

**idp**

15.12.2021

**Nº 05**

**REVISTA**

# **DEBATES EM ECONOMIA APLICADA**

**AVALIAÇÃO DO IMPACTO DO ENSINO EM TEMPO  
INTEGRAL: UMA ANÁLISE DO PROGRAMA NOVO  
MAIS EDUCAÇÃO**

**RENATO COSTA COITINHO**

**AVALIAÇÃO DO IMPACTO DO ENSINO EM  
TEMPO INTEGRAL: UMA ANÁLISE DO  
PROGRAMA NOVO MAIS EDUCAÇÃO**  
**RENATO COSTA COITINHO<sup>1</sup>**

---

<sup>1</sup> Renato Costa Coitinho é Mestre em Economia pelo Instituto Brasileiro de Ensino, Desenvolvimento e Pesquisa (IDP). E-mail: renatoccoitinho@gmail.com.



**IDP**

# DEBATES EM ECONOMIA APLICADA

[www.idp.edu.br](http://www.idp.edu.br)

O IDP é um centro de excelência no ensino, na pesquisa e na extensão nas áreas da Administração Pública, Direito e Economia. O Instituto tem como um de seus objetivos centrais a profusão e difusão do conhecimento de assuntos estratégicos nas áreas em que atua, constituindo-se um *think tank* independente que visa contribuir para as transformações sociais, políticas e econômicas do Brasil.

## DIREÇÃO E COORDENAÇÃO

### Diretor Geral

Francisco Schertel

### Coordenador do Mestrado em Administração Pública

Caio Cordeiro de Resende

### Coordenador do Mestrado em Economia

José Luiz Rossi

## CONSELHO EDITORIAL

### Coordenação

Thiago Costa Monteiro Caldeira

### Supervisão e Revisão

Renan Holtermann, Matheus Gonçalves,  
Mathias Tessmann, Milton Sobrinho e  
Alessandro Freire

### Comunicação e Marketing

Antonio Zaninetti e Daniel Jordão

### Projeto gráfico e diagramação

Juliana Vasconcelos

**Revista Técnica voltada** à divulgação de resultados preliminares de estudos e pesquisas aplicados em desenvolvimento por professores, pesquisadores e estudantes de pós-graduação com o objetivo de estimular a produção e a discussão de conhecimentos técnicos relevantes na área de Economia.

**Convidamos a comunidade** acadêmica e profissional a enviar comentários e críticas aos autores, visando o aprimoramento dos trabalhos para futura publicação. Por seu propósito se concentrar na recepção de comentários e críticas, a Revista Debates em Economia Aplicada não possui ISSN e não fere o ineditismo dos trabalhos divulgados.

**As publicações** da Revista estão disponíveis para acesso e download gratuito no formato PDF. Acesse: [www.idp.edu.br](http://www.idp.edu.br)

**As opiniões emitidas** nesta publicação são de exclusiva e inteira responsabilidade dos autores, não exprimindo, necessariamente, o ponto de vista do IDP.

**Qualquer citação** aos trabalhos da Série só é permitida mediante autorização expressa do(s) autor(es).

# SUMÁRIO

1. Introdução	7
1.1. Objetivos geral e específicos	9
1.2. Delimitação do Escopo do Estudo	9
2. Fundamentação Teórica	10
2.1. Literatura internacional e doméstica acerca do impacto da educação em tempo integral	12
2.2. O Programa Novo Mais Educação	13
2.3. Estratégia de avaliação de impacto de políticas públicas	16
3. Metodologia	18
3.1. Bases de dados	19
3.2. Universo de escolas do estudo	20
3.3. O modelo de propensity score matching	22
3.3.1. Estimação dos escores das escolas	26
3.3.2. Críticas acerca do modelo de propensity score matching	30
3.4. O modelo de Mahalanobis matching	31
3.5. O modelo de genetic matching	32
4. Resultados	33
4.1. Avaliação do pareamento por propensity score matching	34
4.2. Avaliação do pareamento por distância de Mahalanobis	39
4.3. Avaliação do pareamento por genetic matching	43
4.4. Comparação dos pareamentos	47
4.5. Impacto do PNME sobre o desempenho acadêmico das escolas participantes	49
5. Conclusão	52
6. Referências	55

**Resumo:** O debate acerca dos determinantes de uma educação de qualidade vem de longa data. Dentro deste tema, são bastante explorados os retornos acadêmicos associados à educação em tempo integral. No Brasil, vigorou entre 2017 e 2019 o Programa Novo Mais Educação (PNME), que objetivava aprimorar o desempenho escolar no Ensino Fundamental em Língua Portuguesa e Matemática através da ampliação da jornada escolar. De forma inédita, o presente estudo apresenta uma avaliação do impacto do PNME nas notas dos alunos nestas duas disciplinas em 1.265 escolas brasileiras, para os cinco anos iniciais do Ensino Fundamental, e 1.705 escolas, para os quatro anos finais. Utilizaram-se os métodos de *propensity score matching*, *Mahalanobis matching* e *genetic matching*, pareando as escolas contempladas pelo Programa com instituições semelhantes do grupo de controle. Os pareamentos por distância de Mahalanobis e por *genetic matching* apresentaram um balanceamento melhor em suas covariadas. A diferença no desempenho acadêmico para ambas as disciplinas entre as escolas participantes do PNME foi inconclusiva para os anos iniciais do Ensino Fundamental, variando de acordo com o pareamento utilizado. Para os anos finais, porém, observaram-se ganhos significativos em todos os três modelos, variando entre 0,13 e 0,20 desvios-padrão nas notas de Língua Portuguesa e entre 0,17 e 0,22 desvios-padrão nas notas de Matemática.

**Palavras-chave:** Programa Novo Mais Educação; educação integral; métodos de pareamento.

**Abstract:** The discussion surrounding the determinants of quality education is long run. Within this field, the academic returns associated with full-time education are often explored. In Brazil, the *Programa Novo Mais Educação* (PNME) was a nation-wide program which ran from 2017 through 2019, whose goal was to improve elementary school students' grades in Portuguese language and mathematics through the extension of the school day. The present study is the first to present the Program's impact on students' grades for these two subjects in 1,265 Brazilian schools, for the first five years of elementary schooling, and 1,705 schools, for the latter four years. The methods used were propensity score matching, Mahalanobis matching and genetic matching, pairing treated schools against similar institutions from the control group based on data from the local *Ministério da Educação* and *Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*. The schools paired through Mahalanobis distance and genetic matching showed greater balance between their covariates. The change in academic performance for both subjects among treated schools was inconclusive for the first five years of elementary schooling, varying according to the method employed. For the four latter years, however, significant gains were found through all three models, ranging

from 0.13 to 0.20 standard deviations in Portuguese language and from 0.17 to 0.22 standard deviations in Math.

**KEYWORDS:** *Programa Novo Mais Educação*; full-time education; matching methods.

## 1. INTRODUÇÃO

A performance escolar de um indivíduo está associada não só à sua renda futura (HARMON, OOSTERBEEK, WALKER, 2003) como também à sua participação cívica (DEE, 2004), ao crescimento econômico local como um todo (BLUNDELL, DEARDEN, MEGHIR, SIANESI, 1999) e a outros indicadores sociais e econômicos de importância, chegando até a impactar sua probabilidade de cometer um crime (MCEACHIN, LAUEN, FULLER, PERERA, 2020; KARPOWITZ, KENNER, 2016).

A escola e a performance escolar nunca foram tão importantes quanto em tempos atuais, quando o país e o mundo passam por uma crise sanitária sem precedentes na história recente da humanidade. A pandemia do novo coronavírus impactou uma ampla gama de indicadores sociais e econômicos (ONYEMA *et al.*, 2020) e prejudicou o acesso à educação básica no Brasil. A recuperação econômica e social pós-pandêmica requer, entre outras ações, a retomada do acesso à educação, a redução das desigualdades no ensino e o aprimoramento de formas de apoio aos alunos para melhores performances. Assim, o estudo dos impactos de programas que visem melhorar a educação básica é de grande importância e permitirá a elaboração e desenho mais informado de políticas públicas e estratégias para os anos futuros.

No Brasil, pode-se citar o Programa Novo Mais Educação (PNME), instituído pela Portaria MEC nº 1.144/2016 e regido pela Resolução FNDE nº 17/2017. O PNME vigorou entre 2017 e 2019 e envolvia o repasse de recursos da União para que escolas ampliassem suas jornadas em 5 ou em 15 horas semanais, com foco em atividades de acompanhamento pedagógico em Língua Portuguesa e Matemática.

Tal expansão da carga horária escolar, porém, acarreta custos importantes aos cofres públicos, de forma que os retornos acadêmicos associados a esse programa devem ser compatíveis com os investimentos realizados. Assim, o presente estudo responderá especificamente à seguinte pergunta: Quais são os impactos do Programa Novo Mais Educação sobre o desempenho acadêmico nas escolas públicas brasileiras?

Conhecer clara e precisamente os efeitos de políticas públicas, especialmente aquelas com potencial formativo e de longo prazo como políticas educacionais, é do interesse de toda a sociedade. Políticas como o PNME são financiadas através de recursos públicos e seus benefícios sociais e econômicos – ou a falta dos mesmos – afetam toda a população, direta e indiretamente, através de mecanismos como a redução de desigualdades (e, logo, do dispêndio público destinado ao combate a esse problema) e crescimento econômico (BLUNDELL, DEARDEN, MEGHIR, SIANESI, 1999). Ademais, o presente estudo subsidiará o debate acerca das melhores práticas de incentivo à educação.

Como hipótese, espera-se que a maior carga escolar dos alunos e o decorrente aumento do investimento público nas escolas contempladas pelo programa levem a diferenças positivas e significativas nas notas dos alunos das escolas que migraram para o regime de ensino em tempo integral em relação às escolas que não o fizeram. Apesar da literatura internacional e doméstica vigente sobre o tema não respaldar essa suposição (ALFARO, EVANS, HOLLAND, 2015), apresentando impactos muitas vezes insignificantes ou até negativos no desempenho escolar de programas deste tipo, a atenção especial dada pelo PNME às disciplinas de Língua Portuguesa e Matemática, que são avaliadas pelo Sistema de Avaliação da Educação Básica (SAEB), pode contribuir para impactos mais assertivos do Programa nas notas dos alunos participantes.

Esse trabalho se divide em cinco capítulos. O Capítulo 1 apresenta os objetivos geral e específicos com a delimitação do seu escopo. O Capítulo 2 apresenta um resumo da literatura acerca dos retornos da educação em tempo integral no Brasil e no mundo, as normativas que regem o Programa Novo Mais Educação e os conceitos fundamentais por trás de avaliações de impacto de políticas públicas. O Capítulo 3 descreve as bases de dados e as metodologias empregadas. O Capítulo 4 compara os balanceamentos encontrados e impactos estimados através de cada um dos métodos utilizados. Por fim, O Capítulo 5 resume os resultados, discute as limitações do estudo e conclui.

## 1.1. Objetivos geral e específicos

O objetivo geral do presente estudo é avaliar os impactos da promoção da educação em tempo integral no Brasil sobre o desempenho acadêmico dos alunos das escolas públicas nacionais que aderiram ao Programa Novo Mais Educação, medido pelas notas em Língua Portuguesa e Matemática das instituições.

Para tal, foram estabelecidos como objetivos específicos:

- Construir os grupos de tratamento e de controle com base nos dados do Ministério da Educação (MEC) e do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP);
- Construir uma função que defina a probabilidade de uma escola participar ou não do programa, com base nos requisitos estabelecidos pelo MEC;
- Construir modelos de *propensity score matching*, de *Mahalanobis matching* e de *genetic matching* para parear os grupos de controle e tratamento;
- Comparar o balanceamento das covariadas de interesse nos três pareamentos;
- Apresentar os impactos estimados do PNME sobre as notas de Língua Portuguesa e Matemática das escolas tratadas, para cada forma de pareamento.

## 1.2. Delimitação do Escopo do Estudo

Este estudo avaliou apenas os impactos acadêmicos do ensino em tempo integral nas notas de Língua Portuguesa e Matemática de alunos das escolas participantes do Programa Novo Mais Educação entre os anos de 2017 e 2019, a nível nacional, sem adentrar em análises regionais ou de seu custo-benefício ou outras métricas de eficiência. Tais estudos são fundamentais e merecem ser abordados em trabalhos futuros.

## 2. FUNDAMENTAÇÃO TEÓRICA

O desempenho escolar de um indivíduo está associada não só à sua renda futura (HARMON, OOSTERBEEK, WALKER, 2003) como também à sua participação cívica (DEE, 2004), ao crescimento econômico local como um todo (BLUNDELL, DEARDEN, MEGHIR, SIANESI, 1999) e a outros indicadores sociais e econômicos de importância, atingindo até a sua probabilidade de cometer um crime (MCEACHIN, LAUEN, FULLER, PERERA, 2020; KARPOWITZ, KENNER, 2016).

Diversas variáveis têm sido apontadas como determinantes do resultado no desempenho acadêmico de um aluno. Schiefelbein e Simmons (1980), em uma revisão de vinte e seis estudos sobre o tema, listaram 123 fatores que apresentaram resultado estatisticamente significativo no desempenho do aluno na literatura, incluindo variáveis como o número de alunos na sala de aula, a formação e experiência do professor e a situação socioeconômica do aluno, entre outros. Apesar desse grande número de indicadores já identificados, novos determinantes têm sido encontrados e analisados à medida que técnicas de regressão são aprimoradas e refinadas, incluindo, por exemplo, novas técnicas de regressão espacial (RIANI, 2005).

Além dos indicadores mencionados acima, um fator comumente associado ao desempenho acadêmico do aluno é a modalidade de ensino em tempo integral. O ensino em dois turnos implica em uma elevação da carga horária ofertada para o aluno na escola, de forma que se torna natural esperar ganhos acadêmicos (MENEZES-FILHO, 2007). Contudo, as evidências empíricas sobre o tema são heterogêneas (ALFARO, EVANS, HOLLAND, 2015). Além disso, uma particularidade pouco destacada do regime de ensino em tempo integral é o seu custo consideravelmente maior do que o do ensino em turno único (COSTA, 2012; CERDAN-INFANTES, VERMEERSCH, 2007). Assim, torna-se importante a contextualização dos retornos associados à educação em tempo integral, levando em conta não apenas seus efeitos no desempenho acadêmico dos alunos, mas também os custos associados.

É importante, porém, destacar que a educação em tempo integral acarreta mudanças estruturais mais intensa do que os demais determinantes de desempenho acadêmico identificados e, portanto, traz benefícios para os seus alunos que vão além dessa métrica. Silva et al. (2012) destacam, através de um questionário aplicado a 175 alunos do ensino fundamental, melhoras à socialização e organização dos entrevistados atribuídas ao ensino em tempo integral. De forma semelhante, Ostler (2017) ressalta os benefícios de inclusão social de alunos com deficiências relacionados ao melhor atendimento prestado a estes num contexto de ensino em tempo integral.

O debate acerca dos determinantes do desempenho acadêmico e sobre quais estratégias mais influenciam essa métrica se encontra atualmente mais proeminente do que nunca. Os efeitos globais da pandemia do novo coronavírus prejudicaram intensamente os mais amplos indicadores econômicos e sociais tanto no Brasil quanto no mundo, inclusive no que diz respeito à educação básica. Onyema *et al.* (2020) aplicaram um questionário em 200 alunos, familiares, docentes e gestores públicos de diversos países e reportaram perceptíveis reduções de acesso à educação e à infraestrutura escolar devido à pandemia, especialmente em escolas que atendem alunos de menor nível socioeconômico, assim como do aumento do endividamento estudantil. Os autores também destacam sérios choques educacionais que podem provocar impactos de longo prazo no desempenho acadêmico dos alunos.

As dificuldades impostas pela pandemia, porém, não se limitam à esfera social. Alves, Farenzena, Silveira e Pinta (2020) analisam a elasticidade entre as receitas tributárias dos estados e municípios e de suas verbas destinadas à educação básica. Os autores estimaram que a receita aluno-mês (montante do orçamento público despendido mensalmente por aluno) caiu em até 26,9% entre 2018 e 2021 no Brasil em casos extremos, com graves consequências para os alunos e os corpos docentes das instituições públicas nacionais de ensino.

Neste contexto, torna-se fundamental avaliar o impacto dos programas públicos de fomento à educação, em especial daqueles que requerem maior dispêndio como os de

incentivo ao ensino em tempo integral, de forma a subsidiar o debate acerca das melhores estratégias para recuperar as perdas que o coronavírus já trouxe e ainda trará.

## 2.1. Literatura internacional e doméstica acerca do impacto da educação em tempo integral

A literatura internacional acerca do impacto da educação em tempo integral encontra efeitos heterogêneos dessa modalidade de ensino sobre o desempenho acadêmico dos alunos participantes.

Alfaro, Evans e Holland (2015) realizaram uma revisão bibliográfica de 19 estudos sobre o tema na América Latina e no Caribe. Entre os doze estudos que mediram os efeitos da educação em tempo integral sobre as notas dos alunos, apenas cinco encontraram impactos positivos e significantes no rendimento em matemática. Quatro estudos não encontraram efeitos significativos e três estudos estimaram quedas no rendimento nesta disciplina. Já para as notas em línguas, os resultados foram mais positivos. Oito estudos apontaram melhores resultados após a migração para o ensino em tempo integral e apenas dois identificaram quedas de desempenho, embora não significativas. Além disso, entre os sete estudos que analisam outros indicadores, como a taxa de evasão escolar ou a probabilidade de se perseguir uma educação superior, apenas três encontraram impactos positivos da educação em tempo integral em suas variáveis de interesse, com outros três sendo nulos ou não significativos e um trazendo resultados negativos.

Em um estudo no Chile, Bellei (2009), encontrou aumentos significativos tanto nas notas de matemática quanto de línguas, sendo semelhantes entre as disciplinas. O autor destaca que os efeitos da educação em tempo integral foram mais intensos em alunos rurais ou de escolas públicas, assim como em alunos que já possuíam uma alta performance em relação aos demais.

Esse resultado é corroborado por Cerdan-Infantes e Vermeersch (2007), que também encontram efeitos positivos nas notas de matemática e de línguas em escolas públicas de baixa renda no Uruguai que adentraram um programa governamental de subsídio ao ensino em tempo integral. Os ganhos em matemática, porém, foram mais intensos do que em línguas. Os autores ressaltam ainda que os custos associados à migração da escola para a educação em tempo integral são elevados.

Dentro da esfera nacional, os estudos sobre os efeitos do ensino em tempo integral na performance dos alunos são ainda menos otimistas. Em um estudo que analisa os efeitos do Programa Mais Educação (ação precursora ao Programa Novo Mais Educação, objeto do presente estudo) sobre o rendimento escolar dos alunos de escolas do Rio Grande do Sul, Xerxenevsky (2012) não encontrou mudanças nas notas de matemática ou português entre alunos do 8º ano. Entre alunos do 4º ano, houve um efeito positivo nas notas de português, porém negativo em matemática.

Os resultados de Oliveira e Terra (2018), que também examinaram o mesmo programa, porém utilizando uma estratégia de identificação baseada na proporção de alunos participantes do Programa Bolsa Família, mostraram efeitos nulos para todas os anos e disciplinas investigados.

## 2.2. O Programa Novo Mais Educação

O Programa Novo Mais Educação (PNME) foi instituído pela Portaria MEC nº 1.144/2016<sup>2</sup> e foi regido pela Resolução FNDE nº 17/2017, posteriormente alterada pela Resolução FNDE nº 17/2018. O Programa operou entre 2017 e 2019 e teve como objetivo geral o aprimoramento da aprendizagem de Língua Portuguesa e de Matemática no Ensino Fundamental através da ampliação da jornada escolar, podendo esta ser realizada através da complementação de cinco ou de quinze horas semanais, a critério da escola.

---

<sup>2</sup> Disponível em: <<http://portal.mec.gov.br/docman/outubro-2016-pdf/49121-port-1145-11out-pdf/file>>. Acesso em: 16 abr. 2021.

Para as escolas optantes por cinco horas semanais adicionais, instituía-se obrigatoriamente duas atividades designadas “atividades de acompanhamento pedagógico”, com carga semanal de duas horas e meia cada, sendo uma de Língua Portuguesa e a outra de Matemática. Já as escolas optantes por quinze horas semanais adicionais deviam também oferecer as mesmas duas atividades de acompanhamento pedagógico, porém com duração de quatro horas por semana. Além disso, essas escolas selecionavam três atividades complementares dentro das opções fornecidas pelo Programa, como aulas de música, esportes e afins, a serem declaradas através do Sistema PDDE (Programa Dinheiro Direto na Escola) Interativo<sup>3</sup>.

Destaca-se aqui que a adesão ao PNME, na modalidade de quinze horas semanais, qualificava a escola como uma instituição de ensino em tempo integral, atendendo ao requisito de uma carga horária mínima de sete horas.

Para participar do Programa Novo Mais Educação, as escolas eram selecionadas através de três critérios de adesão, sendo eles:

- Ter recebido recursos do PDDE Educação Integral nos últimos três anos (i.e., foram contempladas no Programa Mais Educação, descrito a seguir, que se encerrou em 2016, ou no Programa Novo Mais Educação, a partir de 2017);
- Apresentar Índice de Nível Socioeconômico “baixo” ou “muito baixo”, segundo classificação do INEP;
- Apresentar baixo desempenho no IDEB.

Conforme mencionado, o Programa Novo Mais Educação foi precedido pelo Programa Mais Educação, em vigor de 2008 até 2016, que possuía teor semelhante. O Programa Mais Educação foi instituído pela Portaria Interministerial nº 17/2007<sup>4</sup> e regulamentado pelo

---

<sup>3</sup> O Sistema PDDE Interativo é um portal desenvolvido para o envio de documentação, registro e prestação de contas de recursos do Programa Dinheiro Direto na Escola. Disponível em: <<http://pddeinterativo.mec.gov.br/>>. Acesso em: 16 abr. 2021.

<sup>4</sup> Disponível em: <[http://portal.mec.gov.br/index.php?option=com\\_docman&view=download&alias=2446-port-17-120110&category\\_slug=janeiro-2010-pdf&Itemid=30192](http://portal.mec.gov.br/index.php?option=com_docman&view=download&alias=2446-port-17-120110&category_slug=janeiro-2010-pdf&Itemid=30192)>. Acesso em: 16 abr. 2021.

Decreto 7.083/2010<sup>5</sup> e também tinha como objetivo a melhora da aprendizagem por meio da oferta de ensino em tempo integral. Contudo, diferentemente do PNME, que requeria atividades de acompanhamento pedagógico focadas em Língua Portuguesa e Matemática, o Programa Mais Educação exigia apenas que as escolas participantes oferecessem uma dessas duas disciplinas, com duração de seis horas semanais. Ademais, o Programa Mais Educação também demandava a implementação de três atividades complementares, porém sem delimitação específica de carga horária, contanto que a escola atingisse o requisito de sete horas diárias de aula.

Finalmente, a última diferença de destaque entre os dois programas se refere aos critérios de adesão. O Programa Mais Educação priorizava escolas que já haviam sido contempladas no programa em anos anteriores, similarmente ao Programa Novo Mais Educação. Seus critérios perante o IDEB da escola eram mais específicos, exigindo uma nota inferior a 4,6 nos anos iniciais do Ensino Fundamental e inferior a 3,9 nos anos finais. Ademais, ao invés de requerer baixo nível socioeconômico pelas métricas do INEP, o Programa Mais Educação priorizava escolas com 50% ou mais de seus alunos inseridos no Bolsa Família, além de ter critérios específicos para escolas rurais.

Assim, existiam claras distinções entre o Programa Novo Mais Educação e seu predecessor. O PNME possuía como clara diretriz o fortalecimento da aprendizagem em língua portuguesa e matemática, demandando de suas escolas participantes a realização de atividades de acompanhamento pedagógico obrigatoriamente nessas duas matérias, enquanto o Programa Mais Educação apresentava um escopo mais amplo.

Apesar do PNME ter sido sujeito a críticas quanto à sua efetividade em de fato contribuir para a formação educacional e pessoal dos alunos de forma mais ampla e não apenas melhorar os indicadores educacionais desejados (MATUOKA, 2017), o maior foco do PNME traz consigo indagações interessantes. Afinal, houve uma reformulação e uma

---

<sup>5</sup> Disponível em: <[http://www.planalto.gov.br/ccivil\\_03/\\_Ato2007-2010/2010/Decreto/D7083.htm](http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7083.htm)>. Acesso em: 16 abr. 2021.

mudança de trajetória do Programa com maior atenção dada às disciplinas avaliadas pelo SAEB e que compõem o resultado da escola no IDEB. Kelleghan, Madaus, Airasian (1982) já destacavam na década de 80 que o desvio do foco das instituições de ensino para avaliações seriadas – mesmo fenômeno adotado pelo Programa Novo Mais Educação em relação ao seu predecessor – podia levar à rotulagem dos alunos baseado em seu desempenho acadêmico e aumentar o nível de reprovação de alunos de famílias de menor renda.

Assim, torna-se necessário avaliar se tal mudança de foco se traduziu efetivamente em efeitos positivos tangíveis no desempenho escolar das instituições participantes, sendo esse o objetivo deste estudo.

### 2.3. Estratégia de avaliação de impacto de políticas públicas

Em qualquer avaliação de impacto de um tratamento, seja ele uma política pública, teste clínico, intervenção econômica, ou afins, o maior desafio se resume em isolar adequadamente os efeitos do tratamento das demais variações causadas por outros fatores, como por exemplo, o tempo, as mudanças no perfil da unidade analisada ou as mudanças estruturais no ambiente onde essa unidade existe.

Desta forma, a única avaliação perfeita seria o resultado da comparação entre um cenário onde a unidade analisada recebeu o tratamento e aquele onde a mesma unidade não o recebeu. Tal comparação, porém, é logicamente impossível. O procedimento factível é contrastar a unidade tratada com alguma unidade semelhante que não recebeu o tratamento (CALIENDO, KOPEINING, 2005). Esse fenômeno é denominado o problema do contrafactual não observado.

Considere que  $Y_0$  represente o produto potencial das unidades analisadas que não receberam o tratamento e  $Y_1$  o das unidades que o receberam. Por exemplo,  $Y_0$  pode representar aqui o desempenho acadêmico médio das escolas caso não fossem contempladas pelo Programa Novo Mais Educação. Similarmente,  $Y_1$  representa o desempenho das mesmas

caso recebessem recursos do Programa. As escolas que foram contempladas pelo PNME são designadas como  $T = 1$ , enquanto as não contempladas são identificadas como  $T = 0$ .

Uma avaliação de impacto do Programa objetiva identificar qual foi o impacto sobre o desempenho acadêmico da escola exclusivamente pelo fato dela ter sido contemplada pelo PNME. Assim, seria necessário comparar o desempenho das escolas que receberam recursos do PNME com o desempenho teórico dessas mesmas escolas caso não tivessem sido contempladas pela intervenção. Esse conceito é conhecido como o *Average Treatment Effect on Treated* (ATT), ou o efeito médio do tratamento sobre os tratados, e pode ser escrito como:

$$ATT = E[Y_1 | T = 1] - E[Y_0 | T = 1] \quad (1)$$

Mas conforme destacado anteriormente,  $E[Y_0 | T = 1]$ , ou seja, o desempenho da escola que participou do Programa caso não o houvesse feito, não é observável. Os únicos potenciais mensuráveis são por definição  $E[Y_0 | T = 0]$  e  $E[Y_1 | T = 1]$ . Assim, a única diferença observável se dá por:

$$E[Y_1 | T = 1] - E[Y_0 | T = 0] \quad (2)$$

A diferença entre esses dois resultados (o que deveria ser mensurado, mas não é observável, e o que de fato é observável) gera o chamado “viés de seleção”, obtido ao subtrair (1) de (2):

$$V = (E[Y_1 | T = 1] - E[Y_0 | T = 0]) - (E[Y_1 | T = 1] - E[Y_0 | T = 1]) \quad (3)$$

$$V = E[Y_0 | T = 1] - E[Y_0 | T = 0] \quad (4)$$

Esse viés de seleção representa a heterogeneidade inata nos resultados das unidades de análise mesmo desconsiderando os efeitos do tratamento (ANGRIST, PISCHKE, 2009). A

amplitude desse viés é intensificada quando o grupo de controle é inadequado, normalmente devido a diferenças importantes entre esse grupo e o grupo de tratamento antes da intervenção (BIONDI, VASCONCELLOS, MENEZES-FILHO, 2009). Assim, torna-se essencial adotar medidas que reduzam ao máximo esse viés de seleção, de forma a garantir uma avaliação adequada do efeito da intervenção sobre a variável de interesse, como as que serão descritas na próxima seção.

### 3. METODOLOGIA

Para avaliar o impacto do ensino em tempo integral ofertado pelo Programa Novo Mais Educação, adotou-se como variável de interesse as notas de Língua Portuguesa e Matemática dos alunos dos anos iniciais (1º a 5º ano) e finais (6º a 9º ano) do Ensino Fundamental brasileiro. A seguir, optou-se pelo uso de três modelos para a comparação entre o grupo intervenção, ou seja, as escolas que participaram do PNME, e o grupo controle composto por escolas semelhantes que não participaram do programa.

O primeiro foi o de *propensity score matching* (ROSENBAUM, RUBIN, 1983), que permite comparar cada unidade do grupo de tratamento não apenas com o conjunto de unidades do grupo de controle, mas com a(s) unidade(s) do grupo de controle cujas características mais se assemelham à unidade de tratamento, através da construção de um escore de propensão a receber o tratamento. Assim, as unidades de tratamento são pareadas com unidades do grupo controle com escores semelhantes, reduzindo o viés de seleção da amostra (KALLUS, 2016).

O pareamento por escores de propensão, porém, é alvo de diversas críticas<sup>6</sup>. Assim, utilizaram-se também dois outros modelos de pareamento: pareamentos por distância de Mahalanobis (SEKHON, 2011), que relaciona as unidades de tratamento e controle com base na proximidade de cada uma de suas covariadas (ao invés de um escore único estimado por regressão logística), e por *genetic matching* (SEKHON, 2011; DIAMOND, SEKHON, 2013), um

---

<sup>6</sup> As principais críticas são detalhadas na Seção 3.3.2 deste trabalho.

modelo iterativo que incorpora elementos de ambos *propensity score matching* e *Mahalanobis matching* e que refina iterativamente uma matriz de pesos aplicada ao vetor de covariadas de forma a maximizar o balanceamento do pareamento final.

As Seções 3.1 e 3.2 descrevem os dados utilizados e o universo de escolas analisadas. As Seções 3.3, 3.4 e 3.5 detalham cada um dos métodos utilizados.

### 3.1. Bases de dados

Para realização do estudo foram utilizados os resultados do Índice de Desenvolvimento da Educação Básica (IDEB) de 2015, 2017 e 2019, os resultados do Indicador de Nível Socioeconômico das Escolas de Educação Básica (INSE) de 2015 e os dados provenientes do Censo Escolar de 2016, 2017 e 2019, realizados pelo Instituto Nacional de Estudos e Pesquisas Anísio Teixeira (INEP). Foram utilizadas ainda a relação das escolas contempladas pelo Programa Mais Educação em 2016, obtida a partir do Censo Escolar daquele ano, e a relação daquelas participantes do Programa Novo Mais Educação, fornecida pelo Ministério da Educação (MEC), conforme apresentado na Tabela 1.

Tabela 1 – Síntese dos dados utilizados e suas fontes

Dado	Descrição	Fonte
Censo Escolar	Dados socioeconômicos e de infraestrutura das escolas. Foram utilizados os Censos de 2016, 2017 e 2019.	INEP
IDEB	Resultados por escola em 2015, 2017 e 2019 da taxa de aprovação, notas do SAEB em língua portuguesa e matemática, e nota final do IDEB.	INEP
INSE	Indicador de Nível Socioeconômico das Escolas de Educação Básica no ano de 2015.	INEP
Escolas contemplados pelo PME	Relação de escolas contempladas pelo PME em 2016.	Censo Escolar/INEP
Escolas contempladas pelo PNME	Relação de escolas contempladas pelo PNME em 2017 e 2019.	MEC

Fonte: Elaboração própria.

Os resultados do IDEB trazem as notas finais por escola para as provas de Língua Portuguesa e Matemática do Sistema de Avaliação da Educação Brasileira (SAEB), tanto para

os anos iniciais (1º a 5º ano) como para os anos finais (6º a 9º ano) do Ensino Fundamental. Ademais, a base traz informações sobre as taxas de aprovação em cada ano matricular, para cada escola. A combinação das notas finais nas duas disciplinas mencionadas acima com a taxa de aprovação média da instituição produz o IDEB daquela escola. Foram obtidos os dados do IDEB, tanto do índice final quanto de seus componentes, referentes aos anos de 2017 e 2019, últimos dois anos de divulgação do SAEB, que é bianual, sendo coincidentes com o período de vigência do Programa Novo Mais Educação, assim referentes ao ano de 2015, para fins de construção dos escores das escolas, conforme descrito nas próximas seções.

O Indicador de Nível Socioeconômico das Escolas de Educação Básica (INSE) servia como um dos critérios de adesão ao PNME, sendo assim um indicador importante para o pareamento ideal dos grupos de controle e de tratamento. Foi utilizada a base do INEP de 2015, último ano disponível antes da adesão das escolas ao PNME, que traz os indicadores calculados para 63.258 instituições de todo os país.

Já o Censo Escolar engloba todas as instituições de ensino nacionais, trazendo informações sobre seus números de turmas e alunos, infraestruturas, cursos oferecidos, capacitação de seus professores, e outras, sendo a mais rica fonte de informações socioeconômicas e demográficas sobre os centros educacionais do país. Dado o escopo do estudo, foram extraídos os microdados dos Censos Escolares de 2017 e 2019, de forma a oferecer variáveis de controle para subsidiar o modelo alinhadas aos dados dos resultados do IDEB descritos acima.

### **3.2. Universo de escolas do estudo**

Munidos dos dados desejados das escolas brasileiras, partiu-se para a delimitação do grupo de tratamento, i.e., o conjunto de escolas participantes do Programa Novo Mais Educação. Utilizou-se uma listagem das escolas integrantes do PNME nos anos de 2017 e 2019, fornecida pelo Ministério da Educação, analisando-se assim um prazo largo o suficiente para justificar a existência de impactos tangíveis no desempenho acadêmico das escolas

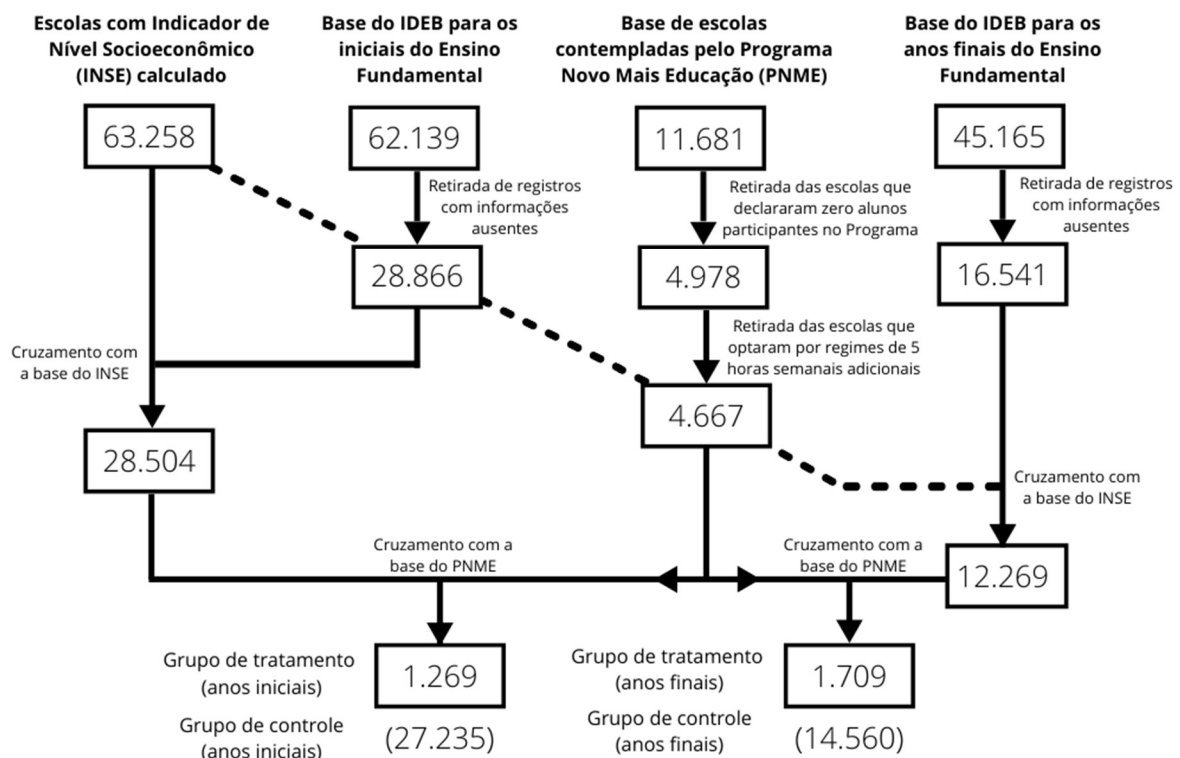
tratadas. Tal base de dados também contém informações sobre a modalidade do Programa optada pela escola (seja de cinco ou de quinze horas semanais adicionais), assim como o número de alunos participantes. No total, 11.681 escolas foram contempladas pelo PNME em ambos os anos. Dessas, 6.703 foram eliminadas do estudo por declarar um número de alunos nulo em um ou mais períodos. Das 4.978 remanescentes, apenas 311 (6,2%) optaram pela modalidade de cinco horas semanais adicionais. Face à restrita dimensão da amostra, optou-se por eliminar essas escolas da análise e avaliar apenas o impacto resultante da participação no PNME com quinze horas semanais adicionais.

Em seguida, obteve-se também a relação das escolas participantes do Programa Mais Educação, predecessor do PNME, em 2016, uma vez que o recebimento de recursos do PME era um dos critérios de seleção especificados na legislação do Programa Novo Mais Educação. Tal identificação foi realizada através do registro das turmas escolares no Censo Escolar daquele ano, uma vez que este inclui um campo que identifica a participação ou não da turma no PME. Foram consideradas como escolas participantes do PME todas as 10.449 escolas que possuíam pelo menos uma turma do Ensino Fundamental listada no Censo Escolar como participante.

Após o levantamento das bases de dados necessárias para a construção dos grupos de controle e tratamento, realizou-se o cruzamento. Entre as 62.139 escolas com informações sobre o resultado do IDEB dos anos iniciais do Ensino Fundamental, 33.273 foram eliminadas devido a campos faltantes, i.e., declarações com valores em branco em um ou mais anos. Outras 362 escolas foram eliminadas por não terem o seu INSE calculado em 2015. Das 28.504 instituições remanescentes, 1.269 foram contempladas com recursos do Programa Novo Mais Educação entre 2017 e 2019. Assim, o grupo de tratamento foi composto por estas 1.269 instituições, enquanto as demais 27.235 compuseram o grupo de controle. Destaca-se que devido à sua natureza censitária, todas as escolas selecionadas foram também encontradas no Censo Escolar.

Finalmente, realizou-se um exercício simétrico para os resultados do IDEB dos anos finais do Ensino Fundamental. De um total de 45.165 escolas constantes na base do IDEB, 28.896 foram eliminadas, com as demais 16.269 sendo divididas entre as 1.709 que foram contempladas pelo PNME e as 14.560 que não foram. Os procedimentos adotados na construção dos grupos de controle e tratamento estão ilustrados na Figura 1.

Figura 1 – Fluxograma dos tratamentos realizados nas bases de dados para a construção dos grupos de controle e tratamento



Fonte: Elaboração própria com base em dados do INEP e do MEC.

### 3.3. O modelo de propensity score matching

O modelo de *propensity score matching* (ROSENBAUM, RUBIN, 1983) busca resolver o problema do contrafactual não observado através do pareamento de unidades do grupo de tratamento com as unidades mais semelhantes existentes no grupo de controle. Tal processo

se dá através do *matching* direto, quando as variáveis de controle são poucas e discretas, ou através de um escore de propensão construído a partir de uma função de probabilidade de a unidade participar do tratamento, dado um vetor de variáveis covariadas de controle mais extenso, sintetizando-as em um único valor (XERXENEVSKY, 2012).

Assim, constrói-se uma função de probabilidade baseada nos requisitos necessários para receber o tratamento e aplica-se a mesma para cada observação, gerando-se uma distribuição dessa probabilidade dentro dos grupos de controle e tratamento. Em seguida, realiza-se um pareamento entre as observações de cada grupo de forma a comparar unidades tratadas com unidades de controle cujos escores se assemelhem o máximo possível. Assim, pareia-se cada observação no grupo de tratamento com uma observação no grupo de controle que possui uma probabilidade semelhante de ter sido tratada, mas não foi.

Formalmente, essa probabilidade condicional  $P(x)$  para uma determinada unidade ser selecionada para tratamento dado um conjunto de covariadas  $x$  pode ser descrita como

$$\hat{P}(x) = P(T_i = 1 | X_i = x) \quad (1)$$

e é estimada através de uma regressão *logit* ou *probit* (BIONDI, VASCONCELLOS, MENEZES-FILHO, 2009). Dessa forma, cada observação possui um escore que varia entre zero (nenhuma chance de tratamento) até um (certeza de tratamento).

É importante destacar aqui que existem três hipóteses fundamentais para a devida aplicação dos escores de propensão de pareamento: a de ignorabilidade do tratamento, a hipótese implícita de suporte comum e, como consequência destas duas primeiras, a de ignorabilidade forte generalizada, conforme Rosenbaum e Rubin (1983).

A hipótese de ignorabilidade do tratamento diz que o resultado esperado da unidade analisada independe do tratamento, dado seu conjunto de covariadas  $x$  (também independentes ao tratamento). Ela pode ser descrita como:

$$Y_{0,1} \perp T \mid x \quad (2)$$

A hipótese implícita de suporte comum implica que a probabilidade condicional nunca será nula ou unitária, ou seja, não existe conjunto de covariadas que permita determinar com exatidão se a unidade receberá ou não intervenção. Segundo Biondi, Vasconcellos e Menezes-Filho (2009), essa propriedade garante comparabilidade entre os grupos de controle e de tratamento. Formalmente,

$$0 < P(T = 1 \mid x) < 1 \quad (3)$$

Finalmente, dada a validade das primeiras duas hipóteses – ou seja, dada a existência da chamada ignorabilidade forte, Rosenbaum e Rubin (1983) apontam também uma terceira constatação: a ignorabilidade forte se mantém para qualquer escore de pareamento  $P(x)$ . Algebricamente,

$$Y_{0,1} \perp T \mid P(x) \quad (4)$$

Estas hipóteses são essenciais para minimizar o viés de seleção mencionado na Seção 2.3, garantido assim uma avaliação adequada do impacto do tratamento analisado. Através da comparação pelos escores de pareamento  $P(x)$  das unidades do grupo de controle e grupo de tratamento, é possível então calcular o *Average Treatment Effect* (ATE), ou efeito médio do tratamento, descrito por Xerxenevsky (2012) como:

$$ATE = E[Y_1 \mid P(x), T = 1] - E[Y_0 \mid P(x), T = 0] = E[Y_1 - Y_0 \mid P(x)] \quad (5)$$

Uma vez estabelecidas as hipóteses fundamentais para uma comparação não-viesada e estimada a função de probabilidade de se receber a intervenção dado o vetor de covariadas

$x$  escolhido, é necessário realizar testes de balanceamento dos grupos, de forma a verificar se as distribuições dos *propensity scores* são semelhantes entre os dois<sup>7</sup>. Em seguida, realiza-se o pareamento em si, comparando unidades do grupo de controle com aquelas do grupo de tratamento cujos escores de propensão ao tratamento sejam próximos.

Existem diversas formas de se realizar esse pareamento (XERXENEVSKY, 2012; KHANDKER, KOOLWAL, SAMAD, 2009). A mais simples é denominada *nearest-neighbour matching* e consiste simplesmente em comparar a unidade tratada com a unidade de controle com *propensity score* mais próxima. Tal exercício, porém, pode levar a diferenças desiguais nos escores de propensão de unidades distintas, i.e., é possível que haja algumas unidades de tratamento com diversas unidades de controle próximas e algumas onde essa distância seja larga.

Assim, outros métodos dividem o conjunto de unidades de controle de forma a prevenir essa possibilidade. O *radius matching* limita o intervalo máximo aceitável para pareamento, de forma que unidades tratadas que não possuam pares dentro desse intervalo sejam descartadas. Já o modelo de *interval matching*, ou modelo estratificado (BIONDI, VASCONCELLOS, MENEZES-FILHO, 2009) separa o conjunto de *propensity scores* observados em intervalos fixos (e.g. de 0,2 a 0,5 e de 0,5 a 0,8, em um exemplo com apenas dois estratos), garantindo que as unidades tanto de controle quanto de tratamento presentes dentro de cada estrato possuam médias semelhantes em seus escores de pareamento e descartando os demais intervalos.

Existe também o método de *kernel*, apontado por Caliendo e Kopeing (2005) como o mais robusto, onde cada unidade tratada é comparada com uma média ponderada de todas as unidades de controle. O peso das unidades de controle é inversamente proporcional à distância entre elas e a unidade tratada, ou seja, o método dá maior peso para as unidades de controle mais próximas.

---

<sup>7</sup> Para uma revisão dos principais testes utilizados na literatura, ver Lee (2013).

Em resumo, constrói-se uma função de probabilidade que gere escores de propensão balanceados e define-se uma metodologia de pareamento entre os grupos de controle e de tratamento, de forma a permitir que a diferença na variável de análise entre os grupos possa enfim ser calculada de forma não-viesada.

Para este trabalho, foi utilizado o método de *nearest-neighbour matching* com pareamento de uma unidade do grupo de tratamento para uma unidade do grupo de controle (*one-to-one matching*), permitindo-se reposição. Os pareamentos por *nearest-neighbour matching* minimizam seu erro quadrado médio quando o número de pares para cada observação do grupo de tratamento é pequeno, normalmente entre um e dois (AUSTIN, 2010), e o uso de reposição, onde uma unidade do grupo de controle pode ser pareada com múltiplas unidades do grupo de tratamento, tende a gerar pareamentos menos viesados em suas covariadas (KALLUS, 2016).

### 3.3.1. Estimação dos escores das escolas

A estimação dos escores de cada escola, ou seja, de sua probabilidade de ser ou não tratada, foi realizada através de um modelo *probit*<sup>8</sup> baseado nos critérios de seleção descritos no normativo do Programa Novo Mais Educação. Assim, foram utilizadas como variáveis independentes a participação da escola em 2016 no precursor Programa Mais Educação, assim como o INSE da instituição e sua nota no IDEB, ambos para o ano de 2015, últimos anos com dados disponíveis antes da adesão das escolas ao PNME. Primeiramente, foi realizado esse exercício para os anos iniciais do IDEB.

Observa-se uma relação negativa entre o INSE da escola e sua probabilidade de ser tratada, condizente com o esperado, dado os critérios de seleção do Programa Novo Mais Educação. Similarmente, as instituições que participaram do antigo Programa Mais Educação

---

<sup>8</sup> Para fins de comparação, foram utilizados também modelos *logit* com as mesmas especificações. Os resultados apresentados neste trabalho independem do modelo utilizado. Os resultados da regressão *logit* estão disponíveis no Quadro A, no Apêndice.

apresentaram maiores chances de participarem de seu sucessor. Já o coeficiente estimado para a nota do IDEB em 2015, porém, aponta uma relação contrária à esperada dado o normativo do PNME. De fato, escolas com notas mais altas no índice apresentaram uma *maior* chance de serem contempladas pelo Programa. Todos os parâmetros estimados foram significativos a um nível de confiança de 99% e o pseudo-R<sup>2</sup> estimado pelos desvios da regressão com covariadas contra a regressão nula (estimada apenas com o intercepto)<sup>9</sup> foi de 0,1797<sup>10</sup>.

Quadro 1 – Parâmetros estimados por regressão *probit* da probabilidade de tratamento, para os anos iniciais do IDEB

Variável preditiva	Variável dependente
	Participação da escola no PNME
Nota final da escola no IDEB em 2015	0.319*** (0.016)
Participação da escola no PME em 2016	0.150*** (0.040)
INSE da escola em 2015	-0.136*** (0.003)
Constante	2.983*** (0.129)
Observações	28.504
AIC	8.521,5

Significância: \*\*\* 1%, \*\* 5%, \* 10%

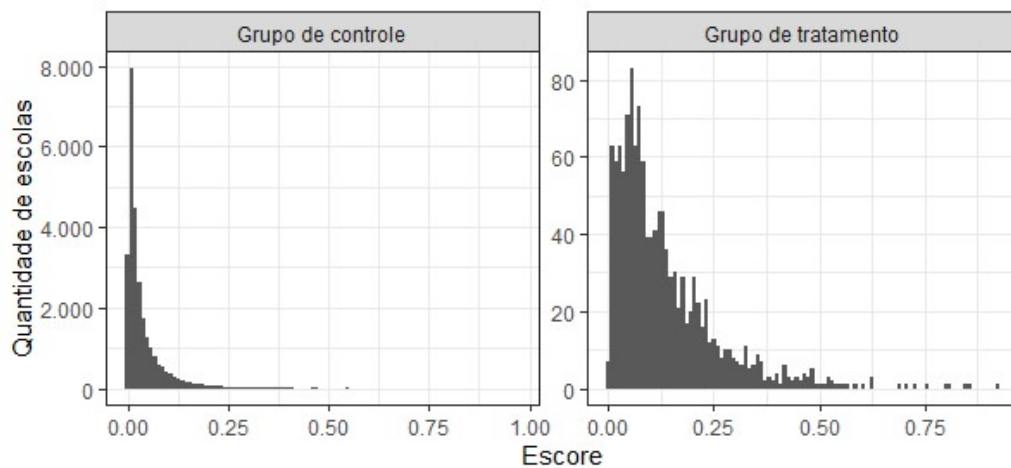
Fonte: Elaboração própria com base em dados do INEP e do MEC.

Assim, o score de cada escola foi dado pela sua probabilidade de receber o tratamento dado o valor de suas covariadas,  $0 < P(T = 1 | x) < 1$ . O Gráfico 1 traz a comparação da distribuição dos escores de cada grupo.

<sup>9</sup> O pseudo-R<sup>2</sup> é uma medida de *fit* para modelos de regressão não-linear, semelhante ao R<sup>2</sup> observado em regressões lineares. Um levantamento acerca das medidas de *fit* não-lineares está disponível em HAGLE, MITCHELL, 1992.

<sup>10</sup> O pseudo-R<sup>2</sup> encontrado para a regressão *logit* foi semelhante, sendo de 0,1732.

Gráfico 1 – Histogramas dos escores dos grupos de controle e de tratamento, para os anos iniciais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

É possível perceber que a curva dos escores dentro do grupo de tratamento pesa mais para a direita do que no grupo de controle, ou seja, as escolas que de fato participaram do PNME apresentaram de forma geral escores mais altos do que aquelas do grupo de controle. Essa constatação corrobora o *fit* do modelo logarítmico utilizado.

Em seguida, foi realizado o mesmo procedimento para os anos finais do IDEB. Os resultados foram bastante semelhantes, com todas as variáveis significativas a um nível de confiança de 99% e o mesmo escore crescente com aumentos na nota do IDEB da escola em 2015. O pseudo- $R^2$  da regressão foi de 0,2060 e os resultados aqui apresentados independem da escolha de modelagem por *probit* ou por *logit*<sup>11</sup>.

Quando avaliados os anos finais do IDEB, percebe-se uma concentração maior de escolas em escores de propensão mais elevados. Essa realidade é observada tanto no grupo de controle quanto no de tratamento, porém é mais intensa nesse segundo. Novamente, esse resultado reforça a adequação da regressão realizada.

<sup>11</sup> Os parâmetros estimados pela regressão *logit* estão disponíveis no Quadro B, no Apêndice. O pseudo- $R^2$  da regressão foi de 0,1986.

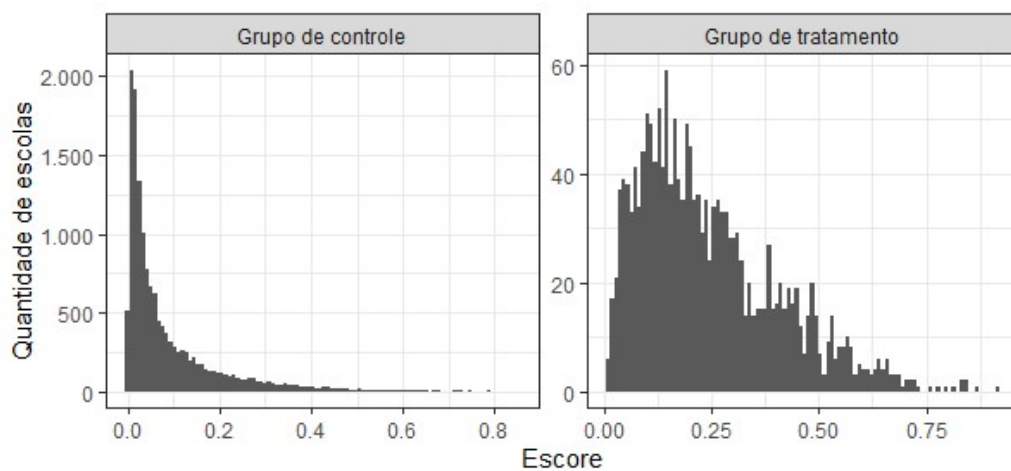
Quadro 2 – Parâmetros estimados por regressão *probit* da probabilidade de tratamento, para os anos finais do IDEB

Variável preditiva	Variável dependente
	Participação da escola no PNME
Nota final da escola no IDEB em 2015	0.335*** (0.020)
Participação da escola no PME em 2016	0.230*** (0.040)
INSE da escola em 2015	-0.145*** (0.003)
Constante	4.017*** (0.136)
Observações	16.269
AIC	8.689,4

© Significância: \*\*\* 1%, \*\* 5%, \* 10%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 2 – Histogramas dos escores dos grupos de controle e de tratamento, para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

### 3.3.2. Críticas acerca do modelo de propensity score matching

O pareamento por *propensity score matching* descrito nas seções anteriores é alvo de algumas críticas no debate acerca de avaliação de impacto de políticas públicas. As principais contestações se referem à hipótese de ignorabilidade do tratamento, afirmando que em cenários reais é muitas vezes discutível assumir que o resultado esperado da unidade analisada, bem como seu conjunto de covariadas, independam completamente da existência de tratamento (KING, NIELSEN, 2019). De fato, em muitas políticas públicas, como o próprio Programa Novo Mais Educação, importantes covariadas (como a nota do IDEB ou o INSE da escola) são utilizadas como critérios de adesão ao programa, de forma que se espera que sejam não independentes do tratamento, mas que possuam na verdade uma relação causal com ele.

Ademais, o pareamento por escores de propensão depende da construção desses mesmos escores, baseados em uma regressão logarítmica que define uma probabilidade de ser contemplado pelo tratamento. A verdadeira relação entre as covariadas e o tratamento raramente é conhecida de fato, sendo apenas estimada com base nas informações disponíveis (ANGRIST, PISCHKE, 2009). Dessa forma, é possível que a análise por *propensity score matching* seja viesada caso a regressão logarítmica realizada não capture adequadamente essa relação, como quando variáveis relevantes são omitidas na estimação dos escores.

Assim, optou-se por estimar os impactos do Programa Novo Mais Educação também através de pareamentos por distância de Mahalanobis e por *genetic matching*, baseados nos trabalhos de Sekhon (2011) e Diamond e Sekhon (2013). Os modelos utilizados, bem como as especificações utilizadas para cada, são detalhados nas seções seguintes.

### 3.4. O modelo de Mahalanobis matching

O pareamento por *propensity score matching* descrito nas seções anteriores transforma um conjunto de covariadas em um único valor (no caso, um escore de propensão ao tratamento) que é então utilizado para parear as unidades do grupo de tratamento com as do grupo de controle. Tal tratamento, porém, calcula a distância entre as unidades baseadas apenas no seus escores de propensão, e não nos valores individuais de cada covariada, de forma que uma ou mais variáveis podem permanecer muito distantes entre as duas unidades mesmo quando seus escores são próximos (SEKHON, 2011; KALLUS, 2016).

Assim, o pareamento por *Mahalanobis matching* surge como opção de pareamento multivariado que calcula a distância entre as covariadas individuais do grupo de tratamento e do grupo de controle (SEKHON, 2011). Assim, o *Mahalanobis matching* procura pares de unidades que tenham a menor distância em todas essas variáveis (a chamada distância de Mahalanobis), e não em um escore de propensão estimado por regressão logarítmica.

A distância de Mahalanobis é descrita por Diamond e Serhon (2013) como:

$$d(x, x') = \sqrt{(x - x')' S^{-1} (x - x')} \quad (6)$$

onde  $x$  e  $x'$  representam os vetores de covariadas das unidades pareadas e  $S$  representa a matriz de covariância amostral de  $x$ .

Uma vez estimada essa distância, realiza-se o pareamento das unidades de controle e de tratamento de forma simétrica ao descrito na seção anterior, porém servindo-se da distância de Mahalanobis calculada ao invés do escore de propensão estimado. Assim, utilizou-se nesse trabalho um *nearest neighbour matching* com reposição, pareando cada unidade do grupo de tratamento com uma única do grupo de controle (*one-to-one matching*).

O uso do pareamento através da distância de Mahalanobis permite circundar algumas das dificuldades associadas a escores de propensão, como a hipótese de ignorabilidade do

tratamento (KING, NIELSEN, 2019) e a aproximação correta de uma função de probabilidade que não é conhecida (ANGRIST, PISCHKE, 2009).

### 3.5. O modelo de genetic matching

O pareamento por *genetic matching* representa um método iterativo baseado em um algoritmo genético que procura maximizar o balanceamento das covariadas pareadas a cada ciclo ao refinar sucessivamente os pesos dados às covariadas (DIAMOND, SEKHON, 2013). Esse modelo é não paramétrico e independe do conhecimento dos escores de propensão de tratamento, semelhantemente ao *Mahalanobis matching*, mas pode incorporar tais escores, caso presentes, para aprimorar seus resultados (SEKHON, 2011).

O modelo de *genetic matching* utiliza uma medida de distância de Mahalanobis generalizada definida por Sekhon (2011) como:

$$d(x, x') = \sqrt{(x - x')' \left(S^{-\frac{1}{2}}\right)' W \left(S^{-\frac{1}{2}}\right) (x_i - x_i)} \quad (7)$$

onde  $W$  é uma matriz quadrada de pesos de tamanho  $k$  e com todos seus elementos não-diagonais nulos e  $S^{-\frac{1}{2}}$  é a decomposição de Cholesky da matrix de covariância  $S$ . Nota-se que, caso incorporados ao modelo, os escores de propensão estimados entram nos vetores de covariadas  $x$  e  $x'$ . Em um cenário onde esses escores sejam a única informação necessária para realizar um pareamento perfeitamente balanceado, os pesos dados a todas as demais covariadas seria nulo, de forma que a distância final seria simétrica à calculada por *propensity score matching*. Semelhantemente, caso a distância de Mahalanobis seja a mais adequada, o *genetic matching* vai apresentar uma matriz de pesos com todos os elementos de sua diagonal igual à unidade (exceto o referente ao escore de propensão, caso incorporado). Assim, tanto o *propensity score matching* quanto o *Mahalanobis matching* podem ser considerados casos específicos de *genetic matching* (SEKHON, 2011).

Efetivamente, o modelo de *genetic matching* encontra iterativamente o conjunto de parâmetros  $k$  de forma a maximizar o balanceamento dos pares finais, medidos através de testes  $T$  e de testes Kolmogorov-Smirnov não-paramétricos. Assim, o pareamento por *genetic matching* incorpora características tanto de *propensity score matching* quanto de *Mahalanobis matching*, sendo capaz de reduzir o erro quadrado médio de suas estimações em relação a ambos. Existem ganhos tangíveis em combinar *matchings* de escores de propensão com *matchings* de distância de Mahalanobis, uma vez que o primeiro é eficaz em minimizar a distância dos pares em relação aos seus escores (i.e., a um valor único que incorpora informações de seu conjunto de covariadas) enquanto o segundo minimiza a distância em relação aos valores individuais dessas covariadas (ROSENBAUM, RUBIN, 1985 *apud* DIAMOND, SEKHON, 2013).

## 4. RESULTADOS

Este capítulo apresenta e compara o balanceamento dos pareamentos de cada modelo e descreve os impactos estimados do Programa Novo Mais Educação para as notas de Língua Portuguesa e Matemática das escolas contempladas, para os anos iniciais e finais do Ensino Fundamental. A construção das bases de dados utilizadas e dos grupos de controle e tratamento, assim como os testes de balanceamento, o pareamento das unidades de análise e a estimação dos efeitos do tratamento foram realizados em R<sup>12</sup>. Para os testes de balanceamento, foi utilizado o pacote *Matching* (SEKHON, 2011), enquanto que para a estimação de impacto do Programa, abordada na seção 4.5, foram utilizados tanto o pacote *Matching* quanto o *MatchIt* (HO, IMAI, KING, STUART, 2011), tendo os coeficientes estimados se mostrado robustos à escolha de pacote.

---

<sup>12</sup> O R é uma linguagem de programação voltada para análises estatísticas e visualização de dados. Para maiores detalhes, acesse <<https://www.r-project.org/about.html>>. Acesso em: 16 abr. 2021.

## 4.1. Avaliação do pareamento por propensity score matching

Uma vez construídos os escores de probabilidade de tratamento, realizou-se o pareamento por *nearest-neighbour matching*, pareando cada escola do grupo de tratamento com uma única escola do grupo de controle cujo escore mais se aproximou do seu (AUSTIN, 2010). Conforme mencionado na Seção 3.3., foi permitido o uso de reposição, de forma que as escolas do grupo de controle podiam ser pareadas com múltiplas escolas do grupo de tratamento (KALLUS, 2016).

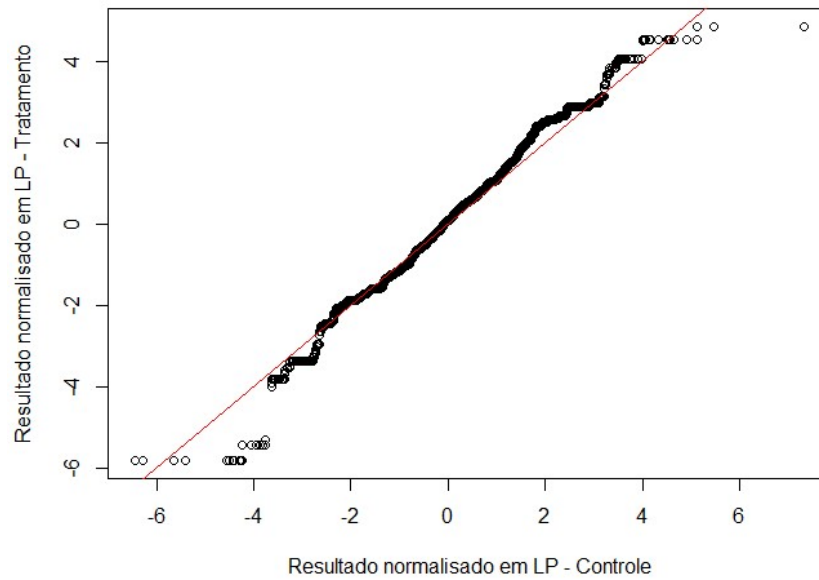
As variáveis de interesse foram as notas em Língua Portuguesa e Matemática, medidas pelo IDEB, das instituições observadas. De forma a capturar os efeitos do tratamento, optou-se pela diferença entre as notas da escola em cada uma das duas matérias em 2017, ano de adesão ao Programa Novo Mais Educação, e em 2019, após 3 anos de participação no mesmo<sup>13</sup>. Essa diferença foi normalizada e apresentada no restante deste trabalho em desvios-padrão da média de sua respectiva matéria.

Primeiramente, verificou-se a qualidade do pareamento para os anos iniciais do IDEB. Os Gráficos 3 e 4 trazem as visualizações Q-Q para cada uma das duas disciplinas avaliadas, ilustrando qual o resultado da unidade do grupo de tratamento, no eixo das ordenadas, contra o resultado de seu par no grupo de controle, no eixo das abscissas. Um pareamento perfeito traria todos seus pares localizados na diagonal do gráfico Q-Q, onde os resultados são iguais para ambos os grupos.

Gráfico 3 – Gráfico Q-Q do pareamento PSM para a diferença em Língua Portuguesa (em desvios-padrão), para os anos iniciais do IDEB

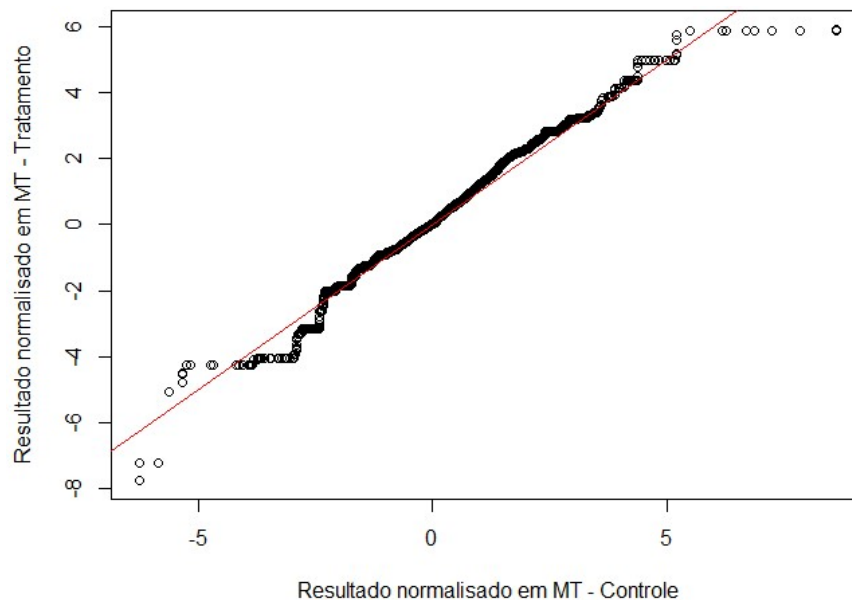
---

<sup>13</sup> Optou-se por não se calcular a diferença contra as notas em 2017, mesmo já tendo aderido ao PNME, e não contra as notas de 2015, para descartar ao máximo variações nas variáveis de interesse independentes ao Programa, como poderiam ser observadas no largo período entre as avaliações do IDEB em 2015 e adesão das escolas no PNME dois anos depois.



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 4 – Gráfico Q-Q do pareamento PSM para a diferença em Matemática (em desvios-padrão), para os anos iniciais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

É possível verificar que os pares tiveram resultados em sua maioria próximos em ambas as disciplinas. Perto das pontas, porém, onde as unidades apresentaram notas muito acima ou abaixo da média de todas as escolas, observou-se um distanciamento mais intenso.

A comparação das covariadas antes e depois do pareamento corrobora essa visualização, apontando perceptível aproximação em suas médias após o pareamento por *propensity score matching* (Tabela 3). Esse foi o caso para duas das três variáveis utilizadas na construção dos escores de probabilidade, com a exceção sendo as notas do IDEB das instituições em 2015, que manteve uma distância semelhante, porém com sinal inverso. Ainda assim, tanto as notas do IDEB quanto o INSE das escolas apontaram diferenças significativas em suas médias, com as a participação no PME em 2016 sendo a única covariada a apresentar um bom pareamento.

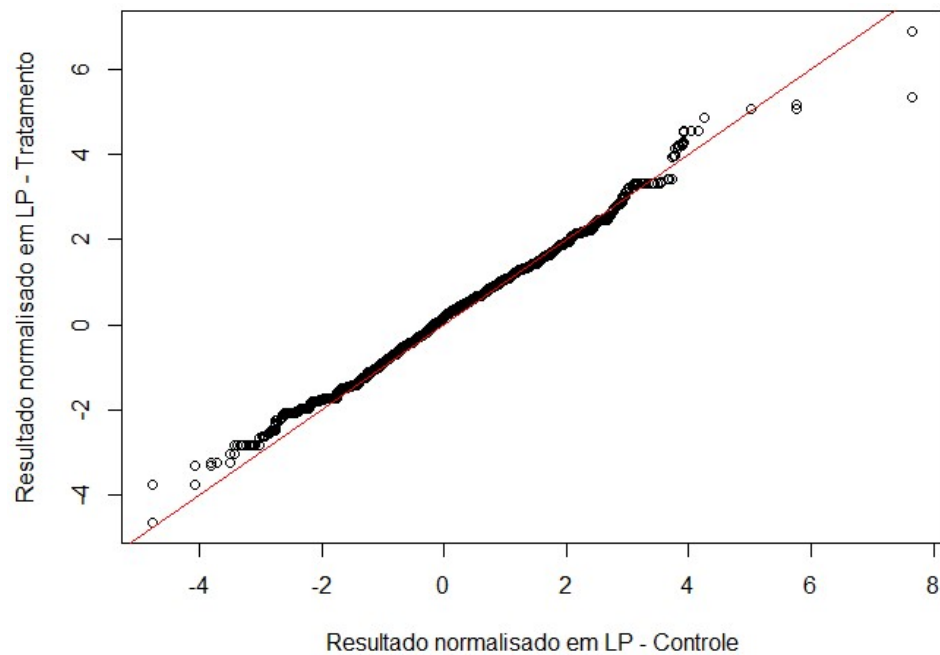
Tabela 2 – Análise do pareamento por PSM, para os anos iniciais do IDEB

Estatística	Nota do IDEB em 2015		Participação no PME		INSE da escola em 2015	
	Sem parear	Pareado	Sem parear	Pareado	Sem parear	Pareado
Média do tratamento	5,2552	5,2552	0,1734	0,1734	43,681	43,681
Média do controle	5,4194	5,0911	0,1182	0,1799	49,399	43,306
Diferença média padronizada	-0,1933	0,1931	0,1456	-0,0172	-1,622	0,106
Razão das variâncias	0,61	0,47	1,38	0,97	0,45	0,51
p-valor Teste T das médias	< 0,00	< 0,00	< 0,00	0,6691	< 0,00	< 0,00
p-valor KS Bootstrap das médias	< 0,00	< 0,00			< 0,00	< 0,00

Fonte: Elaboração própria com base em dados do INEP e do MEC.

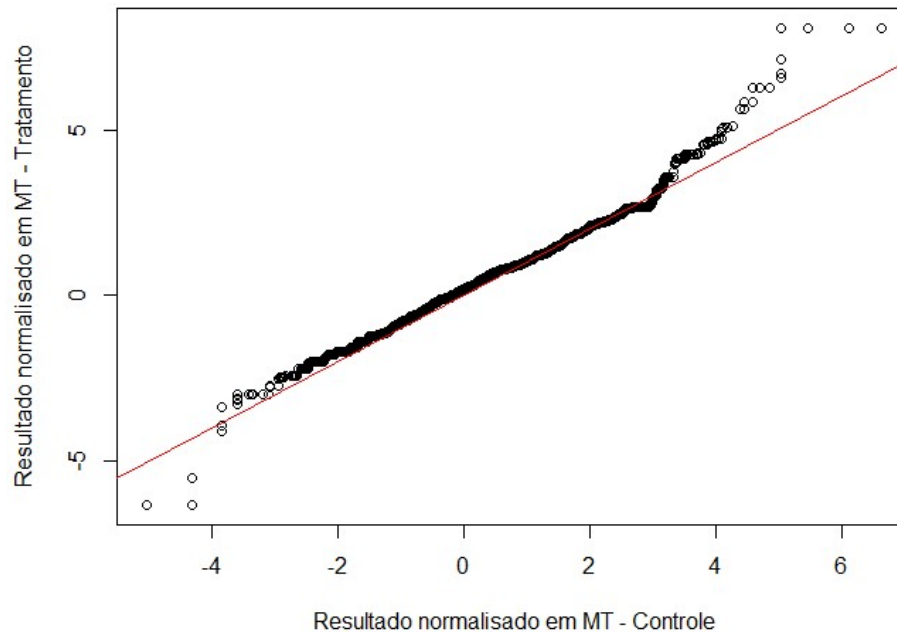
A mesma análise foi em seguida realizada para os anos finais do IDEB. As visualizações Q-Q foram bastante semelhantes (Gráficos 5 e 6), com resultados em ambas as disciplinas mais próximas entre os pares de menor desvio e mais distantes nas extremidades.

Gráfico 5 – Gráfico Q-Q do pareamento PSM para a diferença em Língua Portuguesa (em desvios-padrão), para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 6 – Gráfico Q-Q do pareamento PSM para a diferença em Matemática (em desvios-padrão), para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Por fim, a Tabela 4 traz as variáveis de balanceamento para o pareamento por *propensity score matching* para os anos finais do IDEB. Novamente, destaca-se uma notável aproximação em relação às médias não-pareadas, porém ainda muito distantes para que suas diferenças fossem significativamente nulas (com exceção, mais uma vez, da participação no antigo Programa Mais Educação).

Tabela 3 – Análise do pareamento por PSM, para os anos finais do IDEB

Estatística	Nota do IDEB em 2015		Participação no PME		INSE da escola em 2015	
	Sem parear	Pareado	Sem parear	Pareado	Sem parear	Pareado
Média do tratamento	4,2198	4,2198	0,1908	0,1908	43,325	43,325
Média do controle	4,3650	4,0872	0,1174	0,1958	49,284	43,029
Diferença média padronizada	-0,2055	0,1878	0,1867	-0,0129	-1,604	0,080
Razão das variâncias	0,64	0,63	1,49	0,98	0,49	0,68
p-valor Teste T das médias	< 0,00	< 0,00	< 0,00	0,7097	< 0,00	< 0,00
p-valor KS Bootstrap das médias	< 0,00	< 0,00			< 0,00	< 0,00

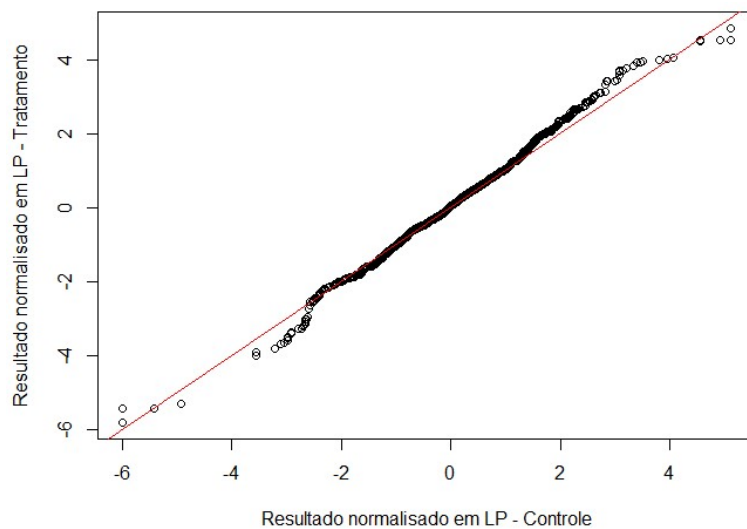
Fonte: Elaboração própria com base em dados do INEP e do MEC.

## 4.2. Avaliação do pareamento por distância de Mahalanobis

O pareamento por distância de Mahalanobis resultou, de forma geral, em uma maior aproximação entre as unidades do grupo de tratamento e as unidades do grupo de controle. As visualizações Q-Q para as diferenças normalizadas em Língua Portuguesa (Gráfico 7) e Matemática (Gráfico 8) nas escolas participantes do PNME e as demais pareadas dessa forma apontaram uma maior adesão à diagonal, embora ainda seja perceptível um descolamento próximo às extremidades. Destaca-se que as diferenças em Matemática apresentaram um distanciamento maior nos pares com desvios negativos do que as diferenças em Língua Portuguesa.

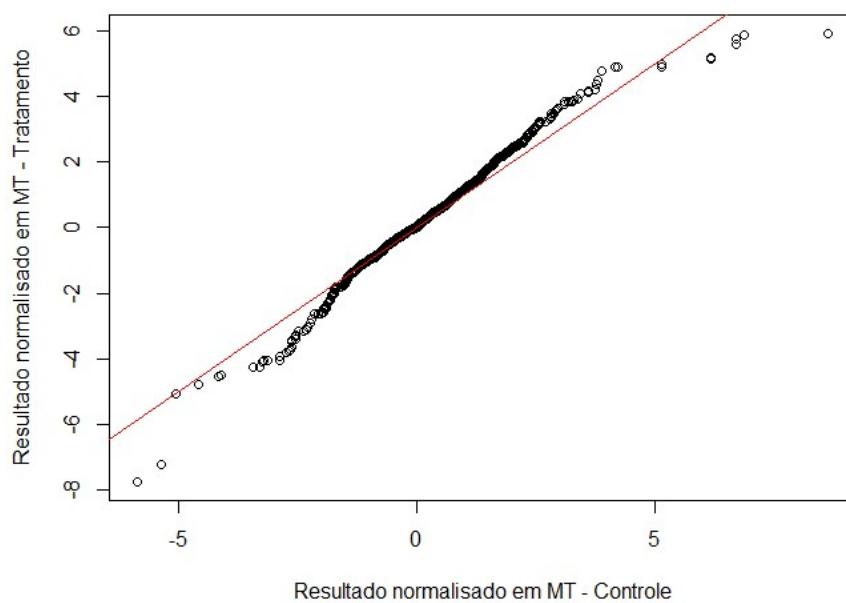
A análise das variáveis de balanceamento (Tabela 5) corrobora essa maior proximidade em relação ao pareamento por *propensity score matching*. Há uma aproximação nas médias das variáveis depois do pareamento por distância de Mahalanobis, cujas diferenças passaram a ser todas significativamente nulas. A razão das variâncias de todas as covariadas se aproxima perceptivelmente da unidade, o que aponta mais um indício de bom pareamento. Destaca-se, porém, que essa maior aproximação é de certa forma esperada, uma vez que o pareamento por distância de Mahalanobis relaciona as unidades dos grupos de controle e tratamento de forma a minimizar diretamente a distância entre suas covariadas (DIAMOND, SEKHON, 2013). Na Seção 4.4., mais adiante, é apresentada a comparação mais detalhada do balanceamento de cada pareamento.

Gráfico 7 – Gráfico Q-Q do pareamento por distância de Mahalanobis para a diferença em Língua Portuguesa (em desvios-padrão), para os anos iniciais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 8 – Gráfico Q-Q do pareamento por distância de Mahalanobis para a diferença em Matemática (em desvios-padrão), para os anos iniciais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

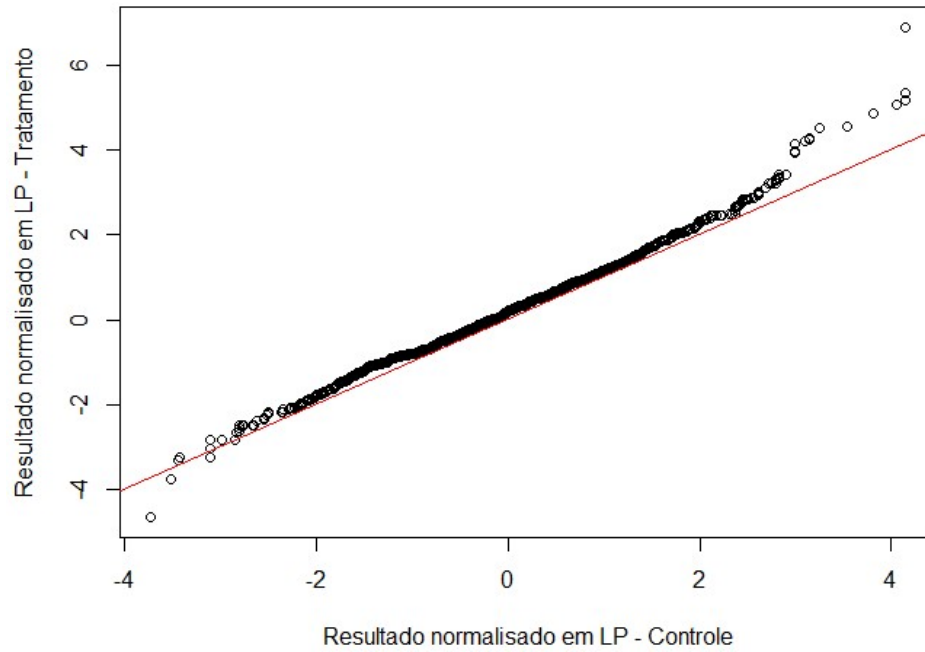
Tabela 4 – Análise do pareamento por distância de Mahalanobis, para os anos iniciais do IDEB

Estatística	Nota do IDEB em 2015		Participação no PME		INSE da escola em 2015	
	Sem parear	Pareado	Sem parear	Pareado	Sem parear	Pareado
Média do tratamento	5,2552	5,2552	0,1734	0,1734	43,681	43,681
Média do controle	5,4194	5,2544	0,1182	0,1734	49,399	43,688
Diferença média padronizada	-0,1933	0,0009	14,5580	0,0000	-1,622	-0,002
Razão das variâncias	0,61	1,00	1,38	1,00	0,45	1,01
p-valor Teste T das médias	< 0,00	0,63292	< 0,00	1	< 0,00	0,39815
p-valor KS Bootstrap das médias	< 0,00	1			< 0,00	1

Fonte: Elaboração própria com base em dados do INEP e do MEC.

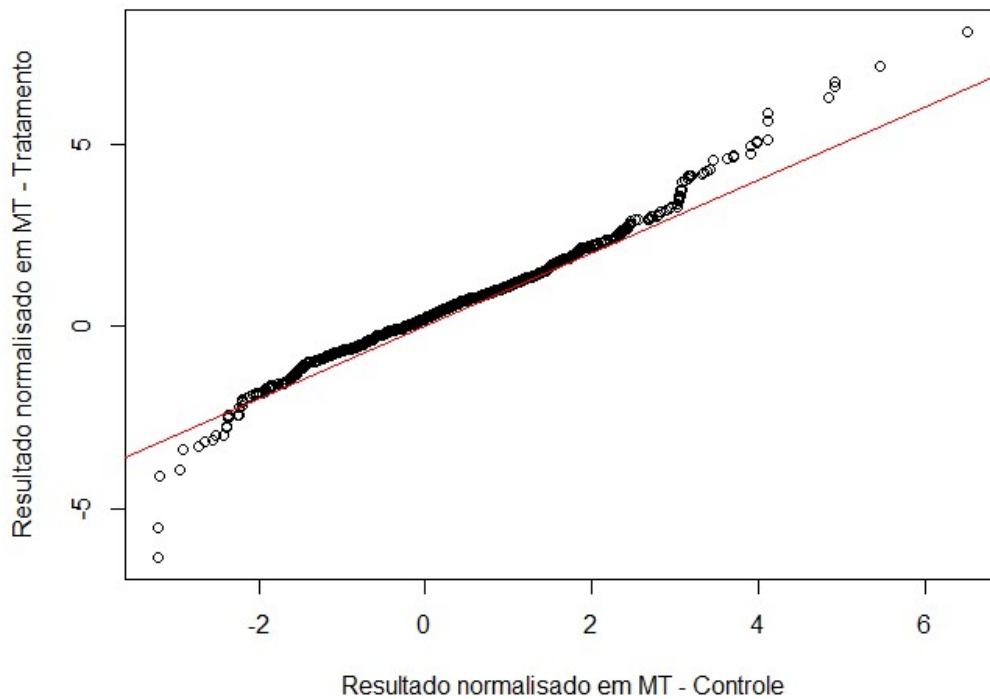
Estas constatações se mantêm também para os anos finais do IDEB. As visualizações Q-Q apontam novamente boa adesão à diagonal, apesar de apresentarem uma inversão em seus descolamentos; onde, na análise para os anos iniciais do IDEB, havia uma degradação da qualidade do pareamento nos pares com desvios negativos, para os anos finais do índice esse afastamento se configura nos pares com desvios positivos de cerca de três ou mais desvios-padrão da média. Esse fenômeno é observado tanto para as diferenças em Língua Portuguesa quanto em Matemática.

Gráfico 9 – Gráfico Q-Q do pareamento por distância de Mahalanobis para a diferença em Língua Portuguesa (em desvios-padrão), para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 10 – Gráfico Q-Q do pareamento por distância de Mahalanobis para a diferença em Matemática (em desvios-padrão), para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Ainda assim, a proximidade nas médias das variáveis de balanceamento após o pareamento por *Mahalanobis matching* é novamente visível (Tabela 6), apesar de apresentar maior descolamento que o observado para os anos iniciais. A diferença entre as médias do grupo de tratamento e de controle se mantêm significativamente nulas para as notas do IDEB em 2015 e a participação no PME em 2016, enquanto o INSE das escolas, apesar de mais próximo do que no pareamento por PSM, se mostra significativamente não-nula a um nível de confiança de 95%.

Tabela 5 – Análise do pareamento por distância de Mahalanobis, para os anos finais do IDEB

Estatística	Nota do IDEB em 2015		Participação no PME		INSE da escola em 2015	
	Sem parear	Pareado	Sem parear	Pareado	Sem parear	Pareado
Média do tratamento	4,2198	4,2198	0,1908	0,1908	43,325	43,325
Média do controle	4,3650	4,2205	0,1174	0,1908	49,284	43,339
Diferença média padronizada	-0,2055	-0,0009	0,1867	0,0000	-1,604	-0,004
Razão das variâncias	0,64	0,99	1,49	1,00	0,49	1,01
p-valor Teste T das médias	< 0,00	0,5374	< 0,00	1,0000	< 0,00	0,0392
p-valor KS Bootstrap das médias	< 0,00	1,0000			< 0,00	0,9980

Fonte: Elaboração própria com base em dados do INEP e do MEC.

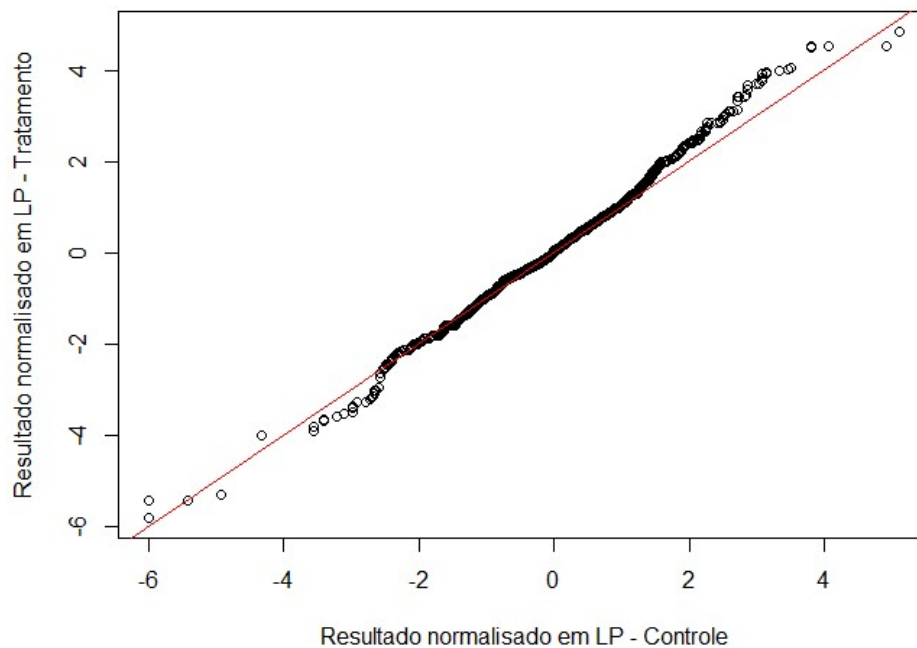
### 4.3. Avaliação do pareamento por genetic matching

O balanceamento das covariadas quando as unidades de tratamento e de controle foram pareadas por *genetic matching* foi bastante próximo para os anos iniciais do IDEB. Os Gráficos 11 e 12 trazem as visualizações Q-Q para Língua Portuguesa e Matemática, respectivamente.

Os dois grupos apresentam uma grande proximidade até mesmo perto das pontas, onde os desvios são grandes em relação à média de cada matéria. Há novamente o viés de descolamento para cima da diagonal nos desvios intensamente positivos e para baixo nos desvios fortemente negativos, em especial nas notas de Matemática, semelhantemente ao observado no pareamento por distância de Mahalanobis.

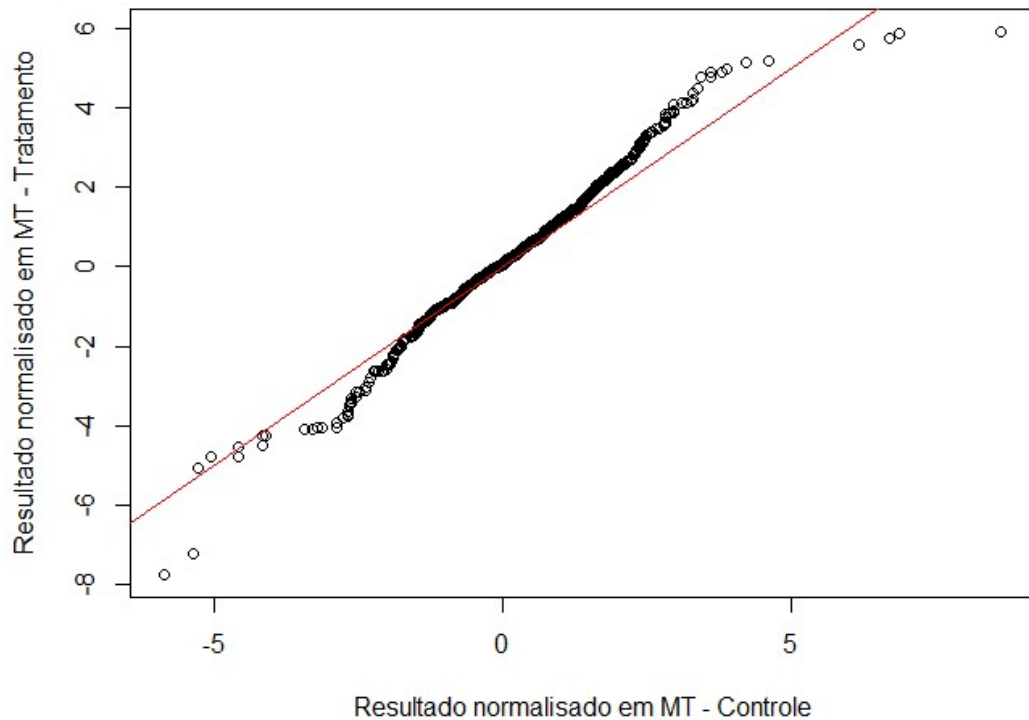
A Tabela 7 traz os resultados do pareamento por *genetic matching* nas covariadas de balanceamento para os anos iniciais do IDEB. A diferença entre as médias de controle é significativamente nula para todas as covariadas, com as razões das variâncias se aproximando da unidade. Os resultados, de forma geral, são muito semelhantes aos obtidos no pareamento por *Mahalanobis matching*.

Gráfico 11 – Gráfico Q-Q do pareamento por *genetic matching* para a diferença em Língua Portuguesa (em desvios-padrão), para os anos iniciais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 12 – Gráfico Q-Q do pareamento por *genetic matching* para a diferença em Matemática (em desvios-padrão), para os anos iniciais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

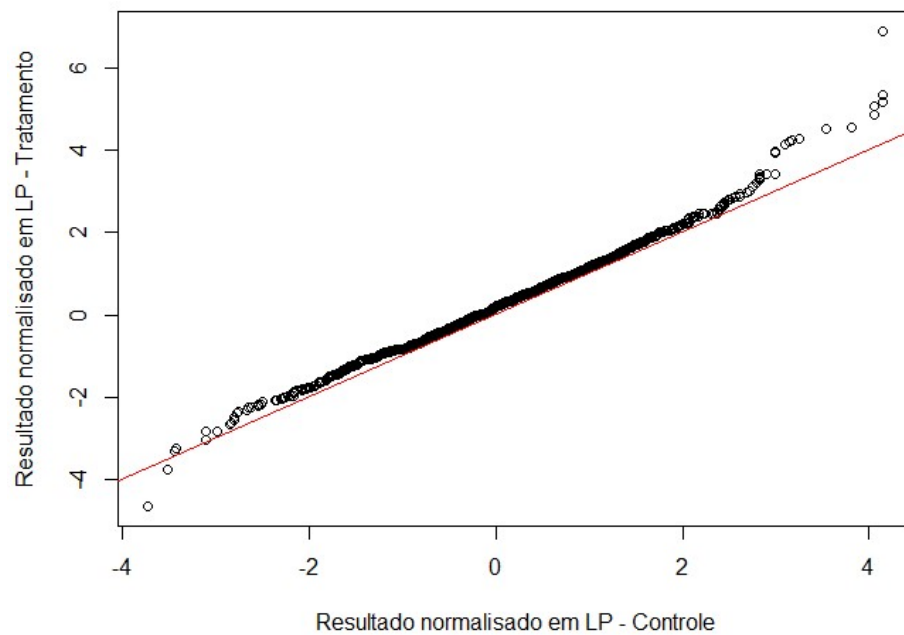
Tabela 6 – Análise do pareamento por *genetic matching*, para os anos iniciais do IDEB

Estatística	Nota do IDEB em 2015		Participação no PME		INSE da escola em 2015	
	Sem parear	Pareado	Sem parear	Pareado	Sem parear	Pareado
Média do tratamento	5,2552	5,2552	0,1734	0,1734	43,681	43,681
Média do controle	5,4194	5,2545	0,1182	0,1734	49,399	43,690
Diferença média padronizada	-0,1933	0,0007	0,1456	0,0000	-1,622	-0,003
Razão das variâncias	0,61	1,00	1,38	1,00	0,45	1,01
p-valor Teste T das médias	< 0,00	0,3938	< 0,00	1,0000	< 0,00	0,3620
p-valor KS Bootstrap das médias	< 0,00	1,0000			< 0,00	1,0000

Fonte: Elaboração própria com base em dados do INEP e do MEC.

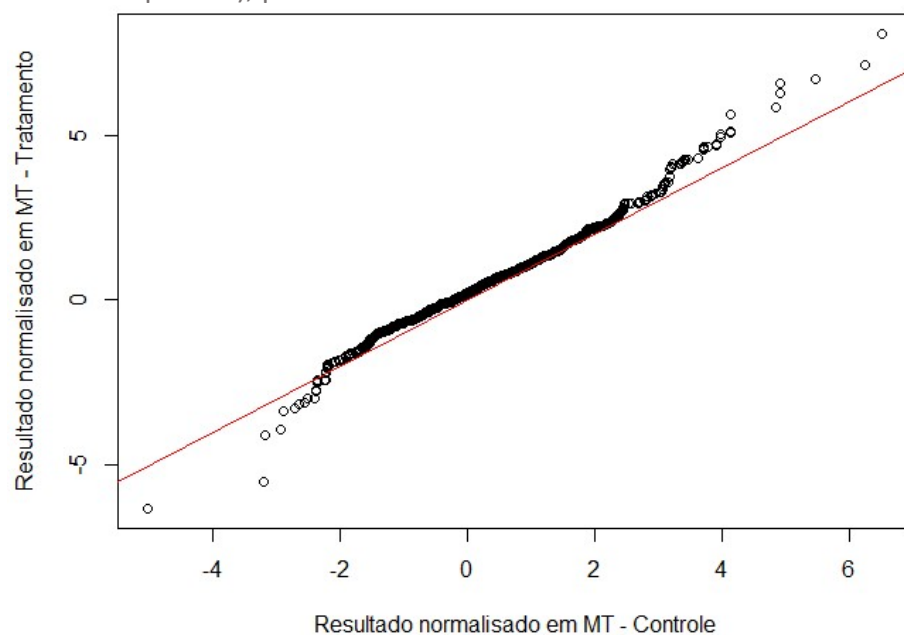
Para os anos finais do IDEB, os resultados apresentam novamente um balanceamento um pouco inferior ao obtido nos anos iniciais. Ainda assim, o pareamento por *genetic matching* aponta um elevado grau de proximidade entre as médias das covariadas entre o grupo de controle e o de tratamento.

Gráfico 13 – Gráfico Q-Q do pareamento por *genetic matching* para a diferença em Língua Portuguesa (em desvios-padrão), para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Gráfico 14 – Gráfico Q-Q do pareamento por *genetic matching* para a diferença em Matemática (em desvios-padrão), para os anos finais do IDEB



Fonte: Elaboração própria com base em dados do INEP e do MEC.

Por fim, a Tabela 8 traz o balanceamento do pareamento realizado por *genetic matching*. Simetricamente ao resultado para os anos iniciais do IDEB, encontram-se diferenças significativamente nulas nas médias das notas do IDEB e na participação do PME. O INSE das escolhas mantém-se como exceção, sendo menos próximo que as demais covariadas e apontando médias significativamente diferentes a um nível de confiança de 10%.

Tabela 7 – Análise do pareamento por *genetic matching*, para os anos finais do IDEB

Estatística	Nota do IDEB em 2015		Participação no PME		INSE da escola em 2015	
	Sem parear	Pareado	Sem parear	Pareado	Sem parear	Pareado
Média do tratamento	4,2198	4,2198	0,1908	0,1908	43,325	43,325
Média do controle	4,3650	4,2186	0,1174	0,1908	49,284	43,335
Diferença média padronizada	-0,2055	0,0017	0,1867	0,0000	-1,604	-0,003
Razão das variâncias	0,64	1,01	1,49	1,00	0,49	1,01
p-valor Teste T das médias	< 0,00	0,1832	< 0,00	1,0000	< 0,00	0,0981
p-valor KS Bootstrap das médias	< 0,00	1,0000			< 0,00	0,9990

Fonte: Elaboração própria com base em dados do INEP e do MEC.

#### 4.4. Comparação dos pareamentos

Uma vez verificado o balanceamento resultante de cada método de pareamento, torna-se possível compará-los e verificar qual apresentou maior proximidade entre as escolas contempladas pelo Programa Novo Mais Educação e as demais instituições de ensino da amostra.

A Tabela 9 resume essa comparação para cada covariada de balanceamento para os anos iniciais do IDEB. É perceptível a aproximação entre o grupo de tratamento e o de controle nas comparações sem pareamento, com pareamento por *propensity scores* e com pareamento por distância de Mahalanobis. Os resultados por pareamento genético (SEKHON, 2011) se aproximaram bastante dos obtidos por distância de Mahalanobis.

No balanceamento dos pareamentos para os anos finais do IDEB (Tabela 10), o pareamento por distância de Mahalanobis apresentou a maior proximidade entre as unidades do grupo de tratamento e de controle para a nota do IDEB em 2015, enquanto o pareamento por *genetic matching* balanceou melhor o INSE das escolas. Em ambos os casos, ainda foi observada uma diferença significativa nas médias do INSE das escolas a um nível de confiança de 90%.

Tabela 8 – Comparação dos modelos de pareamento, para os anos iniciais do IDEB

Covariada	Indicador	Pareamento			
		Sem pareamento	PSM	Mahalanobis	Genetic
Nota do IDEB em 2015	Média do tratamento	5,2552	5,2552	5,2552	5,2552
	Média do controle	5,4194	5,0911	5,2544	5,2545
	Diferença média padronizada	-0,1933	0,1931	0,0009	0,0007
	Razão das variâncias	0,61	0,47	1,00	1,00
	p-valor Teste T das médias	< 0,00	< 0,00	0,6329	0,3938
	p-valor KS Bootstrap das médias	< 0,00	< 0,00	1,0000	1,0000
Participação no PME em 2016	Média do tratamento	0,1734	0,1734	0,1734	0,1734
	Média do controle	0,1182	0,1799	0,1734	0,1734
	Diferença média padronizada	0,1456	-0,0172	0,0000	0,0000
	Razão das variâncias	1,38	0,97	1,00	1,00
	p-valor Teste T	< 0,00	0,6691	1,0000	1,0000
	p-valor KS Bootstrap				
INSE da escola em 2015	Média do tratamento	43,681	43,681	43,681	43,681
	Média do controle	49,399	43,306	43,688	43,690
	Diferença média padronizada	-1,622	0,106	-0,002	-0,003
	Razão das variâncias	0,45	0,51	1,01	1,01
	p-valor Teste T	< 0,00	< 0,00	0,3982	0,3620
	p-valor KS Bootstrap	< 0,00	< 0,00	1,0000	1,0000

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Tabela 9 – Comparação dos modelos de pareamento, para os anos finais do IDEB

Covariada	Indicador	Pareamento			
		Sem pareamento	PSM	Mahalanobis	Genetic
Nota do IDEB em 2015	Média do tratamento	4,2198	4,2198	4,2198	4,2198
	Média do controle	4,3650	4,0872	4,2205	42,2186
	Diferença média padronizada	-0,2055	0,1878	-0,0009	0,0017
	Razão das variâncias	0,64	0,63	0,99	1,01
	p-valor Teste T das médias	< 0,00	< 0,00	0,5374	0,1832
	p-valor KS Bootstrap das médias	< 0,00	< 0,00	1,0000	1,0000
Participação no PME em 2016	Média do tratamento	0,1908	0,1908	0,1908	0,1908
	Média do controle	0,1174	0,1958	0,1908	0,1908
	Diferença média padronizada	0,1867	-0,0129	0,0000	0,0000
	Razão das variâncias	1,49	0,98	1,00	1,00
	p-valor Teste T	< 0,00	0,7097	1,0000	1,0000
	p-valor KS Bootstrap				
INSE da escola em 2015	Média do tratamento	43,325	43,325	43,325	43,325
	Média do controle	49,284	43,029	43,339	43,335
	Diferença média padronizada	-1,604	0,08	-0,004	-0,003
	Razão das variâncias	0,49	0,68	1,04	1,01
	p-valor Teste T	< 0,00	< 0,00	0,0392	0,0981
	p-valor KS Bootstrap	< 0,00	< 0,00	0,9980	0,9990

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Assim, este trabalho contribui com a literatura acerca da avaliação de políticas educacionais também com o presente contraste entre o pareamento por *propensity score matching*, distância de Mahalanobis ou *genetic matching*, trazendo indícios de que estes dois últimos trazem benefícios para o balanceamento dos grupos de tratamento e controle na análise do desempenho acadêmico das escolas públicas brasileiras.

## 4.5. Impacto do PNME sobre o desempenho acadêmico das escolas participantes

Conforme destacado no segundo capítulo deste trabalho, o impacto da migração para regimes de educação em tempo integral no desempenho acadêmico das escolas é heterogêneo na literatura. Essa realidade é evidenciada dentro do próprio cenário nacional, com tanto Xerxenevsky (2012) quanto Oliveira e Terra (2018) encontrando efeitos nulos da educação em tempo integral nas notas de matemática e português de alunos dos anos finais

do Ensino Fundamental, quando implementada através do antigo Programa Mais Educação. Para os anos iniciais, o primeiro estudo consegue apontar efeitos significantes, porém com sinal positivo para notas de Língua Portuguesa e negativo em Matemática, enquanto o segundo permanece apontando impacto nulo do Programa.

Existem, no entanto, algumas peculiaridades do Programa Novo Mais Educação que devem ser consideradas. O PNME objetiva melhorar não o aprendizado dos alunos como um todo, como seu antecessor, mas suas notas de Língua Portuguesa e Matemática, critérios avaliados pelo SAEB. É de se esperar então que os impactos do Programa se concentrem nessas duas disciplinas e sejam, portanto, mais observáveis. Ademais, o PNME possibilita às escolas participantes optarem por regimes de cinco ou quinze horas semanais adicionais. No presente trabalho, porém, as escolas optantes pelo primeiro foram eliminadas da amostra por questões amostrais, de forma que o tratamento aqui analisado se refere exclusivamente às escolas que adicionaram quinze horas à sua grade horária semanal.

Assim, a Tabela 11 resume os impactos do PNME sobre as notas de Língua Portuguesa e Matemática para os anos iniciais do IDEB nas escolas contempladas, diferenciado pelo método de pareamento utilizado. Para fins de robustez, as estimações foram realizadas através de dois pacotes, *Matching* (SEKHON, 2011) e *MatchIt* (HO, IMAI, KING, STUART, 2011), e os resultados obtidos com cada se encontram ilustrados nas Tabelas 11 (para os anos iniciais do Ensino Fundamental) e 12 (para os anos finais).

Utilizando um *propensity score matching* com escores de probabilidade de tratamento baseados nos critérios de seleção do Programa e pareamento realizado por *nearest neighbour matching* unitário com reposição, houve uma melhora significativa<sup>14</sup> das notas em ambas as matérias nas escolas participantes. O impacto foi de ganhos de 0,10 a 0,11 desvios-padrão nas notas de Língua Portuguesa, dependendo do pacote utilizado, e de 0,14 a 0,17 desvios-padrão nas notas de Matemática nessas escolas.

---

<sup>14</sup> O ATT foi diferente de zero a um nível de confiança de 95%, para as notas em Língua Portuguesa, e de 99%, para as notas de Matemática, conforme descrito na Tabela 11.

Tabela 10 – Estimações do impacto do PNME nas notas normalizadas de Língua Portuguesa e Matemática, por tipo de pareamento, para os anos iniciais do IDEB

Estatística	PSM				Mahalanobis				Genetic			
	Língua Portuguesa		Matemática		Língua Portuguesa		Matemática		Língua Portuguesa		Matemática	
	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt
ATT estimado	0,106	0,095	0,169	0,139	0,068	0,075	0,123	0,128	0,085	0,096	0,161	0,153
Erro-padrão	0,045	0,051	0,049	0,055	0,053	0,053	0,058	0,057	0,055	0,050	0,059	0,053
Estatística T	2,35	1,84	3,44	2,52	1,30	1,43	2,11	2,24	1,57	1,93	2,72	2,86
p-valor	0,0187	0,0660	0,0006	0,0120	0,1946	0,1537	0,0347	0,0249	0,1175	0,0539	0,0065	0,0043
Significância <sup>1</sup>	**	*	***	**	-	-	**	**	-	*	***	***

<sup>1</sup> Grau de significância: - Acima de 10%, \* Entre 10% e 5%, \*\* Entre 5% e 1%, \*\*\* Abaixo de 1%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Esses ganhos em Língua Portuguesa, porém, não se repetiram quando o pareamento foi realizado através de *Mahalanobis matching*. Nessa análise, o ATT foi significativo apenas para os resultados em Matemática, sendo de 0,13 a 0,14 desvios-padrão para os anos iniciais do IDEB.

Por fim, o pareamento por *genetic matching* apontou novamente um impacto significativo do Programa Novo Mais Educação apenas nas notas de Matemática das escolas participantes, sendo estas superiores às das demais instituições de ensino em 0,15 a 0,16 desvios-padrão.

Por outro lado, quando analisados os anos finais do IDEB (Tabela 12) verificaram-se ganhos altamente significativos em ambas as matérias independentemente da forma de pareamento ou pacote utilizado. No caso das notas de Língua Portuguesa, observou-se um ganho nas escolas contempladas pelo PNME que variou de 0,13 a 0,19 desvios-padrão, sendo mais elevado quando o pareamento foi realizado por distância de Mahalanobis. Já as notas de Matemática apresentaram ganhos ainda mais intensos, variando de 0,17 a 0,22 desvios-padrão (esse último novamente quando estimado por distância de Mahalanobis).

Tabela 11 – Estimações do impacto do PNME nas notas normalizadas de Língua Portuguesa e Matemática, por tipo de pareamento, para os anos finais do IDEB

Estatística	PSM				Mahalanobis				Genetic			
	Língua Portuguesa		Matemática		Língua Portuguesa		Matemática		Língua Portuguesa		Matemática	
	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt
ATT estimado	0,132	0,149	0,182	0,167	0,198	0,198	0,222	0,220	0,193	0,182	0,203	0,194
Erro-padrão	0,034	0,039	0,037	0,042	0,042	0,039	0,044	0,013	0,042	0,038	0,044	0,041
Estatística T	3,86	3,83	4,88	4,00	4,74	5,01	5,03	5,14	4,57	4,80	4,58	4,69
p-valor	0,0001	0,0001	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Significância <sup>1</sup>	***	***	***	***	***	***	***	***	***	***	***	***

<sup>1</sup> Grau de significância: - Acima de 10%, \* Entre 10% e 5%, \*\* Entre 5% e 1%, \*\*\* Abaixo de 1%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Finalmente, destaca-se também que foram realizadas estimações dos efeitos do Programa utilizando variáveis adicionais nos pareamentos por distância de Mahalanobis e por *genetic matching*, uma vez que a natureza não-paramétrica dessas distâncias facilita o uso de um conjunto maior de covariadas sem incorrer em problemas de correlação ou qualidade da regressão realizada (DIAMOND, SEKHON, 2013). Foram inseridas nos modelos as notas em Língua Portuguesa e Matemática das escolas em 2015, e os resultados das estimações se encontram disponíveis no Apêndice (Tabelas A e B). As conclusões se mantiveram simétricas para todos os modelos, pacotes, disciplinas e anos analisados, com única exceção das notas de Matemática nos anos iniciais no pareamento por distância de Mahalanobis, que não foram significativos nessa análise.

## 5. CONCLUSÃO

Este estudo estimou o impacto do ensino em tempo integral quando fomentado pelo Programa Novo Mais Educação sobre o desempenho escolar em Língua Portuguesa e Matemática nas escolas públicas brasileiras. O PNME, que vigorou entre 2017 e 2019, aportou recursos às escolas participantes para que adicionassem cinco ou quinze horas semanais focadas nestas duas matérias em suas grades horárias.

Para essa estimação, foram analisadas apenas as escolas que optaram pela modalidade de quinze horas adicionais. As variáveis de interesse foram as diferenças normalizadas nas notas de Língua Portuguesa e de Matemática das escolas entre 2017 e 2019. A estimação do impacto do Programa foi realizada através *propensity score matching*, *Mahalanobis matching* e *genetic matching*. Os pareamentos por distância de Mahalanobis e por *genetic matching* apresentaram melhor balanceamento nas covariadas estudadas.

Os resultados encontrados foram inconclusivos para os anos iniciais do Ensino Fundamental. Apontam-se ganhos significativos em ambas as matérias quando as escolas foram pareadas por *propensity score matching*, sendo de 0,11 desvios-padrão em Língua Portuguesa e 0,17 em Matemática. Quando analisadas por *Mahalanobis matching* ou por *genetic matching*, porém, houve ganhos significativos apenas nas notas de Matemática (0,12 e 0,15 desvios-padrão, respectivamente). No caso do pareamento por distância de Mahalanobis, porém, esses ganhos não se repetiram após a inserção de covariadas adicionais na matriz de Mahalanobis.

Para os anos finais do Ensino Fundamental, porém, todos os três métodos apontaram ganhos significativos nas notas de ambas as disciplinas. Esse resultado independe do pacote utilizado para a estimação ou da adição de covariadas na matriz de Mahalanobis. O impacto do Programa Novo Mais Educação variou de 0,13 a 0,17 desvios-padrão nas notas em Língua Portuguesa, a depender do pareamento realizado, e de 0,17 a 0,22 desvios-padrão nas notas em Matemática.

Este resultado difere bastante do encontrado na literatura nacional acerca dos retornos do ensino em tempo integral. Os trabalhos de Xerxenevsky (2012) e Oliveira e Terra (2018) ambos estimam impactos não-significantes da educação em tempo integral para os anos finais do Ensino Fundamental, quando implementada pelo Programa Mais Educação, precursor do PNME. Já a presente análise do Programa Novo Mais Educação aponta valores positivos e significantes para as duas disciplinas. Tal resultado pode estar associado ao maior foco dado pelo PNME especificamente às matérias Língua Portuguesa e Matemática, com a

obrigatoriedade da instituição de ensino oferecer horas adicionais em ambas (ao invés de apenas uma, conforme regia o antigo PME) e aos normativos do programa que objetivaram diretamente melhoras no resultado acadêmico dos alunos (no lugar do foco mais amplo do PME em uma melhor educação).

Para os anos iniciais do Ensino Fundamental, porém, os resultados deste trabalho se aproximam mais do que é encontrado na literatura. O impacto do Programa Novo Mais Educação para esta faixa de idade apresentou coeficientes não-significativos ou cuja significância variou em função do pareamento ou escolha de covariadas. Em seus trabalhos de avaliação do antigo PME, Oliveira e Terra (2018) encontraram resultados similarmente não-significativos, enquanto Xerxenevsky (2012) aponta um aumento de 2,34 pontos nas notas de Língua Portuguesa, porém uma queda de 3,04 pontos nas notas de Matemática dos alunos participantes<sup>15</sup>.

Uma possível explicação para o presente trabalho encontrar impactos mais intensos nos anos finais do Ensino Fundamental é que a base de dados obtida não é capaz de desagregar quantos alunos de cada faixa etária participaram do Programa Novo Mais Educação nas escolas contempladas. Assim, é possível que as instituições de ensino tenham ofertado mais turmas de ensino em tempo integral para as turmas dos anos finais do Ensino Fundamental.

É importante, também, que outras ressalvas sejam feitas ao escopo desse trabalho. O presente estudo se limitou em avaliar o impacto do PNME nas notas em Língua Portuguesa e Matemática dos alunos das escolas participantes. Um programa de incentivo ao ensino em tempo integral, porém, pode impactar alunos, suas famílias e sua comunidade em outras dimensões que transcendem suas notas. Além disso, não foram considerados aqui os custos associados ao Programa nem discutidas outras formas de fomentar o desempenho acadêmico escolar que sejam possivelmente menos onerosas aos cofres públicos. Estudos de

---

<sup>15</sup> Os valores apresentados por Xerxenevsky (2012) não estão padronizados. Para propósitos de comparação, o resultado de +0,13 desvios-padrão em Língua Portuguesa encontrado nesse trabalho representa +1,85 pontos na nota dessa disciplina, enquanto +0,17 desvios-padrão em Matemática equivale a +2,47 pontos na matéria.

custo-efetividade são fundamentais para uma análise mais completa dos benefícios de uma política pública desta dimensão.

Ainda assim, este estudo demonstrou, apesar dessas limitações, um potencial impacto positivo do incentivo ao ensino em tempo integral na performance acadêmica das escolas públicas nacionais.

Finalmente, o estudo destaca-se por ser o primeiro, até a redação desta dissertação, que procura realizar uma avaliação quantitativa do impacto do PNME, abrindo a porta para outros que subsidiem importantes debates acerca dos retornos das políticas educacionais, de forma ampla, e da educação em tempo integral, de forma específica, no cenário nacional, para o desenvolvimento mais sustentável do país e para a necessária redução das desigualdades sociais e econômicas que ainda persistem no Brasil.

## 6. REFERÊNCIAS

ALFARO, Pablo; EVANS, David K.; HOLLAND, Peter. *Extending the school day in Latin America and the Caribbean. World Bank Policy Research Working Paper no. 7309*, jun. 2015. Disponível em: <<http://documents.worldbank.org/curated/en/477421467986293530/pdf/WPS7309.pdf>>. Acesso em: 15 abr. 2021.

ALVES, Thiago; FARENZENA, Nalú; SILVEIRA, Adriana; PINTO, José Marcelino. Implicações da pandemia da COVID-19 para o financiamento da educação básica. **Rev. Adm. Pública**, Rio de Janeiro, v. 54, n. 4, ago. 2020, p. 979-993. DOI: 10.1590/0034-761220200279.

ANGRIST, Joshua; PISCHKE, Jörn-Steffen. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press, 2009. DOI: 10.2307/j.ctvcm4j72.

AUSTIN, Peter C. *Statistical criteria for selecting the optimal number of untreated subjects matched to each treated subject when using many-to-one matching on the propensity score. American journal of epidemiology*, vol. 172, n. 9, ago. 2010, p. 1092-1097. DOI: 10.1093/aje/kwq224

BELLEI, Cristián. *Does lengthening the school day increase students' academic achievement? Results from a natural experiment in Chile*. **Economics of Education Review**, v. 28, n. 5, p. 629-640. ISSN 0272-7757. DOI: 10.1016/j.econedurev.2009.01.008.

BIONDI, Roberta Loboda; VASCONCELLOS, Lígia; MENEZES-FILHO, Naércio Aquino. Avaliando o impacto da Olimpíada Brasileira de Matemática das Escolas Públicas (OBMEP) no desempenho de matemática nas avaliações educacionais. **Anais da Sociedade Brasileira de Econometria**, Rio de Janeiro, 2009. Disponível em: <[https://www.researchgate.net/publication/229048378\\_Avaliando\\_o\\_impacto\\_da\\_Olimpiada\\_Brasileira\\_de\\_Matematica\\_das\\_Escolas\\_Publicas\\_OBMEP\\_no\\_desempenho\\_de\\_matematica\\_nas\\_avaliacoes\\_educacionais](https://www.researchgate.net/publication/229048378_Avaliando_o_impacto_da_Olimpiada_Brasileira_de_Matematica_das_Escolas_Publicas_OBMEP_no_desempenho_de_matematica_nas_avaliacoes_educacionais)>. Acesso em: 16 abr. 2021.

BLUNDELL, Richard; DEARDEN, Lorraine; MEGHIR, Costas; SIANESI, Barbara. (1999), *Human Capital Investment: The Returns from Education and Training to the Individual, the Firm and the Economy*. **Fiscal Studies**, v. 20, 1999, p. 1-23. DOI: 10.1111/j.1475-5890.1999.tb00001.x

BRASIL. Ministério da Educação e Cultura. **Programa Mais Educação**. Disponível em: <<http://portal.mec.gov.br/programa-mais-educacao/apresentacao?id=16689>>. Acesso em: 15 abr. 2021.

BRASIL. Ministério da Educação e Cultura. **Programa Novo Mais Educação**. Disponível em: <<http://portal.mec.gov.br/programa-mais-educacao>>. Acesso em: 15 abr. 2021.

CALIENDO, Marco; KOPEINING, Sabine. *Some Practical Guidance for the Implementation of Propensity Score Matching*. IZA Discussion Paper no. 1588, May 2005. **Journal of Economic Surveys**, v. 22, n. 1, 2008, p. 31-72. Disponível em: <<https://www.iza.org/publications/dp/1588/some-practical-guidance-for-the-implementation-of-propensity-score-matching>>. Acesso em: 16 abr. 2021.

CERDAN-INFANTES, Pedro; VERMEERSCH, Christel. *More time is better: an evaluation of the full time school program in Uruguay*. **World Bank Policy Research Working Paper no. 4167**, mar. 2007. Disponível em: <<https://ssrn.com/abstract=970809>>. Acesso em: 15 abr. 2021.

COSTA, Nilce; GOMES, Cândido A. C. Um olho no custo, outro no benefício: a escola de tempo integral no campo é viável? **Meta: Avaliação**, Rio de Janeiro, v. 4, n. 11, mai./ago. 2012, p. 188-209. DOI: 10.22347/2175-2753v4i11.147.

DEE, Thomas S. *Are there civic returns to education?* **Journal of Public Economics**, v. 88, n. 9-10, 2004, p. 1697-1720. ISSN 0047-2727. DOI: 10.1016/j.jpubeco.2003.11.002.

DIAMOND, Alexis; SEKHON, Jasjeet. *Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies*. **The Review of Economics and Statistics** 2013, v. 95, n.3, jul. 2013, p. 932–945. DOI: 10.1162/REST\_a\_00318.

FERNANDES, Reynaldo. **Índice de Desenvolvimento da Educação Básica (Ideb)**. Brasília: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2007. ISSN 1414-0640. Disponível em: <[http://portal.inep.gov.br/documents/186968/485287/%C3%8Dndice+de+Desenvolvimento+da+Educa%C3%A7%C3%A3o+B%C3%A1sica+\(Ideb\)/26bf6631-44bf-46b0-9518-4dc3c310888b?version=1.4](http://portal.inep.gov.br/documents/186968/485287/%C3%8Dndice+de+Desenvolvimento+da+Educa%C3%A7%C3%A3o+B%C3%A1sica+(Ideb)/26bf6631-44bf-46b0-9518-4dc3c310888b?version=1.4)>. Acesso em: 16 abr. 2021.

HAGLE, Timothy M.; MITCHELL, Glenn E. *Goodness-of-Fit Measures for Probit and Logit*. **American Journal of Political Science**, vol. 36, n. 3, 1992, p. 762-784. DOI: 10.2307/2111590.

HARMON, Colm; OOSTERBEEK, Hessel; WALKER, Ian. *The Returns to Education: Microeconomics*. **Journal of Economic Surveys**, v. 17, n. 2, 2003, p. 115-156. DOI: 10.1111/1467-6419.00191.

HO, Daniel; IMAI, Kosuke; KING, Gary; STUART, Elizabeth A. *MatchIt: Nonparametric Preprocessing for Parametric Causal Inference*. **Journal of Statistical Software**, v. 42, n. 8, p. 1–28, 2011. DOI: 10.18637/jss.v042.i08.

KALLUS, Nathan. *Causal Inference by Minimizing the Dual Norm of Bias: Kernel Matching & Weighting Estimators for Causal Effects*. **Causation: Foundation to Application Workshop**, Jersey City, jun. 2016. Disponível em: <<http://www.its.caltech.edu/~fehardt/UAI2016WS/papers/Kallus.pdf>>. Acesso em: 27 jun. 2021.

KARPOWITZ, Daniel; KENNER, Max. *Education as Crime Prevention: The Case for Reinstating Pell Grant Eligibility for the Incarcerated*. **Bard Prison Initiative**, Annandale-on-Hudson, 2016. Disponível em: <<https://bpi.bard.edu/wp-content/uploads/2019/09/KarpowitzKenner2003.pdf>>. Acesso em: 16 abr. 2021.

KELLAGHAN, Thomas; MADAUS, George F.; AIRASIAN, Peter W. **The Effects of Standardized Testing**. Dordrecht: Springer Netherlands, 1982. DOI: 10.1007/978-94-009-7386-2.

KHANDKER, Shahidur; B. KOOLWAL, Gayatri; SAMAD, Hussain. **Handbook on Impact Evaluation: Quantitative Methods and Practices**. The World Bank, 2009. Disponível em: <<http://elibrary.worldbank.org/doi/book/10.1596/978-0-8213-8028-4>>. Acesso em: 16 abr. 2021.

KING, Gary; NIELSEN, Richard. *Why Propensity Scores Should Not Be Used for Matching*. **Political Analysis**, vol. 27, n. 4, out. 2019, p. 435-454. DOI: 10.1017/pan.2019.11.

LEE, Wang-Sheng. *Propensity score matching and variations on the balancing test*. **Empirical Economics**, v. 44, n. 1, p. 47-80, 2013. DOI: 10.1007/s00181-011-0481-0.

MATUOKA, Ingrid. Os entraves para a continuidade do Novo Mais Educação. Reportagem. **Centro de Referências em Educação Integral**, set. 2019. Disponível em: <<https://educacaointegral.org.br/reportagens/os-entraves-para-continuidade-do-novo-mais-educacao/>>. Acesso em: 16 abr. 2021.

MATUOKA, Ingrid. Por que o Novo Mais Educação não dialoga com a educação integral. Reportagem. **Centro de Referências em Educação Integral**, ago. 2017. Disponível em: <<https://educacaointegral.org.br/reportagens/novo-mais-educacao-nao-dialoga-educacao-integral/>>. Acesso em: 16 abr. 2021.

MCEACHIN, Andrew; LAUEN, Douglas; FULLER, Sarah; PERERA, Rachel. *Social returns to private choice? Effects of charter schools on behavioral outcomes, arrests, and civic participation*. **Economics of Education Review**, v. 76, 2020. ISSN 0272-7757. DOI: 10.1016/j.econedurev.2020.101983.

MENEZES-FILHO, Naercio. Os determinantes do desempenho escolar do Brasil. **4º Seminário de Economia de Belo Horizonte**, Centro de Pesquisa em Economia Internacional, set. 2007. Disponível em: <[http://www.cepe.ecn.br/seminarioiv/download/menezes\\_filho.pdf](http://www.cepe.ecn.br/seminarioiv/download/menezes_filho.pdf)>. Acesso em: 15 abr. 2021.

OLIVEIRA, Luiz Felipe Batista; TERRA, Rafael. Impacto do Programa Mais Educação em indicadores educacionais. **FEA-USP**, jun. 2018. Disponível em: <[https://www.fea.usp.br/sites/default/files/anexo-evento/impacto\\_do\\_pme\\_em\\_indicadores\\_educacionais\\_-\\_feausp2018.pdf](https://www.fea.usp.br/sites/default/files/anexo-evento/impacto_do_pme_em_indicadores_educacionais_-_feausp2018.pdf)>. Acesso em: 15 abr. 2021.

ONYEMA, Edeh Michael; EUCHERIA, Nwafor Chika; OBAFEMI, Faith Ayobamidele; SEN, Shuvro; ATONYE, Fyeface Grace; SHARMA, Aabha; ALSAYED, Alhuseen Omar. *Impact of Coronavirus Pandemic on Education*. **Journal of Education and Practice**, 2020. Disponível em: <<https://iiste.org/Journals/index.php/JEP/article/view/52821>>. Acesso em: 16 abr. 2021.

OSTLER, Denise de Almeida. **Impacto da implantação do Programa Ensino Integral no atendimento a alunos com deficiência: desdobramentos em duas escolas públicas da Baixada Santista**. Tese de Doutorado em Educação. Universidade Metodista de São Paulo, São Bernardo do Campo, 266 fls., 2017. Disponível em: <<http://tede.metodista.br/jspui/handle/tede/1676>>. Acesso em: 15 abr. 2021.

RIANI, Juliana de Lucena Ruas. **Determinantes do resultado educacional no Brasil: famílias, perfil escolar dos municípios e dividendo demográfico numa abordagem hierárquica e espacial.** Tese de Doutorado em Demografia. Universidade Federal de Minas Gerais, Belo Horizonte, 218 fls., 2005. Disponível em: <<https://repositorio.ufmg.br/handle/1843/MCCR-6VTMGE>>. Acesso em: 15 abr. 2021.

ROSENBAUM, Paul R.; RUBIN, Donald B. *The central role of the propensity score in observational studies for causal effects.* **Biometrika**, v. 70, n. 1, abr. 1983, p. 41-55. DOI: 10.1093/biomet/70.1.41.

SCHIEFELBEIN, Ernesto & SIMMONS, John. Os determinantes do desempenho escolar: uma revisão de pesquisas nos países em desenvolvimento. **Cad. Pesq.**, São Paulo, v. 35, nov. 1980, p. 53-71. Disponível em: <<http://publicacoes.fcc.org.br/ojs/index.php/cp/article/view/1649/1636>>. Acesso em: 15 abr. 2021.

SEKHON, Jasjeet S. *Multivariate and Propensity Score Matching Software with Automated Balance Optimization: The Matching Package for R.* **Journal of Statistical Software**, v. 42, n. 7, 2011. DOI: 10.18637/jss.v042.i07.

SILVA, Daianne P.; COURA, Maritza A. S.; NÓBREGA, Roberto; MORAES, José F. V. N.; CAMPBELL, Carmen S. G. Impacto das atividades esportivas, informática e música sobre a vida de alunos de escola com educação em tempo integral. **Motrivivência**, v. 39, 2012. DOI: 10.5007/2175-8042.2012v24n39p142.

XERXENEVSKY, Lauren Lewis. **Programa mais educação: avaliação do impacto da educação integral no desempenho de alunos no Rio Grande do Sul.** Dissertação de Mestrado em Economia do Desenvolvimento. Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre, 143 fls., 2012. Disponível em: <<http://tede2.pucrs.br/tede2/handle/tede/3924>>. Acesso em: 15 abr. 2021.



## Apêndice

Quadro A – Parâmetros estimados por regressão *logit* da probabilidade de tratamento, para os anos iniciais do IDEB

Variável preditiva	Variável dependente
	Participação da escola no PNME
Nota final da escola no IDEB em 2015	0.641*** (0.031)
Participação da escola no PME em 2016	0.319*** (0.081)
INSE da escola em 2015	-0.270*** (0.007)
Constante	6.109*** (0.259)
Observações	28.504
AIC	8.589,0

Significância: \*\*\* 1%, \*\* 5%, \* 10%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Quadro B – Parâmetros estimados por regressão *logit* da probabilidade de tratamento, para os anos finais do IDEB

Variável preditiva	Variável dependente
	Participação da escola no PNME
Nota final da escola no IDEB em 2015	0.635*** (0.037)
Participação da escola no PME em 2016	0.425*** (0.073)
INSE da escola em 2015	-0.262*** (0.006)
Constante	7.229*** (0.253)
Observações	16.269
AIC	8.770,4

Significância: \*\*\* 1%, \*\* 5%, \* 10%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Tabela A – Estimações do impacto do PNME nas notas normalizadas de Língua Portuguesa e Matemática, por tipo de pareamento, utilizando duas covariadas adicionais para os pareamentos por distância de Mahalanobis e por *genetic matching*, para os anos iniciais do IDEB

Estatística	PSM				Mahalanobis (5 covariadas)				Genetic (6 covariadas)			
	Língua Portuguesa		Matemática		Língua Portuguesa		Matemática		Língua Portuguesa		Matemática	
	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt
ATT estimado	0,106	0,095	0,169	0,139	0,053	0,053	0,072	0,072	0,040	0,073	0,122	0,130
Erro-padrão	0,045	0,051	0,049	0,055	0,053	0,052	0,058	0,056	0,054	0,052	0,058	0,057
Estatística T	2,35	1,84	3,44	2,52	1,00	1,03	1,24	1,29	0,74	1,39	2,10	2,28
p-valor	0,0187	0,0660	0,0006	0,0120	0,3164	0,3035	0,2134	0,1981	0,4620	0,1633	0,0359	0,0225
Significância <sup>1</sup>	**	*	***	**	-	-	-	-	-	-	**	**

<sup>1</sup> Graus de significância: - Acima de 10%, \* Entre 10% e 5%, \*\* Entre 5% e 1%, \*\*\* Abaixo de 1%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

Tabela B – Estimações do impacto do PNME nas notas normalizadas de Língua Portuguesa e Matemática, por tipo de pareamento, utilizando duas covariadas adicionais para os pareamentos por distância de Mahalanobis e por *genetic matching*, para os anos finais do IDEB

Estatística	PSM				Mahalanobis (5 covariadas)				Genetic (6 covariadas)			
	Língua Portuguesa		Matemática		Língua Portuguesa		Matemática		Língua Portuguesa		Matemática	
	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt	Matching	MatchIt
ATT estimado	0,132	0,149	0,182	0,167	0,161	0,161	0,198	0,198	0,190	0,175	0,195	0,182
Erro-padrão	0,034	0,039	0,037	0,042	0,042	0,039	0,045	0,043	0,041	0,039	0,044	0,043
Estatística T	3,86	3,83	4,88	4,00	3,86	4,14	4,36	4,64	4,64	4,43	4,43	4,25
p-valor	0,0001	0,0001	0,0000	0,0001	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
Significância <sup>1</sup>	***	***	***	***	***	***	***	***	***	***	***	***

<sup>1</sup> Graus de significância: - Acima de 10%, \* Entre 10% e 5%, \*\* Entre 5% e 1%, \*\*\* Abaixo de 1%

Fonte: Elaboração própria com base em dados do INEP e do MEC.

conselho  
editorial

COORDENAÇÃO

**José Luiz Rossi Junior**

**Thiago Costa Monteiro Caldeira**

SUPERVISÃO E REVISÃO

**Renan Silveira Holtermann**

**Mathias Schneid Tessmann**

**Milton de Souza Mendonça Sobrinho**

**Alessandro de Oliveira Gouveia Freire**

COMUNICAÇÃO & A MARKETING

**Antonio Aleixo Zaninetti Silva**

**Daniel Jordão**

PROJETO GRÁFICO E DIAGRAMAÇÃO

**Juliana Vasconcelos**



The background features a dark blue color scheme with various data visualization elements. On the left, there is a bar chart with five bars of increasing height, with values 178, 180, 175, 172, and 190. A line graph with two upward-trending lines is overlaid on the bars. At the top center, a circular gauge shows 60%. To the right, another circular gauge shows 75%. A dotted line with a downward-pointing triangle is also visible. The overall aesthetic is modern and technical.

# idp

SGAS Quadra 607 - Módulo 49  
Via L2 Sul, Brasília-DF  
CEP: 70200-670

  /sejaidp  
 (61) 3535-6565  
 idp.edu.br