrem em alguns critérios. Desse modo, para utilizar a metodologia de dados em painel, a análise será realizada usando dois conjuntos de dados, isto é, um painel desbalanceado de escolas públicas para os anos de 2011, 2013, 2015 e 2017, separado pelas áreas do conhecimento, língua portuguesa e matemática. É imprescindível ressaltar que foram excluídas do painel escolas que apareciam somente uma vez nos quatro anos compreendidos no período e somente foram consideradas escolas com no mínimo uma turma no 5º ano em cada escola.

2.3.2 Função de Produção Educacional (FPE)⁸

O paradigma dominante que é utilizado para relacionar os efeitos dos recursos educacionais sobre os resultados dos alunos nas últimas décadas tem sido a função de produção educacional. Na literatura empírica, os estudos trouxeram a analogia da teoria da firma buscando compreender a tecnologia de combinar os insumos escolares e familiares, de modo que o resultado educacional seja maximizado. Nesse caso, a estrutura conceitual da

⁸ A resenha clássica está nos trabalhos de Hanushek. Para além da literatura apresentada, a estrutura da função de produção da educação também se encontra no estudo de Hanushek (2002).

função de produção da educação considera as escolas como “fábricas” que produzem “conhecimento” e as demais características ligadas à escola e aos professores são os “insumos” (Todd; Wolpin, 2003; Glewwe *et al.*, 2011).

Como já foi citado anteriormente, a literatura que circunda o teor da pesquisa procurou estimar a função de produção educacional incorporando variáveis que tenham efeito sobre a qualidade, ou seja, de dimensão interna e externa. Segundo Dourado, Oliveira e Santos (2007), a dimensão externa contempla as variáveis relacionadas a cor/raça, gênero, instrução dos pais, renda familiar, número de filhos e quantidade de horas de estudo em casa. Outrossim, poder-se-ia adicionar aspectos regionais para captar, por exemplo, a quantidade de capital humano da localidade e nível cultural.

Para Kroth e Gonçalves (2019), a dimensão interna circunda vários aspectos como: infraestrutura escolar, em que se destaca salas de aulas adequadas, bibliotecas e laboratórios; nível dos professores, relacionado a capacitação, titulação, salários, plano de cargo, e carga horária para realizar o planejamento das atividades; gestão do trabalho escolar, que envolve métodos pedagógicos e programa curricular; e espaço social, relacionado à segurança na escola e a programas de permanência dos alunos.

Hanushek e Woessmann (2010)⁹ enfatizam, nesse contexto, que os testes de proficiência são uma boa *proxy* para estimar a qualidade de educação, pois tem a finalidade de mensurar habilidades em leitura, matemática e ciências, e conseguem agregar muito do conhecimento cognitivo dos alunos, não somente aqueles conquistados no ambiente escolar. Com isso, a função de aprendizagem pode ser descrita como:

$$Y_{it} = f(F_i^t, P_i^t, S_i^t, H_i^t) \quad (1)$$

em que: Y_{it} corresponde ao desempenho escolar do i -ésimo estudante no tempo t ; F_i^t é um vetor de antecedentes familiares do i -ésimo aluno acumulado no tempo t ; P_i^t é o efeito dos pares (amigos dentro e fora da escola) do i -ésimo estudante no período t ; S_i^t são os insumos que aumentam a qualidade da escola (características da escola e dos professores) do i -ésimo estudante acumulado no tempo t ; e H_i^t são as habilidades inatas do i -ésimo estudante. Perante o exposto, o resultado educacional é explicado pela acumulação desses componentes, de modo que possam ter efeito duradouro (Hanushek, 1979, 1997; Todd; Wolpin, 2003, 2007; Rivkin; Hanushek; Kain, 2005; Glewwe; Kremer, 2006; Menezes Filho; Oliveira, 2014).

⁹ Para medir a qualidade, os autores construíram uma *proxy* a partir resultados de testes de proficiência realizados por agências internacionais, como a OCDE.

2.3.3 Modelo de dados em painel

O modelo de dados em painel, também conhecido como modelo de dados longitudinais ou modelo de efeitos fixos/aleatórios, é utilizado para examinar dados provenientes de unidades observadas ao longo do tempo, combinando dados em *cross section* (várias unidades em um único ponto no tempo) e dados de séries temporais (observações repetidas durante o período de tempo para as mesmas unidades) (Wooldridge, 2006). Dada a disponibilidade de um conjunto de dados em painel, a aplicação do modelo básico de efeitos não observados oferece a vantagem de remover da regressão quaisquer efeitos fixos no tempo que possam influenciar a variável resposta, mas que não foram explicitamente incorporados.

Posto isso, a estratégia de identificação consistirá na aplicabilidade da natureza dos dados em painel para controlar os efeitos não observados das escolas fixos no tempo. Para tal fim, o intuito é analisar o impacto de um agregado de insumos exercem sobre a proficiência dos alunos em língua portuguesa e matemática. É plausível supor que haja atributos não observáveis das escolas que tenham influência sobre o aprendizado e que podem estar relacionados com os insumos aqui examinados. Ademais, na literatura sobre o teor da pesquisa, as características socioeconômicas dos alunos e o *background* familiar contribuem fortemente para explicar o aprendizado. Por isso, serão inclusas como variáveis de controle.

Em um panorama econométrico, será estimada a seguinte equação que modela o desempenho escolar nos testes de língua portuguesa e matemática do Saeb:

$$Y_{it} = \beta_0 + \beta_1 A_{it} + \beta_2 P_{it} + \beta_3 D_{it} + \beta_4 E_{it} + c_i + u_{it} \quad (2)$$

em que: Y_{it} é a proficiência em português e matemática na escola i no tempo t ; A_{it} é um vetor de características do aluno na escola i no tempo t ; P_{it} é o vetor de atributos dos professores na escola i no tempo t ; D_{it} são as particularidades do diretor na escola i no tempo t ; E_{it} é um vetor de característica de infraestrutura da escola i no tempo t ; c_i é o efeito específico não observado da escola e constante no tempo; e u_{it} são os distúrbios idiossincráticos, pois se alteram tanto em t como em i . Os parâmetros de interesse $\beta_1, \beta_2, \beta_3, \beta_4$ equivalem as relações condicionais dos insumos sobre o desempenho educacional.

2.3.4 Testes de especificação

Nesta parte serão destacados os testes de especificações que atribuem uma melhor escolha para o tipo de painel. Logo, são descritos os testes de Breush-Pagan, Teste F e Hausman.

2.3.4.1 Teste Breush-Pagan: efeitos aleatórios versus pooled

O teste de Breush-Pagan é usado para determinar qual dos modelos é o mais preferível, o modelo *pooled* ou modelo de efeitos aleatórios. Simbolicamente, $H_0: \sigma_n^2 = 0$ (constante comum - *pooled*) e a $H_1: \sigma_n^2 \neq 0$ (efeitos aleatórios), nesse caso o teste se dará pela seguinte relação:

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N (\sum_{t=1}^T \widehat{w}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T \widehat{w}_{it}^2} \right] \sim \chi_1^2,$$

Logo, N representa o número de escolas, T os períodos de tempo. Se $LM > \chi_1^2$, considera-se os efeitos aleatórios.

2.3.4.2 Teste F: pooled versus efeitos fixos

Este teste é utilizado para decidir qual dos modelos é mais adequado na estimação entre o modelo *pooled* ou modelo de efeitos aleatórios. Na hipótese nula (hipótese *pooled*), admite-se a homogeneidade na constante e , na hipótese alternativa, afirma-se a heterogeneidade na constante (efeitos fixos). Então, a $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_n$ e $H_1: \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_n$, a estatística F aplicada para testar essa hipótese é:

$$F_{stat} = \frac{\frac{R_{fe}^2 - R_{pool}^2}{N-1}}{\frac{1 - R_{fe}^2}{(NT - N - K)}} \sim F_{(N-1, NT-N-K)}$$

em que R_{fe}^2 e R_{pool}^2 são os respectivos coeficientes de determinação do modelo com constante comum e com efeitos fixos, N é o número de escolas, T é os períodos de tempo e K é o número de variáveis explicativas. Desse modo, os efeitos fixos são preferíveis se $F_{stat} > F_{(N-1, NT-N-K)}$.

2.3.4.3 Teste de Hausman: efeitos fixos versus efeitos aleatórios

O teste de Hausman é utilizado para decidir entre os modelos de efeitos aleatórios H_0 e efeitos fixos H_1 é o mais adequado. Na hipótese nula, os estimadores do modelo com efeitos aleatórios (estimação GLS) são consistentes e eficientes ($H_0: Cov(\eta, X_{it}) = 0$), já sob a hipótese alternativa, o modelo de efeitos aleatórios não é adequado, mas os efeitos fixos são consistentes ($H_1: Cov(\eta, X_{it}) \neq 0$). Dito isso, a estatística aplicada para realizar o teste de hipóteses é a seguinte:

$$H = (\hat{\beta}_{fe} - \hat{\beta}_{re})' [Var(\hat{\beta}_{fe}) - Var(\hat{\beta}_{re})]^{-1} (\hat{\beta}_{fe} - \hat{\beta}_{re}) \sim \chi_K^2$$

em que $\hat{\beta}_{fe}$ e $\hat{\beta}_{re}$ são os respectivos vetores dos estimadores dos efeitos fixos e aleatórios, $Var(\hat{\beta}_{fe})$ e $Var(\hat{\beta}_{re})$ são as matrizes de variância-covariância dos estimadores e K é o número de regressores. Se $H > \chi_k^2$, rejeita-se o modelo com efeitos aleatórios, logo o modelo com efeitos fixos é mais apropriado.

2.4 Resultados e discussão

Nesta seção, serão apresentados os principais resultados encontrados por meio da aplicabilidade dos métodos estabelecidos. Serão discutidas algumas evidências iniciais a partir da estatística descritiva e, em seguida, a análise econométrica, a qual mostra a relação das variáveis que compõem a FPE sobre o rendimento das escolas nos testes de proficiência.

2.4.1 Estatística descritiva

Esta subseção tem o intuito de apresentar a estatística descritiva das variáveis utilizadas no presente estudo. A Tabela 1 apresenta a frequência absoluta, médias e desvios padrões das escolas nas notas da Prova Brasil em língua portuguesa e matemática entre 2011 e 2017, em que se ressaltam as pontuações cada vez maiores nas médias das respectivas disciplinas ao longo do período. Desse modo, destaca-se a maior frequência de escolas presentes na amostra no ano de 2017, já em relação aos resultados nas duas avaliações do Saeb, houve aumentos mais expressivos, em média, nas notas de matemática ao longo do tempo.

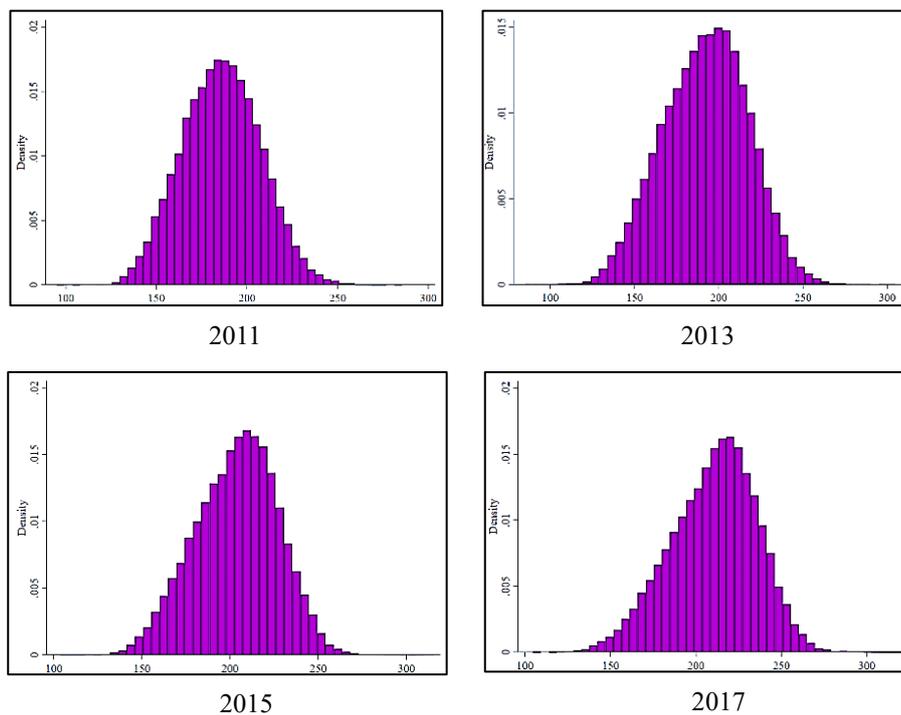
Tabela 1 – Média e desvio padrão da proficiência em língua portuguesa e matemática para o 5º ano entre 2011 e 2017

Ano	Frequência absoluta	Língua Portuguesa		Matemática	
		Média	Desvio padrão	Média	Desvio padrão
2011	29.470	185,932	21,793	204,816	25,435
2013	36.811	192,074	25,399	207,732	28,541
2015	36.855	203,625	23,573	215,481	24,509
2017	41.397	210,461	25,295	219,689	26,527

Fonte: Elaboração própria a partir dos dados do Saeb de 2011 a 2017.

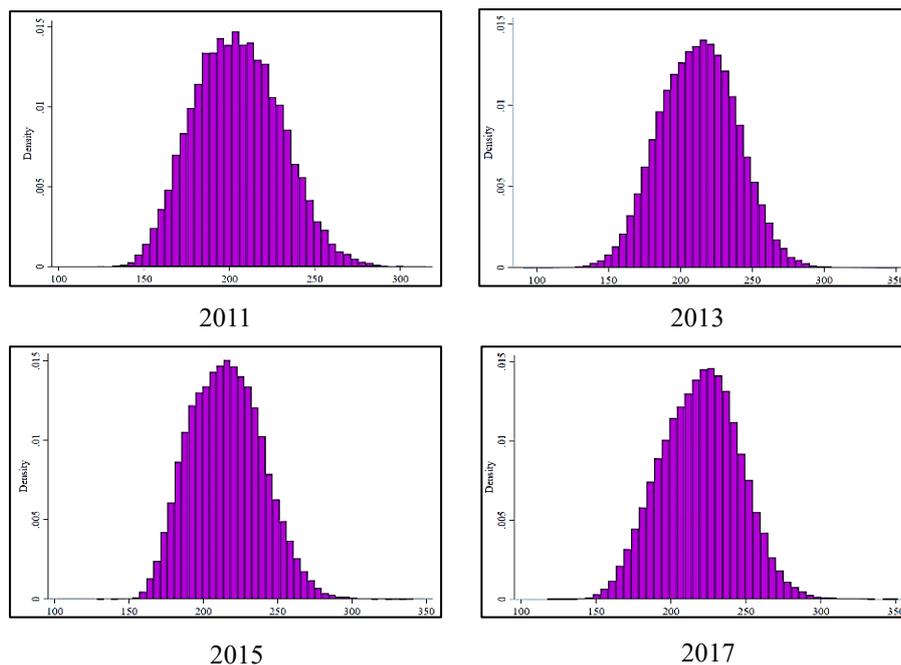
Para visualizar de forma mais detalhada as notas do 5º ano do ensino fundamental, os Gráficos 1 e 2 apresentam, respectivamente, a distribuição das pontuações em língua portuguesa e matemática em 2011, 2013, 2015 e 2017. Pelos histogramas, tem-se um melhor entendimento da dispersão das notas, confirmando o que foi exposto na Tabela 1 sobre o aumento da média e também a respeito das escolas terem um maior desempenho em matemática. Nota-se que, ao passar de cada ano, o pico, o qual concentra a maior quantidade de observações mais comum, se situa em uma média de desempenho maior.

Gráfico 1 – Histograma das notas do 5º ano do ensino fundamental em língua portuguesa entre 2011 e 2017



Fonte: Elaboração própria a partir dos dados do Saeb de 2011 a 2017.

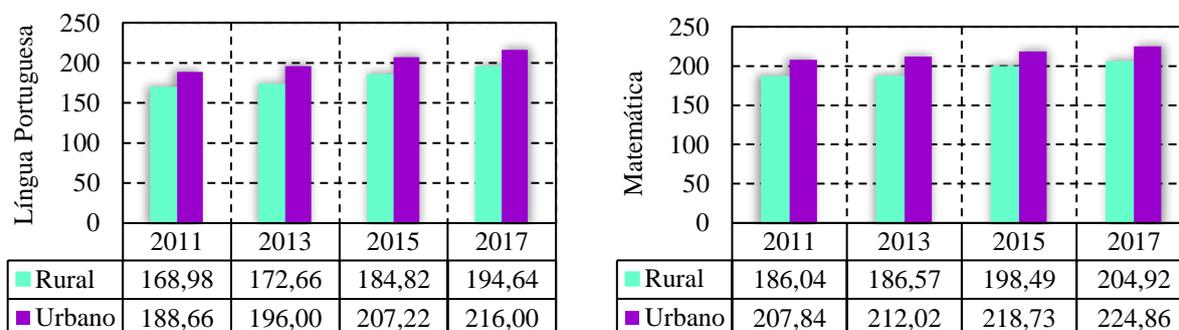
Gráfico 2 – Histograma das notas do 5º ano do ensino fundamental em matemática entre 2011 e 2017



Fonte: Elaboração própria a partir dos dados do Saeb de 2011 a 2017.

De forma mais específica, o Gráfico 3 delinea o desempenho médio das escolas públicas nos dois testes supracitados por localização no intervalo em estudo. Em um primeiro momento, é perceptível que as escolas localizadas na área rural possuem desempenho médio inferior em português e matemática quando equiparadas com as escolas públicas situadas na área urbana, com um *gap* de aproximadamente 20 pontos nas avaliações. As escolas rurais obtiveram em português e matemática uma variação de aproximadamente 2%, nas notas médias dos testes entre 2011 e 2017, da mesma forma para as escolas urbanas.

Gráfico 3 – Desempenho médio das escolas em língua portuguesa e matemática por localização entre 2011 e 2017



Fonte: Elaboração própria a partir dos dados do Saeb de 2011 a 2017.

A Tabela 2 traça uma análise inicial correlato as características individuais dos alunos e seu *background* familiar, professores, diretores e escolas entre período de 2011 e 2017. De início, em todos os anos, as escolas públicas possuem em média 50% de alunos do sexo masculino, já no caso da proporção de alunos que trabalham nas escolas, nota-se um declínio irrisório a partir do período de 2013, se situando em 15,8% e chegando a aproximadamente 12% para o ano de 2017.

A proporção de alunos pretos é bem inferior, em torno de 9% de representatividade nas escolas, já em relação aos alunos que possuem computador em casa, o ano de 2015 foi o que obteve o maior quantitativo dessa variável, com mais de 56%. Consoante a escolaridade da mãe, é perceptível que proporção de mães sem instrução é menor quando comparadas aos outros níveis de escolaridade, nesse caso, para todos os anos, há uma proporção maior de mães com ensino médio comparado aquelas que tem ensino superior.

Tabela 2 – Média e desvio padrão das referente as características individuais dos alunos e seu *background* familiar, professores, diretores e escolas entre os anos de 2011 e 2017

	(Continua)			
Dimensão / Variáveis	2011	2013	2015	2017
Aluno / <i>background</i> familiar				
Masculino	0,5030 (0,0925)	0,4995 (0,1063)	0,5038 (0,0935)	0,4984 (0,1103)
Trabalha	0,1404 (0,0967)	0,1589 (0,1050)	0,1307 (0,0945)	0,1214 (0,1011)
Já reprovou	0,3105 (0,1728)	0,2987 (0,1725)	0,2533 (0,1557)	0,2212 (0,1523)
Já abandonou	0,0715 (0,0639)	0,0809 (0,0778)	0,0708 (0,0661)	0,0671 (0,0695)
Preta	0,0908 (0,0746)	0,0930 (0,0821)	0,0929 (0,0809)	0,0985 (0,0914)
Computador	0,4716 (0,2473)	0,5689 (0,2452)	0,5530 (0,2279)	0,4712 (0,2260)
Mãe nunca estudou	0,0311 (0,0446)	0,0342 (0,0537)	0,0286 (0,0438)	0,0259 (0,0457)
Mãe com ensino médio	0,1114 (0,0722)	0,1253 (0,0819)	0,1350 (0,0789)	0,1275 (0,0838)
Mãe com ensino superior	0,0950 (0,0701)	0,0983 (0,0766)	0,1097 (0,0767)	0,1146 (0,0853)
Professor				
Ensino superior	0,9078 (0,2567)	0,8316 (0,2848)	0,8944 (0,2273)	0,5283 (0,2305)
Leciona entre 6 e 15 anos	0,4296 (0,4171)	0,3868 (0,3465)	0,3672 (0,3453)	0,2037 (0,2319)
Leciona há mais de 15 anos	0,4367 (0,4204)	0,4268 (0,3592)	0,4696 (0,3659)	0,2913 (0,2538)

Dimensão / Variáveis	(Conclusão)			
	2011	2013	2015	2017
Professor				
Possui entre 6 a 10 SM	0,2189 (0,3489)	0,1622 (0,2707)	0,2735 (0,3274)	0,0823 (0,1650)
Possui mais de 10 SM	0,0285 (0,1397)	0,0105 (0,0675)	0,0594 (0,1694)	0,0033 (0,0318)
Diretor				
Ensino superior	0,9235 (0,2657)	0,9107 (0,2852)	0,9431 (0,2316)	0,9293 (0,2563)
Exerce a função entre 6 e 10 anos	0,2532 (0,4348)	0,1808 (0,3849)	0,2119 (0,4087)	0,1812 (0,3852)
Exerce a função há mais de 10 anos	0,1995 (0,3996)	0,1521 (0,3592)	0,1678 (0,3737)	0,1499 (0,3571)
Escola				
Biblioteca	0,5393 (0,4985)	0,5169 (0,4997)	0,5218 (0,4995)	0,4723 (0,4992)
Laboratório de ciências	0,1156 (0,3198)	0,0979 (0,2973)	0,1006 (0,3007)	0,0795 (0,2706)
Laboratório de informática	0,7685 (0,4218)	0,7866 (0,4097)	0,7664 (0,4231)	0,6395 (0,4802)
Quadra de esportes	0,5316 (0,4990)	0,5344 (0,4988)	0,5639 (0,4959)	0,5332 (0,4989)
Docentes no 5º ano do EF	13,2584 (7,7061)	12,6148 (7,6379)	13,4088 (8,0155)	12,2319 (8,0924)
Turmas no 5º ano do EF	11,6924 (6,5995)	11,0523 (6,2721)	11,2100 (6,3094)	10,0357 (6,3889)

Fonte: Elaboração própria a partir dos dados do Saeb e Censo Escolar de 2011 a 2017.

Nota: Desvio padrão entre parênteses.

Analisando a nível professor, evidencia-se uma queda cada vez mais acentuada na proporção de professores que possuem ensino superior, destaque para o ano de 2011 com 90,7% do corpo docente com esse grau de instrução. No que condiz à experiência, ao longo do tempo vai reduzindo a proporção de professores que leciona entre 6 a 10 anos, do mesmo modo para os que lecionam a mais de 10 anos. Assim como ocorre na variável que mensura a experiência, acontece uma queda substancial relacionada aos professores que ganham entre 6 a 10 salários mínimos e aqueles que ganham mais de 10 salários mínimos, com apenas 8,2% e 0,3%, respectivamente, para o ano de 2017.

No que concerne às características dos diretores, existe uma concentração bem maior de diretores que possuem ensino superior nesses quatro anos investigados, pois a média tende a 1 ou 100%. Em relação ao período que exerce a função de diretor, o ano de 2011 concentrou o maior quantitativo de diretores tanto com 6 a 10 anos de experiência, quanto os que tem mais de 10 anos de execução da função. Com ênfase na infraestrutura escolar,

verifica-se que poucas escolas possuem laboratório de ciências, em contrapartida, mais de 70% das escolas possuem laboratório de informática.

Seguidamente, mais de 50% das escolas possuem quadra de esportes em todos os períodos. Em outra ótica, a quantidade média de docentes que lecionam no 5º ano do ensino fundamental foi maior em 2011 e 2015 com aproximadamente 13 professores em atividade. Enfim, sabendo que foram consideradas na amostra somente escolas que tinha pelo menos uma turma do 5º do ensino fundamental, a média dessa durante os anos é por volta de 11 turmas por escola.

2.4.2 Análise Econométrica

Após apresentar o comportamento das variáveis por meio da estatística descritiva, a Tabela 3 traça os resultados das estimações realizadas por POLS (*Pooled Ordinary Least Squares*) e efeitos fixos (EF) usando-se de um painel de escolas públicas desbalanceado. As estimações do POLS foram incluídas somente para equiparar com as estimações obtidas por EF, pois, de acordo com Franco e Menezes Filho (2017), esse método é enviesado na presença de variável omitida que esteja correlacionada com o erro. Além disso, não foram consideradas as estimações dos efeitos aleatórios (EA) em detrimento do teste de Hausman apresentar *p-valores* menores que 5%, levando a rejeitar a hipótese nula sobre a preferência do modelo, como mostra a Tabela 3.

Tabela 3 – Testes de especificação

Testes de especificação	Língua Portuguesa	Matemática
Breush-Pagan	119,32 (0,000)	168,61 (0,000)
Teste F	3,1211 (0,000)	3,0581 (0,000)
Hausman	10631 (0,000)	12195 (0,000)

Fonte: Elaboração própria a partir dos dados do Saeb e Censo Escolar de 2011 a 2017.

Nota: *p-valores* entre parênteses

A Tabela 4 mostra que os aumentos na proporção de alunos do sexo masculino na escola tendem a elevar o desempenho nos testes de matemática, pois os coeficientes apresentaram sinais positivos e significantes a 1% no POLS. O modelo de efeitos fixos sugeriu que a proporção de homens na escola possui uma relação negativa com as notas de

português, em contrapartida, essa variável eleva as pontuações de matemática em 6,7 pontos. Esse resultado vai ao encontro da literatura sobre os homens terem mais aptidão para matemática (Franco; Menezes Filho, 2017).

O coeficiente da variável que representa os alunos que trabalham esboça efeitos negativos e significantes a 1% sobre a performance nas avaliações para os dois modelos realçados. É perceptível que o modelo de efeitos fixos suaviza essa correlação negativa tanto em português quanto em matemática, com respectivos 32,5 e 27,1 pontos. Isso significa que um acréscimo na proporção de estudantes que trabalham fora de casa diminui o desempenho nos exames. Todavia, este fator pode ser amenizado garantindo a permanência do aluno na escola através de políticas públicas de apoio, por exemplo: transporte escolar, merenda-escolar e bolsas de auxílio (Hanushek, 1979; Kroth; Gonçalves, 2019).

Da mesma forma, aumentos na proporção de estudantes que já reprovaram e abandonaram uma ou mais vezes possui coeficientes estatisticamente significantes e negativos, ou seja, tendem a reduzir o rendimento nas avaliações. Em língua portuguesa, os efeitos negativos foram na ordem de 40,4 e 12,1 pontos, por sua vez, em matemática equivaleram a respectivos 32,8 e 15,9 pontos. Esses resultados podem ser explicados pelo fato de que, geralmente, os alunos que já reprovaram e/ou abandonaram têm dificuldades nos estudos, o que faz com que obtenham desempenho abaixo nos testes.

Relativamente ao aspecto racial, alunos que se declaram de cor/raça preta, os seus coeficientes mostraram sinais negativos e significantes sobre o desempenho em matemática, tanto para o modelo POLS quanto para o de efeitos fixos, e estatisticamente insignificante sobre a proficiência em português. Vale ressaltar que a inserção dos efeitos fixos atenua a relação dessa característica em ambos os testes. Os resultados corroboram o que já é discutido na literatura sobre as questões referente às características raciais, pois, geralmente, indivíduos de cor/raça preta tendem a apresentar desempenho escolar inferior.

De acordo com os resultados socioeconômicos explanados pelos efeitos fixos, o fato de o aluno possuir computador em casa está correlacionado positiva e significativamente com o desempenho dele nas notas de português e matemática em 7,8 e 6,9 pontos, respectivamente. No que tange ao *background* familiar, alunos que possuem mãe que nunca estudou afeta negativamente o desempenho nas duas disciplinas na ordem de 17,7 e 18,2, para o painel de efeitos fixos. Em contrapartida, alunos com mães que possuem ensino médio e superior contribuem de forma positiva no aprendizado.

Observou-se que alunos com mães que tem ensino médio geram efeito nas notas de

português e matemática em 33,6 e 27,8 pontos, respectivamente, e aquelas que possuem ensino superior influenciam no desempenho escolar dos alunos na ordem de 24,4 e 14,7 pontos. Essas evidências denotam a importância da escolaridade familiar para uma melhor performance nos testes padronizados, ao qual já é destacado na literatura econômica. Conforme Häkkinen, Kirjavainen e Uusitalo (2003), um dos fatores que mais explicam o desempenho escolar é a escolaridade dos pais.

Tabela 4 – Resultados da estimação da função de produção educacional para língua portuguesa e matemática no 5º ano do ensino fundamental entre 2011 e 2017

(Continua)

Dimensão / Variáveis explicativas	Variável dependente			
	Língua Portuguesa		Matemática	
	(POLS)	(EF)	(POLS)	(EF)
Aluno / <i>background</i> familiar				
Masculino	0,569 (0,443)	-4,502*** (0,452)	13,781*** (0,477)	6,702*** (0,436)
Trabalha	-48,743*** (0,511)	-32,509*** (0,548)	-44,336*** (0,550)	-27,152*** (0,527)
Já reprovou	-41,523*** (0,326)	-40,487*** (0,418)	-39,942*** (0,351)	-32,804*** (0,402)
Já abandonou	-43,526*** (0,729)	-12,153*** (0,787)	-53,100*** (0,785)	-15,907*** (0,758)
Preta	-19,983*** (0,553)	0,980 (0,674)	-28,997*** (0,596)	-2,096*** (0,649)
Computador	28,099*** (0,274)	7,867*** (0,396)	32,410*** (0,295)	6,994*** (0,381)
Mãe que nunca estudou	-30,434*** (1,060)	-17,766*** (1,190)	-29,088*** (1,141)	-18,220*** (1,146)
Mãe com ensino médio	32,849*** (0,631)	33,658*** (0,670)	24,750*** (0,679)	27,835*** (0,646)
Mãe com ensino superior	14,613*** (0,642)	24,449*** (0,746)	3,645*** (0,690)	14,779*** (0,719)
Professor				
Ensino superior	-5,896*** (0,173)	-9,384*** (0,177)	-2,142*** (0,186)	-5,139*** (0,170)
Leciona entre 6 a 15 anos	-7,129*** (0,182)	-8,116*** (0,183)	-3,852*** (0,196)	-4,625*** (0,176)
Leciona há mais de 15 anos	-5,497*** (0,180)	-7,044*** (0,189)	-2,431*** (0,194)	-3,839*** (0,182)
Possui entre 5 e 10 SM	-1,782*** (0,164)	1,097*** (0,172)	-2,422*** (0,177)	0,670*** (0,166)
Possui mais de 10 SM	0,279 (0,393)	5,355*** (0,387)	-2,443*** (0,423)	2,884*** (0,373)
Diretor				
Ensino superior	3,325*** (0,173)	1,607*** (0,178)	2,838*** (0,187)	1,260*** (0,171)
Exerce a função entre 6 a 10 anos	0,889*** (0,114)	-0,340*** (0,115)	1,189*** (0,123)	0,162 (0,111)
Exerce a função há mais de 10 anos	1,195*** (0,124)	-0,092 (0,146)	1,826*** (0,135)	0,513*** (0,140)
Escola				

(Conclusão)

Dimensão / Variáveis explicativas	Variável dependente			
	Língua Portuguesa		Matemática	
	(POLS)	(EF)	(POLS)	(EF)
Biblioteca	2,253*** (0,095)	0,374** (0,156)	1,824*** (0,103)	0,087 (0,151)
Laboratório de ciências	-2,426*** (0,158)	-0,804*** (0,310)	-2,863*** (0,171)	-0,823*** (0,298)
Laboratório de informática	0,246** (0,110)	-1,595*** (0,156)	0,847*** (0,118)	-1,319*** (0,150)
Quadra de esportes	2,022*** (0,103)	4,246*** (0,186)	2,706*** (0,110)	2,710*** (0,179)
Docentes no 5º ano do EF	0,230*** (0,013)	0,686*** (0,020)	0,268*** (0,014)	0,355*** (0,020)
Turmas no 5º ano do EF	-0,285*** (0,015)	-1,003*** (0,029)	-0,280*** (0,016)	-0,458*** (0,028)
Localização				
Rural	0,329*** (0,136)	-3,528*** (0,995)	0,762*** (0,147)	-3,476*** (0,958)
Constante	206,243*** (0,364)	-	207,437*** (0,391)	-
Observações	144.531	144.531	144.531	144.531
R ²	0,580	0,360	0,548	0,243
F-Statistic	8306,308*** (df=24; 144506)	2166,519*** (df=24; 92385)	7313,143*** (df=24; 144506)	2166,519*** (df=24; 92385)

Fonte: Elaboração própria a partir dos dados do Saeb e Censo Escolar de 2011 a 2017.

Nota: (1) Erros padrões robustos para heteroscedasticidade e autocorrelação entre parênteses. (2) *p<0,1; **p<0,05; ***p<0,01.

Adicionalmente, Menezes Filho (2007) ao avaliar os determinantes do desempenho escolar dos alunos da 4ª e 8ª séries do ensino fundamental e da 3ª série do ensino médio nos testes de proficiência em matemática, encontrou que as variáveis que melhor explicam a performance escolar são as características familiares e do aluno, como o nível educacional da mãe, cor/raça, atraso escolar e reprovação prévia, presença de computador em casa, número de livros e trabalho fora de casa.

Em consonância com os resultados para as características dos professores, a escolaridade e a experiência demonstraram impactos negativos sobre a proficiência escolar nos dois modelos. Somente as variáveis relacionadas ao nível salarial, apresentaram um efeito positivo e irrisório no desempenho escolar quando controlado para efeitos fixos. Como exposto na seção anterior, existe dissensos na literatura que envolvem os efeitos características dos professores. A saber Hedges, Laine e Greenwald (1994) encontraram que os efeitos medianos para a formação dos professores são negativos em cada uma das suas amostras de estudos e os autores Lee e Barro (2001) mostram que o salário médio dos professores tem efeito positivo no desempenho educacional, embora sejam fracos.

Ademais, Hanushek (1979) argumenta que a formação dos professores não tem efeito sobre o desempenho e pode ser interpretada como uma indicação de que as instituições não alteram, em média, as competências dos docentes. Portanto, dependendo da forma como os professores são atribuídos às escolas, isso não implicará necessariamente que os aumentos médios de experiência e capacitação numa escola influenciem para melhorar o desempenho educacional.

Possivelmente, se fossem estimados uma relação da formação específica do professor, ou seja, associar apenas os docentes com nível superior em matemática com o desempenho dos alunos nessa disciplina, talvez trouxesse resultados divergentes aos encontrados nesse estudo. Silva Filho (2019) aborda que uma titulação acadêmica mais elevada não torna, necessariamente, um professor eficiente, pois dado o cenário em que se encontra, o docente depende não só do conhecimento, mas também de práticas essenciais ao desempenho dos alunos. Aparentemente, isso poderia explicar um possível efeito positivo da formação específica do professor e os melhores meios de gerenciamento da sala de aula.

Em outro cenário, as características observáveis dos diretores demonstraram que o fato do mesmo possuir ensino superior tem uma relação positiva e tímida com o desempenho em português e matemática na ordem de 1,6 e 1,2 pontos no modelo de efeitos fixos. Por sua vez, diretores que exercem a função entre 6 a 10 anos, tem relação negativa na avaliação de português (-0,34) e insignificante em matemática. Para aqueles que tem mais de 10 anos de experiência controlados por efeitos fixos, o parâmetro não foi estatisticamente significativo para proficiência em português, somente teve uma correlação positiva e ínfima de 0,51 sobre o desempenho em matemática.

Elucidando os resultados correlatos às características escolares, o fato de a escola possuir biblioteca apresentou sinal positivo sobre o teste de língua portuguesa com significância de 5%, e foi insignificante sobre a performance em matemática. De acordo com os modelos de POLS e efeitos fixos, a escola, por possuir laboratório de informática e de ciências, tem efeito negativo, porém não tão relevante sobre as duas avaliações. Somente possuir quadra de esportes tem uma relação mínima, mas positiva e significativa a 1%. No que tange à quantidade de docentes que lecionam no ensino fundamental, seus coeficientes apresentaram efeitos positivos, mas pequenos, não chegando a pelo menos a 1 ponto em ambos os painéis.

Em outro enfoque, a quantidade de turmas do 5º ano do ensino fundamental denotou coeficientes negativos e significantes a 1% sobre as notas nos testes. O estudo de Franco e

Menezes Filho (2017) encontrou resultados negativos e insignificantes do número de turmas, tanto diurnas quanto noturnas, sobre o desempenho educacional. É plausível ressaltar que escolas que possuem mais turmas, possivelmente tem uma quantidade menor de alunos por sala. Contudo, quanto maior a quantidade de turmas, maior o desgaste dos professores. Nesse caso, o tamanho da turma é algo constantemente discutido na literatura, mas os achados divergem sobre a sua importância no aprendizado dos alunos.

Nas estimações do painel de efeitos fixos, as escolas que se localizam na área rural têm uma correlação negativa e significativa de aproximadamente 3,5 pontos no desempenho escolar das duas disciplinas supracitadas. Normalmente, as pesquisas que mensuram efeitos da variável localização realizam uma decomposição para analisar o diferencial de desempenho entre o setor rural e urbano. Particularmente, as discussões centram-se na ideia de que não a localização em si que afetaria as diferenças de resultados, mas aspectos como características dos alunos, das famílias, das escolas, dentre outros fatores.

Alguns estudos explanam que as diferenças de desempenho escolar urbano-rural estão relacionadas com as características socioeconômicas da família que, conseqüentemente, afeta os incentivos educacionais, e também questões relacionadas ao características das escolas, visto que as escolas rurais possuem uma infraestrutura inferior em comparação as urbanas (Ramos; Duque, Nieto, 2012; Amini; Nivorozhkin, 2015). Complementarmente, as escolas vivem diferentes realidades, especialmente quando é considerado aquelas situadas nas periferias ou em localidades mais afastadas do país. Tais escolas recebem alunos de alta vulnerabilidade, e, além de possuir um corpo docente instável, não usufruem de uma infraestrutura escolar que tenham condições adequadas de funcionamento (Setúbal, 2010).

2.5 Considerações finais

Com o intento de analisar os determinantes do desempenho escolar no Brasil, este trabalho utilizou os microdados do Saeb, advindos da Prova Brasil para o 5º ano do ensino fundamental e do Censo Escolar nos anos de 2011, 2013, 2015 e 2017. Por meio disso, foram incorporadas as variáveis relacionadas aos alunos, professores, diretor e escola, na qual se desenvolveu uma função de produção educacional com dados em painel considerando o modelo de efeitos fixos que controla as características não observáveis exclusivas de cada escola e constantes no tempo.

Em geral, observou-se que quase todas as variáveis foram altamente significativas para explicar o desempenho escolar, mesmo que grande parte delas veio a apresentar coeficientes bem pequenos. Os resultados mais importantes presentes na análise, tanto positivo quanto negativo, foram as características dos alunos e o *background* familiar. A saber, escolas com maior proporção de alunos que trabalham fora de casa, que já reprovaram e abandonaram uma ou mais vezes, indicam resultados negativamente mais severos sobre a proficiência em língua portuguesa e matemática. Por outro lado, a educação da mãe tem fortes efeitos positivos e significantes, tornando-se um fator imprescindível para a melhoria do desempenho escolar.

Constatou-se, significativamente, que o fato de os alunos possuírem computador em casa ajuda no seu desempenho, com coeficientes bem expressivos. Para tanto, as particularidades dos professores não demonstraram tão eficazes para desempenho nas avaliações, com exceção do aspecto salarial que gerou uma influência irrisória. No caso dos diretores, o tempo que ele exerce a função na escola constatou alguns coeficientes insignificantes e negativos, apenas o fato de possuir ensino superior contribui de forma positiva e significativa para a proficiência em língua portuguesa e matemática.

No tocante às análises relacionadas ao ambiente escolar, alguns destaques afetam positivamente o desempenho, como a escola deter quadra de esportes e o aumento na quantidade de professores em atividade. Por outro lado, o fato de a escola ter laboratório de informática e ciências pioram o desempenho dos alunos, da mesma forma que aumentar quantidade de turmas. Logo, assim como grande parte da literatura defende, os aspectos que fazem um bom professor, um bom diretor e uma boa escola não são suscetíveis a mensuração e não são explicadas por fatores observáveis como escolaridade e anos de experiência.

Em síntese, uma possível limitação do estudo é não poder ter usufruído de uma quantidade de período maiores e mais recentes para que os resultados encontrados possam representar a situação atual do país. No entanto, a conclusão do presente estudo é que os insumos escolares voltados aos aspectos dos professores, como o nível de formação e experiência, não agregaram valor, mas docentes com maiores níveis de salário tendem a contribuir no aumento dos resultados educacionais. Do mesmo modo, a infraestrutura escolar não é relevante para explicar o desempenho escolar. Não obstante, as características socioeconômicas dos alunos e o nível de escolaridade da mãe são as principais variáveis que maior explicam desempenho escolar no 5º ano avaliados nos testes de língua portuguesa e matemática.

Finalmente, embasando-se nas evidências apresentadas, políticas públicas voltadas a permanência dos alunos na escola devem ser incentivadas para melhoria dos resultados educacionais, sobretudo em escolas com estudantes em situações desvantajosas, como é o caso de escolas rurais. Por meio disso, mitigar o quantitativo de alunos que trabalham fora de casa, que já reprovaram ou abandonaram a escola. Logo, são necessários mais investimentos em merenda escolar, transporte escolar, bolsas de auxílio e programas de transferências de renda para alunos com maior vulnerabilidade social.

3 A RELAÇÃO DOS GASTOS PÚBLICOS POR ALUNO E QUALIDADE EDUCACIONAL DOS MUNICÍPIOS BRASILEIROS

3.1 Considerações iniciais

Nas últimas décadas, o sistema educacional brasileiro promoveu avanços significativos no acesso à educação para a população. De acordo com dados do Anuário da Educação Básica de 2021, o ensino fundamental praticamente alcançou a universalização, com 98% das crianças e jovens entre 6 e 14 anos frequentando a escola. Todavia, em números absolutos, o quantitativo de alunos que ainda não tem acesso à educação formal é excepcionalmente significativo. Houve também um aumento relevante do acesso à educação para a população brasileira de 18 a 29 anos com escolaridade média de 11,8 anos de estudo (Todos Pela Educação).

No entanto, mesmo com o avanço considerável na frequência escolar, muitos problemas ainda persistem, especialmente relacionados à qualidade do ensino. Os resultados do Pisa¹⁰, em 2022, comparado com 81 países, posicionam o Brasil entre a 44^a e 57^a posição no *ranking* de desempenho em leitura, entre a 62^o e 69^a em matemática e entre a 53^o e 64^o em ciências (MEC, 2023). Perante isso, uma das mudanças mais discutidas para contornar esse cenário é a destinação de recursos direcionados ao sistema educacional, uma vez que, se acredita que aumentos nos gastos refletem em melhorias no aprendizado escolar. Contudo, é essencial destacar que essa abordagem não se resume em apenas aumentar o orçamento, mas demanda investimentos estratégicos e políticas bem direcionadas, caso contrário, poder-se-á obter resultados controversos.

O investimento do Brasil em educação equivale a 6,3% do PIB¹¹ anual, porém, quando se refere ao valor absoluto por aluno na Educação Básica ainda é notavelmente inferior ao observado nos sistemas educacionais dos países desenvolvidos. Conforme os dados da OCDE, o gasto médio por aluno no ensino infantil e fundamental é 2,3 vezes maior nos países que englobam a organização do que no Brasil. No que se refere à educação superior, o Brasil possui gastos substancialmente maiores, visto que, enquanto os países da OCDE gastam 1,8 vez mais com as universidades, no Brasil, a diferença aumenta em

¹⁰ É uma avaliação trienal aplicada a alunos de 15 anos em cerca de 80 países, a maior parte deles é da própria OCDE. O Brasil participa como país convidado e a aplicação do exame, em 2018, envolveu 597 escolas públicas e privadas e 10.961 alunos brasileiros.

¹¹ Produto Interno Bruto

aproximadamente quatro vezes. Não obstante, o intuito não é gastar menos nessa área, mas expandir os recursos destinados à educação básica sem distorcer a alocação eficiente (Todos Pela Educação, 2021).

Um dos principais pesquisadores contemporâneos em Economia da Educação, Eric Hanushek, enfatiza consistentemente a ausência de correlação positiva entre o desempenho dos alunos e o volume de recursos destinados à educação (Card; Krueger, 1996). O autor já conduziu diversos estudos acerca da qualidade da educação e mantém um posicionamento constante quando avalia o grau de importância dos gastos educacionais na performance estudantil. Em contrapartida, existe uma corrente que aborda uma relação positiva entre esses fatores (Greenwald; Hedges; Laine, 1996).

Para Hanushek e Woessmann (2010), a melhoria nos resultados educacionais pode ser adquirida por intermédio do desempenho em avaliações padronizadas, uma vez que um ano de escolaridade não produz as mesmas habilidades cognitivas em todos os lugares, e usar essas medidas poderia assumir que elas são equivalentes. Além do mais, conhecimento e habilidades cognitivas indicam uma série de fatores fora da escola, como famílias, colegas, dentre outros. Ignorar tais fatores poderá causar um erro de medição nas análises de crescimento econômico.

Posto isso, analisar a relação entre gastos em educação e desempenho escolar é crucial para identificar se a aplicação de recursos públicos está sendo efetiva, de forma que reflita na qualidade da educação. Nesse contexto, com o intuito de contribuir para estudos na área que não encontram concordância em suas análises, esta pesquisa tem como objetivo analisar se os gastos públicos por aluno no ensino fundamental influenciam na qualidade educacional dos municípios brasileiros na rede pública de ensino, medido pelos testes de proficiência em língua portuguesa e matemática no 5º e a 9º ano no Saeb. Para tanto, foi construído um painel de dados abrangendo os anos de 2013, 2015, 2017 e 2019, com controle de efeitos fixos e constante no tempo.

A contribuição da pesquisa ou as lacunas que podem ser preenchidas condizem, por exemplo: em considerar quatro períodos mais recentes de análise, pois alguns estudos além de serem com anos mais antigos, utilizam menores períodos de tempo; fazer o uso somente das despesas com ensino fundamental em sua totalidade como variável explicativa do desempenho para o 5º e 9º ano; e a estratégia de compilação dos dados para todos os anos aplicados, de modo que foram utilizadas três fontes de dados.

Este capítulo contempla, além dessa seção introdutória, a seção 3.2 com as principais

literaturas que envolve o teor da pesquisa. A seção 3.3 apresenta os procedimentos metodológicos, com a descrição da base de dados e da estratégia empírica. A seção 3.4 realça os principais resultados da estatística descritiva e análise econométrica. Por fim, a seção 3.5 realiza as considerações finais do estudo.

3.2 A discussão acerca da qualidade educacional: Consensos e dissensos na literatura existente

A extensa literatura que versa sobre os desafios do Estado em aprimorar a qualidade da educação pública traz consigo conhecimentos consolidados. A ideia de que aumentar o investimento em educação é a solução para melhorar a baixa qualidade de ensino, não demonstra aderência nos estudos empíricos. Por meio dessa percepção, a presente seção está dividida em duas subseções. A primeira mostra as principais pesquisas que relacionam os recursos destinados à educação e ao desempenho escolar. A segunda avalia políticas públicas implementadas para a melhoria do desempenho escolar.

3.2.1 A relação entre recursos educacionais e desempenho escolar

Existe, na realidade, vários estudos que tiveram como objetivo explorar os fatores que melhor explicam os resultados educacionais, os quais são medidos por meio de testes padronizados, sendo uma das maneiras de representar a qualidade da educação. Explicitamente sobre as pesquisas que norteiam uma relação entre gastos e desempenho escolar, busca-se responder se o aumento de tais gastos tem algum efeito sobre a proficiência dos alunos, de forma que, mesmo com uma série de estudos que realizaram essa relação, não há consenso nos resultados encontrados. Portanto, independentemente das variáveis e dos métodos empregados, existe conclusões diferentes na literatura econômica.

Do ponto de vista empírico, Hanushek (1989) traça um arcabouço das primeiras evidências empíricas que relacionaram gastos com educação e resultados em avaliações externas. Embora grande parte dos estudos abordem uma associação positiva entre gastos e desempenho escolar, esse efeito desaparece quando se controla pelas diferenças das características da família. Portanto, considera que as variações nos gastos não estão sistematicamente correlacionadas com as variações no desempenho dos estudantes; a discrepância de aptidão dos professores não está relacionada aos antecedentes educacionais

ou tempo de experiência; e os professores mais capacitados não são mais remunerados do que os menos habilitados. O autor conclui que essas evidências mostram que as decisões escolares devem se distanciar de políticas tradicionais “destinadas aos insumos” e focar naquelas que oferecem incentivos de desempenho.

Opondo-se à essa abordagem, Greenwald, Hedges e Laine (1996) revisitaram uma gama de estudos sobre a função de produção educacional para usufruir de métodos meta-analíticos e, com base nisso, avaliar a magnitude da relação entre insumos escolares e performance escolar. Os artigos selecionados compilaram dados a nível de distrito ou unidades menores e empregavam característica socioeconômicas ou utilizavam-se de um desenho longitudinal. Por conseguinte, a análise indicou que uma variedade de recursos estava correlacionada positivamente com o sucesso educacional, com efeitos suficientes para aludir que aumentos moderados nos gastos estão relacionados com aumentos significativos no desempenho.

Em consonância com Hanushek e Kimko (2000), ao analisarem o efeito da qualidade da força de trabalho sobre o crescimento econômico, investigaram umnexo causal entre gastos escolares e performance dos alunos em diversas avaliações internacionais de desempenho acadêmico em matemática e ciências. Por meio de um painel com aproximadamente 70 países, foram utilizados dados do *International Association for Evaluation of Educational Achievement* (IEA) e do *International Assesment of Educational Progress* (IAEP). Em síntese, os autores evidenciaram que as variações dos gastos escolares não estão correlacionadas com a diferença de desempenho dos alunos.

Com o objetivo de explorar os principais determinantes da qualidade educacional, avaliada por meio de pontuações em testes internacionais, taxas de repetência e taxas de abandono, Lee e Barro (2001) empregaram um painel que incorpora medidas de *output* e *input* em diversos países. Os pesquisadores observaram que características familiares, como a renda e a educação dos pais, apresentam associação com o desempenho dos alunos. Descobriram também que a razão professor-aluno tem uma forte e positiva correlação com os resultados escolares. Por outro lado, os efeitos mais sutis, mas ainda positivos, surgem em relação ao salário médio dos professores e ao período escolar. Além disso, o PIB per capita apresentou uma correlação insignificante com o desempenho em avaliações de matemática e ciências, mas uma correlação positiva nos testes de leitura.

Por sua vez, Woessmann (2003) desenvolveu um modelo para elucidar o funcionamento dos sistemas educacionais. O autor buscou estimar o efeito do contexto

familiar, dos recursos e instituições escolares no desempenho em matemática e ciências em 39 países no período de 1994 e 1995. Utilizando dados individuais do *Trends in International Mathematics and Science Study* (TIMSS) com uma amostra representativa de 260.000 estudantes, foi empregado uma função de produção educacional estimada por Mínimos Quadrados Ponderados. Os resultados destacaram que as discrepâncias de desempenho entre países, em termos de qualidade educacional, não se correlacionavam com as diferenças de gastos com educação, mas com as divergências dos sistemas de ensinos, por exemplo: testes centrais; mecanismos de controle em assuntos curriculares e orçamentais; e autonomia escolar em processos e decisões de pessoal.

Em âmbito nacional, Amaral e Menezes Filho (2008) analisaram se os gastos com educação elevaram o desempenho dos alunos das 4ª e 8ª séries do ensino fundamental. Para mensurar esse efeito, os autores utilizaram dados das notas da Prova Brasil em língua portuguesa e matemática e das despesas com educação fundamental. Ao controlar pela escolaridade média da população, proporção de docentes com ensino superior, quantidade de horas-aula e *dummies* de unidade da federação, os autores encontraram efeitos pequenos e estatisticamente insignificantes dos gastos sobre o desempenho nos testes. Além disso, por meio de regressões quantílicas, encontraram que o efeito dos gastos sobre a performance escolar na quarta série ocorre somente nos municípios situados no quantis mais altos.

Com foco na qualidade educacional dos municípios da região Nordeste, Gonçalves e França (2013) verificaram a capacidade das políticas públicas educacionais em promover aumentos no desempenho no teste de matemática da Prova Brasil no ano de 2005 para os alunos dos anos finais de ensino. Para tal finalidade, foi utilizado uma função de produção educacional estimando um modelo multinível, isto é, com variáveis individuais dos alunos, variáveis relacionadas às escolas e características dos municípios que administram as escolas. Em suma, os autores observaram que os programas como Merenda Escolar e TV Escola impactam positiva e significativamente no desempenho dos alunos, assim também como variáveis referentes a democracia, crescimento e gastos em educação.

Menezes Filho e Oliveira (2014) conduziram uma pesquisa que analisou a relação entre gastos por aluno e a qualidade educacional nos municípios brasileiros. Nessa perspectiva, os autores aplicaram os dados do Saeb nos anos de 2005, 2007, 2009 e 2011, empregando métodos de regressão de dados em painel na análise da função de produção educacional. Em síntese, os resultados destacaram a existência de uma correlação positiva entre gastos por aluno e a performance dos estudantes em língua portuguesa e matemática.

No que concerne à Monteiro (2015), procurou analisar se o aumento dos gastos públicos em educação, proporcionados pelo aumento das receitas de *royalties*, impactou na quantidade e qualidade do ensino nos municípios brasileiros entre os anos de 2000 e 2010. Aplicando uma análise de regressão linear, os achados mostraram que o acréscimo das despesas está relacionado com a ampliação da população, mensurada em termos de anos de estudo e redução de analfabetismo. Porém, não houve reflexos de melhora na qualidade do sistema de ensino medida pelas notas da Prova Brasil no 5º e 9º ano do ensino fundamental.

Panassol (2018) buscou identificar se os maiores gastos no ensino fundamental melhoram a qualidade de ensino nos municípios do estado do Rio Grande do Sul. Para isso, utilizaram os resultados do Ideb no ano de 2015, aplicando variáveis de controle relacionadas aos alunos, às escolas, aos professores e à gestão municipal. A estratégia empírica correspondeu a uma estimação de modelos MQO e regressão quantílica baseados na função de produção de Hanushek e Luque (2003), atribuindo diferentes variáveis em cada um dos modelos. Os resultados mostraram que os gastos podem apresentar correlação positiva ou não com o desempenho escolar, dependendo das variáveis explicativas utilizadas. Precisamente, um aumento de 1% na média nas despesas com ensino fundamental, aumenta o Ideb em aproximadamente 0,009%, ou seja, não existe uma relevância econômica significativa no que concerne a relação entre gastos e desempenho educacional.

Já Kroth e Gonçalves (2019) buscaram analisar o impacto dos gastos com educação e gastos sociais (assistência social, saúde e cultura) sobre a qualidade educacional nos anos de 2007, 2009 e 2011 em 4655 municípios brasileiros. Como estratégia de identificação, os autores utilizaram dados dos testes de proficiência da Prova Brasil para o ensino fundamental aplicando variáveis instrumentais (GMM – *Generalized Method of Moments*) de dados em painel. Por conseguinte, os principais resultados apresentaram que os gastos educacionais e sociais municipais geraram efeitos positivos e significantes na determinação da qualidade da educação.

Após o aumento do investimento em educação proporcionados pelo o Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação (Fundeb), Alves e Frio (2022) analisaram a eficiência dos gastos municipais em educação dos municípios brasileiros sobre o Índice de Desenvolvimento da Educação Básica (Ideb). Metodologicamente, os autores empregaram fronteira estocástica variante e invariante no tempo para os anos iniciais e anos finais do ensino fundamental, considerando um painel de dados nos anos de 2011, 2013, 2015 e 2017. Os resultados enfatizaram uma

relação positiva entre os gastos defasados nos quatro anos e o Ideb, da mesma forma que a porcentagem de professores com ensino superior e mães com ensino superior. Contrariamente, variáveis como distorção série/idade e porcentagem de pessoas situadas na área rural reduzem a eficiência.

3.2.2 Políticas públicas e performance estudantil

Analisar a relação entre recursos e desempenho escolar pode gerar implicações bastante profundas para possíveis políticas públicas, de modo que a evidência ou não de sua importância poderá levar a diversos tipos de intervenções do poder público em algum setor específico. Partindo dessa percepção, sob à luz de políticas públicas em educação que visam estipular um tamanho máximo de alunos por sala, Krueger (1999) avaliou o Projeto STAR (*Student/Teacher Achievement Ratio*) na cidade de Connecticut, nos Estados Unidos, em que alunos e professores foram distribuídos de forma aleatória em turmas de tamanhos distintos. Os resultados demonstraram que, em média, a nota dos alunos nos testes padronizados aumenta em 4 pontos percentuais no primeiro ano onde foram realocados em turmas menores.

Nesse sentido, em Israel, Angrist e Lavy (1999) utilizaram dados de alunos da 4ª e 5ª série em 1991 e da 3ª série em 1992, aplicando regressão descontínua de inscrição conhecida como regra de Maimonides como um instrumento para o tamanho das salas. Os resultados apresentaram que uma redução no tamanho da turma induz um aumento significativo nas notas de leitura e matemática para os alunos da 5ª série e um aumento menor nas pontuações de leitura para os alunos da 4ª série. Logo, os resultados dos testes da 3ª série mostraram pouca evidência de qualquer relação com o tamanho da turma.

Em um estudo adicional, Angrist e Lavy (2002) examinaram o impacto de um programa que disponibilizou computadores e treinamento para professores do ensino fundamental e médio em escolas israelenses. Os autores usaram um teste aplicado aos alunos em junho de 1996, administrado pelo *National Institute for Testing and Evaluation* (NITE) utilizando estimativas de *Ordinary Least Squares* (OLS) para investigar o efeito do *Computer-Aided Instruction* (CAI), além de ter sido desenvolvido um estimador de Variáveis Instrumentais não-lineares. Resumidamente, os resultados sugeriram um efeito negativo não significativo no desempenho escolar nas avaliações de matemática para alunos da 4ª série e 8ª série.

No mesmo panorama, tendo em consideração as políticas de inclusão digital visando

a melhoria do processo de aprendizagem nas escolas, Goolsbee e Guryan (2006) avaliaram o impacto de um subsídio conhecido como *E-Rate* implementado pelo Governo Federal dos Estados Unidos nas escolas públicas do estado da Califórnia em 1998. Estimando uma equação de investimentos com dados de todas as escolas entre o período de 1996 e 2000, os resultados do estudo apresentaram que o programa teve impacto significativo no que condiz à expansão da infraestrutura de internet. Por outro lado, tão pouco foram encontradas evidências de efeito sobre o aprendizado dos alunos.

Uma política que foi adotada em diversas partes do mundo relacionada ao aumento de gastos em educação, foi as políticas de incentivos financeiros ligados ao desempenho dos professores. Nos Estados Unidos, Hanushek e Raymond (2004) analisaram a eficácia dessas políticas de responsabilização (*accountability*) no crescimento de desempenho dos alunos do *National Assessment of Educational Progress* (NAEP) na década de 1990. Os autores concluem que os sistemas *accountability* melhoram as pontuações dos estudantes, porém tende a ampliar o hiato entre brancos e negros nas pontuações.

Retornando à literatura de Tecnologias da Informação e Comunicação (TIC's), Leuven *et al.* (2007) avaliaram uma política de subsídios para estudantes desfavorecidos que foi implementada no ensino fundamental na Holanda. Diante disso, o primeiro subsídio tinha a finalidade de melhorar as condições de trabalho dos professores, ao mesmo tempo que o segundo fornecia computadores e *softwares* para as escolas com mais de 70% dos estudantes em situação de vulnerabilidade. Para mensurar tal efeito, os autores empregaram estratégias de regressão descontínua e diferenças em diferenças para obter estimativas mais precisas. Concisamente, os resultados indicaram efeitos negativos dos computadores na performance dos alunos em testes escolares.

Dando ênfase as políticas de valorização salarial ocorridas no Brasil em 1998, Menezes Filho e Pazello (2007) procuraram aferir se o aumento dos salários dos professores, proporcionado pelo Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (Fundef), impactou na proficiência dos alunos da 8ª série do ensino fundamental das escolas públicas. Para isso, utilizaram dados do Saeb baseando-se na metodologia de diferenças em diferenças comparando o desempenho de alunos entre escolas públicas e privadas nos anos de 1997 e 1999. Os resultados apresentaram que o aumento dos salários relativos entre os municípios da rede pública impactou positivamente no desempenho dos alunos.

Em paralelo ao estudo de Angrist e Lavy (1999), Matavelli e Menezes Filho (2020)

avaliaram uma política pública que delimitou um número máximo de alunos por turma nos estados de São Paulo, Santa Catarina e Minas Gerais no ano de 2015. Metodologicamente, utilizaram dados do Saeb para o 5º e 9º ano usando o método de regressão descontínua *fuzzy*, com o tamanho da sala predito pela função de Maimonides como instrumento para o tamanho da sala observado. Os achados demonstraram que não há efeito estatisticamente significativo do tamanho da sala sobre a nota dos alunos.

Haja vista a divergência da literatura no que tange os efeitos do aumento de recursos sobre a performance dos estudantes, Rocha, Belluzzo e Nicolella (2022) avaliaram o impacto do aumento dos recursos financeiros proporcionados pelo Programa de Transferência de Renda Financeira (PTRF) sobre o desempenho dos alunos no Saeb no ano de 2007 das escolas públicas na cidade de São Paulo. Com a aplicação do método de regressão descontínua, os resultados mostraram que a política impactou de forma positiva e significativamente o desempenho da 8ª série em língua portuguesa e matemática para as escolas com menores números de matrículas. Por outro lado, não foram encontrados impactos significantes para a 4ª série, mesmo para as escolas menores, e tão pouco para as escolas com maiores números de matrículas.

3.3 Metodologia

Esta seção destina-se a apresentar as principais fontes de dados aplicados na pesquisa juntamente com a descrição das variáveis, bem como descrever o modelo econométrico para os dados em painel e os testes de especificação.

3.3.1 Base de dados e descrição das variáveis

A base de dados foi compilada a partir de três fontes principais referentes aos anos de 2013, 2015, 2017 e 2019. Elas incluem: o Sistema de Avaliação da Educação Básica (Saeb), principal mecanismo de análise da qualidade do ensino fundamental e do ensino médio em âmbito municipal, estadual e federal; o Censo Escolar, é o principal instrumento de coleta de informações da educação básica e mais importante pesquisa estatística educacional brasileira; e o Sistema de Informações Contábeis e Fiscais do Setor Público

Brasileiro (Siconfi)¹², do Tesouro Nacional, que disponibiliza as despesas municipais por função.

Levando em consideração essas informações, a descrição das variáveis dependentes e de controle estão explícitas no Quadro 2. Nesse contexto, a qualidade da educação é mensurada pelo resultado dos testes de proficiência no Saeb em língua portuguesa e matemática para os 5º e 9º ano do ensino fundamental. Embora a qualidade possa envolver diversas dimensões, essa escolha se dá em detrimento da disseminação de seu uso na literatura internacional e nacional, como é destacado por Hanushek e Woessmann (2010). Os testes de proficiência são uma boa *proxy* para estimar a qualidade de educação. Vale ressaltar que os dados foram compilados em nível municipal, considerando os resultados das escolas públicas municipais e estaduais¹³, desconsiderando, dessa forma, as escolas federais e privadas.

Para construir a variável de gastos públicos municipais por aluno, foram utilizados os dados fornecidos pelo Siconfi/Tesouro Nacional, precipuamente das despesas municipais por função, considerando exclusivamente as despesas pagas¹⁴ com o ensino fundamental de cada ano estudado. Após a obtenção dos gastos em educação, estes foram ajustados pelo quantitativo de alunos matriculados, conforme os dados do Censo Escolar. Cabe enfatizar que serão aplicados nas estimações o mesmo quantitativo de despesas para explicar o desempenho em português e matemática para as turmas de 5º e 9º ano do ensino fundamental.

A variável de gastos públicos em saúde per capita foi elaborada utilizando dados do Siconfi/Tesouro Nacional, nos quais os gastos foram divididos pela população total dos municípios, obtida por meio dos dados do IBGE. O intuito é que um investimento maior no sistema de saúde possa resultar em melhorias no bem-estar, na capacidade física e intelectual, e, por conseguinte, no aprendizado dentro do município. De acordo com Hanushek e Woessmann (2010), saúde e nutrição são fatores que impactam as habilidades cognitivas dos indivíduos.

A razão professor-aluno no ensino fundamental é derivada dos dados do Censo Escolar para os anos investigados. Os dados dessa fonte estão em nível escolar, sendo realizado uma média a nível município contemplando todos os professores que lecionam na

¹² Disponível em: <https://siconfi.tesouro.gov.br/siconfi/index.jsf>

¹³ Visto que há escolas estaduais que oferecem ensino fundamental.

¹⁴ Corresponde às despesas para as quais os pagamentos foram efetuados. Então, após uma despesa ser empenhada e, posteriormente, liquidada, o pagamento é processado para finalizar a transação. Assim, quando um pagamento é feito, a despesa é considerada paga.

educação fundamental. Ampla e frequentemente usada em estudos que analisam os impactos do tamanho das turmas no desempenho dos alunos, essa variável é essencial para compor a função de produção educacional. Os resultados obtidos a partir dessa variável possibilitam informar políticas educacionais e direcionar a alocação de recursos visando melhorias na qualidade do sistema educacional.

Quadro 2 – Descrição das variáveis

Variáveis	Descrição	Fonte
Variáveis dependentes		
nota_mt5	Proficiência em matemática para o 5º ano do ensino fundamental	Saeb
nota_mt9	Proficiência em matemática para o 9º ano do ensino fundamental	Saeb
nota_lp5	Proficiência em língua portuguesa para o 5º ano do ensino fundamental	Saeb
nota_lp9	Proficiência em língua portuguesa para o 9º ano do ensino fundamental	Saeb
Variáveis explicativas		
GEA	Gastos públicos por aluno no ensino fundamental	Siconfi e Censo Escolar
GSpC	Gastos públicos com saúde per capita	Siconfi e IBGE
prof_aluno	Razão professor-aluno no ensino fundamental	Censo Escolar
mae_sup	Quantidade de mães com ensino médio/superior no 5º e 9º ano	Saeb

Fonte: Elaboração própria com base nos dados da pesquisa.

A quantidade de mães com ensino médio ou superior foi extraída dos dados do Saeb, para o 5º e 9º ano do EF, ou seja, será controlada mediante qual for a variável dependente. Desse modo, a literatura econômica esboça que a educação dos pais e o ambiente familiar desempenham papéis significativos na trajetória escolar e desenvolvimento dos estudantes. Com essa afirmação, Carneiro e Heckman (2003) mostram que a família afeta tanto as características cognitivas quanto não cognitivas dos seus membros, com efeitos sobre o sucesso escolar e socioeconômico.

Para tanto, é fundamental destacar que as variáveis em termos monetários foram deflacionadas de acordo com o Índice de Preços ao Consumidor Amplo (IPCA) referente ao ano de 2019. Ademais, todas as variáveis passaram por um processo de linearização para lidar com problemas de normalidade, especialmente devido à presença de *outliers*, como é o caso de grandes municípios. Essa transformação permite examinar os coeficientes em termos percentuais facilitando a interpretação dos resultados.

3.3.2 *Função de Produção Educacional*

A questão educacional não se refere apenas ao quantitativo de anos de estudo, mas especialmente à qualidade dos anos de escolaridade existentes. A forma mais geral para se medir os determinantes do desempenho escolar é por meio da função de produção educacional, apresentada, a princípio, por Coleman (1968). Desde esse período, tanto a literatura internacional quanto a literatura nacional têm se dedicado em estabelecer qual a finalidade ou produto de um sistema de ensino: maximizar o desempenho médio dos estudantes, reduzir as disparidades de conhecimento entre aluno, oferecer capacidades cognitivas gerais ou desenvolver competências e habilidades necessárias para o mercado de trabalho (Waltenberg, 2006).

Os estudos trouxeram analogia da teoria da firma, buscando compreender a tecnologia de combinar os insumos escolares e familiares, de modo que o resultado educacional seja maximizado. A literatura econômica apresenta uma variedade de variáveis que podem ser incorporadas na função de produção da educação que podem exercer influência nos resultados educacionais (Todd; Wolpin, 2003). Nessa percepção, Hanushek e Luque (2003) atribuem a seguinte estrutura geral da estimativa dessa função:

$$Y = f(X, R) \quad (3)$$

em que, Y equivale ao resultado educacional, R representa os *inputs* associados aos recursos escolares e X corresponde ao conjunto de elementos que influenciam a educação, especialmente, nível de escolaridade da família (Lee; Barro, 2001).

A literatura empírica constata os testes de proficiência como uma boa *proxy* para mensurar a qualidade educacional de um país, pois conseguem agregar muito do conhecimento cognitivo dos alunos, não somente aqueles conquistados no ambiente escolar (Hanushek, 2006; Hanushek; Wossmann, 2007, 2010). Segundo Hanushek (2006), essa forma de medir desempenho está correlacionada com a produtividade, rendimentos individuais e o crescimento do produto de um país.

3.3.3 *Modelo de dados em painel*

O modelo de dados em painel, também conhecido como modelo de dados longitudinais ou modelo de efeitos fixos/aleatórios, é utilizado para analisar dados advindos de unidades observadas ao longo do tempo, combinando dados em *cross section* (várias

unidades em um único ponto no tempo) e dados de séries temporais (observações repetidas durante o período de tempo para as mesmas unidades) (Wooldridge, 2006). Dada a disponibilidade de um conjunto de dados em painel, a estimação do modelo básico de efeitos não observados traz a vantagem de se eliminar da regressão efeitos fixos no tempo que podem influenciar a variável resposta que não foram incorporadas. Desse modo, o modelo pode ser descrito como:

$$Y_{it} = X_{it}\beta + c_i + u_{it}, \quad \text{para } t = 1, 2, \dots, T. \quad (4)$$

Para analisar a relação entre gastos públicos por aluno e qualidade educacional, emprega-se o modelo econométrico de dados em painel com efeitos fixos, pois traz vantagens de se controlar diferenças não observadas e invariáveis ao longo do tempo. É válido mencionar que as variáveis foram logaritmizadas, então, os seus coeficientes representarão suas respectivas elasticidades. Dito isso, a equação segue a seguinte estrutura:

$$\ln(nota_{it}) = \beta_1 \ln(GEA_{it}) + \beta_2 \ln(GSpc_{it}) + \beta_3 \ln(prof_aluno_{it}) + \beta_4 \ln(mae_sup_{it}) + c_i + u_{it} \quad (5)$$

em que, $nota_{it}$ são as notas em matemática e língua portuguesa para o 5º e 9º ano do ensino fundamental no município i no tempo t ; GEA_{it} é um vetor dos gastos por aluno no município i no tempo t ; $GSpc_{it}$ são os gastos públicos com saúde per capita no município i no tempo t ; $prof_aluno_{it}$ é a razão professor-aluno no município i no tempo t ; mae_sup_{it} é a quantidade de mães com ensino médio ou superior no município i no tempo t ; c_i é o efeito específico não observado do município; e u_{it} são os distúrbios idiossincráticos, pois se alteram tanto em t como em i . Os parâmetros de interesse $\beta_1, \beta_2, \beta_3, \beta_4$, equivalem às relações condicionais dos insumos sobre o desempenho educacional.

3.3.4 Testes de especificação do painel

Nesta parte serão destacados os testes de especificações que atribuem uma melhor escolha para o tipo de painel. Logo, são descritos os testes de Breush-Pagan, Teste F e Hausman.

3.3.4.1 Teste Breush-Pagan: efeitos aleatórios versus pooled

O teste de *Breush-Pagan* é usado para determinar qual dos modelos é o mais preferível, o modelo *pooled* ou modelo de efeitos aleatórios. Simbolicamente, $H_0: \sigma_n^2 = 0$

(constante comum – *pooled*) e a $H_1: \sigma_n^2 \neq 0$ (efeitos aleatórios), nesse caso o teste se dará pela seguinte relação:

$$LM = \frac{NT}{2(T-1)} \left[\frac{\sum_{i=1}^N (\sum_{t=1}^T \widehat{w}_{it})^2}{\sum_{i=1}^N \sum_{t=1}^T \widehat{w}_{it}^2} \right] \sim \chi_1^2,$$

Logo, N representa o número de municípios, T os períodos de tempo. Se $LM > \chi_1^2$, considera-se os efeitos aleatórios.

3.3.4.2 Teste F : *pooled versus efeitos fixos*

Este teste é utilizado para decidir qual dos modelos é mais adequado na estimação entre o modelo *pooled* ou modelo de efeitos aleatórios. Na hipótese nula (hipótese *pooled*), admite-se a homogeneidade na constante e , na hipótese alternativa, afirma-se a heterogeneidade na constante (efeitos fixos). Então, a $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_n$ e $H_1: \alpha_1 \neq \alpha_2 \neq \dots \neq \alpha_n$, a estatística F aplicada para testar essa hipótese é:

$$F_{stat} = \frac{\frac{R_{fe}^2 - R_{pool}^2}{N-1}}{\frac{1 - R_{fe}^2}{(NT - N - K)}} \sim F_{(N-1, NT-N-K)}$$

em que R_{fe}^2 e R_{pool}^2 são os respectivos coeficientes de determinação do modelo com constante comum e com efeitos fixos, N é o número de municípios, T é os períodos de tempo e K é o número de variáveis explicativas. Desse modo, os efeitos fixos são preferíveis se $F_{stat} > F_{(N-1, NT-N-K)}$.

3.3.4.3 Teste de Hausman: *efeitos fixos versus efeitos aleatórios*

O teste de Hausman é utilizado para decidir entre os modelos de efeitos aleatórios H_0 e efeitos fixos H_1 é o mais adequado. Na hipótese nula, os estimadores do modelo com efeitos aleatórios (estimação GLS) são consistentes e eficientes ($H_0: Cov(\eta, X_{it}) = 0$), já sob a hipótese alternativa, o modelo de efeitos aleatórios não é adequado, mas os efeitos fixos são consistentes ($H_1: Cov(\eta, X_{it}) \neq 0$). Dito isso, a estatística aplicada para realizar o teste de hipóteses é a seguinte:

$$H = (\hat{\beta}_{fe} - \hat{\beta}_{re})' [Var(\hat{\beta}_{fe}) - Var(\hat{\beta}_{re})]^{-1} (\hat{\beta}_{fe} - \hat{\beta}_{re}) \sim \chi_K^2$$

em que $\hat{\beta}_{fe}$ e $\hat{\beta}_{re}$ são os respectivos vetores dos estimadores dos efeitos fixos e aleatórios,

$Var(\hat{\beta}_{fe})$ e $Var(\hat{\beta}_{re})$ são as matrizes de variância-covariância dos estimadores e K é o número de regressores. Se $H > x_k^2$, rejeita-se o modelo com efeitos aleatórios. Logo, o modelo com efeitos fixos é mais apropriado.

3.4 Resultados e discussão

Nesta seção, serão apresentados os principais resultados encontrados por meio da aplicabilidade dos métodos estabelecidos. Discutir-se-á, inicialmente, algumas evidências iniciais a partir da estatística descritiva e, em seguida, os resultados das estimações para as disciplinas de português e matemática.

3.4.1 Estatística descritiva

Dando início às investigações, a Tabela 6 fornece um resumo estatístico abrangente do desempenho em português e matemática no 5º e 9º do ensino fundamental, com a frequência absoluta e relativa, média e desvio padrão entre o período de 2013 e 2019. Perante o exposto, é perceptível que os resultados nos exames de proficiência estão, em média, mais elevados com o passar do tempo nas respectivas disciplinas. A variação das notas correspondeu a um aumento de 6,4% e 4,4% no 5º ano do EF, e 3,6% e 2,5% no 9º ano do EF para português e matemática, respectivamente. Embora a variação do rendimento educacional seja consideravelmente pequena, a turma do 5º ano do EF possui um maior percentual de evolução na performance no Saeb dentro do intervalo em estudo.

É factível compreender que a quantidade de municípios presentes na base de dados é diferente para cada ano de análise, com destaque para o período de 2017, que contém a maior quantidade observações. Desse modo, de um quantitativo de 5570 municípios, houve uma perda irrisória de algumas observações em detrimento de dados ausentes em algumas variáveis, com ênfase principalmente em municípios mascarados, ou seja, aqueles que decidiram não revelar seus resultados nos testes de proficiência. Nesse caso, os dados estão desbalanceados, mas compõem quase todos os municípios em sua totalidade.

Tabela 5 – Estatística descritiva da proficiência em Língua Portuguesa e Matemática para o 5º e 9º ano entre 2013 e 2019

Ano/ série	Frequência absoluta	Frequência relativa	Língua Portuguesa		Matemática	
			Média	Desvio padrão	Média	Desvio padrão
5º ano						
2013	5.349	96,03%	189,25	23,48	206,45	27,61
2015	5.287	94,92%	200,79	21,88	214,58	23,14
2017	5.505	98,83%	207,46	22,41	218,53	24,17
2019	5.485	98,47%	207,85	22,07	222,35	23,86
Total	21.626	-	201,42	23,69	215,55	25,44
9º ano						
2013	5.415	97,22%	237,67	17,76	244,25	19,80
2015	5.245	94,17%	244,90	17,27	249,80	17,40
2017	5.473	98,25%	250,59	17,94	251,15	19,89
2019	5.393	96,82%	252,10	16,73	256,39	19,16
Total	21.526	-	246,33	18,34	250,40	19,58

Fonte: Elaboração própria a partir dos dados do Saeb de 2013 a 2019.

A Tabela 6 realça o número de observações *missing* contidas na base de dados entre o hiato de 2013 e 2019. Para tanto, as variáveis de desempenho no 5º e 9º ano do EF demonstraram que o ano de 2015 foi o que realçou a maior quantidade de informações ausentes, com 283 e 325 observações, respectivamente. O maior número de *missings* das despesas com educação e saúde foram em 2013 com 318 e 228 observações, já a educação da mãe correspondeu ao ano de 2015 com 263 e 415 informações para o 5º e 9º ano, respectivamente. Vale mencionar que a variável razão professor-aluno contemplou todos os municípios da amostra.

Tabela 6 – Quantidade de *missing* para cada variável entre 2013 e 2019

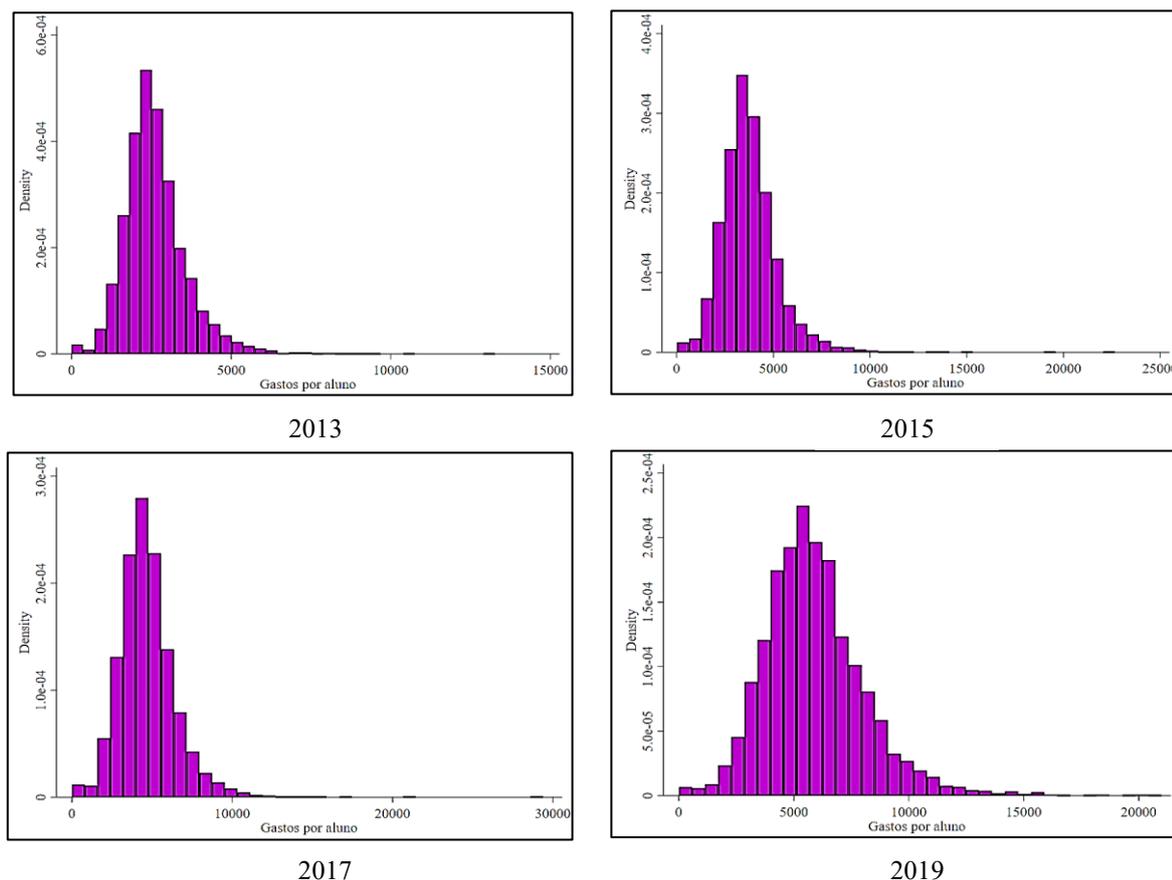
Variável	2023	2015	2017	2019	Total
Língua portuguesa 5º ano	221	283	65	85	654
Matemática 5º ano	221	283	65	85	654
Língua portuguesa 9º ano	155	325	97	177	754
Matemática 9º ano	155	325	97	177	754
Gastos com EF por aluno	318	281	151	260	1.010
Gastos com saúde per capita	228	164	69	198	659
Educação da mãe 5º ano	211	263	58	86	618
Educação da mãe 9º ano	282	415	100	280	1.077
Razão professor-aluno	0	0	0	0	0

Fonte: Elaboração própria com base nos dados da pesquisa.

O Gráfico 4 realça a distribuição dos gastos municipais por aluno no ensino fundamental entre o período de 2013 e 2019. O sumário estatístico referente a essa variável apresentou que a média em 2013 foi de R\$ 2.629,60 unidades monetárias e desvio padrão

R\$ 1.017,04; em 2015 equivaleu a R\$ 3.773,34 e desvio padrão R\$ 1.481,56; já em 2017 foi de R\$ 4.687,14 e desvio padrão R\$ 1.777,89; e em 2019, essa média correspondeu a R\$ 5904,24 e desvio padrão R\$ 2.225,45. A distribuição assimétrica significa que a maioria das observações está concentrada em valores mais baixos de gastos com educação por aluno, enquanto há poucas observações em valores mais altos.

Gráfico 4 – Histograma da distribuição dos gastos por aluno entre 2013 e 2019



Fonte: Elaboração própria com base nos dados da pesquisa.

Para fortalecer os resultados encontrados, no Apêndice A, são encontrados os resultados para uma amostra de municípios balanceada com 4.027 observações para cada ano, totalizando em 16.180 observações no painel. Por conseguinte, serão apresentados os resultados da estimação do modelo para o painel desbalanceado para usufruir de uma maior quantidade de municípios presentes na amostra.

3.4.2 Estimação do modelo

Os resultados das estimações dos modelos POLS, efeitos aleatórios e efeitos serão apresentados lado a lado para permitir equiparar entre as diferentes abordagens de dados em painel. Com isso, a escolha do modelo mais adequado depende de várias considerações, incluindo a estrutura dos dados e a relevância teórica. É de praxe na literatura o uso de efeitos fixos, visto que é mais robusto à presença de correlação entre os efeitos individuais e as variáveis explicativas. Outrossim, a significância do teste Hausman também é essencial para validar a escolha entre os modelos, o qual mostrou que o modelo de efeitos fixos é mais apropriado.

Tabela 7 – Testes de especificação

Testes	5º ano		9º ano	
	Língua Portuguesa	Matemática	Língua Portuguesa	Matemática
Breush-Pagan	125,45 (0,000)	127,61 (0,000)	106,85 (0,000)	115,85 (0,000)
Teste F	14,331 (0,000)	15,80 (0,000)	8,5144 (0,000)	11,256 (0,000)
Hausman	995,79 (0,000)	1.308 (0,000)	381,5 (0,000)	729,91 (0,000)

Fonte: Elaboração própria.

Nota: *p-valores* entre parênteses.

Diante disso, é importante ressaltar, primeiramente, que as variáveis estão logaritmizadas, significando que os coeficientes representam as suas respectivas elasticidades. Sendo assim, a Tabela 7 revela os resultados dos modelos para dados em painel em português e matemática no 5º ano do EF, em que se destaca, de imediato, que todos os coeficientes foram significantes a um nível de 1%. Como uma forma de comparação para os demais modelos, os resultados do modelo POLS denotaram que os gastos com EF por aluno influenciam de forma negativa e significativa os resultados em português e matemática em aproximadamente 0,021% quando controlado pelas demais variáveis.

Em seguida, os resultados do modelo de efeitos aleatórios tratam das diferenças individuais de cada município como uma variável aleatória, em que são frequentemente utilizados quando se supõe que as características específicas não observadas estão correlacionadas com a variável explicativa, mas não são de interesse direto. Diante disso, os coeficientes dos gastos com EF por aluno apresentaram sinais positivos, implicando que aumentos nas despesas induz a um incremento de 0,017% e 0,012% no desempenho em português e matemática, respectivamente.

No modelo de efeitos fixos, isto é, quando se controla as características não

observáveis e constantes no tempo de cada município, os resultados foram um pouco mais expressivos quando equiparado ao modelo anterior. A estimativa demonstrou uma relação positiva entre os gastos com EF por aluno e as notas em português e matemática de aproximadamente 0,02%, ou seja, aumentos nas despesas tendem a elevar o desempenho municipal nessas disciplinas. De todo modo, embora os coeficientes indiquem estatísticas significantes diferentes de zero, a magnitude desses efeitos é consideravelmente pequena.

Vale ressaltar também os coeficientes das demais variáveis de controle nas especificações do modelo de efeitos fixos. Os gastos com saúde per capita possui coeficientes estatisticamente significantes e positivos, apresentando, por sua vez, que o reflexo da melhora desses gastos influencia na performance municipal em ambas as disciplinas avaliadas, com 0,065% em português e 0,049% em matemática. Após o controle dos fatores não observáveis e constantes no tempo, os coeficientes da variável razão professor-aluno mostraram uma influência positiva sobre o rendimento nas avaliações em 0,023% e 0,016% em ambas as disciplinas.

Em sequência, com a finalidade de controlar pela escolaridade dos pais, devido ao respaldo na literatura, os coeficientes da quantidade de mães com ensino médio/superior se manifestaram de forma positiva e significativa em 0,031% e 0,025% em português e matemática, respectivamente, no modelo de efeitos fixos. Tais resultados refletem o efeito geracional proveniente da educação, que sinalizam que pais mais instruídos tendem a ter filhos mais educados.

Tabela 8 – Resultados do modelo para a proficiência em língua portuguesa e matemática no 5º ano do ensino fundamental

Variáveis	Língua portuguesa			Matemática		
	POLS	Efeitos aleatórios	Efeitos fixos	POLS	Efeitos aleatórios	Efeitos fixos
Gastos com EF por aluno	-0,021*** (0,001)	0,017*** (0,001)	0,027*** (0,001)	-0,025*** (0,002)	0,012*** (0,001)	0,022*** (0,001)
Gastos com saúde per capita	0,119*** (0,002)	0,077*** (0,001)	0,065*** (0,001)	0,112*** (0,002)	0,062*** (0,001)	0,049*** (0,001)
Razão professor/aluno	0,094*** (0,003)	0,058*** (0,003)	0,023*** (0,004)	0,093*** (0,003)	0,052*** (0,003)	0,016*** (0,004)
Educação da mãe	0,026*** (0,001)	0,027*** (0,001)	0,031*** (0,001)	0,022*** (0,001)	0,021*** (0,001)	0,025*** (0,001)
Constante	4,898*** (0,015)	4,751*** (0,012)	-	5,052*** (0,015)	4,956*** (0,012)	-
Observações	20.548	20.548	20.548	20.548	20.548	20.548
R ²	0,304	0,870	0,405	0,266	0,864	0,296

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Erros padrões robustos de White. (2) *p<0.1; **p<0.05; ***p<0.01.

A Tabela 8 traça os resultados dos modelos para o desempenho em português e matemática no 9º ano do EF, enfatizando que os resultados encontrados para essa turma apresentam uma correlação mais fraca quando se equipara com o 5º ano do EF. Da mesma forma que nas evidências anteriores, os resultados para o modelo POLS mostraram uma relação negativa e significativa dos gastos com EF por aluno e a performance educacional nos exames de proficiência. Ao controlar por efeitos fixos, os achados sugeriram que um aumento de 1% nas despesas com educação eleva o desempenho dos alunos em 0,017% e 0,014% em português e matemática, respectivamente.

As outras descobertas continuam destacando a relevância das despesas com saúde e da educação das mães para melhorar a qualidade educacional, com coeficientes positivos e significantes. Nessa perspectiva, um aumento de 1% nos gastos com saúde aumenta o desempenho em 0,04% em português e 0,03% em matemática quando controlado por efeitos fixos e constantes no tempo de cada município. A quantidade de mães que possuem ensino médio e superior influenciam em 0,014% e 0,012% nos resultados dos testes das respectivas disciplinas. Em contrapartida, a relação professor-aluno apresentou coeficientes positivos, porém insignificantes, tanto em português quanto em matemática, não podendo, dessa forma, tirar conclusões sobre tais resultados.

Tabela 9 – Resultados do modelo para a proficiência em língua portuguesa e matemática no 9º ano do ensino fundamental

Variáveis	Língua portuguesa			Matemática		
	POLS	Efeitos aleatórios	Efeitos fixos	POLS	Efeitos aleatórios	Efeitos fixos
Gastos com EF por aluno	-0,005*** (0,001)	0,010*** (0,001)	0,017*** (0,001)	-0,011*** (0,001)	0,007*** (0,001)	0,014*** (0,001)
Gastos com saúde per capita	0,065*** (0,001)	0,050*** (0,001)	0,042*** (0,001)	0,065*** (0,001)	0,040*** (0,001)	0,030*** (0,001)
Razão professor/aluno	0,040*** (0,002)	0,026*** (0,002)	0,004 (0,003)	0,050*** (0,002)	0,030*** (0,002)	0,002 (0,003)
Educação da mãe	0,013*** (0,0004)	0,012*** (0,001)	0,014*** (0,001)	0,009*** (0,0004)	0,008*** (0,001)	0,012*** (0,001)
Constante	5,204*** (0,010)	5,137*** (0,009)	-	5,307*** (0,011)	5,268 (0,009)	-
Observações	20.073	20.073	20.073	20.073	20.073	20.073
R ²	0,234	0,951	0,294	0,211	0,948	0,207

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Erros padrões robustos de White. (2) *p<0.1; **p<0.05; ***p<0.01.

Em síntese, os resultados das estimações demonstram a significância das variáveis

independentes para explicar a qualidade da educação. Mesmo com a significância, os coeficientes das despesas com educação apresentaram-se muito baixos, já as despesas com saúde denotaram um retorno maior. Essa evidência pode induzir a uma complementariedade desses dois tipos de gastos para proporcionar melhoria na qualidade educacional dos municípios, na medida em que estimula melhores condições aos alunos e à escola.

A partir das evidências explícitas, permite-se inferir que as despesas com educação fundamental influenciam no desempenho escolar. Independentemente dos métodos e variáveis aplicadas nas análises, tais resultados não vão ao encontro de alguns estudos supracitados no primeiro momento sobre a relação dessas duas variáveis. O sinal dos coeficientes positivos e a significância estatística não estão de acordo com aqueles autores que não encontraram uma relação positiva entre gastos educacionais e qualidade no ensino (Hanushek, 1986, 1989, 1996, 2002; Hanushek; Kimko, 2000; Woessmann, 2003; Amaral; Menezes Filho, 2008; Monteiro, 2015). Mas estão correlatos com alguns estudos nacionais (Menezes Filho; Pazello, 2007; Menezes Filho; Oliveira, 2014; Kroth; Gonçalves, 2019; Alves, Frio, 2022).

Adicionalmente, a Tabela 9 traça os resultados dos gastos com educação fundamental defasados em um período ($t-1$) sobre o desempenho do 5º ano do EF em português e matemática. Essa abordagem visa identificar se as despesas tendem a ter um efeito cumulativo ao longo do tempo, ou seja, os recursos investidos podem afetar o desempenho educacional com o passar do tempo. Entretanto, ao se utilizar o estimador de efeitos fixos, as estimativas de todas as variáveis tem uma correlação menor quando comparado aos resultados sem defasagem.

Com coeficientes um pouco mais baixos, a associação entre os gastos com EF defasados para o 5º ano no modelo de efeitos fixos correspondeu a 0,016% e 0,013% em português e matemática, respectivamente. Os coeficientes dos gastos com saúde se apresentam análogos nas duas disciplinas em 0,03% de correlação com a variável resposta, já a educação das mães influencia em 0,017% em português e um efeito quase nulo em matemática. Logo, apesar das estimativas serem quase nulas, ainda continuam positivas e significantes a um nível de 1%, com exceção da razão professor-aluno que não demonstrou estatisticamente significativa.

Tabela 10 – Resultados do modelo para a proficiência em língua portuguesa e matemática no 5º ano do ensino fundamental com gastos em educação defasados

Variáveis	Língua portuguesa			Matemática		
	POLS	Efeitos aleatórios	Efeitos fixos	POLS	Efeitos aleatórios	Efeitos fixos
Gastos com EF por aluno _(t-1)	-0,018*** (0,002)	0,007*** (0,001)	0,016*** (0,001)	-0,017*** (0,002)	0,005*** (0,001)	0,013*** (0,001)
Gastos com saúde per capita	0,109*** (0,002)	0,054*** (0,002)	0,035*** (0,002)	0,107*** (0,002)	0,055*** (0,001)	0,038*** (0,002)
Razão professor/aluno	0,082*** (0,003)	0,046*** (0,003)	0,003 (0,004)	0,077*** (0,003)	0,040*** (0,003)	0,006 (0,004)
Educação da mãe	0,024*** (0,001)	0,017*** (0,001)	0,017*** (0,001)	0,020*** (0,001)	0,012*** (0,001)	0,009*** (0,001)
Constante	4,914*** (0,018)	4,986*** (0,014)	-	4,974*** (0,018)	5,056*** (0,013)	-
Observações	15.210	15.210	15.210	15.210	15.210	15.210
R ²	0,244	0,916	0,140	0,232	0,919	0,140

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Erros padrões robustos de White. (2) *p<0.1; **p<0.05; ***p<0.01.

As estimativas para o 9º ano do EF presentes na Tabela 10 enfatizam uma relação ainda menor das variáveis explicativas sobre a variável dependente. Visivelmente, os desfechos dessas estimações sugeriram que os coeficientes dos gastos com educação no período anterior estão correlacionados com aumentos nas notas de português e matemática, sendo mais expressivos no modelo de efeitos fixos. Pode-se perceber que os resultados nas duas disciplinas são parecidos, pois se destaca a maior influência dos gastos com saúde per capita dentre todas as variáveis.

Tabela 11 – Resultados do modelo para a proficiência em língua portuguesa e matemática no 9º ano do ensino fundamental com gastos em educação defasados

Variáveis	Língua portuguesa			Matemática		
	POLS	Efeitos aleatórios	Efeitos fixos	POLS	Efeitos aleatórios	Efeitos fixos
Gastos com EF por aluno _(t-1)	-0,002* (0,001)	0,007*** (0,001)	0,013*** (0,001)	-0,005*** (0,001)	0,003*** (0,001)	0,007*** (0,001)
Gastos com saúde per capita	0,062*** (0,001)	0,042*** (0,001)	0,027*** (0,001)	0,066*** (0,001)	0,042*** (0,001)	0,027*** (0,001)
Razão professor/aluno	0,036*** (0,002)	0,028*** (0,003)	-0,003 (0,004)	0,045*** (0,002)	0,035*** (0,003)	0,004 (0,004)
Educação da mãe	0,013*** (0,001)	0,013*** (0,001)	0,018*** (0,001)	0,010*** (0,001)	0,010*** (0,001)	0,016*** (0,001)
Constante	5,190*** (0,012)	5,221*** (0,011)	-	5,231*** (0,013)	5,299*** (0,011)	-
Observações	14.820	14.820	14.820	14.820	14.820	14.820
R ²	0,194	0,969	0,151	0,198	0,967	0,126

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Erros padrões robustos de White. (2) *p<0.1; **p<0.05; ***p<0.01.

Em resumo, os resultados indicam que os gastos públicos destinados ao ensino fundamental são importantes tanto nas estimações sem defasagem quanto nas com defasagens. No entanto, os resultados revelam uma relação fraca com o desempenho educacional, sugerindo a possibilidade de o volume de recursos investidos pelo país não estarem sendo suficientes para garantir a qualidade do ensino. Além disso, é plausível que ocorram problemas na gestão desses recursos, levando a uma alocação ineficiente deles.

Em linha com esses achados, Kroth e Gonçalves (2019) aborda duas possíveis explicações para os gastos municipais com educação terem apresentado coeficientes baixos, a saber: mesmo com o crescimento do montante de recursos para a educação no decorrer dos últimos anos, estes ainda estão a um nível baixo, aquém de promover ganhos de qualidade esperados. Nesse caso, os recursos financeiros recebidos pela escola podem não estar garantindo o mínimo para sua manutenção, isto é, com condições adequadas de infraestrutura, material didático e salas de aula; e à baixa capacitação dos professores, pois somente com o aumento salarial não promoveria ensino de qualidade.

Para agregar robustez às evidências encontradas nesta pesquisa, foram estimadas regressões com um painel fortemente balanceado contidas no Apêndice A.1 e A.2. Em detrimento de informações ausentes em algumas variáveis, como visto anteriormente, foi perdida uma certa quantidade de observações após o balanceamento dos dados. Nesse contexto, pode-se notar que os coeficientes dos gastos com educação apresentam o mesmo nível de correlação dos que as evidências delineadas na Tabela 7 e 8, ou seja, demonstram a consistência dos resultados perante o número de observações presentes na base de dados.

3.5 Considerações finais

Em linhas gerais, o investimento em educação, quando aplicado corretamente, torna-se fundamental para o desenvolvimento dos países, tanto em questões econômicas quanto em termos sociais e culturais. Nessa perspectiva, o investimento em educação pode ser a solução para vários problemas presentes na sociedade, logicamente que não é somente destinar uma fração de recursos e isso ocasionará o aumento da qualidade educacional. Torna-se essencial que exista uma alocação de recursos da forma mais eficiente possível.

Nesse contexto, o objetivo deste capítulo foi analisar se o gasto por aluno no ensino fundamental, em termos de valores liquidados, exerce influência na qualidade da educação

dos municípios brasileiros entre o período de 2013 e 2019. Utilizando diversas fontes de dados para a construção das variáveis, foram estimados um modelo de painel com efeitos fixos. De fato, todas as especificações do painel realçaram coeficientes positivos e significantes, indicando que aumentos nos gastos elava o desempenho educacional. Esses resultados demonstraram, todavia, uma pequena correlação.

Essa baixa relação indica que, mesmo com a evolução dos gastos destinados ao ensino fundamental nos últimos anos, o montante de recursos por aluno parece ainda ser insuficiente para garantir maiores retornos no que condiz a qualidade de ensino. Outras descobertas reforçam a importância dos gastos com saúde, sugerindo, de certa forma, como um complemento dos gastos com educação para melhorar o desempenho nos testes de proficiência. A educação da mãe, como já é de praxe na literatura econômica, tende a ter um papel importante em influenciar positivamente o rendimento escolar.

Portanto, a principal conclusão é que os gastos por aluno no ensino fundamental são essenciais para aumentar os resultados nas avaliações no Saeb em português e matemática no 5º e 9º ano. Contudo, é válido reconhecer que essa relação ainda é consideravelmente pequena. Nesse caso, resulta importante encontrar meios mais eficientes para fazer com que mais recursos maximize a qualidade de ensino no país. Posto isso, a questão para a melhoria do rendimento escolar parece estar ligada à qualidade de gestão dos recursos, pois os esforços tem se concentrado na quantidade da educação, em vez de considerarem a qualidade educacional.

4 DESIGUALDADES DE DESEMPENHO EDUCACIONAL ENTRE ALUNOS DE ESCOLAS RURAIS E URBANAS NO BRASIL

4.1 Introdução

Em busca de desenvolvimento, independentemente do contexto a ser analisado, o Brasil precisa contar com alguns pilares fundamentais para chegar a tal desfecho, por exemplo: uma indústria nacional robusta, agricultura eficiente, uma infraestrutura sólida e um sistema educacional de qualidade. Dentre diversos pré-requisitos, é possível que o país tenha apresentado menos progresso na área da educação, de forma que, mesmo com o aumento da frequência escolar ao longo do tempo, a qualidade do sistema educativo é algo que ainda não alcançou sucesso, conforme apontado pelos resultados dos indicadores educacionais (Pereira; Castro, 2021).

Sob essa perspectiva, tanto o acesso às instituições de ensino quanto a qualidade do serviço não são uniformes em todo o território brasileiro, seja âmbito regional, estadual ou municipal. Essas disparidades são particularmente marcantes quando se verifica os indicadores educacionais entre áreas rurais e urbanas, pois há diferenças significativas nas taxas de reprovação, abandono escolar e desempenho final dos alunos. Isso ocasiona mudanças na taxa de distorção idade-série¹⁵, na qual, para o exercício de 2021, essa taxa foi de 14,5% ensino fundamental na área urbana e 21,2% na área rural. Essa defasagem é ainda mais abrangente no ensino médio, onde a taxa para área urbana e rural equivaleu a 27,3% e 38,2%, respectivamente (INEP, 2023).

Embora exista uma extensa literatura que trata da importância de diversos fatores para a aprendizagem dos alunos, como características individuais e familiares e recursos escolares existe, no entanto, uma escassez de estudos que analisam a qualidade educacional no contexto geográfico, ou seja, se o aluno reside em áreas rurais ou urbanas. Ignorar essas questões pode trazer distorções nas investigações, uma vez que, os resultados de alunos localizados em áreas urbanas são, notavelmente, mais superiores aos da área rural, o que pode ser explicado por diversos fatores.

¹⁵ É um indicador utilizado para mensurar a defasagem entre a idade do aluno e a série em que ele está matriculado. Isso pode ocorrer quando um estudante está em uma série que não corresponde à sua idade adequada, em que tal distorção pode ser indício de reprovações e abandonos.

Segundo Lounkaew (2013), esse *gap* existe porque as escolas urbanas usufruem de mais recursos do que as escolas rurais, o que faz com que os alunos possam desfrutar de mais benefícios dessas dotações. No Brasil, ainda que com a implementação de diversas políticas públicas¹⁶ nos últimos anos que atingiram o meio rural, as condições das escolas rurais são precárias quando comparadas com as escolas urbanas. As escolas rurais ainda possuem recursos escassos como bibliotecas, internet, computadores, entre outros (Pereira; Castro, 2021).

Partindo desse contexto, o objetivo deste capítulo consiste em analisar a dimensão da lacuna de desempenho educacional entre alunos que da área rural e urbana, destacando o quanto esse *gap* continua significativo quando está sob o controle empírico de outras variáveis. Em outras palavras, busca-se responder as seguintes questões: as diferenças de desempenho entre estudantes que frequentam escolas rurais e urbanas continuam grandes mesmo após o controle das covariadas relacionadas as características individuais e socioeconômicas do aluno e da família? Essas diferenças tendem a aumentar com o passar dos anos, especialmente em um período de pós-pandemia?

A motivação dessa temática é a escassez na literatura em estudar o componente geográfico como determinante no aprendizado escolar, especialmente com uma análise a nível aluno em um cenário de pós-pandemia. Para isso, como estratégia de investigação, *a priori* será utilizado os dados do Saeb para as avaliações de língua portuguesa e matemática para o exercício de 2017, 2019 e 2021, restringindo as investigações para alunos do ensino médio da rede pública de ensino. Metodologicamente, será aplicado mínimos quadrados utilizando efeitos fixos tanto de estados como de municípios para controlar possíveis características não observáveis das escolas. Os resultados encontrados serão equiparados à luz de pesquisas que propuseram essa temática.

Além desta seção introdutória, o artigo dispõe da seção 4.2 com a revisão de literatura que circunda o propósito da pesquisa. Na seção 4.3 é realizado os procedimentos metodológicos, ao qual se insere a descrição das variáveis e o método de investigação. Posteriormente, a seção 4.4 esboça os principais resultados encontrados e, por fim, a seção seguinte realiza as conclusões da pesquisa.

¹⁶ Programa Nacional de Alimentação Escolar (Pnae), Programa Nacional de Transporte Escolar (PNTE), Programa Escola Ativa, Programa de Formação de Professores em Exercício (Proformação), o Programa Nacional de Educação na Reforma Agrária (Pronera), dentre outros.

4.2 Diferenças de desempenho educacional no meio urbano-rural

O arcabouço dos estudos empíricos que trata sobre os determinantes do aprendizado escolar é excepcionalmente amplo. Contudo, o desempenho da educação dando ênfase à condição residencial dos alunos é modestamente investigado. Ou seja, pouco se sabe se eles são residentes rurais ou urbanos. O debate centra-se na noção de que a localização em si não influencia as diferenças de resultados escolares, pois são outros fatores os mais influentes, como os atributos dos alunos, das famílias, das escolas, etc. Nesse sentido, serão apresentadas algumas pesquisas internacionais e nacionais, que se esforçaram para aprofundar algumas percepções relacionadas a essa questão, trazendo conclusões do desempenho escolar inferior dos estudantes de escolas rurais perante àqueles de área urbana.

Nesse contexto, iniciando um panorama que aborda as diferenças de desempenho entre alunos que frequentam escolas urbanas e rurais, Ramos, Duque e Nieto (2012) examinaram essas disparidades na Colômbia, utilizando os microdados do Pisa, entre os anos de 2006 e 2009, com os resultados das avaliações em matemática, ciências e leitura. Por meio de uma função de produção educacional, foi utilizada uma ampliação da metodologia desenvolvida em Juhn-Murphy-Pierce (1993) para verificar os efeitos da variação no tempo. Os resultados mostraram que a maior parte do diferencial de desempenho educacional entre as escolas rurais e urbanas está associada às características familiares. Nesse sentido, há uma relação modesta diante dos insumos da escola.

Com base nos resultados das avaliações do Pisa na Tailândia, Lounkaew (2013) procurou explicar as diferenças do sucesso educacional entre alunos das áreas rural e urbana no ano de 2009. Utilizando uma decomposição pelo método de Oaxaca-Blinder (1973), o autor estimou uma função de produção da educação ao nível do aluno em pontos diferentes ao longo das distribuições de desempenho. O autor identificou que o efeito das características do aluno, da família e da escola variam conforme a distribuição, e os exercícios de decomposição revelaram que em torno de 45% e 48% das diferenças de desempenho urbano-rural são explicadas pelas características não mensuráveis da escola.

Na mesma linha, Amini e Nivorozhkin (2015) analisam os determinantes do desempenho escolar rural-urbano para os alunos da Rússia. Com isso, estimaram uma função de produção de educação empregando técnica de variável instrumental com o intento de controlar a endogeneidade dos insumos escolares por intermédio de um modelo multinível de dados hierárquicos. Os resultados para o método de decomposição apontaram que as

disparidades das escolas rurais e urbanas no desempenho escolar está relacionada às particularidades socioeconômicas da família que influencia incentivos educacionais.

Em um estudo na China, Zhang, Li e Xue (2015) buscaram comparar o desempenho educacional de crianças rurais, crianças migrantes de áreas rurais para urbanas e crianças urbanas entre o período de 2009 e 2010. Utilizando dados do Painel Familiar da China (*China Family Panel Survey – CFPS*) e da pesquisa *Rural–Urban Migration in China* (RUMiC), os resultados enfatizaram que o aprendizado educacional das crianças rurais e dos migrantes rurais-urbanos é, significativamente, menor do que aquelas situadas no meio urbano. Mesmo após controlar as diferenças das características individuais, a lacuna educacional inexplicada ainda é substancial.

Em outro estudo na Colômbia, Gonzalez, Gómez e Gómez (2021) verificaram os principais determinantes do desempenho dos estudantes na avaliação do Pisa. Para tanto, primeiramente, estimaram uma função de produção educacional em nível aluno em diversos pontos da distribuição, depois aplicaram a decomposição por Oaxaca-Blinder para compreender o quanto as diferenças de desempenho podem ser explicadas por características dos alunos, da família e da escola. Com efeito, os fatores observados são responsáveis por quase todo o *gap* de desempenho, pois as características da escola explicaram 40%; o *background* familiar é crucial, explicando 32% em leitura e 20% em matemática e ciências; e atributos individuais dos alunos explicam 10% da lacuna de desempenho, todavia, não foi significativo em leitura.

Com foco nos países de baixo e médio rendimento, Betancur, Miller, Votruba-Drzal (2024) investigam as diferenças urbano-rurais das crianças desfavorecidas da Índia, Peru e Vietnã. De início, o estudo centra-se nas primeiras competências escolares das crianças, controlado por aspectos infantis e familiares e, posteriormente, considera-se se tais disparidades explicam a participação das crianças na educação de primeira infância pública e privada, dada a disponibilidade limitada nas zonas rurais. Os desfechos da pesquisa indicaram que as características da família explicam metade do *gap* no desempenho no Peru e no Vietnã, e explicam completamente essas lacunas na Índia. Resultados adicionais sugerem que residir em áreas urbanas está positivamente relacionada ao desempenho escolar por meio do aumento da frequência nas escolas de primeira infância.

No Brasil, Bezerra e Kassouf (2006) avaliaram os fatores que influenciam o desempenho educacional, equiparando estudantes de escolas localizadas em áreas rurais e urbanas. A partir do método de mínimos quadrados, os autores utilizaram os microdados do

Saeb para os alunos da 4ª série/5º ano do ensino fundamental no ano de 2003. Em síntese, concluíram que as variáveis que contribuem para o sucesso educacional do aluno da área urbana é a educação da mãe, a renda familiar, a rede de ensino, o atraso escolar e o trabalho infantil. Já no meio rural, alguns fatores que contribuem são a renda familiar, a região onde a escola se localiza, recursos educacionais (computadores, biblioteca, etc), o número de pessoas no domicílio e atraso escolar.

Rodrigues *et al.* (2020) procuraram examinar os fatores que contribuem para as diferenças de desempenho dos alunos que frequentam escolas na localidade rural e urbana no Brasil. Os autores estimaram uma Regressão de Função de Influência Recentrada (RIF) proposto por Firpo, Fortin e Lemieux (2007) e na metodologia de decomposição de Oaxaca-Blinder aplicada aos quantis, utilizando dados do Saeb para o 5º ano do ensino fundamental nos testes de português e matemática. Os resultados realçaram que as pontuações em português e matemática são mais elevadas nas escolas urbanas do que nas rurais e, além disso, as características das escolas são os principais impulsionadores da lacuna de desempenho, especialmente nos quantis superiores, sobretudo, nas notas de matemática.

Do mesmo modo, Alencar *et al.* (2021) verificaram se as diferentes características dos professores explicam o *gap* no desempenho educacional dos estudantes do 9º do ensino fundamental no ano de 2017. Por meio da função de produção educacional, os autores adotaram a estratégia proposta por Firpo, Fortin e Lemieux (2018) que permite a decomposição detalhada de Oaxaca-Blinder (1973) nos quantis de distribuição das pontuações. Os resultados indicaram que os professores com licenciatura na área contribuem de forma positiva e significativa para a proficiência em português e matemática, com efeito maior para alunos da área rural. Nada obstante, essa influência descrese na proporção que aumenta o quantil em ambos os grupos.

Ainda no Brasil, Cruz, Moura e Esperidião (2022) procuraram avaliar se o fator de localidade na área rural e urbano é significativo para explicar o desempenho inferior neste setor nos municípios entre o intervalo de 2013 e 2017. Com base nos microdados do Saeb para o 5º ano do ensino fundamental, os autores usaram o método proposto por Firpo, Fortin e Lemieux (2007) estimando Regressões Quantílicas Incondicionais e o Tratamento de Efeito de Quantil para captar o impacto distributivo da variável categórica de localização da escola mantendo constante a distribuição de outros fatores. As evidências apontaram, de maneira robusta, que existe uma penalidade de desempenho aos estudantes que frequentam

escolas rurais, em que mesmo após controlar os potenciais impactos das variáveis preconizadas na literatura, essa diferença não é mitigada.

4.3 Metodologia

Esta seção destina-se a apresentar a base de dados empregada na pesquisa juntamente com a descrição das variáveis. Ademais, descreverá também a estratégia empírica com a aplicabilidade do modelo de MQO com efeitos fixos.

4.3.1 Base de dados e descrição das variáveis

A base de dados é advinda do Sistema de Avaliação da Educação Básica (Saeb), conduzido pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). O Saeb é um conjunto de avaliações externas em larga escala que tem como objetivo realizar um diagnóstico da educação básica brasileira e também fatores que podem interferir no desempenho dos alunos. Com isso, o presente estudo vai recorrer aos microdados dos anos de 2017, 2019 e 2021 para o 3º ano do ensino médio nos testes de língua portuguesa e matemática das escolas da rede pública.

Ademais, o Saeb é uma avaliação censitária bianual que envolve os alunos do 5º e 9º ano do ensino fundamental regular das escolas públicas que possuem, no mínimo, 20 alunos matriculados nos anos avaliados. Ademais, é realizado também uma avaliação para o 3º ano do ensino médio, mas somente começou a ser censitária a partir de 2017, o que justifica o início do período a ser analisado. Vale ressaltar também a importância do estudo para o ano de 2021, pois é um período no contexto pós-pandemia ao qual exigiu uma adaptação dos sistemas educacionais para oferecer opções de aprendizado.

Os testes têm a finalidade de fazer uma avaliação do desempenho escolar nas áreas do conhecimento em língua portuguesa (com foco em leitura) e matemática (com ênfase na resolução de problemas). Além desses instrumentos, são aplicados quatro tipos de questionários: o de alunos, que respondem perguntas sobre o ambiente e nível socioeconômico familiar, hábitos de estudo e de leitura, dentre outros aspectos; o dos professores e o do diretor, que fornecem informações sobre sua formação profissional, nível socioeconômico e cultural, formas de gestão, práticas pedagógicas, entre outros. Enfim, informações sobre a escola, em que possuem perguntas relacionadas ao ambiente físico,

infraestrutura, recursos, etc.

Dada essas especificações, o Quadro 4 realça a descrição das variáveis a serem utilizadas nas estimações. As variáveis dependentes correspondem as notas padronizadas em língua portuguesa e matemática, ou seja, com média 0 e desvio padrão 1 para que se possa equiparar os resultados com estudos internacionais. Referente às variáveis explicativas¹⁷, o foco de estudo é a *dummy* de localização para observar as diferenças de desempenho entre alunos rurais e urbanos. Além disso, foram consideradas características individuais do aluno como a cor/raça, utilizando a categoria “brancos” como base; variáveis familiares, como a quantidade de membros que residem no domicílio e a educação da mãe.

O Nível Socioeconômico do Aluno (NSE) é um índice construído utilizando *Principal Components Analysis* (PCA)¹⁸, tendo como base o estudo de Carnoy, Rosa e Simões (2022). O PCA é uma técnica estatística que reduz a dimensionalidade dos dados, ou seja, permite diminuir a quantidade de variáveis em um conjunto de dados mantendo-se o máximo de informações possíveis. Assim, a finalidade é transformar variáveis correlacionáveis em grupo de variáveis que não estão correlacionadas (denominadas de componentes principais) por meio da combinação linear das variáveis originais.

Quadro 3 – Descrição das variáveis

Variáveis	Descrição
Variáveis dependentes	
Nota de língua portuguesa	Proficiência em língua portuguesa no 3º do ensino médio
Nota de matemática	Proficiência em matemática no 3º do ensino médio
Variáveis explicativas	
Rural	<i>Dummy</i> de localização: 1 rural, 0 urbano
Branco	<i>Dummy</i> de cor/raça: 1 branco, 0 caso contrário
Pardos	<i>Dummy</i> de cor/raça: 1 pardos, 0 caso contrário
Pretos	<i>Dummy</i> de cor/raça: 1 pretos, 0 caso contrário
Indígenas	<i>Dummy</i> de cor/raça: 1 indígenas, 0 caso contrário
Amarelos	<i>Dummy</i> de cor/raça: 1 amarelos, 0 caso contrário
Quant. de pessoas no	Quantidade de pessoas residentes no domicílio dos alunos
NSE	Nível Socioeconômico do Aluno
Educação da mãe	<i>Dummy</i> para educação da mãe: 1 ensino superior/ensino médio, 0 caso contrário

Fonte: Elaboração própria com base nos dados da pesquisa.

¹⁷ Não foram adicionadas variáveis relacionadas a escola, pois seria necessário trabalhar com dados do Censo Escolar. Desse modo, o Saeb possui uma estrutura diferente, ou seja, parte das escolas são mascaradas, então optou-se por utilizar o Saeb como um todo, sem utilizar variáveis que estão no Censo Escolar.

¹⁸ Tradução: Análise de Componentes Principais.

Partindo disso, a PCA procura verificar os eixos onde há mais variação dos dados, na qual o primeiro componente principal captura uma variância maior. Nesse caso, um dos maiores desafios é encontrar o melhor grupo de variáveis que estejam fortemente correlacionados com o NSE. Para Carnoy, Rosa e Simões (2022) esse grupo tem que ser suficientemente representativo a fim de influenciar nos resultados educacionais, e devidamente pequeno para facultar uma redução adequada da dimensionalidade, preservando, por sua vez, a informação. Dito isto, foi construído um índice baseado em constituintes do que o aluno relatou que existe em seu domicílio, como: número de carros, banheiros, quartos, computadores, televisões e máquinas de lavar.

4.3.2 Mínimos Quadrados Ordinários (MQO) com efeitos fixos

a literatura demonstra que o desempenho dos alunos é, especialmente, explicado por três fatores básicos, a saber, características individuais, da família e da escola¹⁹, conforme exposto anteriormente (Hanushek, 1997; Nieto; Ramos, 2014; Woessmann, 2016; Gonzalez; Gómez; Gómez, 2021). Para desenvolver as relações de interesse, primeiramente, serão estimadas as regressões múltiplas por MQO para os anos de 2017, 2019 e 2021, individualmente, para obter um panorama geral do efeito das variáveis explicativas sobre a variável resposta. Além disso, serão incluídos efeitos fixos²⁰ de estados e municípios em modelos separados para captar efeitos não observados e constantes no tempo.

Diante disso, a ênfase das estimativas concentra-se no *gap* de desempenho rural-urbano, dado o controle das demais covariadas. Assim sendo, o modelo linear completo pode ser explícito da seguinte forma:

$$Y_{p,m} = \beta_0 + \beta_1 rural + \beta_2 pardos + \beta_3 pretos + \beta_4 indígenas + \beta_5 amarelos + \beta_6 NSE + \beta_7 qt_pmd + \beta_8 educ_mae + c + \varepsilon \quad (8)$$

em que, $Y_{p,m}$ são as notas padronizadas em língua portuguesa e matemática; *rural* é a *dummy* de localização; *pardos*, *pretos*, *indígenas*, *amarelos*, equivalem as *dummies* de cor/raça; *NSE* é o Nível socioeconômico do Aluno; *qt_pmd* é o número de pessoas que moram no domicílio; *educ_mae* é uma *dummy* para mães com ensino médio e superior;

¹⁹ Infelizmente as análises desse estudo não consideram as características da escola (ver a seção 4.3.1).

²⁰ Uma provável explicação para não adicionar efeitos fixos de escola é por conta de um provável conflito entre este e a *dummy* rural e urbano, ou seja, tais efeitos estão sendo dropados.

c é os efeitos fixos para município ou unidades da federação (UF); ε é o termo de erro; e os β 's equivalem aos parâmetros ligados as variáveis explicativas.

4.4 Resultados e discussão

Esta seção reporta os principais resultados encontrados por meio da aplicabilidade do método de análise. Discutir-se-á, inicialmente, algumas evidências iniciais a partir da estatística descritiva e, em seguida, a análise econométrica.

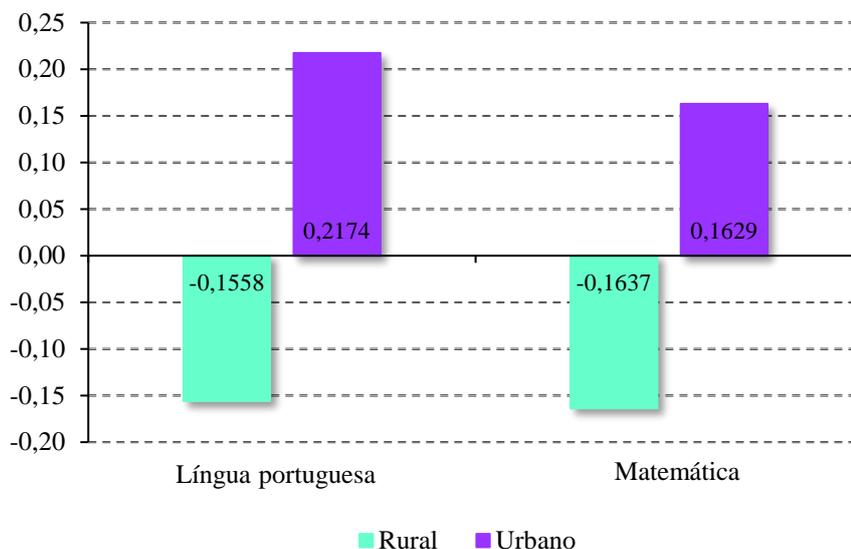
4.4.1 Estatística descritiva

No contexto das análises, as evidências iniciais²¹ estão presentes no Gráfico 5, com as médias referentes ao desempenho nas disciplinas de português e matemática para todo o período de investigação. Como as notas estão padronizadas, a interpretação se dá por um ponto de referência, com média 0 e desvio padrão 1. Nesse caso, valores positivos indicam que os resultados estão acima da média e valores negativos sugerem que a nota está abaixo da média. Dessa forma, quanto maior o valor absoluto, maior a distância em desvios padrão da média.

Com essas especificações, é nítido que alunos residentes na área rural possuem rendimento educacional inferior aos alunos que moram na área urbana, pois os resultados estão abaixo da média, ou seja, são negativos. Com isso, as médias dos estudantes urbanos correspondem a aproximadamente 0,217 DP em português e 0,163 DP em matemática, já os estudantes rurais possuem médias equivalentes a -0,156 DP e -0,164 DP em português e matemática, respectivamente.

²¹ As estatísticas por ano estão encontradas no Apêndice A, onde mostra que as variáveis foram padronizadas.

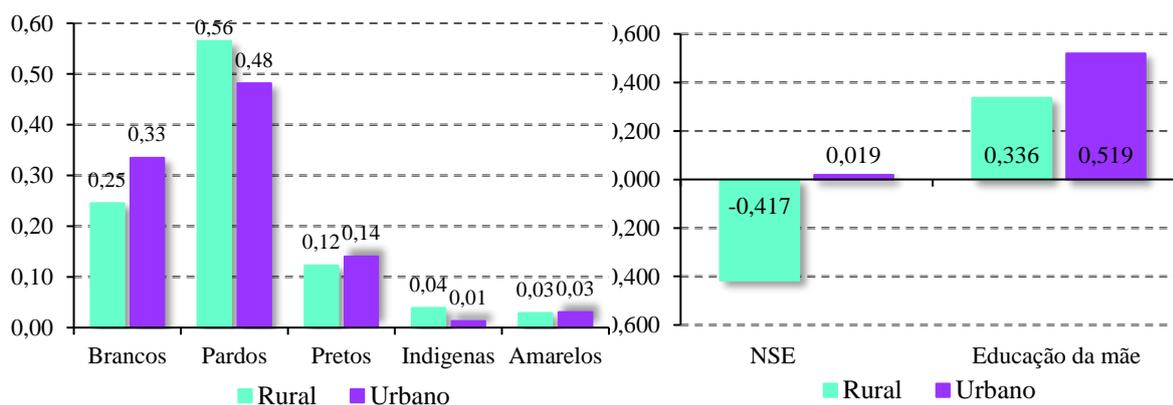
Gráfico 5 – Média total das notas padronizadas em língua portuguesa e matemática por localidade



Fonte: Elaboração própria baseado nos dados da pesquisa

O Gráfico 6 apresenta as estatísticas totais dos atributos individuais e socioeconômicos dos estudantes rurais e urbanos. Observa-se, em média, que há mais brancos e pretos na área urbana e existe mais pardos e indígenas na área rural. Já em relação ao nível socioeconômico, os alunos rurais estão abaixo da média com -0,417 DP, e quanto a educação da mãe, o quantitativo que possuem ensino médio e superior é maior na área urbana.

Gráfico 6 – Média total das características individuais e socioeconômicas dos alunos por localidade



Fonte: Elaboração própria baseado nos dados da pesquisa.

Sumariamente, a próxima subseção destaca os primeiros resultados da estimação do modelo econométrico. Embora a estatística descritiva seja importante para verificar o

comportamento dos dados, não fornece evidências sólidas sobre a desigualdade educacional entre alunos rurais e urbanos. Nesse caso, serão apresentados, em seguida, as regressões filtradas nos três períodos de tempo e, a partir disso, tirar conclusões da dimensão da lacuna de desempenho ao longo do tempo para os testes de língua portuguesa e matemática.

4.4.2 Análise Econométrica

De início, a Tabela 10 delinea os resultados das estimações de quatro cenários para o ano de 2017. A finalidade da inserção gradativa das covariadas e dos efeitos fixos é para analisar o comportamento da variável de interesse para com os modelos, em que se pode mencionar, de imediato, a significância estatística de todos os coeficientes. No modelo I, tem-se o efeito total da variável rural sobre as duas disciplinas, em que se nota que os alunos da área rural possuem menor desempenho comparado aqueles da área urbana, apresentando uma lacuna de 0,342 DP em português e 0,291 DP em matemática.

Em seguida, no modelo II é apresentada a relação do *gap* controlado pelas características observáveis dos alunos para mitigar o viés, inserindo as *dummies* de cor/raça, a quantidade de pessoas residentes no domicílio, o nível socioeconômico do aluno e a educação da mãe. Após esse controle, as diferenças entre os grupos são equivalentes a 0,244 DP em português e 0,182 DP em matemática. Já o modelo III, há um controle dos efeitos fixos dos estados com o intuito de capturar características não observáveis específicas de cada unidade da federação. Desse modo, este equivaleu a uma disparidade entre alunos da área rural e urbana em torno de 0,183 DP e 0,119 DP nas respectivas disciplinas.

O melhor panorama que atenua as diferenças nos resultados educacionais dos alunos rurais e urbanos é o modelo IV, pois neste o efeito de localidade é controlado pelas características dos alunos e efeitos fixos dos municípios. Em tal situação, os resultados demonstraram que o desempenho dos estudantes rurais é inferior aos estudantes urbanos em 0,165 DP e 0,117 DP em português e matemática, respectivamente. Diante de todos esses achados, é importante destacar que as desigualdades entre alunos que frequentam as escolas rurais e urbanas, dado os quatro modelos, foram mais tênues na proficiência em matemática.

Tabela 12 – Resultados das estimações em língua portuguesa e matemática em 2017

Variáveis	Língua Portuguesa				Matemática			
	I	II	III	IV	I	II	III	IV
Rural	-0,342*** (0,005)	-0,244*** (0,005)	-0,183*** (0,047)	-0,165*** (0,016)	-0,291*** (0,005)	-0,182*** (0,005)	-0,119** (0,045)	-0,117*** (0,015)
Pardos		-0,212*** (0,002)	-0,166*** (0,022)	-0,140*** (0,005)		-0,191*** (0,002)	-0,151*** (0,020)	-0,114*** (0,004)
Pretos		-0,280*** (0,003)	-0,241*** (0,028)	-0,208*** (0,009)		-0,284*** (0,003)	-0,252*** (0,030)	-0,199*** (0,007)
Indígenas		-0,418*** (0,008)	-0,370*** (0,026)	-0,310*** (0,012)		-0,375*** (0,008)	-0,330*** (0,027)	-0,264*** (0,011)
Amarelos		-0,194*** (0,006)	-0,161*** (0,025)	-0,138*** (0,007)		-0,182*** (0,006)	-0,162*** (0,022)	-0,124*** (0,009)
Qt_pmc		-0,116*** (0,002)	-0,086*** (0,008)	-0,071*** (0,003)		-0,110*** (0,002)	-0,076*** (0,007)	-0,060*** (0,003)
NSE		0,077*** (0,001)	0,042*** (0,003)	0,025*** (0,002)		0,118*** (0,001)	0,088*** (0,004)	0,070*** (0,002)
Educ_mae		0,226*** (0,003)	0,237*** (0,018)	0,229*** (0,006)		0,211*** (0,002)	0,228*** (0,014)	0,230*** (0,005)
Constante	0,014*** (0,001)	0,140*** (0,002)	-	-	0,012*** (0,001)	0,124*** (0,002)	-	-
Covariadas		✓	✓	✓		✓	✓	✓
EF Mun				✓				✓
EF UF			✓				✓	
R ² adj	0,005	0,05	0,075	0,11	0,003	0,055	0,084	0,13
Observações	1.136.021	1.011.601	1.011.601	1.011.601	1.136.021	1.011.601	1.011.601	1.011.601

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Todas as estimações tem erros-padrão clusteurizados a nível de município. (2) *p<0,1; **p<0,05; ***p<0,01.

A título de investigação, o *gap* correspondente ao aspecto racial mostrou-se negativo em todos os grupos, enfatizando que os indivíduos que se declaram brancos possuem desempenho superior em comparação aos outros. Com isso, o maior diferencial consiste na categoria de alunos indígenas com 0,31 DP e 0,264 DP de desempenho inferior aos alunos brancos nas disciplinas dotadas para o último modelo; há também uma diferença significativa para estudantes pretos de 0,208 DP e 0,199 DP em português e matemática, respectivamente.

Em outro contexto, os *status* socioeconômico do aluno também desempenha um papel importante para explicar as pontuações nos testes de proficiência, impactando em 0,077 DP em português e 0,118 DP em matemática no modelo II, ao inserir os efeitos fixos de município, o efeito decaiu para 0,025 DP e 0,07 DP. Como já era esperado, mães que têm ensino médio e superior gera efeito positivo nos resultados educacionais em 0,23 DP. Diante disso, Ramos, Duque e Nieto (2012) quando analisaram as diferenças entre os estudantes

rurais e urbanos, encontraram nas suas estimativas que a escolaridade da mãe e o índice socioeconômico impactam positiva e significativamente o desempenho dos alunos.

É interesse enfatizar que a quantidade de pessoas no domicílio afeta de forma negativa, mas significativa, os resultados educacionais. Todavia, essa evidência suaviza quando se controla os efeitos fixos dos municípios, constatando uma influência de -0,071 DP em português e -0,06 DP em matemática. Os achados de Bezerra e Kassouf (2006) relataram que uma maior quantidade de pessoas residentes no domicílio prejudica o desempenho escolar, sendo esse maior efeito nas notas de português quando comparado as de matemática.

As estimações para o ano de 2019, contidos na Tabela 11, mostram resultados maiores para a variável de interesse quando comparado ao período anterior. As diferenças entre os grupos após o controle das covariadas e dos efeitos fixos dos municípios estão mais fortes, em que alunos da área rural tem desempenho menor do que alunos da área urbana em 0,184 DP e 0,141 DP em português e matemática, respectivamente. Em uma visão diferente, mesmo que haja um acréscimo do *gap* para este ano, percebe-se, entretanto, que o declínio na lacuna de desempenho entre o modelo I e IV em 2019 é maior. Ou seja, o controle empírico consegue atenuar mais essas desigualdades.

Tabela 13 – Resultados das estimações em língua portuguesa e matemática em 2019
(Continua)

Variáveis	Língua Portuguesa				Matemática			
	I	II	III	IV	I	II	III	IV
Rural	-0,421*** (0,005)	-0,322*** (0,006)	-0,234*** (0,047)	-0,184*** (0,017)	-0,392*** (0,005)	-0,272*** (0,006)	-0,176*** (0,056)	-0,141*** (0,017)
Pardos		-0,242*** (0,003)	-0,194*** (0,018)	-0,171*** 0,005		-0,234*** (0,003)	-0,182*** (0,018)	-0,148*** (0,004)
Pretos		-0,330*** (0,004)	-0,292*** (0,024)	-0,265*** 0,007		-0,351*** (0,004)	-0,311*** (0,028)	-0,261*** (0,007)
Indígenas		-0,509*** (0,010)	-0,449*** (0,028)	-0,391*** (0,013)		-0,495*** (0,010)	-0,428*** (0,030)	-0,358*** (0,013)
Amarelos		-0,243*** (0,007)	-0,211*** (0,020)	-0,191*** (0,009)		-0,241*** (0,007)	-0,210*** (0,022)	-0,176*** (0,009)
Qt_pmc		-0,167*** (0,004)	-0,137*** (0,014)	-0,124*** (0,005)		-0,174*** (0,004)	-0,136*** (0,014)	-0,118*** (0,005)
NSE		0,110*** (0,001)	0,076*** (0,004)	0,056*** (0,002)		0,157*** (0,001)	0,124*** (0,004)	0,104*** (0,003)
Educ_mae		0,185*** (0,002)	0,198*** (0,015)	0,180*** (0,004)		0,197*** (0,002)	0,217*** (0,012)	0,212*** (0,004)
Constante	0,316*** (0,001)	0,468*** (0,002)	-	-	0,268*** (0,001)	0,430*** (0,002)	-	-
Covariadas		✓	✓	✓		✓	✓	✓
EF Mun				✓				✓

(Conclusão)

Variáveis	Língua Portuguesa				Matemática			
	I	II	III	IV	I	II	III	IV
EF UF			✓				✓	
R ² adj	0,007	0,061	0,084	0,121	0,006	0,075	0,105	0,152
Observaçõ	1.030.007	702.017	702.017	702.017	1.030.007	702.017	702.017	702.017

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Todas as estimações tem erros-padrão clusteurizados a nível de município. (2) * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$.

O ano de 2021 foi um período delicado tanto no cenário brasileiro quanto mundialmente, com adaptações do sistema de ensino causado pela pandemia da Covid-19. É compreensível que as disparidades de desempenho entre alunos localizados em rurais e urbanos estejam mais acentuadas, visto as dificuldades encontradas para prosseguir estudando. Com as aulas suspensas durante essa crise sanitária, diversos estudantes tiveram dificuldades no aprendizado remoto, principalmente na área rural, devido à péssima qualidade da internet e escassez de recursos para garantir o acesso as aulas em casa.

Com essas especificações, introduzindo os resultados para esse período através da Tabela 12, percebe-se que as desigualdades de desempenho estudantil aumentaram tanto para o contexto de localidade, objeto de estudo, como para o aspecto racial. Em termos de localidade, o *gap* total equivaleu a 0,422 DP em português e 0,145 DP em matemática no modelo I, já no último exercício, depois do controle das covariadas e efeito fixos, essa lacuna de desempenho entre estudantes rurais e urbanos correspondeu a 0,195 DP em português e 0,145 DP em matemática.

Comparativamente, a diferença bruta de desempenho das crianças rurais e urbanas do estudo de Zhang, Li e Xue (2015) corresponde, em média, a 0,54 DP para os testes de matemática, ou seja, as disparidades dessa pesquisa são menores quando equipara as estimativas. Embora esses autores concentrem-se na análise de crianças e essa pesquisa seja a nível de ensino médio, eles explanaram que essa diferença não muda significativamente para aqueles estudantes matriculados em escola primária e ensino fundamental. Sendo assim, isso implica que a desigualdade educacional entre áreas rurais e urbanas pode não diminuir na proporção em que essas crianças crescem e obtêm mais educação formal.

Já com a inserção de variáveis como características individuais, familiares e qualidade da escola, Zhang, Li e Xue (2015) demonstraram que as desigualdades nos resultados escolares ainda persistem, com -0,196 DP em matemática. No que concerne à Ramos, Duque e Nieto (2012), grande parte da lacuna rural-urbana está correlata com as

características familiares. Dessa forma, é imprescindível a adequação e a formulação política educacional direcionadas em aprimorar as condições da família.

De um ângulo alternativo, ainda que a pesquisa de Betancur, Miller, Votruba-Drzal (2024) seja para um público diferente, os resultados são bastante elevados quando equiparado aos achados desse estudo para o ensino médio. No Peru, as diferenças do resultado educacional para os testes de matemática equivaleram a 0,66 DP e 0,79 DP para crianças de cinco e oito anos, respectivamente. Moderadamente, no Vietnã *gap* entre crianças rurais e urbanas são 0,61 DP aos cinco anos e 0,65 aos oito anos, já na Índia essas desigualdades são menores, apresentando valores de 0,36 DP e 0,19 DP para as crianças de cinco e oito anos, respectivamente.

Tabela 14 – Resultados das estimações em língua portuguesa e matemática em 2021

Variáveis	Língua Portuguesa				Matemática			
	I	II	III	IV	I	II	III	IV
Rural	-0,422*** (0,005)	-0,312*** (0,006)	-0,228*** (0,042)	-0,195*** (0,016)	-0,365*** (0,005)	-0,248*** (0,006)	-0,156*** (0,046)	-0,145*** (0,016)
Pardos		-0,246*** (0,003)	-0,194*** (0,008)	-0,179*** (0,004)		-0,225*** (0,003)	-0,169*** (0,012)	-0,145*** (0,004)
Pretos		-0,338*** (0,004)	-0,284*** (0,015)	-0,271*** (0,006)		-0,361*** (0,004)	-0,301*** (0,022)	-0,267*** (0,009)
Indígenas		-0,546*** (0,011)	-0,477*** (0,030)	-0,433*** (0,013)		-0,479*** (0,011)	-0,404*** (0,032)	-0,354*** (0,015)
Amarelos		-0,293*** (0,008)	-0,247*** (0,023)	-0,233*** (0,010)		-0,263*** (0,008)	-0,219*** (0,026)	-0,196*** (0,011)
Qt_pmc		-0,191*** (0,004)	-0,166*** (0,008)	-0,159*** (0,005)		-0,192*** (0,004)	-0,161*** (0,008)	-0,150*** (0,005)
NSE		0,107*** (0,001)	0,075*** (0,004)	0,062*** (0,002)		0,151*** (0,001)	0,119*** (0,003)	0,106*** (0,003)
Educ_mae		0,176*** (0,002)	0,185*** (0,014)	0,170*** (0,004)		0,185*** (0,002)	0,199*** (0,009)	0,199*** (0,005)
Constante	0,281*** (0,001)	0,441*** (0,003)	-	-	0,154*** (0,001)	0,310*** (0,003)	-	-
Covariadas		✓	✓	✓		✓	✓	✓
EF Mun				✓				✓
EF UF			✓				✓	
R ² adj	0,008	0,064	0,086	0,113	0,006	0,076	0,103	0,135
Observações	910.346	640.460	640.460	640.460	910.346	640.460	640.460	640.460

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Todas as estimações tem erros-padrão clusteurizados a nível de município. (2) * $p < 0,1$; ** $p < 0,05$; *** $p < 0,01$.

Concisamente, é perceptível que a inserção gradativa das covariadas e o controle dos efeitos fixos ajudam a mitigar as desigualdades de desempenho dos alunos rurais perante os alunos urbanos, mas não anulam completamente essa desigualdade. Esperava-se, no entanto,

que essas diferenças fossem praticamente eliminadas com o controle potencial das variáveis, o que não ocorreu em nenhuma das estimativas. Além disso, essa lacuna continua aumentando aos poucos durante o hiato de estudo, o que é compreensível dado o cenário pandêmico ocorrido dentro desse período.

De maneira geral, não há como afirmar precisamente fatores que expliquem esse aumento no diferencial, mas um ponto importante a ser destacado são as condições socioeconômicas do aluno que, como já foi observado nos resultados, é um aspecto relevante no processo de aprendizagem. Posto isso, vale ressaltar que muitos estudantes residentes rurais são desprovidos de dispositivos eletrônicos, como computadores ou tablets, e tampouco um acesso à internet de alta velocidade para acompanhar as aulas. Durante a crise causada pela Covid-19, não possuir esses equipamentos se tornou um grande empecilho para o acesso as aulas.

De outra forma, em algumas áreas rurais, principalmente naquelas economicamente desfavorecidas, existe alunos que frequentam a escola não por motivos educacionais, mas exclusivamente por causa das condições socioeconômicas precárias. Diante disso, a escola serve como um local onde os estudantes tem acesso a refeições regulares, tornando-se crucial para a nutrição e alimentação. Com a crise sanitária, ocorreu o fechamento das escolas, esses alunos sofreram perda significativa neste ponto, afetando a nutrição e o bem-estar. Como já foi supramencionado, a literatura destaca que tais fatores impactam o desempenho cognitivo do aluno (Hanushek; Woessmann, 2010).

Em resumo, os resultados encontrados nas estimações mostram que, em média, as desigualdades de rendimento educacional nos exames de proficiência entre estudantes residentes rurais e urbanos, são notavelmente elevados. Embora o público-alvo utilizado neste estudo seja um pouco destoante das pesquisas internacionais, em termos de desvios padrões, esses valores realçados nos três anos são considerados altos quando são comparados com esses estudos.

4.5 Considerações finais

Este estudo buscou contribuir para a discussão na literatura sobre a qualidade da educação básica ao examinar as desigualdades no desempenho escolar entre alunos que frequentam escolas localizadas na zona rural e urbana, especificamente no ensino médio. Para realizar esse objetivo, foram utilizados os dados do Saeb advindo da Prova Brasil nos

anos de 2017, 2019 e 2021, pois foi a partir desse período que tal avaliação se tornou censitária para o ensino médio. Dessa forma, como estratégia empírica, foi empregado estimativas de MQO com efeitos fixos para a obtenção dos resultados.

Em um primeiro momento, as análises para o modelo de MQO revelaram que as diferenças de rendimento escolar entre alunos rurais e urbanos diminuem quando há inserção das covariadas referente as características do aluno, da família e efeitos fixos de município. No entanto, ao examinar a trajetória desse *gap* durante os três anos, torna-se evidente que tais diferenças no desempenho rural-urbano aumentam no decorrer do tempo, principalmente no ano de 2021, período marcado pela crise da Covid-19, que pode ter agravado suficientemente as condições socioeconômicas dos alunos que residem na área rural. Outra descoberta interessante é que as desigualdades de rendimento escolar são mais sutis nas avaliações de matemática para todos os períodos investigados.

Uma limitação do estudo é a não utilização dos dados do Censo Escolar para extrair variáveis relacionadas a escola, como infraestrutura e características dos professores. Visto que, em detrimento da análise ser a nível aluno, gera maior dificuldade em ligar variáveis cujo os dados não são identificados. Portanto, é uma vantagem trabalhar com dados em nível de indivíduo, usufruindo do máximo de observações possíveis. A desvantagem está na complexidade de ligar uma base em nível de aluno e outra em nível de escola.

Dado o contexto de desigualdade no que concerne à localidade residencial do aluno, torna-se essencial a formulação de políticas públicas mais eficazes no tocante a redução das diferenças de desempenho, possibilitando, assim, um sistema de ensino mais igualitário. Logo, aumentar os investimentos ligados à distribuição de renda para a melhoria das condições socioeconômicas das famílias dos alunos que residem na área rural é importante a fim de impulsionar um melhor bem-estar e, conseqüentemente, melhorar os resultados educacionais. Outra questão importante é destinar mais recursos para o transporte escolar, visto que os alunos de tal área dependem disso para o deslocamento até às escolas.

5 CONCLUSÃO

Fundamentado no arcabouço teórico sobre Economia da Educação, com ênfase na educação básica, esta dissertação foi estruturada em três ensaios, por meio dos quais teve o intento de identificar os principais fatores que influenciam nos resultados educacionais. A literatura econômica existente, que é especializada na temática dos determinantes da qualidade educacional, aponta, de maneira unânime, as características familiares, o contexto socioeconômico e as habilidades intrínsecas dos alunos como fatores que influenciam significativamente no desempenho escolar. A discordância maior consiste na importância dos recursos educacionais nesse cenário.

De acordo com as evidências do Capítulo 2, tendo como base a função de produção educacional, o objetivo foi identificar os principais fatores observáveis que estão associados ao desempenho escolar no 5º ano do ensino fundamental em língua portuguesa e matemática entre o período de 2011 e 2017. Com a inserção de insumos voltados às características individuais do aluno, da família e da escola controlado por efeitos fixos, os resultados sugeriram que os aspectos escolares como características dos professores e infraestrutura não são tão essenciais para explicar o aprendizado escolar. Todavia, os atributos individuais dos alunos e a educação da mãe são importantes, trazendo ilação do que é discutido na literatura.

Visto a presente heterogeneidade na literatura no que condiz aos recursos empregados em educação, o Capítulo 3 teve a finalidade examinar a relação entre gastos com educação por aluno e qualidade educacional nos municípios brasileiros, com foco no 5º e 9º ano do ensino fundamental entre o hiato de 2013 e 2019. A partir da aplicabilidade de um painel de efeitos fixos, foi possível inferir que, em todas as especificações do painel, os coeficientes foram positivos e significantes, indicando que os aumentos nos gastos elavam a qualidade educacional nos municípios. Todavia, esses resultados demonstraram uma pequena correlação.

Finalmente, o Capítulo 4 buscou compreender a magnitude do diferencial de desempenho cognitivo entre os alunos do ensino médio que frequentam as escolas rurais e urbanas entre o período de 2017 e 2021. Como estratégia de identificação, foi empregado o modelo de MQO com efeitos fixos de estado e município, com a finalidade de captar atributos não observáveis. Os resultados em torno da média apresentaram que a inserção gradual das covariadas e efeitos fixos reduzem a lacuna de desempenho entre os alunos rurais

e urbanos com diferenças menores na proficiência em matemática. Todavia, as desigualdades aumentam com o passar dos anos.

Portanto, analisar os determinantes do desempenho escolar ajudam a compreender as áreas de melhoria no sistema educacional, como promover métodos de ensino adequado, melhor alocação de recursos à educação e políticas escolares. Ademais, tais investigações podem revelar desigualdades entre diferentes grupos de alunos, sendo eles socioeconômicos, étnicos ou geográficos, permitindo, dessa forma, o desenvolvimento de estratégias que promovam a equidade e a redução das disparidades. Enfim, o presente estudo serve para nortear os formuladores de políticas públicas no sentido de aprimorar ou elaborar políticas e programas de intervenção mais eficazes, que atendam às necessidades específicas dos alunos.

REFERÊNCIAS

ALBERNAZ, Ângela; FERREIRA, Francisco H.G.; FRANCO, Creso. **Qualidade e equidade na educação fundamental brasileira**. Rio de Janeiro: Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio), Departamento de Economia, 2002. Texto para discussão, n. 455.

ALENCAR, Nataniele dos Santos.; ARAUJO, Jair Andrade de; JUSTO, Wellington Ribeiro; SOBREIRA, Diogo Brito. Diferentes características dos professores explicam o gap educacional entre escolas urbanas e rurais no Brasil? **Estudios Económicos**, Bahía Blanca, v. 38, n. 76, p. 45-68, 2021.

ALVES, Pedro Jorge Holanda; FRIO, Gustavo Saraiva. Uma Análise dos resultados educacionais dos municípios brasileiros. **Planejamento e Políticas Públicas**, Brasília, n. 61, p. 285-316, 2022. Disponível em: <https://www.ipea.gov.br/ppp/index.php/PPP/article/view/1465>. Acesso em: 15 dez. 2023.

AMARAL, Luiz Felipe Leite Estanislau do; MENEZES FILHO, Naércio Aquino. A relação entre gastos educacionais e desempenho escolar. *In*: ENCONTRO NACIONAL DE ECONOMIA, 36., 2008, Salvador. **Anais [...]**. Rio de Janeiro: ANPEC, 2008. Disponível em: <https://en.anpec.org.br/previous-editions.php?r=encontro-2008>. Acesso em: 21 dez. 2023.

AMINI, Chiara; NIVOROZHKIN, Eugene. The urban–rural divide in educational outcomes: Evidence from Russia. **International Journal of Educational Development**, Amsterdam, v. 44, [s.n.], p. 118-133, 2015.

ANGRIST, Joshua D.; LAVY, Victor. Using Maimonides' rule to estimate the effect of class size on scholastic achievement. **The Quarterly journal of economics**, Cambridge, v. 114, n. 2, p. 533-575, 1999.

ANGRIST, Joshua; LAVY, Victor. New evidence on classroom computers and pupil learning. **The Economic Journal**, Cambridge, v. 112, n. 482, p. 735-765, 2002.

BARROS, Ricardo Paes de; MENDONÇA, Rosane; SANTOS, Daniel Domingues dos; QUINTAES; Giovani. **Determinantes do desempenho educacional no Brasil**. Rio de Janeiro, 2001. Texto para discussão, n. 834.

BECKER, Gary S. Investment in human capital: A theoretical analysis. **Journal of political economy**, Chicago, v. 70, n. 5, p. 9-49, 1962.

BECKER, Gary S. **Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education**. 2. ed. Chicago: University of Chicago Press, 2009.

BETANCUR, Laura; MILLER, Portia; VOTRUBA-DRZAL, Elizabeth. Urban-rural achievement gap in low-and middle-income countries: The role of early childhood education. **Early Childhood Research Quarterly**, Amsterdam, v. 66, p. 11-23, 2024.

BEZERRA, Marcio Garcia; KASSOUF, Ana Lucia. Análise dos fatores que afetam o desempenho escolar nas escolas das áreas urbanas e rurais do Brasil. *In*: CONGRESSO SOCIEDADE BRASILEIRA DE ECONOMIA, ADMINISTRAÇÃO E SOCIOLOGIA RURAL (SOBER), 44., 2006, Fortaleza. **Anais [...]**. Fortaleza: SOBER, 2006. Disponível em: <https://sober.org.br/anais/>. Acesso em: 10 jan. 2024.

BLINDER, Alan S. Wage discrimination: Reduced form and structural estimates. **Journal of Human resources**, Wisconsin, v. 8, n. 4, p. 436-455, 1973.

BOARDMAN, Anthony E.; MURNANE, Richard J. Using panel data to improve estimates of the determinants of educational achievement. **Sociology of education**, Washington, D.C., v. 52, n. 2, p. 113-121, 1979.

CARD, David; KRUEGER, Alan B. Labor market effects of school quality: Theory and evidence. **NBER Working Paper**, Cambridge, n. 5450, 1996.

CARNEIRO, Pedro Manuel; HECKMAN, James J. Human capital policy. **IZA Discussion Paper**, Chicago, n. 821, 2003.

CARNOY, Martin; ROSA, Leonardo; SIMÕES, Alexandre. Trends in the academic achievement gap between high and low social class children: The case of Brazil. **International Journal of Educational Development**, Califórnia, v. 94, p. 102650, 2022.

COLEMAN, James S. Equality of educational opportunity. **Integrated education**, Illinois, v. 6, n. 5, p. 19-28, 1968.

CURI, Andréa Zaitune; SOUZA, André Portela de. Medindo a qualidade das escolas: evidências para o Brasil. **Economia Aplicada**, São Paulo, v. 19, n. 3, p. 541-574, 2015.

CRUZ, Italo Spinelli da; MOURA, Fábio Rodrigues; ESPERIDIÃO, Fernanda. Diferença do desempenho educacional entre zonas rurais e urbanas: uma análise quantílica incondicional. *In*: ENCONTRO NACIONAL DA ASSOCIAÇÃO BRASILEIRA DE ESTUDOS REGIONAIS E URBANOS (ENABER), 20., 2022, Salvador. **Anais [...]**. Salvador: ABER, 2022. Disponível em: <https://brsa.org.br/enaber-2022/#artigos>. Acesso em: 10 jan. 2024.

DOURADO, Luiz Fernandes; DE OLIVEIRA, João Ferreira; DE ALMEIDA SANTOS, Catarina. **A qualidade da educação conceitos e definições**. Brasília: Instituto de Pesquisa Econômica Aplicada (IPEA), p. 69-69, 2007. Textos para discussão, 24.

FIRPO, S.; FORTIN, N.; LEMIEUX, T. Regressões quantílicas incondicionais, **Econometrica**, Nova York, v. 77, n.3, p. 953-973, 2009.

FRANCO, Ana Maria de Paiva; MENEZES FILHO, Naércio Aquino. Os determinantes do aprendizado com dados de um painel de escolas do SAEB. **Economia aplicada**, São Paulo, v. 21, n. 3, p. 525-548, 2017.

GLEWWE, Paul; HANUSHEK, Eric; HUMPAGE, Sarah; RAVINA, Renato. School

Resources and Educational Outcomes in Developing Countries: A Review of the Literature from 1990 to 2010. **NBER Working Papers**, Cambridge, n. 17554, 2011.

GLEWWE, Paul; KREMER, Michael. Schools, teachers, and education outcomes in developing countries. **Handbook of the Economics of Education**, Amsterdam, v. 2, p. 945-1017, 2006.

GOMEZ-GONZALEZ, Jose Eduardo.; RODRÍGUEZ-GÓMEZ, Wilson. RODRÍGUEZ-GÓMEZ, Efrén. Explaining the rural-urban student performance gap for different distribution quantiles in Colombia. **Documentos de Trabajo**, Chía, v. 74, [s,n], 2021.

GONÇALVES, Flávio de Oliveira; FRANÇA, Marco Túlio Aniceto. Qualidade educacional nos municípios nordestinos: evidências a partir do Prova Brasil 2005. **Planejamento e Políticas Públicas**, Brasília, n. 40, p. 107–139, 2013.

GOOLSBEE, Austan; GURYAN, Jonathan. The impact of Internet subsidies in public schools. **The Review of Economics and Statistics**, Cambridge, v. 88, n. 2, p. 336-347, 2006.

GREENWALD, Rob; HEDGES, Larry V.; LAINE, Richard D. The effect of school resources on student achievement. **Review of educational research**, Washington, D.C, v. 66, n. 3, p. 361-396, 1996.

HÄKKINEN, Iida; KIRJAVAINEN, Tanja; UUSITALO, Roope. School resources and student achievement revisited: new evidence from panel data. **Economics of Education Review**, Amsterdam, v. 22, n. 3, p. 329-335, 2003.

HANUSHEK, Eric. Teacher characteristics and gains in student achievement: Estimation using micro data. **The American Economic Review**, Tennessee, v. 61, n. 2, p. 280-288, 1971.

HANUSHEK, Eric A. Conceptual and empirical issues in the estimation of educational production functions. **Journal of human Resources**, Wisconsin, v. 14, n. 3, p. 351-388, 1979.

HANUSHEK, Eric A. The economics of schooling: Production and efficiency in public schools. **Journal of Economic Literature**, Tennessee, v. 24, n. 3, p. 1141-1177, 1986.

HANUSHEK, Eric A. The impact of differential expenditures on school performance. **Educational Researcher**, Califórnia, v. 18, n. 4, p. 45-62, 1989.

HANUSHEK, Eric A. Assessing the effects of school resources on student performance: An update. **Educational evaluation and policy analysis**, Washington, D.C, v. 19, n. 2, p. 141-164, 1997.

HANUSHEK, Eric A. Publicly provided education. **Handbook of public economics**, Amsterdam, v. 4, [s,n], p. 2045-2141, 2002.

HANUSHEK, Eric A. Alternative school policies and the benefits of general cognitive

- skills. **Economics of Education Review**, Amsterdam, v. 25, n. 4, p. 447-462, 2006.
- HANUSHEK, Eric A. Education production functions. **The economics of education**. 2.ed. Amsterdam: Academic Press, 2020. p. 161-170.
- HANUSHEK, Eric A.; KIMKO, Dennis D. Schooling, labor-force quality, and the growth of nations. **American economic review**, Tennessee, v. 90, n. 5, p. 1184-1208, 2000.
- HANUSHEK, Eric A.; LUQUE, Javier A. Efficiency and equity in schools around the world. **Economics of Education Review**, Amsterdam, v. 22, n. 5, p. 481-502, 2003.
- HANUSHEK, Eric A.; RAYMOND, Margaret E. The effect of school accountability systems on the level and distribution of student achievement. **Journal of the European Economic Association**, Tessengerlo, v. 2, n. 2-3, p. 406-415, 2004.
- HANUSHEK, Eric A.; RIVKIN, Steven G.; TAYLOR, Lori L. Aggregation and the estimated effects of school resources. **NBER Working Paper**, Cambridge, n. 5548, 1996.
- HANUSHEK, Eric A.; WOESSMANN, Ludger. The role of education quality for economic growth. **World Bank policy research working paper**, Washington, D.C, n. 4122, 2007.
- HANUSHEK, Eric A.; WOESSMANN, Ludger. The role of cognitive skills in economic development. **Journal of Economic Literature**, Tennessee, v. 46, n. 3, p. 607-668, 2008.
- HANUSHEK, Eric A.; WOESSMANN, Ludger. Education and economic growth. **Economics of Education**, Amsterdam, v. 60, [s,n], p. 67, 2010.
- HECKMAN, James J.; LOCHNER, Lance J.; TODD, Petra E. Earnings functions, rates of return and treatment effects: The Mincer equation and beyond. **Handbook of the Economics of Education**, Amsterdam, v. 1, p. 307-458, 2006.
- HEDGES, Larry V.; LAINE, Richard D.; GREENWALD, Rob. Does money matter? A meta analysis of studies of the effects of differential school inputs on student achievement. **Educational Researcher**, Washington, D.C, v. 23, [s,n], p. 9-10, 1994.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Taxa de distorção idade-série, 2023. Disponível em: <https://www.gov.br/inep/pt-br/aceso-ainformacao/dados-abertos/indicadores-educacionais/taxas-de-distorcao-idade-serie>. Acesso em: 18 set. 2023.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Sistema de Avaliação da Educação Básica, 2023. Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 18 set. 2023.
- KROTH, Darlan Christiano; GONÇALVES, Flavio de Oliveira. O impacto dos gastos públicos municipais sobre a qualidade da educação: Uma análise de variáveis instrumentais entre 2007 e 2011. **Planejamento e Políticas Públicas**, Brasília, n. 53, 2019.

- KRUEGER, Alan B. Experimental estimates of education production functions. **The Quarterly Journal of Economics**, Cambridge, v. 114, n. 2, p. 497-532, 1999.
- LAVY, Victor. Expanding school resources and increasing time on task: Effects on students' academic and noncognitive outcomes. **Journal of the European Economic Association**, Tessengerlo, v. 18, n. 1, p. 232-265, 2020.
- LEE, Jong-Wha; BARRO, Robert J. Schooling quality in a cross-section of countries. **Economica**, Londres, v. 68, n. 272, p. 465-488, 2001.
- LEUVEN, Edwin; LINDAHL, Mikael; OOSTERBEEK, Hessel; WEBBINK, Dinand. The effect of extra funding for disadvantaged pupils on achievement. **The Review of Economics and Statistics**, Cambridge, v. 89, n. 4, p. 721-736, 2007.
- LOUNKAEW, Kiatanantha. Explaining urban-rural differences in educational achievement in Thailand: Evidence from PISA literacy data. **Economics of Education Review**, Amsterdam, v. 37, p. 213-225, 2013.
- MATAVELLI, Ieda Rodrigues; MENEZES FILHO, Naércio Aquino. Efeitos de tamanho da sala no desempenho dos alunos: Evidências usando regressões descontínuas no Brasil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 74, n. 3, p. 352-401, 2020.
- MENEZES-FILHO, Naércio Aquino. **Os determinantes do desempenho escolar do Brasil**. São Paulo: Instituto Fernand Braudel de Economia Mundial, 2007.
- MENEZES FILHO, Naercio; OLIVEIRA, Alison Pablo de. A relação entre gastos em educação e desempenho escolar nos municípios brasileiros: uma análise com dados em painel. In: FERNANDES, Reynaldo; SOUZA, André Portela Fernandes de; BOTELHO, Fernando; SCORZAFAVE, Luiz Guilherme (org.). **Políticas públicas educacionais e desempenho escolar dos alunos da rede pública de ensino**. Ribeirão Preto: Funpec, 2014. Disponível em: <https://docplayer.com.br/7278363-Politicass-publicas-educacionais-e-desempenho-escolar-dos-alunos-da-rede-publica-de-ensino.html>. Acesso em: 15 dez. 2023.
- MENEZES FILHO, Naércio; PAZELLO, Elaine. Do teachers' wages matter for proficiency? Evidence from a funding reform in Brazil. **Economics of Education Review**, Amsterdam, v. 26, n. 6, p. 660-672, 2007.
- MINCER, Jacob. Investment in human capital and personal income distribution. **Journal of Political Economy**, Chicago, v. 66, n. 4, p. 281-302, 1958.
- MINCER, Jacob. Schooling, Experience, and Earnings. **National Bureau of Economic Research**, New York, v. 2, n. 2, 1974.
- MINISTÉRIO DA EDUCAÇÃO (MEC). Divulgados resultados do Brasil no Pisa 2022. 2023. Disponível em: <https://www.gov.br/mec/ptbr/assuntos/noticias/2023/dezembro/divulgados-os-resultados-do-pisa-2022>. Acesso em: 15 dez. 2023.
- MONTEIRO, Joana. Gasto público em educação e desempenho escolar. **Revista**

Brasileira de Economia, Rio de Janeiro, v. 69, n. 4, p. 467-488, 2015.

NIETO, Sandra; RAMOS, Raúl. Decomposition of differences in PISA results in middle income countries. **IZA Discussion Paper**, Bonn, n. 8028, 2014.

OAXACA, Ronald. Male-female wage differentials in urban labor markets. **International Economic Review**, Filadélfia, v. 14, n. 3, p. 693-709, 1973.

PANASSOL, Paulo Eduardo. **Gastos educacionais e desempenho escolar em municípios do Rio Grande do Sul**. Dissertação (Mestrado em Economia) – Universidade do Rio Grande do Sul, Rio Grande do Sul, 2018.

PEREIRA, Caroline Nascimento; CASTRO, César Nunes de. **Educação no meio rural: Diferenciais entre o rural e o urbano**. Brasília: Instituto de Pesquisa Econômica Aplicada (IPEA), n. 2632, 2021. Texto para Discussão.

RAMOS, R., DUQUE, J.C. e NIETO, S. Decomposing the rural-urban differential in student achievement in Colombia using PISA microdata. **IZA Discussion Paper**, Bonn, n. 6515, 2012.

RIVKIN, Steven G.; HANUSHEK, Eric A.; KAIN, John F. Teachers, schools, and academic achievement. **Econometrica**, New York, v. 73, n. 2, p. 417-458, 2005.

ROCHA, Vanderson Amadeu da; BELLUZZO, Walter; NICOLELLA, Alexandre C. Gastos e desempenho escolar: Evidências do Programa de Transferência de Recursos Financeiros do município de São Paulo. **Revista Brasileira de Economia**, Rio de Janeiro, v. 75, n. 4, p. 442-463, 2022.

RODRIGUES, Luciana de Oliveira; COSTA, Edward Martins; SILVA, Vitor Hugo Miro Couto; MARIANO, Francisca Zilania; JESUS FILHO, Jaime de. A note on performance differences between urban and rural schools in Brazil. **Revista Brasileira de Economia**, Rio de Janeiro, v. 74, n. 4, p. 494-507, 2021.

SCHULTZ, Theodore W. Investment in human capital. **The American Economic Review**, Tennessee, v. 51, n. 1, p. 1-17, 1961.

SETÚBAL, Maria Alice. Equidade e desempenho escolar: é possível alcançar uma educação de qualidade para todos?. **Revista brasileira de Estudos Pedagógicos**, Brasília, v. 91, n. 228, p. 345-366, 2010.

SILVA FILHO, Geraldo Andrade da. Efeito da formação docente sobre proficiência no início do ensino fundamental. **Revista Brasileira de Economia**, Rio de Janeiro, v. 73, n. 3, p. 385-411, 2019.

SOUZA, Dayane da Silva Rodrigues de; SAMPAIO, Luciano Menezes Bezerra; SAMPAIO, Raquel Menezes Bezerra. A modalidade de formação dos professores de matemática impacta no desempenho dos alunos do ensino fundamental. *In*: ENCONTRO NACIONAL DE ECONOMIA, 49., 2021, Online. **Anais [...]**. São Paulo: ANPEC, 2021. Disponível em: <https://en.anpec.org.br/previous-editions.php?r=encontro-2021>. Acesso

em: 15 dez. 2023.

TODD, Petra E.; WOLPIN, Kenneth I. On the specification and estimation of the production function for cognitive achievement. **The Economic Journal**, Londres, v. 113, n. 485, p. F3-F33, 2003.

TODD, Petra E.; WOLPIN, Kenneth I. The production of cognitive achievement in children: Home, school, and racial test score gaps. **Journal of Human Capital**, Chicago, v. 1, n. 1, p. 91-136, 2007.

TODOS PELA EDUCAÇÃO. Anuário Brasileiro da Educação Básica, 2021. Editora Moderna. Disponível em: <https://todospelaeducacao.org.br/anuario-da-educacao/>. Acesso em: 16 dez. 2023.

WALTENBERG, Fabio Domingues. **Análise econômica de sistemas educativos: Uma resenha crítica da literatura e uma avaliação empírica da iniquidade do sistema educativo brasileiro**. 2003. Dissertação (Mestrado em Economia) – Faculdade de Economia, Administração e Contabilidade, Universidade de São Paulo, São Paulo, 2003.

WOESSMANN, Ludger. Schooling resources, educational institutions and student performance: the international evidence. **Oxford bulletin of economics and statistics**, Oxford, v. 65, n. 2, p. 117-170, 2003.

WOESSMANN, Ludger. The importance of school systems: Evidence from international differences in student achievement. **Journal of Economic Perspectives**, Tennessee, v. 30, n. 3, p. 3-32, 2016.

WOOLDRIDGE, Jeffrey M. **Introductory econometrics: a modern approach**. 3. ed. South-Western Cengage Learning, 2006.

ZHANG, Dandan; LI, Xin; XUE, Jinjun. Education inequality between rural and urban areas of the People's Republic of China, migrants' children education, and some implications. **Asian Development Review**, Manila, v. 32, n. 1, p. 196-224, 2015.

APÊNDICE A – RESULTADOS DO MODELO COM BALANCEAMENTO

Tabela A.1 – Resultados do modelo para a proficiência em língua portuguesa e matemática no 5º ano do ensino fundamental

Variáveis	Língua portuguesa			Matemática		
	POLS	Efeitos aleatórios	Efeitos fixos	POLS	Efeitos aleatórios	Efeitos fixos
Gastos com EF por aluno	-0,021*** (0,002)	0,018*** (0,001)	0,028*** (0,001)	-0,026*** (0,002)	0,013*** (0,001)	0,022*** (0,001)
Gastos com saúde per capita	0,119*** (0,002)	0,077*** (0,001)	0,066*** (0,002)	0,113*** (0,002)	0,061*** (0,001)	0,049*** (0,002)
Razão professor/aluno	0,099*** (0,003)	0,054*** (0,004)	0,020*** (0,004)	0,100*** (0,004)	0,049*** (0,004)	0,014*** (0,004)
Educação da mãe	0,024*** (0,001)	0,025*** (0,001)	0,031*** (0,001)	0,019*** (0,001)	0,019*** (0,001)	0,024*** (0,004)
Constante	4,931*** (0,017)	4,746*** (0,014)	-	4,962*** (0,014)	4,962*** (0,014)	-
Observações	16.099	16.099	16.099	16.099	16.099	16.099
R ²	0,305	0,383	0,424	0,267	0,283	0,307

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Erros padrões robustos de White. (2) *p<0.1; **p<0.05; ***p<0.01.

Tabela A.2 – Resultados do modelo para a proficiência em língua portuguesa e matemática no 9º ano do ensino fundamental

Variáveis	Língua portuguesa			Matemática		
	POLS	Efeitos aleatórios	Efeitos fixos	POLS	Efeitos aleatórios	Efeitos fixos
Gastos com EF por aluno	-0,005*** (0,001)	0,011*** (0,001)	0,018*** (0,001)	-0,011*** (0,001)	0,007*** (0,001)	0,014*** (0,001)
Gastos com saúde per capita	0,065*** (0,001)	0,050*** (0,001)	0,043*** (0,001)	0,065*** (0,001)	0,040*** (0,001)	0,031*** (0,001)
Razão professor/aluno	0,041*** (0,002)	0,022*** (0,003)	0,00001 (0,003)	0,051*** (0,002)	0,025*** (0,003)	-0,0002 (0,003)
Educação da mãe	0,012*** (0,0004)	0,011*** (0,001)	0,013*** (0,001)	0,008*** (0,001)	0,008*** (0,001)	0,011*** (0,001)
Constante	5,216*** (0,011)	5,127*** (0,010)	-	5,324*** (0,012)	5,258*** (0,010)	-
Observações	16.103	16.103	16.103	16.103	16.103	16.103
R ²	0,234	0,300	0,312	0,206	0,223	0,220

Fonte: Elaboração própria com base nos dados da pesquisa.

Nota: (1) Erros padrões robustos de White. (2) *p<0.1; **p<0.05; ***p<0.01.

APÊNDICE B – ESTATÍSTICA DESCRITIVA DAS VARIÁVEIS ENTRE 2017 E 2021

Tabela B.1 – Estatística descritiva do ano de 2017

Variável	Média	Desvio padrão	Mínimo	Máximo
Língua portuguesa	0,02663	0,99587	-2,389	2,904
Matemática	0,02598	0,99965	-2,304	3,908
Rural	0,04163	0,19974	0	1
NSE	0,00007	0,99995	-1,983	3,05
Branços	0,31413	0,46417	0	1
Pardos	0,50141	0,5	0	1
Pretos	0,13378	0,34041	0	1
Indígenas	0,01592	0,12518	0	1
Amarelos	0,03476	0,18317	0	1
Quant. de pessoas no domicílio	0,33805	0,47304	0	1
Educação da mãe	0,44585	0,49706	0	1

Fonte: Elaboração própria baseado nos dados da pesquisa.

Tabela B.2 – Estatística descritiva do ano de 2019

Variável	Média	Desvio padrão	Mínimo	Máximo
Língua portuguesa	0,31645	0,98959	-2,404	2,953
Matemática	0,26841	1,01452	-2,309	3,584
Rural	0,04251	0,20176	0	1
NSE	-0,00002	1,02388	-2,498	2,49
Branços	0,32857	0,4697	0	1
Pardos	0,48747	0,49984	0	1
Pretos	0,14038	0,34738	0	1
Indígenas	0,01416	0,11817	0	1
Amarelos	0,02941	0,16895	0	1
Quant. de pessoas no domicílio	0,08483	0,27863	0	1
Educação da mãe	0,52698	0,49927	0	1

Fonte: Elaboração própria baseado nos dados da pesquisa.

Tabela B.3 – Estatística descritiva do ano de 2021

Variável	Média	Desvio padrão	Mínimo	Máximo
Língua portuguesa	0,31645	0,98959	-2,404	2,953
Matemática	0,26841	1,01452	-2,309	3,584
Rural	0,04251	0,20176	0	1
NSE	-0,00002	1,02388	-2,498	2,49
Branços	0,32857	0,4697	0	1
Pardos	0,48747	0,49984	0	1
Pretos	0,14038	0,34738	0	1
Indígenas	0,01416	0,11817	0	1
Amarelos	0,02941	0,16895	0	1
Quant. de pessoas no domicílio	0,08483	0,27863	0	1
Educação da mãe	0,52698	0,49927	0	1

Fonte: Elaboração própria baseado nos dados da pesquisa.