



Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Economia

Economia da Educação: Os efeitos do gosto e da internet como insumos para proficiência

Elias das Neves Freire

João Pessoa–PB, Setembro de 2021

Elias das Neves Freire

Economia da Educação: Os efeitos do gosto e da internet como insumos para proficiência

Tese apresentada ao Programa de Pós Graduação em Economia da Universidade Federal da Paraíba - UFPB, em cumprimento às exigências do curso de Doutorado em Economia.

Universidade Federal do Paraíba – UFPB

Centro de Ciências Sociais

Programa de Pós-Graduação em Economia

Orientador: Profa. Dra. Mércia Santos da Cruz

João Pessoa–PB

Setembro de 2021

Catálogo na publicação
Seção de Catalogação e Classificação

F866e Freire, Elias das Neves.

Economia da educação : os efeitos do gosto e da internet como insumos para proficiência / Elias das Neves Freire. - João Pessoa, 2021.
72 f. : il.

Orientação: Mércia Santos da Cruz.
Tese (Doutorado) - UFPB/CCSA.

1. Economia da educação. 2. Proficiência. 3. Gosta de estudar. 4. Prova Brasil. 5. Computador. 6. Internet.
I. Cruz, Mércia Santos da. II. Título.

UFPB/BC

CDU 33:37(043)



Universidade Federal da Paraíba
Centro de Ciências Sociais Aplicadas
Programa de Pós-Graduação em Economia

Campus Universitário I – Cidade Universitária – CEP 58.059-900 – João Pessoa – Paraíba
Tel: (83) 3216-7482 – <https://sigaa.ufpb.br/sigaa/public/programa/portal.jsf?id=1875> – E-mail: ppge.ccsa@gmail.com

Ata da reunião da Banca Examinadora designada para avaliar o trabalho de tese do doutorando **Elias das Neves Freire**, submetida para obtenção do grau de doutor em Economia, área de concentração em **Economia Aplicada**.

Aos vinte e oito dias do mês de setembro do ano dois mil e vinte e um, às quinze horas, no Programa de Pós-Graduação em Economia, do Centro de Ciências Sociais Aplicadas, da Universidade Federal da Paraíba, reuniram-se, em cerimônia pública, os membros da Banca Examinadora, constituída por **Profª. Drª. Mércia Santos da Cruz** (Orientadora), da Universidade Federal da Paraíba; **Prof. Dr. Paulo Aguiar do Monte** (Examinador Interno), da Universidade Federal da Paraíba; **Prof. Dr. Wallace Patrick Santos de Farias Souza** (Examinador Interno), da Universidade Federal da Paraíba; **Prof. Dr. Francisco Soares de Lima** (Examinador Externo), da Universidade Estadual do Rio Grande do Norte; e **Profª. Drª. Fabrícia Jóiisse Vitorino Carvalho** (Examinadora Externa), da Universidade Federal do Mato Grosso do Sul, a fim de examinarem o candidato ao grau de doutor em Economia, área de concentração em **Economia Aplicada**, **ELIAS DAS NEVES FREIRE**. Além dos examinadores e do examinando, compareceram também, representantes do Corpo Docente e do Corpo Discente. Iniciando a sessão, a professora **Profª. Drª. Mércia Santos da Cruz**, na qualidade de presidente da Banca Examinadora, comunicou aos presentes a finalidade da reunião e os procedimentos de encaminhamento desta. A seguir, concedeu à palavra ao candidato, para que fizesse oralmente a exposição do trabalho, apresentado sob o título: **“Economia da Educação: Os efeitos do gosto e da internet como insumos para proficiência”**. Concluída a exposição, a senhora presidente solicitou que fosse feita a arguição por cada um dos examinadores. A seguir foi concedida a palavra ao candidato, para que respondesse e esclarecesse às questões levantadas. Terminadas as arguições, a Banca Examinadora passou a proceder à avaliação e ao julgamento do candidato. Em seguida, a senhora presidente comunicou aos presentes que a Banca Examinadora, por unanimidade, **aprovou** a tese apresentada e defendida com o conceito **APROVADO**, concedendo assim, o grau de **Doutor em Economia**, área de concentração em **Economia Aplicada**, ao doutorando **ELIAS DAS NEVES FREIRE**. E, para constar, eu, Ricardo Madeira Cataldi, secretário *ad hoc* do Programa de Pós-Graduação em Economia, lavrei a presente ata, que assino junto com os membros da Banca Examinadora. João Pessoa, 28 de setembro de 2021.

Profª. Drª. Mércia Santos da Cruz
Orientadora – UFPB

Paulo Aguiar do Monte
Prof. Dr. Paulo Aguiar do Monte
Examinador Interno – UFPB

Wallace Patrick Santos de Farias Souza
Prof. Dr. Wallace Patrick Santos de Farias Souza
Examinador Interno – UFPB

Francisco Soares de Lima
Prof. Dr. Francisco Soares de Lima
Examinador Externo – UERN

Fabrícia Jóiisse Vitorino Carvalho
Profª. Drª. Fabrícia Jóiisse Vitorino Carvalho
Examinadora Externa – UFMS

Ricardo Madeira Cataldi
Ricardo Madeira Cataldi
Secretário Ad Hoc – PPGE/ UFPB

Emitido em 19/10/2021

ATA Nº 1/2021 - PPGE (11.01.13.38)
(Nº do Documento: 1)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado digitalmente em 19/10/2021 10:41)
MERCIA SANTOS DA CRUZ
PROFESSOR DO MAGISTERIO SUPERIOR
1646137

Para verificar a autenticidade deste documento entre em <https://sipac.ufpb.br/documentos/> informando seu número: **1**,
ano: **2021**, documento (espécie): **ATA**, data de emissão: **19/10/2021** e o código de verificação: **c2b43f7a1b**

A minha esposa Vanubia, Maria Luzia (filha), Pedro Henrique (filho), meus irmãos e meus pais in memoriam, por sempre estarem comigo em todos os momentos.

Agradecimentos

À minha orientadora, Prof. Dra. Mércia Santos da Cruz, pela paciência na orientação e incentivo, que tornaram possível a realização desta tese.

Aos demais conselheiros da banca examinadora, professores Dr. Paulo Aguiar do Monte, Dr. Wallace Patrick Santos de Farias Souza, Dr. Francisco Soares de Lima e Dra. Fabrícia Jóisse Vitorino Carvalho, pelas contribuições e sugestões dadas em prol do aprimoramento desta tese.

Aos meus colegas do DINTER-UFPB/UERN/UFERSA pelo apoio, carinho, incentivo e parceria nos momentos difíceis do decorrer deste doutorado.

Aos professores do departamento de economia da UFPB que ministraram às disciplinas do doutorado com maestria.

Aos meus amigos e colegas professores do IFRN-Mossoró, que me incentivaram à realizar essa pós-graduação.

Aos colegas professores do DME-UERN, pelo apoio e incentivo dado à mim na busca por mais uma qualificação profissional.

Aos meus amigos matemáticos Ênio, Jeovanizélio, Laudelino, Otoniel, Rafael, Amorim, Braz, Djair, Fagno, Rildo, Hélio(DME), Valderir, Damiano Aquino, Ronaldo Garcia e Walter Martins pela amizade e incentivo acadêmico sempre presente.

Aos amigos especiais: Aldiere, Glauber, Hailson, Stefeson, Vladimir, Django, Djair, Fagno, Rildo, Paulo Henrique, Marvininho, Valderir e primos: Mailsom, Matsom e Erinaldo pelo carinho e paciência nos meus momentos de “loucura” durante o desenvolvimento desse trabalho.

Aos meus irmãos, cunhados e esposas: Eliel, Marta, Luana, Camila, Aldemir, Raniere, Flávio, Fábio, Francinaldo, Cleide e Alexsandra aos quais tenho enorme carinho e admiração, pelo apoio sempre presente.

A minha querida e amada esposa Vanubia, por acreditar sempre na minha capacidade de vencer obstáculos.

Aos meus filhos Maria Luzia e Pedro Henrique, que sempre têm estado ao meu lado nos momentos de alegria e dificuldade.

Ao meu pai Valmir e a minha mãe Elza, *in memoriam*, que sempre incentivaram seus filhos à estudarem.

À CAPES pelo apoio financeiro para realização desta tese.

*“Educação nunca foi despesa. Sempre foi
investimento com retorno garantido.”
(Arthur Lewis)*

Resumo

Esta tese é composta por dois ensaios independentes relacionados a economia da educação. O primeiro ensaio tem como objetivo mensurar o efeito de gostar de estudar matemática sobre o desempenho em matemática. Para tanto, foi usado os microdados da Prova Brasil de 2017, para uma amostra de alunos do 9º ano do ensino fundamental das escolas públicas. Como estratégia empírica utilizou-se o propensity score matching (pareamento pelo escore de propensão) e o método de Lewbel (2012). Os resultados mostraram um efeito positivo na proficiência em Matemática, confirmando a hipótese de que gostar de estudar Matemática tem-se um efeito positivo no desempenho. O segundo ensaio buscou avaliar o efeito do computador com internet nas escolas sobre o desempenho em Matemática dos alunos do 9º ano da rede pública das escolas brasileiras. Utilizando os microdados da Prova Brasil (2019) e do Censo (2019) e uma estratégia empírica baseada nos modelos multiníveis lineares, encontrou-se evidências de efeitos positivos no desempenho em Matemática, mesmo após o controle por características do aluno, da turma e da escola. Conclui-se, dessa forma, que mais atenção deve ser dada a essa questão, que demanda políticas públicas eficazes para a promoção de um ambiente mais propício ao aprendizado e ao acúmulo de capital humano.

Palavras-chaves: Proficiência; Gosta de estudar; Prova Brasil; Computador; Internet; Economia da educação.

Abstract

This thesis is composed of two independent essays related to the economics of education. The first essay objective to measure the effect of enjoying studying mathematics on performance in mathematics. For this purpose, microdata from Prova Brasil of 2017. was used for a sample of students from the 9º year of public elementary school. The empirical strategy used was propensity score matching (propensity score matching) and the Lewbel method (2012). The results showed a positive effect on proficiency in Mathematics, confirming the hypothesis that enjoying studying Mathematics has if a positive effect on performance. The second essay sought to evaluate the effect of computers with internet in schools on the performance in Mathematics of students in the 9º year of public schools in Brazilian schools. Using the microdata from the Prova Brasil (2019) and from the Census (2019) and an empirical strategy based on linear multilevel models, evidence of positive effects on performance in Mathematics was found, even after controlling for characteristics of the student, of the class and school. It is therefore concluded that more attention should be given to this issue, which requires effective public policies to promote an environment more conducive to learning and the accumulation of human capital.

Keywords: Proficiency; Like study; Brazil exam; Computer; Internet; Education economics..

Lista de ilustrações

Figura 1 – Distribuição das notas de matemática por gostar de estudar	10
Figura 2 – Função densidade do escore de propensão	23
Figura 3 – Função densidade do escore de propensão para os reprovados	23

Lista de quadros

Quadro 1 – Descrição e metodologia das variáveis utilizadas nas análises economé-	
tricas	9
Quadro 2 – Descrição e metodologia das variáveis utilizadas nas análises multiníveis	33

Lista de tabelas

Tabela 1 – Proporção de alunos que gostam e não gostam de matemática por regiões	11
Tabela 2 – Estatísticas descritivas das variáveis utilizadas nos modelos	12
Tabela 3 – Resultados do modelo logit que estima a probabilidade do aluno gostar de estudar Matemática	20
Tabela 4 – Estimação do efeito médio de tratamento	21
Tabela 5 – Estimação do efeito médio de tratamento dos reprovados	22
Tabela 6 – Limites de Rosenbaum a análise de sensibilidade de tendências ocultas devido a variáveis omitidas	24
Tabela 7 – ATT estimado - Método Lewbel (2012)	26
Tabela 8 – Estatística descritiva das variáveis usadas nos modelos	35
Tabela 9 – Modelo de covariâncias e correlações implícitas para os quatro diferentes pares possíveis de alunos	40
Tabela 10 – Resultado da estimação dos modelos nulo e ancova para a proficiência em Matemática	44
Tabela 11 – Resultado da estimação dos modelos com intercepto aleatório para a proficiência em Matemática	49

Sumário

1	INTRODUÇÃO	1
2	PROVA BRASIL: O QUE FAZ A DIFERENÇA NO DESEMPENHO ? O EFEITO DE GOSTAR DE ESTUDAR MATEMÁTICA SOBRE A PROFICIÊNCIA	3
2.1	Introdução	3
2.2	Revisão da literatura	5
2.3	Aspectos Metodológicos	7
2.3.1	Dados	7
2.3.2	Estratégia Empírica	13
2.3.2.1	Pareamento por Escore de Propensão	13
2.3.2.2	Método de Lewbel (2012)- Identificação com Heterocedasticidade	16
2.3.2.3	Análise de sensibilidade	17
2.4	Resultados	18
2.4.1	Pareamento pelo escore de propensão	18
2.4.2	Análise de Robustez: Método de Lewbel	25
2.5	Considerações Finais	26
3	O EFEITO DE COMPUTADORES COM INTERNET NAS ESCOLAS NA PROFICIÊNCIA DOS ESTUDANTES: UMA ANÁLISE MULTINÍVEL	28
3.1	Introdução	28
3.2	Aspectos metodológicos	32
3.2.1	Dados	32
3.2.2	Estratégia Empírica	36
3.2.2.1	Modelo Incondicional ou Nulo	36
3.2.2.2	Modelo Condicional	40
3.3	Resultados e discussões	42
3.3.1	Modelo Nulo (Modelo 1) e Modelo Ancova (Modelo 2)	42
3.3.2	Modelos com intercepto aleatório e com variáveis explicativas no nível do aluno, da escola e da turma - Modelo 3 e Modelo 4	46
3.4	Considerações finais	50
	REFERÊNCIAS	52

1 Introdução

A função de produção educacional apresentada nos dois ensaios desta tese tem como produto o desempenho do aluno e como insumos as características dos alunos, turmas e escolas. Sabendo que essa função matemática é de várias variáveis, isto é, apresentam vários insumos, notamos a importância de testar outros insumos que não se encontra na literatura para explicar o desempenho escolar do aluno. Para [Hanushek \(2020\)](#) o modelo representado por uma função de produção educacional relaciona o desempenho de cada aluno, por exemplo, com sua capacidade de aprendizagem. Vale ressaltar que em ambos os ensaios a variável dependente é o desempenho em matemática. Essa escolha de se trabalhar com a proficiência em Matemática parte de que alguns autores [Andrade e Soares \(2008\)](#), [Rodrigues, Rios-Neto e Pinto \(2011\)](#), [Palermo, Silva e Novellino \(2014\)](#) por apontarem que esta disciplina seria mais afetada pela qualidade das aulas e da própria escola do que a proficiência em língua portuguesa. Segundo [Palermo, Silva e Novellino \(2014\)](#) as habilidades de leitura, por sua vez, seriam fortemente influenciadas pelo capital cultural que os alunos trazem de casa.

A matemática é uma disciplina considerada distante por parte de um grupo de alunos. Um dos principais desafios para o professor de matemática é fazer com que os alunos tenham uma atitude positiva em relação à aprendizagem da matemática. Portanto, os professores devem estar atentos às crenças afetivas dos alunos e às inter-relações daqueles que estão aprendendo matemática, a fim de empregar estratégias mais eficazes no ensino e melhorar a aprendizagem dos alunos em matemática, reduzindo suas crenças negativas. ([GAFOOR; KURUKKAN, 2015](#)).

O gosto por matemática é medido a partir do gosto dos alunos de fazer e aprender matemática ([KUPARI; NISSINEN, 2013](#)). Por outro lado o gosto dos alunos durante a aprendizagem pode influenciar seu comportamento ou aspecto cognitivo da atitude ([SYYEDA, 2016](#)). Ainda no tocante ao gosto por matemática, [Dowker, Bennett e Smith \(2012\)](#) afirmam que a maioria dos estudos de alunos do ensino médio e adultos têm sido consistentes em mostrar que o desempenho real em matemática tende a estar relacionado positivamente ao gosto por matemática e às altas auto-avaliações em matemática, e negativamente à ansiedade matemática por exemplo, ([AIKEN, 1974; HEMBREE, 1990; MA; KISHOR, 1997; ASHCRAFT; KRAUSE, 2007](#)).

Dito isto, o objetivo do primeiro ensaio é mensurar o efeito do aluno gostar de estudar matemática no desempenho médio controlado por variáveis associadas ao aluno, à escola, ao professor e ao diretor utilizando os microdados da Prova Brasil de 2017. ¹

¹ Vale salientar que o não uso dos microdados da Prova Brasil mais recentes (2019), deu-se pelo fato da

A inserção da informática na escola como recurso pedagógico deve partir da constatação feita pela própria comunidade escolar da necessidade de mudança no processo educacional, a fim de adequar o ensino às novas demandas sociais. Para que os recursos e os benefícios da informática possam ser utilizados de forma consciente, eficaz e crítica, é necessário haver mobilização, discussão e reflexão. Uma melhor compreensão de como a tecnologia computacional afeta os resultados educacionais é crítica porque esclarece se essa tecnologia é um insumo importante no processo de produção educacional e se as disparidades no acesso se traduzirão em desigualdade educacional. (BULMAN; FAIRLIE, 2016)

Com a utilização do computador com internet para fins educacionais é possível ao professor e à escola dinamizarem o processo de ensino-aprendizagem com aulas mais criativas, mais motivadoras e que despertem, nos alunos, a curiosidade e o desejo de aprender, conhecer e fazer descobertas.

O acesso a computadores nas escolas por aluno varia em alguns países. Em média, nos países da OCDE em 2018, havia quase um computador disponível na escola para fins educacionais para cada aluno de 15 anos (proporção computador-aluno igual a 0,8). Na Áustria, Islândia, Luxemburgo, China, Nova Zelândia, Reino Unido e Estados Unidos, a proporção de estudantes de informática era de 1,25 ou mais, enquanto na Albânia, Brasil, Grécia, Kosovo, Montenegro, Marrocos, Turquia e Vietnã ou seja, havia apenas um computador disponível para cada 4 alunos (proporção = 0,25) ou menos. (PEÑA-LÓPEZ et al., 2020)

Ainda de acordo Peña-López et al. (2020) o Brasil é o segundo país com a menor quantidade de computadores por estudante na escola em um ranking de 79 países e territórios feito pela Organização para Cooperação e Desenvolvimento Econômico (OCDE). De acordo com o levantamento realizado, na classificação dos países pelo percentual de computadores nas escolas conectadas à internet, o Brasil aparece na 52ª colocação.

Dito isto, o segundo ensaio explora a teoria e a literatura sobre os impactos da tecnologia nos resultados educacionais. Isto é, a literatura concentra-se no contexto primário no qual a tecnologia pode ser usada para fins educacionais no uso do computador com internet nas escolas pelos alunos do 9º ano do ensino fundamental das escolas públicas brasileiras. Assim, o objetivo desse segundo ensaio é avaliar o efeito do computador com internet nas escolas sobre o desempenho em Matemática dos alunos do 9º ano da rede pública das escolas brasileiras utilizando os microdados da Prova Brasil (2019) e do Censo (2019).

2 PROVA BRASIL: O que faz a diferença no desempenho ? O efeito de gostar de estudar Matemática sobre a proficiência

2.1 Introdução

A matemática é uma disciplina única e é uma parte fundamental do currículo escolar. É um instrumento para o desenvolvimento de todas as outras ciências. Consciente ou não, estamos usando a matemática em todas as facetas da vida. No entanto, a maioria dos alunos em todo o mundo não gosta de matemática. [Scarpello \(2007\)](#) relata que setenta e cinco por cento dos americanos param o estudo da matemática e ficam longe de muitas carreiras relacionadas à matemática. Ele identifica a ansiedade matemática como uma das principais razões para isso.

O gosto por matemática é medido a partir do gosto dos alunos de fazer e aprender matemática ([KUPARI; NISSINEN, 2013](#)). Por outro lado o gosto dos alunos durante a aprendizagem pode influenciar seu comportamento ou aspecto cognitivo da atitude ([SYYEDA, 2016](#)). Assim, de acordo com os resultados do PISA 2012 publicados pela OCDE (2013), os alunos podem aprender matemática porque a consideram agradável e interessante. Eles afirmam ainda que o interesse e o gosto afetam o grau e a continuidade do envolvimento no aprendizado e a profundidade do entendimento. Para [Mazana, Montero e Olifage \(2019\)](#) isso significa que quanto mais os alunos gostam de matemática, mais eles tendem a se envolver na resolução de problemas, melhorando assim seu aprendizado e desempenho. Uma vez que o gosto, a aprendizagem e o desempenho dos alunos estão relacionados, vale a pena avaliar o status dos alunos em relação ao gosto pela matemática, a fim de acompanhar os outros dois fatores elucidados.

Ainda no que tange ao gosto por matemática, [Markovits e Forgasz \(2017\)](#) afirmam que os alunos do ensino fundamental relatam altos níveis de prazer em matemática, embora haja algumas exceções. No entanto, o gosto pode se deteriorar com a idade de acordo com ([ADELSON; MCCOACH, 2011](#); [UPADYAYA; ECCLES, 2014](#); [GOTTFRIED et al., 2007](#)). Por outro lado, [Beek et al. \(2017\)](#) descobriram que o desempenho em matemática se correlacionou positivamente com o gosto por matemática e negativamente com ansiedade em matemática em uma amostra de adolescentes holandeses. [Pinxten et al. \(2014\)](#) realizaram um estudo longitudinal com crianças belgas do ensino fundamental, e o prazer em matemática foi um indicador positivo moderado de desempenho em matemática posterior e esforço percebido posteriormente em matemática. Tem havido comparativamente

poucas pesquisas sobre os fatores que fazem as crianças gostarem ou não de matemática, mas [Upadyaya e Eccles \(2014\)](#) descobriram que as avaliações dos professores sobre o esforço e a capacidade dos alunos do ensino fundamental em matemática prevêm não apenas seu desempenho subsequente e auto-avaliação matemática, mas seu subsequente desfrute dela.

No que tange ao exposto, podemos perceber que o desempenho escolar é explicado por variáveis associadas, por exemplo, ao aluno. Sendo assim, o objetivo desse estudo é mensurar o efeito do aluno gostar de estudar matemática no desempenho médio controlado por variáveis associadas ao aluno, à escola, ao professor e ao diretor utilizando os microdados da Prova Brasil de 2017. Vale salientar que a preferência de se trabalhar com a proficiência em Matemática parte de que alguns autores [Andrade e Soares \(2008\)](#), [Rodrigues, Rios-Neto e Pinto \(2011\)](#), [Palermo, Silva e Novellino \(2014\)](#) apontarem que esta disciplina seria mais afetada pela qualidade das aulas e da própria escola do que a proficiência em língua portuguesa.

Diante disso, parte-se da hipótese de que gostar de estudar matemática, aumenta o efeito no desempenho em matemática. Assim, para viabilizar o teste da hipótese, a estratégia empírica consiste em estimar um PSM (Pareamento por escore de propensão) pelo método Kernel e do vizinho mais próximo (1 controle para 1 tratado) para obtermos o efeito médio de tratamento nos tratados (alunos que gostam de estudar matemática) sobre a variável resposta Proficiência em Matemática. Dado a ausência de instrumentos externos disponíveis, será usado em seguida um estimador de dois estágios proposto por [Lewbel \(2012\)](#) que explora a heterocedasticidade do primeiro estágio para gerar instrumentos para identificação.

Levando em consideração que o gosto por matemática e o desempenho estão diretamente relacionados, esse estudo justifica-se devido aos dados apresentados no SAEB 2017, onde 65,96% dos alunos afirmaram que gostam de estudar matemática e além do mais do percentual de alunos reprovados, aproximadamente 60% afirmaram ter apreço por essa disciplina.

Os problemas relacionados para o tema apresentado nesse estudo vão desde o mito de que a matemática é considerada uma disciplina difícil de ser compreendida por parte de alguns alunos, a causalidade reversa presente, ou seja, a variável explicativa “Gosto” é endógena, o tema pouco explorado na economia da educação até as informações importantes não presentes no banco de dados.

Dessa forma, é possível notar que o gosto por matemática pode ter efeitos positivos no aluno, na escola e na sociedade, melhorando o desempenho dos alunos nas escolas e trazendo benefícios como: avanços tecnológicos, desenvolvimento econômico de um país por meio de modelos matemáticos e bons resultados em matemática dos alunos em exames

como o PISA¹ e a OBMEP².

Esse capítulo possui cinco seções incluindo esta introdução. Na segunda seção tem-se a revisão de literatura e na terceira os aspectos metodológicos. Na quarta seção tem-se os resultados das estimativas. Finalmente, na quinta seção apresentam-se as considerações finais.

2.2 Revisão da literatura

A revisão da literatura apresentada nessa seção trata de alguns estudos que relaciona gosto por matemática e desempenho. É importante destacar que uma atitude em relação à matemática é definida como uma medida agregada de gostar e não gostar dessa disciplina, uma tendência para evitar ou se envolver em atividades à ela e uma crença de que a matemática é útil ou inútil (MA; KISHOR, 1997).

Em um estudo de Parsons, Croft e Harrison (2009), foi detectado uma relação estatisticamente significativa entre a confiança dos alunos em matemática e o desempenho. Esse estudo descobriu que gostar de matemática estava relacionado ao desempenho. Ao identificar domínios afetivos que se correlacionam e predizem a proficiência nessa disciplina entre alunos em Cingapura, Lim e Chapman (2015) relataram que, entre as variáveis afetivas, a autoconfiança em matemática e o valor percebido por ela estavam correlacionados com o desempenho. Por outro lado, Stankov, Morony e Lee (2014) também observou a importância da confiança como um preditor de desempenho acadêmico. A partir dessas literaturas, pode-se deduzir que o gosto dos alunos pela aprendizagem da matemática, o valor e a confiança são alguns dos fatores afetivos a serem considerados na previsão do desempenho dos alunos.

De acordo com Belbase (2010), Mullis et al. (2011), Tavsancil e Yalcin (2015), Khine, Al-Mutawah e Afari (2015) gostar de matemática, sendo uma das características afetivas orientada para a matemática e relacionada com o desempenho, consiste nas reações afetivas e comportamentais dos alunos ao seu interesse pela matemática. Esses pesquisadores chegaram à conclusão de que os alunos que gostam de aprender à disciplina tem níveis mais altos de aproveitamento dessa matéria.

Ainda em relação ao gosto pela matemática, alguns estudos mostram que a relação entre o desempenho e o gosto é positiva, isto é, alunos que possuem uma aptidão voltada para a matemática tendem a obter melhores resultados em comparação com aqueles que não gostam da disciplina. Mokgwathi, Graham e Fraser (2019) investigaram a relação entre as percepções e atitudes dos professores sul-africanos do 9º ano e dos alunos com o seu desempenho. Nesse estudo os autores usaram uma modelagem multinível e verificaram que

¹ Programa Internacional de Avaliação de Estudantes

² Olimpíadas Brasileiras de Matemática das Escolas Públicas

os alunos que gostam de aprender matemática e à valorizam se sobressaem em relação aos que não gostam e não valorizam. Também foi descoberto que os alunos que estão confiantes em matemática superaram os que não estão confiantes. Uma descoberta interessante foi que a satisfação profissional do professor não tem influência no desempenho em matemática de seus alunos.

No tocante a relação entre desempenho e atitude em relação à matemática, [Ma e Kishor \(1997\)](#), [Papanastasiou \(2000\)](#), [Nicolaidou e Philippou \(2003\)](#), [Bramlett \(2007\)](#) encontraram uma relação positiva. Já de acordo com [Syed \(2016\)](#) a atitude em relação à matemática se desenvolve ao longo do tempo e pode afetar seu autoconceito e comportamento em sala de aula. Acredita-se que os alunos que gostam de Matemática e têm alto autoconceito e autoestima sobre suas habilidades prestam mais atenção, participam ativamente em sala de aula e estão mais envolvidos no processo de aprendizagem.

Por outro lado, [Stipek \(2002\)](#) revela que uma atitude negativa em relação ao assunto pode impedir o progresso dos alunos e afetar sua autoconfiança e moral. O prazer dos alunos durante a aprendizagem e sua utilidade percebida da matemática (cognição) podem influenciar seu comportamento ou aspecto conativo de atitude. Alunos com emoções positivas em relação à matemática mostram um maior nível de comprometimento e maior capacidade de exercer esforços que resultam em melhores resultados. Um estudo de caso conduzido por [Renninger e Hidi \(2002\)](#) afirma que “um aluno com um interesse bem desenvolvido pelo assunto pode ter alto desempenho, enquanto um aluno com um interesse menos desenvolvido tem menos probabilidade de grandes realizações”. Um estudo Internacional de Matemática e Ciências da IEA (TIMSS, 2011) também encontrou correlação positiva generalizada entre gostar de matemática e desempenho.

Ainda no que se refere ao gosto pela matemática, alguns estudos como apontados em [Laros, Marciano e Andrade \(2010\)](#), [Rodrigues \(2017\)](#), [Paula e Couto \(2019\)](#) encontraram uma relação positiva entre desempenho e gosto pela matemática. Por outro lado, [Cazorla et al. \(2008\)](#) constataram que estudantes das séries iniciais do ensino fundamental que afirmam gostar de Matemática obtiveram desempenho médio superior aos demais estudantes.

Cabe ainda ressaltar que a importância das habilidades matemáticas para o sucesso do estudo em economia tem sido amplamente pesquisada. De acordo com [Arnold e Straten \(2012\)](#) na população de alunos com deficiência em matemática, a motivação é importante. Aplicando uma análise fatorial a uma pesquisa, os autores identificaram quatro fatores motivacionais; entre estes, a motivação intrínseca está mais fortemente relacionada ao sucesso do estudo no primeiro ano. Os autores também mostram que a motivação intrínseca pode ajudar a superar a educação matemática preparatória inadequada.

Diante disso, o desempenho do aluno na Matemática merece especial atenção. Segundo [Hanushek e Woessmann \(2012\)](#), o crescimento da produtividade econômica de um país, ou de uma região, é conduzido de forma mais clara pela proficiência na Matemática

dos seus alunos, do terceiro ciclo do ensino básico, que pelo seu desempenho em outras disciplinas. Esse estudo Forneceu evidências de que a associação robusta entre habilidades cognitivas e crescimento econômico reflete um efeito causal das habilidades cognitivas e apóia os benefícios econômicos de uma política escolar eficaz.

2.3 Aspectos Metodológicos

2.3.1 Dados

Para a realização deste ensaio, utilizou-se os microdados da Prova Brasil de 2017, disponíveis no portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). A Prova Brasil é uma das três modalidades que compõe o Sistema de Avaliação da Educação Básica (SAEB), implementado pelo Governo Federal em 1990. Juntamente com a Prova Brasil, fazem parte do SAEB a Avaliação Nacional da Educação Básica (ANEB) e a Avaliação Nacional da Alfabetização (ANA)

Os instrumentos utilizados na Aneb (Avaliação Nacional da Educação Básica) e na Anresc (Prova Brasil) são os testes (provas) e os questionários. Neste ensaio, foi escolhida uma amostra proveniente dos microdados da prova brasil 2017. Os dados refere-se aos alunos do 9º ano³ do ensino fundamental, da escola, do professor e do diretor. Em seguida foi realizado um merge do banco de dados dos alunos, escola, professor e diretor para obtermos um único banco. Realizado o filtro, isto é, eliminando os missing (dados faltantes) e as escolas que foram máscaradas⁴ o conjunto de dados contém uma amostra de 850.434 alunos distribuídos entre as 17.294 escolas públicas brasileiras.

Posto isso, faremos agora uma breve descrição das variáveis utilizadas nesse ensaio antes de prosseguirmos para a apresentação dos resultados. A variável dependente ou variável resposta obtida do questionário é a proficiência em matemática que varia de 0 à 500. A variável de tratamento utilizada nesse trabalho é “Gosto” que assume valor 1 caso o aluno goste de estudar matemática e 0 caso contrário e as variáveis de controle são as apresentadas no Quadro 1.

A variável Inse (indicador de nível sócio econômico) possibilita, de modo geral, situar o público atendido pela escola em um estrato social, apontando o padrão de vida referente a cada um de seus níveis ou estratos. Esse indicador é calculado a partir do nível de escolaridade dos pais e da posse de bens e contratação de serviços pela família dos alunos. No âmbito da escola, o Inse (indicador de nível sócio econômico) foi criado a partir da média aritmética simples da medida de nível socioeconômico dos alunos. Para

³ Dado que a maioria das perguntas no banco de dados do SAEB são alto declaradas, dentre os alunos do ensino fundamental, foi escolhido o nono ano de modo que esses alunos tem mais maturidade em relação aos outros

⁴ Escolas que participaram e tiveram a nota em Matemática publicada pelo SAEB 2017.

o Inse médio das escolas foram consideradas somente as que tinham 10 ou mais alunos pertencentes ao universo de referência. Para melhor descrever o nível socio econômico das escolas, foram criados, a partir da análise de cluster por método hierárquico, seis grupos, de forma que o Grupo 1 congrega as escolas com Inse médio mais baixo e o Grupo 6, com mais alto.

Vale ressaltar que todas as variáveis apresentadas no Quadro 1 são binárias, com exceção da variável de resultado. Porém no dicionário de variáveis do microdados do SAEB 2017 algumas delas são politômicas⁵ como por exemplo a escolaridade da mãe (7 categorias), tarefa (4 categorias), experiência do professor (7 categorias). No entanto essas variáveis politômicas foram transformadas em dummies para serem utilizadas no modelo econométrico apresentado na metodologia adotado nesse ensaio.

⁵ Variáveis de múltiplas categorias

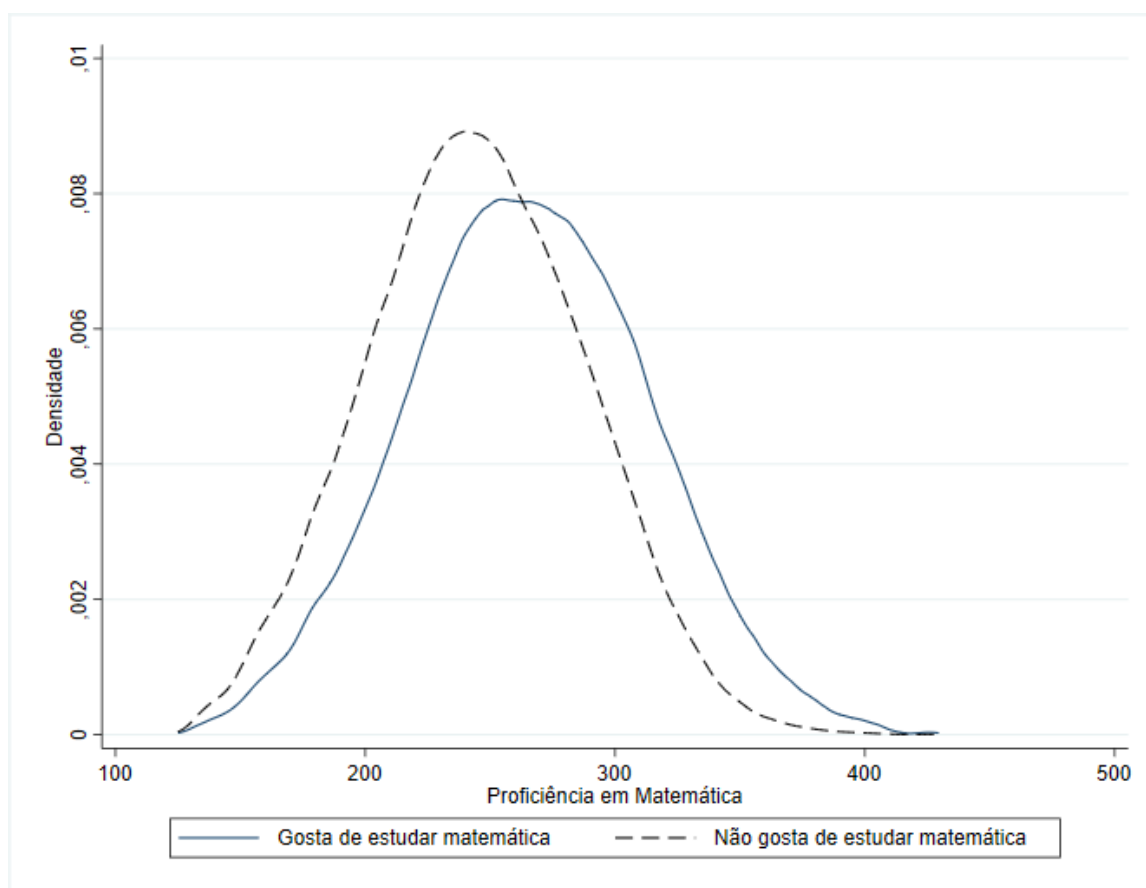
Quadro 1 – Descrição e metodologia das variáveis utilizadas nas análises econométricas

Variáveis	Descrição das Variáveis
Proficiência	Proficiência do aluno em Matemática transformada na escala única do SAEB.
Gosto	1 se o aluno gosta de estudar matemática e 0 caso contrário.
Sexo	1 se o sexo do aluno for Masculino e 0 caso contrário.
Raça	1 se a raça do aluno for branca e 0 caso contrário.
Trabalha	1 se o aluno trabalha fora de casa (recebendo ou não salário) e 0 caso contrário.
Incentivo	1 se os pais do aluno o incentivam a estudar e 0 caso contrário.
Reunião(pais)	1 se os pais dos alunos participam de reuniões escolares e 0 caso contrário.
Reprovado	1 se o aluno foi reprovado e 0 caso contrário.
Escolaridade(mãeI)	1 se a mãe completou o ensino médio e 0 caso contrário.
Escolaridade(mãeII)	1 se a mãe completou a faculdade e 0 caso contrário.
Tarefa	1 se o aluno faz o dever de casa de Matemática sempre ou quase sempre e 0 caso contrário.
Gosto(português)	1 se o aluno gosta de estudar português e 0 caso contrário.
Municipal	1 se a escola é da rede municipal e 0 se é da rede estadual.
Nível(socioeconômicoI)	1 se o indicador de nível socioeconômico é baixo e 0 caso contrário
Nível(socioeconômicoII)	1 se o indicador de nível socioeconômico é médio baixo e 0 caso contrário
Nível(socioeconômicoIII)	1 se o indicador de nível socioeconômico é médio e 0 caso contrário
Nível(socioeconômicoIV)	1 se o indicador de nível socioeconômico é médio alto e 0 caso contrário
Biblioteca	1 se existe biblioteca na escola e 0 caso contrário
Reforço	1 se a escola oferece reforço e 0 caso contrário.
Experiência(diretor)	1 se o tempo no cargo na direção da escola está entre 3 à 5 anos e 0 caso contrário.
Experiência(professor)	1 se o professor possui mais de 20 anos na profissão e 0 caso contrário.
DE(professor)	1 se o professor leciona somente em uma escola e 0 caso contrário.

Fonte: Elaboração própria a partir do dicionário de variáveis dos microdados do SAEB 2017

A Figura 1 mostra o comparativo entre os alunos que gostam de estudar matemática e os que não gostam em relação a distribuição das notas de matemática. Percebe que a distribuição para os alunos que apreciam essa disciplina é mais assimétrica em relação a distribuição das notas dos alunos que não gostam, isto é, à priori nota-se uma diferença entre as médias para os dois grupos.

Figura 1 – Distribuição das notas de matemática por gostar de estudar



Fonte – Elaboração própria a partir dos microdados do SAEB 2017

A Tabela 1 mostra uma relação entre os alunos que gostam de estudar matemática para os que não gostam por região brasileira, ou seja, no banco de dados do SAEB 2017 foi constatado que a menor proporção de alunos que gostam de matemática para os que não gostam está na região sudeste (1,71), enquanto que a maior proporção é encontrada na região Norte (2,42). Para as demais regiões Nordeste, Sul e Centro-Oeste essas proporções são respectivamente, 2,11, 2 e 1,88. Como podemos perceber as regiões menos desenvolvidas economicamente apresentam uma maior proporção de alunos que gostam de estudar matemática, isto é, há uma maior heterogeneidade. Por outro lado nas regiões mais desenvolvidas economicamente essa heterogeneidade diminui.

Tabela 1 – Proporção de alunos que gostam e não gostam de matemática por regiões

Gosto por Matemática	Região					Total
	Norte	Nordeste	Sudeste	Sul	Centro-Oeste	
Não	25997	79914	120415	34634	28505	289465
Sim	62992	169018	205978	69394	53587	560969
Total	88989	248932	326393	104028	82092	850434

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

Sobre a variável dependente “ Proficiência em Matemática⁶”, a média dos alunos corresponde à 258,452. A proficiência mínima encontrada foi de 124,988 e a máxima 429,636. Por outro lado, a média dos alunos que gostam de estudar matemática corresponde à 265,459, enquanto que a média dos alunos que afirmaram que não gostam de estudar essa disciplina é 244,873. Realizado um teste estatístico de comparação entre as médias, observou-se que a diferença entre as médias é estatisticamente significativo ao nível de significância de 1%.

De 1.820.104 alunos que participaram da prova Brasil em 2017 na pesquisa, houve perda de informação (missing) em algumas variáveis, restando, portanto, 850.434 observações. De acordo com a Tabela 2 as principais características dos alunos presentes na amostra são: sexo masculino (47,6%) e (52,4%) feminino. Os alunos são em sua maioria não brancos (70,5%), seguido pelos brancos (29,5%). Em relação ao trabalho: (87,2%) dos alunos responderam que não trabalha, enquanto (12,8%) afirmaram que trabalha. Perguntado sobre se os pais incentivam o aluno a estudar, (1,4%) disseram que não e (98,6%) responderam que sim.

Do total de alunos presentes na amostra, (34%) não gostam de estudar matemática, enquanto que (66%)⁷ afirmaram ter apreço por essa disciplina. Ainda vale salientar que a reprovação representa (26,4%) do universo. Quanto a escolaridade dos pais, (27,5%) das mães completaram o ensino médio, mas não completou a faculdade e (11,3%) completou a faculdade. No que tange ao indicador de nível socioeconômico podemos perceber uma proporção maior para o grupo que apresenta a categoria indicador de nível socioeconômico médio representada no Quadro 1 pela variável “Nível(socioeconômicoIII)”. No tocante a tarefa, (57,9%) dos alunos realizam a tarefa de Matemática e (42,1%) não o fazem. Em relação a rede de ensino, (55,3%) do alunado presente na prova Brasil estudam na rede estadual de ensino e os outros (44,7%) são oriundos da rede municipal de ensino.

No que se refere ao reforço escolar, (89,2%) dos alunos afirmaram que a escola oferece reforço e (10,8%) disseram que não há reforço na escola onde ele estuda. Do universo dos professores, foi constatado que (28,3%) deles, possui mais de 20 anos de experiência

⁶ De acordo com o SAEB 2017 a nota varia de 0 à 500

⁷ Vale salientar que essa proporção é oriunda da pergunta “ Você gosta de estudar matemática?” presente no questionário do aluno que tem opções “Sim” ou “Não” como respostas à essa pergunta.

na profissão. No tocante a dedicação exclusiva, (52,5%) dos professores ensinam apenas em uma escola.

Cabe ainda salientar que a média ponderada das escolas públicas para esse ano específico, no caso 2017 foi de 256,072 pontos. Comparando as escolas públicas municipais com as estaduais em relação à essa média, nota-se que as escolas municipais possui a média inferior à das estaduais, sendo a diferença estatisticamente significativa ao nível de significância de 1%. Em relação as variáveis binárias apresentadas na Tabela 2, a média representa a proporção de observações que foram codificadas com valor um. Essas proporções foram apresentadas nos parágrafos anteriores, assim como a proporção complementar representada pelas categorias que receberam valor zero.

Tabela 2 – Estatísticas descritivas das variáveis utilizadas nos modelos

Variável	Média	Desvio padrão	Mínimo	Máximo
Proficiência	258,452	47,602	124,988	429,636
Gosto	0,660	0,474	0	1
Sexo	0,476	0,500	0	1
Raça	0,295	0,456	0	1
Trabalha	0,128	0,334	0	1
Incentivo	0,986	0,117	0	1
Reunião(pais)	0,574	0,494	0	1
Reprovado	0,264	0,441	0	1
Escolaridade(mãeI)	0,275	0,447	0	1
Escolaridade(mãeII)	0,113	0,316	0	1
Tarefa	0,579	0,494	0	1
Gosto(português)	0,759	0,428	0	1
Municipal	0,447	0,497	0	1
Nível(socioeconômicoI)	0,024	0,155	0	1
Nível(socioeconômicoII)	0,131	0,338	0	1
Nível(socioeconômicoIII)	0,436	0,496	0	1
Nível(socioeconômicoIV)	0,321	0,467	0	1
Biblioteca	0,838	0,369	0	1
Reforço(escolar)	0,892	0,310	0	1
Experiência(diretor)	0,236	0,425	0	1
Experiência(professor)	0,283	0,450	0	1
DE(professor)	0,525	0,499	0	1

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

2.3.2 Estratégia Empírica

Esta seção apresenta a estratégia empírica utilizada no estudo. Os métodos que serão empregados são o Propensity Score Matching (PSM) e o método de Lewbel (2012). A ideia é averiguar se há diferenças entre o desempenho em Matemática dos alunos que gostam de estudar matemática e os que não gostam, isto é, mensurar o efeito de gostar de estudar sobre a proficiência em Matemática.

2.3.2.1 Pareamento por Escore de Propensão

Na literatura de avaliação de impacto, os dados geralmente não vêm de estudos randomizados, mas de estudos observacionais (não randomizados). Em um trabalho seminal, [Rosenbaum e Rubin \(1983\)](#) propuseram o pareamento do escore de propensão como um método para reduzir o viés na estimativa dos efeitos do tratamento com conjuntos de dados observacionais. Esses métodos se tornaram cada vez mais populares em ensaios médicos e na avaliação de intervenções de política econômica.

Visto que em estudos observacionais a atribuição de sujeitos aos grupos de tratamento e controle não é aleatória, a estimativa do efeito do tratamento pode ser enviesada pela existência de fatores de confundimento. O pareamento do escore de propensão é uma maneira de “corrigir” a estimativa dos efeitos do tratamento controlando a existência desses fatores de confundimento com base na ideia de que o viés é reduzido quando a comparação dos resultados é realizada usando indivíduos tratados e de controle que são tão semelhantes quanto possível . ([BECKER, 2009](#))

Estimar o efeito médio de um tratamento binário ou política em um resultado escalar é um objetivo básico de muitos estudos empíricos em economia. Se a atribuição ao tratamento for exógena ou não confundimento (ou seja, independente de resultados potenciais condicionais a covariáveis ou variáveis de pré-tratamento, uma suposição também conhecida como seleção de observáveis), o efeito médio do tratamento pode ser estimado por pareamento ou pela média dentro da subpopulação diferenças de tratamento e controle de médias. Se houver muitas covariáveis, tais estratégias podem não ser desejáveis ou mesmo viáveis. ([HIRANO; IMBENS; RIDDER, 2003](#))

Uma abordagem alternativa para resolver o problema é baseada no escore de propensão. De acordo com [Rosenbaum e Rubin \(1983\)](#) o escore de propensão é definido como sendo a probabilidade condicional de atribuição a um tratamento particular dado um vetor de covariáveis observadas. Ainda segundo o autor, a teoria de amostra grande e pequena mostra que o ajuste para o escore de propensão escalar é suficiente para remover o viés devido a todas as covariáveis observadas. O interesse do presente estudo é estimar o efeito causal da variável Gosto (gosta de estudar Matemática) no desempenho, usando os microdados da Prova Brasil de 2017. Para tanto, como especificação tradicional para

estimar tal efeito, tem-se o modelo descrito abaixo:

$$Y = \alpha + \beta D + \gamma X + \epsilon \quad (2.1)$$

onde Y é variável de resultado, isto é, desempenho em Matemática na Prova Brasil 2017, D é a variável indicadora de tratamento para Gosto (gosta de estudar Matemática), X é um vetor de variáveis de controle que afetam o desempenho em Matemática, e ϵ é o termo de erro. Mensurar β de forma consistente requer que o termo de erro não esteja correlacionado com a variável Gosto, isto é, $\text{COV}(D; \epsilon) = 0$. Se essa condição não for satisfeita as estimativas por OLS serão viesadas e inconsistentes (ANGRIST; PISCHKE, 2008).

O efeito causal da variável “Gosto” no desempenho em Matemática poderia ser mensurado, a partir das diferenças observadas entre a nota obtida pelo indivíduo que gosta de estudar Matemática e a nota auferida por este mesmo indivíduo na situação em que ele não gostasse de estudar Matemática, perceba que não é possível observar essa pessoa nesses dois momentos. O problema fundamental da avaliação de impacto surge porque apenas um dos resultados potenciais é observado para cada indivíduo. O resultado não observado é chamado de resultado contrafactual. Portanto, estimar o efeito do tratamento individual não é possível e deve-se concentrar nos efeitos médios do tratamento.

De outro modo podemos afirmar que quando a participação no tratamento não é por atribuição aleatória, mas depende estocasticamente de um vetor de variáveis observáveis x , como em dados observacionais ou quando o tratamento é direcionado a alguma população definida por algumas características observáveis (como idade, sexo ou status socioeconômico), então o conceito de escores de propensão é útil. Denotando por $p(x)$ o escore de propensão ou a probabilidade condicional de participação no tratamento dado o vetor de covariáveis x , teremos:

$$p(x) = \text{Pr}[D = 1|X = x] \quad (2.2)$$

em que $D = \{0, 1\}$ é a uma variável binária indicadora de tratamento e x é o vetor de características observáveis. A medida do escore de propensão pode ser calculada com base nos dados (D_i, x_i) usando métodos paramétricos ou semiparamétricos (por exemplo, uma regressão logit ou probit).

Rosenbaum e Rubin (1983) mostram que se a exposição ao tratamento é aleatória dentro das células definidas por x , também é aleatória dentro das células definidas pelos valores da variável unidimensional $p(x)$. Como resultado, dada uma população de unidades denotadas por i , se o escore de propensão $p(x_i)$ for conhecido, então o efeito médio do tratamento no tratado (ATT) pode ser estimado da seguinte forma:

$$\beta_{ATT} = E[y_{1i} - y_{0i} | D_i = 1] \quad (2.3)$$

$$= E[E\{y_{1i} - y_{0i} | D_i = 1, p(x_i)\}] \quad (2.4)$$

$$= E[E\{y_{1i} | D_i = 1, p(x_i)\} - E\{y_{0i} | D_i = 0, p(x_i)\} | D_i = 1] \quad (2.5)$$

Segundo [Khandker, Koolwal e Samad \(2009\)](#) o pareamento por escore de propensão (PSM) constrói um grupo de comparação ou controle estatístico baseado em um modelo de probabilidade de participação no tratamento, usando características observadas. Os participantes são então comparados com base nessa probabilidade, ou escore de propensão, aos não participantes. O efeito médio do tratamento do programa é então calculado como a diferença média nos resultados entre esses dois grupos. A validade do PSM depende de duas condições: (a) independência condicional (ou seja, que fatores não observados não afetam a participação) e (b) presença de um suporte comum considerável ou sobreposição nos escores de propensão nas amostras de participantes e não participantes. Uma suposição que desempenha um papel importante na avaliação do tratamento é a condição de balanceamento, que afirma que

$$D \perp x | p(x). \quad (2.6)$$

Isso pode ser expresso alternativamente, dizendo que, para indivíduos com o mesmo escore de propensão, a atribuição ao tratamento é aleatória e deve parecer idêntica em termos de seu vetor x . A condição de equilíbrio é uma hipótese testável.

Um resultado útil sobre a independência condicional dado $p(x)$ devido a [Rosenbaum e Rubin \(1983\)](#) afirma que

$$y_0, y_1 \perp D | x \implies y_0, y_1 \perp D | p(x), \quad (2.7)$$

onde, y_0 representa o resultado para o grupo controle (alunos que não gostam de estudar matemática) e y_1 o resultado para o grupo tratado (alunos que gostam de estudar matemática). Isso implica que a suposição de independência condicional dado x implica independência condicional dado $p(x)$, isto é, independência de y_0 , y_1 e D dado $p(x)$.

Para obter o β na equação 3.1 estima-se o escore de propensão utilizando um modelo logit ou probit, fazendo o matching (pareamento) do grupo de tratamento com o grupo de controle a partir dos métodos do vizinho mais próximo (cada unidade de tratamento é pareada à unidade de controle com o escore de propensão mais próximo), e estimação por kernel (cada tratado é pareado com vários controles com peso proporcional à distância entre o tratado e cada controle).

2.3.2.2 Método de Lewbel (2012)- Identificação com Heterocedasticidade

Esta seção descreve o estimador de dois estágios proposto por [Lewbel \(2012\)](#), que explora a heterocedasticidade do primeiro estágio da regressão para gerar instrumentos internamente que possibilitem a identificação, quando não existem restrições de exclusão disponíveis, dada que muitas vezes não existem variáveis instrumentais externas que possam ser usadas diretamente nos modelos estimados na subseção anterior.

Seguindo [Lewbel \(2012\)](#), o sistema triangular de equações como apresentado nas equações 2.8 - 2.10 mostra que quando as correlações dos erros são causadas devido a fatores não observados, a identificação se dá tendo regressores não correlacionados com o produto dos erros heterocedásticos.

Para o propósito desse artigo, supondo U ser uma variável omitida que possivelmente afeta a variável “gosta de estudar matemática” G (variável endógena de tratamento) tanto quanto afeta a proficiência, e denotando v_1 e V_2 como erros idiossincráticos. Portanto, é possível identificar o efeito causal de gostar de estudar matemática, denotado por β , através do Método dos Momentos Generalizados (GMM) ou de Mínimos Quadrados Modificados em Dois Estágios (2SLS).

$$Y = X'\beta_1 + \gamma G + \epsilon_1 \quad (2.8)$$

$$G = X'\beta_2 + \epsilon_2 \quad (2.9)$$

$$E[X\epsilon_1] = 0, \quad E[X\epsilon_2] = 0, \quad \text{Cov}[Z, \epsilon_1\epsilon_2] = 0, \quad (2.10)$$

onde $\epsilon_1 = \alpha_1 U + V_1$, $\epsilon_2 = \alpha_2 U + V_2$ e $Z \subseteq X$ ⁸. A equação (2.10) apresenta as condições necessárias para a identificação e estimação, juntamente com alguma heterocedasticidade em ϵ_i . Assim, a classe de modelos que satisfazem os pressupostos subjacentes ao método de Lewbel são aqueles que as correlações dos erros nas equações em cross-section são por conta da presença de um fator comum não observado.

A equação auxiliar ou regressão do primeiro estágio pode ser usada para fornecer os elementos necessários para o método de Lewbel. Em sua versão mais simples, instrumentos podem ser construídos a partir dos resíduos das regressões auxiliares multiplicado por cada uma das variáveis exógenas centrada na média. Esses resíduos tem covariância zero com cada um dos regressores usados para construí-lo, o que implica que a média dos instrumentos gerados são zero.

Por outro lado, o produto desses resíduos com os elementos dos regressores centrados na média não será zero, se houver evidência considerável de heterocedasticidade de escala com relação às variáveis explicativas.⁹ Assim, quanto maior o grau da heterocedasticidade

⁸ Z é um subconjunto de X , e com isso, nenhuma informação de fora do modelo especificado é necessária.

⁹ A heterocedasticidade pode ser analisada por testes tal como o de Breusch-Pagan

no processo de erro, maior será a correlação dos instrumentos gerados com as variáveis endógenas incluídas (que são os regressandos das regressões auxiliares).

Em suma, quando se quer corroborar os resultados encontrados por IV tradicionais ou na falta de uma variável instrumental externa válida, como é o caso de muitas aplicações empíricas, o método de Lewbel é uma alternativa interessante para estimar o efeito causal de interesse, tal como no presente caso de um aluno gostar de estudar matemática sobre o desempenho.

É importante salientar que as mesmas variáveis usadas no efeito de tratamento médio são usadas nessa etapa, isto é, tratamento binário igual a 1 para os alunos que gostam de estudar Matemática e 0 caso contrário, uma variável de resultado no caso o desempenho e um conjunto de covariáveis de controle que foi descrito na subseção de dados.

2.3.2.3 Análise de sensibilidade

O pareamento tornou-se um método popular para estimar os efeitos médios do tratamento. Esse método é baseado na suposição de independência condicional ou não confundimento, que afirma que o pesquisador deve observar todas as variáveis que influenciam simultaneamente a decisão de participação e as variáveis de resultado. Essa é uma premissa de identificação forte e deve ser justificada caso a caso.¹⁰ Portanto, verificar a sensibilidade dos resultados estimados com relação a desvios dessa premissa de identificação torna-se um tópico cada vez mais importante na literatura de avaliação de impacto.

Sabemos que os fatores que não são observados na estimação podem enviesar os resultados do efeito do tratamento sobre os tratados. Sendo assim, a análise de limites mensura o potencial impacto do viés de seleção que surge devido as variáveis não observadas. Nesse estudo foi utilizado o método conhecido como limites de Rosebaum¹¹. A finalidade do teste é estimar qual deve ser a influência de uma eventual variável omitida sobre o viés de seleção existente na probabilidade de participação no tratamento, e, que eventualmente possa prejudicar as conclusões a respeito dos efeitos causais. Essa análise de sensibilidade pode ser utilizada para testar a robustez dos resultados à presença de uma covariável omitida. Nesse estudo, o teste visa avaliar o impacto de uma eventual variável omitida sobre as notas dos alunos.

Os limites de Rosenbaum podem ser expressos pela desigualdade:

$$\frac{1}{e^\gamma} \leq \frac{P_i(1 - P_j)}{P_j(1 - P_i)} \leq e^\gamma \quad (2.11)$$

¹⁰ Caliendo e Kopeinig (2008) fornecem um levantamento das etapas necessárias ao implementar (propensity score) métodos de pareamento

¹¹ Maiores detalhes ver Rosenbaum (2002)

em que i e j são dois indivíduos com características observáveis dentro de uma distribuição logística e γ expressa a medida do grau de ruptura de uma amostra livre do viés de seleção. Assim, ambos os indivíduos pareados têm a mesma probabilidade de participar apenas se $e^\gamma = 1$. Caso contrário, os indivíduos que parecem ser semelhantes (em termos de covariáveis) podem diferir em suas chances de receber o tratamento. Nesse sentido, e^γ é uma medida do grau de desvio de um estudo que está livre de viés oculto.

2.4 Resultados

2.4.1 Pareamento pelo escore de propensão

Esta subseção apresenta o efeito de gostar de estudar matemática sobre a proficiência com base no pareamento pelo escore de propensão (propensity score matching) utilizando o estimador Kernel e o vizinho mais próximo. Antes disso, a Tabela 3 apresenta as variáveis que explicam as diferenças de probabilidade de gostar de estudar matemática e de não gostar, através de um escore de propensão (propensity score) logístico. É importante relembrar que a variável binária Gosto (gosta de estudar matemática) foi construída de modo a assumir o valor 1 para os alunos que responderam que gostam de estudar matemática e 0 caso contrário. Dessa forma coeficientes positivos indicam que a variável de tratamento impacta para aumentar a proficiência enquanto que coeficientes negativos impactam em reduzir a proficiência.

O modelo logit apresentado na Tabela 3 abaixo foi utilizado como modelo preditivo, isto é, para obtermos o valor da probabilidade de o aluno gostar de estudar matemática dado um conjunto de covariáveis observadas. No entanto, vejamos a interpretação do efeito marginal de algumas variáveis que apresentaram coeficientes significativos. Se o aluno for do sexo masculino sua probabilidade de gostar de estudar matemática aumenta em média 8,6 pontos percentuais (p.p). Esse fato pode ser explicado pelo distanciamento das meninas em relação a matemática. No tocante ao incentivo a estudar por parte dos pais, a probabilidade aumenta em média 6,73 p.p, mostrando que o incentivo é fundamental para o desenvolvimento cognitivo dos alunos. Caso o aluno tenha sido reprovado, a probabilidade dele gostar da disciplina em análise, cai em média 6,73 p.p. Tal resultado segue em linha com [Jacomini \(2009\)](#), [Correa \(2013\)](#), e pode ser explicado pela experiência da reprovação ao menos uma vez. Já o fato da mãe ter completado o ensino médio, a chance do aluno gostar de matemática aumenta em 0,2 p.p. Por outro lado, fazer a tarefa de casa aumenta em média a probabilidade do aluno gostar de estudar matemática em aproximadamente 32 p.p. Estudar em escola pública municipal eleva em média a probabilidade do “gosto por estudar matemática” em aproximadamente 2 p.p.

Ainda de acordo com a Tabela 3, podemos perceber que professores que trabalham à mais de 20 anos, faz a probabilidade do aluno gostar da disciplina cair em média em 0,4 p.p. É importante enfatizar neste ponto que a variável “Experiência do professor” foi significativa, porém com sinal negativo. Tal evidência corrobora com o encontrado em Biondi e Felício (2007), Krieg (2006), onde um período longo de permanência de um professor numa escola pode desencorajá-lo de suas atividades pedagógicas deixando lacunas no aprendizado do aluno. No que tange a participação dos pais em reuniões escolares percebemos que o coeficiente da regressão logística foi significativo e corrobora com o encontrado em (BIONDI; VASCONCELLOS; MENEZES-FILHO, 2009). Por outro lado, gostar de estudar português aumenta as chances do gosto por matemática em 7,83 p.p. Isso mostra que a leitura é fundamental, principalmente quando os alunos resolvem tarefas matemáticas com suporte escrito como aponta (MENEZES, 2011).

Estudos recentes geralmente relatam que a experiência do professor tem um efeito positivo nos resultados dos testes dos alunos (ROCKOFF, 2004; JEPSEN, 2005; RIVKIN; HANUSHEK; KAIN, 2005; CLOTFELTER; LADD; VIGDOR, 2006; GOLDBABER; ANTHONY, 2007). O efeito positivo também parece ser de natureza não linear, conforme demonstrado por melhorias substanciais na habilidade de ensino durante os primeiros 3 – 5 anos na sala de aula, com os efeitos geralmente diminuindo por volta do quinto ano (RIVKIN; HANUSHEK; KAIN, 2005).

Por outro lado, a variável “Reprovado” foi significativa com sinal esperado pela literatura, isto é, o fato do aluno ser reprovado reduz o desempenho médio em 6,73 p.p. Outra covariável que foi estatisticamente significativa foi o indicador de nível socioeconômico que apresentou quatro categorias: baixo, médio baixo, médio e médio alto. A Tabela 3 mostra que os efeitos marginais foram significativos e decrescentes, isto é, todos os indicadores de nível socioeconômico apresentados tiveram o sinal esperado pela literatura, ou seja, o fato do aluno pertencer a qualquer um desses níveis, aumenta a chance dele gostar de Matemática. Esses resultados corroboram com a literatura encontrada em Palermo, Silva e Novellino (2014), Soares (2004), Lobo, Cassuce e Cirino (2017), que usaram essa variável como controle para explicar o desempenho escolar.

Para obtermos o efeito médio de tratamento nos tratados, utilizou-se o modelo logit para obtenção do escore de propensão. Com o intuito de fazermos comparações, foi estimado mais de um modelo utilizando o pareamento pelo escore de propensão pelo algoritmo Kernel e do vizinho mais próximo. Como podemos observar, a Tabela 4 apresenta os resultados para a proficiência em matemática, o ATC (efeito médio de tratamento nos não tratados) apresentou diferença média de 15,616 pontos na nota do desempenho entre o grupo de tratamento (alunos que gostam de estudar matemática) e o grupo de controle (alunos que não gostam de estudar matemática) pelo método de pareamento Kernel. Já em relação ao método de pareamento pelo vizinho mais próximo o ATC obtido corresponde à

Tabela 3 – Resultados do modelo logit que estima a probabilidade do aluno gostar de estudar Matemática

Gosto	(1)	(2)
	coeficiente	Efeito marginal
Sexo	0,438*** (0,005)	0,0946*** (0,001)
Incentivo	0,230*** (0,020)	0,0519*** (0,005)
Reunião(pais)	0,099*** (0,005)	0,0216*** (0,001)
Reprovado	-0,298*** (0,005)	-0,0663*** (0,001)
Tarefa	1,447*** (0,005)	0,317*** (0,001)
Gosto(português)	0,350*** (0,006)	0,0783*** (0,001)
Municipal	0,090*** (0,005)	0,020*** (0,001)
Nível(socioeconômicoI)	0,473*** (0,019)	0,0936*** (0,003)
Nível(socioeconômicoII)	0,304*** (0,011)	0,063*** (0,002)
Nível(socioeconômicoIII)	0,151*** (0,009)	0,0327*** (0,002)
Nível(socioeconômicoIV)	0,065*** (0,009)	0,0141*** (0,002)
Experiência(prof)	-0,019*** (0,005)	-0,004*** (0,001)
Observações	863.491	863.491

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

Nota: Erro padrão entre parênteses. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

15,913. Por outro lado, o efeito médio de tratamento sobre os tratados (ATT'S), indicou resultado favorável ao grupo de alunos que gostam de estudar matemática, com diferença média de 17,261 pelo estimador Kernel e 17,267 pelo estimador do vizinho mais próximo (NN). Isso mostra que em termos de robustez o método do vizinho mais próximo é mais adequado para explicar o efeito de gostar de estudar matemática sobre o desempenho.

Diante disso, percebemos que gostar de estudar matemática tem um efeito positivo no desempenho, o que corrobora com os resultados encontrados em (LAROS; MARCIANO; ANDRADE, 2010; RODRIGUES, 2017; MOKGWATHI; GRAHAM; FRASER, 2019). Por exemplo, Laros, Marciano e Andrade (2010) utilizaram dados do SAEB 2001 e uma modelagem multinível para explicar os fatores que afetam o desempenho e chegaram num efeito positivo e significativo de aproximadamente 8 pontos para a variável explicativa

“Gosta de estudar matemática”. Vale ressaltar que a variável “Gosta de estudar matemática” nos dados do SAEB 2001 é a mesma dos dados do SAEB 2017, ou seja, binária¹².

Já Mokgwathi, Graham e Fraser (2019) utilizaram também modelagem multinível e verificaram que os alunos que gostam de aprender matemática e valorizam a matemática se sobressaem em relação aos que não gostam e não valorizam. Esse estudo investigou a relação entre as percepções e atitudes dos professores sul-africanos do 9º ano e dos alunos em relação ao seu desempenho em matemática. O estudo utilizou uma análise secundária dos dados matemáticos do TIMSS 2015. Outro estudo que corrobora com o resultado destacado nesse ensaio foi o achado em Barbosa e Fernandes (2000a) que encontraram uma correlação positiva entre gostar de estudar matemática e a proficiência utilizando dados do SAEB 1997.

Tabela 4 – Estimação do efeito médio de tratamento

Proficiência em matemática			
	OLS	Kernel	NN
ATE	16,326*** (0,105)	16,700*** (0,127)	16,806*** (0,111)
ATT		17,261*** (0,161)	17,267*** (0,123)
ATC		15,616*** (0,164)	15,913*** (0,107)
NATE		20,586*** (0,099)	20,586*** (0,108)
Observações	850.434		

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

Notas: Desvios-padrão robustos a heterocedasticidade entre parênteses, com bootstrap de 1000 replicações. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

ATE: efeito médio de tratamento, ATT: efeito médio de tratamento nos tratados

ATC: efeito médio de tratamento nos não tratados e NATE: diferença média bruta.

Por outro lado, a Tabela 5 mostra o efeito médio de tratamento nos tratados para o grupo de reprovados como teste de robustez, visto que os dados apresentados mostra uma proporção grande para os alunos que afirmaram que gostam de estudar matemática. Diante disso, foi feita uma estimação para o grupo de reprovados para identificar o efeito do gosto nesse grupo. Assim, de acordo com a Tabela 5 o efeito médio de tratamento nos tratados corresponde à 13,068. Em comparação com o ATT de toda à amostra, notamos uma redução de 4,133 no efeito médio. Isso mostra que mesmo o aluno sendo reprovado, o gosto por matemática aumenta a proficiência.

¹² Assume valor 1 se o aluno gosta de estudar Matemática e 0 caso contrário

Tabela 5 – Estimação do efeito médio de tratamento dos reprovados

Proficiência em matemática			
	OLS	Kernel	NN
ATE	12,698*** (0,189)	12,853*** (0,197)	12,869*** (0,190)
ATT		13,068*** (0,230)	12,997*** (0,215)
ATC		12,531*** (0,228)	12,678*** (0,194)
NATE		13,621*** (0,182)	13,621*** (0,178)
Observações	224.189		

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

Notas: Desvios-padrão robustos a heterocedasticidade entre parênteses, com bootstrap de 1000 replicações. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

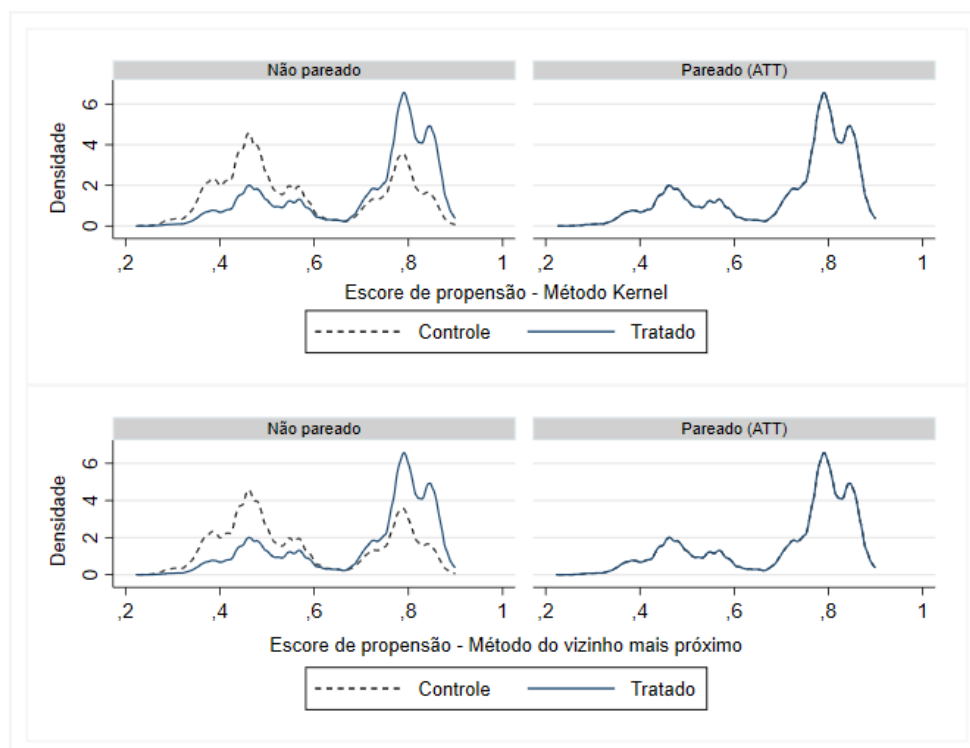
ATE: efeito médio de tratamento, ATT: efeito médio de tratamento nos tratados

ATC: efeito médio de tratamento nos não tratados e NATE: diferença média bruta.

Diferentes estratégias podem ser usadas para identificar a condição de sobreposição entre os grupos. Nesse modelo foi utilizada uma análise visual dos gráficos com a distribuição de densidade do escore de propensão para os dois grupos e para a amostra dos reprovados, com a sobreposição das distribuições.

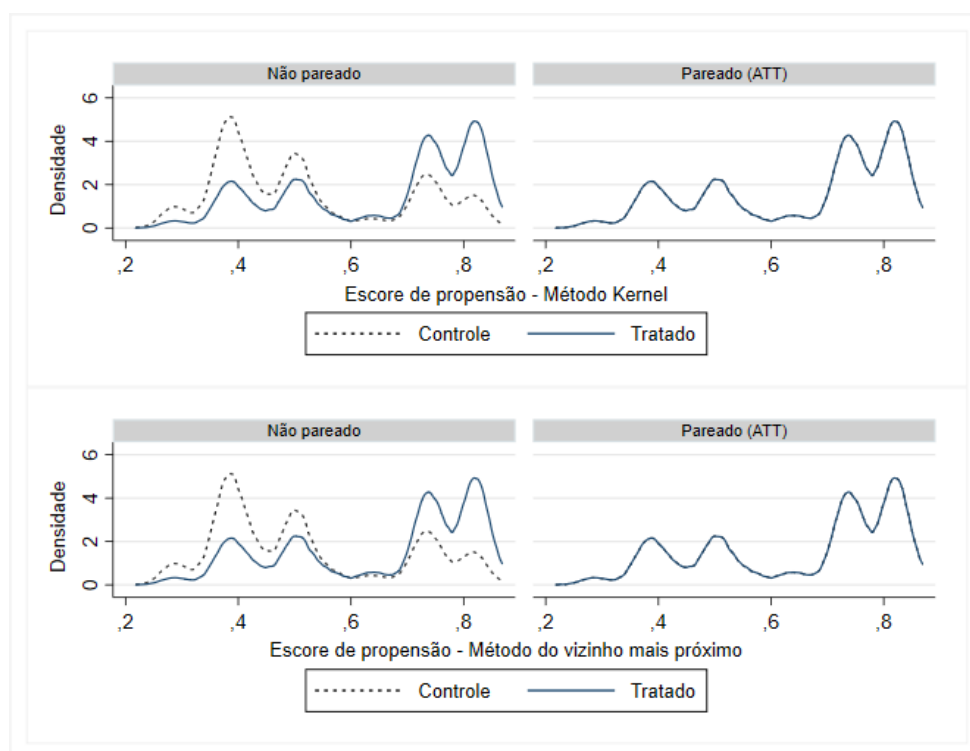
Os resultados referentes as proficiências em matemática mostram que, no gráfico representado pelas Figuras 2 e 3 é possível observar a diferença na distribuição dos escores de propensão entre os alunos no grupo de tratamento e controle. Antes do pareamento, é possível perceber uma alternância de valores estimados, isto é, unidades do grupo de controle apresentavam valores estimados abaixo numa certa área do grupo de tratamento e valores superiores em outra região tanto pelo pareamento do escore de propensão pelo algoritmo Kernel como pelo algoritmo do vizinho mais próximo. Após o pareamento podemos observar a sobreposição desses escores de propensão dentro da região de suporte comum.

Figura 2 – Função densidade do escore de propensão



Fonte – Elaboração própria a partir dos microdados do SAEB 2017

Figura 3 – Função densidade do escore de propensão para os reprovados



Fonte – Elaboração própria a partir dos microdados do SAEB 2017

Em termos metodológicos, os resultados reportados na Tabela 4 partem da hipótese de independência condicional, cuja presença de variáveis omitidas na especificação do modelo de pareamento pode resultar em um viés na estimação dos efeitos médios sobre o grupo tratado. Dessa forma, é feita a análise de sensibilidade através dos limites de Rosenbaum, com o objetivo de verificar até que ponto as variáveis não observáveis podem ser prejudiciais para os resultados encontrados.

A análise de sensibilidade foi utilizada nesse estudo para testar a robustez dos resultados à presença de viés devido a uma covariável omitida. De acordo com [Cechin et al. \(2015\)](#), a análise de sensibilidade proposta por Rosenbaum não exclui a possibilidade de fatores não observáveis existirem, quando o viés de seleção deve estar presente para que seja possível aceitar a hipótese nula de ausência de efeito de tratamento. A Tabela 6 apresenta os resultados da análise de sensibilidade de Rosenbaum sobre o desempenho em Matemática, mostrando o nível crítico do Γ igual a 2. Isso significa que se a presença de variáveis não observadas levarem a uma diferença na odds ratio de receber tratamento entre os alunos dos grupos de tratamento e controle por um fator de 2 pode-se questionar o efeito de gostar de estudar matemática sobre a proficiência.

Tabela 6 – Limites de Rosenbaum a análise de sensibilidade de tendências ocultas devido a variáveis omitidas

Gamma	sig+	sig-	t-hat+	t-hat-	CI+	CI-
1	0	0	17,492	17,492	17,334	17,651
1,25	0	0	11,648	23,338	11,489	23,497
1,5	0	0	6,895	28,093	6,735	28,253
1,75	0	0	2,906	32,085	2,744	32,247
2	1	0	-0,521	35,515	-0,684	35,679
2,25	1	0	-3,516	38,514	-3,682	38,681
2,5	1	0	-6,172	41,174	-6,340	41,342
2,75	1	0	-8,553	43,558	-8,723	43,729
3	1	0	-10,707	45,716	-10,879	45,889

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

De outro modo, dizemos que os resultados do teste de análise de sensibilidade de Rosenbaum indicam ausência de viés em possíveis características não observáveis para um grau de liberdade de até 2, o que implica, portanto, que o pareamento é robusto. Assim, com base nesse estudo, verifica-se que existe uma associação entre gostar de estudar Matemática e o desempenho do aluno do 9º ano que realizou a Prova Brasil 2017.

2.4.2 Análise de Robustez: Método de Lewbel

Os resultados apresentados na Tabela 7 mostram um impacto positivo do aluno gostar de estudar matemática no desempenho. Esta subseção começa pela apresentação dos resultados do efeito de tratamento médio estimado por Generalized Method of Moments (GMM) a partir do Método de Lewbel.

Especificamente, a causa reversa pode ser um problema, pois é provável que os alunos que afirmaram que gostam de estudar matemática não alcancem bons resultados e o desempenho também possa impactar no gosto por matemática. Assim, embora esperemos que o gosto por matemática influencie o desempenho, é provável que o desempenho também possa influenciar o gosto por matemática. Além disso, existem fatores não observados que podem influenciar o desempenho, que também são responsáveis pelo gosto de estudar matemática, causando também endogeneidade. Sem controlar a endogeneidade, as estimativas relatadas são susceptíveis de ser estatisticamente tendenciosas e enganosas.

Consequentemente, adotamos a abordagem de Lewbel (2012) de mínimos quadrados de dois estágios (2SLS) para lidar com a endogeneidade e para garantir que nossos resultados sejam robustos para problemas de endogeneidade. O uso da abordagem Lewbel 2SLS em vez da abordagem 2SLS tradicional se deve à indisponibilidade de instrumentos externos que se enquadram na restrição de exclusão em nosso conjunto de dados. A abordagem Lewbel 2SLS é eficaz em tais casos e tem sido amplamente utilizada na literatura.

As estimativas corroboram os resultados anteriores de um efeito positivo de gostar de estudar Matemática sobre o desempenho. No entanto, o método de Lewbel evidencia uma possível superestimação do efeito médio quando não considerados os instrumentos para a variável de tratamento “Gosta de estudar Matemática”. Tal fato pode ser visto pela redução dessa variável, que passa de 17,267 (efeito de tratamento médio no tratado) para 9,822 (efeito médio de tratamento no tratado pelo método de Lewbel). Esse resultado foi superestimado em virtude da falta de instrumentos adequados que expliquem o fato do aluno gostar de estudar matemática. Sendo assim, o modelo estimado pelo escore de propensão é mais válido.

A Tabela 7 apresenta o teste de subidentificação do modelo, o qual se rejeita a hipótese nula indicando que o modelo é identificado. Por outro lado, o teste de Cragg-Donald rejeita a hipótese nula de instrumentos fracos, sugerindo que os instrumentos internos gerados são relevantes para o tratamento¹³

¹³ Os resultados do teste podem ser solicitados junto aos autores.

Tabela 7 – ATT estimado - Método Lewbel (2012)

Variável dependente: Proficiência em Matemática	
	Total
ATT _{GMM}	9,822*** (0,437)
Observações	850434
Teste de Kleibergen-Paap LM	43,1***

Fonte: Elaboração própria a partir dos microdados do SAEB 2017

Notas: Desvios-padrão robustos a heterocedasticidade entre parênteses, com bootstrap de 1000 replicações. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

2.5 Considerações Finais

Este estudo teve como objetivo investigar a relação entre gostar de estudar matemática e o desempenho obtido na prova de matemática. O estudo foi realizado, utilizando os dados da Prova Brasil do ano de 2017, com alunos matriculados no 9º ano do ensino fundamental das escolas públicas brasileiras. Para essa pesquisa o método utilizado foi o Propensity Score Matching (Escore de Propensão) utilizando o algoritmo Kernel e do vizinho mais próximo, bem como o método de Lewbel.

Como podemos observar, o efeito de gostar de estudar matemática sobre o desempenho em matemática, controlado por variáveis do aluno, da escola e do professor, teve um diferencial positivo esperado tanto pelo pareamento por kernel quanto pelo método do vizinho mais próximo. No primeiro método o ATT alcançado foi de aproximadamente 16,701 pontos, enquanto no segundo esse diferencial foi de aproximadamente 16,783 pontos. Por outro lado, sabendo que a variável “Gosta de estudar Matemática” é endógena, aplicou-se a metodologia de Lewbel como análise de robustez e o efeito esperado dessa variável sobre o desempenho foi positivo e estatisticamente significativo.

Em vista disso, notamos que gostar de estudar matemática pode trazer benefícios para o aluno como desenvolver o raciocínio lógico e aplicá-lo em diferentes situações do cotidiano. Outro benefício está relacionado ao ganho de conhecimentos matemáticos principalmente no aprimoramento de cálculo, útil para quem deseja ingressar em determinadas profissões que exigem do profissional domínio na área de matemática. Como exemplo podemos destacar algumas profissões como: engenharias, informática, estatística, medicina e etc.

Com base na suposição de independência condicional, o pareamento tornou-se uma abordagem popular para estimar os efeitos médios do tratamento. Assim, verificar a sensibilidade dos resultados estimados com relação a desvios dessa premissa de identificação tem se tornado um tópico cada vez mais importante na literatura de avaliação de impacto.

Isto é, se houver variáveis não observadas que afetam a atribuição ao tratamento e a variável de resultado simultaneamente, um viés oculto pode surgir para o qual os estimadores de pareamento pelo escore de propensão não são robustos. Sendo assim, os resultados do teste de análise de sensibilidade de Rosenbaum indicam ausência de viés em possíveis características não observáveis para um grau de liberdade de até 1,5, o que implica, portanto, que o pareamento é robusto.

Diante disso, podemos perceber a importância de implantar políticas públicas que incentivem o aluno a criar uma cultura voltada a preferência ao estudo. Como por exemplo, gostar de estudar matemática aumenta a proficiência média e desenvolve habilidades para novos horizontes do conhecimento. Já existe um consenso entre os pesquisadores sobre a idéia de que quando os alunos demonstram uma competência percebida ajustada as suas características pessoais e sociais estarão também melhor adaptadas, independentemente do nível de realização ou da capacidade individual.

Por fim, esse estudo diferente de outros, explica o quanto uma variável afetiva impacta o desempenho em matemática ou em outras disciplinas. Com isso, sabendo que o desempenho escolar é o produto de uma função de produção educacional e que essa função é de várias variáveis, fica cada vez mais evidente a importância de testar um insumo diferente dos já utilizados na economia da educação.

3 O EFEITO DE COMPUTADORES COM INTERNET NAS ESCOLAS NA PROFICIÊNCIA DOS ESTUDANTES: UMA ANÁLISE MULTINÍVEL

3.1 Introdução

O uso do computador e da internet é considerado cada vez mais importante nas sociedades modernas. De acordo com [Lindbeck e Snower \(2000\)](#) os computadores e a internet introduziram mudanças significantes nos processos de trabalho e na organização das estruturas corporativas nos últimos anos. Para [Bulman e Fairlie \(2016\)](#) o uso de tecnologia é onipresente no sistema educacional da maioria dos países desenvolvidos. Por exemplo, essencialmente todas as salas de aula de instrução nas escolas públicas dos EUA têm computadores com acesso à Internet (Departamento de Educação dos EUA, 2012).

No Brasil os dados do Censo Escolar 2020, divulgados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep), mostram a situação das escolas da educação básica brasileira no que diz respeito à disponibilidade de equipamentos de tecnologia da informação e comunicação. Assim, esse estudo tem relevância particular por mostrar a importância do uso de computadores com internet nas escolas no desempenho dos alunos, em contexto anterior à pandemia de COVID-19. Com a necessidade de implantar políticas públicas educacionais baseadas em tecnologia para diminuir os impactos negativos nos processos de ensino e aprendizagem, o uso de computadores e da internet passaram a ser mais visto na literatura da economia da educação. Os dados do SAEB 2019 e do Censo 2019 mostra que, no 9º ano do ensino fundamental, 50,3% das escolas públicas brasileiras possui internet com computador. Já na rede municipal, esse percentual passa a ser de aproximadamente 36% quando comparado com as escolas que dispõem de computadores com internet para acesso aos alunos.

O acesso a computadores nas escolas pode melhorar os resultados dos alunos de várias maneiras. O software de computador tem o potencial de fornecer instruções individualizadas, o que normalmente é difícil de obter em instrução em grupo [Koedinger et al. \(1997\)](#). Da mesma forma, o conteúdo da instrução pode ser individualizado de acordo com os pontos fortes e fracos do aluno. Como os alunos podem usar programas de ensino sem a supervisão direta de um professor, as TICs e a instrução auxiliada por computador prometem aumentar a quantidade geral de instrução que os alunos recebem, enquanto

permitem que os pais e professores monitorem o progresso dos alunos (CUBAN, 1993; ROUSE; KRUEGER, 2004; BARROW; MARKMAN; ROUSE, 2009).

Para Bulman e Fairlie (2016) o uso do computador nas escolas pode beneficiar os alunos de duas maneiras: por meio da aquisição de conhecimentos de informática úteis para o mercado de trabalho; e por meio da aquisição de habilidades básicas, como matemática, leitura e escrita. Assim, a literatura econômica fornece diferentes justificativas para enfocar a eficácia dos computadores como uma ferramenta pedagógica para a aquisição de habilidades básicas.

Além do mais a Internet representa um recurso potencialmente valioso para descobrir informações sobre uma ampla gama de tópicos educacionais para reduzir os custos de coordenação de projetos em grupo. Para Cuban (1993) computadores, internet, software e outras tecnologias, devido à sua natureza interativa, podem envolver os alunos em um ambiente de aprendizagem diferente dos métodos tradicionais.

Por outro lado, na busca de melhoria da qualidade do sistema educacional, aliada a um melhor desempenho dos alunos em exames de proficiência padronizados, o governo federal criou políticas públicas para integrar a informática ao processo de aprendizagem. De acordo com Linden, Banerjee e Duflo (2003) o aumento do acesso dos alunos às tecnologias de informação e comunicação (TICs), como o computador e acesso à internet, compensariam outros fatores problemáticos – a exemplo do baixo preparo dos professores –, tornando a escola mais atrativa.

Diante disso, de acordo com Hanushek (1979) medir o desempenho educacional e compreender seus determinantes são importantes para a formulação de políticas com respeito a questões diversas, como responsabilidade do professor, sistemas de financiamento educacional e integração escolar. Um dos problemas que o pesquisador encontra para mensurar o efeito do desempenho escolar sobre as variáveis relacionadas as características dos alunos, escolas e turmas é que a estrutura de certos dados estão inseridos em uma estrutura de clusters. Um tipo de modelagem econométrica utilizada na economia para se trabalhar com dados dispostos em clusters são os modelos lineares multiníveis.

De acordo com Snijders e Bosker (2011) a análise multinível é uma metodologia de análise de dados com padrões complexos de variabilidade, com foco em fontes aninhadas de tal variabilidade por exemplo alunos em turmas, turmas em escolas, funcionários em empresas, etc. Na análise de tais dados, geralmente é esclarecedor levar em conta o fato de que cada nível de aninhamento está associado a uma variabilidade que tem uma interpretação distinta. Há variabilidade, por exemplo, entre alunos, mas também entre classes, e pode-se tirar conclusões incorretas se nenhuma distinção for feita entre essas diferentes fontes de variabilidade. A análise multinível é uma abordagem para a análise de tais dados, incluindo as técnicas estatísticas, bem como a metodologia para seu uso. Assim, considerando que os dados do SAEB 2019 segue essa estrutura percebe-se a importância

do uso dessa metodologia.

A Organização para a Cooperação e Desenvolvimento Econômico (OCDE) divulgou em seu relatório referente ao Programa Internacional de Avaliação de Alunos (PISA) de 2018 que o Brasil oferece em média, apenas um computador para cada quatro estudantes nas instituições de ensino. Esse dado é bastante preocupante, quando se compara o Brasil com países em que a educação é de suma importância para o desenvolvimento econômico.

A presença de computadores com internet seja na escola ou no domicílio, pode apresentar um efeito positivo no desempenho escolar do aluno, porém isso depende de vários fatores associados as características dos alunos. Por exemplo, [Young \(2006\)](#) examinou a condição de uso da Internet e do capital social, bem como seus efeitos no desempenho acadêmico de adolescentes de 361 alunos do ensino médio e constataram que a Internet expande seu alcance para a vida escolar dos adolescentes.

Por outro lado, [Spiezia \(2011\)](#) avaliou se o uso das TIC (Tecnologia da informação e comunicação) tinha um impacto no desempenho dos alunos usando dados do (PISA) 2006. Depois de controlar as características observáveis dos alunos e a auto-seleção, foi encontrado um efeito positivo e significativo da frequência de uso do computador nas pontuações de ciências. Na maioria dos países, entretanto, esse efeito parece maior quando o computador é usado em casa em vez de ser na escola. Esse achado questiona a eficácia das políticas educacionais voltadas para o incentivo ao uso do computador na escola como ferramenta de aprendizagem.

Já [Wenglinsky \(1998\)](#) e [Angrist e Lavy \(2002\)](#) relatam efeitos negativos do uso do computador nas escolas em algumas medidas de desempenho dos alunos. O estudo quase experimental de [Angrist e Lavy \(2002\)](#) descobriram que a introdução da instrução auxiliada por computador nas escolas israelenses exerceu um efeito negativo estatisticamente significativo no desempenho em matemática dos alunos da quarta série e um efeito negativo, mas estatisticamente insignificante no desempenho dos alunos em outras disciplinas e séries superiores.

[Lai et al. \(2015\)](#) em um artigo usaram um experimento de campo randomizado para explorar os efeitos da aprendizagem assistida por computador (CAL) nos resultados acadêmicos e não acadêmicos dos alunos em escolas de migrantes em Pequim. Os resultados mostraram que um programa corretivo de (CAL) realizado fora do horário escolar regular melhora as notas de matemática padronizadas do aluno em 0,15 desvios-padrão e a maior parte do efeito do programa ocorreu dentro de 2 meses após o início do programa. Os alunos com pais mais instruídos se beneficiaram mais com o programa. Além disso, o (CAL) também aumentou significativamente o interesse dos alunos em aprender.

De acordo com [Bulman e Fairlie \(2016\)](#) nos EUA, 3,9 bilhões de dólares é gasto em tecnologia por escolas, famílias e legisladores na esperança de melhorar os resultados

educacionais. Esses autores exploram a literatura teórica e empírica sobre os impactos da tecnologia nos resultados educacionais. A literatura concentra-se em dois contextos primários nos quais a tecnologia pode ser usada para fins educacionais: (i) uso em sala de aula nas escolas e (ii) uso doméstico pelos alunos. Teoricamente, o investimento em tecnologia da informação e comunicação (TIC) e o uso de instrução auxiliada por computador (CAI) pelas escolas e o uso de computadores em casa têm implicações ambíguas para o desempenho educacional: despesas dedicadas à tecnologia necessariamente compensam as entradas que podem ser mais ou menos eficientes e o tempo alocado para o uso da tecnologia pode substituir o ensino tradicional em sala de aula e as atividades educacionais em casa.

No entanto, muitas das evidências na literatura escolar são baseadas em intervenções que fornecem financiamento suplementar para tecnologia ou tempo de aula adicional e, portanto, favorecem a descoberta de efeitos positivos. Por outro lado, estudos de (TIC) e (CAI) em escolas produzem evidências mistas com um padrão de resultados nulos. Exceções notáveis a esse padrão ocorrem em estudos de países em desenvolvimento e intervenções (CAI) que visam matemática em vez de linguagem.

Tendo em vista que a internet é um serviço muito útil no cotidiano das pessoas e tem relação importante no sistema educacional, esse estudo justifica-se devido aos dados apresentados no SAEB 2019, onde 45,9% das escolas públicas brasileiras possuem computadores com internet para uso dos alunos gerando uma preocupação para os formuladores de políticas públicas, no sentido da importância do avanço tecnológico como recurso pedagógico para os alunos.

Os problemas relacionados para o tema apresentado nesse estudo vão desde a carência desse recurso ausente em muitas escolas públicas brasileiras, o tema pouco explorado na economia da educação até as informações importantes não presentes no banco de dados. Dessa forma, tendo em vista a importância de se considerar os efeitos do computador com internet nas escolas sobre o desempenho dos alunos, constatada por vários estudos empíricos, o objetivo do presente artigo é verificar como a presença do computador com internet nas escolas afeta o desempenho em matemática dos alunos do 9º ano das escolas públicas de Ensino Fundamental. Sendo assim, parte-se da hipótese de que o efeito do computador com internet nas escolas aumenta o desempenho em matemática. A estratégia empírica consiste em uma análise multinível com três níveis, sendo o nível 1 referente aos alunos, o nível 2 associado às escolas e o nível 3 referente às turmas.

Esse capítulo possui quatro seções incluindo esta introdução. Na segunda seção tem-se os aspectos metodológicos. Na terceira seção tem-se os resultados das estimações. Finalmente, na quarta seção apresentam-se as considerações finais.

3.2 Aspectos metodológicos

3.2.1 Dados

Para a realização desta pesquisa utilizou-se os microdados da Prova Brasil de (2019), disponíveis no portal do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP). A Prova Brasil é uma das três modalidades que compõe o SAEB, implementado pelo Governo Federal em 1990. Juntamente com a Prova Brasil, fazem parte do SAEB a Avaliação Nacional da Educação Básica (ANEB) e a Avaliação Nacional da Alfabetização (ANA). Além do mais, utilizou-se os microdados do Censo Escolar 2019 com o intuito de coletar informações da escola que não constam no banco de dados do saeb 2019.

Na edição do SAEB 2019 foram extintas as nomenclaturas ANA, Aneb e Anresc e todas as avaliações externas passaram a ser identificadas como SAEB. Esse novo SAEB inclui, pela primeira vez, a avaliação das instituições que oferecem Educação Infantil, em caráter de estudo piloto, e a avaliação dos estudantes do 1º ano do Ensino Fundamental, além de manter a avaliação dos estudantes do 5º e 9º anos do Ensino Fundamental e 3ª e 4ª séries do Ensino Médio. Outra novidade é a aplicação de testes de Ciências da Natureza e Ciências Humanas aos estudantes do 9º ano do Ensino Fundamental.

No total, os microdados da Prova Brasil (2019) referente as observações dos alunos do 9º ano do ensino fundamental e do Censo 2019 contém informações sobre 1.648.468 estudantes brasileiros. Após a junção do banco de dados do SAEB e do Censo escolar e a exclusão de missing (dados faltantes), a base final dispôs de 622.047 observações distribuídas em 24.837 escolas e 47.162 turmas.

Especificado o modelo adotado neste ensaio, faremos agora uma breve descrição dos dados antes de prosseguirmos para a apresentação dos resultados. A variável dependente ou variável resposta obtida do questionário é a proficiência em matemática que varia de 0 à 500.

Como o desempenho escolar pode ser afetado por características dos alunos, da escola e da turma, faz-se necessário controlar por essas características. Sendo assim, optou-se por reportar no Quadro 2 a descrição das variáveis utilizadas, salientando que a escolha por controlar por determinadas características é feita a partir da literatura acerca dos determinantes do desempenho escolar (HANUSHEK, 1970; BROOMHALL; JOHNSON, 1994; KROHN; O'CONNOR, 2005; ALHAJRAF; ALASFOUR, 2014; EGALITE, 2016; HANUSHEK; WOESSMANN, 2017; DICKE et al., 2018; GALVEZ-NIETO et al., 2020; BOBBA; FRISANCHO, 2020; TORRES et al., 2020).

As variáveis referentes à Raça, Incentivo, Reprovado, Rede Wi fi e escolaridade da mãe são utilizadas como controles para o nível do aluno nos modelos multiníveis. Com

relação às variáveis dos demais grupos, alusivas às condições das escolas e características dos professores e diretores, estas têm como objetivo medir os efeitos da escola, da turma, do background dos professores e diretores sobre o desempenho dos estudantes.

A variável explicativa “Computador(internet)¹” que significa computador com internet na escola para uso do aluno assume valor 1 caso a escola tenha computador com acesso à internet e 0 caso contrário. Essa variável foi obtida por meio de uma interação entre as variáveis binárias “Computador” e “Internet” presentes no banco de dados do Censo 2019 e tem a finalidade de captar o efeito sobre a proficiência em matemática.

Quadro 2 – Descrição e metodologia das variáveis utilizadas nas análises multiníveis

Variáveis	Descrição das Variáveis
Nota	Proficiência do aluno em Matemática transformada na escala única do SAEB, com média=250 e desvio=50 (do SAEB/97).
Computador(internet)	1 se a escola possui computador com internet para uso do aluno e 0 caso contrário
Raça	1 se a raça do aluno for branca e 0 caso contrário.
Incentivo	1 se os pais incentivam o aluno a estudar sempre ou quase sempre e 0 caso contrário.
Reprovado	1 se o aluno foi reprovado e 0 caso contrário.
Rede(Wi fi)	1 se existe rede wi fi no domicílio do aluno e 0 caso contrário.
Escolaridade(mãeI)	1 se a mãe possui Ensino Fundamental completo e 0 caso contrário.
Escolaridade(mãeII)	1 se a mãe possui Ensino Médio completo e 0 caso contrário.
Escolaridade(mãeIII)	1 se a mãe possui Ensino Superior e 0 caso contrário.
Municipal	1 se a escola é da rede municipal e 0 se é da rede estadual.
Biblioteca	1 se na escola existe biblioteca e 0 caso contrário
Sala(leitura)	1 se na escola existe sala de leitura e 0 caso contrário
Internet*Municipal	1 se a escola municipal possui internet e 0 caso contrário
Projeto(CT)	1 se a escola possui projeto de ciência e tecnologia e 0 caso contrário
Projeto(MT)	1 se a escola possui projeto mundo do trabalho e 0 caso contrário
Experiência(diretor)	tempo de experiência como diretor variando de 0 à 35 (anos)
DE(professor)	1 se o professor leciona apenas nessa escola e 0 caso contrário.
Especialização	1 se o professor fez especialização e 0 caso contrário.

Fonte: Elaboração própria a partir dos microdados da Prova Brasil (2019) e do Censo (2019).

O computador com acesso a internet quando usado para fins educacionais é uma ferramenta poderosa para o desenvolvimento do ensino de matemática. O uso adequado da

¹ Computadores com internet para uso do aluno no laboratório de informática da escola

Internet como tecnologias de informação e comunicação pode fornecer muitas oportunidades, incluindo a melhoria das realizações acadêmicas Guan e Subrahmanyam (2009); Erdogdu e Erdogdu (2015). As redes da Internet nos permitem interagir com uma ampla gama de pessoas em todo o mundo. Além disso, programas de aprendizagem baseados na Internet, sistemas de educação médica e intervenções de saúde, como terapia cognitiva, tornaram-se generalizados, gerando resultados positivos (WIMBLE, 2016; SINCLAIR et al., 2016).

A Internet representa o maior acervo de conhecimento humano e tem permitido que muitas comunidades no mundo gerenciem a comunicação e resolvam problemas com mais rapidez. Por outro lado, a internet possibilita que alunos, professores e profissionais da educação se comuniquem, criem conteúdo, serviços, aplicativos e inovações para seus empreendimentos básicos. Com o acesso à Internet, um amplo leque de ideias, recursos de ensino e aprendizagem, técnicas e inovações passam a ser amplamente compartilhados. Almasi, Machumu e Zhu (2017) exploraram o uso da Internet entre alunos do ensino médio e seus efeitos na aprendizagem utilizando modelos mistos². Os resultados revelaram que o uso da Internet para fins educacionais apresentaram um efeito positivo.

Em relação as principais características dos alunos presentes no banco de dados do SAEB 2019 e do Censo 2019 tem-se que 29,50% dos alunos declararam ser brancos, 84,03% afirmaram que são incentivados pelos pais à estudarem. Por outro lado, 75,6% possui rede Wi Fi em seu domicílio o que mostra uma proximidade da tecnologia nas residências dos alunos. Além do mais 23,8% do alunado foi reprovado uma ou duas vezes. Em relação a escolaridade dos pais observou-se que 13,3% dos alunos afirmaram que a mãe possui ensino fundamental completo, 27,4% ensino médio completo e 12% responderam que a mãe possui ensino superior completo.

Quanto as características relacionadas as escolas, percebe-se que 51% delas são municipais sendo que 16,6% possuem internet. Além do mais das escolas públicas analisadas, 57% possui biblioteca, 39% possui sala de leitura e 45,9% possui computador com acesso à internet para uso do aluno. Ainda na amostra utilizada nesse artigo podemos perceber que 39% das escolas públicas ofertam projeto de ciências e tecnologia e 39,7% ofertam projetos voltados para o mercado de trabalho. Em relação ao diretor foi usada uma variável discreta assumindo valores inteiros de 0 à 35 para representar a experiência do diretor em anos. Com respeito à qualificação do professor, podemos observar que 24% dos professores tem especialização. Isto significa que o complementar (76%) está distribuído do seguinte modo: 5,47% possui mestrado, 2,02% doutorado e 68,51% não possui pós-graduação. Neste ensaio, optou-se por não incluir as dummies relativas a qualificação do professor referente ao mestrado e doutorado pelo motivo da amostra desses grupos ser muito pequena em relação ao tamanho da amostra em estudo.

² Outra denominação para modelos multiníveis

De acordo com a Tabela 8 percebe-se que a média referente à experiência do diretor foi de aproximadamente 7 anos. Essa variável de controle é importante, pois estudos como o encontrado em Felício e Fernandes (2005), Oliveira e Waldhelm (2016) mostram que diretores com mais anos de experiência numa única escola apresentam efeitos positivos no desempenho do aluno. Em relação a variável dependente “Proficiência em Matemática”, a média global dos alunos corresponde à 262,090. Por outro lado, considerando que a nota média em Matemática no Saeb 2019 varia no intervalo de 0 e 500, percebe-se que a proficiência mínima encontrada foi de 126,327 e a máxima 397,487. Como isso, podemos observar que a média em matemática alcançada pelos alunos foi superior a média nacional que é de 250 pontos. Por outro lado, a média em matemática dos alunos que estudam em escolas que possui computador com internet para o uso dos alunos foi de 266,643, enquanto a média dos alunos que frequentam escolas que não tem computador com acesso a internet para os discentes foi de 257,473.

Como podemos notar, existe uma diferença entre as médias desses dois grupos. Para sabermos se essa diferença foi estatisticamente significativa, realizou-se um teste de diferença de médias, utilizando um nível de significância de 5%. Realizado o teste t, observou-se que a hipótese nula foi rejeitada, isto é, a diferença observada é estatisticamente significativa ao nível de significância considerado.

Tabela 8 – Estatística descritiva das variáveis usadas nos modelos

Variável	Média	Desvio padrão	Mínimo	Máximo
Nota	262,090	47,872	126,327	397,487
Computador(internet)	0,459	0,498	0	1
Raça	0,295	0,456	0	1
Incentivo	0,840	0,366	0	1
Reprovado	0,238	0,426	0	1
Rede(Wi Fi)	0,756	0,429	0	1
Escolaridade(mãeI)	0,133	0,339	0	1
Escolaridade(mãeII)	0,274	0,446	0	1
Escolaridade(mãeIII)	0,120	0,325	0	1
Municipal	0,510	0,500	0	1
Biblioteca	0,570	0,495	0	1
Sala(leitura)	0,390	0,488	0	1
Internet*Municipal	0,166	0,372	0	1
Projeto(CT)	0,390	0,489	0	1
Projeto(MT)	0,397	0,489	0	1
Experiência(diretor)	6,708	6,161	0	35
DE(professor)	0,550	0,497	0	1
Especialização	0,240	0,427	0	1

Fonte: Elaboração própria a partir do dicionário de variáveis do Saeb (2019) e do Censo (2019).

3.2.2 Estratégia Empírica

A modelagem multinível (MLM, também conhecida como modelagem linear hierárquica, HLM) é uma estrutura metodológica amplamente usada nas ciências sociais para analisar dados com uma estrutura hierárquica, onde unidades inferiores de agregação são “aninhadas” em unidades superiores, incluindo dados longitudinais. Em economia, entretanto, o MLM é usado muito raramente. Em vez disso, os economistas usam técnicas econométricas separadas, incluindo erros padrão robustos de cluster e modelos de efeitos fixos.

Sobre os modelos multiníveis, [Oshchepkov e Shirokanova \(2020\)](#) revisaram a literatura metodológica e contrastou as técnicas econométricas tipicamente usadas em economia com a análise de dados hierárquicos usando MLM. Nessa revisão os autores sugeriram que as técnicas econômicas são geralmente menos convenientes, flexíveis e eficientes em comparação com o MLM. A limitação importante do MLM, no entanto, é sua incapacidade de lidar com o problema da variável omitida no nível mais baixo de dados, enquanto as técnicas econômicas padrão podem ser complementadas por métodos quase-experimentais que mitigam esse problema. É improvável, entretanto, que essa limitação possa explicar e justificar o raro uso do MLM na economia. No geral, eles concluíram que o MLM tem sido ignorado de forma irracional na economia, e aconselham os economistas a aplicar esta estrutura, fornecendo diretrizes de “quando e como” usar a MLM.

Nesse ensaio, foi utilizado a modelagem multinível pelo fato de os dados estarem organizadas em estrutura hierárquica, isto é, alunos inseridos em turmas e turmas dentro das escolas. Nesse ensaio, a estratégia empírica utilizada para a estimação dos modelos multiníveis de três níveis foi baseada em ([RAUDENBUSH; BRYK, 2002](#); [GOLDSTEIN, 2011](#)).

3.2.2.1 Modelo Incondicional ou Nulo

O modelo mais simples de três níveis é totalmente incondicional; isto é, nenhuma variável preditora é especificada em nenhum nível. Esse modelo representa como a variação em uma medida de resultado é alocada nos três níveis diferentes (aluno, turma e escola).

Modelo em nível do aluno: Modelamos o desempenho em Matemática de cada aluno em função da média da turma mais um erro aleatório:

$$Y_{ijk} = \pi_{0jk} + e_{ijk} \quad (3.1)$$

onde:

Y_{ijk} é o desempenho do aluno i turma j e na escola k ;

π_{0jk} é o desempenho médio da turma j na escola k ; e

e_{ijk} é o efeito aleatório associado ao aluno (nível 1), ou seja, é o desvio do escore do aluno ijk em relação à proficiência média da escola. Esse efeito é assumido como normalmente distribuído com média zero e variância σ^2 .

Os índices i , j e k indicam alunos, turmas e escolas onde existem
 $i = 1, 2, \dots, n_{jk}$ alunos na turma j na escola k ;
 $j = 1, 2, \dots, J_k$ turmas na escola k ; e
 $k = 1, 2, \dots, K$ escolas.

Modelo em nível da turma: Neste nível, cada média da turma, π_{0jk} , é resultado da variação aleatória em torno da média da escola. Assim, obtemos:

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \quad (3.2)$$

onde:

β_{00k} é o desempenho médio na escola k ;

r_{0jk} é o efeito aleatório associado a turma (nível 2), ou seja, é o desvio do escore da escola jk em relação à proficiência média da escola. Esse efeito é suposto como normalmente distribuído com média zero e variância τ_π . Dentro de cada escola a variabilidade entre as escolas é suposta como igual.

Modelo de nível escolar: O modelo de nível 3 representa a variabilidade entre as escolas. Vemos as médias escolares, β_{00k} , variando aleatoriamente em torno de uma grande média:

$$\beta_{00k} = \gamma_{000} + u_{00k} \quad (3.3)$$

onde:

γ_{000} é a média global;

u_{00k} é o efeito aleatório associado à escola (nível 3), ou seja, é o desvio do escore da escola k em relação à média geral da proficiência. Esse efeito é suposto como normalmente distribuído com média zero e variância τ_β .

Substituindo (2.3) e (2.2) em (2.1), obtemos uma única equação para o modelo incondicional dado por:

$$Y_{ijk} = \gamma_{000} + u_{00k} + r_{0jk} + e_{ijk} \quad (3.4)$$

Uma maneira útil de interpretar a magnitude relativa dos componentes de variação, é calcular os coeficientes de partição de variação (VPCs). As estatísticas da VPC mostram

a proporção da variação da resposta que se encontra em cada nível da hierarquia do modelo nulo³. A variação total da resposta para o aluno i é definida como

$$\begin{aligned}\text{var}(Y_{ijk}) &= \text{var}(\gamma_{000} + u_{00k} + r_{0jk} + e_{ijk}) \\ &= \text{var}(\gamma_{000}) + \text{var}(u_{00k}) + \text{var}(r_{0jk}) + \text{var}(e_{ijk}) \\ &= 0 + \tau_\beta + \tau_\pi + \sigma^2 \\ &= \tau_\beta + \tau_\pi + \sigma^2\end{aligned}$$

Assim, a variação total é simplesmente a soma dos três componentes de variação separados. Essa variação total é constante. Observe, no entanto, que isso geralmente não será o caso em modelos mais complexos, como aqueles que incluem coeficientes aleatórios em que a variação total será uma função de todas as covariáveis inseridas no modelo com coeficientes aleatórios.

A VPC no nível da escola é calculada como a razão entre a variação da escola e a variação total, isto é,

$$VPC_u = \frac{\tau_\beta}{\tau_\beta + \tau_\pi + \sigma^2}$$

A VPC no nível da turma é calculada como a razão da variação da turma para a variação total, isto é,

$$VPC_r = \frac{\tau_\pi}{\tau_\beta + \tau_\pi + \sigma^2}$$

Finalmente, a VPC no nível do aluno é calculada como a razão entre a variação do aluno e a variação total

$$VPC_e = \frac{\sigma^2}{\tau_\beta + \tau_\pi + \sigma^2}$$

Uma segunda maneira de interpretar a magnitude relativa dos componentes de variância é calcular os coeficientes de correlação intraclasse (ICCs). As estatísticas do ICC

³ Nos modelos condicionais, as VPCs são baseadas nas respostas residuais e não nas observadas. Portanto, as VPCs baseadas em modelos condicionais medem a proporção da variação do resultado inexplicada pelas variáveis preditoras que se encontram em cada nível da hierarquia do modelo

medem o grau esperado de similaridade (ou homogeneidade) entre as respostas em um determinado cluster (por exemplo, escola)⁴.

De um modo geral, a correlação entre dois alunos i e i' é escrita como

$$\text{corr}(Y_{ijk}, Y_{i'j'k'}) = \frac{\text{cov}(Y_{ijk}, Y_{i'j'k'})}{\sqrt{\text{var}(Y_{ijk})}\sqrt{\text{var}(Y_{i'j'k'})}}$$

onde

$$\begin{aligned} \text{cov}(Y_{ijk}, Y_{i'j'k'}) &= \text{cov}(\gamma_{000} + u_{00k} + r_{0jk} + e_{ijk}, \gamma_{000} + u_{00k'} + r_{0j'k'} + e_{i'j'k'}) \\ &= \text{cov}(u_{00k}, u_{00k'}) + \text{cov}(r_{0jk}, r_{0j'k'}) + \text{cov}(e_{ijk}, e_{i'j'k'}) \end{aligned}$$

por outro lado, já sabemos que:

$$\text{var}(Y_{ijk}) = \text{var}(Y_{i'j'k'}) = \tau_\beta + \tau_\pi + \sigma^2$$

Com isso, a força da correlação depende da extensão em que os alunos i e i' pertencem às mesmas unidades de nível superior. Os alunos que compartilham as mesmas escolas, por exemplo, serão mais parecidos e, portanto, mais correlacionados do que os alunos que frequentam escolas diferentes. Quanto maiores as diferenças entre as escolas, mais alunos semelhantes aparecerão dentro de suas escolas. A Tabela ?? apresenta as covariâncias e correlações implícitas no modelo para os quatro pares de alunos distinguidos pelo modelo.

A Tabela 9 mostra que as covariâncias e correlações são funções dos três componentes de variância. Os componentes de variância são, por definição, zero ou positivo e, portanto, as covariâncias também serão zero ou positivas, enquanto as correlações estarão no intervalo de zero a um.

O par 1 fornece a correlação entre um aluno e ele mesmo. Essa correlação é igual a um.

O par 2 fornece a correlação entre dois alunos que frequentam a mesma turma (e, portanto, a mesma escola). Essa correlação é chamada de ICC da turma.

O par 3 fornece a correlação entre dois estudantes que são ensinados em diferentes turmas, mas que frequentam a mesma escola. Essa correlação é chamada de ICC da escola. Observe que o ICC da escola coincide com o VPC da escola. No entanto, essa equivalência não será

⁴ Nos modelos condicionais, os ICCs são baseados nas respostas residuais, e não nas observadas. Os ICCs nos modelos condicionais, portanto, medem a similaridade nos resultados ajustados para as variáveis preditoras; isto é, a semelhança em resultados inexplicáveis. Os ICCs baseados em modelos condicionais são algumas vezes chamadas de ICCs ajustados

Tabela 9 – Modelo de covariâncias e correlações implícitas para os quatro diferentes pares possíveis de alunos

Par	Alunos	Turmas	Escolas	Covariância	Correlação
1	$i = i'$	$j = j'$	$k = k'$	$\tau_\beta + \tau_\pi + \sigma^2$	1
2	$i \neq i'$	$j = j'$	$k = k'$	$\tau_\beta + \tau_\pi$	$\frac{\tau_\beta + \tau_\pi}{\tau_\beta + \tau_\pi + \sigma^2}$
3	$i \neq i'$	$j \neq j'$	$k = k'$	τ_β	$\frac{\tau_\beta}{\tau_\beta + \tau_\pi + \sigma^2}$
4	$i \neq i'$	$j \neq j'$	$k \neq k'$	0	0

Fonte: Elaboração própria

válida em modelos mais complexos, como aqueles que incluem coeficientes aleatórios.

O par 4 fornece a correlação entre dois alunos que frequentam escolas diferentes (e, portanto, diferentes turmas). Esses dois estudantes não compartilham fontes comuns de influência e, portanto, são considerados independentes; eles têm uma correlação esperada igual a zero.

3.2.2.2 Modelo Condicional

O modelo totalmente incondicional, equações (3.1), (3.2) e (3.3), permitem estimar a variabilidade associada aos três níveis - alunos, turmas e escolas. Presumivelmente, parte da variabilidade em cada nível pode ser explicada pelas variáveis medidas em cada nível. Ou seja, as características do histórico do aluno, da turma e da escola podem ser usadas como preditoras. Além disso, algumas das relações nos níveis de classe e escola podem variar aleatoriamente entre essas unidades. Da mesma forma, os coeficientes de regressão podem variar aleatoriamente no nível da escola. Os interceptos escolares normalmente seriam aleatórios; as inclinações de regressão também podem diferir entre as escolas.

Modelo geral de nível 1. Em cada turma, modelamos o desempenho do aluno em função dos preditores no nível do aluno, além de um erro aleatório no nível do aluno:

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}a_{1ijk} + \pi_{2jk}a_{2ijk} + \dots + \pi_{pjk}a_{pijk} + e_{ijk} \quad (3.5)$$

onde,

Y_{ijk} é o desempenho do aluno i na turma j e na escola k ;

π_{0jk} é a média do desempenho da turma j da escola k ;

a_{pijk} são as $p = 1, 2, \dots, P$ características que predizem o desempenho do aluno;

π_{pjk} são os coeficientes correspondentes às características do aluno e indicam a direção e a força da associação entre cada característica, a_p , e os resultados da turma jk ;

e_{ijk} é o efeito aleatório associado ao aluno e representa o desvio do escore do aluno ijk em relação ao escore predito no modelo do nível do aluno. Os efeitos residuais são supostos normalmente distribuídos com média zero e variância σ^2 .

Modelo geral de nível 2: cada coeficiente de regressão do modelo do nível do aluno (incluindo o intercepto) pode ser visto como um efeito fixo, ou seja, sem variação aleatória. Essas possibilidades conduzem à formulação do modelo para variação entre as turmas dentro das escolas. Para cada efeito da escola, π_{pjk} ,

$$\pi_{pjk} = \beta_{00k} + \sum_{q=1}^{Q_p} \beta_{pqk} X_{qjk} + r_{pjk} \quad (3.6)$$

onde,

β_{00k} é o intercepto para a escola k do efeito π_{pjk} no modelo das turmas;

X_{qjk} é a característica da turma utilizada como preditora do efeito da turma π_{pjk} (observe que cada π_p tem um único conjunto de preditores do segundo nível X_{qjk} , $q = 1, 2, \dots, Q_p$);

β_{pqk} é o coeficiente correspondente que representa a direção e a força da associação entre as características da turma X_{qjk} e π_{pjk} ;

r_{pjk} é um efeito aleatório de nível 2 que representa o desvio do coeficiente de nível 1 da turma $jk's$, π_{pjk} , de seu valor previsto com base no modelo no nível da turma.

Observe que existem $P + 1$ equações no modelo de nível 2 uma para cada um dos coeficientes de nível 1. Presume-se que os efeitos aleatórios nessas equações estejam correlacionados. Formalmente, assumimos que o conjunto de r_{pjk} é multivariado normalmente distribuído, cada um com uma média 0, alguma variação τ_{pp} e alguma covariância entre os elementos r_{pjk} e $r_{p'jk}$ de $\tau_{pp'}$. Coletamos essas variâncias e covariâncias em uma matriz chamada T_{pi} , cuja dimensionalidade depende do número de coeficientes de nível 1 especificados como aleatórios. Por exemplo, se um efeito da turma, π_{pjk} , for especificado como fixo, nenhum preditor de nível 2 seria incluído na Equação (3.6) para esse efeito e o r_{pjk} correspondente seria definido como zero. Se o π_{pjk} for especificado como variação aleatória, as variáveis X aparecerão, mas o π_{pjk} ainda será zero.

Modelo geral de nível 3: Um processo de modelagem semelhante é replicado no nível da escola. Cada "resultado" de nível 3 (ou seja, cada coeficiente de β_{pq}) pode ser previsto por alguma característica no nível da escola. Isto é

$$\beta_{pqk} = \gamma_{pq0} + \sum_{s=1}^{S_{pq}} \gamma_{pqs} W_{sk} + u_{pqk} \quad (3.7)$$

onde,

γ_{pq0} é o termo do intercepto do modelo no nível da escola para β_{pqk} ;

W_{sk} é a característica da escola utilizada para prever o efeito da escola, β_{pqk} (observe que cada β_{pq} tem um único conjunto de preditores do terceiro nível W_{sk} , $s = 1, 2, \dots, S_{pq}$);

γ_{pqs} é o coeficiente correspondente ao terceiro nível que representa a direção e a força de associação entre a característica da microrregião W_{sk} e β_{pqk} ;

u_{pqk} é um efeito aleatório de nível 3 que representa o desvio do coeficiente da escola k ,

β_{pqk} , em relação ao seu valor previsto com base no modelo de nível da escola. Observe que para cada escola existem $\sum_{p=0}^P (Q_p + 1)$ equações no modelo de nível 3.

Os resíduos dessas equações são assumidos como multivariados normalmente distribuídos. Supõe-se que cada um tenha uma média zero, alguma variância e covariância entre todos os pares de elementos. As variâncias e covariâncias são coletadas em uma matriz, T_β . A dimensionalidade de T_β depende do número de coeficientes de nível 2 que são especificados como aleatórios. Como era verdade para os coeficientes π no modelo de nível 2, cada um dos coeficientes β (incluindo o intercepto) pode ser visto como fixo, variando aleatoriamente ou não no modelo de nível 3. Se β_{pq} for especificado como fixo ou com variação não aleatória, o correspondente u_{pqk} será assumido como zero. O modelo totalmente incondicional, equações (3.1), (3.2) e (3.3), permitem estimar a variabilidade associada aos três níveis - alunos, turmas e escolas. Presumivelmente, parte da variabilidade em cada nível pode ser explicada pelas variáveis medidas em cada nível. Ou seja, as características do histórico do aluno, da turma e da escola podem ser usadas como preditoras. Além disso, algumas das relações nos níveis de classe e escola podem variar aleatoriamente entre essas unidades. Da mesma forma, os coeficientes de regressão podem variar aleatoriamente no nível da escola. Os interceptos escolares normalmente seriam aleatórios; as inclinações de regressão também podem diferir entre as escolas.

3.3 Resultados e discussões

3.3.1 Modelo Nulo (Modelo 1) e Modelo Ancova (Modelo 2)

Começamos especificando e ajustando um modelo de componentes de variação de três níveis para a proficiência em Matemática dos alunos. Este modelo inclui apenas um intercepto, efeitos aleatórios da escola, da turma e um termo de erro residual no nível do aluno; o modelo não faz ajustes para variáveis preditoras. O modelo simplesmente decompõe a variação total na proficiência dos alunos em componentes de variação separados da escola, da turma e do aluno. O modelo é escrito como

$$\text{Nota}_{ijk} = \pi_{0jk} + e_{ijk} \quad (3.8)$$

$$\pi_{0jk} = \beta_{00k} + r_{0jk} \quad (3.9)$$

$$\beta_{00k} = \gamma_{000} + u_{00k} \quad (3.10)$$

Podemos representar esse modelo acima por uma única equação, simplesmente substituindo (3.10) em (3.9) e em seguida (3.9) em (3.8). Assim, obtemos:

$$\text{Nota}_{ijk} = \gamma_{000} + u_{00k} + r_{0jk} + e_{ijk}, \quad (3.11)$$

onde,

Nota_{ijk} é a proficiência em matemática do aluno i da turma j e na escola k ;

u_{00k} é o efeito aleatório associado à escola (nível 3), ou seja, é o desvio do escore da escola k em relação à média geral da proficiência;

r_{0jk} é o efeito aleatório associado a turma (nível 2), ou seja, é o desvio do escore da escola jk em relação à proficiência média da escola;

e_{ijk} é o efeito aleatório associado ao aluno (nível 1), ou seja, é o desvio do escore do aluno ijk em relação à proficiência média da escola. E além do mais, $u_{00j} \sim N(0, \tau_\beta)$, $r_{0jk} \sim N(0, \tau_\pi)$ e $e_{ijk} \sim N(0, \sigma^2)$. Isso significa que os efeitos da escola, da turma e os erros residuais no nível do aluno são assumidos de forma independente e normal, distribuídos com médias zero e variâncias constantes.

De acordo com a Tabela 10 a estimativa para o intercepto (efeito fixo) mostra que a proficiência média global dos alunos em Matemática é 259,796 pontos, sendo estatisticamente significativo ao nível de significância de $\alpha = 0,1\%$. Pode-se verificar que da variância total de 2304,653; 372,169 é devido a variabilidade entre escolas, 115,755 é devido a variabilidade entre as turmas e 1816,729 é a variabilidade dentro das turmas. Para compararmos o modelo nulo com um modelo de nível único, sem efeitos na escola e sem efeitos na turma foi realizado um teste LR (teste de razão de verossimilhança). A estatística do teste LR obtida foi ($\chi^2_2 = 93120,40$) com um p-valor efetivamente igual a zero, com isso a hipótese nula de que os interceptos aleatórios sejam iguais a zero foi rejeitada ao nível de significância de 0,1%, o que faz com que a estimação de um modelo tradicional de regressão linear seja descartada para os dados agrupados desse ensaio.

Tabela 10 – Resultado da estimação dos modelos nulo e ancova para a proficiência em Matemática

Variáveis explicativas	Modelo 1		Modelo 2	
Efeito fixo	Estimativa	Erro padrão	Estimativa	Erro padrão
Intercepto	259,796***	(0,16)	255,450***	(0,18)
Raça			5,868***	(0,12)
Reprovado			-23,807***	(0,13)
Rede(Wi fi)			4,283***	(0,14)
Escolaridade(mãeI)			7,465***	(0,17)
Escolaridade(mãeII)			11,855***	(0,13)
Escolaridade(mãeIII)			14,506***	(0,18)
Efeito aleatório	Estimativa	Erro padrão	Estimativa	Erro padrão
var(escola)	372,169	(5,67)	294,685	(4,55)
var(turma)	115,755	(2,41)	73,394	(1,85)
var(Residual)	1816,729	(3,36)	1686,478	(3,12)
Deviance	6484975,7		6432363,3	

Fonte: Elaboração própria a partir dos microdados do SAEB 2019 e do Censo 2019

Nota: Erro padrão entre parênteses. * $p < 0.05$, ** $p < 0.01$ e *** $p < 0.001$.

Utilizando as fórmulas obtidas na subsubseção (3.2.2.1) para o cálculo das estatísticas do VPC (coeficiente de partição de variação), observa-se que 16,15% da variação na proficiência em Matemática dos alunos do nono ano das escolas públicas brasileiras está entre as escolas, 5,02% encontra-se dentro das escolas entre as diferentes turmas, enquanto os 78,83% restantes estão dentro das turmas entre os alunos. Portanto, há uma variação substancial dentro das turmas entre os alunos, mas uma variação relativamente pequena de uma escola para a seguinte. No entanto, vemos que mais de três quartos da variação da variável resposta são atribuíveis aos próprios alunos.

Por outro lado, olhando as estatísticas do ICC (coeficiente de correlação intraclasse), vemos que o ICC da escola é de 0,1615 indicando que aproximadamente 16% da variância total da proficiência em Matemática é devido à alteração entre escolas, representando um primeiro indício de existência de variabilidade na proficiência dos alunos provenientes de escolas distintas. Já o ICC da turma é de 0,2117 indicando que aproximadamente 21% da variância total da proficiência em Matemática é devido à alteração entre turmas.

Isso mostra que a correlação entre dois alunos da mesma escola, mas em turmas diferentes é de 0,1615, enquanto a correlação entre dois alunos da mesma escola e da mesma turma é bastante mais alta, com 0,2117. Assim os alunos ensinados na mesma turma têm, portanto, um desempenho 1,3 vezes maior do que os alunos ensinados na turma adjacente.

Depois de realizado a estimação do modelo nulo, foi inserido variáveis de controle

no nível do aluno com o intuito de obtermos um segundo modelo que possivelmente possa ser mais ajustado. Esse modelo é conhecido como ANCOVA e possui a seguinte forma geral

$$\begin{aligned} \text{Nota}_{ijk} = & \gamma_{000} + \gamma_{100}\text{Raça}_{ijk} + \gamma_{200}\text{Reprovado}_{ijk} + \gamma_{300}\text{Rede(wi fi)} + \\ & + \gamma_{400}\text{Escolaridade(mãeI)}_{ijk} + \gamma_{500}\text{Escolaridade(mãeII)}_{ijk} + \\ & + \gamma_{600}\text{Escolaridade(mãeIII)}_{ijk} \end{aligned}$$

onde, $u_{00j} \sim N(0, \tau_\beta)$, $r_{0jk} \sim N(0, \tau_\pi)$ e $e_{ijk} \sim N(0, \sigma^2)$.

Ainda de acordo com a Tabela 10, o modelo ANCOVA⁵ reduziu a variância de segundo e terceiro nível, conforme esperado, bem como o valor do ICC da escola e da turma. Como podemos observar os efeitos fixos apresentados no modelo 2 (modelo ANCOVA) apresentaram sinal esperado e foram estatisticamente significativos, corroborando com a literatura referente a economia da educação. Sabendo que a variável “Reprovado⁶” assume valor 1 caso o aluno tenha sido reprovado e 0 caso contrário, podemos perceber que o fato do aluno ter sido reprovado diminui sua proficiência em $-23,807$ pontos. No que se refere à raça, observou-se que os alunos brancos têm um desempenho superior, em média, de $5,868$ pontos em comparação àqueles que se declararam não brancos. Isso pode estar relacionado à condição econômica, mas, também, a outras características da família do aluno. Esse resultado já é conhecido na literatura empírica e, conforme afirma [Ferrão et al. \(2001\)](#), [Soares, Alves e Mari \(2002\)](#), [Couri \(2010\)](#), os não brancos, além de possuírem menores níveis de proficiência, ainda enfrentam maior atraso escolar e maiores taxas de repetência.

Quanto a escolaridade da mãe, nota-se que os efeitos fixos obtidos foram significativos, mostrando que a escolaridade da mãe afeta positivamente a nota do aluno conforme a literatura da economia da educação. Em comparação aos alunos com mães que possuem o ensino superior completo com mães que não pertencem a essa categoria, o diferencial em média na proficiência em Matemática é aumentada, em $14,506$ pontos. De acordo com [Barros et al. \(2001\)](#), [Magnuson \(2003\)](#), [Menezes-Filho \(2007a\)](#), acredita-se que mães com maior escolaridade pressionem mais as escolas pela melhoria do ensino ou que isso facilite o aprendizado pelo melhor background familiar dos alunos. Já para [Lobo, Cassuce e Cirino \(2017\)](#) a escolaridade dos pais, principalmente a materna, também foi amplamente identificada como um dos principais fatores do desempenho escolar. Isto porque mulheres mais educadas tendem a se casarem com homens igualmente mais educados, proporcionando,

⁵ Modelo que considera apenas variáveis referentes ao nível do aluno

⁶ Maiores detalhes ver o quadro 1

portanto, um aumento nos rendimentos familiares, que tem efeitos positivos diretos no investimento em capital humano dos filhos.

É importante ressaltar que a modelagem multinível usada nesse ensaio segue as etapas apresentadas em [Hox, Moerbeek e Schoot \(2010\)](#), isto é, estima-se um modelo vazio (sem variáveis explicativas) e em seguida estima-se o modelo ANOVA (variáveis explicativas no nível do aluno) que gera coeficientes mais acurados. O modelo mais ajustado nessa subsecção será comparado com o primeiro modelo apresentado na subsecção (3.3.2). Entre as medidas existentes para se analisar o grau de ajuste dos modelos multiníveis optou-se pela estatística deviance⁷ ($-2 \text{ Log Likelihood}$), a qual segue uma distribuição quiquadrada com graus de liberdade iguais à diferença entre os números de parâmetros (k) usados na estimação dos dois modelos que se quer comparar. De acordo com [Bickel \(2007\)](#) espera-se encontrar valores baixos dessa estatística para que haja melhor ajuste da modelagem. Assim, de acordo com a Tabela 10 o melhor modelo ajustado dentre os listados é o Modelo 2 (Modelo ANOVA, visto que sua estatística deviance é inferior ao apresentado pelo Modelo 1 (Modelo Nulo).

3.3.2 Modelos com intercepto aleatório e com variáveis explicativas no nível do aluno, da escola e da turma - Modelo 3 e Modelo 4

Nesta subsecção apresentamos a estimação do (Modelo 3)⁸ e do (Modelo 4)⁹ com o propósito de encontrarmos o modelo mais robusto, isto é, o que melhor se ajusta aos dados. Para tanto, foi realizado o teste de razão de verossimilhança que resultou em 6, 10. Este valor é obtido pelo quociente da diferença entre os deviances dos dois modelos e a diferença entre o número de parâmetros estimados. Dado que o $p - \text{valor}$ encontrado foi 0,0473, rejeitamos a hipótese nula e concluímos que o modelo 4 é bem mais ajustado que o modelo 3. Esses modelos apresentam a seguinte expressão geral:

Modelo 3

$$\begin{aligned} \text{Nota}_{ijk} = & \gamma_{000} + \gamma_{100}\text{Raça}_{ijk} + \gamma_{200}\text{Reprovado}_{ijk} + \gamma_{300}\text{Rede(wi fi)} + \\ & + \gamma_{400}\text{Escolaridade(mãeI)}_{ijk} + \gamma_{500}\text{Escolaridade(mãeII)}_{ijk} + \\ & + \gamma_{600}\text{Escolaridade(mãeIII)}_{ijk} + \gamma_{700}\text{Municipal}_{ijk} + \\ & + \gamma_{800}\text{Biblioteca}_{ijk} + \gamma_{900}\text{Sala(leitura)}_{ijk} + \gamma_{1000}\text{Computador(internet)}_{ijk} + \\ & + \gamma_{1100}\text{Projeto(CT)}_{ijk} + \gamma_{1200}\text{Projeto(MT)}_{ijk} + \gamma_{1300}\text{Experiência(diretor)}_{ijk} \end{aligned}$$

⁷ É uma estatística que mede a falta de ajuste entre os dados e o modelo

⁸ Modelo com intercepto aleatório e variáveis explicativas no nível do aluno e da escola

⁹ Modelo com intercepto aleatório e variáveis explicativas no nível do aluno, da escola e da turma

Modelo 4

$$\begin{aligned}
 \text{Nota}_{ijk} = & \gamma_{000} + \gamma_{100}\text{Raça}_{ijk} + \gamma_{200}\text{Reprovado}_{ijk} + \gamma_{300}\text{Rede(wi fi)} + \\
 & + \gamma_{400}\text{Escolaridade(mãeI)}_{ijk} + \gamma_{500}\text{Escolaridade(mãeII)}_{ijk} + \\
 & + \gamma_{600}\text{Escolaridade(mãeIII)}_{ijk} + \gamma_{700}\text{Municipal}_{ijk} + \\
 & + \gamma_{800}\text{Biblioteca}_{ijk} + \gamma_{900}\text{Sala(leitura)}_{ijk} + \gamma_{1000}\text{Computador(internet)}_{ijk} + \\
 & + \gamma_{1100}\text{Projeto(CT)}_{ijk} + \gamma_{1200}\text{Projeto(MT)}_{ijk} + \gamma_{1300}\text{Experiência(diretor)}_{ijk} \\
 & + \gamma_{1400}\text{DE(professor)}_{ijk} + \gamma_{1500}\text{Especialização}_{ijk}
 \end{aligned}$$

A Tabela 11 apresenta os resultados estimados referente ao modelo 4 (Modelo com intercepto aleatório e com variáveis explicativas no nível do aluno, da escola e da turma) para a proficiência em Matemática. Vale ressaltar que a escolha de se trabalhar com a proficiência em Matemática parte de que alguns autores [Andrade e Soares \(2008\)](#), [Rodrigues, Rios-Neto e Pinto \(2011\)](#), [Palermo, Silva e Novellino \(2014\)](#) apontam que esta disciplina seria mais afetada pela qualidade das aulas e da própria escola do que a proficiência em língua portuguesa. Segundo [Palermo, Silva e Novellino \(2014\)](#) as habilidades de leitura, por sua vez, seriam fortemente influenciadas pelo capital cultural que os alunos trazem de casa. O primeiro resultado a ser analisado é o ICC¹⁰ da turma e da escola que apresentaram respectivamente os seguintes valores 0,1703 e 0,1342. Ao se estimar o modelo 4, percebe-se uma redução do coeficiente intraclasses da escola em relação aos modelos anteriores para 0,1342. Ou seja, apenas 13,42% da variância total das notas dos alunos do 9º é explicada pela diferença entre escolas públicas. Tal redução era de se esperar ao serem incluídas variáveis independentes, reduzindo assim o percentual da variância total explicada pelo tipo de escola, evidenciando a importância do nível do aluno para explicar o desempenho escolar.

No que tange a nossa variável de interesse computador com internet na escola “Computador(internet)”, o efeito sobre a proficiência em matemática apresentou sinal esperado e foi estatisticamente significativo. Assim, de acordo com a Tabela 11 podemos perceber que o fato do aluno estudar em escolas que tenham computador com internet aumenta a proficiência em 6,728 pontos. Possivelmente, a presença de computadores com internet nas escolas estimula os alunos a desenvolverem habilidades matemática não encontradas nos livros através de jogos e softwares exclusivos para o aprendizado matemático. Esse achado corrobora com o encontrado em [Young \(2006\)](#), [Biondi e Felício \(2007\)](#), [Spiezia \(2011\)](#) que também encontraram um efeito positivo da presença do computador com internet nas escolas sobre o desempenho. Assim o efeito encontrado

¹⁰ Coeficiente de correlação intraclasses

nas estimações confirmam a hipótese de que o computador com internet aumenta a proficiência em Matemática do aluno.

Ainda no tocante ao efeito positivo da variável de interesse, [Woessmann e Fuchs \(2004\)](#) afirmam que as hipóteses positivas chegam ao ponto em que tudo o mais constante, os computadores constituem um insumo no processo de aprendizagem dos alunos que deve ajudar a produzir um melhor resultado de aprendizagem. Assim usar o computador para empregar software educacional pode inferir positivamente o conhecimento para os alunos. Além disso, o acesso à Internet pode ajudar os alunos a explorar enormes possibilidades de informação para fins escolares e aumentar a aprendizagem por meio da comunicação.

Em relação ao resultado encontrado para a variável de interesse, o mesmo contrapõe aos encontrados em ([WENGLINSKY, 1998](#); [ANGRIST; LAVY, 2002](#)). Vale salientar que o estudo quase experimental realizado por [Angrist e Lavy \(2002\)](#) foi com alunos da quarta série em matemática, o que mostra um público diferente do que foi analisado nesse estudo.

O processo de aprendizagem através do computador implica que o aluno, através dele, possa adquirir conceitos sobre o conteúdo em estudo. Entretanto, o computador não é o instrumento que ensina, mas a ferramenta com a qual o aluno desenvolve algo, e, portanto, o aprendizado ocorre pelo fato de estar executando uma tarefa por intermédio do computador. Assim, o uso do computador com internet pode ter um efeito positivo no desempenho acadêmico, desde que sejam cuidadosos e principalmente direcionados para atividades de apoio à aprendizagem ([LEI; ZHAO, 2007](#); [SALOMON; KOLIKANT, 2016](#)).

No tocante a estimação pela modelagem multinível do modelo 4, em relação às informações referentes à infraestrutura da escola, observa-se que escolas que tem sala de leitura e biblioteca apresentam melhor desempenho médio em relação às escolas que não são dotadas destes atributos. Os resultados encontrados na estimação desses efeitos foram estatisticamente significativos e com sinal positivo. Assim, observa-se a importância das condições físicas e de funcionamento dos estabelecimentos de ensino para o bom desempenho acadêmico, em concordância com ([HANUSHEK, 1989](#); [BARBOSA; FERNANDES, 2000b](#); [MENEZES-FILHO, 2007b](#)).

Em relação as variáveis “Projeto(CT)” e “Projeto(mundo trabalho)” observou-se um efeito positivo para ambas as variáveis nos modelos 3 e 4 apresentados na Tabela 11, isto é, escolas que oferecem para os seus alunos um projeto voltado para Ciência e Tecnologia, bem como para o mercado de trabalho apresentam um impacto positivo. A importância de um projeto de ciência e tecnologia em uma escola é desenvolver a alfabetização científica e tecnológica dos cidadãos, auxiliando o aluno a construir conhecimentos, habilidades e valores necessários para tomar decisões responsáveis sobre questões de ciência e tecnologia na sociedade e atuar na solução de tais questões ([SOLOMON, 1988](#); [IGLESIA, 1995](#)).

Tabela 11 – Resultado da estimação dos modelos com intercepto aleatório para a proficiência em Matemática

Variáveis explicativas	Modelo 3		Modelo 4	
Efeito fixo	Estimativa	Erro padrão	Estimativa	Erro padrão
Intercepto	245,958***	(0,44)	245,929***	(0,46)
Raça	5,726***	(0,12)	5,726***	(0,12)
Reprovado	-23,779***	(0,13)	-23,778***	(0,13)
Rede(wi fi)	3,982***	(0,14)	3,983***	(0,14)
Escolaridade(mãeI)	7,414***	(0,17)	7,414***	(0,17)
Escolaridade(mãeII)	11,779***	(0,13)	11,779***	(0,13)
Escolaridade(mãeIII)	14,454***	(0,18)	14,453***	(0,18)
Municipal	0,754**	(0,31)	0,784**	(0,31)
Biblioteca	4,268***	(0,32)	4,284***	(0,32)
Sala(leitura)	3,570***	(0,32)	3,564***	(0,32)
Computador(internet)	6,729***	(0,30)	6,728***	(0,30)
Projeto(CT)	2,506***	(0,31)	2,519***	(0,31)
Projeto(MT)	1,843***	(0,32)	1,839***	(0,32)
Experiência(diretor)	0,179***	(0,02)	0,179***	(0,02)
DE(professor)			0,245	(0,22)
Especialização			-0,549*	(0,25)
Efeito aleatório	Estimativa	Erro padrão	Estimativa	Erro padrão
var(escola)	273,076	(4,30)	272,873	(4,30)
var(turma)	73,207	(1,84)	73,214	(1,84)
var(residual)	1686,201	(3,12)	1686,211	(3,12)
Deviance	6431167		6431160,9	

Fonte: Elaboração própria a partir dos microdados do SAEB 2019 e do Censo 2019

Nota: Erro padrão entre parênteses. * $p < 0.05$. ** $p < 0.01$. *** $p < 0.001$.

No que se refere às características dos diretores, nota-se que, nas escolas que têm diretores mais experientes, para cada ponto adicional, em média, a nota esperada em na proficiência em Matemática aumenta em 0,179 pontos em ambos os modelos analisados na Tabela 11, sendo as outras variáveis mantidas constantes. Isso representa um aumento significativo, mas com pouco impacto positivo. O resultado encontrado corrobora com os achados em Eberts e Stone (1988), Filho e Ribeiro (2009), Bêteille, Kalogrides e Loeb (2012), Benedicto e Teixeira (2020) que também mostram um impacto positivo da experiência do diretor.

3.4 Considerações finais

Este artigo se propôs a estudar os efeitos do computador com internet nas escolas sobre a proficiência em matemática referente aos alunos do 9º ano do ensino fundamental das escolas públicas. Devido ao avanço tecnológico da informática no Brasil, e considerando que as escolas particulares são mais estruturadas do que as escolas públicas estaduais e municipais do ponto de vista tecnológico e fisicamente, foi considerado nesse ensaio apenas os alunos que estudam em escolas públicas brasileiras. Assim, para atender aos objetivos do estudo, utilizou-se a metodologia referente a modelagem multinível usada na economia da educação para dados oriundos de clusters.

Tal procedimento identificou efeitos positivos sobre a proficiência em matemática. Vale salientar que para chegarmos ao modelo final utilizando a estatística deviance e o teste da razão de verossimilhança foi estimado inicialmente um modelo 1 (modelo nulo) para mostrar que o método de estimação por OLS não é adequado. Em seguida foi inserido variáveis no nível do aluno para obtermos o modelo 2 (modelo ANCOVA) que comparado com o modelo 1 foi mais ajustado aos dados. Adiante, foi estimado mais dois modelos onde foi escolhido o modelo 4 (modelo com intercepto aleatório e com variáveis explicativas no nível do aluno, da escola e da turma) como modelo final por ser mais robusto que o modelo 3 (modelo com intercepto aleatório e com variáveis explicativas no nível do aluno e da escola).

Os resultados das estimações encontrados nos modelos para a variável de interesse “Computador(internet)” foi estatisticamente significativa e apresentou o sinal esperado (positivo), isto é, a hipótese de que o efeito do computador com internet nas escolas aumenta o desempenho em matemática foi corroborada com a literatura citada nesse ensaio. Assim, pode-se observar que as escolas que ofertam computadores com internet para o uso dos alunos tem um ganho maior na proficiência em Matemática dos seus alunos.

Cabe ainda destacar que o uso dos computadores com internet nas escolas oferece aos alunos e professores ferramentas sofisticadas para o desenvolvimento de atividades que não se encontram nos livros didáticos. Como exemplo dessas ferramentas podemos citar os jogos matemáticos, sites com conteúdos voltados para olimpíadas de matemática, banco de questões para professores elaborarem provas, e etc. Por outro lado, as variáveis utilizadas como controle no modelo final também tiveram um efeito estatisticamente significativo, mostrando que a escolha para esses controles estão de acordo com a literatura referente a economia da educação.

Diante disso, nota-se a importância de testar novos insumos da função de produção educacional para explicar, por exemplo, a proficiência dos alunos. Assim esse estudo contribui para a literatura da economia como forma de explicar o desempenho dos alunos em matemática por um insumo pouco utilizado na literatura, porém utilizando uma

metodologia de regressão linear multinível para dados estruturados em clusters.

Por fim, a presença de computadores com internet nas escolas influencia positivamente o desempenho em matemática do alunado, mesmo após controlar pelo efeito de características do aluno, da turma e da escola. Assim, é preciso que mais atenção seja dada a essa questão, que demanda políticas públicas eficazes para aquisição por exemplo de computadores com internet nas escolas onde esse equipamento de informática não existe. Também é preciso que a estrutura física da escola, como a construção de laboratórios de informática esteja entre as prioridades no planejamento estrutural das escolas, pois a implantação de um laboratório não só de informática, mas de outras áreas revela um ambiente mais propício ao aprendizado e ao acúmulo de capital humano, essenciais ao desenvolvimento do país.

Referências

- ADELSON, J. L.; MCCOACH, D. B. Development and psychometric properties of the math and me survey: Measuring third through sixth graders' attitudes toward mathematics. *Measurement and Evaluation in Counseling and Development*, Sage Publications Sage CA: Los Angeles, CA, v. 44, n. 4, p. 225–247, 2011. Citado na página 3.
- AIKEN, L. R. Two scales of attitude toward mathematics. *Journal for research in Mathematics Education*, National Council of Teachers of Mathematics, v. 5, n. 2, p. 67–71, 1974. Citado na página 1.
- ALHAJRAF, N. M.; ALASFOUR, A. M. The impact of demographic and academic characteristics on academic performance. *International Business Research*, Canadian Center of Science and Education, v. 7, n. 4, p. 92, 2014. Citado na página 32.
- ALMASI, M.; MACHUMU, H.; ZHU, C. Internet use among secondary schools students and its effects on their learning. In: *Proceedings of INTED2017 Conference 6th-8th March*. [S.l.: s.n.], 2017. Citado na página 34.
- ANDRADE, R. J. de; SOARES, J. F. O efeito da escola básica brasileira. *Estudos em avaliação educacional*, v. 19, n. 41, p. 379–406, 2008. Citado 3 vezes nas páginas 1, 4 e 47.
- ANGRIST, J.; LAVY, V. New evidence on classroom computers and pupil learning. *The Economic Journal*, Oxford University Press Oxford, UK, v. 112, n. 482, p. 735–765, 2002. Citado 2 vezes nas páginas 30 e 48.
- ANGRIST, J. D.; PISCHKE, J.-S. *Mostly harmless econometrics: An empiricist's companion*. [S.l.]: Princeton university press, 2008. Citado na página 14.
- ARNOLD, I. J.; STRATEN, J. T. Motivation and math skills as determinants of first-year performance in economics. *The Journal of Economic Education*, Taylor & Francis, v. 43, n. 1, p. 33–47, 2012. Citado na página 6.
- ASHCRAFT, M. H.; KRAUSE, J. A. Working memory, math performance, and math anxiety. *Psychonomic bulletin & review*, Springer, v. 14, n. 2, p. 243–248, 2007. Citado na página 1.
- BARBOSA, M. E. F.; FERNANDES, C. Modelo multinível: uma aplicação a dados de avaliação educacional. *Estudos em Avaliação Educacional*, n. 22, p. 135–154, 2000. Citado na página 21.
- BARBOSA, M. E. F.; FERNANDES, C. Modelo multinível: uma aplicação a dados de avaliação educacional. *Estudos em Avaliação Educacional*, n. 22, p. 135–154, 2000. Citado na página 48.
- BARROS, R. P. d. et al. Determinantes do desempenho educacional no brasil (texto para discussão n. 834, pp. 1-33). *Rio de Janeiro, RJ: IPEA*, 2001. Citado na página 45.
- BARROW, L.; MARKMAN, L.; ROUSE, C. E. Technology's edge: The educational benefits of computer-aided instruction. *American Economic Journal: Economic Policy*, v. 1, n. 1, p. 52–74, 2009. Citado na página 29.

- BECKER, G. S. *Human capital: A theoretical and empirical analysis, with special reference to education*. [S.l.]: University of Chicago press, 2009. Citado na página 13.
- BEEK, J. P. Van der et al. Self-concept mediates the relation between achievement and emotions in mathematics. *British Journal of Educational Psychology*, Wiley Online Library, v. 87, n. 3, p. 478–495, 2017. Citado na página 3.
- BELBASE, S. Images, anxieties and attitudes toward mathematics. *Online Submission*, ERIC, 2010. Citado na página 5.
- BENEDICTO, B. V.; TEIXEIRA, E. C. O efeito do perfil do diretor escolar sobre a proficiência dos alunos no estado de minas gerais. *Economia Aplicada*, v. 24, n. 1, p. 5–28, 2020. Citado na página 49.
- BÉTEILLE, T.; KALOGRIDES, D.; LOEB, S. Stepping stones: Principal career paths and school outcomes. *Social Science Research*, Elsevier, v. 41, n. 4, p. 904–919, 2012. Citado na página 49.
- BICKEL, R. *Multilevel analysis for applied research: It's just regression!* [S.l.]: Guilford Press, 2007. Citado na página 46.
- BIONDI, R. L.; FELÍCIO, F. d. Atributos escolares e o desempenho dos estudantes: uma análise em painel dos dados do saeb. *Brasília: Inep*, 2007. Citado 2 vezes nas páginas 19 e 47.
- BIONDI, R. L.; VASCONCELLOS, L.; MENEZES-FILHO, N. d. Avaliando o impacto da olimpíada brasileira de matemática das escolas públicas (obmep) no desempenho de matemática nas avaliações educacionais. *São Paulo: Fundação Getúlio Vargas, Escola de Economia de São Paulo*, 2009. Citado na página 19.
- BOBBA, M.; FRISANCHO, V. Self-perceptions about academic achievement: Evidence from mexico city. *Journal of Econometrics*, Elsevier, 2020. Citado na página 32.
- BRAMLETT, D. C. A study of african-american college students'attitudes towards mathematics. 2007. Citado na página 6.
- BROOMHALL, D. E.; JOHNSON, T. G. Economic factors that influence educational performance in rural schools. *American Journal of Agricultural Economics*, Wiley Online Library, v. 76, n. 3, p. 557–567, 1994. Citado na página 32.
- BULMAN, G.; FAIRLIE, R. W. Technology and education: Computers, software, and the internet. In: *Handbook of the Economics of Education*. [S.l.]: Elsevier, 2016. v. 5, p. 239–280. Citado 4 vezes nas páginas 2, 28, 29 e 30.
- CALIENDO, M.; KOPEINIG, S. Some practical guidance for the implementation of propensity score matching. *Journal of economic surveys*, Wiley Online Library, v. 22, n. 1, p. 31–72, 2008. Citado na página 17.
- CAZORLA, I. et al. Relação entre o domínio afetivo e o desempenho em matemática de estudantes das séries iniciais do ensino fundamental. *Revista Brasileira de Estudos Pedagógicos*, v. 89, n. 221, 2008. Citado na página 6.

- CECHIN, L. A. W. et al. O impacto das regras do programa bolsa família sobre a fecundidade das beneficiárias. *Revista Brasileira de Economia*, SciELO Brasil, v. 69, n. 3, p. 303–329, 2015. Citado na página 24.
- CLOTFELTER, C. T.; LADD, H. F.; VIGDOR, J. L. Teacher-student matching and the assessment of teacher effectiveness. *Journal of human Resources*, University of Wisconsin Press, v. 41, n. 4, p. 778–820, 2006. Citado na página 19.
- CORREA, E. *Efeito da repetência nos anos iniciais do Ensino Fundamental: um estudo longitudinal a partir do Geres*. Tese (Doutorado) — PUC-Rio, 2013. Citado na página 18.
- COURI, C. Nível socioeconômico e cor/raça em pesquisas sobre efeito-escola. *Estudos em Avaliação Educacional*, v. 21, n. 47, p. 449–472, 2010. Citado na página 45.
- CUBAN, L. Computers meet classroom: Classroom wins. *Teachers college record*, v. 95, n. 2, p. 185–210, 1993. Citado na página 29.
- DICKE, T. et al. Effects of school-average achievement on individual self-concept and achievement: Unmasking phantom effects masquerading as true compositional effects. *Journal of Educational Psychology*, American Psychological Association, v. 110, n. 8, p. 1112, 2018. Citado na página 32.
- DOWKER, A.; BENNETT, K.; SMITH, L. Attitudes to mathematics in primary school children. *Child Development Research*, Hindawi, v. 2012, 2012. Citado na página 1.
- EBERTS, R. W.; STONE, J. A. Student achievement in public schools: Do principals make a difference? *Economics of Education Review*, Elsevier, v. 7, n. 3, p. 291–299, 1988. Citado na página 49.
- EGALITE, A. J. How family background influences student achievement. *Education Next*, v. 16, n. 2, p. 70–78, 2016. Citado na página 32.
- ERDOGDU, F.; ERDOGDU, E. The impact of access to ict, student background and school/home environment on academic success of students in turkey: An international comparative analysis. *Computers & Education*, Elsevier, v. 82, p. 26–49, 2015. Citado na página 34.
- FELÍCIO, F.; FERNANDES, R. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do ensino fundamental no estado de são paulo. *Encontro Nacional de Economia*, v. 33, p. 2005, 2005. Citado na página 35.
- FERRÃO, M. E. et al. O saeb—sistema nacional de avaliação da educação básica: objetivos, características e contribuições na investigação da escola eficaz. *Revista Brasileira de Estudos de População*, v. 18, n. 1/2, p. 111–130, 2001. Citado na página 45.
- FILHO, N. M.; RIBEIRO, F. P. Os determinantes da melhoria do rendimento escolar. *GIAMBIAGI, F.; HENRIQUES, R.; PESSÔA, S*, p. 171–188, 2009. Citado na página 49.
- GAFOOR, K. A.; KURUKKAN, A. Why high school students feel mathematics difficult? an exploration of affective beliefs. *Online Submission*, ERIC, 2015. Citado na página 1.
- GALVEZ-NIETO, J. L. et al. Multilevel latent class cluster analysis of school climate: individual, family and community factors. *Revista de Psicodidáctica (English ed.)*, Elsevier, v. 25, n. 2, p. 85–92, 2020. Citado na página 32.

- GOLDHABER, D.; ANTHONY, E. Can teacher quality be effectively assessed? national board certification as a signal of effective teaching. *The Review of Economics and Statistics*, The MIT Press, v. 89, n. 1, p. 134–150, 2007. Citado na página 19.
- GOLDSTEIN, H. *Multilevel statistical models*. [S.l.]: John Wiley & Sons, 2011. v. 922. Citado na página 36.
- GOTTFRIED, A. E. et al. Multivariate latent change modeling of developmental decline in academic intrinsic math motivation and achievement: Childhood through adolescence. *International Journal of Behavioral Development*, Sage Publications Sage UK: London, England, v. 31, n. 4, p. 317–327, 2007. Citado na página 3.
- GUAN, S.-S. A.; SUBRAHMANYAM, K. Youth internet use: risks and opportunities. *Current opinion in Psychiatry*, LWW, v. 22, n. 4, p. 351–356, 2009. Citado na página 34.
- HANUSHEK, E. The production of education, teacher quality and efficiency. ERIC, 1970. Citado na página 32.
- HANUSHEK, E. A. Conceptual and empirical issues in the estimation of educational production functions. *Journal of human Resources*, JSTOR, p. 351–388, 1979. Citado na página 29.
- HANUSHEK, E. A. The impact of differential expenditures on school performance. *Educational researcher*, Sage Publications Sage CA: Thousand Oaks, CA, v. 18, n. 4, p. 45–62, 1989. Citado na página 48.
- HANUSHEK, E. A. Education production functions. In: *The economics of education*. [S.l.]: Elsevier, 2020. p. 161–170. Citado na página 1.
- HANUSHEK, E. A.; WOESSMANN, L. Do better schools lead to more growth? cognitive skills, economic outcomes, and causation. *Journal of economic growth*, Springer, v. 17, n. 4, p. 267–321, 2012. Citado na página 6.
- HANUSHEK, E. A.; WOESSMANN, L. School resources and student achievement: A review of cross-country economic research. *Cognitive abilities and educational outcomes*, Springer, p. 149–171, 2017. Citado na página 32.
- HEMBREE, R. The nature, effects, and relief of mathematics anxiety. *Journal for research in mathematics education*, National Council of Teachers of Mathematics, v. 21, n. 1, p. 33–46, 1990. Citado na página 1.
- HIRANO, K.; IMBENS, G. W.; RIDDER, G. Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, Wiley Online Library, v. 71, n. 4, p. 1161–1189, 2003. Citado na página 13.
- HOX, J. J.; MOERBEEK, M.; SCHOOT, R. Van de. *Multilevel analysis: Techniques and applications*. [S.l.]: Routledge, 2010. Citado na página 46.
- IGLESIA, P. M. Cts en la enseñanza-aprendizaje de las ciencias experimentales. *Alambique: didáctica de las ciencias experimentales*, Graó, n. 3, p. 7–12, 1995. Citado na página 48.
- JACOMINI, M. A. Educar sem reprovar: desafio de uma escola para todos. *Educação e pesquisa*, SciELO Brasil, v. 35, p. 557–572, 2009. Citado na página 18.

- JEPSEN, C. Teacher characteristics and student achievement: Evidence from teacher surveys. *Journal of Urban Economics*, Elsevier, v. 57, n. 2, p. 302–319, 2005. Citado na página 19.
- KHANDKER, S.; KOOLWAL, G. B.; SAMAD, H. *Handbook on impact evaluation: quantitative methods and practices*. [S.l.]: The World Bank, 2009. Citado na página 15.
- KHINE, M. S.; AL-MUTAWAH, M.; AFARI, E. Determinants of affective factors in mathematics achievement: Structural equation modeling approach. *Journal of Studies in Education*, v. 5, n. 2, p. 199–211, 2015. Citado na página 5.
- KOEDINGER, K. R. et al. Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, v. 8, n. 1, p. 30–43, 1997. Citado na página 28.
- KRIEG, J. M. Teacher quality and attrition. *Economics of Education review*, Elsevier, v. 25, n. 1, p. 13–27, 2006. Citado na página 19.
- KROHN, G. A.; O’CONNOR, C. M. Student effort and performance over the semester. *The Journal of Economic Education*, Taylor & Francis, v. 36, n. 1, p. 3–28, 2005. Citado na página 32.
- KUPARI, P.; NISSINEN, K. Background factors behind mathematics achievement in finnish education context: Explanatory models based on timss 1999 and timss 2011 data. In: CITESEER. *IEA CONFERENCE 2013, Proceedings*. [S.l.], 2013. Citado 2 vezes nas páginas 1 e 3.
- LAI, F. et al. Does computer-assisted learning improve learning outcomes? evidence from a randomized experiment in migrant schools in beijing. *Economics of Education Review*, Elsevier, v. 47, p. 34–48, 2015. Citado na página 30.
- LAROS, J. A.; MARCIANO, J. L. P.; ANDRADE, J. M. de. Fatores que afetam o desempenho na prova de matemática do saeb: um estudo multinível. *Avaliação Psicológica: Interamerican Journal of Psychological Assessment*, Instituto Brasileiro de Avaliação Psicológica (IBAP), v. 9, n. 2, p. 173–186, 2010. Citado 2 vezes nas páginas 6 e 20.
- LEI, J.; ZHAO, Y. Technology uses and student achievement: A longitudinal study. *Computers & Education*, Elsevier, v. 49, n. 2, p. 284–296, 2007. Citado na página 48.
- LEWBEL, A. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. *Journal of Business & Economic Statistics*, Taylor & Francis Group, v. 30, n. 1, p. 67–80, 2012. Citado 2 vezes nas páginas 4 e 16.
- LIM, S. Y.; CHAPMAN, E. Identifying affective domains that correlate and predict mathematics performance in high-performing students in singapore. *Educational Psychology*, Taylor & Francis, v. 35, n. 6, p. 747–764, 2015. Citado na página 5.
- LINDBECK, A.; SNOWER, D. J. Multitask learning and the reorganization of work: From tayloristic to holistic organization. *Journal of labor economics*, The University of Chicago Press, v. 18, n. 3, p. 353–376, 2000. Citado na página 28.
- LINDEN, L.; BANERJEE, A.; DUFLO, E. Computer-assisted learning: Evidence from a randomized experiment. *Poverty Action Lab Paper*, v. 5, 2003. Citado na página 29.

- LOBO, G. D.; CASSUCE, F. C. da C.; CIRINO, J. F. Avaliação do desempenho escolar dos estudantes da região nordeste que realizaram o enem: uma análise com modelos hierárquicos. *Revista Espacios*, v. 6, 2017. Citado 2 vezes nas páginas 19 e 45.
- MA, X.; KISHOR, N. Assessing the relationship between attitude toward mathematics and achievement in mathematics: A meta-analysis. *Journal for research in mathematics education*, JSTOR, p. 26–47, 1997. Citado 3 vezes nas páginas 1, 5 e 6.
- MAGNUSON, K. *The effect of increases in welfare mothers' education on their young children's academic and behavioral outcomes: Evidence from the National Evaluation of Welfare-to-Work Strategies Child Outcomes Study*. [S.l.]: Institute for Research on Poverty Madison, WI, 2003. Citado na página 45.
- MARKOVITS, Z.; FORGASZ, H. “mathematics is like a lion”: Elementary students' beliefs about mathematics. *Educational Studies in Mathematics*, Springer, v. 96, n. 1, p. 49–64, 2017. Citado na página 3.
- MAZANA, Y. M.; MONTERO, C. S.; OLIFAGE, C. R. Investigating students' attitude towards learning mathematics. Modestum Limited, 2019. Citado na página 3.
- MENEZES-FILHO, N. A. Os determinantes do desempenho escolar do brasil. IFB, 2007. Citado na página 45.
- MENEZES-FILHO, N. A. Os determinantes do desempenho escolar do brasil. IFB, 2007. Citado na página 48.
- MENEZES, L. Matemática, literatura & aulas. *Educação e Matemática*, n. 115, p. 67–71, 2011. Citado na página 19.
- MOKGWATHI, M. S.; GRAHAM, M. A.; FRASER, W. J. The relationship between grade 9 teacher's and learner's perceptions and attitudes with their mathematics achievement. Eskisehir Osmangazi University, Faculty of Education, 2019. Citado 3 vezes nas páginas 5, 20 e 21.
- MULLIS, I. V. et al. Assessment framework and instrument development. *Methods and procedures in TIMSS and PIRLS*, p. 1–22, 2011. Citado na página 5.
- NICOLAIDOU, M.; PHILIPPOU, G. Attitudes towards mathematics, self-efficacy and achievement in problem solving. *European Research in Mathematics Education III. Pisa: University of Pisa*, Citeseer, p. 1–11, 2003. Citado na página 6.
- OLIVEIRA, A. C. P. d.; WALDHLM, A. P. S. Liderança do diretor, clima escolar e desempenho dos alunos: qual a relação? *Ensaio: Avaliação e Políticas Públicas em Educação*, SciELO Brasil, v. 24, n. 93, p. 824–844, 2016. Citado na página 35.
- OSHCHEPKOV, A. Y.; SHIROKANNOVA, A. Multilevel modeling for economists: Why, when and how. *Higher School of Economics Research Paper No. WP BRP*, v. 233, 2020. Citado na página 36.
- PALERMO, G. A.; SILVA, D. B. d. N.; NOVELLINO, M. S. F. *Fatores associados ao desempenho escolar: uma análise da proficiência em matemática dos alunos do 5º ano do ensino fundamental da rede municipal do Rio de Janeiro*. [S.l.]: SciELO Brasil, 2014. Citado 4 vezes nas páginas 1, 4, 19 e 47.

- PAPANASTASIOU, C. Effects of attitudes and beliefs on mathematics achievement. *Studies in educational evaluation*, ERIC, v. 26, n. 1, p. 27–42, 2000. Citado na página 6.
- PARSONS, S.; CROFT, T.; HARRISON, M. Does students' confidence in their ability in mathematics matter? *Teaching Mathematics and Its Applications: An International Journal of the IMA*, Oxford University Press, v. 28, n. 2, p. 53–68, 2009. Citado na página 5.
- PAULA, K. F. S. de; COUTO, A. C. L. Desempenho dos alunos do 9º ano na prova brasil 2015: Uma análise a partir de modelos multiníveis para o paran. *A Economia em Revista-AERE*, v. 27, n. 1, p. 35–48, 2019. Citado na página 6.
- PEÑA-LÓPEZ, I. et al. Pisa 2018 results. effective policies, successful schools. OECD Publishing, 2020. Citado na página 2.
- PINXTEN, M. et al. Enjoying mathematics or feeling competent in mathematics? reciprocal effects on mathematics achievement and perceived math effort expenditure. *British Journal of Educational Psychology*, Wiley Online Library, v. 84, n. 1, p. 152–174, 2014. Citado na página 3.
- RAUDENBUSH, S. W.; BRYK, A. S. *Hierarchical linear models: Applications and data analysis methods*. [S.l.]: sage, 2002. v. 1. Citado na página 36.
- RENNINGER, K. A.; HIDI, S. Student interest and achievement: Developmental issues raised by a case study. In: *Development of achievement motivation*. [S.l.]: Elsevier, 2002. p. 173–195. Citado na página 6.
- RIVKIN, S. G.; HANUSHEK, E. A.; KAIN, J. F. Teachers, schools, and academic achievement. *Econometrica*, Wiley Online Library, v. 73, n. 2, p. 417–458, 2005. Citado na página 19.
- ROCKOFF, J. E. The impact of individual teachers on student achievement: Evidence from panel data. *American economic review*, v. 94, n. 2, p. 247–252, 2004. Citado na página 19.
- RODRIGUES, C. G.; RIOS-NETO, E. L. G.; PINTO, C. C. d. X. Diferenças intertemporais na média e distribuição do desempenho escolar no brasil: o papel do nível socioeconômico, 1997 a 2005. *Revista Brasileira de Estudos de População*, SciELO Brasil, v. 28, p. 5–36, 2011. Citado 3 vezes nas páginas 1, 4 e 47.
- RODRIGUES, L. d. O. Ensaio sobre diferencial de desempenho escolar entre alunos de escolas rurais e urbanas no brasil. 2017. Citado 2 vezes nas páginas 6 e 20.
- ROSENBAUM, P. R. Overt bias in observational studies. In: *Observational studies*. [S.l.]: Springer, 2002. p. 71–104. Citado na página 17.
- ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, Oxford University Press, v. 70, n. 1, p. 41–55, 1983. Citado 3 vezes nas páginas 13, 14 e 15.
- ROUSE, C. E.; KRUEGER, A. B. Putting computerized instruction to the test: a randomized evaluation of a “scientifically based” reading program. *Economics of Education Review*, Elsevier, v. 23, n. 4, p. 323–338, 2004. Citado na página 29.

- SALOMON, A.; KOLIKANT, Y. B.-D. High-school students' perceptions of the effects of non-academic usage of ict on their academic achievements. *Computers in Human Behavior*, Elsevier, v. 64, p. 143–151, 2016. Citado na página 48.
- SCARPELLO, G. Helping students get past math anxiety. *Techniques: Connecting Education and Careers (J1)*, ERIC, v. 82, n. 6, p. 34–35, 2007. Citado na página 3.
- SINCLAIR, P. M. et al. The effectiveness of internet-based e-learning on clinician behaviour and patient outcomes: a systematic review. *International journal of nursing studies*, Elsevier, v. 57, p. 70–81, 2016. Citado na página 34.
- SNIJDERS, T. A.; BOSKER, R. J. *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. [S.l.]: sage, 2011. Citado na página 29.
- SOARES, J. F. O efeito da escola no desempenho cognitivo de seus alunos. *REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, Red Iberoamericana de Investigación Sobre Cambio y Eficacia Escolar, v. 2, n. 2, p. 6, 2004. Citado na página 19.
- SOARES, J. F.; ALVES, M.; MARI, F. Escola eficaz: um estudo de caso em três escolas da rede pública de ensino do estado de minas gerais. *Belo Horizonte: UFMG*, v. 114, 2002. Citado na página 45.
- SOLOMON, J. Science technology and society courses: Tools for thinking about social issues. *International Journal of Science Education*, Taylor & Francis, v. 10, n. 4, p. 379–387, 1988. Citado na página 48.
- SPIEZIA, V. Does computer use increase educational achievements? student-level evidence from pisa. *OECD Journal: Economic Studies*, OECD, v. 2010, n. 1, p. 1–22, 2011. Citado 2 vezes nas páginas 30 e 47.
- STANKOV, L.; MORONY, S.; LEE, Y. P. Confidence: the best non-cognitive predictor of academic achievement? *Educational psychology*, Taylor & Francis, v. 34, n. 1, p. 9–28, 2014. Citado na página 5.
- STIPEK, D. Good instruction is motivating. In: *Development of achievement motivation*. [S.l.]: Elsevier, 2002. p. 309–332. Citado na página 6.
- SYVEDA, F. Understanding attitudes towards mathematics (atm) using a multimodal model: an exploratory case study with secondary school children in england. Faculty of Education, University of Cambridge, 2016. Citado 3 vezes nas páginas 1, 3 e 6.
- TAVSANCIL, E.; YALCIN, S. A determination of turkish student's achievement using hierarchical linear models in trends in international mathematics-science study (timss) 2011. *The Anthropologist*, Taylor & Francis, v. 22, n. 2, p. 390–396, 2015. Citado na página 5.
- TORRES, R. et al. Determinantes do desempenho dos participantes da prova do enem: Evidências para o rio grande do sul. *Desenvolvimento em Questão*, v. 18, n. 53, p. 352–368, 2020. Citado na página 32.
- UPADYAYA, K.; ECCLES, J. S. How do teachers' beliefs predict children's interest in math from kindergarten to sixth grade? *Merrill-Palmer Quarterly (1982-)*, JSTOR, v. 60, n. 4, p. 403–430, 2014. Citado 2 vezes nas páginas 3 e 4.

WENGLINSKY, H. Does it compute? the relationship between educational technology and student achievement in mathematics. ERIC, 1998. Citado 2 vezes nas páginas 30 e 48.

WIMBLE, M. Understanding health and health-related behavior of users of internet health information. *Telemedicine and e-Health*, Mary Ann Liebert, Inc. 140 Huguenot Street, 3rd Floor New Rochelle, NY 10801 USA, v. 22, n. 10, p. 809–815, 2016. Citado na página 34.

WOESSMANN, L.; FUCHS, T. Computers and student learning: Bivariate and multivariate evidence on the availability and use of computers at home and at school. *Available at SSRN 619101*, 2004. Citado na página 48.

YOUNG, B. A study on the effect of internet use and social capital on the academic performance. *Journal of Asian Sociology*, JSTOR, v. 35, n. 1, p. 107–123, 2006. Citado 2 vezes nas páginas 30 e 47.