

# A FORMA DE INGRESSO IMPORTA? UMA ANÁLISE DO DESEMPENHO DOS ALUNOS DO ENSINO MÉDIO NA REDE FEDERAL

Lauana Rossetto Lazaretti<sup>1</sup>  
Marco Túlio Aniceto França<sup>2</sup>

**Resumo:** Este trabalho tem por objetivo verificar se há efeito seleção no desempenho dos alunos que ingressam na Rede Federal de ensino. Os dados utilizados são referentes aos estudantes do último ano do ensino médio, disponíveis no Sistema de Avaliação da Educação Básica (Saeb), para as disciplinas de matemática e língua portuguesa. Como cada unidade federal de ensino possui independência administrativa no método de seleção, a estratégia de identificação se baseia na designação dos alunos em dois grupos: tratamento (ingresso por sorteio) e controle (com base em notas). Para estimar o efeito são utilizados três métodos de pareamento: o Propensity Score Matching (PSM), o *Coarsened Exact Matching* (CEM) e o balanceamento por entropia. Os resultados empíricos indicam que há um efeito entre 0,14 e 0,40 desvios-padrão em matemática, decorrência do processo seleção, com diferença de gênero e na distribuição de notas. Entre as meninas o efeito foi maior em língua portuguesa e para os meninos em matemática. Os estudantes dos quantis inferiores da distribuição de notas possuem um efeito seleção maior. Como possível explicação para o resultado encontrado, destaca-se a influência dos pares. Por fim, os resultados são robustos a diferentes especificações e métodos, bem como se mantêm em exame para o ingresso no ensino superior.

**Palavras-Chave:** Efeito seleção. Ensino Médio. Desempenho no Saeb.

**Abstract:** This paper aims to verify whether there is a selection effect on the performance of students entering the Federal Education System. The data used refer to students in the last year of secondary school, available in the Basic Education Assessment System (Saeb), for the subjects of Mathematics and Portuguese. As each federal education unit has administrative independence in the method of selection, the identification strategy is based on assigning students into two groups: treatment (entrance by lot) and control (based on grades). To estimate the effect, three matching methods are used: Propensity Score Matching (PSM), Coarsened Exact Matching (CEM) and entropy balancing. The empirical results indicate that there is an effect between 0.14 and 0.40 standard deviation in mathematics, resulting from the selection process, with a difference in gender and in the distribution of grades. Among girls, the effect was greater in Portuguese and for male students in mathematics. Students in the lower quantiles of the grade distribution have a greater selection effect. As a possible explanation for the result found, the influence of peers stands out. Finally, the results are robust to different specifications and methods, as well as remaining under examination for admission to graduation.

**Keywords:** Selection Effect. Secondary School. Performance on Saeb.

**Área ANPEC:** Economia do Setor Público.

**JEL Code:** C31; D04; I21.

---

<sup>1</sup> Doutoranda do Programa de Pós-Graduação em Economia do Desenvolvimento da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

<sup>2</sup> Professor do Programa de Pós-Graduação em Economia do Desenvolvimento da Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS).

## 1. Introdução

A literatura empírica acerca do entendimento dos diferentes resultados educacionais entre as escolas ganhou maior relevância a partir do relatório Coleman, publicado em 1965. Uma semelhança encontrada nos estudos que abrangem esta temática está no fato de que o processo de seleção ou a autoseleção dos alunos são mecanismos importantes para explicar esses diferenciais, na medida em que ocorre ganhos de efeito de pares e segregação das escolas. Desta forma, um grande esforço vem sendo realizado para entender o que de fato se deve ao efeito escola e o que é decorrência do conjunto de características dos alunos que cada escola atende.

O fato de o aluno estar matriculado em determinada escola decorre das políticas utilizadas no sistema educacional e da escolha de escola realizada pelos pais. A preferência dos pais, os quais diferem em *background* socioeconômico, é refletida em características observáveis e não observáveis dos alunos. Assim, em regiões com maior opção de escolha escolar é comum haver maior segregação e estratificação do sistema de ensino (ALLEN, 2007; IVANIUSHINA *et al.*, 2018) e, conseqüentemente, a diferença de desempenho entre as escolas tende a ser mais acentuada (KIM, 2018; PIOPIUNIK, 2014). O mesmo ocorre quando há seleção de alunos para o ingresso nas escolas (EPPLÉ; NEWLON; ROMANO, 2002; HANUSHEK; WÖSSMANN, 2006) ou agrupamento de alunos por habilidades (FRYER; LEVITT, 2006; HANUSHEK; KAIN; RIVKIN, 2009; HANUSHEK; RIVKIN, 2008; ROTHSTEIN, 2015; RUHOSE; SCHWERDT, 2016), visto que alunos com menor habilidade inicial podem ficar em situação pior.

Adicional às evidências de escolha de escola e seleção de alunos, parte dos mecanismos que explicam os melhores resultados são encontrados na literatura de efeito de pares. Para Lazear (2001), há dois pontos fundamentais de organização da turma e efeito de pares: (1) a segregação de estudantes por habilidade acadêmica (*tracking*) maximiza os resultados educacionais, em contraponto, estratifica as escolas e aumenta a desigualdade; e (2) sem *tracking* de alunos, a integração tende a contribuir para que ocorra efeito de pares, ou seja, os alunos com maior *background* podem ajudar os demais a aumentar o desempenho. Porém, em termos empíricos, as evidências não são claras. Para Carreell, Sacerdote e West (2013), a heterogeneidade dentro dos grupos tende a formar subgrupos dentro da turma e isso dificulta a interação e integração entre os alunos. Quando ocorre o agrupamento de alunos por habilidades, os resultados tendem a ser distintos: de um lado, Cummins, 2017 e Epple, Newlon e Romano (2002) consideram que as turmas formadas pelos alunos da cauda inferior de notas perdem em termos de desempenho; de outro, Duflo, Dupas e Kremer (2011) encontraram que, com ajuda e direcionamento do professor, o agrupamento por habilidades contribui positivamente para o desempenho de alunos em diferentes níveis de habilidade.

Em meio a tais aspectos, percebe-se que as características dos alunos que as escolas atendem possuem papel fundamental nos indicadores de desempenho. Não obstante, a literatura de *school sector effects* tem procurado entender quanto do desempenho de determinadas escolas se deve ao chamado “efeito escola” e quanto está relacionado as características dos alunos atendidos (ANGRIST; BETTINGER; KREMER, 2006; BAUDE *et al.*, 2020; BENEVIDES; SOARES, 2014; CHUDGAR; QUIN, 2012; HANUSHEK *et al.*, 2007; HOWELL *et al.*, 2002; HOXBY; ROCKOFF, 2004; LUBIENSKI; LUBIENSKI, 2006; MCEWAN, 2004; SAPELLI; VIAL, 2002; SOMERS; MCEWAN; WILLMS, 2004). E, de fato, quando controladas as características endógenas aos alunos, os resultados têm se mostrado menor que o esperado ou insignificantes. Ou seja, parte dos resultados educacionais refletem as características do grupo de alunos atendidos, conhecido como “efeito seleção”.

Desta forma, o processo de seleção se torna um fator importante para explicar as diferenças de desempenho entre as escolas e isso contribui com a discussão acerca da qualidade da Rede Federal de Ensino (RFE) no Brasil. Em sua primeira avaliação em larga escala e comparável entre as escolas que ofertam o ensino médio, o Sistema de Avaliação da Educação Básica (Saeb) de 2017 aponta diferença de desempenho entre as redes de ensino. Os dados mais recentes, de 2019, apontam que a RFE possui desempenho superior as demais escolas que compõem a rede pública (estadual e municipal) e privada. A média de pontos em Língua Portuguesa (LP) da RFE é de 335, da rede privada é 321,85 e do restante da

rede pública é 271,08<sup>3</sup>. A diferença foi ainda maior para Matemática (MT), sendo que a RFE atingiu 356,13 pontos, enquanto as redes privada e pública pontuaram 334,45 e 267,76, respectivamente.

Em paralelo a esse contexto, na medida em que a RFE se destaca nas avaliações de desempenho, surgem discussões com respeito a federalização de toda a educação básica. O pioneiro da ideia é o Senador e Professor Cristovam Buarque, que ampliou o debate acerca do tema, ao defender a ampliação de escolas federais por todo o país. Um exemplo disso é o projeto de lei para a federalização da educação básica<sup>4</sup>, que possui como objetivo aumentar a qualidade do ensino, em razão de as escolas federais serem referência de modernos equipamentos de ensino, da valorização da remuneração do corpo docente e de qualidade. Ou seja, a diferença de desempenho entre as redes tem sido elencada na discussão argumentativa/normativa.

Porém, dentro desta perspectiva, pouco se tem abordado a existência de um processo de seleção dos alunos na rede federal. Isto posto, surge o questionamento: melhores escolas ou melhores alunos? Nem todo o desempenho pode ser atribuído ao efeito escola, uma vez que o *background* estudantil dos alunos que ingressam na RFE é diferente dos demais alunos que permanecem na rede pública. Em outras palavras, parte do desempenho da RFE pode estar relacionado com o processo de seleção dos alunos.

Neste contexto, o presente estudo busca trazer evidências empíricas sobre qual é a magnitude do “efeito seleção” dos alunos que ingressam na Rede Federal de ensino sobre o desempenho no último ano do ensino médio. A estratégia de identificação se baseia na autonomia das unidades de ensino no processo de seleção e, por conseguinte, de ingresso dos alunos. Na presença de duas formas de ingresso, o processo seletivo e o sorteio, os alunos são designados em dois grupos: tratamento e controle. A ideia é que todos os estudantes tenham realizado a sua inscrição (escolha de escola) e condicional em características observáveis, a diferença de desempenho entre os grupos pode ser atribuída ao processo de seleção.

O estudo preenche uma lacuna importante da discussão argumentativa acerca do desempenho superior da RFE. Além disso, no que tange ao conhecimento dos autores, não foram encontrados outros estudos que medem efeito de seleção no ensino médio e no ensino fundamental também é mais comum investigar a existência de efeito escola. Assim sendo, o presente estudo é o primeiro a trazer evidências empíricas para o nível de ensino médio no Brasil. Cabe destacar que, além da qualidade e gratuidade, outras duas características entram na função utilidade da escolha de escola dos pais: (1) ao estudar em escolas públicas, posteriormente, os alunos podem concorrer a cotas na educação superior; e (2) ganhos de efeito de pares - em turmas mais homogêneas (no caso de processo seletivo), se espera ganhos de desempenho.

Além desta introdução, o artigo está dividido em mais cinco seções. A segunda seção trata do contexto da rede federal de ensino médio no Brasil. A terceira seção resume a literatura que investiga os diferentes resultados educacionais entre as escolas e destaca o efeito seleção. Na sequência, a estratégia empírica e a amostra de dados são descritas. Na quinta seção, os resultados são apresentados e discutidos. Por fim, as conclusões finais são elencadas.

## 2. Background da Rede Federal de ensino médio no Brasil

A Rede Federal de ensino básico no Brasil é composta por 38 unidades de Institutos Federais de Educação (IFs), pela Universidade Tecnológica Federal do Paraná (UTFPR), por dois centros federais de educação tecnológicas (Cefet<sup>5</sup>), por 23 escolas técnicas vinculadas às Universidades Federais e pelo Colégio Pedro II. A estrutura de ensino federal atual decorre, principalmente, do Plano de Expansão da Rede Federal de Educação Profissional e Tecnológica, com dois principais marcos legais. O primeiro advém da Lei nº 11.195 de 18 de novembro de 2005, que propõe a expansão da oferta de educação profissional. No segundo, a partir da Lei nº 11.892 de 29 de dezembro de 2008, foram criados os Institutos Federais de Educação, Ciência e Tecnologia, que, além de novas instituições de ensino, outras escolas da

---

<sup>3</sup> A rede municipal atende menos de 1% da população de estudantes. Quando analisadas separadamente, a média da rede estadual é de 271 e da municipal é 285,29. Então, para fins de simplificação, usaremos o termo rede pública para a rede estadual e municipal.

<sup>4</sup> A PLS 337/2016 foi apresentada pelo senador Cristovam Buarque (PPS-DF) em audiência pública federativa em 27 de novembro de 2018. (Agência Senado, 2018) acesso em: <https://www12.senado.leg.br/noticias/materias/2018/11/23/federalizacao-da-educacao-basica-sera-tema-de-audiencia>.

<sup>5</sup> Centros Federais de Educação Tecnológica Celso Suckow da Fonseca - CEFET-RJ e de Minas Gerais - CEFET-MG.

rede federal passaram a ser assim denominadas. Ao todo, foram criadas 500 novas unidades de ensino distribuídas por todas as regiões do país. Em 2006, havia 144 unidades em funcionamento e, em 2018, chegou a 643 (MEC, 2021).

Por ofertar educação profissional integrada ao ensino médio, a rede federal passa a dividir a responsabilidade da última etapa da educação básica com as redes de ensino estadual e privada. Em 2020, segundo o Censo da Educação Básica no Brasil, 3% do total de matrículas no ensino médio pertencia à rede federal, 84% à rede estadual e 13% à rede privada. Desta forma, ainda que tenha ocorrido uma ampliação na oferta de vagas na rede federal, as matrículas são limitadas a capacidade de atendimento das unidades de ensino e não garante o atendimento de todos os demandantes.

A rede federal difere em valor de investimento por aluno e na formação acadêmica dos professores, o que lhes proporciona remuneração mais elevada. O gasto corrente anual por matrícula da rede federal em 2019 era de R\$15.741,98, enquanto para o restante da rede pública um aluno no ensino médio urbano em tempo integral custa R\$6.564,00 em 2020 (PNP, 2020; SIMCAQ, 2021). O quadro de professores é formado, em sua maioria, por mestres e doutores, 51,39% e 34,09%, respectivamente.

Dada a opção de escolha escolar por uma unidade da rede federal de educação, para ingressar, o aluno passa por uma etapa de seleção. Para a qual, em virtude da Lei nº 11.892 de 29 de dezembro de 2008 – que prevê autonomia das entidades integrantes da Rede Federal de Educação – há a possibilidade de a unidade de ensino adotar diferentes formas de ingresso. Por meio dos Censos Escolares de 2013 a 2018, é possível identificar qual a forma de ingresso dos alunos da rede federal<sup>6</sup>. A mais comum – que representa 86% das matrículas em 2018 – é o exame de seleção por meio de provas de conhecimento ou por análise do histórico escolar da etapa cursada anteriormente (no caso do ensino médio, as notas do ensino fundamental nos anos finais). Em menor proporção, os alunos que ingressam por meio de sorteio representam 1,5% do total de alunos no ensino médio integrado - cerca de 2,7 mil alunos. O restante é advindo de transferências.

As unidades que utilizam como forma de seleção o sorteio das vagas pertencem aos Institutos Federais de Brasília (IFB), do Acre (IFAC), de Santa Catarina (IFSC), do Rio Grande do Sul (IFRS) e do Paraná (IFPR). Entre as instituições de ensino, a forma de seleção pode ocorrer de forma uniforme entre todos os *campi* ou diferir entre eles ou dentro deles. No IFB a única forma de seleção é o sorteio, no IFAC o sorteio foi realizado nos processos seletivos de 2010 a 2015 e nos IFSC, IFRS e IFPR alguns campus ou cursos que entraram em funcionamento a partir de 2016 adotaram o sorteio como forma de seleção dos alunos. Por fim, algumas unidades de ensino adotam o sorteio para vagas remanescentes.

### 3. Estratégia Empírica

Com base no cenário de independência da RFE para a adoção do método de seleção dos alunos, a estratégia de identificação se baseia na designação dos alunos em dois grupos. O grupo de tratamento é formado pelos alunos que fizeram a inscrição e foram selecionados por meio de um sorteio. O grupo de controle também realizou a inscrição, mas para cursar o ensino médio na rede federal passou por algum tipo de seleção.

Essa estratégia possui vantagens e desvantagens. O ponto positivo é minimizar o problema de escolha de escola, a qual representa um dos principais fatores de viés de não observáveis. Consequentemente, independente da forma de ingresso para ser selecionado, o aluno realizou a inscrição, o que sinaliza que a rede federal está no leque de opção escolar. Além disso, por comparar alunos de apenas uma rede, é garantida a homogeneidade da estrutura de ensino. Por outro lado, cabe destacar que os alunos analisados são aqueles que estão no último ano do ensino médio e é necessário ter em mente que a amostra representa os “ganhadores”. Logo, os alunos que efetivamente não conseguiram avançar de ano não estão na amostra e, com isso, ressaltamos o cuidado na interpretação dos resultados.

---

<sup>6</sup> A variável de identificação é dividida em nove categorias de ingresso: (1) Sem processo seletivo; (2) Sorteio; (3) Transferência; (4) Exame de seleção sem reserva de vaga; (5) Exame de seleção, vaga reservada para alunos da rede pública de ensino; (6) Exame de seleção, vaga reservada para alunos da rede pública de ensino, com baixa renda e autodeclarado preto, pardo ou indígena; (7) Exame de seleção, vaga reservada para outros programas de ação afirmativa; (8) Outra forma de ingresso; e (9) Exame de seleção, vaga reservada para alunos da rede pública de ensino, com baixa renda. Cabe destacar que a partir de 2019 a variável foi descontinuada, a última informação sobre o processo de seleção é a do ano de 2018.

Ainda assim, os alunos que compõem os dois grupos – tratamento e controle – tendem a diferir em termos de características observáveis. Na ausência de uma alocação aleatória entre os dois grupos, a simples especificação de um modelo via Mínimos Quadrados Ordinários (MQO) leva a uma estimação inconsistente, visto que pode haver um termo de erro ( $\varepsilon$ ) correlacionado com a variável de interesse ( $T$ ) – neste estudo, ingresso por meio de sorteio ou seleção –, ou seja,  $cov(T, \varepsilon) \neq 0$ .

Neste caso, para estimar um parâmetro consistente é necessário encontrar uma técnica que comporte o problema elencado na estimação via MQO. Entre as técnicas possíveis para resolver, a utilização de um balanceamento das características observáveis é uma ferramenta bastante utilizada na literatura empírica. Neste caso, na existência de seleção condicional às características observáveis, assume-se que o tratamento é aleatório.

Postas tais ressalvas, as próximas seções formalizam o modelo em termos de resultados potenciais e apresentam três estimadores para a realização do balanceamento das características observáveis entre os grupos de tratamento e de controle. O primeiro, amplamente usado na literatura a partir do texto seminal de Rosenbaum e Rubin (1983), é o *Propensity Score Matching* (PSM). O PSM consiste no cálculo da probabilidade do estudante em receber o tratamento condicional às suas características observáveis. O segundo método utilizado é o *Coarsened Exact Matching* (CEM) e o terceiro é o balanceamento por entropia. Ambos buscam reduzir as imperfeições de *matching* por escore de propensão. O CEM avança no modelo de *matching* exato e forma grupos de estratos similares. O balanceamento por entropia faz um pré-tratamento dos dados para reduzir o desequilíbrio das covariáveis entre os grupos de tratamento e de controle e, por consequência, aumenta a independência da variável de tratamento. A utilização dos três métodos visa apresentar resultados robustos quanto ao balanceamento das características observáveis e, conseqüentemente, maior precisão na determinação do efeito causal.

### 3.1 O Modelo de Resultados Potenciais

Ao realizar o teste de desempenho padronizado (Saeb) no terceiro ano do ensino médio, os estudantes da Rede Federal obtêm uma nota  $Y$ . Assim, denota-se que  $y(1)$  é o resultado potencial do estudante que ingressou por meio de sorteio – tratamento – e  $y(0)$  é o resultado potencial do mesmo estudante caso tenha entrado por meio de prova – não tratado. Então, o resultado observado do estudante é:  $y_i = T_i y_i(1) + (1 - T_i) y_i(0)$ . Em que,  $T_i = 1$  indica que o estudante recebeu tratamento e  $T_i = 0$  que não recebeu.

Para identificar o efeito do tratamento para cada estudante é necessário subtrair o resultado potencial das duas possibilidades:  $\tau_i = y_i(1) - y_i(0)$ . No entanto, o estudante apenas é passível de um dos eventos: receber ou não receber o tratamento. Se  $T_i = 1$ , então  $y(0)$  não é observado, da mesma forma, se  $T_i = 0$ , então  $y(1)$  também não é observado. Ou seja, há um problema de dados faltantes. Generalizando entre os grupos, o Efeito Médio do Tratamento sobre os Tratados (*Average Treatment Effect on Treated* – ATT) é:  $ATT = E[y(1)|T = 1] - E[y(0)|T = 1]$ . De igual importância, é necessário encontrar um grupo que representa o contrafactual dos estudantes tratados na ausência de ter recebido o tratamento. Quanto mais próximo esse grupo de controle for do contrafactual dos estudantes, menor é o viés de medida na identificação do efeito causal.

Para representar o contrafactual, utilizamos os estimadores de *matching* para tornar os estudantes mais próximos em características observáveis ( $X$ ) a fim de identificar o efeito do tratamento sobre os tratados.

### 3.2 Pareamento por Escore de Propensão (PSM)

Dado que o  $Y_i(0)$  é observado apenas entre os indivíduos que não são tratados e a designação ao tratamento não é aleatória, Rosenbaum e Rubin (1983) propõem dois pressupostos para garantir que o ATT seja identificado.

A primeira hipótese é chamada de seleção em observáveis ou independência condicional. A ideia é que, condicional as variáveis contidas no vetor de características observáveis  $X$ , o tratamento seja aleatório entre os dois grupos – tratamento e controle. Ou seja, os resultados potenciais são independentes do tratamento (Equação 1).

$$Y_i(0), Y_i(1) \perp T | X_i \quad (1)$$

Além disso, é necessário que os estudantes do grupo de tratamento possuam um par no grupo de controle que represente o caso dele não ter recebido o tratamento. O que implica na segunda hipótese do modelo, a condição de sobreposição ou suporte comum (Equação 2).

$$0 < Pr [T = 1|X] < 1 \quad (2)$$

Isso garante que estudantes com as mesmas características possuam uma probabilidade positiva de ser tratado ou não. Rosenbaum e Rubin (1983) também avançam na dificuldade dimensional de pareamento exato. Quanto mais características incluídas no vetor  $X$ , menor é a chance de encontrar um par compatível as características dos indivíduos tratados. Para resolver, por generalização, a hipótese de seleção em observáveis também é válida quando condicional a probabilidade de receber o tratamento (Equação 3).

$$Y_i(0) \perp T|X_i \Rightarrow Y_i(0) \perp T|p(X_i) \quad (3)$$

Em que,  $p(X_i)$  é o escore de propensão, uma função que representa a probabilidade de receber o tratamento a partir do conjunto de características do vetor  $X$ . Na prática,  $p(X_i)$  é estimado por modelos de probabilidade linear – probit ou logit.

Com isso, para Caliendo e Kopeinig (2008), o estimador do ATT no PSM é a diferença de médias entre os tratados e controles situados dentro do suporte comum e ponderados pela distribuição do escore de propensão (Equação 4).

$$ATT_{PSM} = E_{p(X)|T=1}\{E[Y(1)|T = 1, p(X)] - E[Y(0)|T = 0, p(X)]\} \quad (4)$$

Para definir quais observações fazem parte do grupo de controle, é possível escolher uma medida de proximidade do escore de propensão, como o de  $n$  vizinhos mais próximo, *caliper* e raio ou a matriz *Kernel*. Para este estudo, utilizamos as diversas métricas para a estimação dos resultados, bem como é utilizado o método duplamente robusto com o *Inverse Probability Treatment Weighting* (IPTW) proposto por Emsley, Lunt e Pickles (2008) e Lunceford e Davidian (2004). A técnica duplamente robusta possibilita a estimação com erro padrão robusto e clusterizado, visto que no PSM isso não é possível.

Embora consolidado na literatura, o PSM é suscetível de críticas quanto a grandeza de características não observáveis que podem gerar viés no resultado. Para testar se o modelo está sendo prejudicado por tal problema, a próxima seção apresenta uma técnica de análise de sensibilidade.

### 3.2.1 Análise de sensibilidade

Com o pressuposto de independência condicional na realização do PSM, a ideia é que as variáveis que determinam a condição de tratado estejam contempladas no vetor de características observáveis  $X$ . De acordo com Rosenbaum (2002), formalmente, a probabilidade de receber o tratamento ( $\pi_j$ ) é uma função  $\lambda(x_j)$  desse vetor de características observáveis de cada  $j$  indivíduo. Assim, se dois indivíduos ( $j$  e  $k$ ) possuem as mesmas covariadas ( $x_j=x_k$ ), mas a probabilidade de receber o tratamento é diferente entre eles ( $\pi_j \neq \pi_k$ ), há indício de viés – pode ser que características não observáveis estejam afetando o tratamento.

Embora não seja possível testar a hipótese de independência condicional, Rosenbaum (2002) propôs uma análise de sensibilidade dos resultados. A ideia é buscar entender qual seria o tamanho dessa diferença para alterar as conclusões do estudo.

Dado que as chances de receber o tratamento é  $\pi_j/(1 - \pi_j)$  e ela é válida também para o indivíduo  $k$ , a razão das chances entre  $j$  e  $k$  é um número  $\Gamma \geq 1$  – Equação (5).

$$\frac{1}{\Gamma} \leq \frac{\pi_j(1-\pi_k)}{\pi_k(1-\pi_j)} \leq \Gamma \text{ para todo } j \text{ e } k \text{ com } x_j=x_k \quad (5)$$

Quando as chances de receber o tratamento são iguais,  $\Gamma = 1$  e não há viés. Se  $\Gamma > 1$ , por exemplo,  $\Gamma = 2$ , embora os indivíduos sejam similares em características observáveis, um deles apresenta duas vezes mais chances de receber o tratamento. No último caso, há viés. Logo,  $\Gamma$  representa o grau de afastamento de um resultado sem viés. Para vários valores de  $\Gamma$ , Rosenbaum (2002) aponta que, se há sensibilidade dos resultados para valores muito próximos de  $\Gamma = 1$ , a validade de que não há outras características não observáveis ( $u$ ) que estejam afetando o tratamento pode ser questionada.

### 3.3 Coarsened Exact Matching – CEM

Da mesma forma que no PSM, o *Coarsened Exact Matching* (CEM) possui como pressuposto a independência do tratamento condicional em observáveis:  $P\{T|X, y(0), y(1)\} = P(T|X)$ . Neste caso, a designação do tratamento independe dos resultados potências quando condicionados em  $X$ .

Diferente do PSM que calcula a probabilidade de cada unidade de análise ser tratada, a ideia do CEM é inserir um intervalo (*bins*) entre os valores das variáveis do vetor  $X$  para determinar grupos substancialmente significativos (BLACKWELL *et al.*, 2009; IACUS; KING, 2008; IACUS; KING; PORRO, 2011). Além disso, o CEM possibilita que a escolha de variáveis com *matching* exato – por exemplo, variáveis categóricas, como sexo ou cor/raça. O algoritmo CEM é conhecido pela facilidade de entendimento e de utilização, aplicável para grandes amostras de dados. Em suma, os procedimentos do algoritmo são: 1) fazer uma cópia ( $X^*$ ) das variáveis contidas no vetor  $X$ ; 2) Realizar os agrupamentos com os cortes definidos ou gerados automaticamente pelo algoritmo; 3) Criar um estrato para cada observação  $X^*$  e aloca cada observação em um estrato; e 4) Os estratos são atribuídos para os dados originais ( $X$ ) e são excluídos os estratos que não possuem pelo menos um indivíduo do grupo de tratamento e um do grupo de controle. O ATT pode ser gerado automaticamente pela diferença de médias entre os grupos ou por uma regressão linear, como também o *matching* por CEM pode ser utilizado em conjunto com outros métodos a posteriori.

Como a definição do *bin* é realizada *ex ante* ao *matching*, uma vantagem do CEM é reduzir o viés de desbalanceamento a priori, pois, em grande parte dos métodos de *matching* os testes para verificar o ajustamento é realizado *ex post*. Este atributo é denominado *Monotonic Imbalance Bounding* (MIB), ou seja, o desequilíbrio máximo de alguma característica empírica é limitado antes da realização do *matching*.

### 3.4 Balanceamento por Entropia

Ainda com a ideia de aprimorar o balanceamento, no caso da entropia, o objetivo é tornar a variável de tratamento mais independente (ortogonal) das características dos indivíduos. Para isto, é realizado um pré-tratamento entre as variáveis, de modo que as características das unidades de análise sejam reponderadas na função de peso para garantir um melhor ajustamento dos dois grupos – tratamento e controle.

Com base no mesmo modelo de resultados potenciais, formalmente escrito por Hainmueller (2012), o balanceamento por entropia repondera o grupo de controle (Equação 6).

$$E[y(0)|T = 1] = \frac{\sum_{\{i|T=0\}} Y_i \omega_i}{\sum_{\{i|T=0\}} \omega_i} \quad (6)$$

Em que  $\omega_i$  é o peso escolhido para cada unidade de controle. E a escolha do peso para a reponderação é o  $\min_{\omega_i} H(\omega) = \sum_{\{i|T=0\}} h(\omega_i)$ , sendo  $h(\cdot)$  uma métrica de distância, condicional a menor distância entre as observações,

Em resumo, isto significa que, a partir do conjunto de características do grupo de tratamento, são criados vários momentos da distribuição da amostra e com a reponderação do grupo de controle, a densidade de  $X$  dos dois grupos se torna muito próxima. Logo, diferente do PSM que, primeiro, estima pesos unitários com base em uma distribuição probabilística e, na sequência, o ajustamento é verificado, os pesos são estimados diretamente levando em consideração os momentos da amostra.

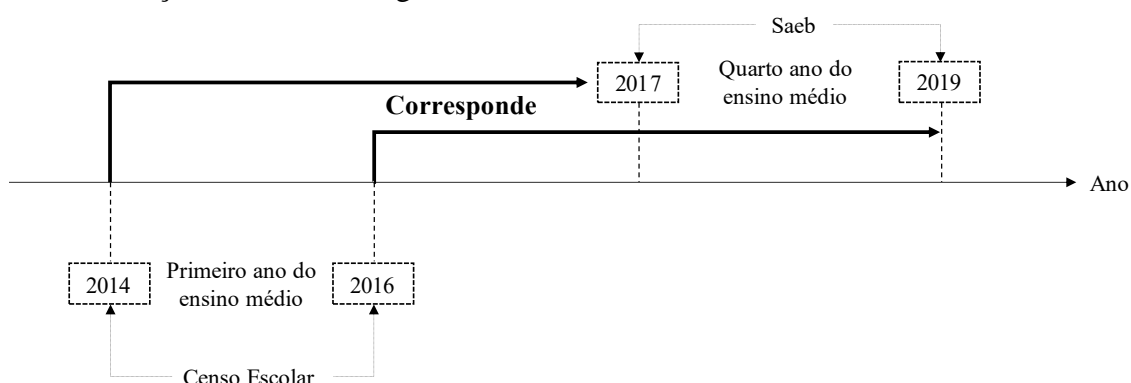
### 3.5 Amostra e Base de Dados

Os dados do desempenho dos alunos brasileiros foram extraídos do Sistema de Avaliação da Educação Básica (Saeb). A pesquisa é uma avaliação em larga escala aplicada a cada dois anos. A partir de 2017, os alunos do terceiro e do quarto ano do ensino médio passaram a ser avaliados nas disciplinas de língua portuguesa e de matemática. Além do teste de conhecimento, o aluno responde a um questionário com informações sobre a sua formação, escolaridade dos pais, características socioeconômicas e perfil demográfico. Adicionalmente, para identificar a forma de seleção de cada unidade de ensino, o Censo Escolar da Educação Básica é utilizado – realizado anualmente e de forma censitária.

Como os alunos da rede federal cursam o ensino médio integrado a um curso técnico, os alunos da análise levam quatro anos para a conclusão da última etapa do ensino básico, enquanto os alunos em cursos

propedêuticos concluem em três anos. Logo, para identificar qual o método de ingresso dos alunos, é necessário verificar qual o processo de seleção que cada unidade de ensino adotava no período de quatro anos anterior à análise (Figura 1).

Figura 1 – Identificação da forma de ingresso dos alunos.



Fonte: Elaborada pelos autores.

A coorte de alunos que realizam a prova do Saeb em 2017 e em 2019 ingressaram no primeiro ano do ensino médio em 2014 e em 2016, respectivamente. Assim, pelo código da unidade de ensino é possível identificar a forma de seleção utilizada e, subsequentemente, a coorte de alunos que ingressaram por meio do sorteio ou do exame de seleção.

De um total de 25.017 alunos que realizam o Saeb no quarto ano do ensino médio da rede federal, 5.638 não possuem identificação da escola e entram apenas no cálculo das médias estadual e nacional<sup>7</sup>. Além disso, para manter um grupo de comparação sem incorrer de haver outros alunos entrantes por sorteio, foram excluídas da amostra aquelas escolas que utilizam o sorteio para vagas remanescentes. Com isso, a amostra final contempla 348 alunos que ingressaram por meio do sorteio e 18.637 entrantes a partir do exame de seleção.

As características da amostra dos alunos tratados e controles antes e depois da realização do pareamento são apresentadas na Tabela 1. As diferenças com valores negativos indicam que os alunos que ingressaram por meio do sorteio possuem valores maiores que os alunos do processo seletivo, e o inverso é válido no caso de valores positivos.

Tabela 1 - Estatísticas descritivas

Variável	Pareamento	Média		Diferença	t-test
		Tratamento	Controle		
Cor Branca	Não	0,23	0,34	0,11***	-3,97
	Sim	0,25	0,26	0,01	-0,28
Cor Preta	Não	0,12	0,09	-0,03*	1,79
	Sim	0,25	0,26	0,01	1,15
Cor Parda	Não	0,49	0,37	-0,12***	4,67
	Sim	0,53	0,55	0,02	-0,65
Cor Amarela	Não	0,04	0,02	-0,02***	2,62
	Sim	0,04	0,04	0,00	0,21
Mãe com Ensino Médio	Não	0,27	0,31	0,04	-1,50
	Sim	0,30	0,31	0,01	-0,27
Mãe com Ensino Superior	Não	0,31	0,26	-0,05**	2,18
	Sim	0,33	0,30	-0,03	0,79

<sup>7</sup> A partir da Portaria Ministerial n.º 931, de 21 de março de 2005, o Saeb passou a avaliar de forma censitária as escolas que possuem no mínimo 30 estudantes matriculados no final de cada etapa de ensino – no último ano do ensino fundamental anos iniciais e anos finais e, desde 2017, no último ano do ensino médio. Além disso, por meio da Lei n.º 13.005, de junho de 2014, os resultados são divulgados para as escolas que possuem uma taxa de presença de 80% dos estudantes matriculados e a datar a Portaria Inep n.º 366, de 29 de abril de 2019, ainda é incluído obedecer cumulativamente ao registro de no mínimo 10 estudantes na hora da aplicação do instrumento avaliativo.



<b>Pai com Ensino Médio</b>	Não	0,24	0,28	0,04	-1,60
	Sim	0,26	0,26	0,00	-0,19
<b>Pai com Ensino Superior</b>	Não	0,18	0,13	-0,05***	2,62
	Sim	0,20	0,20	0,00	-0,10
<b>Trabalha fora</b>	Não	0,15	0,30	0,15***	-6,16
	Sim	0,08	0,08	0,00	-0,15
<b>Cursou creche ou pré-escola</b>	Não	0,52	0,70	0,18***	-6,71
	Sim	0,57	0,57	0,00	-0,16
<b>Ensino público</b>	Não	0,77	0,68	-0,09***	3,46
	Sim	0,83	0,84	0,01	-0,44
<b>EJA</b>	Não	0,32	0,55	0,23***	-8,08
	Sim	0,36	0,36	0,00	-0,08
<b>Mora com pai e mãe</b>	Não	0,54	0,56	0,02	-0,80
	Sim	0,59	0,64	0,05	-1,43
<b>Reprovou</b>	Não	0,23	0,27	0,04**	-1,97
	Sim	0,16	0,19	0,03	-0,86
<b>Abandonou</b>	Não	0,12	0,17	0,05**	-2,15
	Sim	0,06	0,06	0,00	-0,00
<b>Pais incentivam</b>	Não	0,82	0,70	-0,12***	4,93
	Sim	0,89	0,90	0,01	-0,40
<b>Computador</b>	Não	0,70	0,80	0,10***	-4,60
	Sim	0,69	0,68	-0,01	0,35
<b>Nível socioeconômico</b>	Não	0,04	-0,04	-0,08	1,46
	Sim	0,04	0,01	-0,03	0,46

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente.

Primeiramente, cabe destacar que a maioria das características são estatisticamente diferentes entre os dois grupos antes do pareamento, com exceção dos pais possuírem ensino médio, dos estudantes residirem com ambos os pais e do nível socioeconômico. Após o *matching*, todas as médias das características não possuem diferença estatística.

Em termos de cor/raça, o grupo de tratamento possui 12 pontos percentuais (p.p.) a mais de pardos e a menos de brancos que o grupo de controle. Em média, após o pareamento, 53% da amostra é de estudantes autodeclarados pardos, os brancos e os pretos representam 25% cada um e 4% são amarelos. A proporção de mães que possuem ensino médio e ensino superior é maior que a dos pais. Entre as mães cerca de 30% possuem ambos os níveis de escolaridade e entre os pais 26% possuem ensino médio e 20% ensino superior.

Os estudantes do grupo de controle trabalham mais fora de casa que os do grupo de tratamento, 30% e 15%, respectivamente. A proporção do grupo de controle também é maior entre aqueles que cursam creche ou pré-escola (18 p.p) e estudaram em turmas da educação de jovens e adultos (EJA) (13p.p.). Enquanto, os estudantes do grupo de tratamento, em maior proporção, estudaram em escolas públicas (5 p.p.).

Os alunos que residem com os pais (pai e mãe) representam mais da metade da amostra e não há diferença estatisticamente significativa entre os dois grupos. Já no que tange ao rendimento dos alunos, entre aqueles que ingressaram por notas há maior proporção de reprovação (4 p.p.) e de abandono (5 p.p.). Os alunos do grupo de tratamento possuem mais incentivo dos pais para estudar em comparação ao grupo de controle, 82% e 70%, respectivamente.

No que se refere às características socioeconômicas, embora haja mais alunos com computador no grupo de controle (10 p.p.), a variável que indica o nível socioeconômico dos alunos não possui diferença estatística entre os estudantes dos dois grupos. A construção dessa variável é realizada por meio de análise de componente principal, com a extração do primeiro componente, o qual inclui características da residência do estudante, como: possuir geladeira, quantidade de quartos, de televisores, de banheiros, de carros e de máquinas de lavar. Quanto maior o valor, melhor é o nível socioeconômico dos alunos. O Apêndice A apresenta os testes de adequação da análise. O procedimento é realizado devido ao questionário do Saeb não possuir informações sobre a renda familiar do estudante.

Quanto ao desempenho dos alunos nas disciplinas de língua portuguesa e de matemática, o teste de diferença de médias confirma que há disparidades entre os grupos de tratamento e de controle (Tabela 2).

Tabela 2 – Diferença de médias das notas de língua portuguesa e de matemática

Disciplina	Média		Diferença	t-test
	Tratamento	Controle		
Língua Portuguesa	297,73	315,35	17,61***	2,55
Matemática	297,12	326,52	29,39***	2,88

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente.

Os alunos que ingressaram por meio de seleção baseada em notas possuem desempenho superior nas duas disciplinas. Os alunos do grupo de tratamento obtiveram uma média de nota na escala Saeb de 297,73 em Língua Portuguesa e 297,12 em Matemática, enquanto os do grupo de controle alcançaram 315,35 pontos em Língua Portuguesa e 326,52 em Matemática. A diferença entre os grupos é de 17,61 p.p em Língua Portuguesa e 29,39 p.p. em Matemática. Como os alunos analisados são de duas edições do Saeb, para minimizar o efeito da coorte sobre as notas, no que segue, os resultados são apresentados em escore padronizado com média zero e desvio-padrão (dp) igual a um.

#### 4. Resultados e Discussão

Para isolar o efeito seleção dos alunos que ingressam na Rede Federal de ensino, três métodos de pareamento são utilizados e algumas combinações deles são apresentadas. O primeiro deles é o PSM, que conforme apresentado na seção anterior, após a sua utilização, a diferença de médias das características observáveis não foi estatisticamente significativa e representa um indicativo de ajustamento. Adicionalmente, os Apêndices 3.B e 3.C apresentam o teste de tamanho do viés e a distribuição do escore de probabilidades ao tratamento, respectivamente. A redução dos valores do viés médio, do viés mediano e do teste Pseudo  $R^2$ , bem como a sobreposição do escore de propensão dos grupos de tratamento e de controle contribuem para reforçar a hipótese de ajustamento e balanceamento da amostra.

O primeiro passo para identificar o efeito do tratamento sobre os tratados via PSM é o cálculo da probabilidade de receber o tratamento – ter ingressado por meio do sorteio. As variáveis utilizadas para a estimação são aquelas apresentadas na Tabela 1 da seção 4.5. O modelo indica que a probabilidade de receber o tratamento está associada com a região de residência e reduz se o estudante trabalha fora, se cursou creche ou pré-escola e com o fato de ter estudado em escola pública (Apêndice 3.D). Os estudantes residentes nas regiões Norte e Centro-Oeste possuem maior probabilidade de ingressar na rede federal por meio de sorteio, conforme o esperado, visto que são as regiões que mais adotam este critério de seleção<sup>8</sup>. As demais características não possuem influência sobre a forma de ingresso dos alunos, confirmando, de modo geral, que a escolha do tipo de processo de seleção é intrínseca a escola (cada instituição possui autonomia administrativa – Lei nº 11.892 de 29 de dezembro de 2008) e exógena aos alunos. O segundo passo, apresentado na Tabela 3, é a estimação do ATT.

Tabela 3 – Resultados do Efeito do Tratamento sobre os tratados com utilização do PSM

	MQO	NN (1)	NN (1) SR	NN (3)	NN (5)	Kernel	Radius	IPWRA
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)
Língua Portuguesa	-0,09 (0,100)	-0,04 (0,036)	-0,09 (0,068)	-0,04 (0,059)	-0,07 (0,056)	-0,09* (0,050)	-0,20*** (0,048)	0,06 (0,044)
Matemática	-0,24** (0,113)	-0,27*** (0,080)	-0,28*** (0,076)	-0,024*** (0,066)	-0,24*** (0,063)	-0,26*** (0,057)	-0,40*** (0,054)	-0,14*** (0,054)

Fonte: Elaborada pelos autores.

Notas: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente. 2. MQO = Mínimos Quadrados Ordinários; NN (1) = nearest neighbor – o primeiro vizinho mais próximo; NN (1) SR = o primeiro vizinho mais próximo sem reposição; NN (3) = os três primeiros vizinhos mais próximos; NN (5) os cinco primeiros vizinhos mais próximos; Radius = raio de 0,10; IPWRA = Inverse Probability Weighted Regression Adjustment.

<sup>8</sup> O Instituto Federal do Acre é localizado na região Norte e o Instituto Federal de Brasília na região Centro-Oeste.

Os estudantes da RFE que ingressam pelos diferentes tipos de processo de seleção, de forma geral, não possuem diferença de desempenho em língua portuguesa. Com exceção da estimação do PSM com as matrizes de pesos dos tipos Kernel e Raio, que demonstram que os estudantes que entram por meio de sorteio possuem desempenho inferior em Língua Portuguesa. Em contraste, o resultado para a disciplina de matemática é robusto para todas as especificações. Com estimação do efeito entre 0,14dp e 0,40 dp a menos de notas, isso equivale a dizer que os alunos que entram por sorteio possuem notas de matemática inferiores aos que entram por meio de seleção. Inicialmente, a estimativa via MQO mostra que os estudantes que entram por meio do sorteio possuem 0,24 desvio-padrão a menos na disciplina de matemática no exame do Saeb, quando comparados aos estudantes que entram por processo de seleção baseado em notas. O efeito é mantido ou é superior quando a estimação do PSM é utilizada (colunas (b) a (g) da Tabela 3). Apenas na coluna (h), que considera uma abordagem duplamente robusta e avança ao considerar erros padrões robustos e clusterizados por turma – não possível no PSM –, é que a magnitude do efeito se reduz e o ATT passa a ser -0,14 dp.

Os demais métodos de pareamento – o CEM e o balanceamento por entropia – são apresentados nas colunas (a) e (b) da Tabela 4. Adicionais aos efeitos encontrados com as especificações do PSM, o efeito em matemática continua significativo estatisticamente e com magnitude próxima (-0,22 e -0,27 desvios-padrão), enquanto língua portuguesa não apresenta significância estatística com o balanceamento por entropia. Além disso, de acordo com Iacus, King e Porro (2009), o CEM pode ser utilizado em conjunto com outros métodos para ganhar confiabilidade dos resultados. Nas colunas (c) e (d) da Tabela 4 os coeficientes gerados pelo CEM+Entropia e pelo CEM+PSM, respectivamente, reforçam os resultados encontrados. Como contribuição complementar, há inferência estatisticamente significativa para língua portuguesa (-0,13 e -0,18 desvios-padrão).

Tabela 4 – Resultados do Efeito do Tratamento sobre os tratados com utilização do CEM e do Balanceamento por Entropia

	CEM (a)	Entropia (b)	CEM/Entropia (c)	CEM/PSM (d)
Língua Portuguesa	-0,12* (0,068)	-0,07 (0,043)	-0,13*** (0,047)	-0,18*** (0,077)
Matemática	-0,27*** (0,071)	-0,22*** (0,046)	-0,27*** (0,050)	-0,36*** (0,086)

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente.

Contudo, embora com uma magnitude menor na estimação IPWRA, os resultados de matemática são robustos as diferentes especificações e são similares em direção do efeito. Os parâmetros encontrados, corroboram com o que a literatura prévia apresenta sobre o efeito seleção (DUFLO; DUPAS; KREMER, 2011; HANUSHEK; WÖSSMANN, 2006; LARA; MIZALA; REPETTO, 2011), na existência de alguma forma de *tracking* ou escolha por habilidade, parte do desempenho superior de determinadas escolas são decorrência da própria triagem dos alunos.

O fato de haver efeito em matemática e com menor robustez em português, segundo Lubienski e Lubienski (2006), frequentemente está relacionado a como as disciplinas são ensinadas. Linguagens e leitura, em parte, podem refletir o contexto familiar, que neste estudo tem como pressuposto ser balanceado entre os grupos, e matemática é resultado do ensino dentro da escola. Além disso, há uma literatura específica que diferencia as habilidades em linguagens e matemática de acordo com o gênero (CONTINI; TOMMASO; MENDOLIA, 2017; GEVREK; GEVREK; NEUMEIER, 2020). As meninas teriam mais facilidade em português e os meninos em matemática. Para garantir que não há confusão de gênero nos resultados, visto que o Saeb de 2019 não distingue os estudantes entre meninos e meninas e, como consequência, a característica não foi incluída no estudo, na Tabela 5 são apresentados os resultados para o ano de 2017.

Tabela 5 – Resultados para o Saeb de 2017 com inclusão de característica de gênero e efeitos separados por gênero

	Inclusão do gênero			Meninas			Meninos		
	IPWRA	CEM	CEM/ Entropia	IPWRA	CEM	CEM/ Entropia	IPWRA	CEM	CEM/ Entropia
	(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	
LP	-0,10 (0,112)	-0,11 (0,087)	-0,10 (0,067)	-0,14 (0,105)	-0,17* (0,098)	-0,20** (0,087)	-0,12 (0,127)	-0,02 (0,137)	-0,05 (0,094)
MT	-0,34*** (0,100)	-0,32*** (0,077)	-0,27*** (0,076)	-0,23*** (0,073)	-0,24*** (0,077)	-0,27*** (0,093)	-0,49*** (0,176)	-0,38*** (0,112)	-0,34*** (0,119)

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente. 2. LP= Língua Portuguesa, MT=Matemática.

A validade dos resultados, depois de controlada a característica de gênero, foi confirmada nas colunas (a), (b) e (c) da Tabela 5. O efeito continua robusto para a disciplina de matemática e sem efeito do tratamento sobre os tratados em relação a língua portuguesa. Quando analisado o efeito separado para meninas e meninos (colunas (d) a (h)), conforme o esperado, as estudantes autodeclaradas do sexo feminino em 2017 apresentam diferença estatística em LP (-0,17 e -0,20 desvios-padrão) e entre os estudantes autodeclarados do sexo masculino não há efeito. De forma complementar, em matemática, o ATT é maior entre meninos (-0,34 a -0,49 desvios-padrão) quando comparado com o resultado para as meninas (-0,23 a -0,27 desvios-padrão).

#### 4.1 Análise de sensibilidade e ajustamento do balanceamento por entropia

Na medida em que o PSM possui como hipótese a independência do tratamento condicional às características observáveis, para verificar se a estimação é confiável, uma medida de sensibilidade é utilizada. Embora não seja possível testar a hipótese, a utilização dos limites de Rosenbaum possibilita verificar qual a magnitude dos efeitos não observáveis para gerar viés nas estimativas do pareamento. Quanto maior for o valor crítico de  $\Gamma$ , menos os resultados são sensíveis às características não observáveis estarem afetando o tratamento e quanto mais próximo da unidade, maior é a chance.

Na medida em que análise de sensibilidade calcula a diferença dos intervalos de confiança das médias entre os grupos de tratamento e de controle, no Apêndice E são apresentados o valor crítico do teste para os resultados que apresentaram significância estatística. Para língua portuguesa, com a especificação CEM+PSM, o valor crítico de  $\Gamma$  é 1,3. Ou seja, os fatores não observáveis devem aumentar as chances de tratamento em 1,3 vezes para que o efeito do tratamento não fosse significativo. Para matemática, os valores críticos são ainda maiores, 1,5 e 1,7 via PSM e CEM + PSM, respectivamente. Como quanto mais distante da unidade, menor é a chance de os efeitos não observados estarem afetando o resultado, esta estimativa corrobora com a robustez do efeito seleção encontrado neste estudo.

#### 4.2 Mecanismos

Na busca por entender os fatores que contribuem para a redução da nota Saeb entre os alunos que entram por sorteio, esta seção discute e testa possíveis caminhos que levaram a este resultado. Primeiramente, a literatura empírica ressalta a possibilidade de diferentes efeitos entre os quantis de notas (CUMMINS, 2017; EPPLE; NEWLON; ROMANO, 2002). A especificação de regressão quantílica (Tabela 6) permite utilizar uma abordagem duplamente robusta (IPWRA – coluna 1) e em conjunto com o CEM (coluna 2). Em ambas as especificações, matemática possui um efeito maior na distribuição inferior de notas (quantis 10 e 25) e vai perdendo magnitude nos quantis superiores. Ou seja, os alunos com desempenho menor são os que mais possuem diferença entre os dois modos de seleção.

Tabela 6 – Resultados sobre os diferentes quantis da distribuição de notas

	q10		q25		q50		q75		q90	
	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
LP	-0,01 (0,15)	-0,13 (0,128)	0,14*** (0,054)	-0,03 (0,080)	0,04 (0,046)	0,02 (0,049)	-0,02 (0,040)	-0,13*** (0,043)	0,06 (0,490)	0,02 (0,040)
MT	-0,25** (0,133)	-0,32*** (0,119)	-0,31*** (0,114)	-0,32*** (0,072)	-0,11** (0,059)	-0,27*** (0,083)	-0,13*** (0,519)	-0,21*** (0,044)	0,21*** (0,065)	-0,15*** (0,056)

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente. 2. Considerando erros padrão robustos por turma. 3. LP= Língua Portuguesa, MT=Matemática.

Na existência de diferentes efeitos entre as notas e com maior diferença entre os estudantes dos quantis inferiores, a segunda possibilidade a ser testada é a existência de efeito de pares. Se os alunos de maior desempenho contribuem para o aprendizado dos alunos com menor habilidade, é conveniente que os alunos de turmas que ingressam de forma aleatória (dado o sorteio das vagas) possuam menos pares com maior habilidade e, conseqüentemente, tenham ganhos de pares inferiores. Ou seja, as turmas com alunos selecionados com base em notas tendem a terem maior habilidade média e serem mais homogêneas e as turmas compostas por alunos sorteados tendem a ser mais heterogêneas. Neste caso, é provável que há relação positiva entre o desempenho médio da turma e a nota do estudante. Para verificar se há efeitos de pares, uma especificação alternativa<sup>9</sup> é utilizada (Tabela 7).

Tabela 7 – Efeitos médios do tratamento sobre os tratados e *peer effects*

	IPWRA	CEM	CEM/Entropia
Língua Portuguesa	0,05* (0,021)	-0,02 (0,023)	-0,05* (0,029)
Peer Effects	0,70*** (0,048)	0,77*** (0,022)	0,64*** (0,051)
Matemática	-0,04 (0,032)	-0,05** (0,023)	-0,09*** (0,028)
Peer Effects	0,62*** (0,068)	0,77*** (0,020)	0,62*** (0,058)

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente. 2. Considerando erros padrão robustos por turma.

O parâmetro *peer effects* estimado confirma a influência que os pares possuem na nota dos alunos. Em termos de magnitude da associação, os demais alunos da turma contribuem com aproximadamente 0,70 d.p. nas notas das duas disciplinas. O coeficiente do ATT de ambas as disciplinas também perde magnitude e, em alguns casos, significância, o que reforça a forte influência dos pares. A inclusão do controle alternativo de efeito dos pares corrobora com um dos principais mecanismos de diferença de notas no caso de alocação de alunos por desempenho. De acordo com Carrell, Hoekstra e West (2011), Cummins (2017) e Epple, Newlon e Romano (2002), na existência de diferentes notas entre turmas com *tracking*, parte do resultado é explicado pelo efeito dos pares.

Outra explicação para as diferenças de notas é elencada como decorrência da escolha de escola (ALLEN, 2007; IVANIUSHINA *et al.*, 2018; KIM, 2018; PIOPIUNIK, 2014). Embora no desenho de estratégia empírica deste estudo todos os alunos já tenham realizado a sua escolha da escola, uma das principais fonte de viés de variáveis omitidas quando calculado o efeito escola, diz respeito as características não observáveis da coorte da região afetando o resultado (ANGRIST; BETTINGER; KREMER, 2006; BAUDE *et al.*, 2020; BENEVIDES; SOARES, 2014; CHUDGAR; QUIN, 2012; HANUSHEK *et al.*, 2007; HOWELL *et al.*, 2002; HOXBY; ROCKOFF, 2004; LUBIENSKI; LUBIENSKI, 2006; MCEWAN, 2004; SAPELLI; VIAL, 2002; SOMERS; MCEWAN; WILLMS, 2004).

<sup>9</sup> A forma mais tradicional de captar efeito de pares é por meio do modelo linear-em-médias. Formalmente, ela segue a especificação:  $y_i = \alpha + \beta_1 \bar{y}_{-i} + \gamma X_i + \rho \bar{X}_{-1} + \varepsilon_i$ . O que significa que a variável de resultado  $y$  é uma função das características observáveis do aluno  $i$ , da média de notas da turma sem a nota específica do aluno e das características médias da turma (SACERDOTE, 2011).

Para investigar se há algum tipo de efeito intrínseco aos alunos da região ou em relação a qualidade média da coorte de estudantes, a Tabela 8 adiciona como controle a nota média dos alunos do nono ano do ensino fundamental do município correspondente a cada estudante para os anos de 2013 e 2015. Visto que estas são as coortes de estudantes que estariam no ensino médio em 2017 e 2019.

Tabela 8 – Efeitos médios do tratamento controlando pela nota média dos alunos do nono ano no município nos anos de 2013 e 2015

	<b>IPWRA</b>	<b>CEM</b>	<b>CEM/Entropia</b>
Língua Portuguesa	0,05 (0,048)	-0,11 (0,075)	-0,11* (0,064)
Média 9º ano	0,0004 (0,002)	0,01*** (0,001)	0,01* (0,004)
Matemática	-0,10* (0,058)	-0,25*** (0,077)	-0,24*** (0,069)
Média 9º ano	0,01** (0,002)	0,01*** (0,001)	0,005 (0,004)

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente. 2. Considerando erros padrão robustos por turma.

A nota média da coorte do nono ano, ainda que possua relação estatisticamente significativa com a nota dos alunos, possui a magnitude de 0,01 d.p. e não afeta o ATT de interesse. O efeito na disciplina de matemática continua robusto, os alunos que entram por sorteio possuem desempenho inferior.

Em termos ainda de variável não observável, é possível verificar se há efeito do tratamento na incidência de reprovação e de abandono dos estudantes. Ambas as características são controladas nas estimações, mas o efeito direto (Tabela 9) reforça que os dois grupos são muito próximos em características não observáveis.

Tabela 9 – Efeitos médios do tratamento sobre os tratados nas variáveis de reprovação e abandono

	<b>IPWRA</b>	<b>CEM</b>	<b>CEM/Entropia</b>
Reprovou	-0,02 (0,021)	-0,01 (0,03)	-0,01 (0,02)
Abandonou	0,01 (0,009)	0,02* (0,01)	0,02* (0,01)

Fonte: Elaborada pelos autores.

Nota: 1. \* representa confiança de 90%. 2. Considerando erros padrão robustos por turma.

Por último, volta-se com a discussão sobre a análise estar centrada nos concluintes, então denominados vencedores. A ideia é que o Saeb é um exame do tipo *low-stakes* – com baixo risco –, então a utilização de um exame *high-stakes* – com alto risco –, como o Exame Nacional do Ensino Médio (ENEM), pode mudar a conduta dos alunos na prova, visto que ele é parte ou integralmente usado para o ingresso no ensino superior de diversas unidades de ensino. Neste caso, na medida em que o esforço dos alunos pode variar, se a diferença dos resultados se mantiver no ENEM (Tabela 10), há indicativos de que existem efeitos da forma de seleção em etapas posteriores ao ensino médio. Ou seja, no ingresso ao ensino superior e, conseqüentemente, no mercado de trabalho e salários.

Tabela 10 – Efeitos médios do tratamento sobre os tratados nas notas do ENEM

	IPWRA	CEM	CEM/Entropia
Ciências da Natureza	-9,66* (4,990)	-6,46 (4,654)	-8,45 (5,266)
Ciências Humanas	-19,99*** (4,786)	-16,35*** (3,781)	-7,27 (5,152)
Linguagens e Códigos	-11,05*** (3,626)	-10,77*** (3,053)	-5,39 (3,863)
Matemática	-27,46*** (5,662)	-22,39*** (6,426)	-35,50*** (6,995)
Redação	-47,66*** (8,327)	-48,86*** (7,847)	-33,61*** (10,934)

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente. 2. Considerando erros padrão robustos por turma. 3. As variáveis de controle utilizadas são: gênero, cor/raça, escolaridade dos pais, renda, se o estudante reside sozinho, se possui computador, nível socioeconômico e região de residência – informações disponíveis no questionário do ENEM.

As notas do ENEM são divididas em cinco blocos e atribuídas independentemente para cada um deles. Como resultados, o efeito na disciplina de matemática continua robusto às especificações distintas e, adicionalmente, também há diferença nas notas de redação dos candidatos.

## 5. Considerações Finais

As diferentes médias de notas atingidas em exames padronizados por instituições de ensino, normalmente, são vistas como uma medida de eficácia. Poucas vezes as características dos estudantes que ingressam nas escolas são visualizadas diretamente como parte do conjunto de notas que a escola entrega. De forma inovadora e com base em uma estratégia empírica que explora a forma de ingresso dos alunos na rede federal de ensino, este estudo buscou trazer evidências empíricas de como o processo de seleção afeta as notas dos estudantes brasileiros.

Os resultados indicam que há uma diferença na nota de matemática de 0,14 d.p. a 0,40 d.p. quando os alunos são sorteados para entrar na RFE. Em termos de magnitude do efeito, na medida em que o desvio-padrão total das notas na amostra é de 0,94, isso representa uma redução da dispersão entre 14% e 42%. Adicionalmente, as notas médias dos estudantes de 2017 a 2019 aumentou 0,20 d.p., o que indica uma magnitude considerável, visto que o tamanho do impacto equivale de 1 a 4 anos de aumento de notas.

Os efeitos foram distintos entre o gênero dos estudantes e ao longo da distribuição das notas. A magnitude do coeficiente de impacto em matemática é maior entre os meninos, quando comparado com as estudantes autodeclaradas do sexo feminino. Além disso, entre as meninas que ingressam por sorteio, também há indicativos de diferença no desempenho de língua portuguesa. O efeito seleção é maior entre os estudantes da cauda inferior da distribuição de notas.

Uma possível explicação do efeito seleção está no ganho em termos de pares. Conforme a média de desempenho da turma se eleva, há indicativos de associação positiva com a nota dos estudantes nas disciplinas de língua portuguesa e de matemática. Além disso, o efeito seleção pode refletir na nota alcançada para o ingresso no ensino superior, como o ENEM, e, conseqüentemente, pode impactar ao longo da vida dos estudantes.

Os resultados foram robustos a um conjunto de especificações e os testes demonstraram baixa probabilidade de haver efeitos não observados, gerando viés nas estimativas.

Cabe destacar que o trabalho apresenta algumas limitações, visto que os resultados encontrados são para um perfil específico de alunos - do último ano do ensino médio da rede federal de educação - e isso incorre em baixa validade externa. Ainda assim, a evidência reforça que a forma de ingresso dos alunos importa e que isso deve ser considerado em políticas públicas educacionais que levam em consideração a forma de alocação dos alunos entre as escolas e dentro das turmas. Dessa forma, o trabalho contribui para o cenário educacional brasileiro, bem como para a instigar futuras pesquisas sobre o tema.

## REFERÊNCIAS

- ALLEN, Rebecca. Allocating pupils to their nearest secondary school: The consequences for social and ability stratification. *Urban Studies*, v. 44, n. 4, p. 751–770, 2007.
- ANGRIST, Joshua; BETTINGER, Eric; KREMER, Michael. Long-term educational consequences of secondary school vouchers: Evidence from administrative records in Colombia. *American Economic Review*, v. 96, n. 3, p. 847–862, 2006.
- BAUDE, Patrick L. *et al.* The Evolution of Charter School Quality. *Economica*, v. 87, p. 158–189, 2020.
- BENEVIDES, Alesandra de Araújo; SOARES, Ricardo Brito. Diferencial de desempenho das escolas militares: bons alunos ou boa escola? 2014. Disponível em: <[https://www.bnb.gov.br/documents/160445/960917/DIFERENCIAL\\_DE\\_DESEMPENHO\\_DAS\\_ESCOLAS\\_MILITARES.pdf/7ae9ef81-9687-46cb-b501-766ccef1cba2](https://www.bnb.gov.br/documents/160445/960917/DIFERENCIAL_DE_DESEMPENHO_DAS_ESCOLAS_MILITARES.pdf/7ae9ef81-9687-46cb-b501-766ccef1cba2)>.
- BLACKWELL, Matthew *et al.* cem : Coarsened exact matching in Stata. *The Stata Journal*, v. 9, n. 4, p. 524–546, 2009.
- CARRELL, Scott E.; SACERDOTE, Bruce I.; WEST, James E. From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, v. 81, n. 3, p. 855–882, 2013.
- CARRELL, Scott E.; HOEKSTRA, Mark; WEST, James E. Does drinking impair college performance? Evidence from a regression discontinuity approach. *Journal of Public Economics*, v. 95, n. 1–2, p. 54–62, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.jpubeco.2010.08.008>>.
- CHUDGAR, Amita; QUIN, Elizabeth. Relationship between private schooling and achievement: Results from rural and urban India. *Economics of Education Review*, v. 31, n. 4, p. 376–390, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.econedurev.2011.12.003>>.
- COLEMAN, James; HOFFER, Thomas; KILGORE, Sally. Achievement and Segregation in Secondary Schools : A Further Look at Public and Private School Differences. *Sociology of Education*, v. 55, n. 2, p. 162–182, 1982.
- COLEMAN, James S; HOFFER, Thomas. Public and Private High School. *NASSP Bulletin*, 1987.
- CONTINI, Dalit; TOMMASO, Maria Laura Di; MENDOLIA, Silvia. The gender gap in mathematics achievement: Evidence from Italian data. *Economics of Education Review*, v. 58, p. 32–42, 2017. Disponível em: <<http://dx.doi.org/10.1016/j.econedurev.2017.03.001>>.
- CUMMINS, Joseph R. Heterogeneous treatment effects in the low track: Revisiting the Kenyan primary school experiment. *Economics of Education Review*, v. 56, p. 40–51, 2017. Disponível em: <<http://dx.doi.org/10.1016/j.econedurev.2016.11.006>>.
- DUFLO, Esther; DUPAS, Pascaline; KREMER, Michael. Peer effects, teacher incentives, and the impact of tracking: Evidence from a randomized evaluation in Kenya. *American Economic Review*, v. 101, n. 5, p. 1739–1774, 2011.
- EMSLEY, Richard; LUNT, Mark; PICKLES, Andrew. Implementing double-robust estimators of causal effects. n. 3, p. 334–353, 2008.
- EPPLE, Dennis; NEWLON, Elizabeth; ROMANO, Richard. Ability tracking, school competition, and the distribution of educational benefits. *Journal of Public Economics*, v. 83, n. 1, p. 1–48, 2002.
- FRYER, Roland G.; LEVITT, Steven D. The Black-White test score gap through third grade. *American Law and Economics Review*, v. 8, n. 2, p. 249–281, 2006.
- GEVREK, Z. Eylem; GEVREK, Deniz; NEUMEIER, Christian. Explaining the gender gaps in mathematics achievement and attitudes: The role of societal gender equality. *Economics of Education*



*Review*, v. 76, n. April, p. 101978, 2020. Disponível em:  
<<https://doi.org/10.1016/j.econedurev.2020.101978>>.

HAINMUELLER, Jens. Entropy Balancing for Causal Effects : A Multivariate Reweighting Method to Produce Balanced Samples in Observational Studies. *Political Analysis*, v. 20, p. 25–46, 2012.

HANUSHEK, Eric A. *et al.* Charter school quality and parental decision making with school choice. *Journal of Public Economics*, v. 91, p. 823–848, 2007.

HANUSHEK, Eric A.; KAIN, John F.; RIVKIN, Steven G. New evidence about Brown v. Board of Education: The complex effects of school racial composition on achievement. *Journal of Labor Economics*, v. 27, n. 3, p. 349–383, 2009.

HANUSHEK, Eric A.; RIVKIN, Steven G. *HARMING THE BEST: HOW SCHOOLS AFFECT THE BLACK-WHITE ACHIEVEMENT GAP. NATIONAL BUREAU OF ECONOMIC RESEARCH - NBER.* [S.l: s.n.], 2008.

HANUSHEK, Eric A.; WÖSSMANN, Ludger. Does educational tracking affect performance and inequality? differences-in-differences evidence across countries. *The Economic Journal*, v. 116, p. 63–76, 2006.

HOWELL, William G. *et al.* School Vouchers and Academic Performance: Results from Three Randomized Field Trials. *Journal of Policy Analysis and Management*, v. 21, n. 2, p. 191–217, 2002.

HOXBY, Caroline M.; ROCKOFF, Jonah E. The Impact of Charter Schools on Student Achievement. 2004.

IACUS, Stefano M; KING, Gary. Matching for Causal Inference Without Balance Checking. n. 617, 2008.

IACUS, Stefano M; KING, Gary; PORRO, Giuseppe. Causal Inference without Balance Checking : Coarsened Exact Matching. *Political Analysis*, v. August, 2011.

IACUS, Stefano M; KING, Gary; PORRO, Giuseppe. cem : Software for Coarsened Exact Matching. *Journal of Statistical Software*, v. 30, n. 9, 2009.

IVANIUSHINA, Valeria *et al.* School segregation in St. Petersburg—the role of socioeconomic status. *Education Economics*, v. 27, n. 2, p. 1–21, 2018. Disponível em:  
<<https://doi.org/10.1080/09645292.2018.1538408>>.

KIM, Youngran. The effects of school choice on achievement gaps between private and public high schools: Evidence from the Seoul high school choice program. *International Journal of Educational Development*, v. 60, p. 25–32, 2018. Disponível em: <<https://doi.org/10.1016/j.ijedudev.2017.10.019>>.

LANGE, Simon; VON WERDER, Marten. Tracking and the intergenerational transmission of education: Evidence from a natural experiment. *Economics of Education Review*, v. 61, p. 59–78, 2017. Disponível em: <<https://doi.org/10.1016/j.econedurev.2017.10.002>>.

LARA, Bernardo; MIZALA, Alejandra; REPETTO, Andrea. The Effectiveness of Private Voucher Education: Evidence From Structural School Switches. *Educational Evaluation and Policy Analysis*, v. 33, n. 2, p. 119–137, 2011.

LAZEAR, Edward P. Educational Production. *The Quarterly Journal of Economics*, v. 116, n. 3, p. 777–803, 2001. Disponível em:  
<[http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list\\_uids=8508029795338633420related:zECZeQWYEnYJ](http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=pubmed&cmd=Retrieve&dopt=AbstractPlus&list_uids=8508029795338633420related:zECZeQWYEnYJ)>.

LUBIENSKI, Sarah Theule; LUBIENSKI, Christopher. School Sector and Academic Achievement: A Multilevel Analysis of NAEP Mathematics Data. *American Educational Research Journal*, v. 43, n. 4, p.

651–698, 2006.

LUNCEFORD, Jared K; DAVIDIAN, Marie. Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *STATISTICS IN MEDICINE*, v. 23, n. April, p. 2937–2960, 2004.

MALAMUD, Ofer; POP-ELECHES, Cristian. School tracking and access to higher education among disadvantaged groups. *Journal of Public Economics*, v. 95, p. 1538–1549, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.jpubeco.2011.03.006>>.

MCEWAN, Patrick J. The potential impact of vouchers. *Peabody Journal of Education*, v. 79, n. 3, p. 57–80, 2004.

PEKKARINEN, Tuomas; UUSITALO, Roope; KERR, Sari. School tracking and intergenerational income mobility: Evidence from the Finnish comprehensive school reform. *Journal of Public Economics*, v. 93, p. 965–973, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.jpubeco.2009.04.006>>.

PIOPIUNIK, Marc. The effects of early tracking on student performance: Evidence from a school reform in Bavaria. *Economics of Education Review*, v. 42, p. 12–33, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.econedurev.2014.06.002>>.

PNP. *Plataforma Nilo Peçanha*. Disponível em: <<http://plataformanilopecanha.mec.gov.br/>>.

POP-ELECHES, Cristian; URQUIOLA, Miguel. Going to a better school: Effects and behavioral responses. *American Economic Review*, v. 103, n. 4, p. 1289–1324, 2013.

ROSENBAUM, B Y Paul R; RUBIN, Donald B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, v. 70, n. 1, p. 41–55, 1983.

ROSENBAUM, Paul R. *Observational Studies*. Second Edi ed. [S.l.]: Springer Series in Statistics, 2002.

ROTHSTEIN, Richard. The Racial Achievement Gap, Segregated Schools, and Segregated Neighborhoods: A Constitutional Insult. *Race and Social Problems*, v. 7, n. 1, p. 21–30, 2015.

RUHOSE, Jens; SCHWERDT, Guido. Does early educational tracking increase migrant-native achievement gaps? Differences-in-differences evidence across countries. *Economics of Education Review*, v. 52, p. 134–154, 2016.

SACERDOTE, Bruce. Peer Effects in Education: How might they work, how big are they and how much do we know Thus Far? *Handbook of the Economics of Education*. 1. ed. [S.l.]: Elsevier B.V., 2011. v. 3. p. 249–277. Disponível em: <<http://dx.doi.org/10.1016/B978-0-444-53429-3.00004-1>>.

SAPELLI, Claudio; VIAL, Bernardita. THE PERFORMANCE OF PRIVATE AND PUBLIC SCHOOLS IN THE CHILEAN VOUCHER SYSTEM. *Cuadernos de Economía*, n. 118, p. 423–454, 2002.

SIMCAQ. *Simulador de custo aluno qualidade*. Disponível em: <<https://simcaq.c3sl.ufpr.br/simulator/results/caqreport>>.

SOMERS, Marie Andrée; MCEWAN, Patrick J.; WILLMS, J. Douglas. How Effective Are Private Schools in Latin America? *Comparative Education Review*, v. 48, n. 1, p. 48–69, 2004.

## APÊNDICES

### Apêndice A – Testes de KMO e Bartlett

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0,7929
Bartlett's Test of Sphericity	Qui-quadrado	1,8e <sup>+4</sup>
	Sig.	0,0000

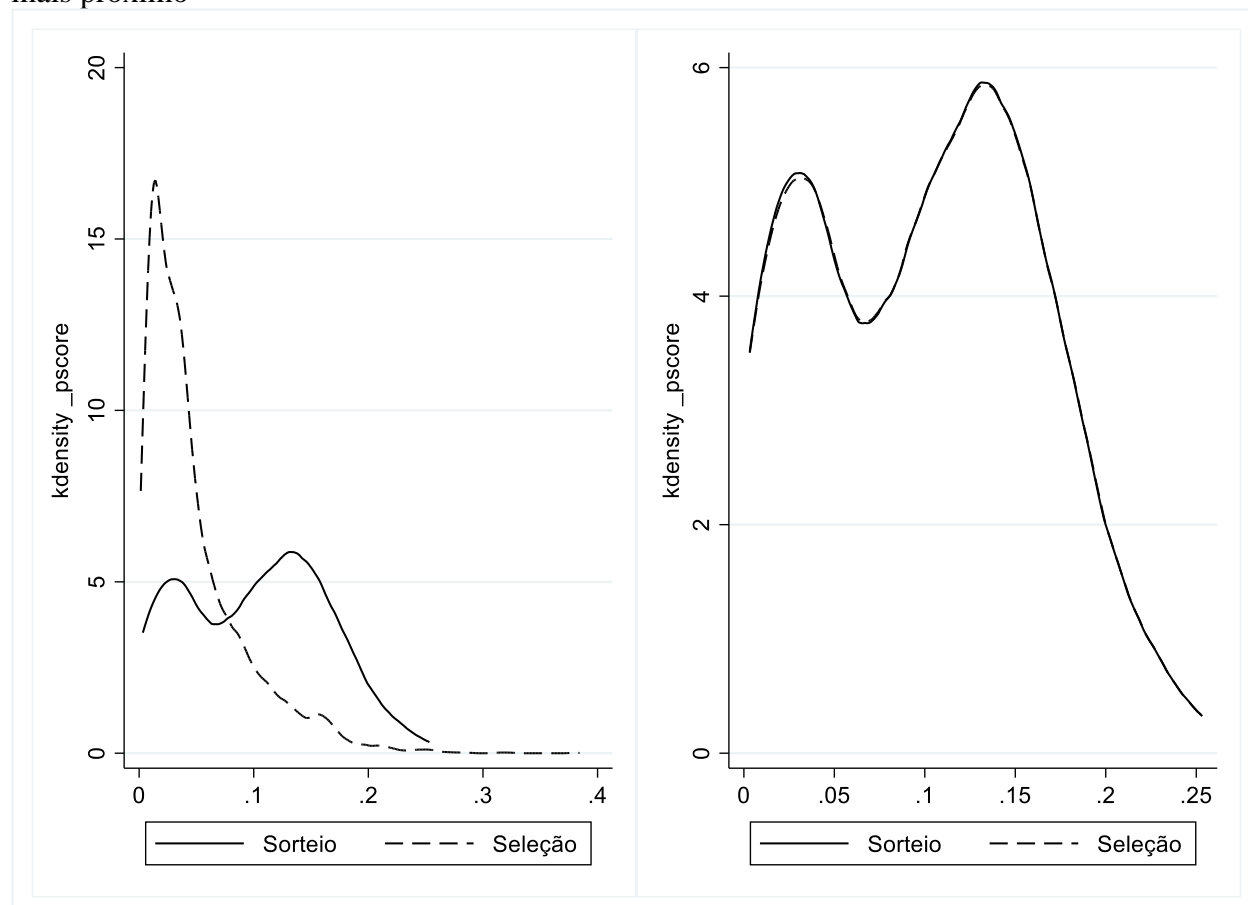
Fonte: Elaborada pelos autores.

Apêndice B - Teste Pseudo  $R^2$ , máxima verossimilhança, viés médio e viés mediano

Amostra	Pseudo $R^2$	Valor p do $\chi^2$	Viés Médio	Viés Mediano
Unmatched	0,095	0,000	27,4	19,2
Matched	0,010	0,995	2,9	2,2

Fonte: Elaborado pelos autores.

Apêndice C - Distribuição do Escore de Propensão para Tratados e para Não Tratados com um vizinho mais próximo



Fonte: Elaborado pelos autores.

Nota: Na estimação não-paramétrica de densidade univariada de *kernel* é usada a função de *epanechnikov* e *bandwidth* automática.

Apêndice D - Resultados estimados para a probabilidade de receber o tratamento

Covariada	Parâmetro Estimado
Região Nordeste (base=Norte)	-
Região Sudeste	-
Região Sul	-0,25*** (0,091)
Região Centro-Oeste	-0,05 (0,071)
Turno vespertino (base=matutino)	0,05 (0,098)
Cor branca	-0,17 (0,149)
Cor preta	0,19 (0,166)
Cor parda	0,13 (0,146)

Cor amarela	0,32 (0,218)
Mãe possui ensino médio	-0,12 (0,076)
Mãe possui ensino superior	-0,07 (0,081)
Pai possui ensino médio	-0,09 (0,073)
Pai possui ensino superior	0,03 0,087
Aluno trabalha fora	-0,29*** 0,095
Aluno curso creche ou pré-escola	-0,24*** (0,063)
Estudou em escola pública	-0,39*** (0,086)
Estudou no EJA	0,21 (0,132)
Mora com pai e mãe	-0,01 (0,063)
Reprovou	-0,085 (0,085)
Abandonou	0,22 (0,149)
Pais incentivam	-0,01 (0,093)
Possui computador	-0,06 (0,071)
Nível socioeconômico	0,03 (0,036)
$R^2$	0,1021

Fonte: Elaborada pelos autores.

Nota: 1. \*, \*\* e \*\*\* representam confiança de 90%, 95% e 99%, respectivamente.

#### Apêndice E – Valores críticos da análise de sensibilidade (limites de Rosenbaum)

Variável		$\Gamma$	p-crítico
Língua Portuguesa	CEM + PSM	1	0,001557
		1.1	0,011663
		1.2	0,050047
		1.3	0,141531
Matemática	PSM	1	0,000054
		1.1	0,000784
		1.2	0,00589
		1.3	0,026721
		1.4	0,082211
		1.5	0,187152
		CEM + PSM	1
	1.1	0,00001	
	1.2	0,000133	
	1.3	0,001022	
	1.4	0,005117	
	1.5	0,018278	
	1.6	0,04982	
	1.7	0,109157	

Fonte: Elaborada pelos autores.