

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO SOCIOECÔNOMICO
DEPARTAMENTO DE ECONOMIA E RELAÇÕES INTERNACIONAIS
CURSO DE GRADUAÇÃO EM CIÊNCIAS ECONÔMICAS

Jhonny Alves Bez Birolo

**Estimando o impacto de um ano extra de Ensino Fundamental na proficiência dos
estudantes brasileiros**

Florianópolis

2023

Jhonny Alves Bez Birolo

**Estimando o impacto de um ano extra de Ensino Fundamental na proficiência dos
estudantes brasileiros**

Trabalho de Conclusão de Curso submetido ao Curso de Graduação em Ciências Econômicas do Centro Socioeconômico da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do título de Bacharel em Ciências Econômicas.

Orientador: Prof. Francis Carlo Petterini Lourenço, Dr.

Florianópolis

2023

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Bez Birolo, Jhonny Alves
Estimando o impacto de um ano extra de Ensino
Fundamental na proficiência dos estudantes brasileiros /
Jhonny Alves Bez Birolo ; orientador, Francis Carlo
Petterini, 2023.
60 p.

Trabalho de Conclusão de Curso (graduação) -
Universidade Federal de Santa Catarina, Centro
Socioeconômico, Graduação em Ciências Econômicas,
Florianópolis, 2023.

Inclui referências.

1. Ciências Econômicas. 2. Impacto de um ano escolar. 3.
Ensino Fundamental Anos Iniciais. 4. SAEB 2009 e 2011. I.
Petterini, Francis Carlo. II. Universidade Federal de
Santa Catarina. Graduação em Ciências Econômicas. III. Título.

Jhonny Alves Bez Birolo

Estimando o impacto de um ano extra de Ensino Fundamental na proficiência dos estudantes brasileiros

Este Trabalho de Conclusão de Curso foi julgado adequado para obtenção do título de Bacharel e aprovado em sua forma final pelo Curso de Graduação em Ciências Econômicas.

Local Florianópolis, 27 de junho de 2023.

Coordenação do Curso

Banca examinadora



Documento assinado digitalmente
Francis Carlo Petterini Lourenco
Data: 03/07/2023 10:12:50-0300
CPF: ***.714.300-**
Verifique as assinaturas em <https://v.ufsc.br>

Prof.(a) Francis Carlo Petterini Lourenço, Dr.(a)
Orientador(a)

Prof.(a) Guilherme Valle Moura, Dr.(a)
Instituição Universidade Federal de Sana Catarina – UFSC.

Prof.(a) Michele Romanello, Dr.(a)
Instituição Universidade Federal de Sana Catarina – UFSC.

Florianópolis, 2023.

AGRADECIMENTOS

Quero primeiro agradecer aos professores do Curso de Graduação em Ciências Econômicas da Universidade Federal de Santa Catarina – UFSC. Principalmente, aos professores Wagner Leal Arienti, Milton Biage, Marcelo Arend e Francis Carlo Petterini Lourenço que me orientaram ou me fizeram renovar o meu interesse pelo curso. Também, um agradecimento especial aos professores Guilherme Valle Moura e Michele Romanello, por aceitarem avaliar este trabalho.

Quero agradecer as pessoas que trabalham comigo no Tribunal de Contas do Estado de Santa Catarina – TCE/SC. Menciono, Akauã Flores Arroyo e Silvio Bhering Sallum. Ao primeiro, por nunca me negar ajuda, sobretudo em programação, e por ser agora um grande amigo; ao segundo, porque sempre fez questão de reconhecer o meu trabalho. Junto aos outros, agradeço o ambiente maravilhoso e inspirador que conseguimos criar diariamente.

Quero agradecer imensamente a minha família, porque me deram o maior presente: uma infância fantástica. Hoje, sempre me apoiam e estão presentes.

Quero agradecer aos meus amigos de infância, meu muito obrigado. Vocês são uma constante importante nas minhas melhores lembranças.

A cada idade corresponde uma forma de vida que tem valor, equilíbrio, coerência que merece ser respeitada e levada a sério; a cada idade correspondem problemas e conflitos reais (...), pois o tempo todo, ela (a criança) teve de enfrentar situações novas (...). Temos de incentivá-la a gostar da sua idade, a desfrutar do seu presente. (SNYDERS, 2009, p. 3).

RESUMO

Este estudo procurou estimar o impacto na proficiência média das escolas brasileiras em Língua Portuguesa e Matemática, para um subconjunto de escolas que realizaram os exames da Prova Brasil com uma escolaridade maior. O Ensino Fundamental de 9 anos deixou uma oportunidade de estimar o impacto do acréscimo de um ano escolar para esse subconjunto, no momento em que os seus alunos realizaram a Prova Brasil no 5º ano contra os alunos no Ensino Fundamental de 8 anos que realizaram os mesmos exames na 4ª série. Para estimar o impacto de maneira consistente, utilizou-se o método de diferenças-em-diferenças (DiD), por suas suposições mais fracas quando comparadas aos estimadores de simples diferença de média na proficiência entre os subconjuntos de escolas, e a mudança na proficiência média antes e depois. Definiu-se como tratamento deste estudo o Ensino Fundamental de 9 anos, e os períodos de tempo, 2009, onde nenhuma escola tinha sofrido o impacto do tratamento (alunos realizando a Prova Brasil no 5º ano), e 2011, primeira edição da Prova Brasil com resultados para os alunos no 5º ano. Os grupos do DiD, de tratamento e de controle, foram formados de maneira a cumprir as suas suposições, nenhuma escola tratada em 2009 e apenas o subconjunto tratado em 2011. Isso foi assegurado, sabendo que o Ensino Fundamental de 9 anos foi aprovado em 2006, considerando o subconjunto de escolas que o implantou nos primeiros dois anos após a sua aprovação como as escolas tratadas e as escolas que o implementarem nos anos subsequentes, as de comparação (controle). O impacto foi mensurado para três dados em painel balanceados: sem controle, com controle e com controle e pareamento. Com dados da Prova Brasil e do Censo Escolar, os resultados sugerem que o impacto médio do tratamento sobre o subconjunto de escolas tratadas (EMPT) foi de 4,89, 2,91 e 1,57 pontos positivos e significativos, respectivamente, na proficiência em Língua Portuguesa; o ganho é um pouco maior em Matemática, chegando a 5,67, 3,84 e 2,61 pontos, na mesma ordem. Comparado aos estimadores com suposições mais fortes, o EMPT aumenta, em média, para 10,92 (diferença de médias) e 11,28 (“antes e depois”) pontos na proficiência em Língua Portuguesa e Matemática, respectivamente, para os dados em painel sem controle. Os estimadores foram interpretados como uma evolução na precisão das estimativas em direção ao painel com controle e pareamento, ancorado na literatura econométrica. O pequeno ganho de proficiência média dos alunos no subconjunto de escolas tratadas, apesar de significativo, encontra apoio no precário acúmulo educacional dos estudantes brasileiros, apresentando como principais justificativas as elevadas taxas de repetência escolar, distorção idade-série e um *mix* de outras variáveis que motivam pouco e sem sucesso a permanência dos alunos nas escolas ou que não entregam uma chance de permanência equitativa.

Palavras-chave: Impacto de um ano escolar; Ensino Fundamental Anos Iniciais; SAEB 2009 e 2011.

ABSTRACT

This study sought to estimate the impact on the average proficiency of Brazilian schools in Portuguese Language and Mathematics, for a subset of schools that took the Prova Brasil exams with a higher level of education. The 9-year Elementary School left an opportunity to estimate the impact of adding one school year for this subset, when its students took the Prova Brasil in the 5th year against students in the 8-year Elementary School who took the same exams in 4th grade. To consistently estimate the impact, the differences-in-differences (DiD) method was used, because of its weaker assumptions when compared to simple mean difference in proficiency estimators between subsets of schools, and change in proficiency mean before and after. The treatment of this study was defined as the Elementary School of 9 years, and the periods of time, 2009, where no school had suffered the impact of the treatment (students taking the Prova Brasil in the 5th year), and 2011, the first edition of the Prova Brasil with results for students in the 5th year. The DiD, treatment and control groups were formed in a way that fulfilled their assumptions, no schools treated in 2009 and only the subset treated in 2011. This was ensured, knowing that the 9-year Elementary School was approved in 2006, considering the subset of schools that implemented it in the first two years after its approval as the treated schools and the schools that implemented it in subsequent years, the comparison (control). Impact was measured for three balanced panel data: no control, with control, and with control and matching. With data from Prova Brasil and Censo Escolar, the results suggest that the mean impact of treatment on the subset of treated schools (EMPT) was 4.89, 2.91 and 1.57 positive and significant points, respectively, on proficiency in Portuguese; the gain is slightly higher in Mathematics, reaching 5.67, 3.84 and 2.61 points, in the same order. Compared to estimators with stronger assumptions, the EMPT increases, on average, to 10.92 (mean difference) and 11.28 (“before and after”) points in proficiency in Portuguese and Mathematics, respectively, for the data in control panel. The estimators were interpreted as an evolution in the precision of the estimates towards the panel with control and pairing, anchored in the econometric literature. The small average proficiency gain of students in the subset of treated schools, despite being significant, finds support in the precarious educational accumulation of Brazilian students, presenting as main justifications the high rates of school repetition, age-grade distortion and a mix of other variables that motivate little and unsuccessfully for students to remain in schools or that do not offer an equal chance of permanence.

Keywords: Impact of a school year; Elementary School Early Years; SAEB 2009 e 2011.

SUMÁRIO

1	INTRODUÇÃO	9
1.1	VIABILIDADE DA ESTIMATIVA	12
2	DADOS E METODOLOGIA	16
2.1	INSUMOS DA PROFICIÊNCIA	16
2.2	DADOS	18
2.3	AMOSTRA.....	19
2.4	METODOLOGIA.....	28
2.5	DIFERENÇAS-EM-DIFERENÇAS	28
2.5.1	Pareamento.....	31
2.5.2	Dados em Painel.....	33
2.5.3	Efeitos Fixos	34
2.6	RESULTADOS	34
2.6.1	Painel 1: sem controle ou pareamento.....	34
2.6.2	Painel 2: com controle	35
2.6.3	Painel 3: com pareamento e controle.....	37
2.6.4	Possível motivo dos resultados encontrados.....	38
2.7	MÉTODOS ERRÔNEOS.....	39
3	CONCLUSÕES.....	41
	REFERÊNCIAS.....	44
	APÊNDICE A – PROFICIÊNCIA E QUESTIONÁRIO	48
	APÊNDICE B – TENDÊNCIA PARALELA.....	49
	APÊNDICE C – MATRIZ DE CORRELAÇÃO DAS VARIÁVIES	53
	APÊNDICE D - DICIONÁRIO DE VARIÁVEIS.....	56

1 INTRODUÇÃO

Muitos estudos se concentram nos determinantes da proficiência dos alunos, através de um conjunto de insumos escolares e socioeconômicos; ver, *e.g.*, Hanushek, Kan e Rivkin (1998), Araújo e Siqueira (2010), Fraco (2008) e Lourenço et al. (2017). Em maior parte, seriam os insumos associados aos próprios alunos que melhor condicionariam o seu desempenho (GRAMANI, 2017; VARANI E SILVA, 2010; FELICIO E FERNANDES, 2005; FRACO, 2008; LOURENÇO ET AL., 2017). Os insumos da escola e dos professores não alcançam semelhante conclusão. As informações hoje disponíveis para esses insumos não conseguiriam esgotar ou controlar razoavelmente as suas influências, o que não aconteceria com as variáveis de alunos. Pritchett e Filmer (1997) também apontam que as variáveis controladas nas funções de produção educacionais apresentam dimensões dadas, segundo o observado. Como são exógenos, possivelmente os insumos associados aos professores, que conseguem promover mudanças em direção aos seus interesses, teriam atingido uma saída maximizada na proficiência dos alunos, enquanto os materiais escolares e outros características ainda conservariam margens positivas de ganhos.

Estimar um impacto de um ano extra na proficiência dos alunos aufere importância na medida que a educação é a fonte de muitos benefícios sociais e econômicos. Os indivíduos com maior escolaridade têm, em média, rendimentos maiores (CARD E KRUEGER, 1996) e menores chances de ter depressão, tanto devido as vantagens oriundas da renda *per capita* quanto do nível educacional (SANTOS E KASSOUF, 2007), por exemplo. Além disso, Friedman (1962) comenta sobre a importância de um eleitorado educado. No entanto, sua relevância consolida-se nas evidências que investimentos em educação estão diretamente associados ao crescimento econômico e consistente nos países. Entre 1890-1990, os gastos *per capita* em alunos nos Estados Unidos saíram de \$164, em 1890, para \$4.622, em 1990 (HANUSHEK, 2003). Observou-se um aumento do capital humano entre as décadas de 1980 e 1990; e das respostas positivas da escolaridade para os estudantes graduados. Diante disso, pouca não é a importância de investimentos em educação para os indivíduos e para onde moram.

Enquanto isso, os investimentos em educação no Brasil aparentam não ter sido suficientes historicamente, isso porque se mantem entre as gerações no Brasil uma coisa em comum: ausência nos alunos de uma proficiência adequada. O Anuário Brasileiro da Educação Básica (2019), divulgado por Todos Pela Educação, uma organização da sociedade civil com o propósito de melhorar a educação pública, apresentou um ciclo da continuidade

dos estudantes brasileiros dentro do sistema escolar em 2018. Para cada 100 estudantes que entraram na escola, 90 concluíram o Ensino Fundamental Anos Iniciais aos 12 anos de idade, com uma distorção idade-série de 2 anos. Destes, 60,70% apresentaram a proficiência esperada em Português e menos da metade, 48,90%, em Matemática. No Ensino Médio, a situação é ainda mais crítica. Dos 100 que entraram na escola, 64 concluíram com uma distorção idade-série de 2 anos, dos quais apenas 29,1% alcançaram a proficiência esperada em português e menos de 10% (6,4 alunos) alcançaram a proficiência adequada em Matemática.

Comparado aos estudantes de outros lugares do mundo, continuam insatisfatórios os rendimentos escolares alcançados pelos alunos brasileiros. Os resultados são do Programa Internacional de Avaliação dos Estudantes (Pisa). É uma avaliação internacional que estima a proficiência média dos estudantes de vários países, com idade entre 15 e 16 anos e devidamente matriculados em instituições de ensino a partir do 7º ano. O primeiro limiar, por ser a idade aproximada em que os estudantes terminam a escolaridade obrigatória entre os países avaliados. Em 2018, as avaliações foram separadas em três etapas, a primeira (núcleo) servia como determinante da dificuldade para as próximas, num modelo dinâmico de avaliação que considerou as diferenças de desempenho entre os estudantes e entre os países. Os estudantes brasileiros participantes alcançaram uma proficiência média em letramento em Leitura de 413 pontos, contra 384 pontos em Matemática. Isso significou, 74 e 104 pontos abaixo da média dos países participantes da OCDE (Organização para a Cooperação e Desenvolvimento Econômico). O mais preocupante, no entanto, é que os 10% dos estudantes brasileiros com melhor desempenho tiveram uma média de 501 pontos, menor que proficiência média de outros países considerando todos os estudantes participantes – Coreia (526), Canadá (512) e Finlândia (507). Entre os quatro níveis de proficiência em Matemática, 41,0% dos estudantes brasileiros ficaram no nível 1, único nível sem especificação da OCDE sobre as habilidades desenvolvidas. Medindo o quanto os estudantes, entre 15 e 16 anos, estão preparados para uma vida participativa em sociedade, através de “conhecimentos e habilidades essenciais” (BRASIL, 2020, p.17), os resultados do Pisa sugerem que os estudantes brasileiros ainda não atingiram uma capacidade de participação em sociedade que esgote os seus benefícios – mesmo em avaliações dinâmicas.

Sasaki et al. (2018) argumentam que o desempenho ruim dos estudantes brasileiros está relacionado, principalmente, as dificuldades dos estudantes brasileiros em responderem as avaliações do Pisa, não distribuindo adequadamente o tempo, por exemplo. Contudo, as avaliações dinâmicas em etapas do Pisa em 2018, levando em consideração o desempenho

dos alunos nas etapas anteriores, muito provavelmente aliviou a intensidade desse efeito. Em termos históricos, os dados apontam como principais problemas do baixo acúmulo de proficiência dos estudantes brasileiros: a repetência escolar, distorção idade-série e evasão escolar (RIBEIRO, 1991; ROJO, 2009; CLEIDE, 2014?). Em 1988, os alunos faziam a primeira série no mínimo duas vezes, onde 52,4% dos alunos iriam reprovar na 1ª série do 1º Grau, levando em média mais de 8,5 anos para terminar 6 séries (RIBERIO, 1990). No mesmo período, evadiram 7,52% da população de 7 a 14 anos. Uma década a frente, em 1996, os alunos demorariam em média 10,4 anos para concluírem as oito séries do Ensino Fundamental, com uma distorção idade-série de 46% (CLEIDE, 2014?). Ribeiro (1991) esclarece, no entanto, que as elevadas taxas de evasão escolar têm origem no processo de matricular novamente um aluno reprovado, sem eliminar a matrícula anterior. Além disso, as repetências causam novas repetências, não preparando melhor o aluno para o próximo ano e caminhando para um baixo acúmulo educacional, distorção idade-série e evasão escolar. Em 2002, Ferrão *et al.* não se distanciam das conclusões encontradas em 1988, sugerindo que os alunos com distorção idade-série têm um desempenho inferior na comparação com a idade adequada e os alunos em risco que reprovam saíram-se melhor se fossem promovidos. Com um acesso ao ensino quase universalizado, a reprovação é, muito provavelmente, a maior problemática da educação brasileira (sem considerar a qualidade na educação ofertada), pois exerceria um alcance também sobre a evasão escolar e a distorção idade-série (consequência imediata).

No entanto, na tentativa de encontrar os motivos da reprovação em massa, que contribuem igualmente para entender os baixos níveis de proficiência, muitas causas estariam supostamente associadas a conservação de valores culturais e ao insucesso em alcançar a proficiência dos alunos pelas metodologias de ensino. Bourdieu e Passeron (1968), estudando o sistema escolar francês, sugerem que os exames tradicionais teriam uma função de conservar, e esconder que conservam, privilégios culturais. Muitos alunos que não tiveram sucesso escolar, na verdade, seriam alunos que foram eliminados através dos exames, em uma determinação cultural por não corresponder ao sucesso escolar esperado da sua classe social. Sistemáticamente, é atribuído um fracasso aos alunos que, supostamente, foram eliminados pelo exame. No Brasil, Ribeiro (1990) concluiu que os alunos brasileiros seriam avaliados no primeiro ano com um fracasso previamente definido pela cultura do sistema escolar, sendo mais intenso em torno das classes menos favorecidas. Outro ponto, é que durante a evolução das metodologias de ensino maior atenção foi dada a escrita, em encargo da leitura. Como consequência, muitos estudiosos apontam para uma alfabetização que ensina o aluno a ler e a

escrever, mas não transforma esse aluno em leitor (Barbosa, 1990). Uma das dificuldades seria a linguagem usada nas escolas, muito diferente da linguagem social das camadas mais pobres. Somado ao aluno não continuar voluntariamente os estudos, isso também fornece novas justificativas para as elevadas taxas de reprovação, distorção idade-série e evasão escolar na menor resistência do aluno em seguir os estudos diante de adversidades.

Em resumo, tem-se muito estudos sobre os determinantes da proficiência dos alunos, mas nenhum do impacto isolado de um ano escolar a mais em relação a alunos com um ano escolar a menos. Os resultados educacionais são preocupantes, encontrando justificativas nas elevadas taxas de reprovação escolar e nas metodologias de ensino, sem muito sucesso em tocar no aluno a curiosidade por aprender. Outros autores sugerem um quadro maior, na medida que o exame elimina os alunos onde é atribuído ao fracasso escolar. Supondo uma diminuição das taxas de reprovação escolar, ou seja, o exame deixaria de ter a sua função conservadora, não ter-se-ia a eliminação da conservação, podendo ser atribuída e alcançada pelos baixos níveis de proficiência com que os estudantes brasileiros terminam a escola.

Com isso, há evidências razoáveis para embasar a suposição deste estudo. Sustenta-se que os ganhos marginais em realizar exames oficiais com um acréscimo de um ano escolar são muito pequenos comparado os alunos que realizam com um ano escolar a menos. Sendo a verificação dessa suposição a intenção maior desta monografia.

1.1 VIABILIDADE DA ESTIMATIVA

Em 6 de fevereiro de 2006, foi aprovada a Lei nº 11.274. Esta regulamentou o Ensino Fundamental de 9 anos, com início aos 6 anos de idade. Como, anteriormente, o Ensino Fundamental tinha 8 anos, com início obrigatório aos 7 anos de idade, essa mudança deixou uma oportunidade de estimar o impacto de um ano escolar. Os alunos no novo Ensino Fundamental passaram a realizar as avaliações oficiais, como o Sistema de Avaliação da Educação Básica (SAEB), com um ano escolar a mais em relação aos alunos, ainda no antigo Ensino Fundamental, que realizariam as mesmas avaliações com um ano escolar a menos.

O novo Ensino Fundamental foi o resultado do Projeto de Lei (PL) da Câmara nº 144, de 2005, que regulou a matrícula obrigatório aos 6 anos de idade e do PL nº 144, de 2005, que o precedeu. Os motivos da aprovação da lei são apresentados no PL nº 3.675, de 2004, versão original do PL nº 144, de 2005. Entre as considerações, estão uma procura por melhorar a educação ofertada, na medida que o acesso à educação se universaliza; acompanhar as metas do Plano Nacional de Educação (PNE), aprovado pela Lei nº 10.172, de

9 de janeiro de 2001; e regulamentar o Ensino Fundamental de 9 anos, que vinha sendo implementado em muitas escolas desde 1998, quando a Câmara de Educação Básica (CEB) do Conselho Nacional de Educação (CNE) recomendou a mesma extensão e antecipação da matrícula para o Ensino Fundamental da rede pública. No entanto, também poderiam argumentar que alunos que entram na escola mais cedo apresentam uma proficiência média maior (ARAÚJO, SIQUEIRA, 2010). Fátima Cleide (2014?) também destaca o atraso das crianças brasileiras em entrar no Ensino Fundamental quando comparado aos de outros países – reduzido com a lei.

O Ensino Fundamental 9 anos foi inclusão de um ano escolar no início (Secretaria de Educação Básica/Ministério da Educação – SEB/MEC). Não houve uma transferência automática dos conteúdos da primeira série para o segundo ano, mas uma reestruturação curricular para incluir no primeiro ano o início da alfabetização (SEB/MEC, 2009). A Lei nº 11.274, de 6 de fevereiro de 2006, determinou um prazo de quatro anos para a adoção do novo regime, entre 2006 e 2010. Apesar das escolas, nesse período, incluíram um ano a mais de ensino com os alunos matriculados aos 6 anos de idade, antes a matrícula obrigatória era aos 7 anos, os alunos do Ensino Fundamental de 8 anos não sofreram mudança de regime, continuando nos prazos e condições do antigo regime (MEC). Não obstante, as escolas tinham turmas matriculadas no novo regime e outras no antigo regime.

No Ensino Fundamental de 9 anos, os alunos deixaram de fazer as avaliações do SAEB na 4ª série, como no Ensino Fundamental de 8 anos, para realizá-las no 5º ano. Com isso, os alunos matriculados no Ensino Fundamental em turmas do 5º ano ou da 4ª série, igualmente com uma idade esperada de 10 anos, tiveram suas proficiências testadas com as mesmas avaliações do SAEB. No entanto, o primeiro subconjunto desses alunos, matriculados nas turmas do 5º ano, com um ano escolar a mais. O mesmo aconteceu nas outras etapas de conclusão onde o SAEB foi, e ainda é, aplicado, mas esse estudo se reserva ao Ensino Fundamental Anos Iniciais.

Para determinar esse impacto, foi empregado o estimador de diferenças-em-diferenças (DiD). Uma ótima descrição do método é vista em Callaway e Sant’Anna (2021):

“A Diferença em Diferenças (DiD) tornou-se um dos projetos de pesquisa mais populares usados para avaliar os efeitos causais das intervenções políticas. Em seu formato canônico, existem dois períodos de tempo e dois grupos: no primeiro período ninguém é tratado, e no segundo período algumas unidades são tratadas (o grupo tratado) e algumas unidades não são (o grupo de comparação). Se, na ausência de tratamento, os resultados médios dos grupos tratados e de comparação tivessem seguido trajetórias paralelas ao longo do tempo (que é a chamada suposição de tendências paralelas), pode-se estimar o efeito médio do tratamento para a

subpopulação tratada (ATT) por comparar a mudança média nos resultados experimentados pelo grupo tratado com a mudança média nos resultados experimentados pelo grupo de comparação.” (p. 200, *traduzido*).

Com isso, fica claro que a intervenção política deste estudo é a Lei nº 11.274, de 6 de fevereiro de 2006, que aprova o Ensino Fundamental de 9 anos. O grupo de tratamento são as escolas que cumpriram a lei nos primeiros dois anos após a sua aprovação, e o grupo de controle as escolas que cumpriam a lei nos anos subsequentes aos dois primeiros. Assim, assegura-se que no primeiro período de tempo ninguém é tratado e no segundo período de tempo apenas o subconjunto de escolas no grupo de tratamento é tratado. No Apêndice C são apresentadas evidências que os grupos gerados por este estudo seguem, possivelmente, uma trajetória paralela na proficiência média das escolas em Língua Portuguesa e Matemática. Isso torna o subconjunto de controle (comparação) um subconjunto consistente da mudança média esperada das escolas tratadas na ausência do tratamento. O efeito médio do tratamento para a subpopulação tratada, sigla ATT em inglês, foi substituído por sua versão em português, EMPT.

O primeiro período de tempo deste estudo, onde nenhum subconjunto de escolas é tratado, é 2009. Em 2005, ano imediatamente anterior ao tratamento e convencionalmente dito como primeiro período de tempo do DiD na literatura, não poderia ser o primeiro período desse estudo, pois o código das escolas na edição de 2005 era mascarado, não tornando possível a necessidade de acompanhar a evolução dessas escolas nas próximas edições sem máscaras do SAEB. Como o primeiro efeito do tratamento acontece apenas em 2011, que é o segundo período de tempo do DiD por isso, as próximas edições anteriores a 2011 poderiam servir como período anterior ao tratamento, no mínimo, dos seus efeitos. Dentro das opções de 2007 e 2009, a edição de 2009 foi escolhida como o primeiro período de tempo, supondo que com uma menor distância entre os dois períodos de tempo do DiD melhor se conservaria as chances de tendência paralela continuada. Com isso, ficam 2009 e 2011 como os períodos de tempo do DiD, respectivamente, anterior e posterior ao efeito do tratamento.

Para confirmar as suposições sobre os subconjuntos de escolas não tratadas no primeiro período (2009) e apenas o subconjunto de escolas do grupo de tratamento tratado no segundo período de tempo (2011), são apresentadas as Tabelas 1 e 2. Descrevem as etapas de ensino ofertadas pelos subconjuntos de escolas tratadas e de controle, condicionada ao momento em que cumprem a Lei nº 11.274, de 6 de fevereiro de 2006. Cada período possível de cumprimento da lei é apresentado junto as etapas possíveis de funcionamento dentro dos subconjuntos de escolas, na vertical, e o novo regime com as etapas iniciando no momento do

cumprimento. Em negrito destaca-se as etapas em que os subconjuntos de escolas realizam o SAEB e as duas linhas em que o ano está em negrito representado os períodos de tempo do DiD, o que juntas indicam as etapas em que os alunos realizaram, efetivamente, o SAEB. Em 2009, nenhuma escola é tratada. Em 2011, o subconjunto de escolas do grupo de tratamento é tratado e o subconjunto de escolas do grupo de comparação continua sem o tratamento.

Tabela 1 – Etapas de ensino ofertadas pelas escolas do grupo de tratamento entre 2005-2011.

Ano	Ano de implementação do Ensino Fundamental de 9 anos			
	2006		2007	
	Antigo Regime (série)	Novo Regime (ano)	Antigo Regime (série)	Novo Regime (ano)
2005	1; 2; 3; 4; 5; 6; 7; 8			
2006	2; 3; 4; 5; 6; 7; 8	1	1; 2; 3; 4; 5; 6; 7; 8	
2007	3; 4; 5; 6; 7; 8	2; 1	2; 3; 4; 5; 6; 7; 8	1
2008	4; 5; 6; 7; 8	3; 2; 1	3; 4; 5; 6; 7; 8	2; 1
2009	5; 6; 7; 8	4; 3; 2; 1	4; 5; 6; 7; 8	3; 2; 1
2010	6; 7; 8	5; 4; 3; 2; 1	5; 6; 7; 8	4; 3; 2; 1
2011	7; 8	6; 5; 4; 3; 2; 1	6; 7; 8	5; 4; 3; 2; 1

Fonte: autoria própria.

Tabela 2 – Etapas de ensino ofertadas pelas escolas do grupo de controle entre 2007-2011.

Ano	Ano de implementação do Ensino Fundamental de 9 anos					
	2008		2009		2010	
	Antigo Regime (série)	Novo Regime (ano)	Antigo Regime (série)	Novo Regime (ano)	Antigo Regime (série)	Novo Regime (ano)
2005						
2006						
2007	1; 2; 3; 4; 5; 6; 7; 8					
2008	2; 3; 4; 5; 6; 7; 8	1	1; 2; 3; 4; 5; 6; 7; 8			
2009	3; 4; 5; 6; 7; 8	2; 1	2; 3; 4; 5; 6; 7; 8	1	1; 2; 3; 4; 5; 6; 7; 8	
2010	4; 5; 6; 7; 8	3; 2; 1	3; 4; 5; 6; 7; 8	2; 1	2; 3; 4; 5; 6; 7; 8	1
2011	5; 6; 7; 8	4; 3; 2; 1	4; 5; 6; 7; 8	3; 2; 1	3; 4; 5; 6; 7; 8	2; 1

Fonte: autoria própria.

A próxima seção trata do referencial teórico, pela necessidade de se estimar o desempenho dos alunos controlando os fatores que o determinam. Portanto, trata das variáveis que influenciam o desempenho dos alunos em exames padronizados no Ensino Fundamental. Em seguida, se mostra a composição da amostra, quais indicadores foram utilizados e empregados para retratar as variáveis que determinam o desempenho dos alunos. A quarta seção, é a metodologia, estendendo o modelo de DiD para além do exposto. Isso significa incluir suas hipóteses adicionais e quando respeitadas, as metodologias adicionais usadas para melhorar as estimativas. Os resultados são então expostos, a partir do DiD com diferentes metodologias adicionais. A última seção, conclusões, reserva-se a confirmação ou rejeição da hipótese, que é o propósito, formulada por esta monografia.

2 DADOS E METODOLOGIA

2.1 INSUMOS DA PROFICIÊNCIA

Como para Pierre Bourdieu (1983), e muitos outros estudiosos do assunto:

“Com 3 ou 4 variáveis, 3 ou 4 propriedades de cada aluno, a profissão de seu pai, de sua mãe, ou seja, o tempo livre da sua mãe, o local de sua residência, a distância da cidade, o sexo, 3 ou 4 variáveis, pode-se levantar, com uma precisão muito forte, as chances de sucesso numa ou outra carreira, num ou outro curso escolar etc.” (p. 285, segundo a versão traduzida por Valle e Soulié (2019)).

Dentre os muitos estudos acerca dos determinantes da proficiência dos alunos no Ensino Fundamental, uma conclusão permanece desde o trabalho pioneiro de Coleman (*Report Coleman*, 1966): são as características associadas ao aluno que determinam a maior margem do seu desempenho escolar (GRAMANI, 2017; VARANI E SILVA, 2010; FELÍCIO E FERNANDES, 2005; FRANCO, 2008; ARAÚJO E SIQUEIRA, 2010; SOARES E SÁTYRO, 2008; LOURENÇO ET AL., 2017).

Controlar os fatores que influenciam o desempenho dos alunos na quarta série (e quinto ano) do Ensino Fundamental é importante tanto para construir um subconjunto de escolas no grupo de comparação consistente com a mudança esperada do subconjunto de escolas do grupo de tratamento na ausência do tratamento, quanto para eliminar a dimensão do efeito que seria atribuído ao tratamento quando na verdade são de outros fatores.

Em 1966, foi publicado *Report Coleman* (Relatório Coleman), uma pesquisa nos Estados Unidos que procurou encontrar os determinantes da diferença entre a proficiência de alunos brancos e negros. O estudo contou com, aproximadamente, 600 mil alunos. *A priori*, se sustentava que os últimos estavam concentrados em escolas piores e isso terminava em ganhos educacionais menores. O estudo, pelo contrário, mostrou que a qualidade das escolas e as características dos professores exerciam pouca influência sobre o desempenho desses alunos e a maior margem seria explicada pelo *background* familiar. O que mais se tentou provar em seguida foi a influência da qualidade da escola e dos professores.

Enquanto o *background* familiar dos alunos e suas características não geram dúvidas, a influência do efeito-professor e efeito-escola ainda são incertas – muito embora, não intuitivamente. Felício e Fernandes (2005) estão entre os apoiadores que ambos os efeitos são significativos. Estudando a proficiência em Língua Portuguesa e Matemática dos alunos na 4ª série do Ensino Fundamental, com base no SAEB (2001), observaram que dentro da

desigualdade total da proficiência dos alunos em Língua Portuguesa e Matemática, 0 a 28,4% e 8 a 34,4%, respectivamente, é explicada pela qualidade da escola.

Franco (2008), numa tentativa ambiciosa de encontrar “os determinantes da qualidade da educação no Brasil” (título da sua tese), analisa os dados em painel de 1997 a 2005 da proficiência em Matemática dos alunos na 4ª série do Ensino Fundamental. O estudo contou com os dados SAEB, com escopo nacional. Dentro das variáveis significativas associadas a escola e não ao aluno, a autora encontra, entre outras, a influência da escolaridade, experiência, sexo e se fez curso de capacitação dos professores; influência da experiência, problemas com professores faltosos e construção do projeto pedagógico dos diretores; e o número de matrículas, localização e se tem biblioteca das escolas – além de variáveis de turmas, como tamanho da turma, horas aula e rotatividade dos professores.

Araújo e Siqueira (2010), com o mesmo objetivo de Franco (2008), estudam a proficiência em Matemática dos alunos na 4ª série do Ensino Fundamental, igualmente, com dados do SAEB (2001). Os autores escolhem um conjunto de variáveis, algumas pouco comuns em trabalhos semelhantes, e estimam por um modelo Probit ordenado – relacionam as variáveis com o desempenho do aluno ordenado pela escala de proficiência do SAEB. Sobre o efeito escola, o trabalho dos autores evidenciou mais uma vez a importância da localização da escola sobre a proficiência, mas a maior contribuição foi a influência positiva que observaram entre a idade em que o aluno começa na escola sobre o desempenho escolar dele.

Soares e Sátyro (2008), usando dados em painel de 1998 a 2005 com efeito fixo, verificaram o impacto de variáveis de insumos escolares sobre a distorção idade-série nas escolas. Eles encontram um impacto significativo desses insumos nas escolas e um aumento dele em função da precariedade da escola. Entre as variáveis avaliadas, estavam: docentes com nível superior, média de alunos por turma, média de horas-aula, biblioteca e fator de infraestrutura.

Pritchett e Filmer (1997) não questionam a influência do professor como insumo escolar, mas demonstram que o professor como o principal agente de mudanças no sistema escolar consegue favorecer os seus interesses, levando os níveis salariais e outras características da sua categoria a uma dimensão que, muito provavelmente, esteja maximizando o seu retorno sobre a função de produção educacional. Os autores mencionam estudos em que os materiais de instrução para os alunos têm um custo-efetividade entre 17 e 34 vezes maiores em relação aos ganhos por acréscimos adicionais nos salários dos professores.

Hanushek, Kain e Rivkin (1998) assinalam que a influência da escola não vem dos seus insumos tradicionais, como estrutura física e outros recursos, mas da heterogeneidade na qualidade entre os insumos dentro e entre as escolas. Os fatores escolares têm efeitos, mas não por seus atributos (físicos e financeiros, por exemplos). Nessa heterogeneidade dos insumos, é a variação na qualidade dos professores a fonte de maior explicação do desempenho dos alunos, em torno, de 7,5% da variação total dos ganhos de desempenho. Os autores ainda destacam para a dificuldade de encontrar tais efeitos pelo simples levantamento de variáveis que descrevam as características das escolas e professores.

Os estudos apresentam argumentos convincentes da importância ou não dos efeitos dos professores e das escolas, mas nenhum discorda da sua atribuição como insumos escolares. Em suma, especificá-los tem maior garantia de controlar seus possíveis efeitos do que os negligenciar. Todas as variáveis de controle incluídas neste estudo estão no Apêndice D, de acordo com a literatura e que, simultaneamente, apresentaram uma associação esperada com a proficiência segundo a mesma; essa última, a correlação entre as variáveis de controle e a proficiência em Língua Portuguesa e Matemática, é demonstrada no Apêndice C.

2.2 DADOS

Os subconjuntos de escolas foram formados através da variável TP_ETAPA_ENSINO, uma das variáveis do Censo Escolar. Na medida em que nenhum indicador educacional apresenta o ano de cumprimento do Ensino Fundamental de 9 anos pelas escolas, a variável mencionada oferece uma indicação indireta sobre isso, descrevendo a etapa de ensino lecionada por cada turma dentro das escolas. As escolas associam cada turma com sua respectiva escolaridade, mas após a reforma do Ensino Fundamental, com as opções do 1º ao 9º ano disponíveis, sem eliminar as opções da 1ª a 8ª série. Dado que as escolas tinham entre 2006 e 2010 para fazer a mudança, sem os alunos do antigo regime, mesmo com a mudança, trocaram de regime (MEC). Portanto, uma escola que implementou o novo Ensino Fundamental em 2010, teriam alunos matriculados no antigo regime do ano anterior, em 2009, estudando na escola até, no mínimo, 2016, quando se formariam no Ensino Fundamental Anos Finais. Com isso, usando os dados do Censo Escolar de 2007 para encontrar a escolaridade das turmas com a variável, os subconjuntos de escolas de cada grupo foram montados. Como a variável não indicava as datas, 2007 foi usado como limiar de divisão na formação dos grupos, considerando as escolas do subconjunto de tratamento as escolas com uma ou mais turmas matriculadas nas etapas do Ensino Fundamental de 9 anos –

com ou sem turmas no outro regime –, e o subconjunto de comparação todas as escolas com turmas somente, e tão somente matriculadas nas etapas do antigo regime.

Os dados sobre as variáveis de controle e proficiência foram coletados do SAEB e do Censo Escolar. Os dois são indicadores educacionais sobre a administração do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), procurando dimensionar a proficiência, as características dos agentes e as condições ambientais em que os estudantes se formam, e as informações mais desagregadas sobre as escolas, respectivamente. O Sistema de Avaliação da Educação Básica (SAEB) existe desde 1990 e começou em 1995 a formular as suas provas com o método de Teoria de Resposta ao Item – TRI, o que tornou as suas edições comparáveis. O motivo de ser, igualmente, os resultados das escolas nas avaliações do SAEB as informações sobre a proficiência das escolas em Língua Portuguesa e Matemática deste estudo.

Além disso, o SAEB sofreu uma mudança. Em 2005, através da Portaria Ministerial nº 931, o sistema de avaliação foi dividido em duas avaliações, a Avaliação Nacional da Educação Básica (Aneb) e Avaliação Nacional do Rendimento Escolar (Anresc). Disso, o antigo SAEB foi continuado com a Aneb, enquanto a inclusão do Anresc representou uma evolução no alcance do indicador, atendendo ao chamado de muitos gestores escolares por informações censitárias sobre as escolas brasileiras, em termos de proficiência (BRASIL, SAEB). Mais conhecida como Prova Brasil, a Anresc é aplicada em todas as escolas com um número de matrículas, maior ou igual, a 20, mas apenas nas etapas de conclusão do Ensino Fundamental e/ou Ensino Médio onde o SAEB estima a proficiência dos alunos e compara se alcançaram a proficiência esperada nessas etapas de encerramento.

Nas edições de 2009 e 2011, as avaliações censitárias, como os questionários, da Prova Brasil não foram ministradas em escolas particulares. O impacto do tratamento, com isso, é sobre as escolas públicas, de dependência municipal, estadual ou federal. A escolha da Prova Brasil como insumo para as variáveis de controle e proficiência deste estudo, foi uma tentativa de considerar o maior número possível de escolas. Não obstante, foi usado como insumo para as variáveis de controle das escolas nos subconjuntos os dados do Censo Escolar. O Apêndice D apresenta e descreve as variáveis de controle e proficiência deste estudo no formato de dicionário.

2.3 AMOSTRA

Os subconjuntos de escolas de tratamento e de controle deste estudo são formados apenas por escolas públicas, de dependência municipal, estadual ou federal. Os dados sobre as variáveis de controle e de proficiência dessas escolas foram coletados para os dois períodos de tempo do DiD, 2009 e 2011. Por consequência, as amostras são dados em painel, com as entidades sendo as escolas e o mínimo de dois tempos, o antes e após do tratamento. Os painéis são igualmente balanceados – contudo, nos dois primeiros painéis o subconjunto de escolas de tratamento tem um número maior de escolas –, com nenhuma escola tratada no primeiro tempo e somente um subconjunto no segundo período de tempo, as escolas tratadas.

Como as estimativas do EMPT é em unidade escolar, as variáveis de alunos e professores coletadas dos questionários foram agregadas em forma de proporção. O cálculo das proporções seguiu rigorosamente a regra de incluir no cálculo somente os alunos e professores de cada escola que preencherem todas as questões de interesse deste estudo, preenchendo ou não as outras questões do questionário. Isso assegurou uma soma 100 entre as proporções das opções disponíveis em cada questão, transformando a resposta de cada questão selecionada em uma variável de controle dos alunos ou professores em unidade escolar. Não obstante, as variáveis de diretor, da escola e de proficiência estavam previamente na unidade adequada, sendo a última a média da proficiência dos alunos participantes da Prova Brasil da escola, em turma do 5º ano ou 4ª série, e não a média da proficiência dos alunos restantes nos painéis, após as regras de inclusão ou exclusão das amostras.

Muito embora, considerou-se necessário e vantajoso manter somente as escolas em que as proporções de variáveis de alunos em unidade escolar foram calculadas com, no mínimo, de 10 aluno sobrando após a imposição de apenas considerar os alunos que preencheram todas as questões de interesse deste estudo. Com isso, tentou-se minimizar as chances de relações espúrias pôr as proporções não serem fieis as características médias dos alunos de cada escola, relacionado a proficiência das escolas a um vetor de características que não seria honesto com a realidade da escola e do desempenho dela.

Além disso, encontrou-se uma associação forte entre a proficiência dos alunos e o preenchimento incompleto ou ausente do questionário pelos alunos, como mostra o Apêndice A. Para contornar essa possível segunda relação espúria por eliminação de estudantes com menores proficiência, dado que a proficiência média das escolas é dos alunos participantes da Prova Brasil e não dos alunos restantes nas amostras, que poderia estar concentrando o impacto sobre os alunos dentro das escolas com rendimento médio maior sem ser o rendimento médio da escola considerado nos painéis, especificou-se na função de produção educacional a variável de proporção de alunos removidos por não responderem as questões de

interesse deste estudo em cada escola. Assim, supostamente o viés seria controlado por essa variável, na medida que estaria relacionada com a proficiência média dos alunos participantes da Prova Brasil na escola e com os alunos removidos de menor proficiência, ajudando a manter a validade interna deste estudo.

O número de escolas eliminadas a cada critério adicional imposto é apresentado na sequência, pela Tabela 3. O painel sem controle ou pareamento termina com mais de 70 mil escolas, caindo para menos de 10 mil no painel com controle e menos de 5 mil escolas no painel com controle e pareamento. A maioria das escolas foram perdidas devido as variáveis de controle dos professores – partindo do seu questionário, as imposições diminuem a amostra para menos da metade (43,12%), contra 87,47% que permanece da amostra inicial dos alunos e 73,52% da amostra inicial dos diretores com os dados de 2009 –, apontando para uma maior omissão desses agentes no preenchimento dos questionários escolares da Prova Brasil.

Tabela 3 – Número de escolas restantes após as restrições impostas pelas amostras.

Critério	Em número de escolas (proporção em %)		
	2.007	2.009	2.011
Censo Escolar*			
*(todas as escolas; filtro por escolas que ofertam a 4ª série/5º ano através da Prova Brasil)	198.507	203.455	242.147
Somente escolas com as variáveis escolares selecionadas disponíveis.		197.468	193.047
Prova Brasil		75.603	72.808
Somente escolas que ofertam a 4ª série/5º ano		43.600 (100)*	41.019 (100)
Critério: variáveis de controle			
Somente escolas que cumpriram todos os critérios relacionados a seleção de variáveis de características (controle) e que tinham a proficiência em Língua Portuguesa e Matemática, simultaneamente, disponíveis.		11.993 (27,51)	12.470 (30,40)
<i>Questionário Aluno</i>			
Somente escolas que ofertam 4ª série/5º ano.		43.886 (100)	41.019 (100)
Somente escolas com 10 ou mais matrículas após o cumprimento do critério das variáveis de alunos: alunos que responderam todas as perguntas selecionadas.		38.389 (87,47)	36.104 (88,02)
<i>Questionário Professor</i>			
Somente escolas que ofertam 4ª série/5º ano.		38.975 (100)	41.019 (100)
Somente escolas que cumpriram o critério das variáveis de professor: professores que responderam todas as perguntas selecionadas.		16.813 (43,14)	21.334 (52,01)
<i>Questionário Diretor</i>			

Critério	Em número de escolas (proporção em %)		
	2.007	2.009	2.011
Número de escolas (que é igual ao número de diretores).		58.374 (100)	56.222 (100)
Somente escolas que cumpriram o critério das variáveis de diretor: diretores que responderam todas as perguntas selecionadas.		42.919 (73,52)	33.619 (59,80)
Escolas com informações completas na Prova Brasil e no Censo Escolar		11.653 (26,73)	11.941 (29,11)
Escolas presentes em 2009 e 2011		4.158	4.158
Painéis Balanceados			
Painel 1: sem pareamento e sem controle			70.616
Escolas Ensino Fundamental Anos Iniciais na Prova Brasil que estavam no Censo Escolar		42.841	39.028
Escolas presentes em 2009 e 2011		35.308 (80,98)	35.308 (86,08)
Painel 2: com controle			8.316
Escolas com informações completas na Prova Brasil e no Censo Escolar		11.653	11.941
Escolas presentes em 2009 e 2011		4.158 (9,54)	4.158 (10,14)
Painel 3: com pareamento			3.580
Pareamento feito sobre o "Painel 2: com controle"		4.158	4.158
Escolas com suporte comum (0,01 desvio padrão de <i>score de proporção</i> de distância e uma escola de controle por escola tratada)		1.790 (4,10)	1.790 (4,36)
* Proporção da amostra, em percentual (%).			

Fonte: elaborado pelo autor, com dados do SAEB e Censo Escolar.

Dos subconjuntos de escolas que tinham as informações demandadas disponíveis, foram levantadas as estatísticas descritivas para ambos os subconjuntos nos dois períodos de tempo do DiD. Com exceção das proficiências, que incluem o valor máximo e mínimo, bem como o primeiro, segundo (mediana) e o terceiro quartil da distribuição, as estatísticas descritivas são as médias das variáveis de controle para o subconjunto de escolas tratadas ou de controle, em 2009 ou 2011. Na Tabela 4, as estatísticas são referentes aos dados em painel das amostras sem controle ou pareamento e com controle.

Destaca-se a evolução das proficiências em Língua Portuguesa e Matemática entre 2009 e 2011 e a diferença entre a média dos grupos, maior para as escolas tratadas. Outras variáveis com diferenças visíveis são TP_CARGH_TOTAL_PROF_B e IN_LABORATORIO_CIENCIAS, respectivamente, com uma média no subconjunto de escolas de controle de 19% e 7% contra a média de 9% e 12% no grupo de tratamento – mantem-se entre as amostras e os períodos de tempo.

Tabela 4 - Estatísticas descritivas das variáveis das amostras sem controle e com controle de características.

Variáveis	Sem Controle				Controle			
	2009		2011		2009		2011	
	Tratados	Controle	Tratados	Controle	Tratados	Controle	Tratados	Controle
MÉDIA_LP								
Min.	92,84	95,6	116,8	121,4	130,5	95,6	133,7	129,7
1º Quartil	166,46	159,1	174,1	162,1	173,6	172,9	180,5	176
Mediana	180,8	174,2	188,8	177,1	186,1	185,3	193,7	188,1
Média	181,13	175,1	188,9	178	187,1	185,7	194,2	188,3
3º Quartil	195,41	190,2	203,8	193,1	199,6	198,4	207,1	200,6
Max.	286,65	294,5	289,5	278,2	280,4	250,4	258,9	251,2
MÉDIA_MT								
Min.	100,6	102,2	132,4	130,6	136,4	102,2	143,3	142,8
1º Quartil	183,1	176	189,9	177,1	190,4	192	196,6	193,1
Mediana	199,8	193,1	207,2	194,3	206,2	208,7	213,2	209,3
Média	200,9	195,3	207,9	196,7	208	208,1	214,2	209,2
3º Quartil	217,6	213,1	225,3	214,4	223,9	224	230	224,8
Max.	316,4	319,3	310,4	314,5	316,4	278,6	291,7	281,3
Escolas Federais	,0006	,0004			,0004	0	,0004	0
Escolas Estaduais	,2579	,2758			,1753	,3635	,1750	,3621
Escolas Municipais	,7415	,7238			,8243	,6365	,8247	,6379
Escolas Urbanas	,8777	,8163			,9518	,9548	,9521	,9555
Escolas Rurais	,1223	,1837			,0482	,0452	,0479	,0445
TP_SEXO_ALUNO					,4907	,4916	,4830	,4885
TP_COR_ALUNO					,3459	,3957	,3425	,4043
TP_IDADE_ALUNO_A					,0008	,0010	,0005	,0007
TP_IDADE_ALUNO_B					,0346	,0475	,0192	,0669
TP_IDADE_ALUNO_C					,4518	,5920	,4183	,6155
TP_IDADE_ALUNO_D					,3290	,2268	,3952	,2045
TP_MORA_MAE					,9093	,9176	,9280	,9359
TP_ESC_MAE_A					,0931	,0897	,1140	,1120
TP_ESC_MAE_B					,2213	,2035	,1480	,1483
TP_ESC_MAE_C					,1490	,1448	,1277	,1329
TP_ESC_MAE_D					,1179	,1305	,1315	,1418
TP_ESC_MAE_E					,0959	,1000	,1082	,1170
TP_MORA_PAÍ					,6692	,6909	,7326	,7494
TP_ESC_PAÍ_A					,0911	,0869	,1119	,1116
TP_ESC_PAÍ_B					,1786	,1637	,1103	,1133
TP_ESC_PAÍ_C					,1281	,1217	,1095	,1189
TP_ESC_PAÍ_D					,0887	,0998	,1040	,1095
TP_ESC_PAÍ_E					,0916	,0950	,1018	,1144
TP_REUNIAO					,6297	,6602	,6180	,6478
TP_INC_EST					,9760	,9763	,9809	,9797
TP_INC_DEVER					,9632	,9637	,9671	,9683
TP_INC_LER					,9514	,9528	,9566	,9557
TP_INC_NAO_FALT					,9667	,9683	,9728	,9711
TP_CONV_ESC					,8365	,8451	,8392	,8536
TP_TRAB_FOR					,1205	,1198	,1098	,1104
TP_REPRÓV					,2812	,2314	,2543	,2320
PROP_ALUNOS_REMOV					,5273	,5237	,4936	,4729
TP_SEXO_PROF					,0768	,0821	,0843	,0825
TP_IDADE_PROF_A					,1280	,1279	,0947	,0950
TP_IDADE_PROF_B					,3521	,3247	,3648	,3415
TP_COR_PROF					,5059	,5964	,4941	,6166
TP_ESC_PROF					,9123	,8922	,9589	,9536
TP_ESC_SUP_PROF_A					,5398	,4758	,6633	,5898
TP_ESC_SUP_PROF_B					,0084	,0094	,0111	,0099
TP_ESC_SUP_PROF_C					,0039	,0056	,0017	,0022
TP_FCONT_PROF					,8837	,8613	,9106	,8699

Variáveis	Sem Controle				Controle			
	2009		2011		2009		2011	
	Tratados	Controle	Tratados	Controle	Tratados	Controle	Tratados	Controle
TP_SALB_TOTAL_PROF_A					,0203	,0279	,0413	,0549
TP_SALB_TOTAL_PROF_B					,1807	,1217	,1341	,0924
TP_SALB_TOTAL_PROF_C					,6220	,5760	,5750	,5214
TP_SALB_TOTAL_PROF_D					,1694	,2580	,2260	,3086
TP_TEMP_LEC_TOTAL_A					,0202	,0098	,0162	,0127
TP_TEMP_LEC_TOTAL_B					,2477	,2385	,2604	,2528
TP_NUM_ESC_TRAB_PROF_A					,3205	,3406	,3473	,3606
TP_NUM_ESC_TRAB_PROF_B					,6699	,6422	,6450	,6322
TP_CARGH_TOTAL_PROF_A					,2011	,1358	,1958	,1180
TP_CARGH_TOTAL_PROF_B					,0861	,1787	,0946	,1769
TP_PROP_LEC_B					,1621	,1097	,0873	,0668
TP_PROP_LEC_C					,4091	,3781	,4027	,4074
TP_PROP_LEC_D					,3827	,4669	,4890	,5029
TP_CAR_INFRA					,2424	,2625	,2529	,2318
TP_CURR_INAD					,2047	,1759	,1309	,1292
TP_CONF_DIRETOR_A					,8212	,8220	,8387	,8447
TP_CONF_DIRETOR_B					,1215	,1241	,1156	,1130
TP_AMB_ESC_A					,9120	,9162	,9160	,9174
TP_AMB_ESC_B					,0625	,0564	,0587	,0604
TP_INSUF_FIN_A					,4398	,4847	,4959	,5388
TP_INSUF_FIN_B					,4747	,4301	,4334	,4006
TP_FALT_PROF_A					,7856	,8054	,7969	,7890
TP_FALT_PROF_B					,1582	,1403	,1572	,1661
TP_FALT_ALUNOS_A					,6353	,6899	,6289	,5883
TP_FALT_ALUNOS_B					,2912	,2537	,2906	,3160
TP_ESC_DIR					,9738	,9804	,9865	,9933
TP_ESC_SUP_DIR_A					,7443	,6918	,8011	,7444
TP_ESC_SUP_DIR_B					,0123	,0202	,0153	,0290
TP_ESC_SUP_DIR_C					,0026	,0067	,0019	,0040
TP_FCONT_DIR					,9129	,9076	,9095	,9312
TP_SALB_DIR_A					,0026	,0000	,0179	,0196
TP_SALB_DIR_B					,0508	,0364	,0355	,0223
TP_SALB_DIR_C					,5462	,4450	,4120	,3318
TP_SALB_DIR_D					,3555	,4666	,4636	,5246
TP_TEMP_FUNC_DIR_A					,5155	,4127	,5062	,4214
TP_TEMP_FUNC_DIR_B					,3211	,3405	,2774	,2933
TP_ASS_DIR_A					,5092	,4383	,5039	,4403
TP_ASS_DIR_B					,4404	,3655	,4426	,3554
TP_PPD_DR_A					,0931	,1632	,1308	,1686
TP_PPD_DR_B					,0022	,0047	,0026	,0020
TP_PPD_DR_C					,6052	,5920	,7443	,6972
TP_PPD_DR_D					,2725	,2016	,1077	,1146
TP_VINC_EST_A					,0692	,1072	,0520	,0735
TP_VINC_EST_B					,0991	,1241	,0950	,1032
TP_VINC_EST_C					,1776	,1962	,1798	,2131
TP_VINC_EST_D					,2628	,2475	,2710	,2637
TP_ROT_CORP_DOC_A					,6994	,6817	,7286	,7411
TP_ROT_CORP_DOC_B					,2355	,2286	,2120	,2009
IN_ESGOTO_INEXISTENTE					,0011	,0013	,0019	,0020
IN_LABORATORIO_INFORMATICA					,6168	,6696	,8243	,8119
IN_LABORATORIO_CIENCIAS					,1136	,0668	,1207	,0728
IN_QUADRA_ESPORTES					,6221	,6514	,6112	,6709
IN_BIBLIOTECA					,6826	,2974	,6557	,3129
IN_COMPUTADOR					,9712	,9703	,9925	,9899
IN_INTERNET					,7944	,8281	,9080	,9096
IN_TRAT	,6596	,3403			,6433	,3567		
QT_SALAS_EXISTENTES					12,2	12,0	12,6	12,1

Variáveis	Sem Controle				Controle			
	2009		2011		2009		2011	
	Tratados	Controle	Tratados	Controle	Tratados	Controle	Tratados	Controle
RAZAO_MAT_DOC					24,35	24,24	22,92	22,38
QT_MAT_BAS					667	691	625	632
QT_MATR*					45,21	62,45	42,76	58,06
QT_DOC_BAS					28	28	28	28
QT_PROF*					1,5	1,7	1,7	1,9

* Depois de condicionado ao critério de seleção.

Fonte: elaborado pelo autor.

As estatísticas descritivas dos dados em painel da amostra com controle e pareamento são apresentadas na Tabela 5. O pareamento é um instrumento que escolhe um subconjunto de escolas de comparação através de um conjunto de variáveis de controle observadas na tentativa de encontrar para cada escola tratada a mais semelhante, por essas características, não tratada. Bem-sucedido, tem-se uma diferença não significativa entre as médias das características dos subconjuntos das escolas, como acontece em experimentos controlados (PEIXOTO *et al.*, 2012).

Para verificar a significância entre a diferença das características médias dos grupos, testou-se a diferença entre as médias para as variáveis de controle no primeiro período de tempo do DiD (2009), localizada na terceira coluna da Tabela 5. Usou-se a regra prática em que resultados menores que 0,25 são não significativos (RUBIN, 2001). Sem nenhum caso de diferença significativa, o pareamento foi consistente com a sua proposta de espelhar um experimento controlado nos observáveis.

Tabela 5 - Estatísticas descritivas das variáveis da amostra pareada e a diferença de média das características entre os grupos de tratamento e controle pós-pareamento.

Variáveis	2009		Teste de Diferença de Média	2011	
	Tratados	Controle		Tratados	Controle
MEDIA_LP					
Min.	130,5	95,6		145,7	137,3
1° Quartil	173,5	170,1		179,6	173,3
Mediana	186,1	184,2		193,6	186
Média	186,4	184		193,5	186,6
3° Quartil	198,6	197,4		205,8	200,2
Max.	280,4	243,2		249,4	251,2
MEDIA_MT					
Min.	148,2	102,2		159,5	152,1
1° Quartil	190,4	188,1		196,4	189,1
Mediana	207,2	205,7		213,1	206,9
Média	207,2	205,8		213,5	207,3
3° Quartil	222,4	222,7		228,9	223,3
Max.	316,4	270,7		283,5	278,4
Escolas Federais	0	0	0	0	0
Escolas Estaduais	,1944	,1866	,0176	,1944	,1866
Escolas Municipais	,8056	,8134	-,0176	,8056	,8134

Variáveis	2009		Teste de Diferença de Média	2011	
	Tratados	Controle		Tratados	Controle
Escolas Urbanas	,9486	,9464	,0104	,9464	,9475
Escolas Rurais	,0514	,0536	-,0104	,0536	,0525
TP_SEXO_ALUNO	,4922	,4931	-,0100	,4824	,4876
TP_COR_ALUNO	,3813	,3811	,0011	,3706	,3848
TP_IDADE_ALUNO_A	,0010	,0010	-,0034	,0003	,0008
TP_IDADE_ALUNO_B	,0507	,0491	,0318	,0278	,0631
TP_IDADE_ALUNO_C	,5333	,5258	,0414	,4644	,5555
TP_IDADE_ALUNO_D	,2473	,2541	-,0453	,3561	,2340
TP_MORA_MAE	,9126	,9112	,0241	,9309	,9306
TP_ESC_MAE_A	,0991	,0971	,0283	,1187	,1235
TP_ESC_MAE_B	,2177	,2144	,0316	,1504	,1528
TP_ESC_MAE_C	,1432	,1441	-,0134	,1263	,1331
TP_ESC_MAE_D	,1189	,1209	-,0293	,1357	,1321
TP_ESC_MAE_E	,0948	,0964	-,0210	,1080	,1138
TP_MORA_PAI	,6920	,6858	,0555	,7474	,7477
TP_ESC_PAI_A	,0954	,0948	,0083	,1203	,1214
TP_ESC_PAI_B	,1739	,1732	,0074	,1103	,1183
TP_ESC_PAI_C	,1205	,1215	-,0147	,1077	,1154
TP_ESC_PAI_D	,0914	,0913	,0029	,1067	,1029
TP_ESC_PAI_E	,0881	,0913	-,0473	,0992	,1096
TP_REUNIAO	,6372	,6299	,0530	,6247	,6189
TP_INC_EST	,9748	,9755	-,0205	,9814	,9781
TP_INC_DEVER	,9623	,9618	,0119	,9663	,9664
TP_INC_LER	,9514	,9509	,0115	,9550	,9532
TP_INC_NAO_FALT	,9673	,9664	,0244	,9728	,9697
TP_CONV_ESC	,8389	,8397	-,0101	,8416	,8454
TP_TRAB_FOR	,1214	,1222	-,0096	,1063	,1182
TP_REPROV	,2731	,2741	-,0065	,2481	,2799
PROP_ALUNOS_REMOV	,5211	,5201	,0084	,4842	,4909
TP_SEXO_PROF	,0899	,0907	-,0033	,1032	,0952
TP_IDADE_PROF_A	,1418	,1325	,0307	,1011	,1021
TP_IDADE_PROF_B	,3492	,3731	-,0556	,3593	,3802
TP_COR_PROF	,5315	,5452	-,0294	,5381	,5649
TP_ESC_PROF	,8886	,8959	-,0282	,9624	,9462
TP_ESC_SUP_PROF_A	,5306	,5255	,0111	,6607	,6461
TP_ESC_SUP_PROF_B	,0081	,0099	-,0211	,0102	,0072
TP_ESC_SUP_PROF_C	,0034	,0029	,0079	,0028	,0021
TP_FCONT_PROF	,8845	,8894	-,0166	,9156	,9036
TP_SALB_TOTAL_PROF_A	,0256	,0305	-,0382	,0349	,0519
TP_SALB_TOTAL_PROF_B	,1593	,1557	,0102	,1517	,1190
TP_SALB_TOTAL_PROF_C	,6057	,6079	-,0052	,5537	,5407
TP_SALB_TOTAL_PROF_D	,1983	,1958	,0072	,2402	,2686
TP_TEMP_LEC_TOTAL_A	,0132	,0127	,0043	,0186	,0123
TP_TEMP_LEC_TOTAL_B	,2484	,2558	-,0191	,2489	,2713
TP_NUM_ESC_TRAB_PROF_A	,3620	,3533	,0203	,3757	,3797
TP_NUM_ESC_TRAB_PROF_B	,6237	,6295	-,0137	,6143	,6149
TP_CARGH_TOTAL_PROF_A	,1835	,1810	,0069	,1857	,1538
TP_CARGH_TOTAL_PROF_B	,1238	,1235	,0012	,1117	,1375
TP_PROP_LEC_B	,1286	,1149	,0408	,0827	,0740
TP_PROP_LEC_C	,4121	,4117	,0010	,3875	,4300
TP_PROP_LEC_D	,4165	,4232	-,0148	,5102	,4697
TP_CAR_INFRA	,2796	,2868	-,0184	,2400	,2681
TP_CURR_INAD	,2013	,1969	,0119	,1210	,1507
TP_CONF_DIRETOR_A	,7975	,8126	-,0426	,8404	,8296
TP_CONF_DIRETOR_B	,1343	,1237	,0356	,1151	,1225
TP_AMB_ESC_A	,9061	,9039	,0083	,9177	,9016

Variáveis	2009		Teste de Diferença de Média	2011	
	Tratados	Controle		Tratados	Controle
TP_AMB_ESC_B	,0636	,0664	-,0129	,0543	,0745
TP_INSUF_FIN_A	,4454	,4365	,0194	,5281	,5211
TP_INSUF_FIN_B	,4512	,4664	-,0333	,4126	,4093
TP_FALT_PROF_A	,7891	,7888	,0009	,7995	,7968
TP_FALT_PROF_B	,1553	,1529	,0072	,1620	,1621
TP_FALT_ALUNOS_A	,6619	,6632	-,0029	,6443	,5955
TP_FALT_ALUNOS_B	,2783	,2814	-,0073	,2859	,3191
TP_ESC_DIR	,9687	,9743	-,0350	,9877	,9933
TP_ESC_SUP_DIR_A	,7251	,7196	,0128	,7944	,7575
TP_ESC_SUP_DIR_B	,0134	,0112	,0202	,0168	,0190
TP_ESC_SUP_DIR_C	,0056	,0022	,0656	,0034	,0011
TP_FCONT_DIR	,9207	,9151	,0198	,9128	,9318
TP_SALB_DIR_A	,0000	,0000	,0000	,0201	,0212
TP_SALB_DIR_B	,0592	,0480	,0509	,0447	,0302
TP_SALB_DIR_C	,5240	,5296	-,0112	,4056	,4179
TP_SALB_DIR_D	,3732	,3754	-,0047	,4670	,4547
TP_TEMP_FUNC_DIR_A	,4894	,4726	,0335	,5095	,4838
TP_TEMP_FUNC_DIR_B	,3464	,3318	,0311	,2804	,3039
TP_ASS_DIR_A	,4447	,4492	-,0089	,4559	,4335
TP_ASS_DIR_B	,4670	,4637	,0068	,4704	,4358
TP_PPD_DR_A	,1352	,1263	,0308	,1520	,1564
TP_PPD_DR_B	,0056	,0045	,0236	,0045	,0034
TP_PPD_DR_C	,6045	,6101	-,0114	,7363	,6961
TP_PPD_DR_D	,2134	,2145	-,0025	,0905	,1251
TP_VINC_EST_A	,0860	,0927	-,0264	,0670	,0648
TP_VINC_EST_B	,1106	,1039	,0224	,0972	,0950
TP_VINC_EST_C	,1687	,1631	,0146	,1810	,2022
TP_VINC_EST_D	,2480	,2447	,0076	,2704	,2514
TP_ROT_CORP_DOC_A	,7017	,7095	-,0171	,7408	,7419
TP_ROT_CORP_DOC_B	,2279	,2190	,0211	,2089	,2078
IN_ESGOTO_INEXISTENTE	,0022	,0011	,0334	,0022	,0022
IN_LABORATORIO_INFORMATICA	,6257	,5944	,0644	,8067	,7899
IN_LABORATORIO_CIENCIAS	,0805	,0726	,0246	,0883	,0793
IN_QUADRA ESPORTES	,5911	,5966	-,0115	,6168	,5978
IN_BIBLIOTECA	,4950	,4525	,0912	,5374	,4447
IN_COMPUTADOR	,9665	,9575	,0535	,9922	,9866
IN_INTERNET	,7888	,7676	,0525	,9263	,8737
IN_TRAT	,50	,50			
QT_SALAS_EXISTENTES	11,7	11,6	,022	12,1	11,8
RAZAO_MAT_DOC	24,37	24,42	-,006	22,80	22,31
QT_MAT_BAS	643	647		608	595
QT_MATR*	46,92	55,80		42,67	51,39
QT_DOC_BAS	27	26		27	26
QT_PROF*	1,5	1,6		1,7	1,8
Média			,00558		

* Depois de condicionado ao critério de seleção.

Fonte: elaborado pelo autor.

Com o pareamento, as variáveis de controle TP_CARGH_TOTAL_PROF_B e IN_LABORATORIO_CIENCIAS deixam de apresentar uma média discrepante entre os subconjuntos de escolas. Para 2009, os valores médios são 12,38% e 8,05% para as escolas de comparação e 12,35% e 7,26% para as escolas tratadas, respectivamente.

2.4 METODOLOGIA

2.5 DIFERENÇAS-EM-DIFERENÇAS

Modelos especificados com variáveis de controle em função da proficiência escolar são chamados de *função de produção educacional* na literatura (PRITCHETT e FILMER, 1997). O estimador do DiD é calculado implicitamente por meio dessa função. Na sequência, a equação (I) apresenta a função de produção educacional do presente estudo, para a proficiência em Língua Portuguesa. Ela é formada por um conjunto de insumos de alunos, professores, diretores e das escolas, sendo o mesmo modelo estimado para a proficiência em Matemática. Com auxílio da álgebra, pode ser escrita como:

$$LP_{it} = \tau + \alpha_1 A_{it} + \alpha_2 P_{it} + \alpha_3 D_{it} + \alpha_4 E_{it} + \gamma T_{it} + \rho t_{it} + \beta(T_{it}t_{it}) + \varepsilon_{it}, \quad (I)$$

em que i é a escola em análise; t é o tempo em que se analisa a escola, antes ou depois do tratamento; LP_{it} é a média da escola i no tempo t em Língua Portuguesa na Prova Brasil; A_{it} , P_{it} , D_{it} e E_{it} são, respectivamente, o vetor de características observadas dos alunos, professores, diretores, e das escolas na escola i no tempo t ; T_{it} é a dummy de tratamento, assumido o valor 1 para escolas no grupo tratado e 0 para as escolas do grupo de controle; t_{it} é a dummy de tempo, assumindo o valor 1 para o período pós-tratamento (2011) e 0 para o período pré-tratamento (2009); $T_{it}t_{it}$ é a interação das dummies de tratamento e tempo, fazendo com que β seja a diferença do efeito do tratamento das escolas no grupo tratado depois do tratamento em relação ao grupo de controle; α_1 , α_2 , α_3 e α_4 são, respectivamente, os coeficientes do efeito das variáveis observadas do vetor de alunos, professores, diretores e das escolas sobre a proficiência média da escola; τ é o termo do intercepto; γ é a diferença de proficiência do grupo tratado e de controle; ρ é a diferença da proficiência das escolas antes e depois do tratamento; ε_{it} é o termo do ruído branco [$\sim RB(0, \sigma_i^2)$], que se espere ser independente das variáveis especificadas.

O impacto médio do tratamento sobre os tratados (EMPT) pode ser encontrado estimando o DiD através da função de produção educacional deste estudo, equação (I). Como mencionado por Callaway e Sant'Anna (2021), o DiD calcula a mudança da proficiência média do subconjunto de tratamento e de comparação entre os tempos do DiD, 2009 e 2011. Em seguida, mensura a diferença na mudança temporal da proficiência média dos grupos. Supondo que o grupo de controle apresente tendência paralela com o grupo de tratamento na

proficiência média, o resultado da segunda subtração é atribuído ao único condicionante presente em um grupo e ausente no outro: o tratamento. Ou melhor, o efeito na proficiência do acréscimo de um ano escolar.

Somado a tendência paralela na mudança de proficiência média no grupo de tratamento e no grupo de controle, o DiD conta com outras duas suposições: (1) que não ocorram mudanças significativas na composição dos grupos e (2) que os grupos não sofram mudanças que os afetam de maneira não homogênea no segundo tempo do DiD (PEIXOTO *et al.*, 2012). Enquanto para a suposição de tendências paralelas são apresentadas possíveis evidências do seu cumprimento no Apêndice B, as outras suposições são tomadas como satisfeitas, por controlar este estudo um conjunto extenso de variáveis em um curto espaço de tempo – o que pode não ser verdade e as estimativas seriam tendenciosas, para mais ou para menos, a depender da influência sofrida.

Com isso, através da equação (I) de proficiência em Língua Portuguesa, e omitindo propositalmente os subscritos e substituindo a soma dos vetores de características ($\alpha_1 A_{it} + \alpha_2 P_{it} + \alpha_3 D_{it} + \alpha_4 E_{it}$) por αX , para melhor visualizar as operações matemáticas, o DiD pode ser estimado pelos seguintes passos, algebricamente:

Grupo Tratado:

Antes do Tratamento,

$$\begin{aligned} E[LP_{it}|T = 1, t = 0] &= \tau + \alpha X + \gamma + \beta + E[\varepsilon|X, T = 1, t = 0] = \\ &= \tau + \alpha X + \gamma. \end{aligned}$$

Depois do tratamento,

$$\begin{aligned} E[LP_{it}|T = 1, t = 1] &= \tau + \alpha X + \gamma + \rho + \beta + E[\varepsilon|X, T = 1, t = 1] = \\ &= \tau + \alpha X + \gamma + \rho + \beta. \end{aligned}$$

Grupo de Controle:

Antes do tratamento,

$$\begin{aligned} E[LP_{it}|T = 0, t = 0] &= \tau + \alpha X + E[\varepsilon|X, T = 0, t = 0] = \\ &= \tau + \alpha X. \end{aligned}$$

Depois do tratamento,

$$\begin{aligned} E[LP_{it}|T = 0, t = 1] &= \alpha X + \rho + E[\varepsilon|X, T = 0, t = 1] = \\ &= \tau + \alpha X + \rho. \end{aligned}$$

Primeira Diferença (mudança temporal das médias dentro de cada grupo):

Grupo Tratado,

$$\begin{aligned} E[LP_{it}|T = 1, t = 1] - E[LP_{it}|T = 1, t = 0] &= \\ = (\tau + \alpha X + \gamma + \rho + \beta) - (\tau + \alpha X + \gamma) &= \\ = \rho + \beta. \end{aligned} \tag{II}$$

Grupo de Controle,

$$\begin{aligned} E[LP_{it}|T = 0, t = 1] - E[LP_{it}|T = 0, t = 0] &= \\ = (\tau + \alpha X + \rho) - (\tau + \alpha X + \rho) &= \\ = \rho. \end{aligned} \tag{III}$$

Segunda Diferença (diferença na mudança temporal entre a média do grupo de tratamento e do grupo de controle):

$$\begin{aligned} \{(II) - (III)\} &= \\ = (\rho + \beta) - (\rho) &= \\ = \beta. \end{aligned}$$

Portanto, β é o efeito médio do tratamento sobre os tratados (EMTP) da proficiência das escolas tratadas em Língua Portuguesa. Além disso, não é necessário fazer a dupla diferença explicitamente, o β faz isso internamente devido ao cruzamento das dummies de tempo e tratamento. No Apêndice B, o impacto médio é representado graficamente e indicado como EMPT. Sua linha vermelha pontilhada é a translação da mudança na proficiência média do subconjunto do contrafactual entre 2009 e 2011, para a escala de proficiência do subconjunto de escolas tratadas em 2009. Como a única coisa que, supostamente, diferencia os dois grupos é o tratamento, a diferença nas mudanças na proficiência média entre 2009 e 2011, a segunda diferença, dos subconjuntos só pode ser atribuída ao efeito de alguma coisa que tem em um subconjunto e não tem no outro, sobrando apenas o tratamento.

2.5.1 Pareamento

Como este estudo não é um experimento controlado, o subconjunto de escolas de comparação pode não ser um contrafactual consiste para o grupo de tratamento. Para contornar isso, foi feito o pareamento entre os subconjuntos de escolas na amostra de controle, tentando alcançar as propriedades ótimas dos experimentos controlados. O pareamento escolhe pares no grupo de controle para cada escola tratada que tem em comum um vetor de características observadas. Depois do pareamento, se espera que as características médias dos dois grupos não sejam diferentes, pelo menos nos observáveis, como acontece nos experimentos controlados.

O pareamento apresenta duas suposições mais fortes, esperando-se que as variáveis de controle sejam todas as variáveis que influenciam tanto a decisão da escola de cumprir ou não cumprir o tratamento em 2006 ou 2007, quanto a sua proficiência média em Língua Portuguesa e Matemática; e que essas variáveis de controle para o grupo de tratamento sejam igualmente representativas das características médias das escolas do subconjunto de comparação – sendo chamado de *suporte comum*, é necessário para cada escola encontrar seu par nos observáveis. Como o vetor das variáveis de controle deste estudo é demasiado extenso, para manter o suporte comum entre as escolas dos subconjuntos, usou-se o pareamento sugerido por Rosenbaum e Rubin (1983). Os pares das escolas entre os conjuntos são formados indiretamente através de suas variáveis de controle, ao calcular as suas chances condicionadas a essas variáveis de cumprirem ou não cumprirem o Ensino Fundamental de 9 anos em 2006 ou 2007; e construir novos subconjuntos formados por cada escola do subconjunto de comparação que encontrou uma probabilidade de participar do tratamento próxima da estimada para uma escola tratada, em seguida outra escola de controle é comparada com as escolas de tratamento até esgotar as escolas do subconjunto de comparação. Na literatura, a probabilidade condicionada ao vetor de variáveis observadas é chamada de *score de propensão*.

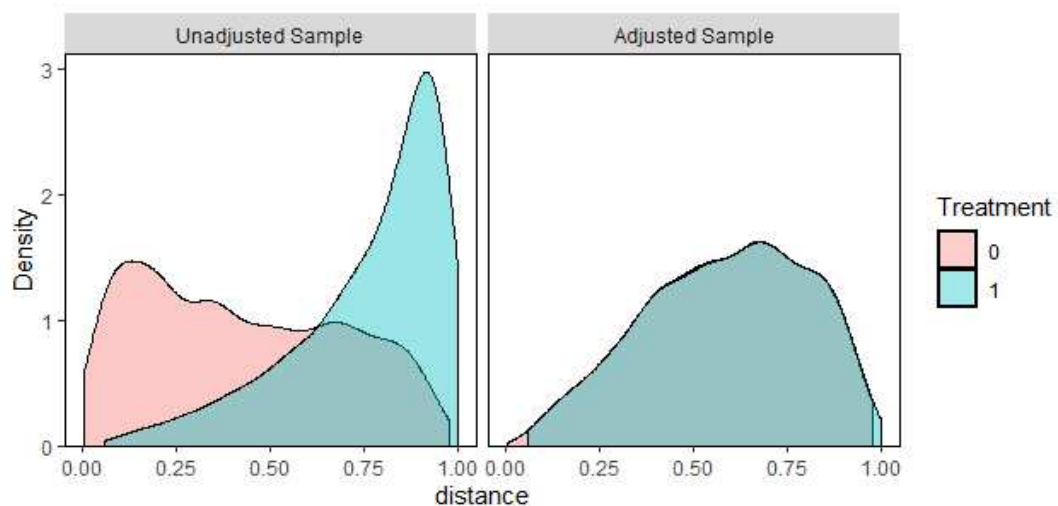
Além disso, para especificar a distância em probabilidade em que as escolas devem estar para formarem os pares, este estudo usou a métrica de *The Nearest Neighbor Matching* (score de propensão por vizinho mais próximo). O *score de propensão* usa uma função de densidade para atingir as restrições ótimas das distribuições de probabilidade, não alcançada por modelos de probabilidade linear. Por sua praticidade, optando-se pela função logística, onde a sua representação algébrica pode ser escrita como:

$$P[T = 1|X] = \frac{\exp(\delta x)}{1 + \exp(\delta x)}, \quad (IV)$$

em que x é o vetor das variáveis de controle; e δ é o vetor de coeficientes estimados. Depois de algumas manipulações, a equação (IV) fica linear nos parâmetros e facilmente pode ser calculada por mínimos quadrados ordinários.

Encontrado o *score* para ambos os grupos, o vizinho mais próximo define a distância, em desvios-padrões de *score*, e o número de escolas no grupo de controle que são ligadas a cada escola tratada. Usando a função logística, pelo método do vizinho mais próximo, os dois grupos foram formados por escolas que estavam, no máximo, a 0,01 desvios-padrão de *score de propensão* uma da outra, e associando apenas uma escola no grupo de controle para cada escola tratada. Nenhuma outra condição foi imposta! Para saber se o pareamento foi bem-sucedido, se comparou a características médias dos dois grupos na Tabela 5. Como analisado, nenhuma diferença média nas variáveis de controle entre os dois subconjuntos foi encontrada, sugerindo que o pareamento foi bem-sucedido na sua tentativa de alcançar as propriedades ótimas dos experimentos controlados.

Figura 1 – Distribuição de densidade do *score de propensão* das escolas, segundo seus respectivos subconjuntos. Igual a 1 para as escolas tratadas e 0 para as escolas no subconjunto de comparação.



Fonte: elaborado pelo autor.

Além disso, como mostra a Figura 1 acima, a suposição de suporte comum é atendida, sendo o vetor de características representativo de ambos os subconjuntos de escolas.

As escolas tratadas estão em azul (*Treatment = 1*), enquanto as escolas de comparação em vermelho (*Treatment = 0*). O gráfico de distribuição de densidade do *score de propensão* à esquerda é anterior ao pareamento, com maiores chances das escolas tratadas de receberem o tratamento. Pode ser descrito como a frequência de escolas que apresentam o respectivo *score de propensão* dentro do intervalo em análise, suficientemente pequeno. Na direita, o gráfico é a distribuição da probabilidade das escolas de serem tratadas após o pareamento, com uma distribuição muito próxima para ambos os subconjuntos de escolas, o que sugere indícios de um *suporte comum* entre eles – escolas de controle com *scores* próximos dos *scores* de escolas tratadas.

2.5.2 Dados em Painel

O estimador do DiD foi estimado através de três dados em painel, igualmente balanceados, para comparar a evolução dos resultados com a suposição de que os painéis seguintes são mais consistentes e precisos, corrigindo possíveis vieses dos painéis anteriores. O primeiro painel, sem controlar as variáveis que influenciam a proficiência média nas escolas, serviu como um *benchmarking* de uma estimativa enviesada e equivocada, por não cumprir as suposições do DiD como um contrafactual adequado. O segundo painel, controlou as variáveis observadas, mas não as variáveis não observadas não constantes no tempo, supostamente chegando a resultados igualmente tendenciosos, mesmo que agora menor. Como *benchmarking* de uma estimativa consistente e não enviesada foi usado os resultados do terceiro painel, que ao alcançar um pareamento que controlou as diferenças nos observáveis, muito provavelmente também controlou nos não observáveis, e mesmo que seja uma suposição muito forte, não seria exagero dizer que controlou melhor que as outras estimativas.

Além disso, os resultados de ambos os painéis foram estimados tanto para a proficiência média das escolas em Língua Portuguesa quanto para Matemática. Os resultados referem-se as variáveis de tempo, tratamento, constante (referência) e do impacto médio do programa. Para os dois últimos painéis, o painel com variáveis de controle e o painel com pareamento, foi estimado o impacto médio do tratamento para seis especificações diferentes: (1) sem especificar as variáveis de controle; (2) apenas as variáveis de alunos são especificadas; (3) apenas as variáveis de professores são especificadas; (4) apenas as variáveis de diretores são especificadas; (5) apenas as variáveis das escolas são especificadas; e, por

último, (6) todas as variáveis de controle são especificadas na função de produção educacional.

Os resultados, assim, são considerados uma evolução entre os painéis e dentro dos próprios painéis, em direção a estimativa (6) do painel com pareamento.

2.5.3 Efeitos Fixos

Os efeitos fixos são variáveis não observadas constantes no tempo, que influenciam na mesma intensidade a proficiência média das escolas nos dois tempos do DiD. Com isso, o estimador de DiD consegue controlar esses efeitos com a primeira diferença do seu método, eliminando os possíveis vieses associados a uma maior concentração de alunos motivados ou professores com melhores didáticas nas escolas de um dos subconjuntos deste estudo, sobre o qual não há informações disponíveis. Em particular, Stock e Watson (2003) mostram que estimar por efeitos fixos ou fazer a diferença “antes e depois” não modifica os resultados, chegando nos mesmos coeficientes sobre as variáveis especificadas (p. 192).

Controlar esse efeito é muito importante, sendo constante em estudos de função de produção educacionais. Muito embora, diferente dos estudos que procuram os determinantes da educação e necessitam especificar os efeitos fixos por remoção de média das variáveis ou inclusão de dummies para cada escola menos uma, esse estudo não necessita especificá-lo por ser controlado o seu efeito implicitamente pelo estimador de DiD.

2.6 RESULTADOS

2.6.1 Painel 1: sem controle ou pareamento

Com mais de 70 mil escolas, o impacto médio do tratamento é apresentado na Tabela 6. Os resultados sugerem um ganho médio de proficiência de 4,89 pontos em Língua Portuguesa e 5,67 pontos em Matemática. Isso representa um ganho médio de 2,59% e 2,73% sobre a proficiência média das escolas do subconjunto de tratamento em 2011, respectivamente, em Língua Portuguesa e Matemática (dados das tabelas 14 e 15). Além disso, comparando a escala de proficiência do Saeb (versão 2001), que apresenta um intervalo constante de 25 pontos entre os diferentes níveis associados as capacidades leitoras ou de resolução de problemas dos alunos, um aumento aproximado de 5 pontos em Língua Portuguesa ou Matemática corresponderia 1/5 na promoção entre as escalas. Caso os ganhos

fossem constantes, seriam necessários 5 anos para uma mudança de escala com os ganhos adicionais estimados nas proficiências por este estudo.

Tabela 6 - EMPT em Língua Portuguesa e Matemática – painel sem controle ou pareamento.

Variáveis	Proficiência	
	Língua Portuguesa	Matemática
Constante	175,09*** (0,195)	195,25*** (0,230)
β (EMPT)	4,89*** (0,340)	5,67*** (0,401)
Tempo	2,91*** (0,276)	1,40*** (0,326)
Tratamento	6,03*** (0,240)	5,60*** (0,284)
R-quadrado Ajustado	0,0555	0,0368

NOTA: Significativo com um tamanho de teste de: 0,1% (**); 1% (*); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Em todo caso, a estimativa não controlou nenhuma das variáveis de insumo dos alunos ou escolares. Por ausência de aleatoriedade, os resultados podem estar sub ou superestimando o efeito médio do tratamento sobre as escolas tratadas. Há, no subconjunto de escolas tratadas, uma concentração de alunos mais ricos? Os professores conseguem entregar uma maior qualidade e estímulos aos alunos? Se isso for verdade, especificar as variáveis diminuiria o efeito do tratamento. Isso foi feito na próxima estimativa.

2.6.2 Painel 2: com controle

Controlado as variáveis parcialmente e em conjunto, os resultados continuam a sugerindo um acréscimo positivo e significativo na proficiência média das escolas, para as duas disciplinas. No entanto, agora com menos de 9 mil escolas e especificando o vetor de variáveis de controle, o impacto médio diminuiu, respectivamente, para 2,91 pontos em Língua Portuguesa e 3,84 pontos em Matemática, para a especificação (6) – Tabelas 7 e 8. Uma variação de menos 40,5% e 47,65% nos ganhos médios, na mesma ordem, em relação ao Painel 1. Supostamente, as variáveis omitidas estavam correlacionadas tanto com a dummy de tratamento quanto com a proficiência média das escolas, provocando um viés do verdadeiro efeito do tratamento. No entanto, o controle dessas variáveis teria sido suficiente para eliminar ou diminuir a endogeneidade na função de produção educacional.

Tabela 7 - EMPT em Língua Portuguesa – Painel com controle.

Variáveis	Especificação					
	(1)	(2)	(3)	(4)	(5)	(6)
Constante	185,68*** (0,4999)	121,92*** (6,0940)	211,64*** (3,2477)	222,58*** (2,2065)	230,53*** (10,8656)	127,30*** (11,060)
β (EMPT)	4,52*** (0,8814)	3,33*** (0,5876)	5,11*** (0,6790)	6,15*** (0,7161)	5,86*** (0,6974)	2,91*** (0,5701)
Tempo	2,60*** (0,7070)	-0,42 (0,4934)	-1,95*** (0,5551)	-2,90*** (0,5836)	-2,97*** (0,5673)	-1,56** (0,4911)
Tratamento	1,38* (0,6232)	3,17*** (0,4214)	3,23*** (0,4873)	2,05*** (0,5120)	1,99*** (0,5181)	2,33*** (0,4289)
Características						
Aluno		Incluído				Incluído
Professor			Incluído			Incluído
Diretor				Incluído		Incluído
Escola					Incluído	Incluído
R-quadrado Ajustado	0,0306	0,5942	0,4312	0,3644	0,3957	0,6248

NOTA: Significativo com um tamanho de teste de: 0,1% (***) ; 1% (**); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Tabela 8 - EMPT em Matemática – painel com controle.

Variáveis	Especificação					
	(1)	(2)	(3)	(4)	(5)	(6)
Constante	208,06*** (0,6065)	115,48*** (7,7501)	234,76*** (3,9087)	249,94*** (2,6972)	266,50*** (13,1741)	129,40*** (10,3950)
β (EMPT)	5,04*** (1,0678)	4,37*** (0,7472)	5,77*** (0,8171)	7,03*** (0,8754)	7,00*** (0,8456)	3,84*** (0,7188)
Tempo	1,18 (0,8565)	-1,78** (0,6275)	-4,37*** (0,6680)	-5,26*** (0,7134)	-5,67*** (0,6877)	-3,34*** (0,6191)
Tratamento	-0,08 (0,7550)	2,32*** (0,5360)	2,19*** (0,5865)	0,51 (0,6259)	0,87 (0,6282)	1,58** (0,5407)
Características						
Aluno		Incluído				Incluído
Professor			Incluído			Incluído
Diretor				Incluído		Incluído
Escola					Incluído	Incluído
R-quadrado Ajustado	0,01365	0,545	0,4288	0,3416	0,3841	0,5865

NOTA: Significativo com um tamanho de teste de: 0,1% (***) ; 1% (**); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Os ganhos de proficiência em Matemática estimados são maiores de maneira consistente. Uma sugestão da causa, seria a objetividade da matemática e a similaridade dos exercícios. Enquanto, a Língua Portuguesa, mesmo tratando-se do mesmo assunto, explora uma diversidade de contextos. Além disso, como na literatura, é vetor de características dos alunos que apresenta o maior grau de explicação da sua própria proficiência. Quando incluído apenas esse vetor, essas características explicam 59,42% e 54,5% das variações observadas na função de produção educacional, respectivamente, em Língua Portuguesa e Matemática. Em seguida, vem os insumos dos professores, respectivamente, 43,12% e 42,18%.

Controlando as várias de insumo, os ganhos em proficiência alcançam apenas 1/8 e 1/6 do intervalo de 25 pontos entre os níveis de proficiência do SAEB, respectivamente, em Língua Portuguesa e Matemática. Por outras palavras, com ganhos constantes, os alunos

demorariam, em média, 8 anos e 6 anos para mudar de nível de proficiência, em sequência. Muito embora, o controle de variáveis observáveis diminua o viés de omissão, pode não conseguir controlar o efeito das variáveis não observadas que variam ou não no tempo. Em experimentos controlados essas suposições são tomadas como satisfeitas, na tentativa de reproduzir essas condições ótimas e controlar esses outros efeitos, estimou-se a seguir o DiD com dados em painel pareados. Supostamente, como os subconjuntos apresentam características médias observáveis não diferentes estatisticamente, não seria um exagero supor que as características médias não observadas também sejam, ainda mais diante do controle dos efeitos fixos pelo estimador de DiD.

2.6.3 Painel 3: com pareamento e controle

Com quase 4 mil escolas, o painel com pareamento também sugere ganhos médios maiores em Matemática – Tabelas 9 e 10. Os resultados apontam para um acréscimo na proficiência média das escolas de 4,49 pontos na especificação (1) para 1,57 pontos e quase não significativo, na especificação (6) em Língua Portuguesa. Diante disso, sai de um ganho semelhante ao Painel 1 para um ganho menor que a especificação (6) do Painel 2. Em termos de variação, e comparando apenas com a especificação (6) do Painel 2, tem-se uma diminuição de 46,05% e 32,03% no ganho médio, respectivamente, em Língua Portuguesa e Matemática. No exercício de comparação com os níveis de proficiência do SAEB, supondo um ganho constante de proficiência, uma escola levaria, em média, 16 anos e 9 anos para subir um nível entre os intervalos de proficiência do SAEB, respectivamente, em Língua Portuguesa e Matemática. Em proporção, 1/16 e 1/9 da escala de 25 pontos.

Tabela 9 - EMPT em Língua Portuguesa – painel com controle e pareamento.

Variáveis	Especificação					
	(1)	(2)	(3)	(4)	(5)	(6)
Constante	184,00*** (06446)	123,64*** (8,7472)	207,71*** (4,7032)	221,16*** (3,2233)	218,40*** (2,3730)	134,20*** (9,7980)
β (EMPT)	4,49*** (1,2892)	2,08* (0,8627)	3,58*** (0,9698)	3,76*** (1,0170)	3,73*** (1,0020)	1,57 . (0,8345)
Tempo	2,60** (0,9116)	1,08 . (0,6489)	-0,46 (0,7013)	-1,28 . (0,7310)	-1,52* (0,7251)	-4,28 (0,6578)
Tratamento	2,37** (0,9116)	2,37*** (0,5851)	2,59*** (0,6841)	2,58*** (0,7189)	2,15*** (0,7080)	2,41*** (0,5631)
Características						
Aluno		Incluído				Incluído
Professor			Incluído			Incluído
Diretor				Incluído		Incluído
Escola					Incluído	Incluído

Variáveis	Especificação					
	(1)	(2)	(3)	(4)	(5)	(6)
R-quadrado Ajustado	0,03164	0,602	0,456	0,3992	0,4171	0,6342

NOTA: Significativo com um tamanho de teste de: 0,1% (***) ; 1% (**); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Tabela 10 - EMPT em Matemática – painel com controle e pareamento.

Variáveis	Especificação					
	(1)	(2)	(3)	(4)	(5)	(6)
Constante	205,78*** (0,7818)	119,51*** (11,0418)	229,53*** (5,6322)	247,37*** (3,9427)	250,35*** (2,8653)	136,10*** (12,2905)
β (EMPT)	4,84** (1,5637)	3,36** (1,0890)	3,69** (1,1614)	3,94** (1,2440)	3,99*** (1,2090)	2,61* (1,0468)
Tempo	1,50 (1,1057)	0,32 (0,8191)	-2,33** (0,8398)	-3,08*** (0,8942)	-3,68*** (0,8754)	-1,86* (0,8251)
Tratamento	1,43 (1,1057)	1,31 (0,7386)	1,73* (0,8192)	1,65 (0,8793)	1,21 (0,8547)	1,42* (0,7063)
Características						
Aluno		Incluído				Incluído
Professor			Incluído			Incluído
Diretor				Incluído		Incluído
Escola					Incluído	Incluído
R-quadrado Ajustado	0,01541	0,5617	0,4608	0,3788	0,4129	0,6022

NOTA: Significativo com um tamanho de teste de: 0,1% (***) ; 1% (**); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Esse painel foi o único painel com uma significância menor, chegando ao tamanho de teste de 10% e 5% em Língua Portuguesa e Matemática, na ordem, contra 0,1% para os outros painéis. Sustenta-se que a diminuição do impacto médio e da sua significância sejam evidências no sentido de maior consistência e precisão nas estimativas do EMPT desse Painel 3, com pareamento. Com isso, possivelmente foram controlados outros efeitos não observados e que variam no tempo, no mínimo, de forma mais adequada que o controle indireto pelas variáveis especificadas, por exemplo. Portanto, os resultados com maior credibilidade deste estudo sugerem que o efeito médio do tratamento sobre os tratados (EMPT) foi de, em média, de 1,57 pontos em Língua Portuguesa e 2,61 pontos em Matemática.

2.6.4 Possível motivo dos resultados encontrados

O novo Ensino Fundamental foi um acréscimo de um ano escolar no início, mas o primeiro ano não passou a ser a primeira série do antigo regime. Apesar do primeiro ano não trabalhar o conteúdo do último ano da pré-escola, também não é uma simples alocação do conteúdo da primeira série do Ensino Fundamental de 8 anos. O SEB/MEC (2009) alerta:

“(…) não se trata de realizar um “arranjo” dos conteúdos da primeira série do Ensino Fundamental de oito anos. Faz-se necessário elaborar uma nova proposta político-

pedagógica e curricular coerente com as especificidades não só da criança de seis anos de idade, como também das demais crianças de sete, oito, nove e dez anos de idade que realizam os cinco anos iniciais do Ensino Fundamental, assim como os anos finais dessa etapa de ensino.” (p. 24)

Contudo, o primeiro ano apresenta os mesmos objetivos básicos dos demais anos dos Anos Iniciais do Ensino Fundamental: a alfabetização e o letramento das crianças, experimentada com outras áreas do conhecimento, não sendo exclusividade do segundo ano. Enquanto a Educação Infantil não tem esse objetivo formalizado, isto é, alfabetizar, o primeiro ano passa a ter (SEB/MEC, 2009). O efeito do tratamento sobre os tratados (EMPT) pode ser entendido, então, como o efeito de uma introdução ao ensino das primeiras letras mais consistente em comparação aos alunos do Ensino Fundamental de 8 anos, que vieram da Educação Infantil sem uma proposta de alfabetização.

A melhor forma de descrever os ganhos de proficiência estimados por este estudo seria: os ganhos de pontuação em função de um começo mais cedo da alfabetização. O baixo efeito da adoção do Ensino Fundamental de 9 anos, muito provavelmente, se deve à similaridade do conteúdo da pré-escola com do primeiro ano. Mesmo o último, explicitamente, ter como objetivo o ensino das primeiras letras, não isenta a pré-escola da oportunidade de introduzi-lo. Contudo, nas conclusões outros motivos são apresentados.

2.7 MÉTODOS ERRÔNEOS

Estimando o EMPT para os métodos que exigem suposições mais fortes, a simples diferença entre os grupos e “antes e depois”, os seus resultados sugerem uma superestimava do EMPT (Tabela 11). O primeiro mensura a diferença de proficiência média em Língua Portuguesa e Matemática entre os grupos de escolas tratadas e o grupo de escolas de controle no segundo período de tempo (2011). Não obstante, o método “antes e depois”, calcula a mudança na proficiência média em Língua Portuguesa e Matemática do grupo de escolas tratadas entre os períodos de tempo (2009 e 2011) – sem incluir as escolas de comparação.

Há suposições, respectivamente, que o grupo de controle é um contrafactual como os gerados em experimentos controlados e, no segundo caso, que apenas o efeito do acréscimo de um ano escolar que influenciou a proficiência média das escolas durante os dois períodos de tempo. Duas hipóteses consideradas fortes demais para ser tomadas como satisfeitas, seja pela ausência do controle de variáveis não observadas que variam no tempo ou pelo conjunto imenso de variáveis que influem a proficiência dos alunos nas avaliações.

Tabela 11 – EMPT em Língua Portuguesa e Matemática com os métodos de diferença de grupos, “antes e depois” e DiD, para os painéis sem controle ou pareamento, com controle e com controle e pareamento.

Proficiência/Amostra	Estimador		
	Diferença entre Grupos	Antes e Depois	DD
Língua Portuguesa			
Sem controle ou pareamento	10,92 (6,9 >)*	7,79 (5,0 >)	4,89 (3,1 >)
Com controle	5,91 (3,8 >)	7,13 (4,5 >)	2,91 (1,8 >)
Com controle e pareamento	6,87 (4,4 >)	7,09 (4,5 >)	1,57 (<i>sem viés</i>)
Matemática			
Sem controle ou pareamento	11,28 (4,3 >)	7,07 (2,7 >)	5,67 (2,2 >)
Com controle	4,96 (2,0 >)	6,23 (2,4 >)	3,84 (1,5 >)
Com controle e pareamento	6,27 (2,4 >)	6,35 (2,4 >)	2,61 (<i>sem viés</i>)

*Ordem do viés (em número de vezes maior) em relação ao resultado do DiD no painel com controle e pareamento (1,57 em Língua Portuguesa e 2,61 em Matemática).

Fonte: elaborado pelo autor.

Tomando os métodos errôneos e mensurado o EMPT para os dados em painel sem controle ou pareamento, seriam necessários, aproximadamente, 2 anos para uma mudança na escala de proficiência do SAEB (25 pontos), contra 4 anos do método “antes e depois” e 5 anos do DiD, para as duas disciplinas. Controlando o vetor de insumos escolares e dos alunos, a diferença entre grupos aponta um ganho de proficiência de 5,91 pontos em Língua Portuguesa e 4,96 pontos em Matemática, enquanto o estimador “antes e depois” mantém os ganhos em torno de 7 pontos para as duas proficiências. O DiD cai para 2,91 e 3,84 pontos, respectivamente. No último painel, os métodos errôneos continuam com os EMPT do modelo com controle, aumentando no caso da diferença entre grupo na proficiência em Matemática, de 4,96 para 6,27 pontos. O resultado do DiD sugere ganhos de 1,57 e 2,61 nesse último painel, respectivamente, em Língua Portuguesa e Matemática, caindo ainda mais em relação ao painel anterior.

Sustentando-se que ganhos do DiD na amostra pareada sejam os ganhos médios mais consistentes e razoáveis, levando as outras estimativas a uma superestimação do impacto de um ano escolar. Isso significa um viés de 4,5 vezes maior em Língua Portuguesa e 2,4 vezes em Matemática dos métodos errôneos para amostra com pareamento, diferença entre grupos e “antes e depois”, respectivamente. Em relação ao efeito do DiD nos painéis sem pareamento, sem controle e com controle, a ordem do viés cai para 3,1 e 1,8 em Língua Portuguesa e 2,2 e 1,5 em Matemática, na sequência. Logicamente, a atribuição do impacto a dimensão das outras estimativas seria um equívoco de, no mínimo, 1,8 vezes em Língua Portuguesa e 1,5 vezes em Matemática enviesados para mais.

3 CONCLUSÕES

Muito embora este estudo evidencie um impacto médio pequeno e significativo, não é possível associá-lo unicamente a qualidade das escolas públicas no Brasil. Principalmente, porque Wachtel (1975) e Card e Krueger (1996) argumentem que mensurar a qualidade escolar do aluno em um ano é uma medida imprecisa da sua qualidade escolar histórica – associado ao retorno da experiência e dos anos de escolaridade para o mercado de trabalho, respectivamente. O mesmo poderia acontecer com os ganhos em proficiência: o rendimento anual do aluno é muito menor que o seu rendimento escolar agregado, pois o aluno acumularia mais conhecimento com o tempo em comparação com um ano isolado. Supostamente, os resultados deste estudo não poderiam ser considerados como os ganhos médios em proficiência por um ano escolar dentro das escolas públicas brasileiras, na dependência municipal, estadual e federal. Seria necessário, condicioná-los ao ano escolar em que se estimou o efeito, a idade dos alunos e sem estender para os demais anos, mesmo assim com um possível viés para menos em relação ao que seria a distribuição dos ganhos históricos ao longo das etapas de ensino.

Os resultados sugerem ganhos positivos e significativos na proficiência tanto em Língua Portuguesa e Matemática, sendo maior na última. Os painéis sem controle, com controle e com pareamento apontam para um aumento médio na proficiência de 4,89, 2,91 e 1,57 pontos em Língua Portuguesa, respectivamente, contra 5,67, 3,84 e 2,61 pontos na proficiência em Matemática. Sabendo que, em Língua Portuguesa, a proficiência média no subconjunto de escolas tratadas, em 2011, foi de 193,46 e em Matemática 213,55 para a amostra com pareamento, esses ganhos de 1,57 e 2,61 pontos correspondem, aproximadamente, a 0,81% e 1,22% da proficiência média dessas escolas tratadas, respectivamente, em 2011.

Os resultados da amostra com pareamento são tomados como os mais consistentes e precisos. O DiD com pareamento controla o efeito médio de outras variáveis que influenciam a proficiência dos alunos melhor que os outros painéis e os métodos que exigem suposições mais fortes, como a simples diferença entre os subconjuntos de escolas em 2011 e a mudança média na proficiência antes e depois. No primeiro caso, uma superestimava na ordem de 4 vezes e, no segundo, em quase 3 vezes em relação a estimativa do DiD com pareamento e controlando todas as variáveis. Somado ao possível viés na estimação do impacto da escola em apenas um ano, quando não considerado o histórico escolar, a literatura também apoia o

pouco rendimento na proficiência dos alunos na educação brasileira, sugerindo ser verdadeira a sustentação deste estudo sobre possíveis ganhos marginais pequenos de um ano escolar a mais.

O estudo apresentou algumas limitações. A primeira e mais simples, foi a opção por não segregar o efeito do tratamento entre as dependências de ensino estadual, municipal e federal. Seria necessário considerar a dimensão da proficiência no ponto de partida entre os grupos segregados por dependência, porque o retorno de programas educacionais tende a ser maior em níveis mais precários de proficiência. O menor tamanho de amostra, também, concentraria ainda mais os resultados para um conjunto de escola, tornando as estimativas menos representativas das suas respectivas populações. Uma segunda limitação foi estimar o impacto em dimensão escolar, enquanto o tratamento é sobre os alunos – são eles que fazem as avaliações. Não obstante, a dimensão escolar aparenta ter a vantagem de incluir outras variáveis que controlam as diferenças socioeconômicas dos alunos, indiretamente também por agrupá-los na escola, mas pouco se pode dizer mais. Uma terceira e quarta limitação, são, respectivamente, a ausência de pesquisas para amparar as conclusões deste estudo e, a quarta e última limitação, o impacto médio nas escolas por ano de implementação do Ensino Fundamental de 9 anos ou o impacto médio para os mesmos alunos nas escolas ao longo do tempo, nas próximas etapas de ensino.

Uma forma fácil de contorna a terceira limitação é usando os dados da Aneb, que ministrou as mesmas avaliações para escolas privadas nos períodos de tempo deste estudo. No entanto, muitas escolas em 2009 poderiam não ter dados para 2011, em função da reamostragem da Aneb. Enquanto isso, a quarta limitação encontra suporte nos avanços recentes no estimador de DiD. Callaway e Sant’Anna (2021) apresentam uma extensão do método, tornado possível estimar o efeito médio de programas para múltiplos períodos de tempo e com tratados que variam no tempo em que são tratados. Como tinham entre 2006 e 2010, as escolas adotaram neste estudo o novo regime em períodos de tempo muito variados, podendo incluir o efeito de um ano escolar sobre as escolas que fizeram isso em 2008, 2009 ou 2010 com os resultados da Prova Brasil de 2013 e 2015, por exemplo. Outrossim, é natural esperar que o efeito do tratamento seja maior com o tempo, levando em conta o histórico escolar do aluno (WACHTEL, 1975; CARD E KRUEGER, 1996). Como os alunos não fariam novamente a Prova Brasil no 5º ano, os efeitos poderiam ser estimados quando alcançassem o 9º ano do Ensino Fundamental e, mesmo, o 3º ano do Ensino Médio. Contudo, como estender o efeito de incluir um ano escolar sobre a Prova Brasil do 5º ano com os

resultados dos mesmos alunos na Prova Brasil do 9º ano, ou 3º do Ensino Médio? Muito provavelmente, retornaria ao efeito de um ano escolar, mas agora nas etapas mais avançadas.

Como sugestões de avanços, as indicações acima de como superar as limitações são aparentemente razoáveis. Não obstante, os principais problemas desse estudo parecem ser os possíveis vieses embutidos na agregação em proficiência média das escolas e o significado de um ano escolar a mais no início. Tendo uma noção que a maioria das crianças fazem o Educação Infantil, o impacto seria uma realocação para a escola de um ensino sem o objetivo alfabetizador anteriormente ofertado pela Pré-Escola. Isso quer dizer, que os alunos da 4ª série, muito provavelmente, também fizeram esse um ano extra, mas na Pré-Escola e sem a obrigação de serem iniciados na alfabetização. No entanto, os resultados foram continuamente significativos e positivos, sugerindo que um ano extra na escola tem retornos médios maiores para proficiência dos alunos, tanto em Língua Portuguesa quanto em Matemática, em relação a iniciar na Pré-escola. Por outras palavras, os alunos que começaram a alfabetização com uma idade menor tiveram um desempenho futuro na escola superior, em termos de proficiência média em Língua Portuguesa e Matemática, comparado aos alunos que começaram a alfabetização um ano mais tarde, na 1ª série.

Uma implicação política do impacto de um ano escolar são novas evidências da dificuldade da educação brasileira em promover um ensino adequado. Junto ao objetivo de estender o Ensino Fundamental e começar a alfabetização mais cedo, é extremamente importante pensar em políticas que alcancem a proficiência dos alunos, seja garantindo a permanência dos alunos nas escolas ou reestruturando e ressignificando a maneira de ensinar.

REFERÊNCIAS

ARAÚJO, Fernando; SIQUEIRA, LBO de. Determinantes do desempenho escolar dos alunos da 4ª série do Ensino Fundamental no Brasil. **Economia e Desenvolvimento**, v. 9, n. 1, p. 70-103, 2010.

BARBOSA, José Juvêncio. Alfabetização e leitura. **Coleção magistério – 2º grau**; 16. São Paulo: Cortez, 1990.

BRASIL. **Brasil no Pisa 2018 [recurso eletrônico]**. Brasília : Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020.

_____. Secretaria de Educação Básica (SEB/MEC). **Ensino Fundamental de Nove Anos: Passo a Passo do Processo de Implementação**. Brasília, DF: Ministério da Educação; 2ª ed., Setembro de 2009. Disponível em:
http://portal.mec.gov.br/dmdocuments/passo_a_passo_versao_atual_16_setembro.pdf. Acesso em: 12 de maio de 2023.

_____. Secretaria de Educação Básica (SEB/MEC). **Ensino Fundamental De Nove Anos: Perguntas Mais Frequentes E Respostas Da Secretaria De Educação Básica (Seb/Mec)**. Brasília, DF: Ministério da Educação. Disponível em:
http://portal.mec.gov.br/seb/arquivos/pdf/Ensfund/ensfund9_perfreq.pdf. Acesso em: 12 de maio de 2023.

_____. Lei nº 11.274, de 6 de fevereiro de 2006. Altera a redação dos arts. 29, 30, 32 e 87 da Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, dispondo sobre a duração de 9 (nove) anos para o Ensino Fundamental, com matrícula obrigatória a partir dos 6 (seis) anos de idade. Brasília, DF: **Casa Civil**, fev., 2006.

_____. Projeto de Lei da Câmara nº 144, de 2005. Altera a redação dos arts. 29, 30, 32 e 87 da Lei nº 9.394, de 20 de dezembro de 1996, que estabelece as diretrizes e bases da educação nacional, dispondo sobre a duração de 9 (nove) anos para o Ensino Fundamental, com matrícula obrigatória a partir dos 6 (seis) anos de idade. Brasília, DF: **Senado Federal**, 2005.

_____. Lei nº 11.114, de 16 de maio de 2005. Altera os arts. 6º, 30, 32 e 87 da Lei nº 9.394, de 20 de dezembro de 1996, com o objetivo de tornar obrigatório o início do Ensino Fundamental aos seis anos de idade. Brasília, DF: **Casa Civil**, mai., 2005.

_____. PL nº 3.675, de 2004. Altera a redação dos art. 6º, art. 29, art. 30, inciso II, art. 32, caput, e art. 87, § 2º e §3º, inciso I, da Lei nº 9.394, de 20 de dezembro de 1996, que "Estabelece as Diretrizes e Bases da Educação Nacional", dispondo sobre a duração mínima de nove anos para o Ensino Fundamental, com matrícula obrigatória a partir dos seis anos de idade. Brasília, DF: **Congresso Nacional**, 2004.

_____. Lei nº 10.172, de 9 de janeiro de 2001. Aprova o Plano Nacional de Educação e dá outras providências. Brasília, DF: **Casa Civil**, mai., 2005.

_____. Portaria nº 931, de 21 de março de 2005. Resolve: Instituir o Sistema de Avaliação da Educação Básica – SAEB. Brasília, DF: **Diário Oficial da União**, seção 1, n. 55, p. 17, mar., 2005.

BOURDIEU, Pierre. À quoi sert le sociologue?. **La marmite: journal des enseignants de Saint-Amand les Eaux**, 1993.

BOURDIEU, Pierre. **La Domination**. [Documento audiovisual do Service audiovisuel de l'Université de Rouen]. Université de Rouen, 1997.

BOURDIEU, Pierre; PASSERON, Jean-Claude. L'examen d'une illusion. **Revue française de sociologie**, p. 227-253, 1968.

CALLAWAY, Brantly; SANT'ANNA, Pedro HC. Difference-in-differences with multiple time periods. **Journal of Econometrics**, v. 225, n. 2, p. 200-230, 2021.

CLEIDE, Fátima. Plano Nacional de Educação: PNE. Brasília, DF: **Senado Federal**, [2014?].

FELÍCIO, Fabiana; FERNANDES, Reynaldo. O efeito da qualidade da escola sobre o desempenho escolar: uma avaliação do Ensino Fundamental no estado de São Paulo. **Encontro Nacional de Economia**, v. 33, p. 2005, 2005.

FERRÃO, Maria Eugenia *et al.* O SAEB–Sistema Nacional de Avaliação da Educação Básica: objetivos, características e contribuições na investigação da escola eficaz. **Revista Brasileira de Estudos de População**, v. 18, n. 1/2, p. 111-130, 2001.

FRANCO, Ana Maria de Paiva. **Os determinantes da qualidade da educação no Brasil**. 2008. Tese de Doutorado. Universidade de São Paulo.

FREIRE, P. **Educação como Prática da Liberdade** (11ª ed.). Paz e Terra, 2011.

FRIEDMAN, Milton. **Capitalism and freedom**. University of Chicago press, 2020.

GRAMANI, Maria Cristina. Análise dos determinantes de eficiência educacional do estado do Ceará. **Ensaio: avaliação e políticas públicas em educação**, v. 25, n. 95, p. 507-526, 2017.

GURAJATI, D. N. **Econometria básica**. Rio de Janeiro. **Campus Elsevier**, 2006.

HAMZE, Amélia. As Legislações do Ensino Fundamental de Nove Anos. **UOL**, São Paulo. Disponível em: <https://educador.brasilecola.uol.com.br/politica-educacional/ensino-fundamental-de-nove-anos.htm>. Acesso em: 16 de março de 2023.

HANUSHEK, Eric A. The failure of input-based schooling policies. **The economic journal**, v. 113, n. 485, p. F64-F98, 2003.

INEP. **Ensino Fundamental Regular - Anos Iniciais (Atualizado em 18/09/2020)**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020. Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/ideb/resultados>. Acesso em: 26 de abril de 2023.

_____. **Microdados do Censo Escolar da Educação Básica 2007**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/cento-escolar>. Acesso em: 17 de fevereiro de 2022.

_____. **Microdados do Censo Escolar da Educação Básica 2009**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/cento-escolar>. Acesso em: 17 de fevereiro de 2022.

_____. **Microdados do Censo Escolar da Educação Básica 2011**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/cento-escolar>. Acesso em: 17 de fevereiro de 2022.

_____. **Microdados do Saeb 2005**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 17 de fevereiro de 2022.

_____. **Microdados do Saeb 2011**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Disponível em: <https://www.gov.br/inep/pt-br/acao-a-informacao/dados-abertos/microdados/saeb>. Acesso em: 17 de fevereiro de 2022.

_____. **Escalas de proficiência do SAEB**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020.

_____. **Sistema de Avaliação da Educação Básica – Documentos de Referência – Versão 1.0**. Brasília, DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2019.

LOURENÇO, Rosenery Loureiro et al. Determinantes sociais e pedagógicos das notas do IDEB. **Revista Pensamento Contemporâneo em Administração**, v. 11, n. 4, p. 27-43, 2017.

MARTINELLI, Selma de Cássia; GENARI, Carla Helena Manzini. Relações entre desempenho escolar e orientações motivacionais. **Estudos de Psicologia (Natal)**, v. 14, p. 13-21, 2009.

PEIXOTO, Betânia et al. **Avaliação econômica de projetos sociais**. Fundação Itaú Social, 2012.

PISTRAK, Moisey M. **A Escola-comuna**. 1 ed. São Paulo: Expressão Popular, 2009

PRITCHETT, Lant; FILMER, Deon. What education production functions really show: a positive theory of education expenditures. **Economics of Education review**, v. 18, n. 2, p. 223-239, 1999.

RIBEIRO, Sérgio Costa. A pedagogia da repetência. **Estudos avançados**, v. 5, p. 07-21, 1991.

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

ROSENBAUM, Paulo R.; RUBIN, Donald B.; O papel central do escore de propensão em estudos observacionais para efeitos causais. **Biometrika**, v. 70, n. 1, p. 41-55, 1983.

ROJO, Roxane. **Letramentos múltiplos, escola e inclusão social**. São Paulo: Parábola Editorial, 2009.

RUBIN, Donald B. Using propensity scores to help design observational studies: application to the tobacco litigation. **Health Services and Outcomes Research Methodology**, v. 2, p. 169-188, 2001.

SANTOS, Marcelo Justus dos; KASSOUF, Ana Lúcia. Uma investigação dos determinantes socioeconômicos da depressão mental no Brasil com ênfase nos efeitos da educação. **Economia aplicada**, v. 11, p. 5-26, 2007.

SASSAKI, Alex Hayato et al. Por Que o Brasil Vai Mal no PISA? uma análise dos determinantes do desempenho no exame. **Inspir. PolicyPaper**, n. 31, 2018.

SOARES, Sergei; SÁTYRO, Natália. **O impacto de infra-estrutura escolar na taxa de distorção idade-série das escolas brasileiras de Ensino Fundamental: 1998 a 2005**. Texto para Discussão, 2008.

STOCK, James H.; WATSON, Mark W. **Econometria**. Boston: Addison Wesley, 2003.

TODOS PELA EDUCAÇÃO. **Anuário Brasileiro da Educação Básica: 2019**. São Paulo: Moderna, 2019.

VALLE, Ione Ribeiro; SOULIÉ, Charles. Pierre Bourdieu: uma sociologia ambiciosa da educação. **Florianópolis: UFSC**, 2019.

VARANI, Adriana; SILVA, Daiana Cristina. A relação família-escola: implicações no desempenho escolar dos alunos dos anos iniciais do Ensino Fundamental. **Revista Brasileira de Estudos Pedagógicos**, v. 91, n. 229, p. 511-527, 2010.

WACHTEL, Paul. The returns to investment in higher education: Another view. In: **Education, income, and human behavior**. NBER, 1975. p. 151-170.

WOOLDRIDGE, Jeffrey M. **Introdução à econometria: uma abordagem moderna**. Pioneira Thomson Learning, 2006.

APÊNDICE A – PROFICIÊNCIA E QUESTIONÁRIO

A seguir, as tabelas 12 e 13 apresentam os resultados das regressões da proficiência média das escolas nas duas disciplinas contra a proporção de alunos removidos após a imposição dos critérios das amostras, para 2009 e 2011, respectivamente. Quando a proporção de alunos removidos aumenta um ponto percentual, a proficiência em Língua Portuguesa e Matemática diminui, respectivamente, entre 70 e 84 pontos e entre 81 e 99 pontos.

Tabela 12 - Proficiência média das escolas em Língua Portuguesa e Matemática na 4ª série do Ensino Fundamental, em 2009 - Amostra de alunos (38.339 escolas e 1.331.547 matrículas tanto tratados quanto controle).

Variáveis	Modelo 1		Modelo 2	
	Língua Portuguesa	Matemática	Língua Portuguesa	Matemática
Constante	221,96*** (0,4260)	249,24*** (0,5080)	208,04*** (0,4773)	231,41*** (0,5656)
PROP ALUNOS_REMOV	-77,82 *** (0,7688)	-91,3*** (0,9169)	-70,23*** (0,7506)	-81,58*** (0,8894)
TP_COR_ALUNO			27,94*** (0,4931)	35,76*** (0,5843)
R-quadrado Ajustado	0,2109	0,2055	0,2718	0,2762

NOTA: Significativo com um tamanho de teste de: 0,1% (***) ; 1% (**); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Tabela 13 - Proficiência média das escolas em Língua Portuguesa e Matemática na 4ª série (controle) e 5º ano (tratados) do Ensino Fundamental, em 2011 - Amostra de alunos (36.104 escolas e 1.229.857 matrículas).

Variáveis	Modelo 1		Modelo 2	
	Língua Portuguesa	Matemática	Língua Portuguesa	Matemática
Constante	229,02*** (0,3930)	255,69*** (0,4618)	212,31 *** (0,4489)	234,83*** (0,5235)
PROP_ALUNOS_REMOV	-84,14*** (0,7546)	-99,48*** (0,8869)	-73*** (0,7322)	-85,56*** (0,8540)
TP_COR_ALUNO			32,76*** (0,4951)	40,89*** (0,5774)
R-quadrado Ajustado	0,2562	0,2584	0,3366	0,3489

NOTA: Significativo com um tamanho de teste de: 0,1% (***) ; 1% (**); 5%(*); e 10%(.).

Fonte: elaborado pelo autor.

Foi incluído a variável TP_COR_ALUNO como *proxy* de renda na segunda metade das tabelas, consistindo, segundo o dicionário no Apêndice D, na proporção de alunos brancos na escola. A variação do coeficiente da proporção de alunos removidos, após a sua introdução, fortalece a suposição que a proporção de alunos removidos por não preencherem o questionário está associado a uma proficiência e condições socioeconômicas menores, tornando vantajoso o seu controle.

APÊNDICE B – TENDÊNCIA PARALELA

As tabelas 14 e 15 apresentam o comportamento médio da proficiência para o subconjunto de escolas tratadas e de controle, respectivamente, em Língua Portuguesa e Matemática, para as edições da Prova Brasil de 2005, 2007, 2009 e 2011. Mesmo a amostra sem controle ou pareamento parece seguir uma razoável tendência paralela para as duas proficiências, o que melhora nas amostras com controle e com pareamento.

Tabela 14 - Proficiência média do grupo de tratamento e de controle em Língua Portuguesa na Prova Brasil de 2005, 2007, 2009 e 2011.

Ano	Painel sem controle ou pareamento		Painel com controle		Painel com controle e pareamento	
	Grupo de Tratamento	Grupo de Controle	Grupo de Tratamento	Grupo de Controle	Grupo de Tratamento	Grupo de Controle
2005	173,29	168,88	176,10	174,38	176,81	173,67
2007	173,16	170,70	176,64	176,41	177,62	175,49
2009	181,16	175,13	187,06	185,74	186,37	184,10
2011	188,92	178,00	194,19	188,28	193,46	186,59

Fonte: elaborado pelo autor. Dados do IDEB.

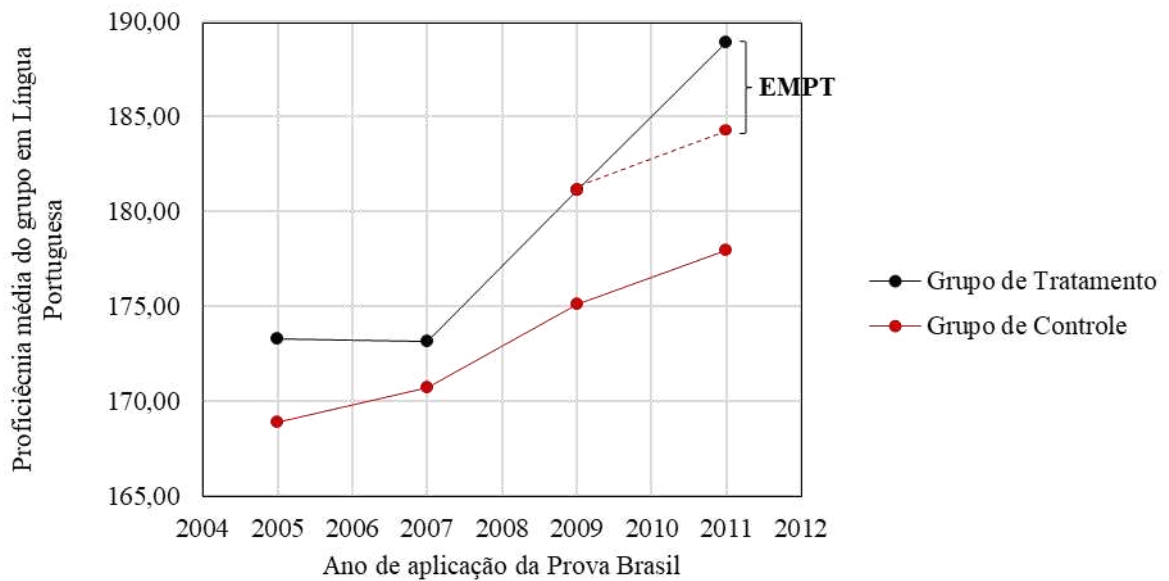
Tabela 15 - Proficiência média do grupo de tratamento e de controle em Matemática na Prova Brasil de 2005, 2007, 2009 e 2011.

Ano	Painel sem controle ou pareamento		Painel com controle		Painel com controle e pareamento	
	Grupo de Tratamento	Grupo de Controle	Grupo de Tratamento	Grupo de Controle	Grupo de Tratamento	Grupo de Controle
2005	180,74	176,91	183,37	182,63	183,88	182,09
2007	190,66	188,05	194,61	194,21	195,45	193,31
2009	200,90	195,30	207,98	208,13	207,20	205,89
2011	207,92	196,65	214,21	209,25	213,55	207,28

Fonte: elaborado pelo autor. Dados do IDEB.

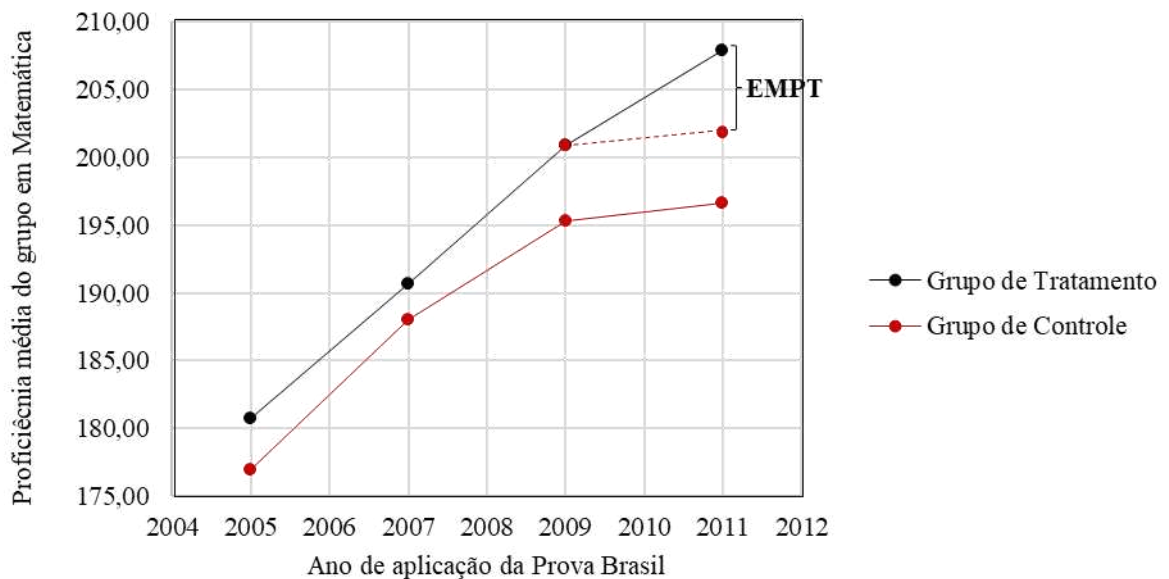
Os passos do DiD, a primeira a e segunda diferença, podem ser vistos nos gráficos a seguir – figuras de 2 a 7. A primeira diferença na mudança na proficiência média entre 2009 e 2011 dos subconjuntos de escolas, e a segunda pela distância vertical entra a curva dos tratados e dos controles, em 2011, (depois da translação da curva para o mesmo ponto em 2009), que é o resultado do DiD. Por outras palavras, a primeira diferença remove a diferença de escala inicial das notas (2009) e as colocam no mesmo ponto de partida, enquanto a segunda diferença estima a diferença de variação entre os grupos. Sem a dupla diferença haveria uma superestimação do efeito do tratamento, seja pelo método “antes e depois” ou pelo método de diferença simples entre os grupos em 2011.

Figura 2 – Tendência da proficiência média em Língua Portuguesa do grupo de tratamento e do grupo de controle em 2005, 2007, 2009 e 2011 – Painel sem controle.



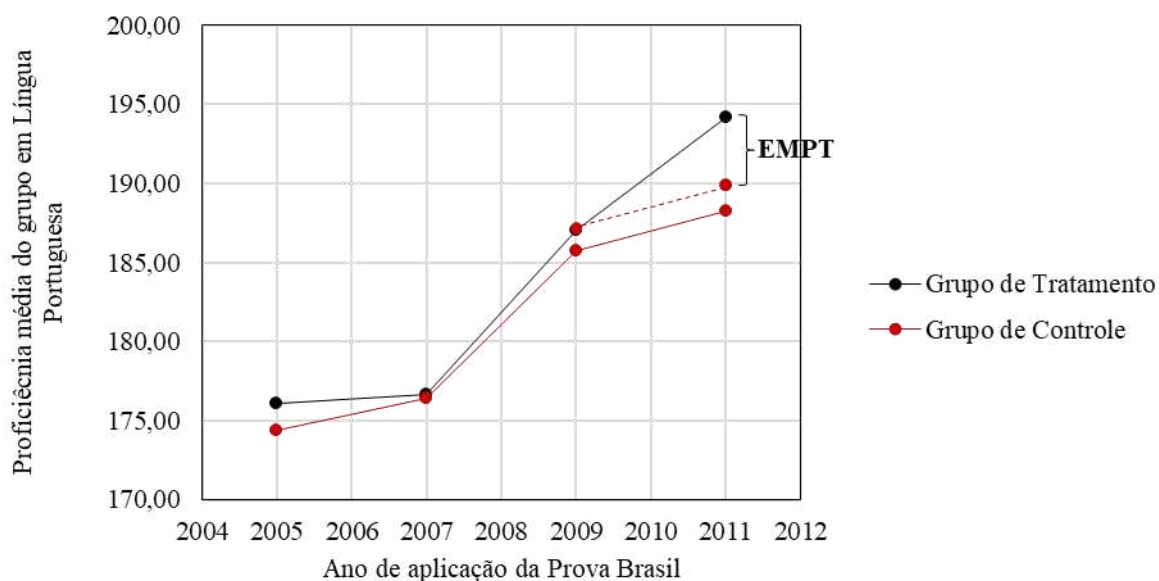
Fonte: elaborado pelo autor. Dados do IDEB.

Figura 3 – Tendência da proficiência média em Matemática do grupo de tratamento e do grupo de controle em 2005, 2007, 2009 e 2011 – Painel sem controle.



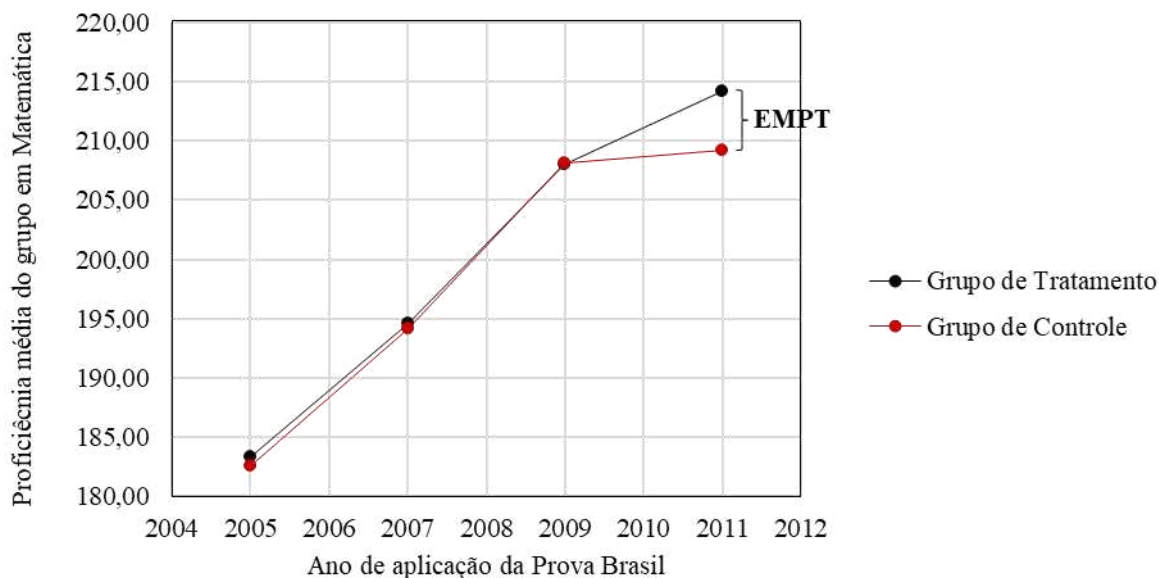
Fonte: elaborado pelo autor. Dados do IDEB.

Figura 4 – Tendência da proficiência média em Língua Portuguesa do grupo de tratamento e do grupo de controle em 2005, 2007, 2009 e 2011 – Painel com controle.



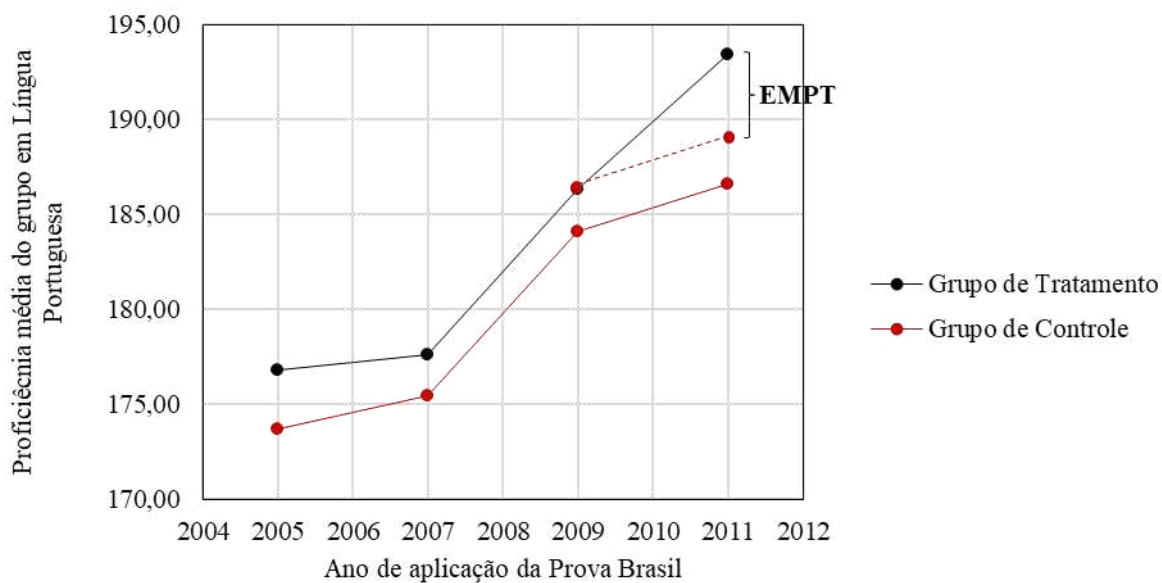
Fonte: elaborado pelo autor. Dados do IDEB.

Figura 5 – Tendência da proficiência média em Matemática do grupo de tratamento e do grupo de controle em 2005, 2007, 2009 e 2011 – Painel com controle.



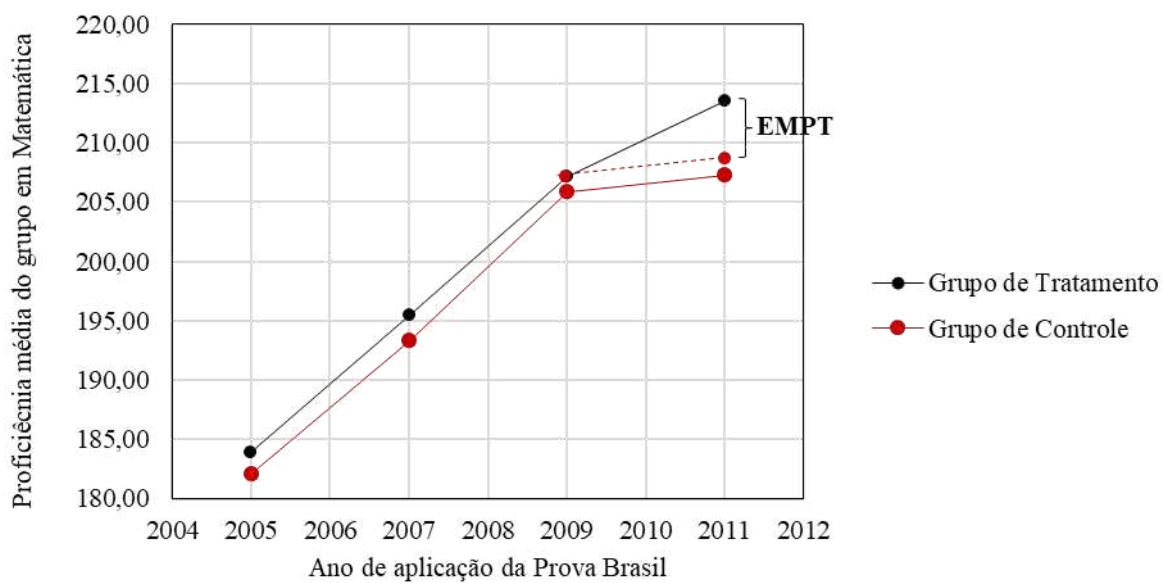
Fonte: elaborado pelo autor. Dados do IDEB.

Figura 6 – Tendência da proficiência média em Língua Portuguesa do grupo de tratamento e do grupo de controle em 2005, 2007, 2009 e 2011 – Painel com controle e pareamento.



Fonte: elaborado pelo autor. Dados do IDEB.

Figura 7 – Tendência da proficiência média em Matemática do grupo de tratamento e do grupo de controle em 2005, 2007, 2009 e 2011 – Painel com controle e pareamento.



Fonte: elaborado pelo autor. Dados do IDEB.

APÊNDICE C – MATRIZ DE CORRELAÇÃO DAS VARIÁVIES

A correlação entre as variáveis de controle de aluno, professor, diretor e da escola em relação a proficiência em Língua Portuguesa e Matemática foram calculadas apenas para a amostra de controle. São, respectivamente, as Tabelas 16, 17, 18 e 19. Devido ao número de variáveis incluídas, omitiu-se a apresentação da correlação para as outras amostras. O estudo se tornaria desnecessariamente extenso e com poucas informações adicionais – se espera que as relações se mantenham nas outras amostras.

Tomando o dicionário do Apêndice D como suporte, para identificar o significado das variáveis, todas apresentaram uma relação esperada com a proficiência de acordo com a literatura, ou com a própria intuição de como deveriam dar-se. O p-valor foi adicionado como complemento para observar a significância estatística das estimativas, mas mesmo quando não significativas as variáveis foram mantidas. Supera a correlação significativa na amostra desse estudo, os diversos estudos que sustentam a importância delas – direta ou indiretamente. Mesmo não apresentando um comportamento comum com a proficiência, até onde se pode dizer que não corrigem uma associação sistemática com uma ou mais das outras variáveis especificadas?

Tabela 16 - Correlação entre a proficiência das escolas em Língua Portuguesa e Matemática e as variáveis de controle do aluno - Painel com controle, 2009 e 2011.

Variáveis	Língua Portuguesa		p-valor		Matemática		p-valor	
	2009	2011	2009	2011	2009	2011	2009	2011
NU_MEDIA_PORT	1	1			0,94	0,93	0	0
NU_MEDIA_MAT	0,94	0,93	0	0	1	1		
QT_MATR	0,28	0,24	0	0	0,27	0,23	0	0
TP_SEXO_ALUNO	-0,03	-0,03	0,0455	0,0422	0,00	0,02	0,810	0,201
TP_COR_ALUNO	0,37	0,39	0	0	0,39	0,42	0	0
TP_IDADE_ALUNO_A	-0,12	-0,09	2,22E-15	1,06E-09	-0,13	-0,09	4,44E-16	8,29E-09
TP_IDADE_ALUNO_B	-0,08	-0,06	2,29E-07	8,44E-05	-0,08	-0,03	2,29E-07	0,043
TP_IDADE_ALUNO_C	0,36	0,26	0	0	0,38	0,28	0	0
TP_IDADE_ALUNO_D	0,02	0,05	0,117	0,001	-0,01	0,01	0,706	0,691
TP_MORA_MAE	0,23	0,20	0	0	0,25	0,22	0	0
TP_ESC_MAE_A	-0,42	-0,38	0	0	-0,37	-0,33	0	0
TP_ESC_MAE_B	-0,14	-0,15	0	0	-0,08	-0,11	1,27E-06	5,30E-12
TP_ESC_MAE_C	0,05	-0,01	0,003	0,435	0,05	0,00	0,003	0,820
TP_ESC_MAE_D	0,36	0,31	0	0	0,32	0,27	0	0
TP_ESC_MAE_E	0,31	0,25	0	0	0,27	0,21	0	0
TP_MORA_PAÍ	0,23	0,21	0	0	0,25	0,24	0	0
TP_ESC_PAÍ_A	-0,40	-0,39	0	0	-0,36	-0,35	0	0
TP_ESC_PAÍ_B	-0,07	-0,01	7,16E-06	0,645	-0,01	0,04	0,589	0,014
TP_ESC_PAÍ_C	0,11	-0,02	8,23E-12	0,130	0,11	-0,01	2,12E-13	0,692853977
TP_ESC_PAÍ_D	0,35	0,33	0	0	0,32	0,29	0	0
TP_ESC_PAÍ_E	0,29	0,19	0	0	0,24	0,15	0	0
TP_REUNIAO	0,32	0,31	0	0	0,32	0,31	0	0

Variáveis	Língua Portuguesa		p-valor		Matemática		p-valor	
	2009	2011	2009	2011	2009	2011	2009	2011
	TP_INC_EST	0,30	0,25	0	0	0,29	0,25	0
TP_INC_DEVER	0,31	0,32	0	0	0,29	0,31	0	0
TP_INC_LER	0,15	0,11	0	1,69E-12	0,14	0,11	0	5,63E-12
TP_INC_NAO_FALT	0,30	0,29	0	0	0,30	0,29	0	0
TP_CONV_ESC	0,07	0,06	3,77E-06	3,82E-05	0,06	0,07	0,000	8,04E-06
TP_TRAB_FOR	-0,43	-0,41	0	0	-0,38	-0,37	0	0
TP_REPROV	-0,48	-0,47	0	0	-0,46	-0,45	0	0
PROP_ALUNOS_REMOV	-0,54	-0,56	0	0	-0,53	-0,57	0	0

Fonte: elaborado pelo autor.

Tabela 17 - Correlação entre a proficiência das escolas em Língua Portuguesa e Matemática e as variáveis de controle do professor - Painel com controle, 2009 e 2011.

Variáveis	Língua Portuguesa		p-valor		Matemática		p-valor	
	2009	2011	2009	2011	2009	2011	2009	2011
	NU_MEDIA_PORT	1	1			0,94	0,93	0
NU_MEDIA_MAT	0,94	0,93	0	0	1	1		
QT_PROF	0,09	0,05	1,54E-09	0,002	0,10	0,04	2,46E-10	0,004
TP_SEXO_PROF	-0,16	-0,15	0	0	-0,14	-0,15	0	0
TP_IDADE_PROF_A	-0,05	-0,03	0,001	0,032	-0,06	-0,04	0,000	0,024
TP_IDADE_PROF_B	-0,04	-0,06	0,010	4,67E-05	-0,04	-0,06	0,010	0,000
TP_COR_PROF	0,29	0,28	0	0	0,32	0,32	0	0
TP_ESC_PROF	0,12	0,07	2,53E-14	3,89E-06	0,12	0,07	6,22E-15	3,02E-06
TP_ESC_SUP_PROF_A	0,11	0,07	1,11E-12	1,52E-06	0,11	0,07	9,11E-13	5,63E-06
TP_ESC_SUP_PROF_B	0,02	0,04	0,288	0,018	0,01	0,04	0,410	0,009
TP_ESC_SUP_PROF_C	0,01	0,00	0,624	0,754	0,01	0,00	0,644	0,793
TP_FCONT_PROF	0,04	0,05	0,014	0,001	0,03	0,05	0,026	0,001
TP_SALB_TOTAL_PROF_A	-0,08	-0,05	1,37E-07	0,002	-0,07	-0,05	3,06E-06	0,001
TP_SALB_TOTAL_PROF_B	-0,14	-0,06	0	3,45E-05	-0,14	-0,05	0	0,002
TP_SALB_TOTAL_PROF_C	0,06	-0,02	9,37E-05	0,314	0,06	-0,01	5,27E-05	0,338
TP_SALB_TOTAL_PROF_D	0,09	0,09	1,19E-09	1,41E-09	0,09	0,08	1,39E-08	1,73E-07
TP_TEMP_LEC_TOTAL_A	-0,04	-0,01	0,004	0,611	-0,05	-0,01	0,003	0,370
TP_TEMP_LEC_TOTAL_B	-0,05	-0,05	0,001	0,003	-0,05	-0,04	0,000	0,008
TP_NUM_ESC_TRAB_PROF_A	-0,06	-0,03	0,000	0,032	-0,05	-0,02	0,001	0,172
TP_NUM_ESC_TRAB_PROF_B	0,06	0,03	9,36E-05	0,047	0,06	0,02	0,000	0,220
TP_CARGH_TOTAL_PROF_A	-0,08	-0,04	2,27E-07	0,013	-0,08	-0,04	1,37E-07	0,014
TP_CARGH_TOTAL_PROF_B	0,05	0,01	0,001	0,366	0,06	0,03	3,37E-05	0,103
TP_PROP_LEC_B	-0,22	-0,21	0	0	-0,22	-0,22	0	0
TP_PROP_LEC_C	-0,13	-0,23	0	0	-0,13	-0,23	0	0
TP_PROP_LEC_D	0,33	0,36	0	0	0,35	0,37	0	0
TP_CAR_INFRA	-0,20	-0,18	0	0	-0,21	-0,19	0	0
TP_CURR_INAD	-0,07	-0,09	8,23E-06	7,45E-09	-0,08	-0,10	3,83E-07	2,99E-10
TP_CONF_DIRETOR_A	0,10	0,08	1,42E-11	7,91E-07	0,11	0,09	5,07E-13	1,26E-08
TP_CONF_DIRETOR_B	-0,08	-0,04	1,09E-06	0,005	-0,08	-0,05	8,03E-07	0,001
TP_AMB_ESC_A	0,08	0,09	1,28E-07	4,32E-09	0,09	0,09	4,58E-09	9,84E-10
TP_AMB_ESC_B	-0,04	-0,09	0,005	1,25E-08	-0,05	-0,09	0,001	6,75E-09
TP_INSUF_FIN_A	0,14	0,13	0	2,22E-16	0,15	0,15	0	0
TP_INSUF_FIN_B	-0,07	-0,08	1,87E-06	9,02E-07	-0,08	-0,09	1,52E-07	5,01E-09
TP_FALT_PROF_A	0,01	0,03	0,649	0,101	0,01	0,03	0,626	0,034
TP_FALT_PROF_B	-0,01	-0,03	0,525	0,081	-0,01	-0,03	0,597	0,034
TP_FALT_ALUNOS_A	0,14	0,13	0	4,44E-16	0,14	0,13	0	0
TP_FALT_ALUNOS_B	-0,11	-0,08	2,21E-12	1,50E-07	-0,11	-0,09	3,02E-12	1,44E-09

Fonte: elaborado pelo autor.

Tabela 18 - Correlação entre a proficiência das escolas em Língua Portuguesa e Matemática e as variáveis de controle de diretor - Painel com controle, 2009 e 2011.

Variáveis	Língua Portuguesa		p-valor		Matemática		p-valor	
	2009	2011	2009	2011	2009	2011	2009	2011
	NU_MEDIA_PORT	1	1			0,94	0,93	0
NU_MEDIA_MAT	0,94	0,93	0	0	1	1		
TP_ESC_DIR	0,08	0,01	5,68E-08	0,611	0,09	0,00	3,60E-08	0,996
TP_ESC_SUP_DIR_A	0,08	0,07	6,49E-08	2,11E-06	0,09	0,07	2,06E-08	1,54E-06
TP_ESC_SUP_DIR_B	0,03	0,02	0,077	0,231	0,01	0,01	0,479	0,645
TP_ESC_SUP_DIR_C	0,02	0,03	0,170	0,080	0,03	0,02	0,064	0,113
TP_FCONT_DIR	0,04	0,04	0,022	0,005	0,03	0,05	0,030	0,002
TP_SALB_DIR_A	0,00	-0,03	0,811	0,094	0,00	-0,02	0,824	0,209
TP_SALB_DIR_B	-0,13	-0,12	0	4,22E-15	-0,13	-0,12	0	3,77E-14
TP_SALB_DIR_C	-0,12	-0,13	1,15E-14	0	-0,10	-0,11	3,81E-11	4,03E-13
TP_SALB_DIR_D	0,15	0,14	0	0	0,14	0,12	0	7,99E-15
TP_TEMP_FUNC_DIR_A	-0,10	-0,09	9,01E-11	1,88E-08	-0,10	-0,08	3,28E-10	5,96E-08
TP_TEMP_FUNC_DIR_B	0,05	0,02	0,002	0,290	0,05	0,01	0,001	0,388
TP_ASS_DIR_A	0,15	0,15	0	0	0,14	0,14	0	0
TP_ASS_DIR_B	-0,20	-0,19	0	0	-0,20	-0,18	0	0
TP_PPD_DR_A	-0,04	-0,01	0,018	0,667	-0,04	-0,01	0,006	0,688
TP_PPD_DR_B	0,00	-0,01	0,838	0,591	0,00	0,00	0,837	0,830
TP_PPD_DR_C	0,04	0,03	0,015	0,088	0,03	0,03	0,029	0,065
TP_PPD_DR_D	0,04	-0,01	0,022	0,634	0,04	-0,01	0,004	0,404
TP_VINC_EST_A	-0,09	-0,07	1,97E-09	6,25E-06	-0,08	-0,06	0,000	0,000
TP_VINC_EST_B	-0,05	-0,06	0,002	0,000	-0,05	-0,06	0,003	0,000
TP_VINC_EST_C	0,02	-0,01	0,146	0,493	0,03	0,00	0,042	0,830
TP_VINC_EST_D	-0,01	0,01	0,713	0,673	-0,01	0,01	0,630	0,686
TP_ROT_CORP_DOC_A	0,05	0,03	0,002	0,025	0,05	0,05	0,001	0,001
TP_ROT_CORP_DOC_B	-0,03	-0,02	0,024	0,298	-0,04	-0,03	0,018	0,032

Fonte: elaborado pelo autor.

Tabela 19 - Correlação entre a proficiência das escolas em Língua Portuguesa e Matemática e as variáveis de controle da escola - Painel com controle, 2009 e 2011.

Variáveis	Língua Portuguesa		P-valor		Matemática		P-valor	
	2009	2011	2009	2011	2009	2011	2009	2011
	MEDIA_LP	1	1			0,94	0,93	0
MEDIA_MT	0,94	0,93	0	0	1	1		
IN_ESGOTO_INEXISTENTE	-0,02	-0,01	0,115	0,41	-0,03	-0,01	0,041	0,53
IN_LABORATORIO_INFORMATICA	0,14	0,13	0	0	0,14	0,12	0	9,55E-14
IN_LABORATORIO_CIENCIAS	0,05	0,05	0,001	0,003	0,05	0,04	0,001	0,014
IN_QUADRA_ESPORTES	0,19	0,21	0	0	0,21	0,22	0	0
IN_BIBLIOTECA	0,08	0,11	3,88E-07	4,35E-13	0,06	0,08	0,0002	9,09E-08
IN_COMPUTADOR	0,17	0,11	0	6,19E-13	0,17	0,10	0	2,79E-10
IN_INTERNET	0,29	0,24	0	0	0,30	0,23	0	0
QT_SALAS_EXISTENTES	0,12	0,10	4,71E-14	2,99E-11	0,12	0,09	5,77E-15	7,18E-09
RAZAO_MAT_DOC	-0,20	-0,13	0	2,22E-16	-0,24	-0,16	0	0

Fonte: elaborado pelo autor.

APÊNDICE D - DICIONÁRIO DE VARIÁVEIS

O dicionário a seguir vai ser usado como suporte para identificação do significado das variáveis nas tabelas no corpo e nos apêndices deste estudo. O código de preenchimento igual a 1 indica que a categoria foi especificidade na regressão de função de produção educacional e 0 que é a categoria de referência (omitida). Em geral, se manteve a descrição original das variáveis do Censo Escolar e das perguntas dos questionários da Prova Brasil, por isso se encontram entres aspas.

Quadro 1 - Dicionário de variáveis selecionadas e incluídas no estudo.

Descrição	Variável	Tipo ¹	Ind. ²	Código de Preenchimento ³
Ano	ANO	Num		
Código da Unidade da Federação	CO_UF	Char		
--- Estimativa ---				
Tempo do Tratamento	IN_TEMPO / Tempo	Num		1 - 2011 0 - 2009
Tratamento (Novo Regime)	IN_TRAT / Tratamento	Num		1 - Escola Tratada 0 - Escola de Controle
Coefficiente de Diferenças-em-Diferenças	IN_EST_DD / β (EMPT)	Num		
--- Escola ---				
Código da Escola	CO_ENTIDADE	Num		
“Dependência Administrativa”	TP_DEPENDENCIA	Num		1 - Federal 2 - Estadual 3 - Municipal 4 - Privada
“Localização”	TP_LOCALIZACAO	Num		1 - Urbana 2 - Rural
“Esgoto sanitário – Inexistente”	IN_ESGOTO_INEXISTENTE	Num		0 - Não 1 - Sim
Dependências existentes na escola - Laboratório de informática	IN_LABORATORIO_INFORMATICA	Num		0 - Não 1 - Sim
Dependências existentes na escola - Laboratório de ciências	IN_LABORATORIO_Ciencias	Num		0 - Não 1 - Sim
“Dependências existentes na escola - Quadra de esportes coberta ou descoberta”	IN_QUADRA_ESPORTES	Num		0 - Não 1 - Sim
“Dependências existentes na escola – Biblioteca”	IN_BIBLIOTECA	Num		0 - Não 1 - Sim
“Equipamentos existentes na escola – Computador”	IN_COMPUTADOR	Num		0 - Não 1 - Sim
“Acesso à Internet”	IN_INTERNET	Num		0 - Não 1 - Sim
Razão Aluno-Professor (total de matrículas / total de docentes da educação básica)	RAZAO_MAT_DOC	Num		
Quantidade de salas existentes na escola	QT_SALAS_EXISTENTES	Num		
Proporção de alunos removidos por não responderem todas as questões	PROP_ALUNOS_REMOV	Num		
--- Aluno ---				
Código do Aluno na Prova Brasil	ID_ALUNO	Num		
“Sexo”	TP_SEXO_ALUNO	Num		1 - Homem

Descrição	Variável	Tipo ¹	Ind. ²	Código de Preenchimento ³
				0 - Mulher
“Como você se considera?”	TP_COR_ALUNO	Num		1 - Branco 0 - Preto/Pardo/Amarelo/Indígena
“Qual a sua idade?”	TP_IDADE_ALUNO	Num	_A _B _C _D	1 - 8 anos ou menos 1 - 9 anos 1 - 10 anos 1 - 11 anos 0 - 12 anos ou mais
“Você mora com sua mãe?”	TP_MORA_MAE	Num		1 - Sim 0 - Não
“Até que série sua mãe ou a mulher responsável por você estudou?”	TP_ESC_MAE	Num	_A _B _C _D _E	1 - Nunca estudou/Não completou a 4ª série 1 - Completou a 4ª série 1 - Completou a 8ª série 1 - Completou o Ensino Médio 1 - Completou a faculdade 0 - O aluno não sabia
“Você mora com seu pai?”	TP_MORA_PAI	Num		1 - Sim 0 - Não
“Até que série seu pai ou o homem responsável por você estudou?”	TP_ESC_PAI	Num	_A _B _C _D _E	1 - Nunca estudou/Não completou a 4ª série 1 - Completou a 4ª série 1 - Completou a 8ª série 1 - Completou o Ensino Médio 1 - Completou a faculdade 0 - O aluno não sabia
“Com que frequência seus pais ou responsáveis vão à reunião de pais?”	TP_REUNIAO	Num		1 - Sempre ou quase sempre. 0 - De vez enquanto/Nunca ou quase nunca.
“Seus pais ou responsáveis incentivam você a estudar?”	TP_INC_EST	Num		1 - Sim 0 - Não
“Seus pais ou responsáveis incentivam você a fazer o dever de casa e os trabalhos da escola?”	TP_INC_DEVER	Num		1 - Sim 0 - Não
“Seus pais ou responsáveis incentivam você a ler?”	TP_INC_LER	Num		1 - Sim 0 - Não
“Seus pais ou responsáveis incentivam você a ir a escola e não faltar às aulas?”	TP_INC_NAO_FALT	Num		1 - Sim 0 - Não
“Seus pais ou responsáveis conversam com você sobre o que acontece na escola?”	TP_CONV_ESC	Num		1 - Sim 0 - Não
“Você trabalha fora de casa?”	TP_TRAB_FORA	Num		1 - Sim 0 - Não
“Você já foi reprovado?”	TP_REPROV	Num		0 - Não 1 - Sim, uma vez ou mais.
--- Professor ---				
“Sexo”	TP_SEXO_PROF	Num		1 - Homem 0 - Mulher
“Idade”	TP_IDADE_PROF	Num	_A _B	1 - Até 29 anos 1 - De 30 a menos de 40 anos 0 - 40 anos ou mais
“Como você se considera?”	TP_COR_PROF	Num		1 - Branco 0 - Preto/Pardo/Amarelo/Indígena

Descrição	Variável	Tipo ¹	Ind. ²	Código de Preenchimento ³
“Qual o seu nível de escolaridade (Até a Graduação).”	TP_ESC_PROF	Num		0 - Até Ensino Médio 1 - Ensino Superior
“Indique a modalidade de cursos de pós-graduação de mais alta titulação que você possui.”	TP_ESC_SUP_PROF	Num	_A _B _C	1 - Atualização/Especialização 1 - Mestrado 1 - Doutorado 0 - Não fez/não completou
“Você participou de alguma atividade de formação continuada (atualização, treinamento, capacitação etc.) nos últimos dois anos?”	TP_FCONT_PROF	Num		1 - Sim 0 - Não
“Qual é, aproximadamente, seu salário bruto (com adicionais, se houver) como professor(a)? (soma de tudo o que você ganha como professor(a))”	TP_SALB_TOTAL_PROF	Num	_A _B _C _D	1 - Até 1 salário mínimo de 2009/2011 1 - De mais de 1 a 2 salários mínimos 1 - De mais de 2 a 5 salários mínimos 1 - De mais de 5 a 10 salários mínimos 0 - Mais de 10 salários mínimos
“Há quantos anos você leciona?”	TP_TEMPO_LEC_TOTAL	Num	_A _B	1 - Menos de 1 ano 1 - Entre 1 e menos de 10 anos 0 - 10 anos ou mais
“Em quantas escolas você trabalha?”	TP_NUM_ESC_TRAB_PROF	Num	_A _B	1 - Apenas nesta escola 1 - Em 2 ou 3 escolas 0 - Em 4 ou mais escolas
“Ao todo, quantas horas-aula você ministra por semana?”	TP_CARGH_TOTAL_PROF	Num	_A _B	1 - Até 24 horas-aula 1 - De 25 a 30 horas-aula 0 - Mais de 30 horas-aula
“Quanto dos conteúdos previstos você conseguiu desenvolver com os alunos desta turma, neste ano?”	TP_PROP_LEC	Num	_B _C _D	0 - Menos de 40% 1 - Entre 40% e 60% 1 - Entre 60% e 80% 1 - Mais de 80%
“(…) possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): ocorrem na escola devido à carência de infra-estrutura física e/ou pedagógica.”	TP_CAR_INFRA	Num		1 - Concordo 0 - Discordo
“(…) aos possíveis problemas de aprendizagem dos alunos da(s) série(s) avaliada(s): estão relacionados aos conteúdos curriculares, que são inadequados às necessidades dos alunos.”	TP_CURR_INAD	Num		1 - Concordo 0 - Discordo
“(…) tenho plena confiança no(a) diretor(a) como profissional.”	TP_CONF_DIRETOR	Num	_A _B	1 - Concordo totalmente e concordo 1 - Neutro 0 - Discordo ou discordo totalmente
“(…) o diretor, os professores e os demais membros da equipe da escola colaboram para fazer esta escola funcionar bem.”	TP_AMB_ESC	Num	_A _B	1 - Concordo totalmente e concordo 1 - Neutro 0 - Discordo ou discordo totalmente
“(…) Ocorreu na escola: insuficiência de recursos financeiros?”	TP_INSUF_FIN	Num	_A _B	1 - Não 1 - Sim, mas não foi grave

Descrição	Variável	Tipo ¹	Ind. ²	Código de Preenchimento ³
				0 - Sim, e foi um problema grave
“(…) Ocorreu na escola: alto índice de faltas por parte dos professores?”	TP_FALT_PROF	Num	_A _B	1 - Não 1 - Sim, mas não foi grave 0 - Sim, e foi um problema grave
“(…) Ocorreu na escola: alto índice de faltas por parte dos alunos?”	TP_FALT_ALUNOS	Num	_A _B	1 - Não 1 - Sim, mas não foi grave 0 - Sim, e foi um problema grave
--- Diretor ---				
“Qual o seu nível de escolaridade (até a graduação).”	TP_ESC_DIR	Num		0 - Até Ensino Médio 1 - Ensino Superior
“Indique a modalidade de cursos de pós-graduação de mais alta titulação que você possui.”	TP_ESC_SUP_DIR	Num	_A _B _C	1 - Atualização/Especialização 1 - Mestrado 1 - Doutorado 0 - Não fez/não completou
“Você participou de alguma atividade de formação continuada (atualização, treinamento, capacitação nos últimos dois anos?etc.)”	TP_FCONT_DIR	Num		1 - Sim 0 - Não
“Qual é aproximadamente o seu salário bruto como diretor(a), com adicionais, se houver?”	TP_SALB_DIR	Num	_A _B _C _D	1 - Até 1 salário mínimo de 2009/2011 1 - De mais de 1 a 2 salários mínimos 1 - De mais de 2 a 5 salários mínimos 1 - De mais de 5 a 10 salários mínimos 0 - Mais de 10 salários mínimos
“Há quantos anos você exerce funções de direção?”	TP_TEMP_FUNC_DIR	Num	_A _B	1 - Menos de 5 anos 1 - Entre 5 e 10 anos 0 - 11 anos ou mais
“Você assumiu a direção desta escola por”	TP_ASSC_DIR	Num	_A _B	1 - Seleção, eleição, ou seleção e eleição 1 - Indicação de técnicos, políticos ou outras indicações 0 - Outra forma
“Quanto ao projeto pedagógico desta escola neste ano (marque apenas uma opção)”	TP_PPD_DR	Num	_A _B _C _D	1 - Encaminhado pela Secretaria de Educação 1 - Elaborado pelo Diretor 1 - Elaborado pelo Diretor e/ou Professores ou Comunidade 1 - De outra maneira ou desconhece 0 - Não existe Projeto Pedagógico
“Qual é o percentual de professores com vínculo estável nesta escola?”	TP_VINC_EST	Num	_A _B _C _D	1 - Menor ou igual a 25% 1 - De 26% a 50% 1 - De 51% a 75% 1 - De 76% a 90% 0 - De 91% a 100%
“(…) Ocorreu na escola rotatividade do corpo docente?”	TP_ROT_CORP_DOC	Num	_A _B	1 - Não 1 - Sim, mas não foi grave 0 - Sim, e foi um problema

Descrição	Variável	Tipo¹	Ind.²	Código de Preenchimento³
				grave
<p>1. Tipo - Num: um regressor métrico (agregado em dimensão escolar); ou um regressor binário/dummies (previamente em dimensão escolar).</p> <p>2. Ind. - Cada variável é adicionada por um indicador quando o número de categorias é igual ou maior a 3: duas ou mais categorias especificadas e uma omitida.</p> <p>3. Descrição: "1 - NOME" é a proporção da categoria(s) especificada(s); e "0 - NOME" é a proporção da categoria omitida.</p>				

Fonte: elaborado pelo autor.