

# IMPACTO DO ENSINO INFANTIL NAS NOTAS DO 9º ANO DOS ALUNOS BRASILEIROS UTILIZANDO DADOS DO SISTEMA NACIONAL DE AVALIAÇÃO DA EDUCAÇÃO BÁSICA 2021

Henrique Diedrich Baitala<sup>1</sup>  
Walcir Soares da Silva Junior<sup>2</sup>

## RESUMO

O propósito deste estudo consiste em avaliar o efeito da política pré-escola sobre o desempenho escolar dos alunos matriculados no 9º ano do ensino fundamental brasileiros. Para diminuir quaisquer potenciais distorções provenientes do processo de seleção, optou-se pela adoção do modelo quasi-experimental Propensity Score Matching. Para essa análise, serão empregados os microdados provenientes do SAEB (Sistema Nacional de Avaliação da Educação Básica) referentes ao ano de 2021. Foram encontrados resultados positivos da pré-escola sobre o desempenho de matemática. Apesar da baixa magnitude, o efeito é estatisticamente significativo a 1%. Entretanto, para a proficiência em português os resultados não tiveram significância estatística. Os resultados corroboram, em parte, as evidências nacionais e internacionais, que sustentam a relação positiva entre a frequência no ensino infantil e o subsequente êxito escolar, mas podem ter sido afetados pela forma desigual com que a pandemia de Covid-19 impactou os resultados escolares.

**Palavras-chave:** ensino infantil, pré-escola, desempenho, avaliação de impacto.

<sup>1</sup> Aluno do curso de Ciências Econômicas da FAE Centro Universitário e orientando do Programa de Apoio à Iniciação Científica (PAIC 2023-2024). *E-mail:* henriquebaitala0303@gmail.com

<sup>2</sup> Professor da FAE Centro Universitário, Doutor em Desenvolvimento Econômico (UFPR) e orientador deste trabalho. *E-mail:* walcir.junior@fae.edu

## INTRODUÇÃO

O presente estudo tem como objetivo a condução de uma pesquisa que pretende avaliar o impacto da política pré-escola no desempenho em português e matemática dos alunos matriculados no 9º ano do ensino fundamental para os alunos de escolas públicas, e uma amostra dos alunos de escolas privadas, em todo território nacional brasileiro. Tais análises são de extrema importância para o avanço da ciência e das próprias políticas públicas, pois, além de trazerem soluções no âmbito científico, ainda verificam as regiões onde esses recursos devem ser implementados (Barros, 2012). Dentro do contexto da análise de políticas públicas, o maior desafio a ser contornado se baseia na aplicação de um grupo contrafactual aos indivíduos tratados (Foguel, 2012).

Métodos convencionais, como Mínimos Quadrados Ordinários, não contornam o desafio da endogeneidade, uma vez que seus resultados não possibilitam discernir se a relação estabelecida é causal ou uma correlação entre variáveis independentes e a variável dependente (Foguel, 2012). Por esse motivo, para realizar o decréscimo do viés de seleção baseado em situações observáveis, a metodologia Propensity Score Matching (PSM) será utilizada, gerando um modelo de seleção que antecede a probabilidade de um aluno frequentar o ensino infantil com base em atributos gerais e contínuos, tais como: características identitárias, características de ensino e desempenho escolar, características familiares e características socioeconômicas. A filtragem destes atributos é originada pela relevância no contexto da adesão ao processo de tratamento (participação ou não no ensino infantil)<sup>3</sup>.

Neste estudo a análise se baseia em um grupo controle, alunos que não iniciaram seus estudos na pré-escola, mas no ensino primário; e um grupo de tratamento, alunos que iniciaram seus estudos na pré-escola. Desse modo, após o pareamento, será avaliado o efeito médio do tratamento sobre os tratados (EMTT) utilizando o método de estratificação. Para avaliar o impacto do ensino infantil, será utilizado o método de Mínimos Quadrados Ponderados pelos escores de propensão (MQP). Esse método permite controlar as variáveis que podem influenciar o desempenho dos alunos, considerando os escores previstos pelo PSM e o modelo duplamente robusto, Regressão Ponderada pelos Escores de Propensão. As estimativas serão baseadas em microdados do Sistema Nacional de Avaliação da Educação Básica (SAEB) de 2021, disponibilizados pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

O objetivo geral deste artigo é avaliar o impacto entre iniciar os estudos na pré-escola e o desempenho de alunos no 9º ano do ensino fundamental brasileiros no teste padronizado SAEB. Os objetivos específicos serão i) determinar a influência

<sup>3</sup> Retirado do capítulo 3 deste artigo e justificado em 3.4.

do ensino infantil sobre as notas dos alunos no 9º ano do ensino fundamental; ii) compreender as características socioeconômicas dos alunos brasileiros no 9º ano do ensino fundamental; iii) discorrer sobre a importância do ensino infantil no desenvolvimento de habilidades e seu estado da arte da literatura internacional.

Uma das principais contribuições deste artigo é capturar como os efeitos desiguais da pandemia de Covid-19 afetaram o impacto da pré-escola. Também será possível avaliar esse efeito após a Lei 12.796 de 5 de abril de 2013, que ampliou o ensino fundamental em um ano considerando o que antes era a pré-escola como o primeiro ano. Alunos que iniciaram seus estudos diretamente no primeiro ano em 2013, estariam no 9º ano em 2021, caso não sofressem reprovação.

Portanto, além da avaliação de impacto, o estudo também explorará características sociais, econômicas, familiares e escolares dos indivíduos. A relevância da educação infantil no contexto do desenvolvimento econômico é incontestável, dada as inúmeras políticas educacionais que circundam as nações do mundo contemporâneo, esse tipo de investimento já foi investigado anteriormente e destaca resultados nos quais o retorno financeiro pode chegar a 13 vezes o custo do investimento em educação pré-escolar (Heckman, 2006).

O presente artigo apresenta cinco seções a contar com a introdução. A segunda seção apresenta uma breve revisão de literatura e um apanhado das principais teorias de capital humano. A terceira seção apresenta as metodologias empregadas, a análise descritiva e apresenta o modelo de seleção empregado na avaliação de impacto. A quarta seção apresenta os principais resultados. A última seção apresenta as considerações finais.

## **1 ALFABETIZAÇÃO NO BRASIL, UMA HEMORRAGIA LONGE DE SER ESTANCADA**

Tendo em conhecimento que o indivíduo durante a infância tem como suas principais características ser ativo e construtor de conhecimento, a pré-escola tem papel fundamental dentro do processo de socialização do indivíduo. Essas instituições apresentam oportunidades para a construção e o desenvolvimento da sua autonomia, da sua socialização, do seu desenvolvimento intelectual, ao mesmo tempo, importa promover a sua integração social e a sua predisposição positiva para ingresso na escolaridade obrigatória (Marchão, 2011).

Em seu artigo seminal, Vygotsky (1934) argumenta que a linguagem é uma ferramenta essencial para o pensamento e o aprendizado. Ela permite comunicar ideias,

resolver problemas e aprender sobre o mundo ao seu redor. No contexto brasileiro, a educação infantil é segmentada em creche (0 a 3 anos) e pré-escola (4 a 5 anos). A partir dos anos 1990, a importância do acesso a esse nível educacional aumentou, como consequência da redemocratização, após a Constituição de 1988, que conferiu não apenas um aumento quantitativo à educação das crianças de zero a seis anos, mas também a estabeleceu como um direito infantil primordial.

Entretanto, a realidade brasileira diverge substancialmente da visão utópica delineada na Constituição. Conforme evidenciado pela Pesquisa Nacional por Amostra de Domicílios (PNAD, 2023), o Brasil, somente para alunos que deveriam estar frequentando o ensino médio, a evasão aproximadamente 9 milhões de brasileiros entre 14 a 29 anos de idade. Os principais motivos subjacentes ao fenômeno de abandono escolar, que culmina em uma taxa de evasão de até 4,1% para este grupo, estão intrinsecamente relacionados à dificuldade de acesso, à necessidade de trabalhar e à dificuldade de conciliar trabalho e estudos. Um percentual de 41,7% dos alunos reportou este último como principal motivo que culminou na sua evasão escolar em 2023. É pertinente destacar que essa dificuldade constitui uma transgressão aos direitos consagrados na Constituição de 1988, uma vez que cabe ao Estado o dever de proporcionar uma infraestrutura educacional que torne o acesso à escola acessível e descomplicado se adaptando a realidade de todos os brasileiros.

Além dessa preocupante estatística, a pesquisa também revela que a parcela mais afetada por essa deficiência do poder estatal corresponde às camadas consideradas como minorias da população. Constatou-se que indivíduos que se autodeclararam como sendo de cor ou raça preta ou parda enfrentam uma taxa de evasão de 6,4%, o que representa uma disparidade de 3,9 pontos percentuais em comparação com indivíduos autodeclarados como de cor ou raça branca.

Com as recentes mudanças, a partir da Lei 12.796 de 5 de abril de 2013, que ampliou o ensino fundamental em um ano considerando o que antes era a pré-escola como o primeiro ano, o aumento dos investimentos na educação infantil demonstra a significância dessa etapa educacional para o desenvolvimento dos alunos. Entretanto, apesar do Brasil investir uma parcela maior do PIB em educação do que a média de países desenvolvidos - 4,2% em comparação a 3,2% - proporcionalmente o valor investido por aluno é menos da metade desses países (OECD, 2016).

Quando se trata do ensino infantil, o investimento brasileiro é de 3,8 mil dólares por aluno, enquanto a média da OCDE de 8,6 mil dólares (OCDE, 2016). Essa discrepância persiste para o resto do ensino obrigatório, e ainda que menos acentuada, para o ensino superior. Estudos do Programa Internacional de Avaliação de Estudantes (PISA, 2018)

demonstram que o desempenho de alunos de 15 anos está correlacionado de maneira positiva ao tempo de exposição ao ensino infantil.

O presente artigo pretende investigar se essa evidência se mantém nos dados mais recentes do Brasil, para a pré-escola, contribuindo para preencher a lacuna entre as pesquisas internacionais e nacionais, corroborando Santos (2015) de que a relação entre ensino infantil e seu impacto é ambígua devido às discrepâncias de qualidade.

Espera-se que os resultados obtidos por Heckman (2006) também sejam congruentes com as conclusões deste estudo, assim, demonstrando que a educação pré-escolar representa uma estratégia eficiente para mitigar a pobreza e reduzir a desigualdade social, por meio desta, preparando crianças para um ensino mais aprofundado na continuidade de sua educação obrigatória.

## 1.1 EVIDÊNCIAS DA TEORIA DO CAPITAL HUMANO

O modelo do capital humano considera o ensino, a saúde e a capacitação profissional como pilares do desenvolvimento econômico das nações (Schultz, 1961). Outro ponto defendido pela tese é de que o produto gerado pelos trabalhadores é um fato determinante da renda individual e da renda nacional (Mincer, 1974). Assim, a renda individual e a nacional são causas do desenvolvimento econômico.

Mesmo sendo amplamente aceita em seus anos de concepção, críticas foram tecidas contra o modelo inicial: como a simplicidade com que a teoria ignora fatores sociais, econômicos e geográficos. Estes fatores tem impacto significativo para o acesso à educação e a capacitação profissional, assim como para o acesso à saúde, impacto este que acaba gerando uma distribuição desigual de capital e proporcionalmente de renda (Gordon, 1967).

Ademais, alguns especialistas apontam que essa teoria tem um embasamento selecionista, justificando o sucesso dos indivíduos unicamente a suas próprias escolhas e decisões. Assim, a teoria ignora fatores relevantes para a tomada de decisão dos indivíduos como, por exemplo, o contexto social e econômico no qual estes indivíduos estão inseridos, ademais, a educação se caracteriza por um processo influenciado por fatores como ambiente familiar e classe social (Bowles, 1971).

Concordando com esses autores, Jencks (1979) realiza uma análise em que a educação pode ser um fator relativamente pequeno na determinação da renda a qual apresenta potencial correlação com sorte, classe social e o ambiente familiar. As respostas trazidas pelos pensadores desta linha de raciocínio a estes questionamentos são vastas. Entretanto, grande parte dos fundadores reconhece que a teoria é reducionista e que

os fatores adicionais que afetam a produtividade e os salários dos trabalhadores são importantes, mas que não é possível negar a importância dos três baluartes da teoria para compreender este tópico (Schultz, 1979).

Existem diversas evidências que confirmam o pensamento apresentado na teoria e forma como os investimentos nessas áreas se relacionam com o crescimento econômico. Os resultados encontrados em Barros (1994) e Levy (1992), além de identificar importantes avanços dentro do pensamento apresentado da teoria, apontam a importância da educação nesse processo.

Nos dias atuais a teoria fornece uma explicação para o papel da educação, da saúde e do treinamento no crescimento econômico, porém, como todas as teorias econômicas é importante ressaltar suas limitações os fatores adicionais que afetam a produtividade e os salários dos trabalhadores. No estado atual da arte, investimentos nessas áreas são motores da evolução econômica. Além disso, quando esses investimentos são realizados de forma ágil e precisa, o resultado encontrado tem maior impacto (Heckman, 2008).

A ressonância dentro do pensamento teórico é presente e frequente autores trazem argumentos construtivos ao debate teórico como a afirmativa de que a qualidade da educação é mais importante do que a quantidade de educação (Hanushek, 2003). Sendo este um fator que pode aumentar os salários dos trabalhadores, segundo Card (1992). Em adesão, que o capital humano é um ingrediente essencial para o crescimento econômico, mas não é suficiente (Romer, 2018).

## 2 METODOLOGIA E ESTATÍSTICAS DESCRITIVAS

O *Propensity Score Matching* (PSM), ou Pareamento pelo Escore de Propensão em português, é uma metodologia estatística com o objetivo de reduzir o viés de seleção na comparação entre grupos de tratamento e controle. A próxima seção descreve essa metodologia.

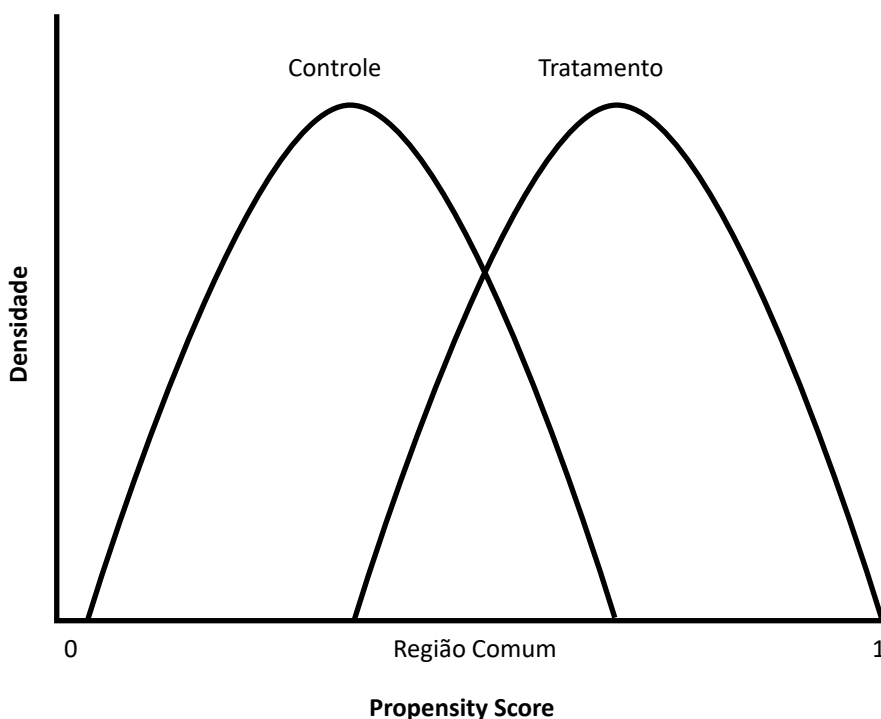
### 2.1 PROPENSITY SCORE MATCHING

A ideia central por trás do PSM é criar grupos de tratamento e controle que sejam semelhantes em todas as características observáveis, exceto na variável de tratamento. Isso é feito calculando o escore de propensão, o quanto o indivíduo tem chances de participar do tratamento, para cada um dos presentes em um conjunto de dados observacionais, este cálculo representa a probabilidade de pertencer ao grupo de tratamento com base em suas características observáveis (Rosenbaum; Rubin, 1983).

Ao que se refere ao grupo controle se faz importante ressaltar que ele seria a parcela da amostra a qual não está inserida no processo estudado, neste artigo a frequência na pré-escola. No que se refere ao grupo tratamento, seria a parcela da amostra a qual está inserida no processo estudado, neste artigo a frequência na pré-escola. O método de escore de propensão é relevante por garantir que a relação entre frequência na pré-escola e desempenho no ensino fundamental seja verdadeira em ambos os grupos. Isso ocorre pois o escore de propensão contém a mesma quantidade de informação que o vetor de características observáveis (Pinto, 2012).

Em resumo, caso o escore de propensão seja confiável e conhecido, traz consigo a possibilidade de estimar o efeito médio do tratamento sobre os tratados (EMTT), assim podendo parear indivíduos tratados e indivíduos não tratados (Pinto, 2012). Para entendimento do método, a figura 1 mostra a região de suporte comum existente entre os grupos de controle e tratamento.

FIGURA 1 — *Propensity Score Matching* (PSM) e a região comum



FONTE: Batista e Domingos (2017)

É possível estimar o efeito médio do tratamento para com os tratados (EMTT) ao realizar o pareamento entre os indivíduos tratados e os indivíduos controle utilizando como base os escores de propensão, mas há a necessidade que sejam estimados (Peixoto et al., 2012). Para a utilização de dados observáveis, são utilizados métodos de regressão e emparelhamento (Cameron, Trivedi, 2005).

## 2.2 MÍNIMOS QUADRADOS PONDERADOS PELOS ESCORES DE PROPENSÃO

O escore de propensão pode ser utilizado como peso, com intuito de balancear indivíduos no grupo controle, assim, tornando-os semelhantes à média do grupo de tratados, em resumo, são ponderados cada unidade do contrafactual pela probabilidade de não se receber tratamento (Cameron, Trivedi, 2005; Peixoto et al., 2012).

Desta maneira, este método é baseado na estimação do escore de propensão, sem a obrigatoriedade de estimar a função de regressão que relaciona os potenciais resultados com as variáveis observáveis, sendo este modelo sensível apenas à modelagem usada para o escore de propensão (Peixoto et al., 2012). Desta maneira, o modelo é duplamente robusto por proporcionar maior confiabilidade aos métodos, neste caso, especificando corretamente o modelo de propensão, a estimativa será consistente com o EMPT (Peixoto et al., 2012).

## 2.3 ESTATÍSTICAS DESCRITIVAS

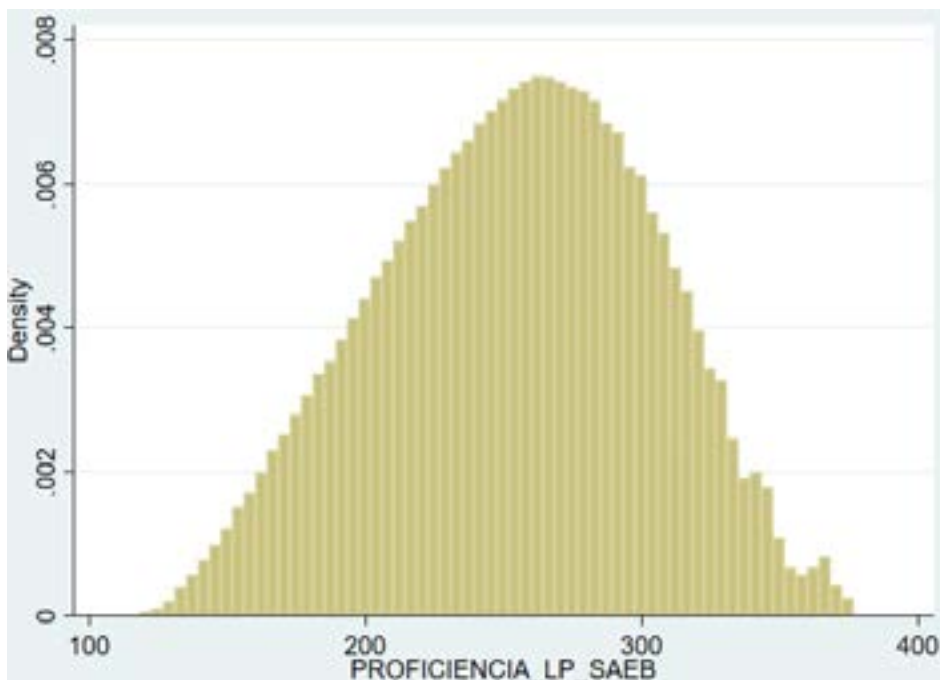
Os dados disponibilizados no SAEB (2021), pelo INEP, registram informações de 2.591.937 alunos matriculados em 39.365 escolas do território brasileiro. Essa pesquisa utilizará duas variáveis dependentes: i) nota de proficiência em língua portuguesa e ii) nota de proficiência em matemática, ambas configuradas segundo a escala do SAEB, que vai de 0 a 500, sendo, na prática, valores acima de 450 e abaixo de 50 muito raros.

A nota de proficiência em ambos os casos representa a capacidade do aluno em responder de maneira correta às questões de seu nível esperado de conhecimento. As variáveis da de resultado da base de dados são *proficiencia\_lp\_saeb* (nota, que consiste nas respostas do exame de proficiência em língua portuguesa) e *proficiencia\_mt\_saeb* (nota, que consiste nas respostas do exame de proficiência em matemática). Para compreender a distribuição de frequência das notas nessas três categorias, os gráficos 1 e 2 apresentam os histogramas das notas.

Nesses gráficos é possível perceber que as distribuições se comportam de maneira semelhante a uma distribuição normal. A média aproximada é de 254 pontos para a proficiência em português e 252 pontos para a proficiência em matemática. Tratando-se de ter iniciado os estudos diretamente na pré-escola, objeto deste estudo, 68,21% dos respondentes assinalaram que iniciaram os estudos na pré-escola. Quando se analisam as médias em português e matemática, para os alunos que iniciaram seus estudos na pré-escola, em comparação com aqueles que iniciaram no primário, observa-se que há uma diferença positiva para ambas disciplinas. Em português, a diferença é de 0,37 pontos. Em matemática de 1,06 pontos, como mostra o gráfico 3.

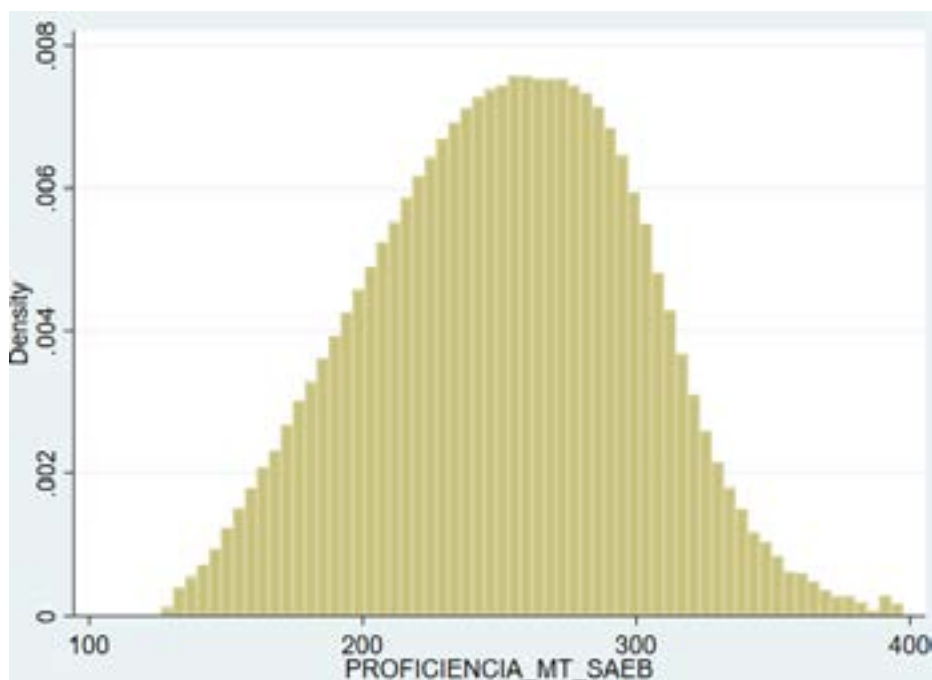


GRÁFICO 1 — Distribuição de frequência Língua Portuguesa SAEB 2



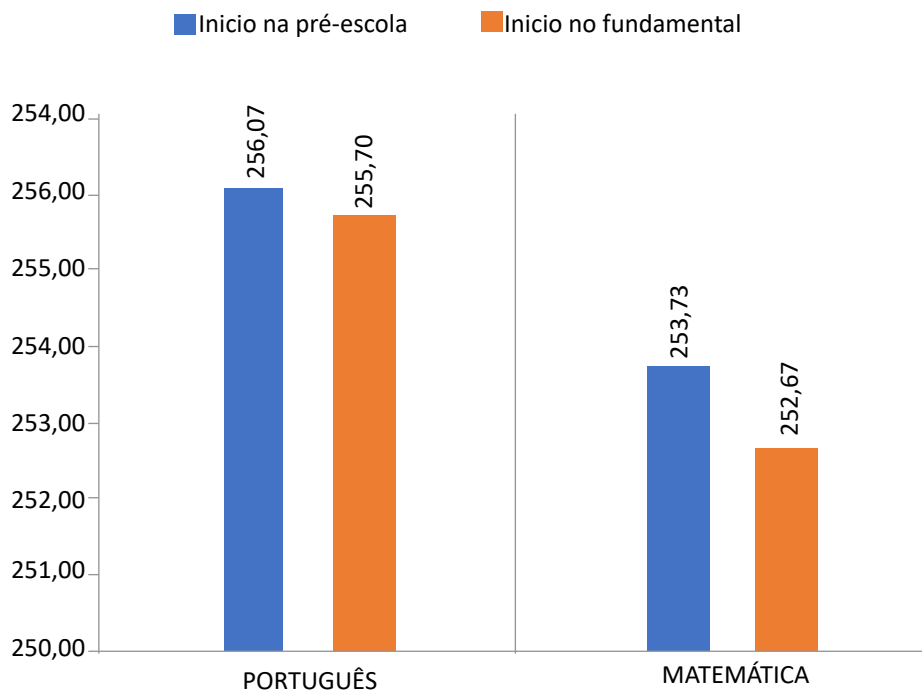
FONTE: SAEB (2021, adaptado)

GRÁFICO 2 — Distribuição de frequência Matemática - SAEB 2021



FONTE: SAEB (2021, adaptado)

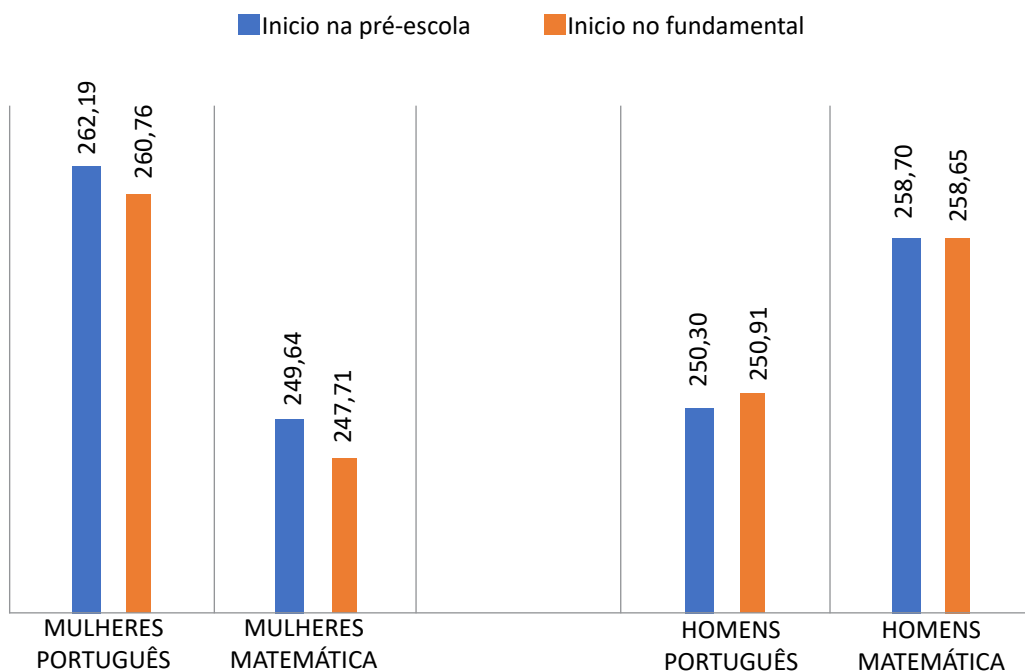
GRÁFICO 3 — Notas médias gerais



FONTE: SAEB (2021, adaptado)

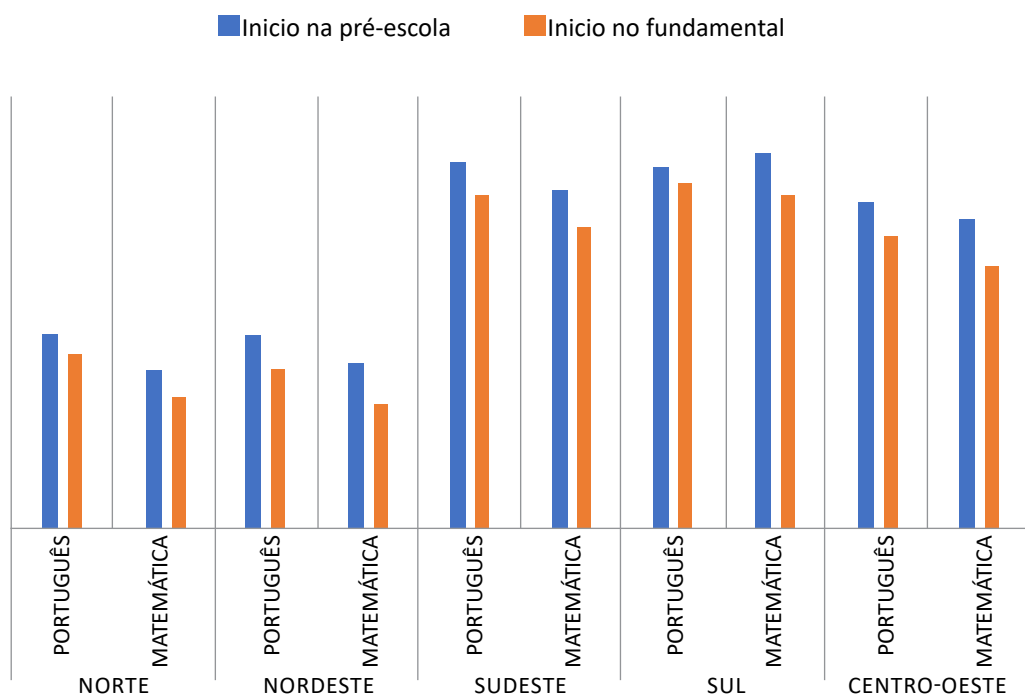
Entre os alunos do 9º ano do ensino fundamental, 50,4% se autodenominam do sexo feminino. O gráfico 4 mostra que, em média, pessoas do sexo feminino tendem a ter um desempenho superior em português e, em contrapartida, pessoas do sexo masculino, um desempenho superior em matemática. Analisando quanto à variável de tratamento, a saber, ter iniciado os estudos na pré-escola, em comparação com ter iniciado diretamente no primário, é possível observar que, no recorte por sexo, alunos que iniciaram na pré-escola têm um desempenho superior àqueles que não. A exceção aqui acontece para indivíduos do sexo masculino, na disciplina de língua portuguesa.

GRÁFICO 4 — Notas médias gerais por sexo



FONTE: SAEB (2021, adaptado)

GRÁFICO 5 — Notas médias gerais por macrorregião



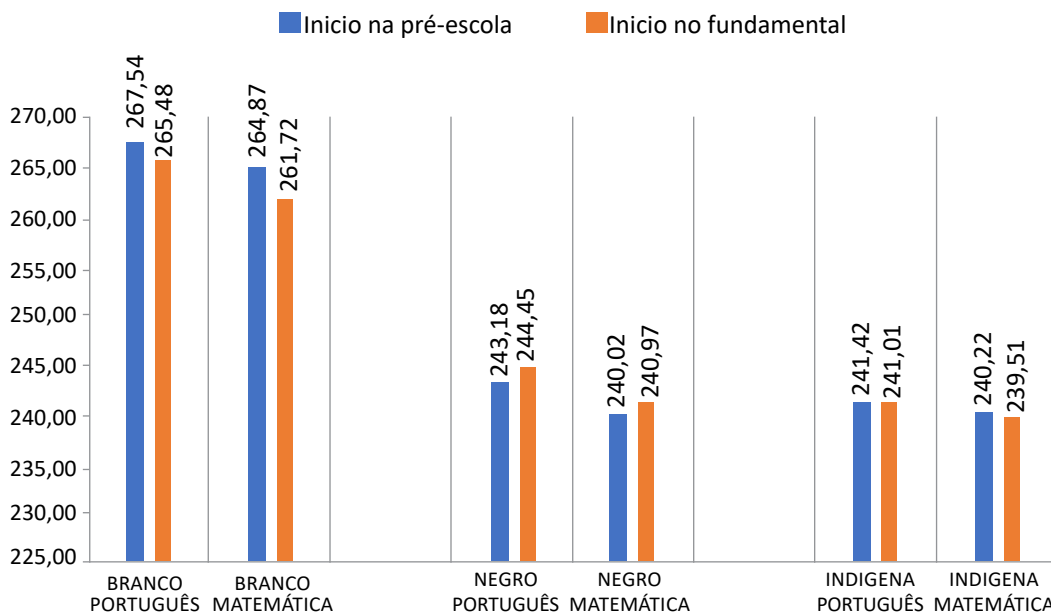
FONTE: SAEB (2021, adaptado)

Em se tratando das regiões brasileiras, as médias por tratamento e disciplina estão representadas no gráfico 5. Em todas as regiões os alunos que iniciaram os estudos escolares na pré-escola obtiveram médias superiores aos que iniciaram seus estudos no primário, obtendo a região Sudeste a maior nota média na proficiência em língua portuguesa. Em relação à proficiência em matemática o melhor desempenho médio foi registrado na região Sul.

Outra questão passível de análise, são as médias vistas pela ótica racial. Pessoas autodeclaradas como brancas e autodeclaradas como indígenas, que iniciaram seus estudos na pré-escola, possuem médias superiores em ambas as proficiências. Entretanto, pessoas autodeclaradas negras têm médias menores, em ambos os exames de proficiência, quando iniciaram seus estudos na pré-escola, o que pode ser uma evidência da falta de acesso de pessoas não brancas a um ensino infantil de qualidade, como pode ser verificado no gráfico 6.

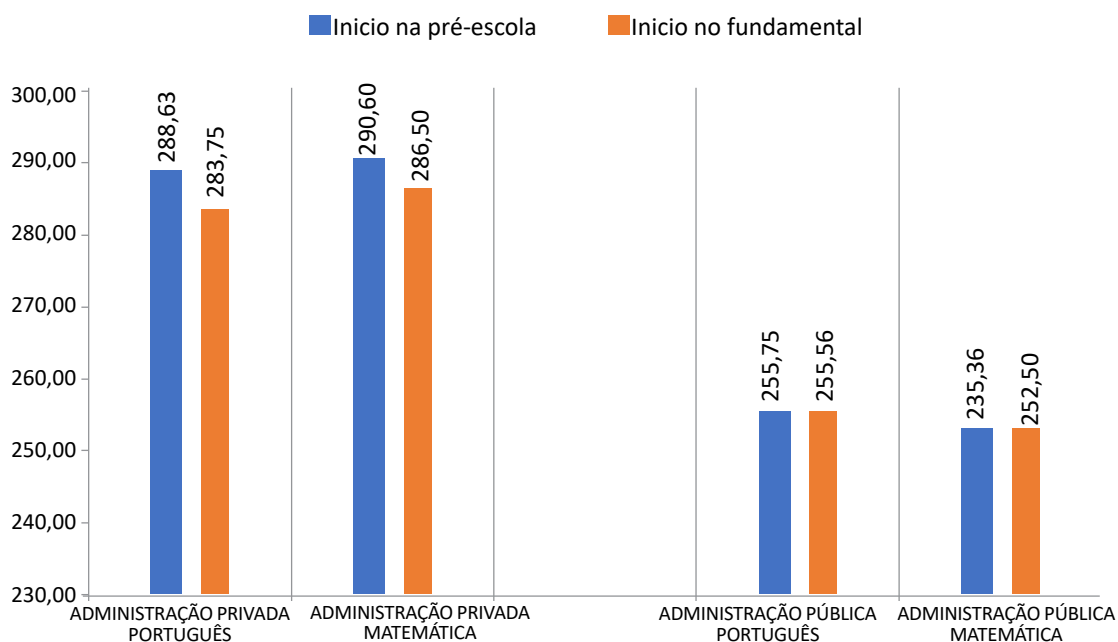
É possível analisar uma discrepância entre o modelo administrativo das instituições de ensino, escolas com administração privada obtém médias maiores nos exames de proficiência, tanto para alunos que iniciaram os estudos na pré-escola tanto para os que iniciaram no ensino fundamental. Tais resultados apontam a ineficiência do controle estatal para com a educação, a qual, como abordado na segunda seção deste trabalho tem o dever constitucional de prover uma educação de qualidade e para todos os usuários da educação pública, como mostra o gráfico 7.

GRÁFICO 6 — Notas médias gerais por etnia



FONTE: SAEB (2021, adaptado)

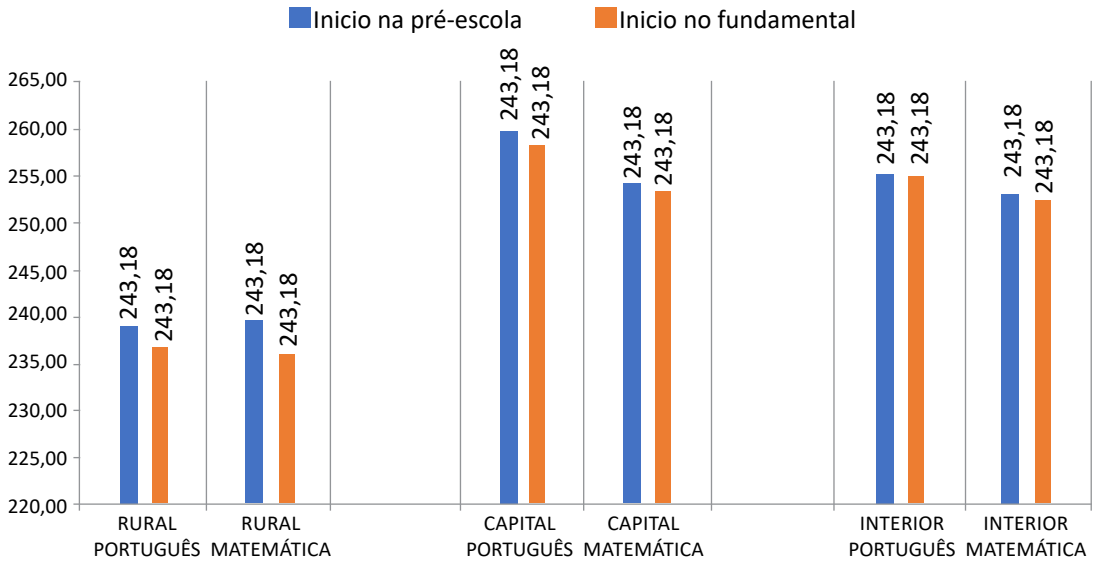
GRÁFICO 7 — Notas médias gerais por tipo administrativo



FONTE: SAEB (2021, adaptado)

Em relação às categorias regionais, classificam-se as escolas em rurais, interior ou capitais. De acordo com a documentação do INEP, escolas rurais podem ser consideradas tanto de administração pública como privada. O gráfico 8 mostra que escolas situadas nas capitais tendem a ter um desempenho maior, em média, para ambas as disciplinas. É possível observar também que, para as três categorias, alunos que iniciaram seus estudos na pré-escola têm um desempenho superior, em média.

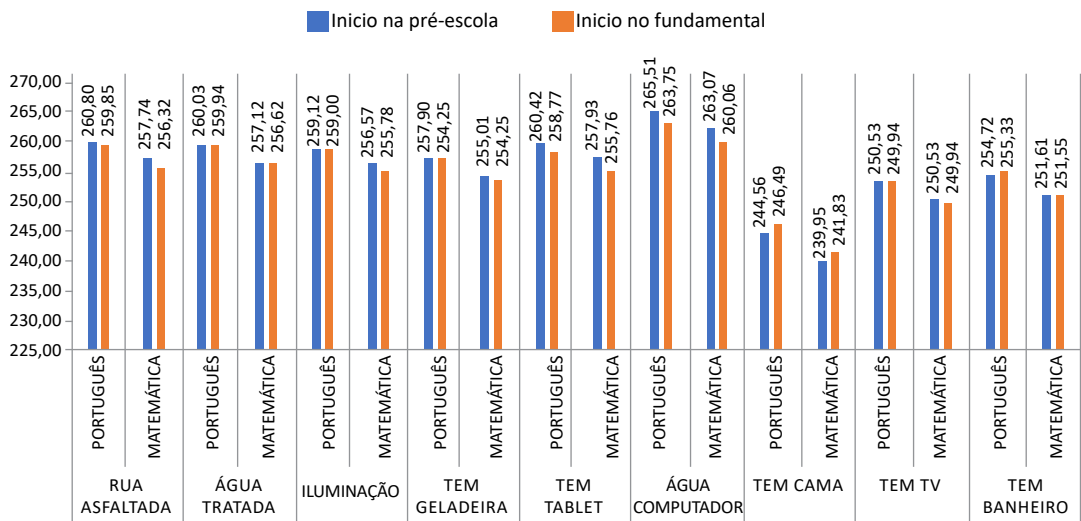
GRÁFICO 8 — Notas médias gerais por categorias regionais



FONTE: SAEB (2021, adaptado)

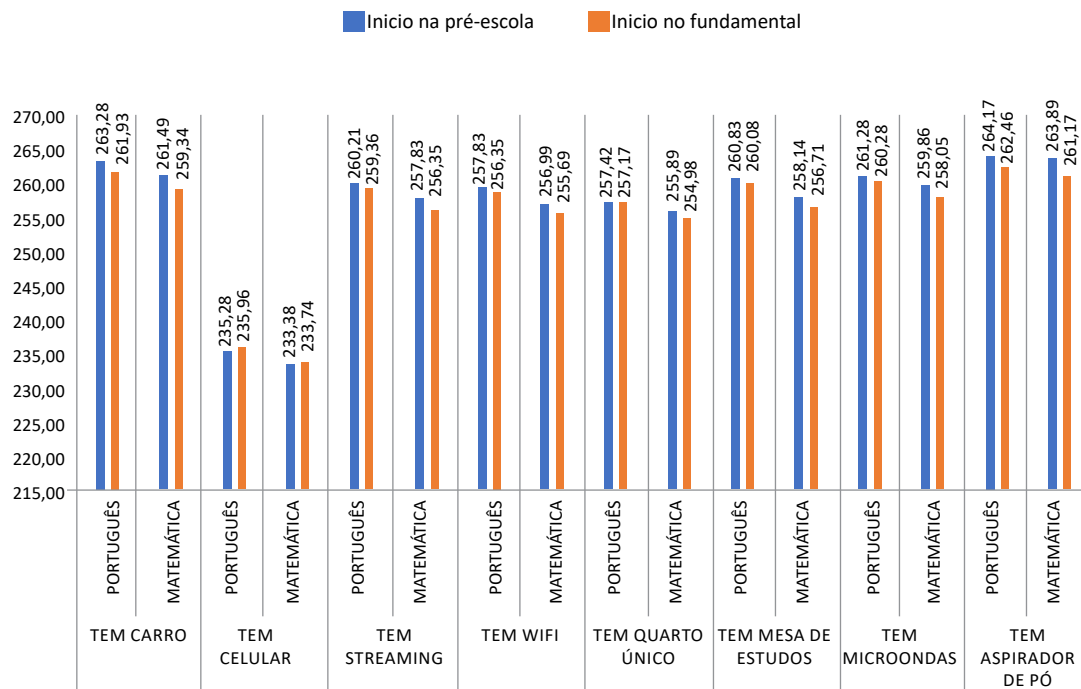
Em relação aos indicadores socioeconômicos – acesso à rua asfaltada, computador, automóvel e outros recursos, as médias são, de forma geral, maiores para os alunos que tiveram acesso a pré-escola, o gráfico 9 e 10 mostram que, para esses indicadores, alunos que iniciaram seus estudos na pré-escola têm um desempenho superior àqueles que iniciaram no primário.

GRÁFICO 9 — Notas médias gerais por características socioeconômicas parte 1



FONTE: SAEB (2021, adaptado)

GRÁFICO 10 — Notas médias gerais por características socioeconômicas parte 2



FONTE: SAEB (2021, adaptado)

Depois da análise dos grupos de estatísticas descritivas, é perceptível que o desempenho do grupo de tratamento, em sua maioria, é em média superior ao desempenho do grupo controle. Entretanto, é importante destacar que valores obtidos pela análise somente de valores médios é imprecisa não condizendo com o real impacto da pré-escola. Então, a metodologia presente neste artigo serve como uma maneira de mensurar o verdadeiro impacto desta política a qual não seria possível por meio de uma regressão linear comum.

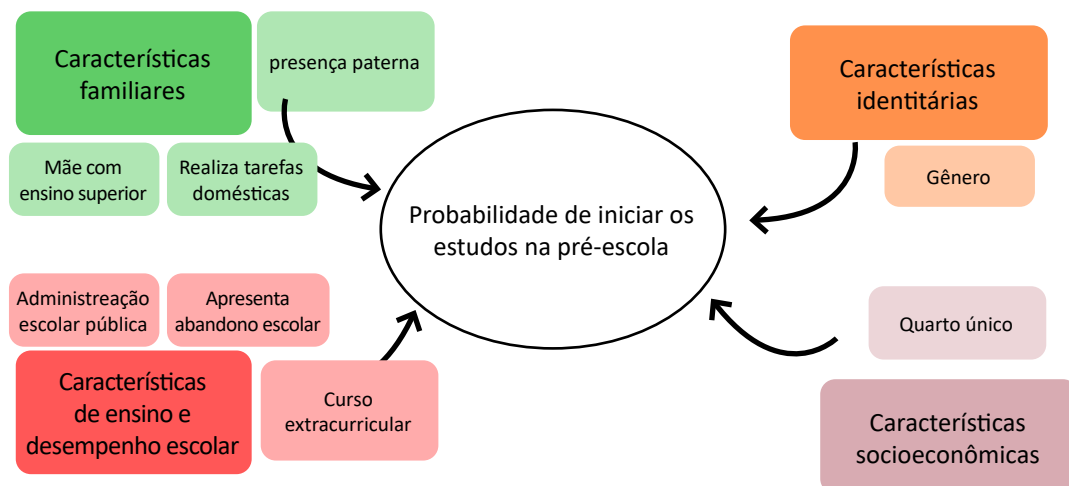
## 2.4 MODELO DE SELEÇÃO

Ao se iniciar a metodologia do PSM a estimação de um modelo de seleção é necessária. Esse pode ser estimado através de um modelo probit em que a variável dependente é binária (Gujarati; Porter, 2000). O modelo de seleção proposto por este trabalho observa que algumas das características do indivíduo têm maior importância que outras quando se refere à probabilidade de o aluno ter acesso à pré-escola ou não.

Os principais fatores que afetam o modelo de seleção para a probabilidade de frequentar a pré-escola foram identificados como sendo i) características familiares;

ii) características de ensino e desempenho escolar; iii) características identitárias e; iv) características socioeconômicas. A identificação desses desses fatores através das variáveis encontradas na base de dados é descrita através da figura 2.

FIGURA 2 — Diagrama das variáveis utilizadas no modelo de seleção



FONTE: Os autores (2024)

As características familiares estão ligadas com a probabilidade de iniciar os estudos na pré-escola, por isso a presença paterna, o acesso da mãe ao ensino superior e o auxílio nas tarefas domésticas estão ligados a este fator. As características de ensino e desempenho escolar estão ligadas ao cotidiano do aluno na realização de atividades acadêmicas e de ensino. É importante ressaltar que se espera uma probabilidade menor de iniciar os estudos na pré-escola, devido ao fato de que, como explanado na seção 2, a evasão escolar é um fenômeno complexo e que está relacionado ao contexto no qual o indivíduo está inserido.

As condições socioeconômicas são entendidas por Alves, Soares e Xavier (2014) como um constructo teórico que aloca indivíduos em classes. Neste estudo, ter um quarto único representa uma visão positiva socioeconômica por representar que o lar do estudante tem um espaço consideravelmente maior, ou seja, uma condição socioeconômica positiva de renda. Por fim, espera-se que características socioeconômicas positivas aumentem a probabilidade de iniciar os estudos na pré-escola.

As características identitárias, em resumo, fazem parte das condições socioeconômicas, porém, se diferenciam do grupo descrito acima por não serem consideradas “positivas” ou “negativas”. No caso deste estudo, a variável de autodeclaração de gênero entre indivíduos autodeclarados meninos ou meninas identifica este fator.



Por estes motivos, as variáveis escolhidas para a seleção do desempenho dos alunos nas provas de proficiência em língua portuguesa e matemática que possuíram ou não acesso à pré- escola foram i) presença paterna; ii) mãe com ensino superior; iii) realiza tarefas domésticas; iv) administração escolar pública; v) apresenta abandono escolar; vi) realiza curso extracurricular; vii) gênero e; viii) apresenta quarto único.

A tabela a seguir apresenta os resultados da estimação por modelo Probit de notas gerais, no qual o grupo tratamento é a “ter iniciado os estudos na pré-escola”. Este modelo, de forma geral, calcula a probabilidade de um aluno ser selecionado para o grupo de tratamento (ter iniciado os estudos na pré-escola) utilizando variáveis observáveis ligadas às características familiares, características de ensino e desempenho escolar, características identitárias e características socioeconômicas.

TABELA 1 — Modelo de escolha binária probit para oportunidade de ter iniciado os estudos na pré-escola

<i>Variáveis de seleção/Grupo de controle</i>	<i>Coefficientes da regressão</i>
<i>Sexo</i>	0,0066
<i>mãe com superior</i>	0,0677***
<i>tem abandono escolar</i>	-,0536***
<i>presença paterna</i>	0,0266***
<i>administração pública</i>	-,3482***
<i>realiza curso extracurricular</i>	0,0343***
<i>quarto único</i>	0,0549***
<i>realiza tarefa domestica</i>	-,0537***
<i>Número de observações</i>	1.099.128
<i>Número de Tratados</i>	749.673
<i>Número de Controles</i>	349.455

FONTE: Enade (2019, adaptado)

Nota: \* significante a 1%. \*\*significante a 5%. \*\*\*significante a 10%

A tabela 1 mostra que o modelo Probit apresentado tem significância de 1% para a maioria das variáveis com exceção da variável sexo. O modelo indica que a administração pública, o abandono escolar e a necessidade de realizar tarefas domésticas possuem probabilidade negativa de serem selecionadas no grupo tratamento. Ademais, a necessidade de realizar tarefa doméstica indica que, por necessidades externas, o aluno tem menos chances de ter iniciado seus estudos na pré-escola. Por fim, alunos com abandono escolar têm também probabilidade negativa, pois o abandono está intrinsecamente ligado ao contexto social no qual estão inseridos.

Em adição, o modelo indica que características familiares positivas aumentam a probabilidade de ter iniciado os estudos na pré-escola e que a realização de cursos

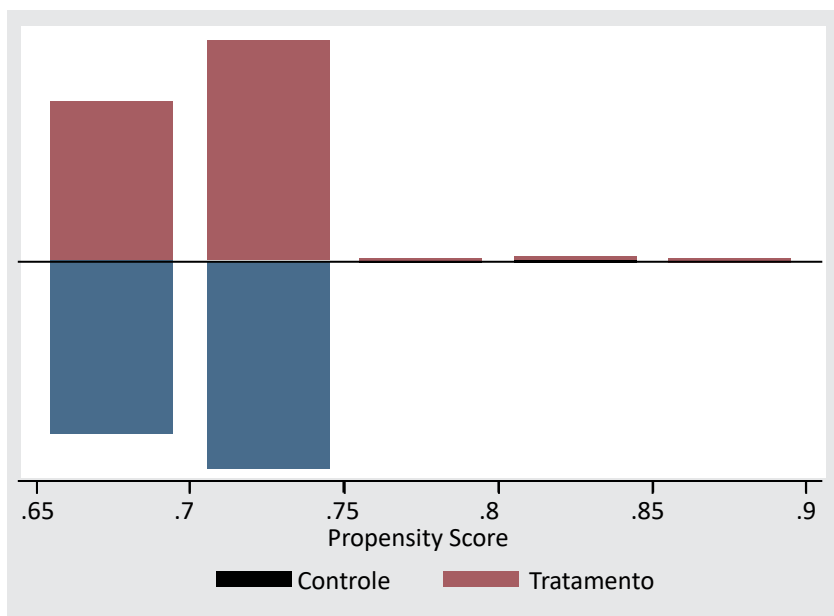
extracurriculares sugere a valorização do estudo por parte dos pais deste indivíduo, assim aumentando suas chances de ser selecionado para iniciar os estudos na pré-escola.

### 3 ESTIMAÇÃO E ANÁLISE DE RESULTADOS

O método para estimação do Efeito Médio do Tratamento sobre os Tratados (EMPT) utilizado foi o stratification matching, que realiza a divisão dos intervalos de variação do escore de propensão de maneira em que cada intervalo, as unidades de tratamento e controle tenham em média, os mesmos escores de propensão (Cameron, Trivedi, 2005). O modelo de Mínimos Quadrados ponderados pelos Escores de Propensão, duplamente robusto, foi utilizado com as variáveis descritas na seção anterior para que possam controlar o desempenho nos testes de proficiência.

Para que seja assegurado que o pareamento tenha qualidade, o gráfico 11 permite identificar quantidades relevantes tanto no grupo controle quanto para o grupo dos tratados. Para facilitar a interpretação, na estimação do Efeito Médio do Tratamento sobre os Tratados (EMPT), foram utilizados os logaritmos naturais das variáveis proficiencia\_lp\_saeb e proficiencia\_mt\_saeb.

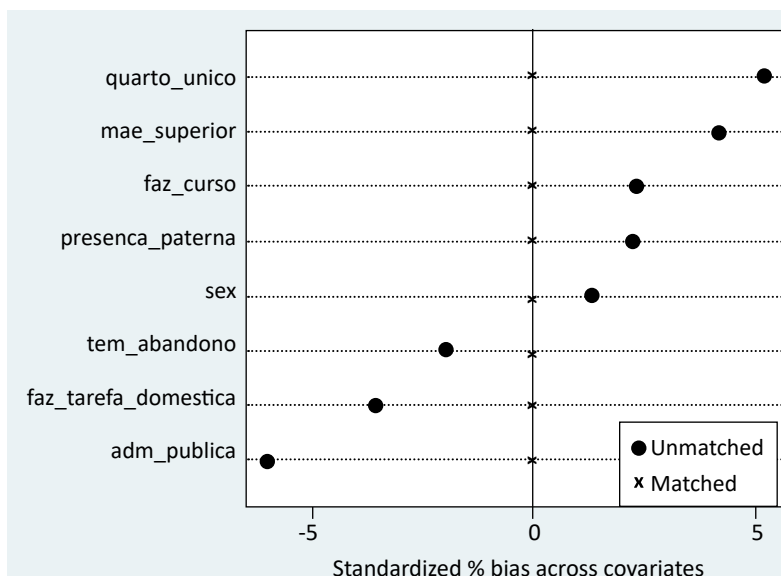
GRÁFICO 11 — Suporte comum entre tratados e controles



FONTE: Os autores (2024)

O gráfico 12 mostra a redução do viés nas variáveis de controle antes e após o pareamento. O gráfico mostra que o viés é aproximadamente zero para todas as variáveis aos o pareamento.

GRÁFICO 12 — Percentual de viés nas variáveis de controle antes e depois do pareamento



FONTE: Os autores (2024)

Por fim, a Tabela 2 apresenta os resultados da avaliação de impacto para ambos os métodos, MQP e EMPT.

TABELA 2 — Resultados da avaliação de impacto (MQP e EMPT)

Tratamento	Iniciou os estudos na pré-escola	
Controle	Presença paterna, Mãe com ensino superior, Realiza tarefas domésticas, Administração escolar pública, Apresentar abandono escolar, Realiza curso extracurricular, gênero e apresentar quarto único.	
Método	MQP	EMPT
Nota exame de proficiência Língua portuguesa	0,0009957 (5,512)	0,000 (0,475)
Nota exame de proficiência matemática	0,00043313*** (5,516)	0,003*** (3,626)
N° de controles	-	80.679
N° tratados	-	188.970

FONTE: Enade (2019, adaptado)

Nota: teste t entre parênteses. \* significativa a 1%. \*\* significativa a 5%. \*\*\*significante a 10%

Em ambos os métodos, os resultados do impacto na nota do exame para proficiência em língua portuguesa não foram estatisticamente significantes. Esse resultado, apesar de divergir dos resultados da literatura, podem ser uma evidência dos efeitos desiguais da pandemia de Covid-19. É possível também que este efeito esteja capturando o impacto da Lei 12.796 de 5 de abril de 2013, que ampliou o ensino

fundamental em um ano considerando o que antes era a pré-escola como o primeiro ano, corroborando os efeitos não significantes da creche (Silva Junior; Gonçalves, 2016).

Para a nota no exame de proficiência em matemática, os resultados foram positivos e com significância de 1%. Quando utilizado o MQP, há um acréscimo de 0,04% para aqueles que iniciaram seus estudos na pré-escola. Pelo método EMPT, essa diferença é numericamente maior, com um aumento de 0,3% no desempenho do aluno quando este iniciou seus estudos na pré-escola.

TABELA 3 — Resultados regressão por mínimos quadrados ponderados

Variáveis independentes	Logaritmo da Nota em Português	Logaritmo da Nota em Matemática
<i>Sexo</i>	-0,0035***	0,0560***
<i>mãe com superior</i>	0,0727***	0,0703***
<i>tem abandono escolar</i>	-0,0516***	-0,0521***
<i>presença paterna</i>	0,0202***	0,0246***
<i>administração pública</i>	-0,098***	-.1115***
<i>realiza curso extracurricular</i>	-0,0144***	-0,0117***
<i>quarto único</i>	0,0028***	0,0056***
<i>realiza tarefa domestica</i>	0,0692***	0,0591***
<i>Constante</i>	5,560***	5,5218***
<i>Observações</i>	269.649	269.649
<i>R-squared</i>	0.0545	0,0553

FONTE: Enade (2019, adaptado)

Nota: teste t entre parênteses. \* significante a 1%. \*\* significante a 5%. \*\*\*significante a 10%.

A análise dos coeficientes das variáveis independentes fornece insights importantes sobre a influência dessas variáveis sobre o desempenho dos alunos. Primeiramente, observa-se que o sexo tem uma influência significativa nas duas variáveis dependentes. Para o desempenho em língua portuguesa o coeficiente é -0,003, indicando que o sexo masculino está associado a uma ligeira diminuição no desempenho. No que se refere ao desempenho em matemática, o coeficiente é 0,06, sugerindo que o sexo masculino tem uma influência positiva nesta variável.

O abandono escolar, como esperado, tem uma influência negativa e significativa em ambas as variáveis dependentes, com coeficientes de aproximadamente -0,052 em ambas as disciplinas. Quanto ao modelo de administração, alunos que estudaram em escolas com administração pública estão associados a desempenhos inferiores, com coeficientes de -0,098 para português e -0,1115 para matemática. Realizar cursos extracurriculares está associado a uma pequena redução em ambas as variáveis dependentes. Embora a influência seja pequena, ela é estatisticamente significativa diferenciando dos resultados encontrados no modelo probit.

Ter uma mãe com ensino superior está associado a aumentos tanto em português quanto em matemática, sugerindo que a educação da mãe influencia positivamente o desempenho acadêmico dos filhos. A presença paterna é, também, uma variável com influência positiva e significativa no desempenho. Ter um quarto único também apresenta uma influência positiva sobre o desempenho, embora com pequena magnitude. Isso indica que essa condição está associada a um aumento no desempenho acadêmico, possivelmente gerado por situações socioeconômicas positivas ao incentivo ao estudo. Realizar tarefas domésticas está associado a um aumento no desempenho acadêmico, diferenciando-se dos resultados encontrados no modelo probit, assim evidenciando uma diferença estatística nos resultados quando utilizado um modelo mais robusto.

Em resumo, a tabela 3 evidencia a influência de diversos fatores no desempenho acadêmico dos estudantes, destacando a importância da escolaridade dos pais, presença paterna, e a realização de tarefas domésticas, enquanto fatores como abandono escolar e o modelo de administração público têm influências negativas.

## CONSIDERAÇÕES FINAIS

Este estudo buscou estimar o impacto da pré-escola no desempenho escolar dos alunos brasileiros. Para garantir a robustez do resultado, foi utilizado o método de pareamento, Propensity Score Matching (PSM), para que fossem pareados grupos de controle e tratamento. E para controlar as características de desempenho dos alunos participantes do SAEB 2021 foi utilizado o método de Mínimos Quadrados Ponderados Pelos Escores de Propensão (MQP).

Alguns avanços no que se refere ao ensino infantil, são muito relevantes. A Lei 12.796 de 5 de abril de 2013 ampliou o ensino fundamental em um ano, considerando o que antes era a pré-escola como o primeiro ano. Segundo dados da OECD (2016), o Brasil investe em educação uma parcela de 4,2% do PIB, superior à média de 3,2% dos países desenvolvidos. No entanto, o valor investido por aluno é menos da metade desses países, de 3,8 mil dólares frente a 8,6 mil dólares para os países da OECD.

Os resultados deste trabalho mostram que ter iniciado os estudos na pré-escola tem um efeito positivo e estatisticamente significativo para a proficiência em matemática. O impacto estimado utilizando o método dos Mínimos Quadrados Ponderados pelo Score de Propensão foi de 0,04%. Ainda que pequenos, esses os resultados apresentam significância estatística de 1%. Estimando o Efeito Médio do Tratamento sobre os Tratados (EMTT) utilizando o método *stratification*, o impacto estimado é de 0,3% no desempenho do aluno quando este iniciou seus estudos na pré-escola. Para a proficiência em língua portuguesa, os resultados não encontraram significância estatística.

Esses resultados, ainda que divirjam dos resultados da literatura, podem ser uma evidência dos efeitos desiguais da pandemia de Covid-19, que levou à hibridização ou digitalização do ensino. É possível, também, que este efeito esteja capturando o impacto da Lei 12.796 de 5 de abril de 2013, que ampliou o ensino fundamental em um ano considerando o que antes era a pré-escola como o primeiro ano, corroborando os efeitos não significantes da creche.

Além da avaliação de impacto, este trabalho apresentou a influência dos diversos fatores sobre o desempenho dos alunos. Características como a mãe ter ensino superior, presença paterna, ter um quarto único e realizar tarefas domésticas estão associadas positivamente ao desempenho dos alunos em português e matemática. Características como ter apresentado abandono escolar, ter estudado em escolas públicas, realizar cursos extracurriculares estão associadas a um desempenho inferior. Quanto ao sexo, meninos apresentam, em média, notas superiores em matemática enquanto meninas possuem melhor média em língua portuguesa.

Este trabalho sugere que estudos futuros possam investigar de maneira mais profunda os efeitos dos anos pandêmicos no ensino brasileiro, especialmente para aquelas crianças em desvantagem, e avaliem de maneira mais robusta, a questão da qualidade escolar para o que hoje se tornou a pré-escola, após a implementação da Lei 12.796 de 5 de abril de 2013.

## REFERÊNCIAS

- ALVES, M. T. G.; SOARES, J. F.; XAVIER, F. P. Índice socioeconômico das escolas de educação básica brasileiras. **Ensaio: Avaliação e Políticas Públicas em Educação**, v. 22, n. 84, p. 671-703, set. 2014.
- BARRO, R. J.; LEE, J.-W. Sources of economic growth. **Carnegie-Rochester Conference Series on Public Policy**, Elsevier, v. 40, n. 1, p. 1-46, 1994.
- BATISTA, M.; DOMINGOS, A. Mais que boas intenções: técnicas quantitativas e qualitativas na avaliação de impacto de políticas públicas. **Revista Brasileira de Ciências Sociais**, v. 32, n. 94, p. 1, 2017.
- BLUESTONE, B.; HARRISON, B. **The deindustrialization of America**: plant closings, community abandonment, and the dismantling of basic industry. [s.l.]: Basic Books, 1982.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Brasil no Pisa 2018**. Brasília: Inep, 2020.
- BRASIL. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). **Resumo técnico: censo escolar da educação básica 2021**. Brasília: Inep, 2021.
- CAMERON, A. C.; TRIVEDI, P. K. **Microeconometrics: methods and applications**. Cambridge: Cambridge University Press, 2005.
- CARD, D. The impact of the minimum wage on employment and earnings of teenagers: new evidence from a pre-employment training program. **American Economic Review**, v. 82, n. 3, p. 747-759, 1992.
- EDWARDS, R. **Contested terrain: the transformation of the workplace in the twentieth century**. [s.l.]: [s.n.], 1979.
- GUJARATI, D. N.; PORTER, D. C. **Econometria básica**. 5. ed. Porto Alegre: AMGH, 2011.
- HANUSHEK, E. A. Schooling, teacher quality, and the economic returns to education. **American Economic Review**, v. 93, n. 2, p. 1184-1208, 2003.
- HECKMAN, J. J. Investing in disadvantaged young children is an economically efficient policy. Nova York: Forum on “Building the economic case for investing in preschool”, 2006.
- HECKMAN, J. J. School, skills and synapses. **NBER working paper**, v. 46, n. 3, p. 289-324, jul. 2008.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA — IBGE. **Abandono escolar: atlas geográfico escolar**. Disponível em: <https://atlascolar.ibge.gov.br/brasil/indicadores-sociais/educacao/21978-abandono-escolar>. Acesso em: 27 jun. 2024.
- INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA — IBGE. **Pesquisa Nacional por Amostra de Domicílios Contínua - PNAD Contínua: Educação 2023**. Rio de Janeiro: IBGE, 2024. Disponível em: <https://loja.ibge.gov.br/pnad-continua-educac-o-2023.html>. Acesso em: 20 jun. 2024.
- LEVY, F.; MURNANE, R. J. U.S. earnings levels and earnings inequality: a review of recent trends and proposed explanations. **Journal of Economic Literature, American Economic Association**, v. 30, n. 3, p. 1333-1381, set. 1992.

MARCHÃO, A. J.; FITAS, A. C. P. A avaliação da aprendizagem na educação pré-escolar: o portfólio da criança. **Revista Iberoamericana de Educación**, n. 64, p. 27-41, 2014.

OECD. **Education at a Glance 2016**: OECD Indicators. Paris: OECD Publishing, 2016. <https://doi.org/10.1787/eag-2016-en>

REICH, M.; GORDON, D. M.; EDWARDS, R. C. Segmented work, divided workers: The historical transformation of labor in the United States. Cambridge: Cambridge University Press, 1973.

ROMER, P. M. The economic growth process: lessons from endogenous growth theory. **Nobel Prize Lecture**, 8 out. 2018. Disponível em: <https://www.nobelprize.org/prizes/economic-sciences/2018/romer/lecture/>. Acesso em: 27 set. 2023.

ROSENBAUM, P. R.; RUBIN, D. The central role of the propensity score in observational studies for causal effects. **Biometrika**, v. 70, n. 1, p. 41-55, 1983.

SANTOS, D. D. **Impactos do ensino infantil sobre o aprendizado**: benefícios positivos, mas desiguais. 2015. 260 f. Tese (Livre Docência em Microeconomia) — Universidade de São Paulo, Ribeirão Preto, 2015.

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

SILVA JUNIOR, W. S.; GONÇALVES, F. O. Evidências da relação entre a frequência no ensino infantil e o desempenho dos alunos do ensino fundamental público no Brasil. **Revista Brasileira de Estudos de População**, v. 33, n. 2, p. 283-301, 2016. <https://doi.org/10.20947/S0102-30982016a0015>

VYGOTSKY, L. S. **Pensamento e linguagem**. São Paulo: Martins Fontes, 1989.