



UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CENTRO SOCIOECONÔMICO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Letícia Spíndola de Faria

**IMPACTOS DA MUNICIPALIZAÇÃO DO ENSINO FUNDAMENTAL SOBRE O  
DESEMPENHO ESCOLAR**

Florianópolis

2024

Letícia Spíndola de Faria

**IMPACTOS DA MUNICIPALIZAÇÃO DO ENSINO FUNDAMENTAL SOBRE O  
DESEMPENHO ESCOLAR**

Dissertação submetida ao Programa de Pós-Graduação  
em Economia da Universidade Federal de Santa Catarina  
para a obtenção do título de mestre em Economia.  
Orientador: Prof. Guilherme Valle Moura, Dr.

Florianópolis

2024

Ficha de identificação da obra elaborada pela autora  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Faria, Leticia Spindola de  
Impactos da Municipalização do Ensino Fundamental sobre  
o Desempenho Escolar / Leticia Spindola de Faria ;  
orientador, Guilherme Valle Moura, 2024.  
80 p.

Dissertação (mestrado) - Universidade Federal de Santa  
Catarina, Centro Socioeconômico, Programa de Pós-Graduação em  
Economia, Florianópolis, 2024.

Inclui referências.

1. Economia. 2. Municipalização. 3. Avaliação de Políticas  
Públicas. 4. Educação. 5. Diferenças-em-Diferenças. I. Moura,  
Guilherme Valle. II. Universidade Federal de Santa  
Catarina. Programa de Pós-Graduação em Economia. III. Título.

Letícia Spíndola de Faria

**IMPACTOS DA MUNICIPALIZAÇÃO DO ENSINO FUNDAMENTAL SOBRE O  
DESEMPENHO ESCOLAR**

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Guilherme Valle Moura, Dr.  
Universidade Federal de Santa Catarina

Prof. Francis Carlo Petterini Lourenço, Dr.  
Universidade Federal de Santa Catarina

Prof. Marco Túlio Aniceto França, Dr.  
Pontifícia Universidade Católica do Rio Grande do Sul

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Economia.

---

Coordenação do Programa de Pós-Graduação em Economia

---

Guilherme Valle Moura, Dr.  
Orientador

Florianópolis, 2024.

## **AGRADECIMENTOS**

Agradeço, primeiramente e principalmente, a Deus pelas vitórias e aprendizados que obtive ao longo desta caminhada.

Aos meus familiares e amigos que me acompanharam nesta intensa jornada e compreenderam a necessidade da minha ausência.

Ao meu orientador e demais professores do Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina, pelo conhecimento compartilhado.

Ao Tribunal de Contas do Estado de Santa Catarina, pelo apoio a mim conferido para que pudesse conciliar trabalho e estudos.

## RESUMO

A oferta do Ensino Fundamental pela rede municipal é uma diretriz presente tanto na Constituição Federal de 1988 quanto em outros importantes instrumentos normativos como a Lei de Diretrizes e Bases da Educação Nacional. Contudo, é muito comum que Estados e Municípios dividam essa atribuição, geralmente ocorrendo a oferta dos Anos Iniciais pela rede municipal e Anos Finais pela rede estadual. Denomina-se política de municipalização do ensino a transferência da gestão de uma escola estadual para a rede municipal, ocorrendo a delegação de todas as responsabilidades e encargos respectivos. Não obstante, em virtude das significativas disparidades existentes entre os Estados e Municípios brasileiros, a implementação dessa política não seguiu diretrizes gerais e unificadas, mas, ao contrário, ocorreu segundo diferentes processos a depender dos entes federados envolvidos, além de ocorrer em momentos distintos do tempo a depender da discricionariedade do gestor. Por conseguinte, a municipalização resultou em processos distintos e repletos de especificidades e peculiaridades. Nesse sentido, este trabalho teve por objetivo avaliar o impacto da municipalização do ensino fundamental sobre o desempenho escolar dos alunos sob a ótica dos resultados na prova Saeb. Para tanto, utilizou-se o modelo de Diferenças-em-Diferenças (DiD) proposto por Callaway e Sant'Anna (2021) que analisa eventos caracterizados por múltiplos períodos de tempo, temporalidade diferenciada do tratamento e permite identificar os efeitos heterogêneos do tratamento. Adicionalmente, esse impacto também foi calculado através da metodologia mais comumente utilizada, isto é, *Two-way Fixed Effects*, a fim de se compreender como as diferentes metodologias tratam do tema. No geral, os resultados indicam efeitos negativos da municipalização, com pouca precisão estatística e em pequena magnitude.

**Palavras-chave:** Educação, Avaliação de Políticas Públicas, Municipalização, Diferenças-em-Diferenças.

## ABSTRACT

The provision of Primary Education by the municipal network is a guideline present both in the Federal Constitution of 1988 and in other important normative instruments such as the Law of Guidelines and Bases of National Education. However, it is very common for States and Municipalities to share this responsibility, usually with the provision of the Initial Years by the municipal network and the Final Years by the state network. The transfer of management of a state school to the municipal network is called the municipalization policy, involving the delegation of all respective responsibilities and burdens. Nevertheless, due to significant disparities between Brazilian States and Municipalities, the implementation of this policy did not follow general and unified guidelines, but rather occurred through different processes depending on the federated entities involved, and occurred at different times depending on the discretion of the manager. Consequently, municipalization resulted in distinct processes full of specificities and peculiarities. In this sense, this study aimed to evaluate the impact of the municipalization of primary education on students' academic performance from the perspective of results on the Saeb test. For this purpose, the Differences-in-Differences (DiD) model proposed by Callaway and Sant'Anna (2021) was used, which analyzes events characterized by multiple periods of time, differentiated temporality of treatment, and allows identifying the heterogeneous effects of treatment. Additionally, this impact was also calculated using the most commonly used methodology, that is, Two-way Fixed Effects, in order to understand how different methodologies address the topic. Overall, the results indicate negative effects of municipalization, with little statistical precision and in small magnitude.

**Keywords:** Education, Public Policy Evaluation, Municipalization, Differences-in-Differences.

## LISTA DE ILUSTRAÇÕES

### FIGURAS

Figura 1 -	Comparação da nota padronizada para 5º e 9º anos das escolas municipalizadas pré e pós tratamento.....	55
Figura 2 -	Comparação da nota de Língua Portuguesa para 5º e 9º anos das escolas municipalizadas pré e pós tratamento.....	55
Figura 3 -	Comparação da nota de Matemática para 5º e 9º anos das escolas municipalizadas pré e pós tratamento.....	56
Figura 4 -	Gráficos de densidade da nota padronizada para 5º e 9º anos das escolas tratadas e controle.....	57
Figura 5 -	Gráficos de densidade da nota de Língua Portuguesa para 5º e 9º anos das escolas tratadas e controle.....	57
Figura 6 -	Gráficos de densidade da nota de Matemática para 5º e 9º anos das escolas tratadas e controle.....	58
Figura 7 -	Efeitos da municipalização nos diferentes grupos.....	62
Figura 8 -	Estimador dinâmico de efeitos médios para nota padronizada (5º e 9º anos).....	65
Figura 9 -	Estimador dinâmico de efeitos médios para nota de Matemática (5º e 9º anos).....	66
Figura 10 -	Estimador dinâmico de efeitos médios para nota de Língua Portuguesa (5º e 9º anos).....	66
Figura 11 -	Estimador dinâmico balanceado (5 anos) de efeitos médios para nota padronizada (5º e 9º anos).....	68
Figura 12 -	Figura 12: Estimador dinâmico balanceado (5 anos) de efeitos médios para nota de Matemática (5º e 9º anos).....	69
Figura 13 -	Estimador dinâmico balanceado (5 anos) de efeitos médios para nota de Língua Portuguesa (5º e 9º anos).....	70

### GRÁFICOS

Gráfico 1 -	Escolas municipalizadas no Brasil (2010-2021).....	23
Gráfico 2 -	Quantidade de matrículas municipalizadas (2010-2021).....	24
Gráfico 3 -	Quantidade média de matrículas municipalizadas (2010-2021).....	24
Gráfico 4 -	Escolas estadualizadas no Brasil (2010-2021).....	25
Gráfico 5 -	Evolução da proficiência média na prova Saeb em Língua Portuguesa no 5º ano do Ensino Fundamental – 2011 a 2021.....	44
Gráfico 6 -	Evolução da proficiência média na prova Saeb em Matemática no 5º ano do Ensino Fundamental – 2011 a 2021.....	45
Gráfico 7 -	Evolução da proficiência média na prova Saeb em Língua Portuguesa no 9º ano do Ensino Fundamental – 2011 a 2021.....	45
Gráfico 8 -	Evolução da proficiência média na prova Saeb em Matemática no 9º ano do Ensino Fundamental – 2011 a 2021.....	46
Gráfico 9 -	Distribuição dos escores de propensão após o pareamento.....	52



## LISTA DE TABELAS

Tabela 1	- Estatísticas descritivas das variáveis utilizadas no PSM (pré pareamento).....	53
Tabela 2	- Estatísticas descritivas das variáveis utilizadas no PSM (pós pareamento).....	53
Tabela 3	- Estimador simples dos efeitos médios de tratamento.....	60
Tabela 4	- Efeitos médios de tratamento para os diferentes grupos e agregado.....	60
Tabela 5	- Estimador de efeitos médios dinâmico agregado.....	64
Tabela 6	- Estimador de efeitos médios dinâmico agregado balanceado para cinco anos.....	68
Tabela 7	- Estimador TWFE.....	72
Tabela 8	- Resumo dos efeitos médios agregados calculados de acordo com a metodologia Callaway e Sant’Anna e o estimador TWFE.....	72

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>11</b>
<b>2</b>	<b>A POLÍTICA DE MUNICIPALIZAÇÃO</b> .....	<b>13</b>
2.1	O ESTADO E AS POLÍTICAS PÚBLICAS EDUCACIONAIS .....	13
2.2	A MUNICIPALIZAÇÃO DO ENSINO .....	15
2.2.1	NO BRASIL .....	17
<b>3</b>	<b>CONTEXTO METODOLÓGICO</b> .....	<b>27</b>
3.1	A ABORDAGEM CALLAWAY E SANT'ANNA (2021).....	34
<b>4</b>	<b>BASE DE DADOS E PROCEDIMENTOS METODOLÓGICOS</b> .....	<b>40</b>
4.1	BASE DE DADOS .....	41
4.2	ESTRATÉGIA EMPÍRICA E ESTATÍSTICAS DESCRITIVAS.....	47
<b>5</b>	<b>RESULTADOS</b> .....	<b>59</b>
<b>5.1</b>	<b>RESULTADOS DO MODELO CALLAWAY &amp; SANT'ANNA</b> .....	<b>59</b>
5.2	ANÁLISE DE ROBUSTEZ.....	71
5.3	COMPARAÇÃO DE ESTIMADORES: CALLAWAY & SANT'ANNA X TWFE.....	72
<b>6</b>	<b>CONCLUSÃO</b> .....	<b>73</b>
<b>REFERÊNCIAS</b>	.....	<b>76</b>
<b>APÊNDICE A</b>	<b>Gráficos de Distribuição de Densidade das variáveis utilizadas no <i>Propensity Score Matching</i> (antes e depois do pareamento).</b>	<b>78</b>
<b>APÊNDICE B</b>	<b>Análise de Robustez: efeitos médios do tempo de tratamento para as diferentes variáveis analisadas</b> .....	<b>80</b>

## 1 INTRODUÇÃO

O direito à educação é reconhecido pela Constituição Federal de 1988 como um direito social fundamental, ou seja, é um direito básico de todos os indivíduos que deve ser ofertado de forma gratuita pelo Estado independente de raça, sexo, idade ou qualquer forma de discriminação.

Além de ser um direito subjetivo, o investimento em educação é caracterizado por gerar externalidades positivas, apresentando benefícios econômicos e não-econômicos. A literatura econômica aponta como efeitos positivos do acesso à educação a expectativa de aumento do nível salarial, aumento de oportunidades no mercado de trabalho, redução da criminalidade, redução do trabalho infantil, dentre outros (D'ATRI, 2007). Devido a seus impactos positivos, o investimento em educação é considerado um instrumento estratégico no âmbito econômico e uma forma efetiva de combater desigualdades sociais.

Contudo, a despeito da promoção e ampliação do acesso à educação e da existência de mecanismos que estimulam a permanência dos estudantes nos sistemas de ensino a fim de efetivar essa garantia constitucional, deve-se analisar outro importante aspecto deste direito: a qualidade do ensino ofertado.

A realidade do ensino público no Brasil é caracterizada por baixos e preocupantes níveis de proficiência e repleta de desafios a serem superados, tais como: altas taxas de evasão escolar e de distorção de idade/série.

Considerando a importância da oferta de educação de qualidade para o desenvolvimento de uma sociedade e da garantia de um ambiente escolar adequado para que os profissionais da educação possam exercer suas funções plenamente, é imprescindível selecionar políticas públicas educacionais baseadas em ações ordenadas e estratégicas com vistas a garantir o direito à educação, concretizado em uma oferta de ensino de qualidade, constituindo-se, também, em uma via para atenuar as desigualdades sociais.

A municipalização do ensino público é uma política educacional, também denominada descentralização do ensino, caracterizada pela transferência da gestão de uma escola estadual para a rede municipal, ocorrendo a delegação de todas as responsabilidades e encargos respectivos.

No Brasil, a Constituição Federal de 1988 e a Lei de Diretrizes e Bases da Educação prescrevem a municipalização de determinadas etapas de ensino, além de existirem instrumentos capazes de financiar o aumento do número de matrículas característico dessa

política, como o Fundeb. Contudo, essa política educacional ainda é, substancialmente, uma decisão política a ser acordada entre a gestão estadual e municipal.

Os defensores da municipalização alegam que a maior proximidade da gestão escolar com a Secretaria Municipal de Educação e com a comunidade favorece a tomada de decisões mais adequadas àquela realidade, de forma mais rápida e flexível, tornando o combate aos problemas individuais de cada escola mais efetivo e, com isso, melhorando a oferta de ensino, além de permitir maior controle social e participação da comunidade e das famílias.

Em contrapartida, os críticos argumentam que muitos municípios não possuem estruturas adequadas para implementar políticas de expansão do ensino público e, diante da ausência de condições técnicas, financeiras e de pessoal, a política de municipalização poderia resultar em ineficiência e impactar negativamente a qualidade do ensino, além de agravar as desigualdades existentes entre os municípios, fragmentando o sistema educacional e dificultando a coordenação de políticas educacionais a nível nacional (LEME; PAREDES; SOUZA, 2009).

Ademais, deve-se considerar que o Brasil é caracterizado por significativa diversidade de realidades econômica, cultural e social, o que resulta em marcantes níveis de desigualdade entre os municípios brasileiros. Aliado a esse fator, destaca-se que o processo de municipalização não seguiu diretrizes gerais e unificadas, mas, ao contrário, ocorreu segundo diferentes processos a depender dos entes federados envolvidos o que evidencia a existência de singularidades e especificidades na implementação dessa política.

Nesse sentido, a presente dissertação visa analisar o impacto da municipalização das escolas sobre o desempenho escolar, a fim de compreender a qualidade do ensino ofertado.

A despeito de todas as demais capacidades e habilidades proporcionadas pela oferta de uma educação de qualidade, nesta pesquisa desempenho escolar será mensurado através de resultados da prova Saeb, isto é, pelo índice nacional de proficiência, utilizando o pressuposto fundamental de que a efetivação de uma educação de qualidade tem como consequência direta resultados superiores em avaliações de aprendizagem.

Sabe-se que o fator com maior impacto sobre o desempenho escolar dos alunos é o denominado *background familiar*, termo utilizado para designar o ambiente socioeconômico em que o estudante está inserido e composto por fatores como: nível educacional e situação financeira dos pais, recursos materiais disponíveis, estabilidade e estrutura familiar, dentre outros. Contudo, o desempenho dos alunos também depende, em alguma medida, da realidade de sua escola, como, por exemplo, infraestrutura escolar; recursos físicos, pedagógicos e

humanos; jornada escolar; experiência e motivação dos professores, dentre outros fatores diretamente ligados à gestão escolar (LEME; PAREDES; SOUZA, 2009).

Dessa forma, é imprescindível que se busque compreender a magnitude do impacto da gestão escolar sobre a aprendizagem de modo a fornecer subsídios para o desenvolvimento de políticas estratégicas capazes de ampliar o nível de aprendizagem dos alunos.

Esta dissertação está dividida em cinco capítulos além desta introdução. Primeiro é apresentado o conceito de municipalização como uma política pública, isto é, como parte de uma estratégia governamental voltada para a melhoria do sistema educacional local, e sua aplicação no Brasil. Na sequência, aborda-se o arcabouço metodológico que justifica a adoção da metodologia utilizada e são apresentadas as bases de dados e a estratégia empírica. Após, os resultados obtidos a partir dos estimadores Callaway e Sant'Anna são analisados conectando-os à literatura relevante e comparando-os com os diferentes resultados obtidos através da metodologia *Two-way Fixed Effects*. Por fim, comentários finais são apresentados na conclusão.

## **2 A POLÍTICA DE MUNICIPALIZAÇÃO**

### **2.1 O ESTADO E AS POLÍTICAS PÚBLICAS EDUCACIONAIS**

O Estado desempenha papel fundamental no desenvolvimento e implementação de políticas públicas, sendo o principal responsável pelas ações, projetos e atividades que visam proporcionar aumento de bem-estar à população e atenuar a desigualdade entre os indivíduos através da provisão de serviços básicos (NASCIMENTO et al., 2023).

Existem inúmeras definições, conceitos e objetivos atrelados ao termo “políticas públicas”, mas, pode-se resumir, conforme apresentado por Souza (2006), que estas se constituem em um conjunto sistematizado de ações governamentais intencionais cujo objetivo principal é assegurar direitos fundamentais aos cidadãos e solucionar problemas sociais, não se limitando a leis e regras e envolvendo processos adicionais à sua proposição, isto é, incluindo as fases de implementação, execução, monitoramento e avaliação.

Nesse contexto, as políticas públicas possuem o propósito de influenciar, modificar, normatizar ou conservar aspectos da estrutura social e, para tanto, envolvem a atuação de diversos agentes em variados níveis de autoridade tratando de temas relacionados a diversos setores, por exemplo, educação, saúde, esporte, justiça, assistência social, dentre outros (FERREIRA; SANTOS, 2014).

Deste modo, políticas públicas educacionais integram o escopo de responsabilidade estatal.

Conforme assevera o artigo 205 da Constituição Federal de 1988 (CF/88), a educação é um direito de todos e dever do Estado e da família, sendo o Estado responsável por sua proteção e promoção de sua oferta gratuita.

Complementarmente ao supramencionado dispositivo, o artigo 208, inciso I e respectivos parágrafos prescrevem que:

Art. 208. O dever do Estado com a educação será efetivado mediante a garantia de:  
I - educação básica obrigatória e gratuita dos 4 (quatro) aos 17 (dezessete) anos de idade, assegurada inclusive sua oferta gratuita para todos os que a ela não tiveram acesso na idade própria; [\(Redação dada pela Emenda Constitucional nº 59, de 2009\) \(Vide Emenda Constitucional nº 59, de 2009\)](#)  
(....)  
§ 1º O acesso ao ensino obrigatório e gratuito é direito público subjetivo.  
§ 2º O não-oferecimento do ensino obrigatório pelo Poder Público, ou sua oferta irregular, importa responsabilidade da autoridade competente.

Portanto, o Estado é responsável pela oferta da educação básica obrigatória e gratuita, incluindo todas as etapas de ensino, quais sejam, Educação Infantil, Ensino Fundamental e Ensino Médio.

Esse direito constitucional se efetiva mediante a implementação de políticas públicas educacionais que envolvem a criação e manutenção de escolas públicas, seleção de docentes qualificados e garantia de formação continuada, valorização profissional, elaboração de currículos e diretrizes pedagógicas adequadas que visem estabelecer um padrão de qualidade, manutenção da infraestrutura escolar, estabelecimento de mecanismos e ações de financiamento da educação, atendimento adequado aos alunos portadores de deficiência, determinação de metas de proficiência a serem atingidas, dentre outros (SANTOS; VASCONCELOS, 2023).

Mais do que isso, as referidas políticas não se restringem ao âmbito escolar, mas contemplam diversas áreas, tratando também do enfrentamento dos problemas que afetam diretamente a qualidade da educação dos estudantes como, por exemplo, fome, drogas e violência.

De acordo com Santos e Vasconcelos (2023), pode-se classificar políticas públicas educacionais em quatro campos complementares: distributivas, redistributivas, regulatórias e constitutivas.

As políticas públicas distributivas são instrumentos pelos quais o Estado objetiva alocar recursos e benefícios de forma mais igualitária a fim de reduzir desigualdades e promover o

acesso à educação. Os referidos autores citam como exemplo construção de escolas em áreas carentes, programas de merenda escolar e transporte gratuito.

As políticas públicas redistributivas objetivam a transferência de recursos de grupos ou setores mais privilegiados para os mais desfavorecidos visando reduzir disparidades socioeconômicas existentes e garantir equidade no sistema educacional. Pode-se citar como exemplo: concessão de bolsas de estudo e auxílio financeiro para famílias de baixa renda para que mantenham seus filhos na escola (SANTOS; VASCONCELOS, 2023).

As políticas públicas regulatórias visam criar normas, diretrizes e padrões que norteiam, melhoram ou supervisionam a qualidade dos serviços educacionais. Pode-se citar como exemplo o estabelecimento de critérios de qualificação dos docentes (SANTOS; VASCONCELOS, 2023).

Por fim, as políticas públicas constitutivas objetivam criar ou transformar estruturas e instituições educacionais, visando estabelecer os fundamentos e bases de determinadas áreas. São exemplos: a elaboração de marcos legais para a educação, definição de políticas de financiamento, criação de órgãos de supervisão e avaliação e reforma curricular (SANTOS; VASCONCELOS, 2023).

Em síntese, políticas públicas educacionais são ações governamentais que visam proporcionar o acesso a uma educação de qualidade e promover equidade no âmbito escolar, permitindo que todos os estudantes tenham oportunidade de desenvolvimento e abordando também temas alheios à educação, mas que impactam diretamente a aprendizagem dos alunos.

Dessa forma, é por meio das políticas públicas educacionais que o Estado busca cumprir seu dever constitucional de assegurar o direito à educação de qualidade a todos, independentemente de sua condição socioeconômica, raça, sexo, idade ou qualquer forma de discriminação, conseqüentemente, promovendo desenvolvimento social.

## 2.2. A MUNICIPALIZAÇÃO DO ENSINO

A municipalização, também denominada descentralização do ensino, é uma política de transferência da gestão de uma unidade escolar pertencente à rede estadual para a rede municipal, englobando todas as responsabilidades e os encargos respectivos.

Deve-se destacar que a decisão de municipalizar uma unidade escolar é influenciada diretamente pelo projeto político vigente no Estado e no Município. Dessa forma, essa política pode ser essencialmente caracterizada como um processo de colaboração e cooperação entre os entes federados (BAPTISTA, 2022).

A municipalização sob a ótica de se constituir em uma estratégia governamental voltada para a melhoria do sistema educacional local, pode ser entendida como uma política pública educacional cujo objetivo principal é aproximar a tomada de decisões e a prestação dos serviços públicos educacionais da comunidade, tornando a atuação governamental, em tese, mais eficiente e sensível às necessidades e peculiaridades locais.

No Brasil, em alguns casos ocorre a municipalização de determinados anos/séries e/ou etapas de ensino. Neste trabalho considera-se municipalização a transferência integral de uma unidade escolar pertencente à rede estadual para a rede municipal, envolvendo toda a sua estrutura física, de pessoal e demais encargos.

De forma geral, a política de municipalização é uma alteração da forma organizacional do ensino sendo um processo de transferência de gestão e de responsabilidades de entidades centrais para entidades locais, isto é, de Estados para Municípios, constituindo-se em uma política descentralizadora e envolvendo as dimensões política, legal e socioeconômica. Conforme apresenta Mendes (2022), a descentralização do ensino é uma das abordagens mais adotadas pelas políticas educacionais em escala global, e acompanha a tendência de redistribuição e descentralização do poder.

Os defensores dessa política alegam que, devido à proximidade da gestão escolar da comunidade, as políticas desenvolvidas na unidade de ensino seriam, em tese, mais assertivas e adequadas uma vez que a gestão municipal tem maior conhecimento da realidade social, econômica e cultural daquela comunidade, impactando positivamente a qualidade educacional. Dessa forma, a municipalização estaria fundamentada em argumentos de eficácia e eficiência na oferta de ensino, bem como de racionalidade no uso dos recursos públicos. Adicionalmente, essa proximidade facilitaria o acompanhamento e avaliação das ações, de forma a garantir a implementação da política em sua totalidade (BOTH, 1997). Complementarmente, autores como Franco (2023) defendem que a municipalização é uma das principais formas de democratizar o acesso à educação, pois promove a participação da comunidade nas decisões e permite maior controle social.

Em contrapartida, os críticos argumentam que muitos municípios não possuem estruturas adequadas para implementar políticas de expansão do ensino público e, diante da ausência de condições técnicas, financeiras e de pessoal, a política de municipalização poderia resultar em ineficiência e impactar negativamente a qualidade do ensino, além de agravar as desigualdades existentes entre municípios, fragmentando o sistema educacional e dificultando a coordenação de políticas educacionais a nível nacional (LEME; PAREDES; SOUZA, 2009).



Em suma, a municipalização ocorre com a transferência de responsabilidades de uma instância em âmbito central para a respectiva instância em nível local, mais próxima da comunidade escolar, o que resulta em uma maior aproximação do cidadão do poder de decisão e permite maior participação da sociedade democratizando o sistema de ensino. No entanto, é fundamental analisar fatores estruturais, administrativos, técnicos e financeiros para que essa política possa ser efetivamente implementada.

### 2.2.1 NO BRASIL

Para compreender a trajetória da política de municipalização no Brasil, deve-se considerar que esta não é uma política recente, mas, ao contrário, foi abordada em diversos momentos da história da educação nacional e refletiu as transformações ocorridas na sociedade brasileira ao longo dos anos.

De acordo com Soares e Vidal (2022), Anísio Teixeira, ainda na década de 50, foi o precursor do termo “municipalização” no Brasil. Uma das propostas defendidas em sua “Tese sobre a Municipalização da Escola” foi a adoção da política de municipalização do ensino primário, assegurando que o município possuía condições necessárias para gerenciar e assumir esta oferta de ensino, devendo estar subordinado às leis estaduais e em conformidade com o previsto nas leis de diretrizes e bases federais (MENDES, 2022).

Para analisar mais detalhadamente a cronologia da política de municipalização no Brasil, é fundamental investigar as legislações que abordaram o tema, tais como a Constituição Federal de 1988 e a Lei de Diretrizes e Bases da Educação, assim como alguns mecanismos de financiamento da educação, essencialmente, Fundef e Fundeb.

Na década de 70, a Lei nº 5.692/71, que fixa Diretrizes e Bases para o ensino, em seu artigo 58, parágrafo único, previu que os municípios assumissem progressivamente a responsabilidade pelo ensino de primeiro grau, tratando diretamente da descentralização da educação e induzindo a adoção dessa política através da transferência de responsabilidades, encargos e serviços educacionais dos Estados para seus Municípios.

Na década de 80, a promulgação da Constituição Federal de 1988 tornou-se um marco significativo na implementação de políticas públicas educacionais ao caracterizar a educação como um direito de todos e um dever do Estado, reconhecendo a educação como um instrumento essencial para o desenvolvimento humano, social e econômico, visando a equidade e igualdade de oportunidade para todos os indivíduos, o que permitiu que se compreendesse a

educação básica como integrante do conteúdo do mínimo existencial e tornando juridicamente exigível o desenvolvimento de políticas públicas educacionais (BAPTISTA, 2022).

Ademais, a CF/88 inovou, em seu artigo 18, ao promover a reorganização político-administrativa do Brasil, conferindo autonomia aos municípios e caracterizando-os como entes federados, o que resultou na ampliação das responsabilidades municipais e uma redefinição da atuação desses entes dentro do cenário nacional.

Essa previsão é de suma importância para a descentralização do ensino, uma vez que permitiu que os municípios tivessem autonomia para organizar e administrar seus sistemas de ensino e maior poder de decisão e atuação, o que impulsionou a política de municipalização.

Além disso, os municípios passaram a contar com maiores transferências governamentais e, portanto, houve aumento significativo do volume de recursos financeiros, fator necessário frente à expansão de seus gastos devido às novas atribuições. Dessa forma, a maior autonomia dada aos municípios pela Carta Magna reverberou positivamente no processo de municipalização.

De forma geral, esta reforma administrativa ocorrida através da promulgação da Constituição, consolidou um processo generalizado de descentralização dos serviços básicos não apenas no âmbito educacional, mas também em áreas como saúde e serviço social, resultando em um aumento da participação dos municípios no fornecimento de serviços. Nessa configuração, o Governo Federal seria responsável por coordenar, formular e direcionar a aplicação das políticas públicas, já os municípios seriam responsáveis por executá-las.

Adicionalmente, o artigo 30, inciso VI, da referida Carta Magna, determina que é competência municipal: manter, com a cooperação técnica e financeira da União e do Estado, programas de educação infantil e de ensino fundamental.

Além do mais, o artigo 211 previu expressamente a existência de um regime de colaboração entre a União, Estados, Distrito Federal e Municípios em relação à organização dos sistemas de ensino. Nesse contexto, nota-se que a articulação e colaboração entre os entes federados é essencial para criar condições favoráveis para implementação da política de municipalização.

Outrossim, o referido artigo definiu que os municípios atuarão prioritariamente no ensino fundamental e na educação infantil, enquanto os Estados e o Distrito Federal atuarão prioritariamente no ensino fundamental e médio. Assim determina a Constituição:

§ 3º Os Estados e o Distrito Federal atuarão prioritariamente no ensino fundamental e médio. [\(Incluído pela Emenda Constitucional nº 14, de 1996\)](#)

Dessa forma, o Texto Constitucional não instituiu exclusividade na oferta do ensino fundamental pelos municípios, não havendo impedimento à essa prestação pela rede estadual.

Transitando para a década de 90, a edição da Lei nº 9.394/1996 que estabelece as diretrizes e bases da educação nacional (LDB) reforçou a política de municipalização do ensino fundamental. Destaca-se que, ao organizar o desenvolvimento da educação no país, esta Lei se tornou o principal mecanismo norteador das políticas públicas educacionais.

O artigo 11 do supramencionado dispositivo, ao tratar das incumbências dos municípios, é mais específico que a Constituição Federal e determina que:

Art. 11. Os Municípios incumbir-se-ão de:

(...)

V - **oferecer** a educação infantil em creches e pré-escolas, e, com prioridade, o ensino fundamental, permitida a atuação em outros níveis de ensino somente quando estiverem atendidas plenamente as necessidades de sua área de competência e com recursos acima dos percentuais mínimos vinculados pela Constituição Federal à manutenção e desenvolvimento do ensino. [\(Redação dada pela Lei nº 12.061, de 2009\)](#)

Já o artigo 10, determina as atribuições dos Estados:

Art. 10. Os Estados incumbir-se-ão de:

(...)

VI - **assegurar** o ensino fundamental e **oferecer**, com prioridade, o ensino médio a todos que o demandarem, respeitado o disposto no art. 38 desta Lei; [\(Redação dada pela Lei nº 12.061, de 2009\)](#)

De acordo com a terminologia utilizada no referido normativo, é atribuição dos Municípios oferecer educação infantil em creches e pré-escolas e, com prioridade, ensino fundamental, enquanto é incumbência dos Estados assegurar o ensino fundamental e oferecer, com prioridade, o ensino médio. Conclui-se, portanto, que a oferta do ensino fundamental é de responsabilidade dos Municípios, cabendo à administração estadual a oferta do ensino médio.

Dessa forma, assim como previsto na Carta Magna, não há óbice à oferta do ensino fundamental pelo Estado, mas sim uma diretriz de que este deve ser ofertado prioritariamente pelos Municípios, devendo o Estado assegurar esta prestação.

Em suma, a LDB impulsionou a municipalização do ensino, além de corroborar as previsões constitucionais.

Nesse sentido, conforme as atribuições definidas pela Constituição Federal e pela LDB, apenas faz sentido analisar a política de municipalização em relação às unidades escolares que

ofertam ensino fundamental, uma vez que representa a passagem da atribuição conferida pela legislação nacional ao responsável pela sua oferta, isto é, dos Estados para seus Municípios.

Por conseguinte, conforme previsto no ordenamento jurídico nacional, a política de municipalização do ensino fundamental é uma diretriz que deve ser cumprida, cabendo à administração pública estadual definir a forma de execução considerando, para tanto, as especificidades de cada município e o momento oportuno. Em regra, por consequência, existe apenas algum grau de discricionariedade em como realizar o processo de transferência, não havendo mérito de escolha quanto a sua realização (BAPTISTA, 2022).

Devido a essa discricionariedade, o processo de municipalização das unidades escolares é repleto de especificidades e peculiaridades a depender do Estado e do Município envolvidos, além de ocorrer em momentos distintos do tempo a depender da decisão do gestor.

Além das referidas previsões normativas, deve-se destacar as legislações que abordam os instrumentos de financiamento da educação, uma vez que se constituíram em grandes estímulos ao processo de municipalização ocorrido no Brasil.

A Emenda Constitucional nº 14/96 criou o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério (Fundef), de natureza contábil, que foi implantado a nível nacional em 1998 e apresentou a estrutura de financiamento do Ensino Fundamental.

De acordo com o artigo 2º da Lei nº 9.424 de 1996, que regulamentou o Fundef, a distribuição dos recursos do Fundo, no âmbito de cada Estado e Distrito Federal, deveria ocorrer entre governo estadual e governos municipais a depender da proporção do número de alunos matriculados no ensino fundamental, vinculado ao custo mínimo anual por aluno<sup>1</sup>, o que se constituiu em grande estímulo à municipalização visto que, ao expandirem suas redes, os municípios fariam jus à uma contrapartida financeira a ser recebida devido ao aumento do número de matrículas.

Com isso, tornou-se financeiramente atrativo para os Municípios ampliarem sua oferta de ensino aumentando o número de alunos matriculados em escolas municipais na medida em que a lógica de distribuição dos recursos do Fundef estava associada ao número de matrículas, isto é, quanto mais alunos matriculados no ensino fundamental, maior a quantidade recebida de recursos.

---

<sup>1</sup> O “custo mínimo por aluno” é um indicador que estabelece o valor mínimo a ser investido por aluno matriculado no ensino fundamental por ano. Esse valor é definido com base em critérios estabelecidos na legislação e visa garantir o investimento de recursos capazes de assegurar ensino e aprendizagem de qualidade.

Dessa forma, o Fundef pode ser caracterizado como um instrumento incentivador da política de municipalização, constituindo-se em um dos principais fatores explicativos da adoção dessa política e representando uma revolução na forma de financiamento da educação. Como resultado, a política de municipalização foi amplamente adotada após sua instituição (D'ATRI, 2007).

Ademais, a Lei instituidora do Fundef definiu que os recursos do Fundo seriam compostos por um percentual expressivo da arrecadação de impostos e que seriam utilizados exclusivamente para custear a manutenção e desenvolvimento do ensino<sup>2</sup> fundamental público e na valorização de seu Magistério.

Outro fator relevante instituído pela referida Lei foi a determinação de que compete aos Estados e ao governo central a complementação dos recursos quando os municípios não atingirem o valor mínimo por aluno/ano definido nacionalmente, o que se coaduna com a previsão legal da existência de um regime de colaboração entre as esferas administrativas no âmbito da educação brasileira.

Além do estímulo financeiro positivo, a estrutura do Fundef, em tese, incitaria a competição entre os municípios e os Estados a fim de atrair o maior número de matrículas e consequentemente, maior repasse. Como consequência dessa competição, poderia ocorrer uma melhoria da qualidade do ensino e maiores investimentos na educação em busca da ampliação da rede (MENDES SEGUNDO; RABELO, 2005).

O Banco Mundial, em relatório elaborado em 2003<sup>3</sup>, analisou a importância da reforma educacional ocasionada pelo Fundef e concluiu como exitosa a estratégia de correlacionar o aumento de matrículas com aumento de recursos. A perspectiva adotada pelo Banco considerou a melhoria na eficiência dos serviços públicos, concluindo que o ensino fundamental deveria ser ofertado integralmente pelos municípios por apresentarem melhor custo-benefício na oferta de educação (MENDES SEGUNDO; RABELO, 2005). Com isso, pode-se afirmar que o estímulo à política de municipalização e seus impactos positivos foram ratificados por organismos internacionais.

O Fundef vigorou entre 1998 e 2006 e foi substituído pelo Fundo de Manutenção e Desenvolvimento da Educação Básica e de Valorização dos Profissionais da Educação

---

<sup>2</sup> O rol das despesas consideradas manutenção e desenvolvimento do ensino estão listadas no artigo 70 da Lei nº 9.394/1996 (LDB). O artigo 71 da referida Lei lista as despesas que não consideradas de manutenção e desenvolvimento do ensino.

<sup>3</sup> Este relatório denomina-se “Educação Municipal no Brasil: recursos, incentivos e resultados” e foi produzido em 2003 pelo Banco Mundial. Disponível em: <https://documents1.worldbank.org/curated/en/263001468235465175/pdf/244130vol020Portuguese.pdf>

(Fundeb), criado pela Emenda Constitucional nº 53 de 2006, que manteve a lógica original do Fundef com algumas alterações, continuando a ser conceituado como um Fundo contábil.

O Fundeb foi regulamentado pela Lei nº 11.494/2007 e pelo Decreto nº 6.253/2007. A diferença primordial é que o novo modelo passou a contemplar toda a Educação Básica<sup>4</sup>, e não apenas o ensino fundamental. Dessa forma, o total de alunos matriculados na rede pública de ensino passou a ser considerado como parâmetro para a distribuição do recurso e estes se tornaram vinculados exclusivamente à aplicação na educação básica. Além disso, houve aumento significativo do volume de recursos destinados ao Fundo devido ao fato da legislação adicionar novas fontes de receita.

Dessa forma, o Fundeb teve como objetivo manter e desenvolver a educação básica pública no Brasil dando continuidade e ampliando o processo iniciado com o Fundef.

A Emenda Constitucional n.º 108, de 27 de agosto de 2020, transformou o Fundeb em instrumento permanente de financiamento da educação, regulamentado pela Lei n.º 14.113 de 2020. Com isso, uma das principais fontes de financiamento da educação no país foi tornada permanente além de ter ocorrido a ampliação da quantidade de recursos destinados pela União a título de complementação.

Adicionalmente, o novo Fundeb inovou ao correlacionar uma parte das transferências nacionais aos resultados obtidos em educação, obrigando os Estados a implementarem um mecanismo baseado em resultados para incentivar o aprimoramento da oferta de educação de seus municípios.

Em síntese, o processo de municipalização do ensino determinado pela legislação e induzido por mecanismos de financiamento governamental através do Fundef e, posteriormente, pelo Fundeb, ocorreu em praticamente todos os Estados da Federação.

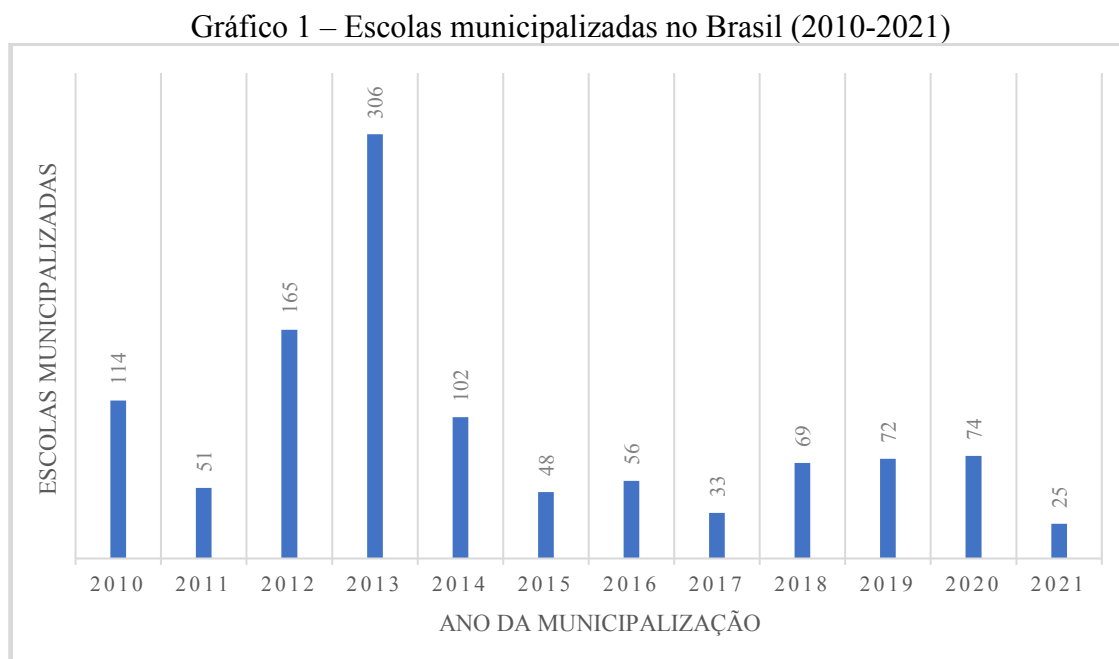
No entanto, conforme afirmado anteriormente, a implementação dessa política não sucedeu de forma padronizada e em cada Estado foi caracterizada por peculiaridades e especificidades devido às suas características sociais, educacionais e políticas substancialmente distintas. Em outras palavras, não houve diretrizes gerais para normatizar a transferência de gestão e, em consequência, ocorreram diferentes configurações e estruturas da adoção da política de municipalização no Brasil. Com isso, questões importantes, tais como a situação

---

<sup>4</sup> A educação básica obrigatória compreende três etapas: Ensino Infantil (crianças até 5 anos, compreendendo atendimento em creches, até 3 anos de idade, e em pré-escolas, de 4 a 5 anos de idade); Ensino Fundamental (com duração de 9 anos, obrigatório para crianças e adolescentes de 6 a 14 anos); e Ensino Médio (com duração de três anos, obrigatório para adolescentes de quinze a dezessete anos).

profissional dos docentes de uma escola municipalizada, não foram definidas e uniformizadas em todo o território nacional.

No Brasil, 1.115 escolas foram municipalizadas no período 2010-2021<sup>5</sup>. Esse histórico está apresentado no Gráfico 1 abaixo.



Fonte: Elaboração própria a partir de dados do Censo Escolar.

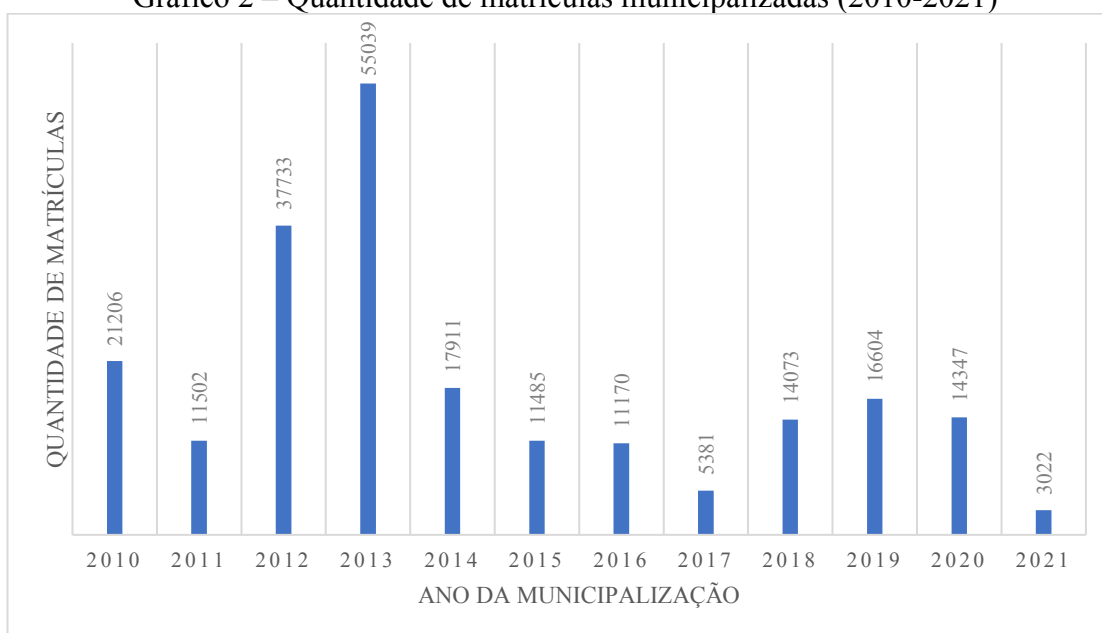
Tão importante quanto conhecer a quantidade de unidades de ensino que passaram pelo processo de municipalização é mensurar a quantidade de alunos envolvidos e sujeitos aos impactos dessa transformação organizacional.

O Gráfico 2 apresenta o total de matrículas municipalizadas por ano, enquanto o Gráfico 3 retrata o número médio de matrículas, isto é, considerando o total de matrículas e a quantidade de escolas municipalizadas.

Nota-se que, em termos médios, a quantidade de matrículas municipalizadas foi relativamente semelhante entre os anos analisados, o que não ocorre quando se analisa a quantidade total de matrículas municipalizadas.

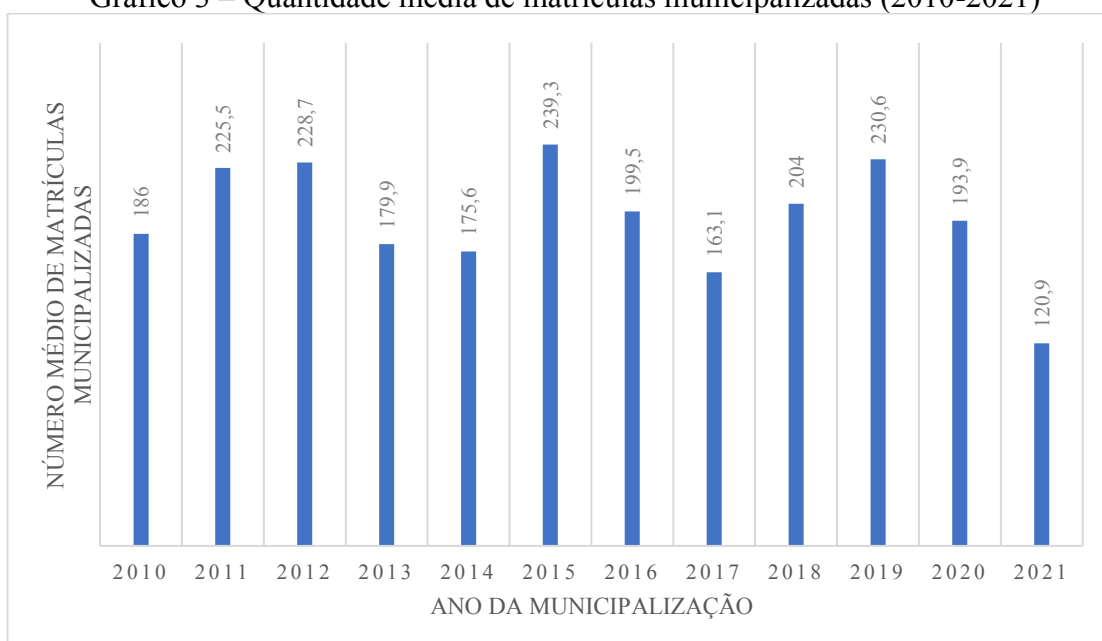
<sup>5</sup> As informações acerca da alteração da gestão escolar podem ser obtidas através da análise dos microdados do Censo Escolar.

Gráfico 2 – Quantidade de matrículas municipalizadas (2010-2021)



Fonte: Elaboração própria a partir de dados do Censo Escolar.

Gráfico 3 – Quantidade média de matrículas municipalizadas (2010-2021)



Fonte: Elaboração própria a partir de dados do Censo Escolar.

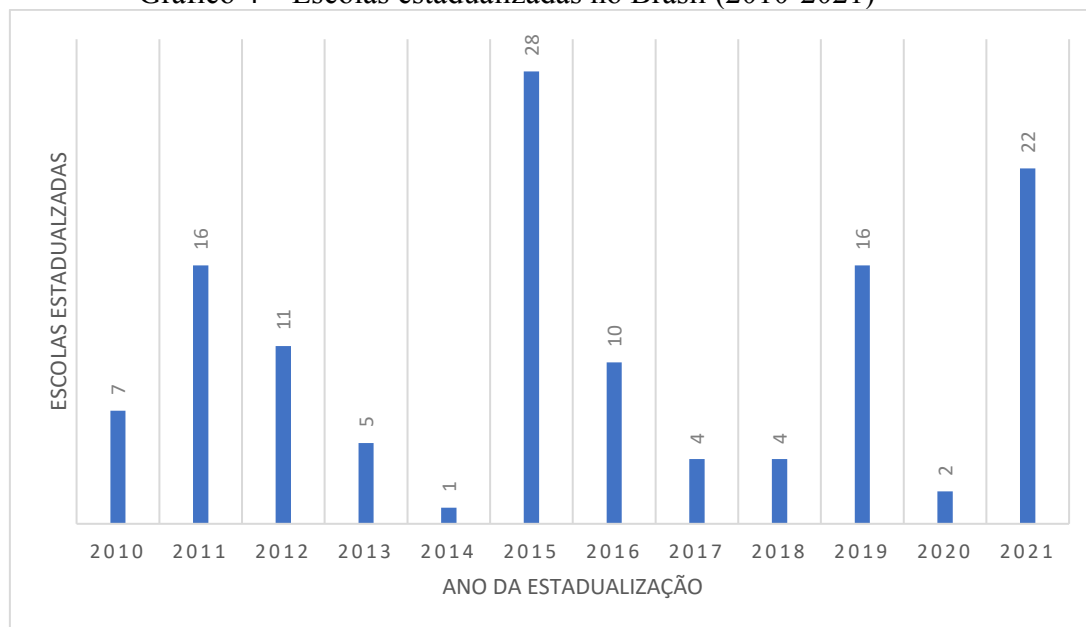
Em contrapartida, apesar dos instrumentos normativos determinarem a municipalização do ensino fundamental como uma diretriz, no período em análise 126 unidades escolares realizaram o movimento inverso, isto é, deixaram de ser municipais para pertencer à rede estadual. Em outras palavras, foram constatados casos de estadualização do ensino.

A ocorrência desse evento pode ser justificada em situações em que a prestação dos serviços educacionais pelo município esteja deficitária e em níveis precários de qualidade



educacional, o que ampara a intervenção do Estado e a oferta de ensino fundamental pela rede estadual.

Gráfico 4 – Escolas estadualizadas no Brasil (2010-2021)



Fonte: Elaboração própria a partir de dados do Censo Escolar.

Compreender o fenômeno da municipalização requer investigar a extensão da influência da gestão escolar na experiência educacional dos estudantes. Nesse sentido, na literatura existem estudos com a finalidade de mensurar o impacto da municipalização sobre o desempenho escolar dos alunos.

Leme, Paredes e Souza (2009), em seu estudo intitulado “A Municipalização do Ensino Fundamental e seu Impacto sobre a proficiência no Brasil”, objetivam avaliar as consequências da municipalização do ensino fundamental sobre os níveis de aprendizagem, considerando “qualidade da educação” como a proficiência de matemática e português dos alunos da 4ª série do ensino fundamental (atualmente denominado 5º ano) e “efeito da municipalização” como o impacto da gestão municipal sobre a proficiência média dos alunos. A metodologia utilizada é Diferenças-em-Diferenças com base nos resultados da prova Saeb dos anos 1997, 1999, 2001, 2003 e 2005 e da Prova Brasil ocorrida em 2005. Essa pesquisa não considera o tempo decorrido após a municipalização, isto é, o “efeito da municipalização” é estimado considerando-se o efeito médio nas escolas que foram municipalizadas sem mensurar o tempo em que a escola permaneceu sob a gestão municipal. A conclusão do estudo sugere que o efeito da municipalização sobre a proficiência dos alunos do 5º ano do ensino fundamental é praticamente nulo.

Outro estudo denominado “O Nível de Governo Importa para a Qualidade da Política Pública? A Municipalização da Educação Fundamental no Brasil” (Ceneviva, 2012) utiliza os resultados da Prova SAEB e da prova Brasil, além de informações do Censo Escolar dos anos de 1999 a 2005, em que são aplicadas técnicas de Mínimos Quadrados Ordinários e Efeitos Fixos. Os resultados sugerem que a expansão da oferta de ensino da rede municipal não impactou a proficiência dos alunos, isto é, o nível de governo não possui impactos diretos sobre o desempenho escolar, de modo que os efeitos da municipalização na proficiência média dos alunos são praticamente nulos, tanto em português como em matemática.

Em suma, em ambos os estudos foi constatado que o efeito da municipalização foi insignificante, isto é, os alunos pertencentes às escolas municipalizadas não apresentaram proficiências distintas das escolas estatisticamente semelhantes que permaneceram sob a gestão estadual.

Em contrapartida, uma pesquisa intitulada “Os Efeitos da Organização do 5º e 9º Ano do Ensino Fundamental nos Municípios sobre a Proficiência dos Alunos” da Fundação Itaú Social (2022)<sup>6</sup> examinou como os efeitos da oferta pelo Estado ou Município do ensino fundamental em escolas distintas afetam a nota dos alunos no Brasil, utilizando dados do Censo Escolar e da Prova Brasil para os anos 2007, 2011, 2015, 2019 além de dados presentes no Atlas do Desenvolvimento Humano nos Municípios. A metodologia utilizada foi modelo de dados em painel. Os resultados indicaram que a municipalização de todo o ensino fundamental trouxe benefícios à evolução da performance dos alunos do 5º e 9º ano nas notas de matemática e português.

Outro objetivo desta pesquisa foi examinar qual estrutura escolar é mais benéfica: ofertar Anos Iniciais e Finais na mesma unidade ou em escolas diferentes. Também foi possível concluir que é mais efetivo separar Anos Iniciais de Anos Finais em duas escolas municipais distintas, a fim de se evitar externalidades negativas. Em suma, os autores demonstraram que a municipalização teve um impacto positivo no desempenho escolar do município.

Contudo, a despeito da qualidade da gestão educacional, o fator de maior influência sobre o desempenho acadêmico dos estudantes e nas oportunidades educacionais de um indivíduo é o denominado *background familiar*, termo utilizado para designar o ambiente socioeconômico em que o estudante está inserido e composto por fatores como: nível

---

<sup>6</sup> O referido estudo está disponível em: <https://www.itausocial.org.br/wp-content/uploads/2022/09/Artigo-Os-efeitos-da-organizacao-do-5o-e-9o-ano-do-ensino-fundamental-nos-municipios-sobre-a-proeficiencia-dos-alunos.pdf>

educacional e situação financeira dos pais, recursos materiais disponíveis, estabilidade e estrutura familiar, dentre outros (D'ATRI, 2007).

Essas condições socioeconômicas apresentam potencial de afetar positivamente o desempenho dos alunos uma vez que aqueles estudantes inseridos em famílias com melhores condições de renda e com estrutura, em tese, terão oportunidade de se dedicar mais ao estudo, ter acesso a reforço escolar, além de possuírem infraestrutura e recursos pedagógicos que facilitam o processo de aprendizagem.

Contudo, o desempenho dos alunos também depende, em alguma medida, da realidade de sua escola, como por exemplo, infraestrutura escolar; recursos físicos, pedagógicos e humanos; jornada escolar; experiência e motivação dos professores, dentre outros fatores diretamente ligados à gestão escolar (LEME; PAREDES; SOUZA, 2009). Nesse sentido, é relevante investigar o impacto da alteração da gestão estadual e municipal sobre o nível de proficiência.

Alia-se a isso o pressuposto de que uma política pública educacional bem-sucedida é aquela eficiente independentemente do *background familiar*, isto é, capaz de mitigar a importância do contexto socioeconômico e familiar e de dar aos estudantes igualdade de oportunidades e equidade na possibilidade de desenvolvimento, constituindo-se em uma alavanca de transformação e ascensão social (D'ATRI, 2007).

Dessa forma, é imprescindível que se busque compreender a magnitude do impacto da gestão escolar sobre a aprendizagem, a fim de se fornecer subsídios para desenvolver políticas estratégicas capazes de aprimorar a educação ofertada aos alunos.

### 3 CONTEXTO METODOLÓGICO

O objetivo principal da maioria dos estudos empíricos no campo da Economia, bem como de outras Ciências Sociais, é determinar se alterações em uma variável específica exercem influência sobre outras variáveis de interesse. Como exemplo, pode-se citar o problema abordado por esta pesquisa, qual seja: identificar o impacto da alteração da gestão escolar sobre o desempenho dos estudantes.

A Econometria é uma disciplina que contempla métodos e instrumentos capazes de realizar essa investigação de forma apropriada e foi desenvolvida com objetivo de testar a validade de teorias econômicas, analisar os impactos de políticas públicas, prever séries de tempo e fornecer informações para o governo e setor privado (WOOLDRIDGE, 2006).

No contexto da avaliação de políticas públicas, os métodos econométricos ganham destaque e são essencialmente úteis. A partir da análise dos dados, pode-se estimar, explicar e avaliar os efeitos concretos de uma intervenção política o que permite identificar a forma como os diversos agentes a ela submetidos são impactados. Dessa forma, os referidos métodos auxiliam os formuladores de políticas públicas a desenvolverem políticas estratégicas capazes de atingir o público-alvo determinado, além de embasar as decisões subsequentes.

Contudo, atribuir causalidade a um tratamento não é trivial no âmbito das Ciências Sociais, uma vez que esses problemas envolvem maiores graus de complexidade e diversos fatores influenciam os resultados simultaneamente.

Uma das vantagens dos métodos econométricos é que possibilitam simular experimentos *ceteris paribus*, isto é, permitem isolar variáveis de interesse e manter constantes todos os demais fatores relevantes a fim de investigar os efeitos de alterações específicas. Se essa condição não ocorre, isto é, se outras variáveis importantes não são mantidas fixas, então não é possível conhecer o efeito causal de uma variável sobre a outra (WOOLDRIDGE, 2006).

Dessa forma, simular um experimento *ceteris paribus* é crucial para a avaliação de políticas públicas e sua importância decorre da possibilidade de examinar o impacto de determinada variável no resultado geral, o que é de fundamental importância em estudos que visam identificar relações de causa e efeito considerando diferentes cenários. É, portanto, especialmente relevante para a formulação e compreensão de políticas públicas, uma vez que ajuda a prever e entender consequências de mudanças específicas em variáveis-chave.

Em Econometria, o modelo de Diferenças em Diferenças (Difference-in-Differences - DiD) é um dos métodos mais populares para realizar estudos e análises de efeitos causais.

O modelo canônico de DiD é configurado de modo que existem duas populações estatisticamente comparáveis, denominadas tratamento e controle, observadas em dois períodos de tempo distintos: antes e depois do tratamento. O grupo de tratamento recebe o tratamento ao mesmo tempo a partir do segundo período, enquanto o grupo de controle não o recebe. Essa configuração considera a existência de um número significativo de observações independentes (clusters) e permite o desenvolvimento de um estudo comparativo, de modo que o estimador DiD decompõe a evolução dos resultados médios obtidos entre os dois grupos.

Deve-se ressaltar que o grupo de controle exerce papel fundamental no referido modelo, uma vez que é a representação do contrafactual para a população tratada, ou seja, é a referência de como esta teria se desenvolvido na ausência do tratamento – daí a exigência dos grupos serem estatisticamente comparáveis.

Nesse cenário, a estimativa DiD será a diferença entre os resultados médios pós e pré-tratamento no grupo de tratamento *versus* grupo de controle, isto é, será o cálculo de uma dupla diferença. O benefício dessa configuração é que o estimador DiD explicita quais comparações são utilizadas para gerar a estimativa (GOODMAN-BACON, 2021).

O estimador DiD é sintetizado pela equação abaixo que é expressa em termos de médias amostrais  $\bar{y}$ :

$$DiD = (\bar{y}_{pós;tratamento} - \bar{y}_{pré;tratamento}) - (\bar{y}_{pós;controle} - \bar{y}_{pré;controle}) \quad (1)$$

Em suma, o modelo canônico apresenta uma configuração 2x2 (dois períodos e dois grupos) e tem por objetivo identificar o efeito médio do tratamento na subpopulação tratada (“Average Treatment Effect on the Treated” - ATT). Portanto, o ATT compara a mudança média nos resultados ocorrida no grupo de tratamento com a mudança média dos resultados ocorrida para o grupo de controle. Para isso, algumas hipóteses centrais são admitidas, dentre elas: hipótese de tendências paralelas e não antecipação do tratamento.

De acordo com a suposição de tendências paralelas, o resultado médio entre as duas populações de interesse (tratada e de controle) teria evoluído em paralelo na ausência do tratamento. Em outras palavras, esse pressuposto assume que, diante da ausência do tratamento, a diferença entre a população tratada e controle seria constante ao longo do tempo. A validade do modelo de diferenças em diferenças depende da hipótese de tendências paralelas (ROTH et al, 2022).

Já a hipótese de não antecipação do tratamento implica que o tratamento não gera efeito causal no período anterior à sua implementação, isto é, não é possível antecipar o efeito do tratamento.

O desafio principal do estimador DiD reside no fato de não ser possível observar a trajetória real do grupo de tratamento na ausência do tratamento, isto é, não é possível identificar os resultados do grupo de tratamento caso não fosse tratado. Sob as hipóteses de tendências paralelas e não antecipação do tratamento, é possível superar este problema, uma vez que ao assumir as referidas hipóteses, a metodologia DiD imputa os resultados médios contrafactuais do grupo de controle para o grupo tratado, de forma que o ATT pode ser identificado (ROTH et al, 2022).

Conforme apresentado por Goodman-Bacon (2021), o estimador DiD padrão, isto é, ATT, é numericamente equivalente ao estimador obtido através da especificação de regressão denominada “Two-Way Fixed Effects” (TWFE), se houver dois períodos e o tratamento for

aplicado a uma parcela das unidades apenas no segundo período. Na tradução, esta metodologia denomina-se "efeitos fixos bidirecionais" porque inclui tanto efeitos fixos no tempo quanto efeitos fixos de unidade.

Esta relação é apresentada pela equação abaixo, que regride o resultado  $Y_{i,t}$  em um efeito fixo individual  $\alpha_i$ , um efeito fixo no tempo  $\phi_t$ , e um indicador demonstrando se a unidade  $i$  é tratada no período  $t$ ,  $D_{i,t}$ . Nessa configuração canônica, o estimador de Mínimos Quadrados Ordinários  $\beta$  da regressão TWFE é equivalente ao ATT (ROTH et al, 2022). Esta configuração também é denominada TWFE estático.

$$Y_{i,t} = \alpha_i + \phi_t + D_{i,t}\beta + \epsilon_{it} \quad (2)$$

Adicionalmente, pesquisas recentes demonstraram que este estimador TWFE é não viesado para o ATT se os efeitos do tratamento forem constantes entre os grupos e em relação ao tempo. Em outras palavras, a especificação estática produz uma estimativa razoável quando não há heterogeneidade nos efeitos do tratamento ao longo do tempo ou das unidades, isto é, todas as unidades têm o mesmo efeito de tratamento e o tratamento tem o mesmo efeito independentemente de quanto tempo tenha se passado desde o início do tratamento. Para a validade desse resultado, é necessário destacar que as hipóteses de tendências paralelas e não-antecipação devem ser combinadas com a hipótese de amostragem independente (ROTH et al, 2022).

Contudo, a maioria das aplicações empíricas do estimador de DiD divergem da configuração canônica 2x2 porque os tratamentos geralmente ocorrem em momentos diferentes. É comum, por exemplo, que governos apliquem políticas públicas em diferentes momentos e que legislações e decisões judiciais afetem indivíduos de forma diferente e em períodos distintos (GOODMAN-BACON, 2021).

Dessa forma, torna-se necessário analisar o cenário caracterizado por adoção escalonada de tratamento, isto é, quando as unidades podem receber o tratamento de interesse em períodos distintos (ou seja, ocorre "temporalidade diferenciada" do tratamento). Para examinar esse caso, deve-se considerar o tratamento como um estado "absorvente", o que significa que, uma vez que a unidade recebe o tratamento, ela permanece tratada.

Conforme apresentado por Roth et al (2022), a suposição de tendências paralelas canônica pode ser estendida para a configuração escalonada, de modo que seja válida para todas as combinações de períodos e todas as combinações de grupos tratados em momentos

diferentes, o que implica que, no cenário em que não houve a aplicação do tratamento, os resultados para todos os grupos apresentariam trajetórias paralelas.

De igual modo, os referidos autores demonstraram que a hipótese de não antecipação também pode ser ampliada para o caso escalonado. Nesse caso, uma unidade apenas apresenta efeitos do tratamento a partir do momento em que é submetida a ele. Desse modo, se uma unidade não recebe o tratamento em determinado período, sua trajetória de resultados independe de quando o receberá no futuro (ROTH et al, 2022).

Contudo, o resultado obtido anteriormente acerca da equivalência entre o ATT e o estimador linear de regressão TWFE não se generaliza para a configuração em que o tempo de aplicação do tratamento é variável (GOODMAN-BACON, 2021).

Em decorrência da temporalidade diferenciada do tratamento, existem inúmeras configurações de estimadores 2x2 que podem ser formados. Para exemplificar essa ideia e considerando o cenário mais simples depois da configuração canônica, isto é, em que há dois grupos submetidos ao tratamento em momentos distintos e um grupo que nunca é tratado, é possível resumir três estimadores 2x2 simples relevantes: o primeiro compara qualquer grupo tratado com o grupo nunca tratado; o segundo compara o grupo tratado preliminarmente com o grupo que ainda não foi tratado; e o terceiro compara o grupo tratado posteriormente com grupo tratado preliminarmente (caso em que as unidades tratadas atuam como grupo de controle) (CUNNINGHAM, 2021).

Nesse caso, conforme comprovado por Goodman-Bacon (2021), o estimador de diferenças em diferenças calculado através da regressão TWFE constitui-se em uma média ponderada de todas as combinações de estimadores DiD possíveis configurados na forma 2x2.

Nesse bojo, incluem-se combinações permitidas, que utilizam unidades tratadas em um determinado momento como grupo de tratamento e unidades não tratadas como grupo de controle, e combinações proibidas, que utilizam unidades tratadas preliminarmente como grupo de controle e unidades tratadas posteriormente como grupo de tratamento. Em síntese, o problema decorre da utilização de unidades já tratadas que, a depender de quando recebem o tratamento, atuam como grupo de tratamento ou de controle.

Nesse caso, os pesos nos estimadores DiD 2x2 serão proporcionais ao tamanho dos grupos e o efeito médio será ponderado pela variância do efeito do tratamento nos tratados. Portanto, diante de eventos caracterizados por temporalidade diferenciada, o tempo que um grupo passa sob tratamento é determinante para sua variância de tratamento, fator que impacta o peso que este estimador simples 2x2 assumirá na soma final. Dessa forma, quando a metodologia TWFE é utilizada para estimar o ATT no cenário escalonado, o resultado obtido é

influenciado diretamente pelo tamanho do painel, de modo que modificá-lo adicionando ou subtraindo anos pode alterar a estimativa final obtida apenas por alterar a variância do tratamento nos grupos (CUNNINGHAM, 2021).

Em um cenário caracterizado por homogeneidade dos efeitos do tratamento, isto é, se os efeitos do tratamento não mudam ao longo do tempo, o estimador TWFE produzirá uma média ponderada pela variância dos efeitos do tratamento entre os grupos e todos os pesos serão positivos. Contudo, diante de heterogeneidade nos efeitos de tratamento, isto é, no caso de os efeitos médios do tratamento variarem ao longo do tempo, podem ocorrer pesos negativos em virtude das unidades tratadas primariamente comporem o grupo de controle e as alterações geradas pelo tratamento serem subtraídas (GOODMAN-BACON, 2021).

Em outras palavras, diante da possibilidade dos efeitos de tratamento serem heterogêneos, as comparações denominadas “proibidas” resultam em possíveis erros devido a problemas de “ponderação negativa” o que permite concluir que o estimador TWFE é viesado em eventos caracterizados por adoção diferenciada do tratamento e efeitos de tratamento heterogêneos (ROTH et al, 2022).

O efeito torna-se viesado porque o grupo de controle experimenta uma tendência nos resultados (em virtude dos efeitos heterogêneos), e esse viés se propaga para os estimadores 2x2 calculados posteriormente (CUNNINGHAM, 2021).

Como resultado de sua análise, Goodman-Bacon (2021) apresenta o Teorema da Decomposição de Diferenças-em-Diferenças a fim de demonstrar o estimador TWFE como uma média ponderada de todos os possíveis estimadores DiD na configuração simples 2x2:

*Theorem (Difference-in-Differences Decomposition Theorem). Assume that the data contain  $k = 1, \dots, K$  timing groups of units ordered by the time when they receive a binary treatment,  $k \in (1, T]$ . There may be one timing group,  $U$ , that includes units that never receive treatment. The OLS estimate,  $\hat{\beta}^{DD}$ , in a two-way fixed-effects regression ( $y_{it} = \alpha_i + \alpha \cdot t + \beta^{DD} D_{it} + \text{eit.}$ ) is a weighted average of all possible two-by-two DD estimators. (Goodman-Bacon, 2021, p. 257)*

Além do Teorema da Decomposição, outros autores também abordam a inaptidão deste estimador para analisar cenários que envolvem múltiplos períodos de tempo e heterogeneidade do efeito do tratamento.

Em síntese, o problema da utilização da metodologia TWFE para o caso escalonado é que, ainda que as hipóteses fundamentais de tendências paralelas e não-antecipação possam ser estendidas para essa configuração, ela permite estimadores de Diferenças-em-Diferenças que utilizam como grupo de controle unidades já tratadas em períodos anteriores e que podem obter



pesos negativos quando os efeitos do tratamento são heterogêneos ao longo do tempo ou das unidades, potencialmente induzindo a um viés. Nesse caso, a estimativa de TWFE pode ser negativa, ainda que o efeito de tratamento tenha sido positivo para todas as unidades.

A fim de sanar as dificuldades encontradas na utilização do estimador TWFE bem como adequar o estimador DiD a situações cotidianas, atualmente existe um debate na literatura de Diferenças-em-Diferenças em que alguns pesquisadores sugerem flexibilizações nas hipóteses centrais do modelo canônico.

Roth et al (2022) classificam três vertentes principais de novas metodologias a depender de qual hipótese do modelo canônico está sendo flexibilizada: múltiplos períodos de tempo e adoção escalonada do tratamento; violações da hipótese de tendências paralelas; e estruturas alternativas de inferência (considerando um pequeno número de clusters).

A primeira vertente da metodologia DiD inclui a análise de múltiplos períodos de tempo e variação no tempo de tratamento, caso em que existem mais de dois períodos e as unidades recebem o tratamento em diferentes momentos. Uma das alternativas é a proposta apresentada por Callaway e Sant’Anna (2021) que consiste em utilizar métodos que especifiquem as comparações a serem feitas entre os grupos de tratamento e controle, bem como os pesos desejados no parâmetro-alvo, de forma a permitir a estimação de um parâmetro causal bem definido, com pesos e grupos de comparação transparentes eliminando as comparações indesejáveis identificadas na metodologia TWFE para o caso escalonado (ROTH et al, 2022).

A segunda vertente flexibiliza a hipótese de tendências paralelas. Em muitos casos, a verificação e a validade dessa suposição é um aspecto crítico durante a análise por ela permitir fatores de confusão que variam no tempo e que afetam o estado do tratamento. Adicionalmente, deve-se destacar que essa suposição é sensível à forma da função selecionada, isto é, as tendências paralelas são sensíveis às alterações funcionais. Nesse caso, uma alternativa é utilizar a configuração em que a hipótese de tendências paralelas é válida apenas condicionada à covariáveis observáveis. A hipótese canônica de tendências paralelas pode ser ampliada para condicionada à covariáveis observáveis, sendo necessário incluir um vetor de covariáveis pré-tratamento e impondo uma condição de sobreposição a fim de garantir que para cada unidade tratada com as covariáveis definidas, existem pelo menos algumas unidades não tratadas da população com os mesmos valores das covariáveis selecionadas.<sup>7</sup> (ROTH et al, 2022)

---

<sup>7</sup> Nesse caso, é preciso garantir que as covariáveis sejam medidas antes do tratamento e, portanto, não sejam afetadas por ele. No caso em que as covariáveis podem ser afetadas pelo tratamento, ocorre um problema denominado “mau controle” que pode enviesar o estimador (ROTH et al, 2022).

A terceira vertente trata de suposições alternativas de amostragem, permitindo analisar cenários com um pequeno número de clusters tratados. A suposição canônica justifica a aleatoriedade nos dados em virtude da amostra ser obtida a partir de um grande número de clusters independentes de uma superpopulação, considerando, para tanto, o Teorema do Limite Central<sup>8</sup>. Contudo, em casos em que há poucos clusters tratados, este teorema pode não ser aplicável e a metodologia canônica pode não apresentar resultados intuitivos. Uma alternativa, por exemplo, é a realização de inferência “baseada em design” para DiD, que considera a população como fixa e a atribuição do tratamento como estocástica. Dessa forma, a incerteza dos dados seria justificada pela aleatoriedade na atribuição do tratamento (ROTH et al, 2022).

De acordo com a breve apresentação acima, foram identificados problemas na utilização do modelo denominado “*Two-way Fixed Effects*” envolvendo cenários caracterizados por temporalidade diferenciada e efeitos de tratamento heterogêneos ao longo do tempo em virtude de estimadores simples 2x2 que comparam unidades tratadas posteriormente com unidades tratadas primariamente.

O Teorema da Decomposição de Goodman-Bacon demonstrou essa problemática e se consolidou como um fator de extrema relevância na compreensão da referida metodologia uma vez que sinalizou a necessidade do desenvolvimento de novos métodos capazes de superar os problemas apontados.

Nesse contexto, a literatura atual de Diferenças-em-Diferenças se expandiu a partir da flexibilização das hipóteses centrais do modelo canônico e esses novos métodos passaram a permitir a análise de configurações específicas sendo capazes de melhorar a validade, razoabilidade e interpretabilidade dos resultados obtidos a partir da aplicação do método DiD.

### 3.1. A ABORDAGEM CALLAWAY E SANT’ANNA (2021)

Conforme apresentado no tópico anterior, muitas aplicações atuais de DiD objetivam flexibilizar hipóteses da configuração canônica e desenvolver estruturas capazes de contemplar as especificidades de casos concretos apresentando resultados mais robustos e consistentes com a realidade.

---

<sup>8</sup> O Teorema do Limite Central é um princípio fundamental da teoria estatística que afirma que, quando se realiza uma grande quantidade de amostras aleatórias de uma população, a distribuição das médias dessas amostras se aproxima de uma distribuição normal, independentemente da distribuição da população original, desde que o tamanho da amostra seja suficientemente grande.

Callaway e Sant'Anna (2021) apresentam uma extensão da metodologia de DiD que compreende a análise de múltiplos períodos de tempo, variação no tempo de tratamento e quando a suposição de tendências paralelas se mantém válida somente se condicionada a covariáveis observáveis. Além disso, consideram que as unidades adotam o tratamento de interesse em determinado momento e permanecem expostas a ele ao longo dos períodos posteriores.

Nessa metodologia, somente são consideradas covariáveis no período anterior à aplicação do tratamento e que potencialmente são capazes de afetar a evolução da variável de interesse nos períodos posteriores. Desse modo, não são consideradas covariáveis pós-tratamento, uma vez que poderiam apresentar efeitos do tratamento.

Diferentemente da configuração tradicional de DiD, na metodologia proposta pelos referidos autores, o  $ATT_{(g,t)}$  reflete o efeito médio do tempo de tratamento na subpopulação tratada, ou seja, o efeito médio do tratamento para um grupo específico “g” em determinado período “t”. Nessa estrutura, um grupo é definido pelo momento em que as unidades são tratadas pela primeira vez. Além disso, os autores consideram duas possibilidades para composição do grupo de controle: unidades que nunca são tratadas e unidades que ainda não foram tratadas - o que adiciona transparência e objetividade à análise.

Como o foco é o cálculo de ATTs específicos de grupos por tempo, em algumas aplicações, pode haver diversos parâmetros de efeito de tratamento. Em virtude disso, os autores apresentam como uma das etapas da metodologia a agregação de todos esses efeitos de tratamento colapsando-os em diferentes medidas resumidas de efeito causal mais interpretáveis e intuitivos, como um ATT maior e representativo dos demais (CUNNINGHAM, 2021).

O destaque do modelo de Callaway e Sant'Anna é a busca pela superação dos problemas fundamentais identificados na utilização do estimador TWFE: o modelo desenvolvido pelos autores é capaz de destacar as diferentes fontes de heterogeneidade nos efeitos do tratamento entre grupos e períodos de tempo. A partir da análise das medidas agregadas, o método permite analisar como, em média, o efeito do tratamento varia de acordo com o tempo de exposição ao tratamento, como os efeitos médios do tratamento variam entre grupos e qual é o efeito cumulativo médio do tratamento em todos os grupos até determinado ponto no tempo (CALLAWAY; SANTA'ANNA, 2021).

Para tanto, é possível utilizar diferentes métodos de estimação a depender das características do caso em estudo: regressões de resultados, ponderação de probabilidade inversa e métodos duplamente robustos – que resultarão em diferentes tipos de estimadores.

Adicionalmente, também é permitido considerar a possibilidade ou não de antecipação do efeito de tratamento.

Essa abordagem alternativa possui duas vantagens principais sobre as regressões TWFE padrão (estática ou com temporalidade diferenciada de tratamento). A primeira vantagem é que fornece estimativas razoáveis mesmo diante de eventos caracterizados por heterogeneidade dos efeitos do tratamento. A segunda é que explicita de maneira transparente exatamente quais unidades estão sendo utilizadas para comporem o grupo de controle utilizado para inferir os resultados potenciais que não são observados. Isso contrasta com os modelos TWFE padrão, que fazem comparações não intuitivas no caso escalonado (ROTH et al, 2022).

Os autores apresentam um procedimento que contempla 3 fases distintas: identificação de parâmetros relevantes; agregação desses parâmetros e formação de medidas resumidas dos efeitos causais (agregando diferentes efeitos médios de tempo de tratamento); e estimação e inferência sobre esses diferentes parâmetros.<sup>9</sup>

Na etapa de identificação dos parâmetros relevantes, o ponto central é a análise do  $ATT_{(g,t)}$ , isto é, o parâmetro causal desagregado denominado “efeito médio de tratamento para grupos-tempo”, que, como afirmado anteriormente, é o efeito médio de tratamento para um grupo específico em um ponto determinado no tempo, em que um “grupo” é definido pelo período de tempo em que as unidades são tratadas pela primeira vez.

Este parâmetro é uma generalização do efeito médio do tratamento sobre os tratados (ATT) calculado sob a configuração canônica de DiD com dois períodos de tempo e duas populações e não impõe restrições à heterogeneidade do efeito do tratamento entre grupos ou ao longo do tempo, o que permite analisar como os efeitos médios do tratamento variam em diferentes dimensões.

Para a identificar o  $ATT_{(g,t)}$ , os autores adotam as seguintes hipóteses:

1. Irreversibilidade do tratamento: a unidade recebe o tratamento e permanece tratada nos períodos subsequentes;
2. Amostragem aleatória: cada unidade é selecionada aleatoriamente de uma grande população de interesse;
3. Antecipação limitada do tratamento: configura uma restrição à antecipação do tratamento para os grupos “eventualmente tratados” (quanto maior for a antecipação do tratamento permitida, maior deverá ser o retrocesso temporal na análise);

---

<sup>9</sup> Além disso, os autores disponibilizam gratuitamente um pacote no software R (“R did” package) com os códigos utilizados para o cálculo dos estimadores agregados – o que torna mais fácil a aplicação dos métodos propostos.

4. Tendências paralelas condicionais considerando um grupo de unidades que nunca são tratadas compondo o grupo de controle;
5. Tendências paralelas condicionais considerando um grupo de unidades que ainda não foram tratadas compondo o grupo de controle;
6. Extensão da hipótese de sobreposição para a configuração de múltiplos grupos e períodos.

As suposições 4 e 5 são versões alternativas de tendências paralelas condicionais e generalizam a hipótese canônica para o caso em que há vários períodos e múltiplos grupos de tratamento. A suposição 4 considera que, condicionado às covariáveis, os resultados médios para o grupo de tratamento e para o grupo de controle formado por unidades “nunca tratadas” seguiriam trajetórias paralelas diante da ausência de tratamento. Nesse caso, é possível utilizar um único grupo de comparação fixo para todas as unidades que eventualmente serão tratadas. Já a suposição 5 considera a validade de tendências paralelas condicionais entre o grupo tratado e o grupo de controle formado por unidades “ainda não tratadas”.

Deve-se destacar que a possibilidade de optar entre as suposições 4 e 5 é importante principalmente em aplicações em que existem tendências específicas de covariáveis nos resultados ao longo do tempo e quando a distribuição de covariáveis é diferente entre grupos - o que pode resultar em vieses significativos (CALLAWAY; SANTA'ANNA, 2021).

Sob as seis suposições elencadas, é possível realizar a identificação dos efeitos médios do tratamento grupo-tempo a partir da equação abaixo:

$$ATT_{(g,t)} = E \left[ \left( \frac{G_g}{E[G_g]} - \frac{\frac{P_g(X)C}{1-P_g(X)}}{E\left[\frac{P_g(X)C}{1-P_g(X)}\right]} \right) (Y_t - Y_{g-1}) \right] \quad (3)$$

Esse problema é configurado de forma que: supondo que haja T períodos em que  $t = 1, \dots, T$ , sendo  $G_g$  uma variável binária igual a 1 quando uma unidade é tratada pela primeira vez no período e  $C$  uma variável binária igual a 1 para unidades nunca tratadas (podem ser utilizadas unidades ainda não tratadas). Para cada unidade, ou  $\{G_1, \dots, G_T\}$  ou  $C$  será igual a 1. O escore de propensão generalizado é dado por  $P_g(X) = P(G_g = 1 | X, G_g + C = 1)$ , sendo a probabilidade de uma unidade ser tratada dado que possui covariáveis  $X$ .  $Y$  é a variável resultado. Dessa forma, é possível identificar o efeito médio do tratamento no grupo-tempo  $ATT_{(g,t)}$ .

Adicionalmente, um dos principais resultados encontrados pelos autores é que, para encontrar os valores do  $ATT_{(g,t)}$ , pode-se utilizar diferentes estimadores:

- Regressão de resultados, que depende da modelagem correta da expectativa condicional da evolução do resultado para os grupos de comparação;
- Ponderação de Probabilidade Inversa, que evita modelar a evolução do resultado do grupo de comparação, mas depende da modelagem correta da probabilidade condicional de uma unidade pertencer ao grupo tratado dado suas covariáveis e que esteja no grupo tratado ou no grupo de controle apropriado; e
- Estimadores Duplamente Robustos, uma abordagem que combina componentes dos dois métodos anteriores, uma vez que depende tanto da modelagem da evolução do resultado do grupo de comparação quanto da propensão de pertencer ao grupo de tratamento. Destaca-se, no entanto, que é exigido apenas que se especifique corretamente a evolução do resultado para o grupo de comparação ou o modelo de escores de propensão, mas não necessariamente ambos. Esse fator confere um grau de robustez adicional.

Os estimadores calculados sob essas diferentes formas não se diferenciam na etapa de identificação, contudo, isso não ocorre quando se deseja estimar e fazer inferências sobre o  $ATT_{(g,t)}$ , caso em que resultam em diferentes tipos de estimadores DiD.

Dessa forma, é possível identificar os  $ATTs_{(g,t)}$  e desenvolver análises a fim de compreender a heterogeneidade dos efeitos do tratamento em diferentes grupos, momentos e períodos de exposição ao tratamento. Nesse caso, os  $ATTs_{(g,t)}$  identificados são os parâmetros causais fundamentais de interesse.

Contudo, a metodologia proposta por Callaway e Sant'Anna (2021) contempla outra fase que permite agregar os  $ATTs_{(g,t)}$  identificados e formar parâmetros causais resumidos.

Com o propósito de resumir os efeitos médios de tratamento identificados e combiná-los de diferentes formas, os autores apresentam formas alternativas de agregação resultando em diversas medidas resumidas dos efeitos causais. De modo geral, são realizadas agregações parciais dos efeitos médios de tratamento grupo-tempo para sintetizar diferentes dimensões da heterogeneidade nos efeitos do tratamento. O objetivo é compreender três medidas fundamentais:

1. A variação dos efeitos médios do tratamento em relação ao tempo de exposição ao tratamento: com o propósito de destacar a heterogeneidade nos efeitos do tratamento em relação ao tempo de exposição a fim de compreender se os efeitos médios aumentam ou diminuem com o tempo decorrido desde o início do tratamento;
2. A variação dos efeitos médios do tratamento entre os diferentes grupos: com o objetivo de mensurar se grupos tratados primariamente possuem, em média, efeitos de tratamento mais altos ou baixos do que os grupos tratados posteriormente, isto é, evidenciar efeitos heterogêneos entre os grupos;
3. O efeito médio cumulativo do tratamento em todos os grupos até um período determinado, útil para destacar a heterogeneidade nos efeitos de tratamento em relação ao tempo calendário.

Além disso, é possível encontrar um único parâmetro geral que sumariza o efeito da participação no tratamento. Pode-se calcular uma média ponderada de cada  $ATT_{(g,t)}$  (calculado em todos os períodos de tempo), atribuindo mais peso aos  $ATT_{(g,t)}$  com tamanhos de grupo maiores. Nesse caso, a interpretação desse parâmetro corresponde a do ATT na configuração canônica de dois períodos e dois grupos. Ao contrário da especificação de regressão TWFE, essa combinação elimina a ocorrência de pesos negativos de modo que, se o efeito de participar do tratamento é positivo para todas as unidades, então o parâmetro agregado não será negativo.

De modo geral, nenhum desses parâmetros serão iguais, a não ser o caso em que o  $ATT_{(g,t)}$  seja idêntico para todos os grupos e todos os períodos de tempo.

A etapa de estimação e inferência ocorre com a escolha do tipo de abordagem que deva ser utilizada: regressão de resultado, ponderação de probabilidade inversa ou estimadores duplamente robustos. Conforme apresentado, ainda que essas abordagens sejam equivalentes na etapa de identificação, isso não ocorre na etapa de estimação/inferência. De modo geral, pode-se afirmar que os estimadores DiD calculados utilizando estimadores duplamente robustos geralmente apresentam maior robustez contra especificações incorretas do modelo do que os demais métodos.

Dessa forma, sob as suposições identificadoras assumidas e considerando cenários caracterizados por heterogeneidade nos efeitos de tratamento e temporalidade diferenciada do tratamento, a escolha do método de estimação poderá levar a conclusões diferentes.

Em suma, Callaway e Sant'Anna apresentam uma nova vertente da metodologia Diferenças-em-Diferenças possível de ser utilizada em casos em que há mais de dois períodos

e as unidades são tratadas em momentos diferentes - uma configuração mais próxima aos experimentos cotidianos. Nessa metodologia, o ATT é considerado um parâmetro de efeito de tratamento bem definido podendo ser agregado em parâmetros resumidos que detalham a heterogeneidade dos efeitos em relação a diferentes dimensões de interesse ou em um único parâmetro geral. Com isso, diferencia-se substancialmente da metodologia TWFE, considerada a mais popular para analisar efeitos causais, e os resultados obtidos apresentam uma interpretação mais intuitiva.

#### **4 BASE DE DADOS E PROCEDIMENTOS METODOLÓGICOS**

Com objetivo de analisar o impacto da política de municipalização de escolas sobre o desempenho escolar dos alunos no período de 2010 a 2021, é essencial compreender que os processos de municipalização ocorridos no Brasil apresentaram características peculiares e que não houve uma diretriz unificada que determinasse a forma como a transferência de gestão das unidades escolares deveria ocorrer.

Ademais, a municipalização ocorreu em múltiplos períodos de tempo, de modo que as observações entraram no tratamento em momentos distintos ocorrendo a formação de diferentes grupos de tratamento. Dessa forma, tornou-se necessário investigar os impactos dessa política através de uma metodologia que considere essa característica de maneira apropriada.

Conforme apresentado nos tópicos anteriores, não é adequado utilizar a metodologia TWFE para investigar o problema em análise visto que poderá produzir resultados enviesados. A alternativa será utilizar a metodologia de Diferenças-em-Diferenças mais robusta para o cenário de múltiplos períodos de tempo proposta por Callaway e Sant'Anna (2021).

A título de comparação, serão calculados os estimadores utilizando tanto a metodologia TWFE quanto a de Callaway e Sant'Anna a fim de confrontar os diferentes resultados obtidos.

Nesta pesquisa, o grupo de tratamento será formado pelas escolas estaduais que se tornaram municipais ao longo do período em análise, ou seja, foram municipalizadas, e o grupo de controle será composto por escolas estaduais que permaneceram sob gestão estadual.

Para que a comparação entre os grupos de tratamento e de controle seja possível, é fundamental que as escolas selecionadas sejam estatisticamente semelhantes, ou seja, possuam características equivalentes. Para tanto, será necessário utilizar técnicas de pareamento. Os critérios dessa seleção considerarão: contexto socioeconômico dos alunos, questões geográficas, quantidades de matrículas e de docentes, infraestrutura da escola, dentre outros.



Dessa forma, será possível estruturar um estudo baseado na metodologia atual e adequada de Diferenças-em-Diferenças que permita mensurar os efeitos da política pública da municipalização no indicador de proficiência considerando as características específicas desse processo.

#### 4.1 BASE DE DADOS

Os indicadores de desempenho escolar são instrumentos que permitem analisar qualitativa e quantitativamente uma escola, tais como o índice de proficiência, taxa de aprovação, taxa de abandono, distorção de idade/série, entre outros. Nesta pesquisa, desempenho escolar será mensurado através de resultados da prova Saeb (Sistema de Avaliação da Educação Básica), isto é, pelo índice nacional de proficiência que avalia alunos de 2º (apenas amostral), 5º e 9º anos do Ensino Fundamental e da 3ª série do Ensino Médio. O Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) é responsável por elaborar e aplicar a prova, bem como divulgar os resultados obtidos.

A fim de avaliar a qualidade da educação ofertada aos estudantes a nível nacional, a prova Saeb atualmente é aplicada a cada dois anos com cobertura censitária das escolas públicas e amostral das privadas<sup>10</sup>. São avaliados conhecimentos básicos em Língua Portuguesa, Matemática e Ciências Humanas e da Natureza - estas últimas apenas amostrais. Trata-se de um exame padronizado que utiliza a Teoria de Resposta ao Item (TRI), de modo que a pontuação final obtida pelo estudante é explicada pelas características das perguntas que respondeu corretamente e não pelo número total de acertos, portanto o que se avalia é a coerência das respostas dos estudantes.

Além disso, a prova Saeb é elaborada a partir de matrizes de referência desenvolvidas pelo Inep, que são instrumentos norteadores para a construção de itens avaliados e alinhados à legislação educacional e ao entendimento de professores, pesquisadores e especialistas acerca das competências e habilidades consideradas essenciais em cada etapa da educação básica.

---

<sup>10</sup> A Portaria nº 250, de 5 de julho de 2021, regulamentou a prova Saeb 2021. O referido normativo estabeleceu os critérios definidos para seleção das escolas públicas e privadas que participaram da avaliação, qual seja: no mínimo 10 alunos matriculados na série a ser avaliada. Outros critérios foram definidos para a publicação dos resultados: no mínimo 10 alunos presentes no dia da prova representando, no mínimo, 80% dos alunos matriculados. Contudo, conforme Portaria nº 399/2022, posteriormente esse percentual foi reduzido para 50%. Com isso, foram definidas quais escolas participaram do sistema de avaliação e tiveram seus resultados divulgados.

Outra característica importante desta avaliação ser ressaltada é a denominada “Escala Saeb”<sup>11</sup>. Trata-se de uma “régua” desenvolvida com base nos parâmetros estabelecidos para os itens aplicados na prova Saeb que relaciona a pontuação a níveis de entendimento (por exemplo: insuficiente, básico, proficiente ou avançado). Costuma ser apresentada como uma escala de 0 a 500 pontos, mas que, na prática, valores acima de 450 e abaixo de 50 dificilmente ocorrem.

Dessa forma, cada disciplina é avaliada de acordo com uma matriz de referência própria, e as notas obtidas são caracterizadas em uma escala que permite compreender as competências e habilidades alcançadas pelos alunos em cada faixa de desempenho.

Nesta pesquisa, serão utilizados os resultados obtidos em Língua Portuguesa e Matemática. Adicionalmente, também será utilizada a nota média padronizada dos estudantes, que consiste na média simples da proficiência em Língua Portuguesa e Matemática, padronizada para um indicador entre 0 e 10.

Outro ponto relevante que deve ser destacado é a exigência de uma taxa de participação mínima no dia da aplicação da prova. Um dos requisitos para que ocorra a aplicação da prova e posterior divulgação dos resultados é que o número de alunos matriculados e presentes no dia da prova obedeça uma quantidade específica e previamente determinada. Sabe-se que a exigência de taxas de participação não representa uma medida de confiabilidade dos resultados encontrados, mas configura uma estratégia para mitigar a manipulação de resultados por meio da seleção intencional de alunos. Ou seja, o intuito é que o resultado do exame seja, de fato, representativo da educação ofertada naquela unidade de ensino favorecendo que os resultados sejam estatisticamente robustos.

Sabe-se que o mero resultado em avaliações que mensuram proficiência não é capaz de inferir com absoluta certeza o desempenho de alunos em virtude de todas as questões inerentes a uma avaliação (por exemplo, nervosismo e ansiedade dos alunos no dia da prova). Aliado a isso, deve-se considerar o elevado grau de complexidade que envolve o conceito de “educação de qualidade”. Em outras palavras, deve-se ter cautela e considerar que diversas outras habilidades podem ser desenvolvidas além de conhecimentos básicos em português e matemática e que não são avaliados pela prova Saeb, tais como habilidades socioemocionais, pensamento crítico e inovador, ética e responsabilidade dentre vários outros. Contudo os resultados da prova Saeb são considerados uma informação relevante capaz de subsidiar a

---

<sup>11</sup> A caracterização da Escala Saeb para cada disciplina e ano correspondente pode ser encontrada em: [https://download.inep.gov.br/publicacoes/institucionais/avaliacoes\\_e\\_examenes\\_da\\_educacao\\_basica/escalas\\_de\\_proficiencia\\_do\\_saeb.pdf](https://download.inep.gov.br/publicacoes/institucionais/avaliacoes_e_examenes_da_educacao_basica/escalas_de_proficiencia_do_saeb.pdf)

elaboração, avaliação e aprimoramento de políticas públicas educacionais e possuem os requisitos necessários para serem utilizados em uma pesquisa científica.

Outra ressalva que deve ser feita acerca da utilização de notas de exames padronizados de proficiência é que estes não discriminam o tempo que cada estudante esteve exposto à educação ofertada pela referida escola. Deve-se considerar que nem todos os alunos permanecem na mesma escola durante toda a sua vida escolar e mudanças são comuns em qualquer período do ano letivo. Com isso, deve-se observar que caso o estudante tenha pertencido por anos a uma escola A que oferta uma educação precária, seja transferido para uma escola B com níveis superiores de educação no meio do ano letivo e, posteriormente realize o exame de proficiência, este indivíduo não pode ser considerado representativo para análise da educação ofertada pela escola B, mas os resultados não fazem essa ponderação. Para compreender melhor a qualidade da educação ofertada por uma escola seria necessário investigar não apenas os resultados obtidos pelos alunos, mas também há quanto tempo eles estiveram expostos à referida educação.

Adicionalmente, deve-se destacar que algumas variáveis de extrema relevância no âmbito educacional não podem ser observadas e mensuradas objetivamente tais como a motivação dos professores, competência e eficiência dos gestores escolares, engajamento e participação da comunidade escolar, qualidade do ensino e apoio destinados a alunos com laudo, dentre outras. Infere-se, portanto, que uma análise considerando apenas o resultado em uma avaliação não é ideal, mas permite a validade deste experimento, principalmente em virtude de apresentar resultados comparáveis desde a década de 1990, o que possibilita acompanhar a evolução histórica da educação ofertada por uma escola.

Dessa forma, a despeito das críticas válidas que podem ser feitas pela utilização de resultados na prova Saeb como indicador de qualidade educacional, essa base de dados apresenta características altamente desejáveis como a comparabilidade dos dados o que permite o incremento das séries históricas.

Nesta pesquisa, serão utilizados os resultados das provas realizadas em 2009, 2011, 2013, 2015, 2017, 2019 e 2021.

Salienta-se que os resultados de 2021 foram os primeiros dados capazes de evidenciar os impactos da pandemia de Covid-19 na educação brasileira. Ressaltar a ocorrência desse fator é de fundamental relevância, uma vez que a pandemia alterou radicalmente a forma de ensino tradicional. Além do período de suspensão das atividades de ensino, a pandemia obrigou que novos métodos de aprendizagem fossem estabelecidos de forma imediata abrangendo não apenas todo o território nacional, como também todo o mundo. Essa crise sanitária teve,

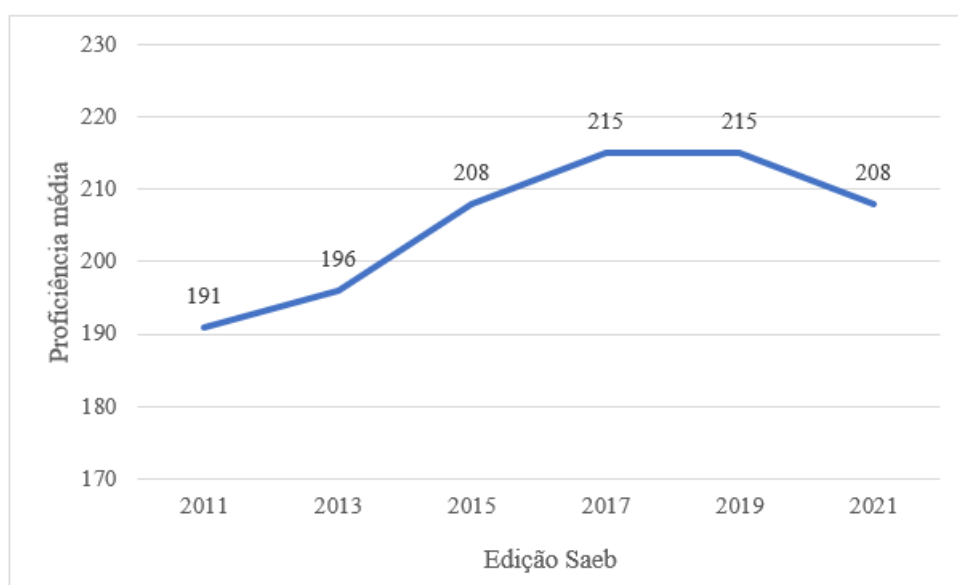
portanto, importantes reflexos na qualidade educacional ofertada pelas escolas principalmente nos anos de 2020 e 2021.

Ainda assim, a prova Saeb foi estruturada de forma a manter a comparabilidade com as edições anteriores e em 2021 foi aplicada em mais de 72 mil escolas públicas e privadas de todas as unidades da Federação, avaliando aproximadamente 5,3 milhões de estudantes<sup>12</sup>.

Deste modo, foram analisados os resultados obtidos em Língua Portuguesa e Matemática para o 5º e 9º anos.

Em relação à evolução dos resultados obtidos no 5º ano no período de 2011 a 2021, a proficiência média de Língua Portuguesa apresentou uma queda de 7 pontos em relação a 2019, enquanto a queda na proficiência média de matemática foi mais acentuada, representando a perda de 11 pontos. Ambas as trajetórias estão apresentadas nos Gráficos 5 e 6.

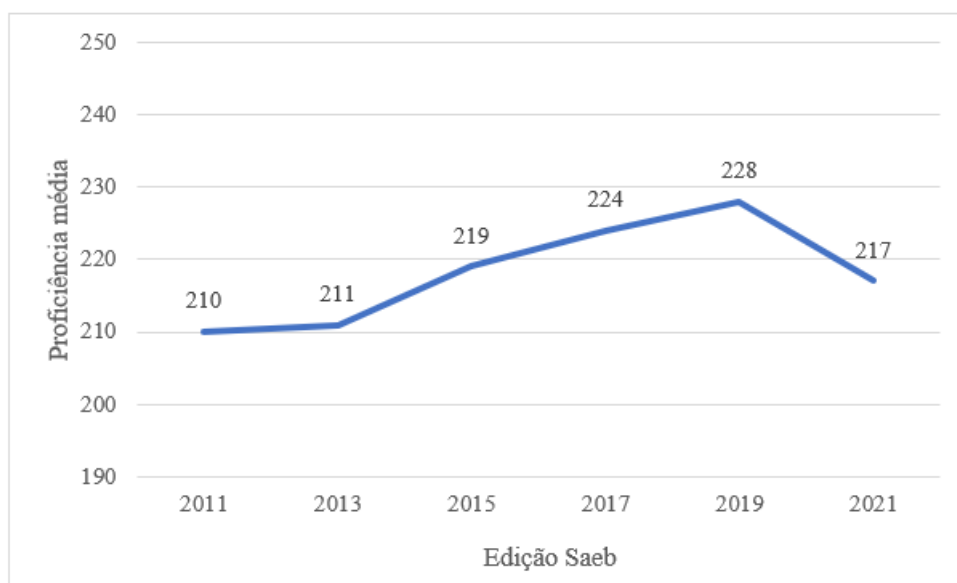
Gráfico 5: Evolução da proficiência média na prova Saeb em Língua Portuguesa no 5º ano do Ensino Fundamental – 2011 a 2021.



Fonte: Elaboração própria a partir de dados do INEP.

<sup>12</sup> Mais informações sobre a prova Saeb 2021 podem ser encontradas em: <https://www.gov.br/inep/pt-br/assuntos/noticias/saeb/mec-e-inep-divulgam-resultados-do-saeb-e-do-ideb-2021#:~:text=Em%202021%2C%20a%20taxa%20caiu,89%2C8%25%20em%202021.>

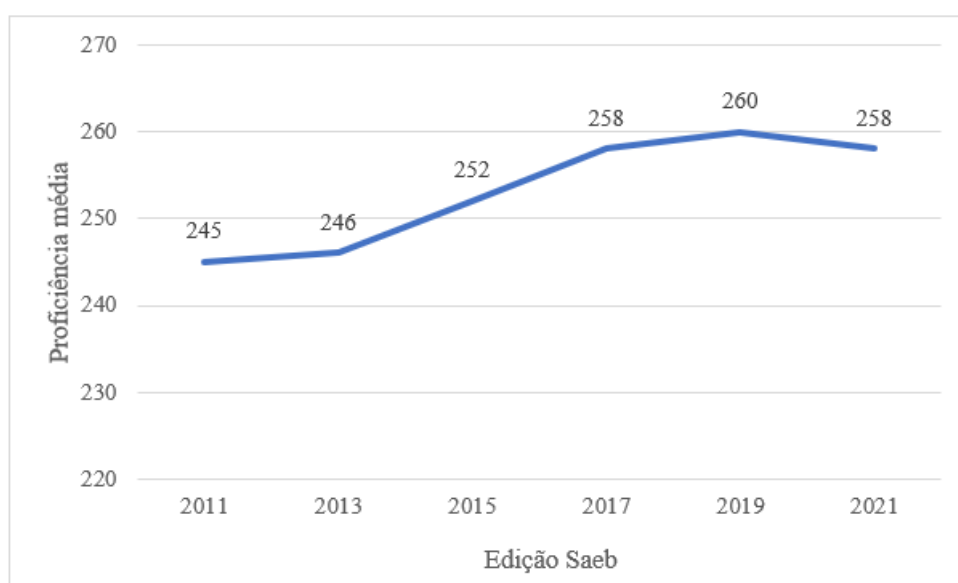
Gráfico 6: Evolução da proficiência média na prova Saeb em Matemática no 5º ano do Ensino Fundamental – 2011 a 2021.



Fonte: Elaboração própria a partir de dados do INEP.

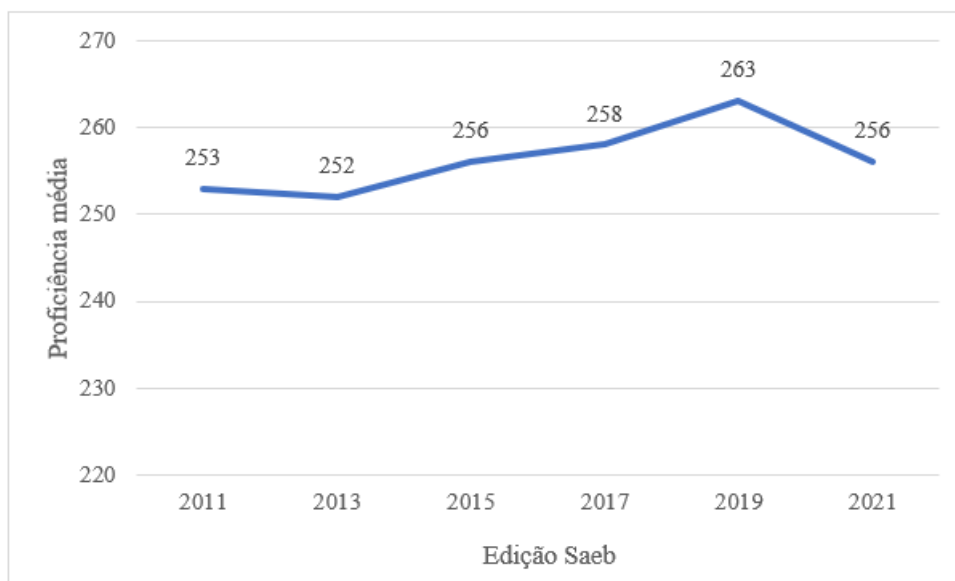
Em relação aos resultados obtidos pelo 9º ano, a evolução histórica das notas médias de Língua Portuguesa apresentou queda de 2 pontos no último exame realizado, enquanto os resultados de Matemática indicaram uma queda de 7 pontos. As referidas trajetórias estão dispostas nos gráficos 7 e 8.

Gráfico 7: Evolução da proficiência média na prova Saeb em Língua Portuguesa no 9º ano do Ensino Fundamental – 2011 a 2021.



Fonte: Elaboração própria a partir de dados do INEP.

Gráfico 8: Evolução da proficiência média na prova Saeb em Matemática no 9º ano do Ensino Fundamental – 2011 a 2021.



Fonte: Elaboração própria a partir de dados do INEP.

De modo geral, conclui-se que nas disciplinas avaliadas e nos anos selecionados, o último ano de realização do exame foi caracterizado por uma queda na proficiência média, tanto em Língua Portuguesa quanto em Matemática, possivelmente em decorrência da pandemia e seus demais desdobramentos, sendo o impacto mais expressivo em Matemática.

Adicionalmente, notou-se que o impacto negativo foi mais significativo no 5º do que no 9º ano, provavelmente devido aos diferentes graus de maturidade dos alunos, em que os primeiros apresentam maior dependência do ensino ofertado na unidade escolar para se desenvolver, enquanto os últimos possuem certa autonomia no aprendizado e são capazes de, em alguma medida, suplantar as debilidades na oferta de ensino.

Em suma, os resultados desse exame de proficiência são uma importante ferramenta a ser utilizada na melhoria da qualidade educacional do País, uma vez que avalia o desempenho dos estudantes permitindo identificar pontos fortes e fracos do sistema de ensino de modo a orientar o desenvolvimento, aprimoramento e monitoramento de políticas públicas educacionais com base em evidências e dados concretos.

Além de mensurar o nível de proficiência dos alunos, a prova Saeb também contempla a aplicação de questionários que são utilizados para calcular o Indicador de Nível Socioeconômico das Escolas de Educação Básica (Inse). Os alunos devem responder questões relacionadas a sua escolaridade, raça, se são pessoas com deficiência, escolaridade dos pais ou responsáveis, quantidade de pessoas que moram em sua casa e se esta é distante da escola dentre

outras informações relevantes a fim de quantificar informações do contexto socioeconômico do aluno e de sua família.

Dessa forma, é realizada uma análise multidimensional que permite contextualizar os resultados de proficiência dos estudantes com o *background* familiar. Esse indicador pode ser utilizado para compreender a realidade social das escolas e destacar desigualdades educacionais. Após a análise das respostas ao questionário e através de uma metodologia estatística que torna os resultados comparáveis ao longo do tempo, os dados são categorizados em oito níveis socioeconômicos diferentes da escala Inse.

Em virtude das competências e atribuições decorrentes da Constituição e da legislação nacional para os Estados e Municípios, somente faz sentido analisar os resultados da prova Saeb e do Inse para as escolas municipalizadas ofertantes de Ensino Fundamental. Das 1.115 escolas municipalizadas no período analisado, apenas 659 apresentaram resultados e puderam ser utilizadas nesta pesquisa.

Adicionalmente, também foram utilizados dados populacionais fornecidos pelo Censo Demográfico do Instituto Brasileiro de Geografia e Estatística (IBGE) e microdados do Censo Escolar da Educação Básica, coletados pelo INEP e atualizados anualmente que detalham informações importantes da unidade escolar e possuem caráter declaratório. Desta base de dados, foram selecionadas variáveis relevantes, por exemplo, a quantidade de matrículas no ensino fundamental (anos finais e iniciais), quantidade de docentes, se houve alteração na gestão escolar (municipalização), quantidade de matrículas de alunos pretos e pardos (com intuito de mensurar o contexto socioeconômico dos estudantes), quantidade de turmas, unidades escolares pertencentes à rede estadual e à rede municipal presentes no Município, dentre outras.

Essas informações foram combinadas de forma a destacar sua complementaridade e permitir uma análise fidedigna da qualidade educacional das unidades escolares selecionadas nesta pesquisa.

#### 4.2 ESTRATÉGIA EMPÍRICA E ESTATÍSTICAS DESCRITIVAS

O problema fundamental de estudos que buscam aferir efeitos causais é identificar e mensurar a trajetória de resultados potenciais do grupo de tratamento caso este não participasse do tratamento. Isso faz com que a estrutura desses estudos envolva, necessariamente, a realização de algumas suposições e algum grau de especulação em virtude da ausência de dados.

Para que se construa uma estimativa confiável, pode-se utilizar o método de pareamento. Este método tem por objetivo organizar um grupo de controle semelhante ao grupo de

tratamento com base em um vetor de características observáveis com a finalidade de estimar o efeito médio do tratamento.

Em um cenário ideal, unidades tratadas e de controle seriam exatamente pareadas em todas as covariáveis relevantes selecionadas, de forma que as distribuições amostrais nos dois grupos seriam idênticas. Contudo, na prática, estudos observacionais requerem o uso de técnicas estatísticas para otimizar o pareamento. Além do mais, deve-se considerar que as unidades nunca são absolutamente idênticas antes da aplicação dos tratamentos.

O critério primordial utilizado pelo método de pareamento é um vetor de variáveis observáveis mensuradas antes da aplicação do tratamento contendo as características que apresentam impactos significativos tanto na decisão de participar do tratamento quanto nos resultados potenciais alcançados pelas unidades tratadas. A partir da semelhança identificada nessas variáveis, realiza-se o pareamento das unidades.

Dessa forma, unidades semelhantes estariam sendo comparadas com a única diferença de que uma participou do tratamento enquanto a outra não. Da forma como é estruturado, o estimador de pareamento é extremamente sensível à seleção das variáveis observáveis. Destaca-se que os indivíduos que não apresentam grau de similaridade são excluídos da amostra.

A hipótese principal desse método é que existe um vetor de variáveis observáveis que contém todas as informações relevantes para a alocação do tratamento e resultados potenciais. Ao controlar por esse vetor, as unidades tornam-se independentes da atribuição do tratamento, isto é, a atribuição do tratamento torna-se aleatória e não está correlacionada com características não observadas capazes de impactar os resultados. Essa hipótese é denominada ignorabilidade (ROSENBAUM; RUBIN, 1983).

Sua implicação é que o resultado de uma unidade do grupo de controle será um parâmetro adequado para estimar o resultado potencial de uma unidade tratada com o mesmo vetor de variáveis observáveis diante da ausência de tratamento. Do mesmo modo, o resultado de uma unidade tratada será um parâmetro adequado para o resultado potencial de uma unidade do grupo de controle com as mesmas características caso tivesse sido selecionada para o tratamento.

Para estimar o efeito médio do tratamento nos tratados via pareamento, é necessário que a região do vetor que engloba as características dos indivíduos tratados também represente as características das unidades do grupo de controle. Essa hipótese é denominada sobreposição (também conhecida como “suporte comum”) e garante que as características dos tratados estejam presentes no grupo de controle e vice-versa (ROSENBAUM; RUBIN, 1983).



Ocorre que a participação ou não no tratamento muitas vezes está associada a variáveis relevantes não observáveis ou impossíveis de serem mensuradas. No cenário em que variáveis não observáveis forem fundamentais para definir a participação no tratamento bem como apresentarem impactos significativos nos resultados potenciais das unidades tratadas, o método do pareamento não conseguirá eliminar o viés de seleção.

De todo modo, pode-se calcular o efeito médio do tratamento para os tratados, com determinada combinação de características definidas em um vetor de variáveis observáveis, comparando a média observada dos resultados potenciais para os tratados com a média observada dos resultados potenciais para os indivíduos do grupo de controle com as mesmas características.

Em síntese, o objetivo do método de pareamento é, essencialmente, desenvolver critérios que permitam identificar as unidades com maior grau de similaridade entre os grupos de tratamento e controle com o propósito de estimar o contrafactual para o grupo de tratamento, o que permite calcular o efeito médio do tratamento sobre os tratados e controlar/mitigar vieses de seleção. Dessa forma, esse método é muito utilizado em estudos que visam analisar impactos de políticas públicas.

A despeito da ideia simplista de identificar unidades semelhantes do grupo de tratamento e controle a partir de variáveis observáveis e compará-las para identificar o efeito do tratamento, sabe-se que, diante da complexidade de alguns eventos, diversas variáveis possuem impactos significativos tanto na decisão de participar do tratamento quanto nos resultados alcançados pelas unidades tratadas e quanto maior for a dimensão desse vetor de variáveis observadas mais difícil será a implementação do estimador de pareamento.

Rosenbaum e Rubin (1983) apresentaram uma alternativa ao uso do pareamento das unidades com base em todo o vetor de variáveis observáveis. Os autores desenvolveram um método em que o pareamento ocorre com base em uma função capaz de resumir todas as características relevantes, intitulada “escore de propensão”. A metodologia denominada “Propensity Score Matching” (PSM) compreende a estimação do “escore de propensão” para cada unidade e, posteriormente, pareamento entre unidades tratadas e de controle com base em escores semelhantes. Com as unidades pareadas, é possível calcular o efeito causal médio do tratamento como se fosse um experimento randomizado.

O escore de propensão é definido como a probabilidade condicional de uma unidade receber o tratamento, dado o conjunto de variáveis selecionadas. Ocorre que, na maioria dos estudos, o escore de propensão não é conhecido e precisa ser estimado, dessa forma, esse

estimador apresenta sensibilidade à especificação correta do modelo do escore de propensão. Em suma, para estimar escores de propensão, alguma modelagem estatística será necessária.

Em seu artigo, os autores também apresentam o conceito de “escore de equilíbrio” sendo uma função  $b(x)$  das variáveis observáveis  $x$ , de modo que a distribuição condicional dessas variáveis dado  $b(x)$  é igual para unidades tratadas e de controle (ROSENBAUM; RUBIN, 1983). Se os escores de propensão estiverem equilibrados entre os grupos, indica que as variáveis observáveis estão distribuídas de forma semelhante nos grupos de tratamento e controle.

Um dos principais resultados de Rosenbaum e Rubin (1983) é que, em qualquer valor de um escore de equilíbrio, se a atribuição do tratamento for fortemente ignorável, a diferença entre as médias de tratamento e controle será uma estimativa imparcial do efeito médio do tratamento nesse valor do escore de equilíbrio. Em outras palavras, com atribuição de tratamento fortemente ignorável, o pareamento de pares em um escore de equilíbrio é capaz de produzir estimativas imparciais dos efeitos do tratamento.

Assim como o pareamento baseado no vetor de características observáveis, o pareamento baseado no escore de propensão também depende de uma definição capaz de quantificar o quão semelhantes são os escores de propensão de unidades tratadas e de controle para relacioná-las. Um dos estimadores mais comumente utilizados é o estimador de pareamento do “vizinho mais próximo correspondente” (“The Nearest Neighbor Matching”) que utiliza os resultados das unidades no grupo de controle que possuem escores de propensão mais próximos ao escore de propensão de um indivíduo tratado para estimar qual seria o resultado desse indivíduo caso ele não recebesse tratamento.

Como vantagens da utilização do método PSM, pode-se citar a redução do viés de seleção, a aproximação dos resultados ao que seria obtido em um experimento randomizado e que a variância da estimativa do efeito médio do tratamento será menor nas amostras pareadas do que seria em amostras aleatórias, em virtude das distribuições apresentarem maiores semelhanças.

Em contrapartida, existem importantes limitações, como a correta especificação do modelo do escore de propensão, a sensibilidade a variáveis não observáveis e a disponibilidade das variáveis relevantes para as unidades tratadas e de controle suficientes para obter resultados consistentes e interpretações intuitivas.

Em síntese, a utilização de métodos de pareamento e escores de propensão decorrem da necessidade de mitigar o viés de seleção de estudos observacionais em que a alocação do tratamento não é aleatória, ou seja, não é um experimento randomizado.

Em experimentos randomizados, o grupo de tratamento possui uma distribuição determinada por um mecanismo aleatório especificado, dessa forma, o escore de propensão é uma função conhecida. Já em experimentos não randomizados, na maioria dos casos, a função do escore de propensão é desconhecida e deve ser estimada a partir de dados observados. Além disso, em um experimento randomizado, o vetor de variáveis relevantes é especificado e contém todas as variáveis observadas que são utilizadas tanto na atribuição do tratamento quanto nos resultados – sendo a atribuição de tratamento e os resultados condicionalmente independentes dado o vetor de variáveis selecionadas (RUBIN, 1974).

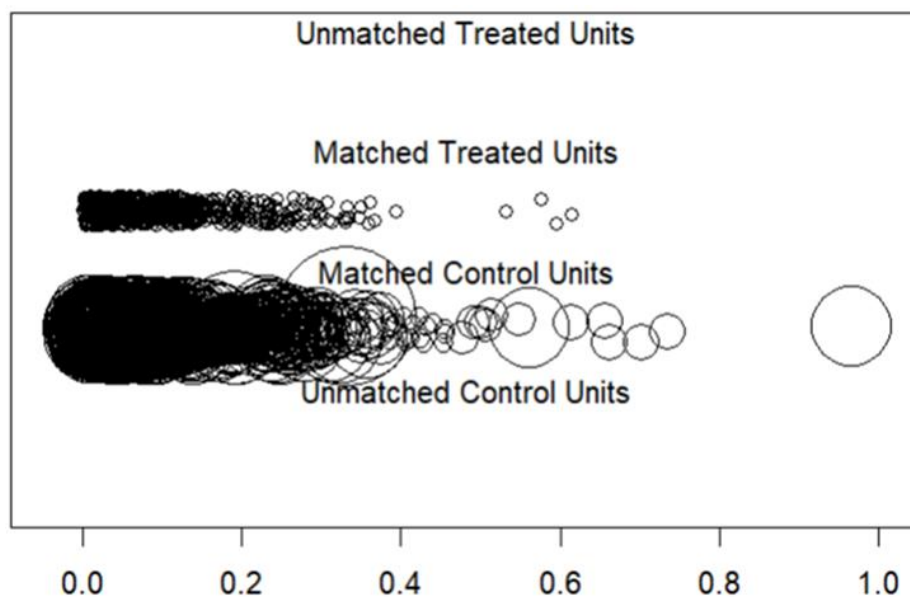
Rubin (1974) apresenta um estudo que conclui que a randomização deve ser utilizada sempre que possível, mas que o uso de dados controlados e não randomizados para estimar efeitos causais é um procedimento razoável e necessário em muitos experimentos. Isto é, não são apenas experimentos adequadamente randomizados que podem levar a estimativas úteis de efeitos causais.

O autor defende que se os pares pareados forem muito semelhantes no sentido de que, antes do início dos tratamentos, o pesquisador controlou as variáveis que poderiam afetar os resultados potenciais e atribuição do tratamento, então a diferença em cada par pareado é o efeito causal do tratamento nos tratados, e pode-se ter confiança que a média de todas as diferenças será o efeito causal médio do tratamento (RUBIN, 1974).

De modo geral, o “Propensity Score Matching” é uma técnica estatística utilizada para reduzir vieses de seleção. É especialmente útil para comparar grupos que não foram aleatoriamente designados e dos quais é necessário compreender efeitos causais através de estatísticas coerentes a fim de mensurar, por exemplo, impactos de políticas públicas.

Nesta pesquisa, o tratamento é a municipalização e o grupo de tratamento é composto por 659 escolas municipalizadas que apresentaram resultado na prova Saeb para o período de 2010 a 2021. Foi estimado o escore de propensão para todos os indivíduos utilizando 23.178 escolas para compor o grupo de controle. A distribuição do Propensity Score está apresentada no Gráfico 9 abaixo.

Gráfico 9: Distribuição dos escores de propensão após o pareamento.



Fonte: Elaboração própria.

Nota-se que, neste pareamento, todas as unidades de controle e de tratamento foram pareadas, de modo que não houve necessidade de descartar unidades. Foi realizada uma correspondência completa em que todas as unidades tratadas e de controle possuem pelo menos uma correspondência.

Em relação às variáveis selecionadas para o pareamento, foram utilizadas características relacionadas ao porte da escola, contexto socioeconômico dos alunos e características geográficas, tais como quantidade de matrículas no ensino fundamental (anos iniciais e finais), quantidade de docentes e de turmas, tamanho populacional, quantidade de matrículas de alunos pretos e pardos (com intuito de mensurar o contexto socioeconômico dos estudantes), unidades escolares pertencentes à rede estadual e à rede municipal presentes no Município.

Deve-se destacar que, possivelmente, a referida seleção de covariáveis não inclui todas as características que influenciam diretamente na municipalização, uma vez que esse processo não possui características objetivas mas, ao contrário, cada caso é caracterizado por singularidades que dificultam a compreensão e qualificação das unidades escolares que foram selecionadas para serem municipalizadas.

As Tabelas 1 e 2 apresentam as variáveis utilizadas no pareamento e algumas estatísticas descritivas das unidades tratadas e de controle antes e após o pareamento. Nota-se que houve uma aproximação significativa entre os valores médios das unidades tratadas e de controle após o pareamento em todas as variáveis utilizadas nesta pesquisa, o que pode ser comparado através da análise da coluna “Diferença média padronizada”.

Tabela 1: Estatísticas descritivas das variáveis utilizadas no PSM (pré pareamento).

Covariáveis	Média das unidades tratadas	Média das unidades de controle	Diferença média padronizada
distance	0,0946	0,0257	0,7995
Quantidade de matrículas Anos Finais	167,9865	264,6297	-0,5702
Quantidade de matrículas Anos Iniciais	120,4163	100,7091	0,1715
Quantidade de matrículas de alunos autodeclarados pretos	15,5051	20,7038	-0,1823
Quantidade de matrículas de alunos autodeclarados pardos	128,5144	192,6479	-0,5445
Quantidade de docentes Anos Iniciais	5,1847	5,0341	0,0304
Quantidade de docentes Anos Finais	9,4197	17,2828	-1,0203
Quantidade de turmas Anos Finais	5,7740	8,9612	-0,6319
Quantidade de turmas Anos Iniciais	4,8418	4,1048	0,1677
População	281.504,9	887.456,2	-1,0334
Quantidade de escolas estaduais no município	32,1820	97,7510	-1,2255
Quantidade de escolas municipais no município	91,3452	144,4000	-0,5696

Fonte: Elaboração própria.

Notas: A Diferença Média Padronizada é usada para comparar a média de uma variável entre dois grupos, considerando a variabilidade dentro de cada grupo. É calculada subtraindo a média do grupo de controle da média das unidades tratadas e dividindo o resultado pelo desvio padrão combinado dos dois grupos.

Tabela 2: Estatísticas descritivas das variáveis utilizadas no PSM (pós pareamento).

Covariáveis	Média das unidades tratadas	Média das unidades de controle	Diferença média padronizada
distance	0,0946	0,0952	-0,0073
Quantidade de matrículas Anos Finais	167,9865	204,543	-0,2157
Quantidade de matrículas Anos Iniciais	120,4163	113,377	0,0613
Quantidade de matrículas de alunos autodeclarados pretos	15,5051	15,8926	-0,0136
Quantidade de matrículas de alunos autodeclarados pardos	128,5144	153,1033	-0,2087
Quantidade de docentes Anos Iniciais	5,1847	4,8193	0,0738
Quantidade de docentes Anos Finais	9,4197	10,3632	-0,1224

Quantidade de turmas Anos Finais	5,7740	6,9127	-0,2258
Quantidade de turmas Anos Iniciais	4,8418	4,4981	0,0782
População	281.504,9	297.384,8	-0,0271
Quantidade de escolas estaduais no município	32,1820	35,295	-0,0582
Quantidade de escolas municipais no município	91,3452	94,1982	-0,0306

Fonte: Elaboração própria.

Deve-se observar a variável denominada “distance”, antes e depois do pareamento. Essa variável calcula a distância entre os escores de propensão atribuída às unidades tratadas e de controle indicando o grau de similaridade ou de diferença. A interpretação a ser realizada é tal que no cenário em que os escores de propensão das unidades tratadas e de controle forem semelhantes, então as unidades serão “próximas” e serão consideradas comparáveis.

Em síntese, essa variável é uma medida da distância entre os escores de propensão dos indivíduos em grupos de tratamento e controle, sendo uma ferramenta essencial para verificar se o pareamento está ocorrendo entre unidades com características semelhantes e comparáveis.

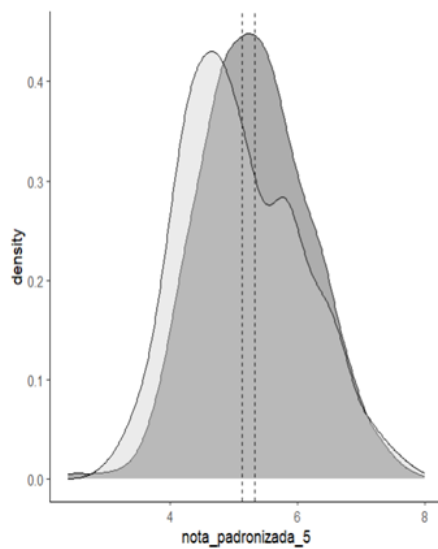
Uma vez que a metodologia do Propensity Score Matching busca identificar para cada unidade tratada pelo menos uma unidade de controle com escore de propensão semelhante, a análise dessa variável é um componente crítico e sinaliza o quão bem as unidades estão pareadas.

Os gráficos contendo a distribuição de densidade de cada uma das variáveis selecionadas, antes e depois do pareamento, encontram-se no Apêndice A.

A fim de melhor caracterizar os grupos selecionados para tratamento e controle, serão analisadas a distribuição de densidade das notas padronizadas, de Língua Portuguesa e Matemática para o 5º e 9º anos.

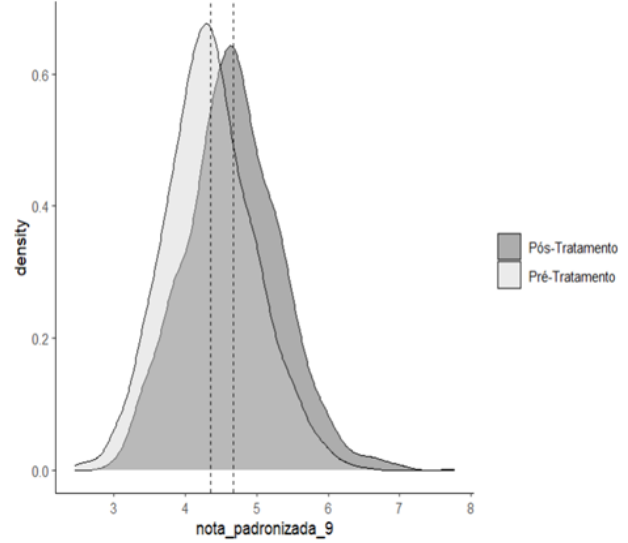
Analisando apenas o grupo de tratamento, isto é, escolas que foram municipalizadas no período de 2010 a 2021, antes e depois do tratamento, tem-se que ocorreu um aumento pouco significativo na nota padronizada tanto para o 5º ano (+ 0,2 pontos) quanto para o 9º ano (+ 0,33 pontos). Os gráficos de densidade com a referida comparação estão na Figura 1.

Figura 1: Comparação da nota padronizada para 5º e 9º anos das escolas municipalizadas pré e pós tratamento.



Pré-Tratamento: **5.13**

Pós-Tratamento: **5.33**



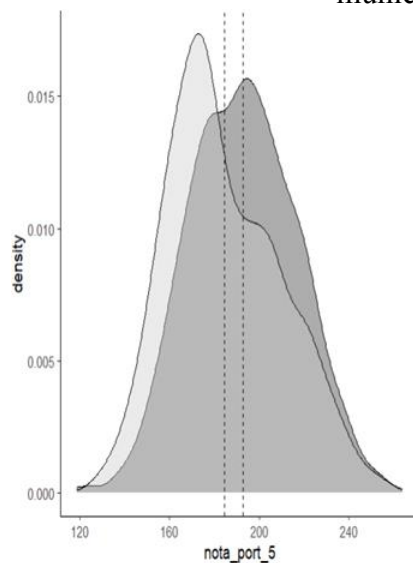
Pré-Tratamento: **4.35**

Pós-Tratamento: **4.68**

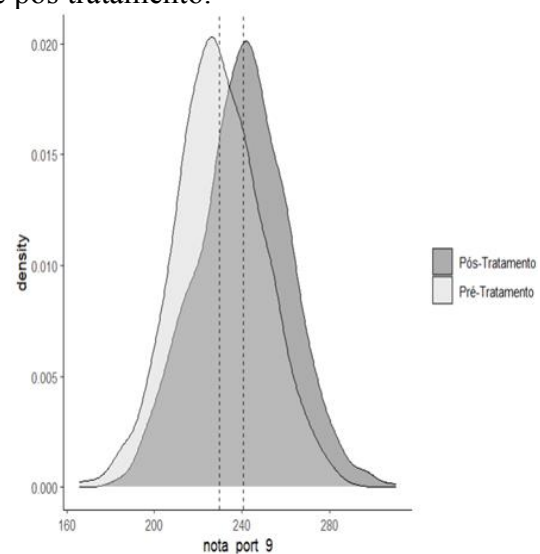
Fonte: Elaboração própria.

A Figura 2 detalha os gráficos de densidade das notas de Língua Portuguesa referentes ao 5º e 9º anos das escolas municipalizadas. Nota-se que, de forma semelhante ao identificado na análise anterior, houve um aumento absoluto pouco significativo nas notas, com impacto menor no 5º ano (+ 9 pontos) que no 9º ano (+ 11 pontos).

Figura 2: Comparação da nota de Língua Portuguesa para 5º e 9º anos das escolas municipalizadas pré e pós tratamento.



Pré-Tratamento: **184**



Pré-Tratamento: **230**

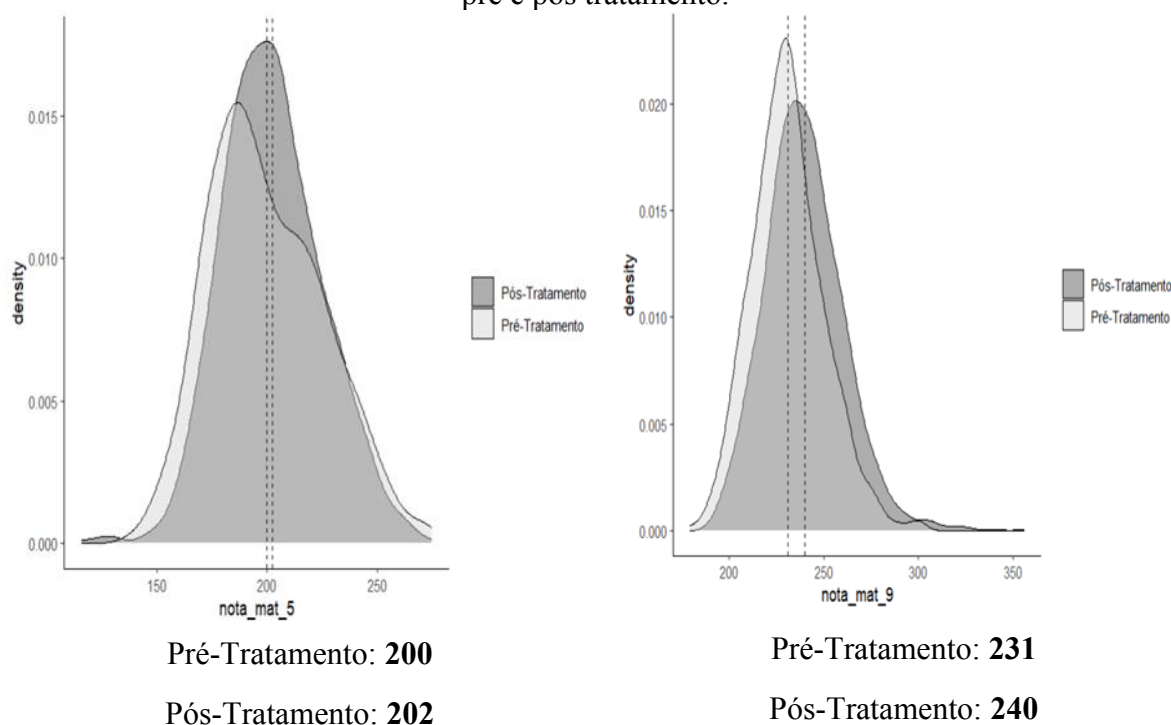
Pós-Tratamento: **193**

Pós-Tratamento: **241**

Fonte: Elaboração própria.

Em relação às notas de Matemática, de mesmo modo, pode-se constatar um aumento nas notas do 5º ano (+ 2 pontos) inferior ao aumento verificado nas notas do 9º ano (+ 9 pontos).

Figura 3: comparação da nota de matemática para 5º e 9º anos das escolas municipalizadas pré e pós tratamento.



Fonte: Elaboração própria.

Dessa forma, após a análise básica da distribuição da densidade das notas médias para as escolas municipalizadas antes e depois de passarem pelo processo de municipalização, percebe-se que, em todos os âmbitos e sob todos os aspectos, houve um modesto aumento de nota média para os 5º e 9º anos avaliados, sendo que os maiores impactos foram observados nas turmas de 9º ano.

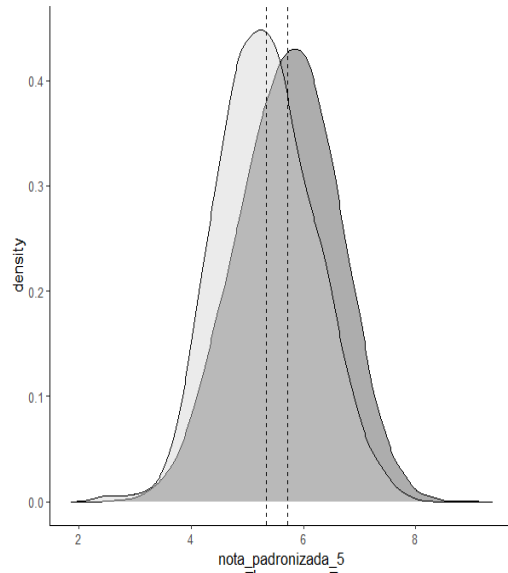
Realizando análise semelhante à anterior, mas dessa vez comparando as notas médias das escolas tratadas após o tratamento com as unidades de controle, tem-se as Figuras 4, 5 e 6.

Ao comparar a densidade das notas padronizadas para os 5º e 9º anos, observou-se que em ambos os casos o grupo de controle, isto é, escolas estaduais que permaneceram sob gestão estadual, apresentou nota média superior ao grupo tratado. Em relação ao 5º ano, a nota média



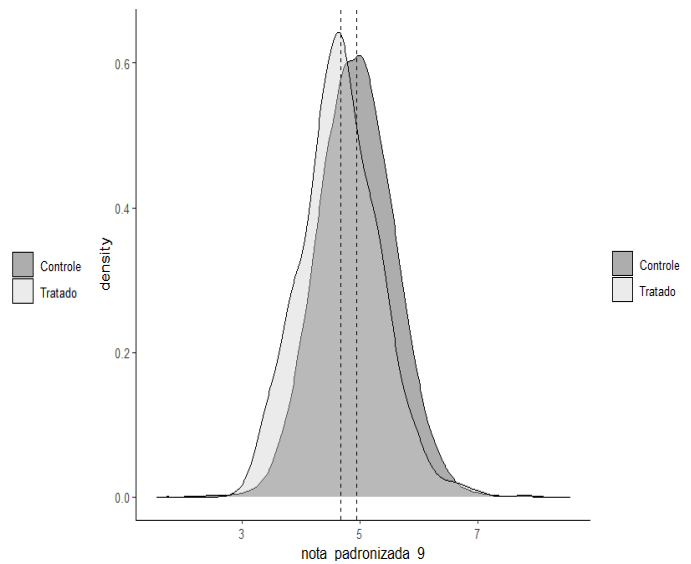
do grupo de controle foi 0,39 pontos superior à nota média do grupo tratado, enquanto no 9º ano constatou-se 0,26 pontos a mais.

Figura 4: Gráficos de densidade da nota padronizada para 5º e 9º anos das escolas tratadas e controle.



**Controle: 5.72**

**Tratado: 5.33**



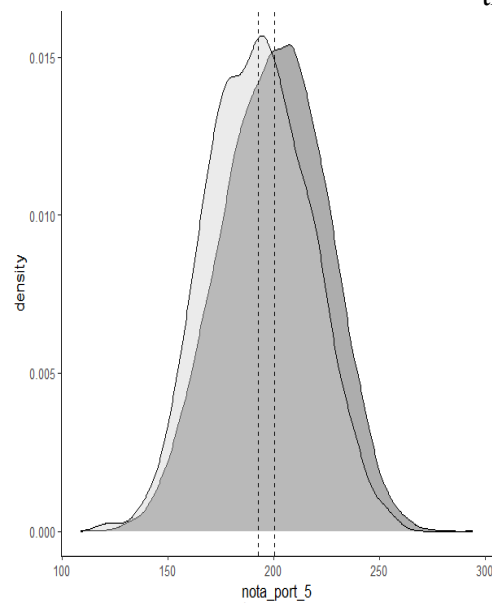
**Controle: 4.94**

**Tratado: 4.68**

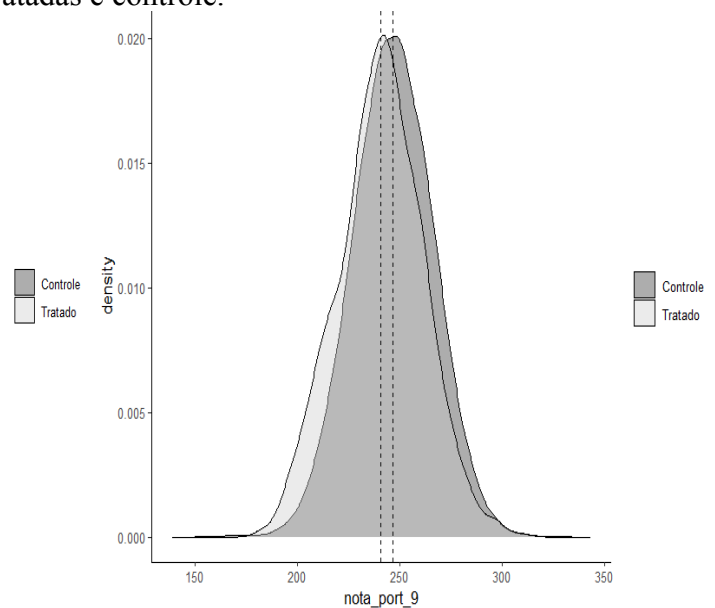
Fonte: Elaboração própria.

Em relação às notas de Língua Portuguesa, também pode-se notar nota média superior para o grupo de controle tanto no 5º (+ 7 pontos) quanto no 9º ano (+ 6 pontos).

Figura 5: Gráficos de densidade da nota de Língua Portuguesa para 5º e 9º anos das escolas tratadas e controle.



**Controle: 200**



**Controle: 247**

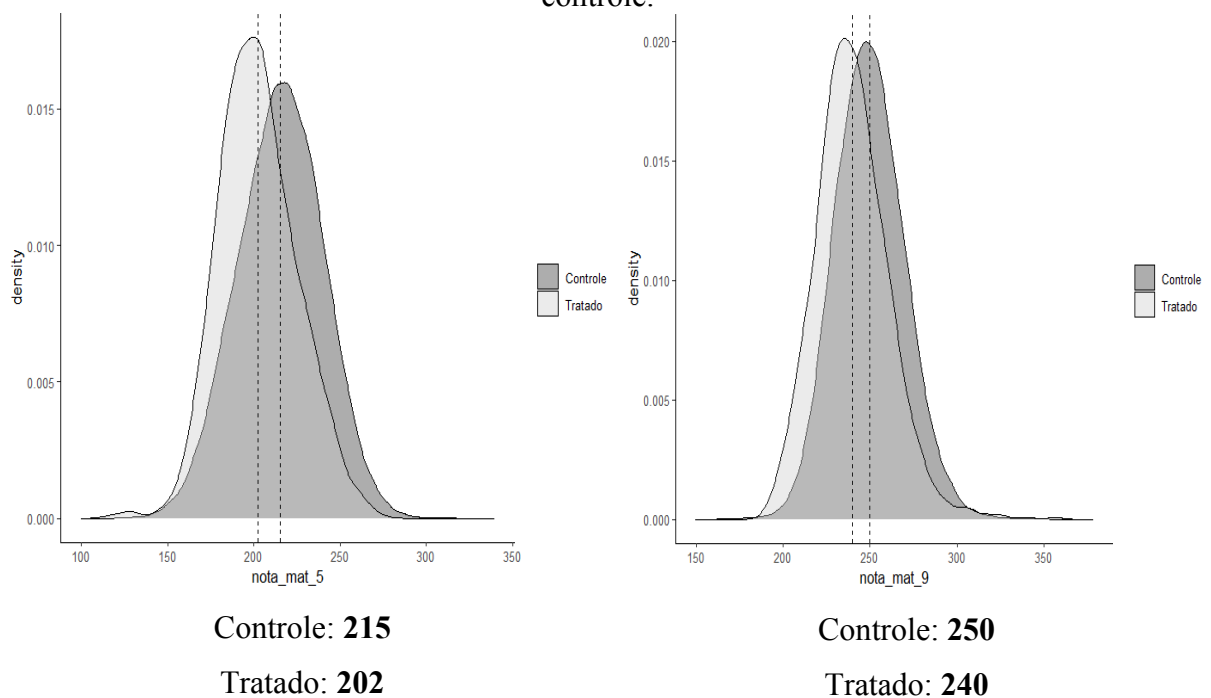
Tratado: **193**

Tratado: **241**

Fonte: Elaboração própria.

Em relação às notas de matemática, o mesmo cenário foi constatado: notas médias do grupo de controle superiores ao grupo tratado, sendo que para o 5º ano foram observados 13 pontos a mais nas unidades de controle e no 9º ano 10 pontos acima da média obtida para as unidades tratadas.

Figura 6: Gráficos de densidade da nota de Matemática para 5º e 9º anos das escolas tratadas e controle.



Fonte: Elaboração própria.

Em síntese, a partir de uma análise básica dos gráficos de densidade, notou-se que, ainda que tenha sido observado um aumento pouco significativo ao se comparar exclusivamente unidades tratadas antes e depois do tratamento, quando se contrapõe unidades tratadas *versus* unidades de controle, estas últimas apresentaram, em todos os casos e sob todos os aspectos, notas médias padronizadas, de Matemática e de Língua Portuguesa superiores tanto para o 5º quanto para o 9º anos, apresentando impactos mais significativos nas turmas de 5º ano.

Essa análise superficial de médias já sinaliza que a municipalização não apresentou impactos significativos na proficiência dos alunos. Contudo, uma análise mais robusta se faz necessária a fim de analisar o problema utilizando técnicas estatísticas adequadas capazes de investigar a questão sob o prisma de suas características fundamentais.

## 5 RESULTADOS

Nesta pesquisa, serão combinadas as técnicas de Propensity Score Matching e a metodologia de Diferenças-em-Diferenças desenvolvida por Callaway e Sant’Anna (2021) o que, em tese, resulta em estimadores menos enviesados. Adicionalmente, a partir da mesma base de dados, serão calculados estimadores utilizando a metodologia TWFE.

Dessa forma, a seção 5.1 analisa o efeito médio do tratamento nas unidades escolares tratadas, o efeito médio específico por grupos e agregado, o efeito dinâmico do tratamento e o efeito dinâmico balanceado para cinco anos obtidos através dos estimadores propostos por Callaway e Sant’Anna. Como o problema investigado é caracterizado pela adoção temporalmente diferenciada do tratamento, o objetivo é identificar e evidenciar a existência de efeitos heterogêneos do tratamento.

No item 5.2 será realizada uma análise da robustez dos resultados obtidos recalculando os estimadores sob a ótica da mesma metodologia, porém alterando a composição do grupo de controle para unidades “ainda não tratadas” ao invés de utilizar unidades “nunca tratadas”.

Por último, no tópico 5.3 serão apresentados os resultados do estimador TWFE a fim de comparar os resultados obtidos através das diferentes metodologias.

### 5.1 RESULTADOS DO MODELO CALLAWAY E SANT’ANNA

Na presente seção serão analisados: o estimador simples, resultante da média ponderada de todos os efeitos médios do tratamento com pesos proporcionais ao tamanho do grupo; os efeitos médios do tratamento nos diferentes grupos e o respectivo efeito agregado, em que um grupo é definido em virtude do momento em que foi submetido à municipalização; os efeitos dinâmicos do tratamento, com vistas a apresentar os efeitos médios da municipalização em diferentes períodos de exposição ao tratamento, ou seja, apresentando efeitos agregados do tratamento por tempo relativo – semelhante a um estudo de evento; por último, o estimador dinâmico balanceado que levará em consideração as unidades escolares expostas ao tratamento a pelo menos 5 anos.

As estimações aqui apresentadas foram efetuadas utilizando as seguintes especificações: grupo de controle composto por unidades “nunca tratadas”, método resultado de regressão, nenhuma antecipação permitida e sendo permitido um painel desbalanceado.

Conforme exposto, o estimador simples calcula uma média ponderada de todos os efeitos médios de tratamento do grupo-tempo, com pesos proporcionais ao tamanho do grupo.

Este estimador é, portanto, a agregação dos efeitos médios e identifica o efeito geral da participação do tratamento independentemente do tempo de exposição ao tratamento.

Tabela 3: Estimador simples dos efeitos médios de tratamento.

Variável Dependente	Nota	Nota	Língua	Língua	Matemática	Matemática
	Padronizada 5º Ano	Padronizada 9º Ano	Portuguesa 5º Ano	Portuguesa 9º Ano	5º Ano	9º Ano
ATT	-0,1176 *	-0,0136	-4,8308 *	-0,6526	-4,8661 *	-0,1624
<b>Simples</b>	(0,0353)	(0,0392)	(1,1611)	(1,1231)	(1,2681)	(1,3236)

Fonte: Elaboração própria.

Notas: \* Significativo a 5%. Entre parênteses está apresentado o desvio padrão.

Nota-se que, a despeito de todos os estimadores sinalizarem um impacto negativo da municipalização nos indicadores de desempenho escolar, apenas os resultados relacionados ao 5º ano apresentaram significância estatística. Neste caso, o impacto foi de -0,11 pontos na nota padronizada e os efeitos em Língua Portuguesa e Matemática foram semelhantes: em ambos os casos o efeito geral da municipalização foi de -4,8 pontos.

Em relação ao 9º ano, além dos estimadores simples não apresentarem significância estatística, sugeriram que o efeito geral da municipalização é negativo e próximo a zero: nota padronizada (-0,01), Língua Portuguesa (-0,65) e Matemática (-0,16).

Callaway e Sant'Anna enfatizam que apesar do estimador simples evitar os problemas com pesos negativos identificados na utilização da metodologia TWFE, essa possivelmente não é a melhor forma de agregação, sugerindo a análise do estimador calculado para discriminar os efeitos por grupos, sendo a principal escolha dos autores para analisar o efeito agregado da participação no tratamento.

Este estimador calcula efeitos médios de tratamento entre diferentes grupos, sendo o efeito agregado a média do efeito entre diferentes grupos.

A Tabela 4 resume os efeitos médios de tratamento para os diferentes grupos e em sua última linha são apresentados os resultados agregados, isto é, o efeito global médio do tratamento para todos os grupos.

Tabela 4: Efeitos médios de tratamento para os diferentes grupos e agregado.

Grupo	Nota	Nota	Língua	Língua	Matemática	Matemática
	Padronizada 5º Ano	Padronizada 9º Ano	Portuguesa 5º Ano	Portuguesa 9º Ano	5º Ano	9º Ano
<b>2010</b>	-0,1702 (0,0893)	0,0687 (0,115)	-5,6972 (2,4393)	0,9485 (3,7259)	-6,2416 (2,838)	3,1716 (4,2503)
<b>2011</b>	0,0430 (0,2198)	0,2306 (0,149)	-1,0993 (5,4181)	5,6019 (3,9230)	-1,1747 (6,7221)	8,2316 (6,2507)
<b>2012</b>	-0,2657 *	-0,0861	-10,4688 *	-3,3442	-8,7883 *	-1,8248

	(0,0779)	(0,0762)	(2,1682)	(2,4931)	(2,1189)	(2,4428)
<b>2013</b>	-0,2355 *	-0,1960 *	-9,8556 *	-5,6536 *	-9,1446 *	-6,1060 *
	(0,0611)	(0,0636)	(1,7867)	(1,8827)	(2,0539)	(2,1395)
<b>2014</b>	0,1156	0,1200	5,1067	4,4543	3,3022	2,7478
	(0,1508)	(0,1203)	(3,6694)	(3,5424)	(3,9605)	(3,6515)
<b>2015</b>	0,0670	-0,0394	2,8356	-0,7654	1,5083	-1,5986
	(0,1768)	(0,1055)	(5,1734)	(3,0159)	(4,9545)	(3,6643)
<b>2016</b>	0,3309	0,1462	7,5596	5,2726	8,1058	3,4991
	(0,1503)	(0,1821)	(4,6832)	(5,2159)	(4,0905)	(5,3393)
<b>2017</b>	-0,1583	0,4320	-5,0988	9,4348	-5,1685	16,4882 *
	(0,1384)	(0,1877)	(3,9202)	(6,6232)	(4,0157)	(5,4604)
<b>2018</b>	0,3885	0,4423 *	14,8193	11,7665 *	10,4819	14,7737 *
	(0,1961)	(0,128)	(6,0950)	(4,0844)	(4,9277)	(4,0593)
<b>2019</b>	0,1631	0,4322 *	10,530 *	12,3194 *	3,9878	13,6153 *
	(0,1187)	(0,1071)	(3,7834)	(3,0134)	(3,3645)	(3,2952)
<b>2020</b>	-0,1584	-0,1059	-5,9426	-3,2344	-3,9046	-3,1180
	(0,1579)	(0,1319)	(4,3106)	(3,9936)	(4,1157)	(4,1299)
<b>2021</b>	0,1657	-0,3174	6,6996	4,7886	1,4368	-23,8343
	(0,2343)	(0,7274)	(7,0415)	(13,443)	(7,3987)	(52,296)
<b>Agregado</b>	-0,0743 *	0,0288	-2,9932 *	0,6872	-3,4365 *	1,0413
	(0,034)	(0,0356)	(1,0882)	(1,2121)	(1,1518)	(1,3065)

Fonte: Elaboração própria.

Notas: \* Significativo a 5%. Entre parênteses está apresentado o desvio padrão.

Os efeitos individuais para todos os grupos e em todas as variáveis analisadas apresentaram resultados com pouca precisão estatística e que oscilam entre impactos negativos e positivos em diferentes graus.

Como as escolas foram municipalizadas em diferentes anos, os impactos do tratamento podem ser heterogêneos entre os grupos. Em outras palavras, poderia ocorrer o cenário em que o efeito na proficiência fosse maior e estatisticamente significativo nas escolas municipalizadas primariamente, uma vez que a exposição à nova gestão escolar e toda a alteração que isto envolve em termos de qualidade educacional exige certo lapso temporal para surtir efeitos. Isso seria condizente com resultados das escolas municipalizadas nos primeiros anos da análise indicando impactos mais robustos enquanto as unidades municipalizadas tardiamente apresentariam efeitos imperceptíveis no impacto da proficiência.

Contudo, a partir da análise dos resultados, não foi possível perceber um efeito significativo para unidades tratadas primariamente ou tardiamente, isto é, contrapondo a ideia de que a unidade escolar municipalizada preliminarmente apresentaria maiores efeitos sobre o desempenho dos alunos.

Conforme verificado, os diversos grupos apresentaram estimativas desagregadas em sua maioria com ausência de significância estatística. Disso decorre a necessidade e importância da análise dos efeitos sob a ótica agregada.

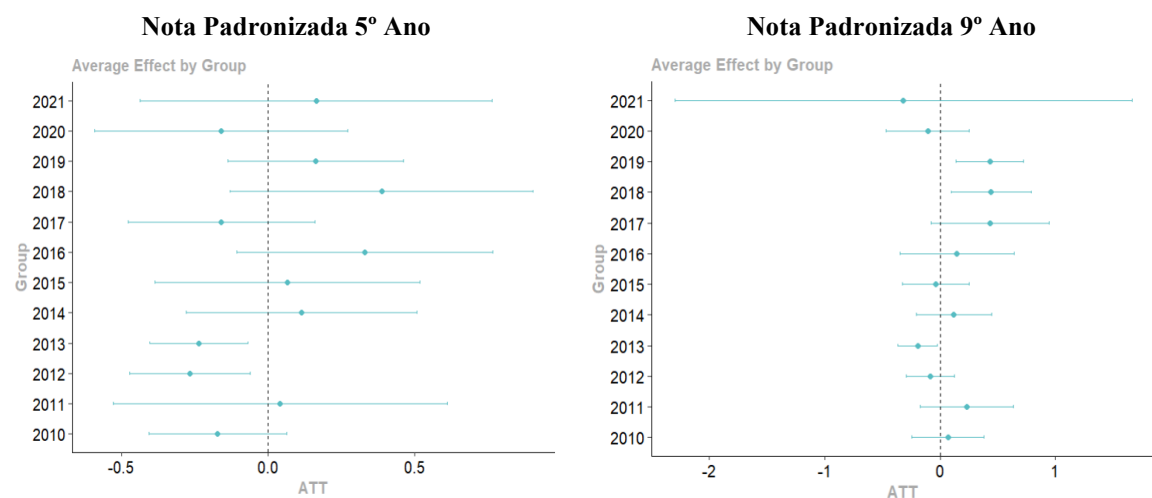
Sob a ótica do efeito global médio, constatou-se resultado significativo estatisticamente e negativo apenas para o 5º ano, isto é, para a nota padronizada (-0,07), Língua Portuguesa (-

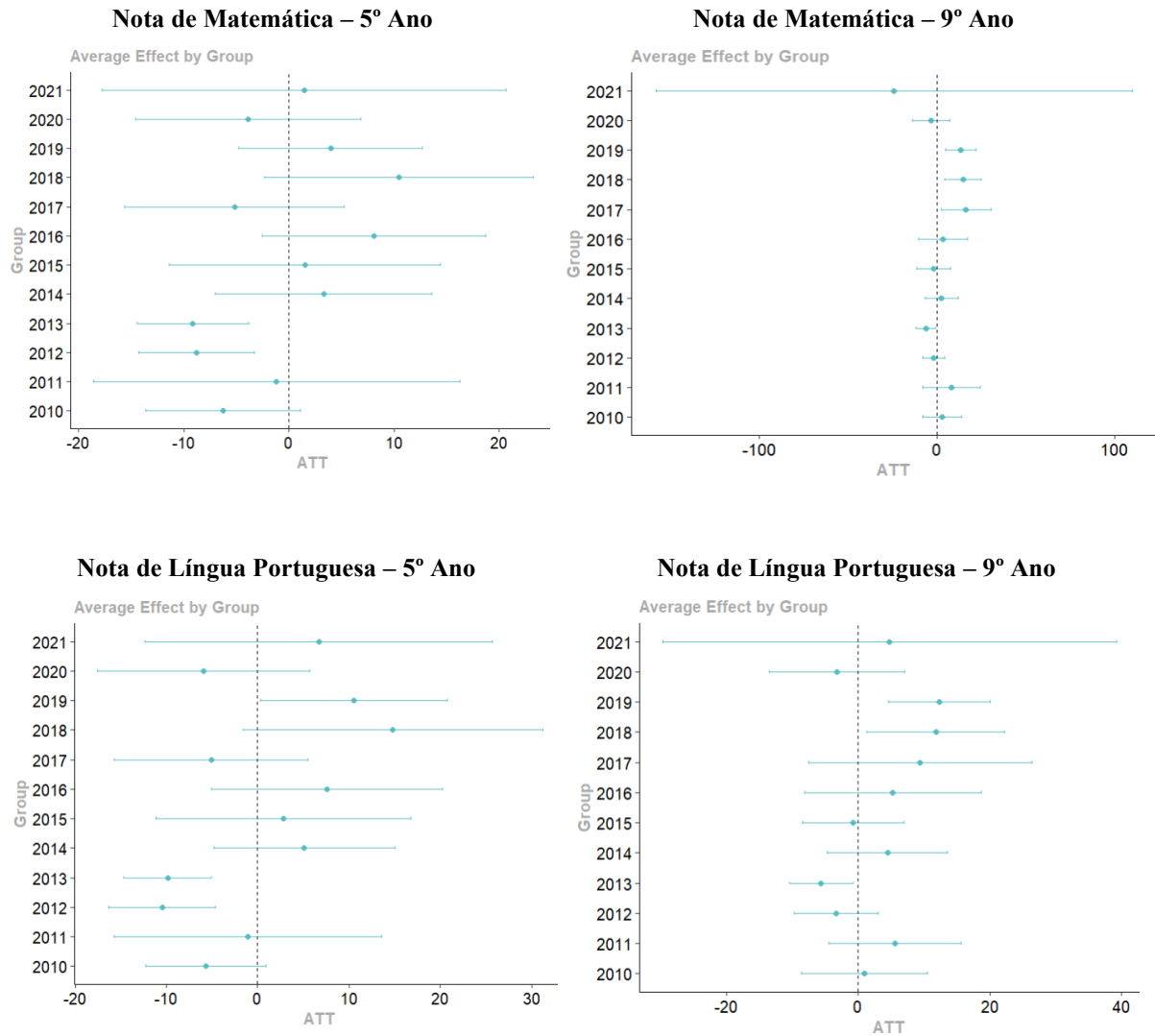
2,99) e Matemática (-3,43). Em relação às avaliações no 9º ano, apesar de não apresentarem significância estatística, indicaram resultados positivos e em pouca magnitude: nota padronizada (0,02), Língua Portuguesa (0,68) e Matemática (1,04). Em suma, os resultados indicam que os grupos experimentaram efeitos ambíguos da municipalização, embora a maior parte das estimações sejam pouco precisas e, no geral, de baixa magnitude.

Assim como verificado pelo estimador simples, o estimador agregado por grupos identificou impactos estatisticamente significativos e negativos para o 5º ano, mas diferentemente deste, constatou resultados positivos para as turmas de 9º ano, embora sem significância. Esses resultados se alinham com o verificado por Leme, Paredes e Souza (2009) e Ceneviva (2012) que identificaram não haver impactos significativos da municipalização sobre a proficiência.

Os gráficos relacionados aos efeitos de grupos estão apresentados na Figura 7. As linhas azuis mostram o efeito do programa para cada grupo, sendo significativas as estimativas que não incluem o zero em seu intervalo de confiança.

Figura 7: Efeitos da municipalização nos diferentes grupos.





Fonte: Elaboração própria.

Em suma, os parâmetros resumidos agregados por grupo são consistentes com a ideia de que a municipalização gerou impactos pouco precisos e de pouca magnitude nos indicadores de proficiência, sendo que para o 5º ano os estimadores sinalizaram impactos negativos e com significância estatística e para o 9º ano obteve-se resultados próximos a zero e sem significância.

A metodologia também permite a estimativa dos efeitos agregados do tratamento por tempo relativo (ou seja, na forma de estudo de evento). Essas estimativas decompõem a diferença média das diferenças entre unidades tratadas e de controle em cada período em relação ao início do tratamento e auxiliam a análise de validade da suposição de tendências paralelas.

Nesse sentido, o estimador dinâmico calcula os efeitos médios em diferentes durações de exposição ao tratamento e o efeito agregado considera o efeito do tratamento em todas as durações positivas de exposição.

A Tabela 5 discrimina o ATT dinâmico agregado para cada variável analisada e em sua maioria apresentou resultados com significância estatística.

Tabela 5: Estimador de efeitos médios dinâmico agregado.

<b>Variável Dependente</b>	<b>Nota Padronizada 5º Ano</b>	<b>Nota Padronizada 9º Ano</b>	<b>Língua Portuguesa 5º Ano</b>	<b>Língua Portuguesa 9º Ano</b>	<b>Matemática 5º Ano</b>	<b>Matemática 9º Ano</b>
<b>ATT Dinâmico Agregado</b>	-0,0285 (0,0438)	0,1205 * (0,053)	-2,3671 (1,4182)	3,1484 * (1,4883)	-2,4204 * (1,3981)	4,0804 * (1,9202)

Fonte: Elaboração própria.

Notas: \* Significativo a 5%. Entre parênteses está apresentado o desvio padrão.

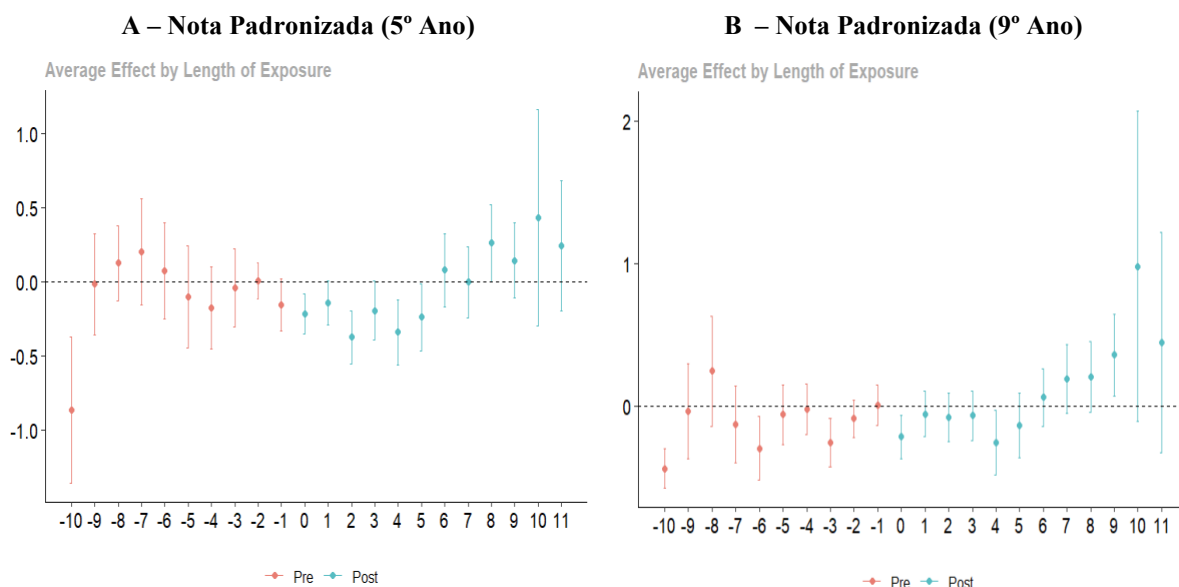
Para o 5º ano, todos os impactos agregados foram negativos, sendo nota padronizada (-0,02), Língua Portuguesa (-2,36) e Matemática (-2,42) - apenas este último estatisticamente significativo. Em relação ao 9º ano, todos os estimadores apresentaram resultados positivos e com significância estatística, quais sejam: nota padronizada (0,12), Língua Portuguesa (3,14) e Matemática (4,08). Ademais, sob essa ótica foi possível diagnosticar impactos de maiores magnitudes no 9º ano em relação ao 5º ano.

Os resultados do estimador dinâmico agregado sinalizam o mesmo evidenciado pelo estimador de efeitos médios agregado por grupos: resultados negativos para o 5º ano e positivos para o 9º ano, todos em pequena magnitude.

Abaixo estão apresentados os gráficos que identificam os efeitos médios ao longo do período de exposição ao tratamento. As linhas verticais representam os intervalos de confiança, sendo que as azuis indicam o efeito médio após a municipalização e são consideradas estimativas estatisticamente significativas as que não incluem o valor zero. Já as linhas em vermelho indicam as estimativas pré-tratamento, isto é, referem-se ao período em que a unidade escolar ainda era pertencente à rede estadual de ensino. Destaca-se que não será possível rejeitar a hipótese de tendências paralelas quando todos os intervalos de confiança para todos os grupos em todos os períodos pré-tratamento incluírem o valor zero.



Figura 8: Estimador dinâmico de efeitos médios para nota padronizada (5º e 9º anos).



Fonte: Elaboração própria.

No gráfico A, percebe-se que a nota padronizada do 5º ano das escolas tratadas já apresentava trajetória decrescente antes do tratamento. Após a municipalização e até o quinto ano de tratamento, nota-se impactos negativos alternados em maiores e menores magnitudes, tornando-se positivo apenas após seis anos da municipalização. Verificou-se significância estatística apenas para 4 períodos após o tratamento, quais sejam: no ano do tratamento (-0,21), dois anos após (-0,37), quatro anos após (-0,34) e cinco anos após o tratamento (-0,23). O restante dos intervalos de confiança abrange o valor zero, do que se infere não haver significância estatística nesses resultados.

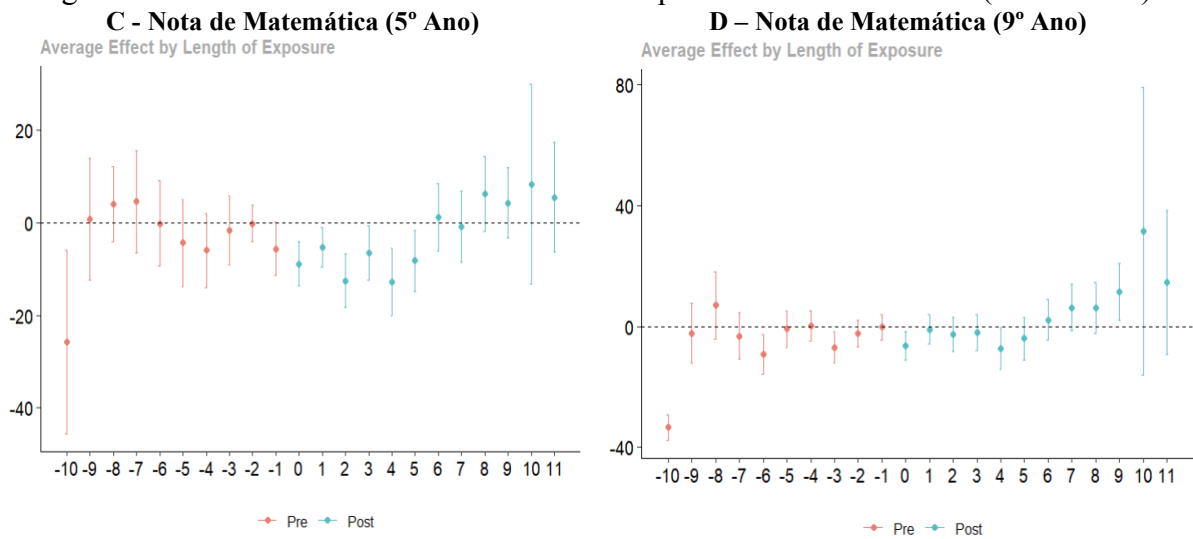
Já no gráfico B refere-se à nota padronizada do 9º ano, do mesmo modo, constatou-se pouca significância estatística nos resultados tanto antes quanto depois do tratamento. Destaca-se que após o tratamento houve apenas dois períodos com significância: após 4 anos da municipalização (-0,25) e nove anos após a municipalização (0,36).

Adicionalmente, pode-se perceber uma evolução positiva dos efeitos médios do tratamento mais acentuada do 9º ano do que para o 5º ano.

Em ambos os casos, constata-se poucos efeitos estatisticamente significativos e os primeiros anos apresentando efeitos negativos, com impactos positivos posteriores.

A Figura 9 apresenta os resultados identificados para os estimadores dinâmicos referentes às notas de Matemática.

Figura 9: Estimador dinâmico de efeitos médios para nota de Matemática (5º e 9º anos).

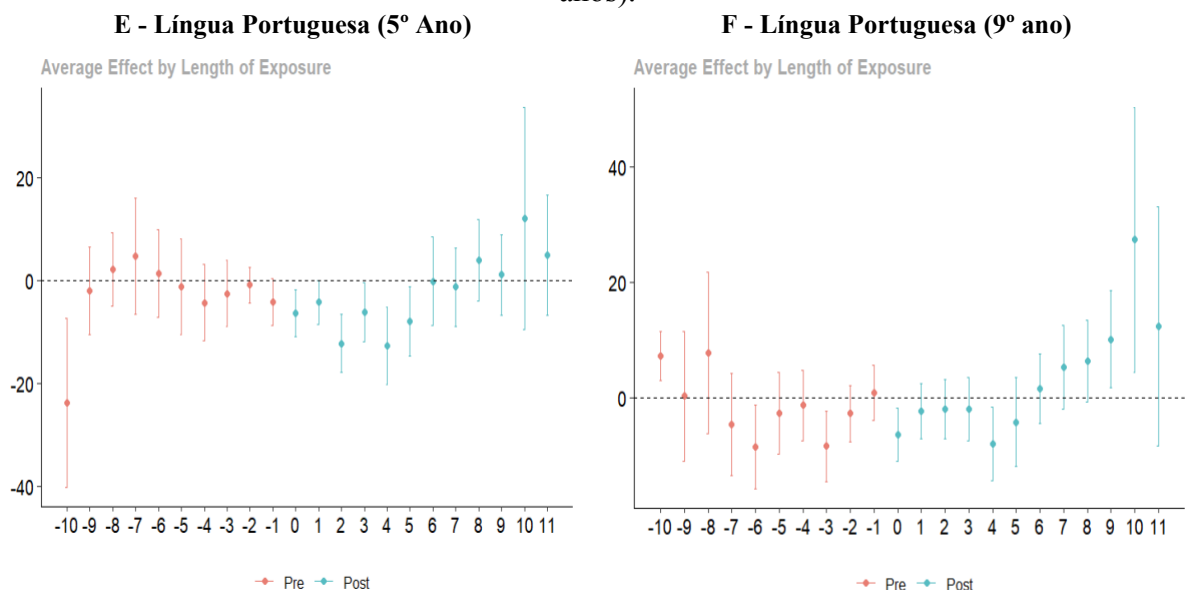


Fonte: Elaboração própria.

O gráfico C apresenta o estimador dinâmico em relação às notas de matemática do 5º ano. Destaca-se que os primeiros 5 anos após tratamento apresentaram resultados estatisticamente significativos e negativos. Posteriormente ocorreram resultados positivos, porém nenhum com significância estatística. Já o gráfico D refere-se às notas de matemática do 9º ano, em que se percebe trajetória crescente após o quarto ano, porém com significância estatística apenas para três períodos após o tratamento: o ano do tratamento, o quarto ano (-7,32) e o nono ano (11,51).

Os gráficos referentes à nota de Língua Portuguesa estão na Figura 10.

Figura 10: Estimador dinâmico de efeitos médios para nota de Língua Portuguesa (5º e 9º anos).



Fonte: Elaboração própria.

O gráfico E apresenta o estimador dinâmico para as notas de Língua Portuguesa do 5º ano. Com formato semelhante às demais análises prévias do 5º ano, os resultados foram estatisticamente significativos e negativos após o tratamento para o ano em que houve a municipalização (-6,32), o segundo ano (-12,19), terceiro ano (-6,18), quarto ano (-12,64) e quinto ano (-7,87).

Já o gráfico F apresenta as notas de Língua Portuguesa para o 9º ano. Percebe-se uma trajetória crescente após o quarto ano do tratamento com efeitos positivos após o sexto ano, porém com menor precisão, sendo estatisticamente significativo o resultado do ano do tratamento (-6,45), o quarto ano (-8,08), o nono ano (10,12) e o décimo ano (27,39).

Em síntese, o estimador dinâmico caracterizou-se por apresentar efeitos negativos, em maiores e menores graus, com significância estatística nos primeiros anos da municipalização. Posteriormente, evidenciou-se uma trajetória de efeitos positivos nos anos que representam as maiores exposições ao tratamento, porém com pouca precisão estatística.

Também pode-se concluir desta análise que, para todas as variáveis analisadas, o formato do efeito foi diferente a depender da série/ano analisado: enquanto para o 5º ano o efeito médio apresentou formato semelhante a U “invertido”, os resultados do 9º ano apresentaram trajetória positiva mais acentuada e sustentada com o decurso do tempo. Adicionalmente, deve-se destacar que todas as disciplinas avaliadas no 5º ano já apresentavam trajetória decrescente no período anterior à municipalização, o que não ocorreu para o 9º ano.

Em suma, a análise do efeito do tratamento por tempo de exposição permite captar se os efeitos médios aumentam ou diminuem com o decorrer do tempo em que as unidades estão expostas ao tratamento, nesse sentido, observa-se que os efeitos médios por tempo decorrido desde a alteração da gestão da unidade escolar tendem a ser negativos nos primeiros anos, posteriormente apresentando resultados positivos.

Contudo, um possível problema com as estimativas do efeito dinâmico é que a composição dos grupos é alterada de acordo com a exposição do tratamento. A fim de resolver essa problemática, podem ser utilizadas as estimativas com os grupos balanceados.

De acordo com Callaway e Sant’Anna (2021), o estimador dinâmico balanceado é composto apenas por grupos que são expostos ao tratamento por pelo menos um período definido de tempo sendo analisados apenas os efeitos dinâmicos nesses períodos selecionados.

Nesta pesquisa, a estatística referente ao estimador dinâmico balanceado considera apenas as unidades escolares que foram municipalizadas há pelo menos cinco anos, a fim de se detalhar a percepção verificada através do estimador dinâmico de que os primeiros anos da municipalização apresentam efeitos negativos.

Tabela 6: Estimador de efeitos médios dinâmico agregado balanceado para cinco anos.

Variável Dependente	Nota Padronizada 5º Ano	Nota Padronizada 9º Ano	Língua Portuguesa 5º Ano	Língua Portuguesa 9º Ano	Matemática 5º Ano	Matemática 9º Ano
<b>ATT</b>						
<b>Dinâmico</b>	-0,3105 *	-0,2282 *	-10,5546 *	-6,798 *	-10,9331 *	-6,8936 *
<b>Agregado</b>	(0,0433)	(0,0407)	(1,2411)	(1,3508)	(1,3952)	(1,2971)
<b>Balanceado</b>						

Fonte: Elaboração própria.

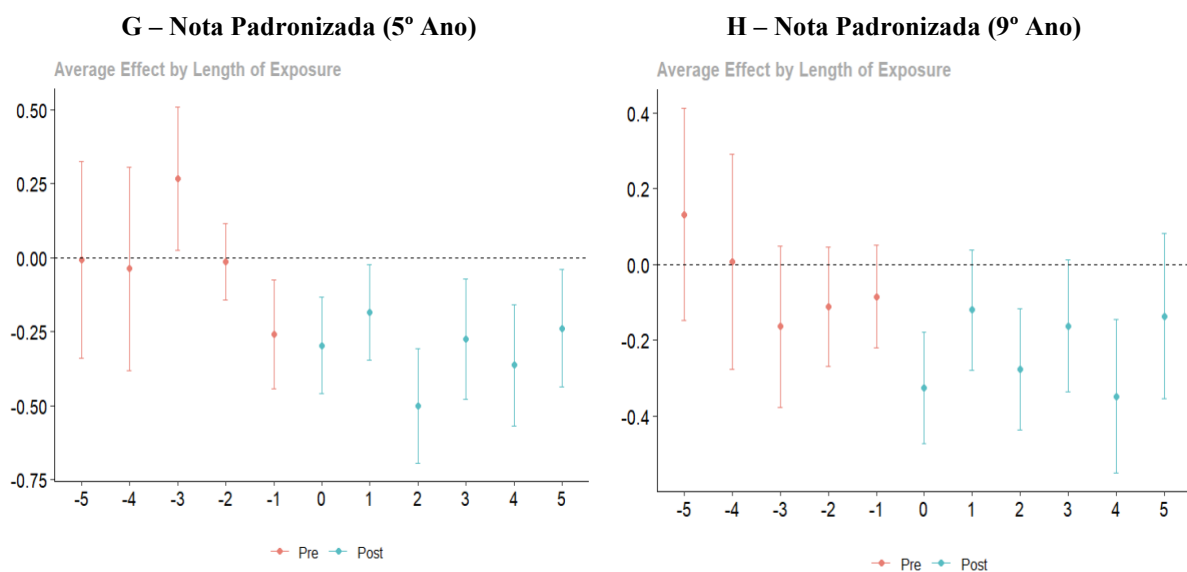
Notas: \* Significativo a 5%. Entre parênteses está apresentado o desvio padrão.

Sob a ótica do estimador dinâmico balanceado para cinco anos, todos os resultados foram negativos e significativos estatisticamente o que ratifica a conclusão obtida através do estimador dinâmico de que os primeiros anos após a municipalização geram impactos negativos no desempenho dos alunos.

Esse estimador apresentou os resultados do efeito da municipalização sobre as notas dos alunos em maior magnitude que os anteriores.

Em relação ao 5º ano, foram constatados impactos semelhantes para Língua Portuguesa (-10,55) e Matemática (-10,93) enquanto a nota padronizada apresentou efeito de -0,31. No 9º ano também foram equivalentes os efeitos em Língua Portuguesa (-6,79) e Matemática (-6,89) e o impacto na nota padronizada foi de -0,22. Em síntese, observaram-se impactos de maior magnitude nas turmas de 5º ano em relação as de 9º ano.

Figura 11: Estimador dinâmico balanceado (5 anos) de efeitos médios para nota padronizada (5º e 9º anos)

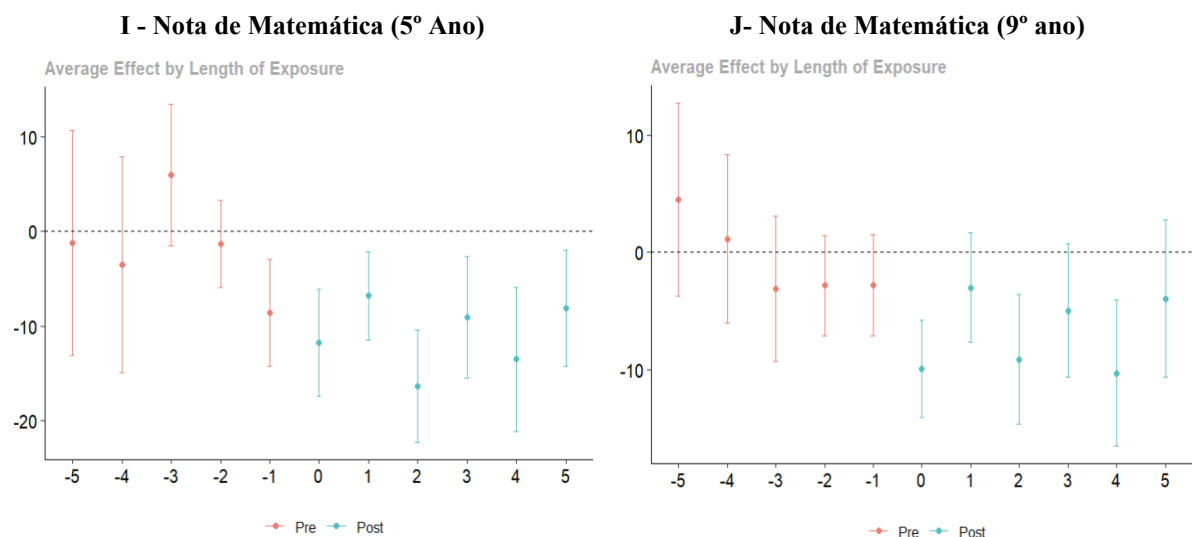


Fonte: Elaboração própria.

O gráfico G apresenta os efeitos dinâmicos balanceados para 5 anos referentes à nota padronizada do 5º Ano. Destaca-se que todos os resultados após o tratamento foram estatisticamente significativos e negativos, com impacto mais expressivo no segundo ano de tratamento (-0,5) e no quarto (-0,36). Em relação à nota padronizada para o 9º ano, os dados presentes no gráfico H indicam que todos os resultados após o tratamento foram negativos, mas apenas alguns apresentaram significância estatística: o ano do tratamento (-0,35), dois anos após (-0,27) e quatro anos após (-0,34).

Disso pode-se concluir que, ainda que com pouca precisão estatística e com resultados modestos, os primeiros cinco anos após a municipalização resultam em impactos negativos na nota padronizada tanto para o 5º quanto para o 9º ano.

Figura 12: Estimador dinâmico balanceado (5 anos) de efeitos médios para nota de Matemática (5º e 9º anos).

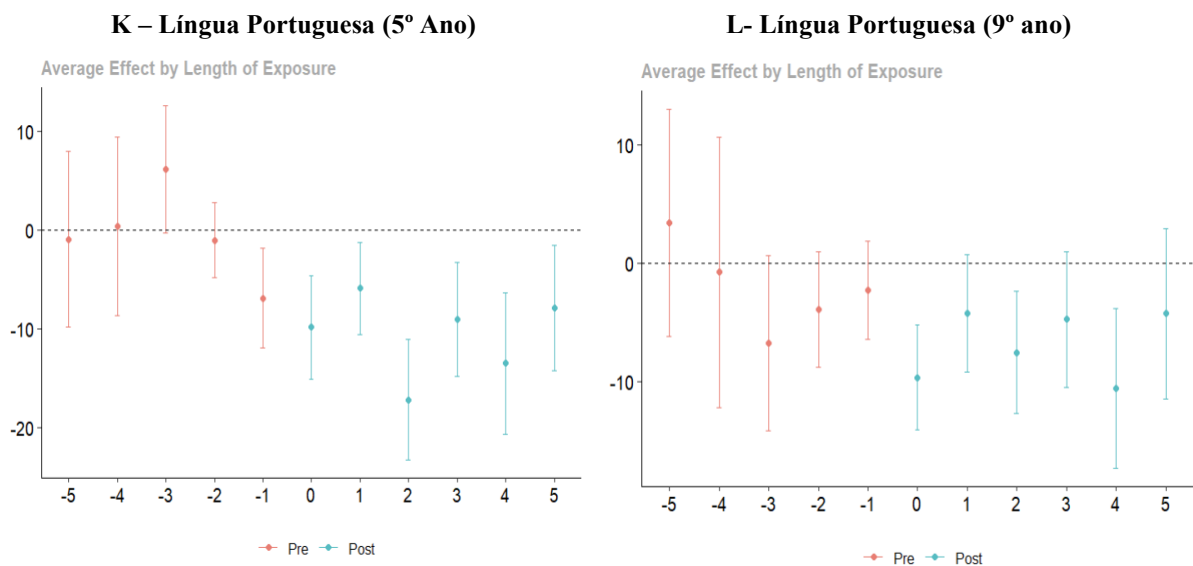


Fonte: Elaboração própria.

O gráfico I apresentou os resultados do estimador dinâmico balanceado para a nota de Matemática do 5º ano através do qual pode-se notar todos os resultados estatisticamente significativos e negativos após o tratamento, com efeitos de maiores magnitudes no ano do tratamento (-11,7), no segundo ano (-16,3) e no quarto ano (-13,5). O gráfico J é a representação para a nota Matemática do 9º ano, em que todos os resultados são negativos após o tratamento, mas há significância estatística apenas no momento do tratamento (-9,9), no segundo ano (-9,1) e no quarto ano (-10,3).

Portanto, percebe-se que há indícios de que, nos primeiros cinco anos, a municipalização impacte negativamente os indicadores de desempenho dos alunos na nota de Matemática do 5º e 9º, apresentando impacto negativo significativo no ano em que a escola é municipalizada.

Figura 13: Estimador dinâmico balanceado (5 anos) de efeitos médios para nota de Língua Portuguesa (5º e 9º anos).



Fonte: Elaboração própria.

O gráfico K representa os efeitos médios obtidos para as notas em Língua Portuguesa para o 5º ano. Todos os resultados foram estatisticamente significativos e negativos após o tratamento, com ênfase no ano do tratamento (-9,8), o segundo ano após o tratamento (-17,1) e no quarto ano (-13,5). Já o gráfico L apresenta os resultados para o 9º ano, em que todos os resultados são negativos após o tratamento, mas há significância apenas no ano do tratamento (-9,6), no segundo ano após o tratamento (-7,4) e no quarto ano (-10,5).

Assim como verificado na análise da nota padronizada e de matemática, há indícios de que nos cinco primeiros anos a municipalização afete, em alguma magnitude, negativamente os resultados nas avaliações de proficiência dos alunos em Língua Portuguesa tanto do 5º quanto do 9º ano.

Em síntese, a análise do estimador dinâmico balanceado para cinco anos corroborou o identificado pela análise do estimador dinâmico: em tese, os efeitos nos primeiros cinco anos após a municipalização são negativos independente da etapa de ensino e da disciplina avaliada.

Em suma, foram apresentados os resultados estimados para quatro medidas de efeito médio: efeito agregado simples, efeito por grupo individual e agregado, efeito agregado

dinâmico e efeito agregado dinâmico balanceado para cinco anos. Desses, muitas estimativas apresentaram ausência de significância estatística, não sendo possível rejeitar a hipótese de o impacto da municipalização ser igual a zero, efeitos consistentes com o indicado pela literatura que diagnosticou pouco ou nenhum impacto da política de municipalização sobre o desempenho escolar dos alunos.

Contudo, os resultados também sinalizam notas piores imediatamente nos primeiros anos após a municipalização tanto para Língua Portuguesa quanto Matemática, para o 5º e 9º ano.

Deve-se ressaltar que existem limitações importantes a serem consideradas. Primeiro, algumas das estimativas dos efeitos médios de tratamento de grupos em períodos pré-tratamento foram significativamente diferentes de zero, o que fornece alguma evidência sugestiva contra a suposição de tendências paralelas. Segundo, existe heterogeneidade na forma com que cada Estado e Município realizam a municipalização e as diretrizes que são assumidas variam a depender dos entes envolvidos, o que pode complicar a interpretação dos resultados. Com isso, sugere-se cautela ao interpretar os resultados aqui apresentados.

Outrossim, não se pode ignorar o fato de que avaliar a qualidade educacional ofertada pelas unidades educacionais e o respectivo aprendizado dos alunos envolvem alto grau de complexidade em virtude dos fatores capazes de afetar diretamente a aprendizagem e o desenvolvimento de um aluno estarem ligados ao seu background familiar.

## 5.2 ANÁLISE DE ROBUSTEZ

Robustez estatística é um conceito fundamental para compreender e interpretar métodos estatísticos, principalmente em estudos em que se utiliza dados incompletos, que podem conter outliers ou qualquer outro tipo de atipicidade.

Um método robusto é aquele em que sua performance não é altamente influenciada por pequenas alterações na base de dados ou por violações dos pressupostos assumidos, mantendo-se preciso, válido e confiável.

Conforme sugere Callaway e Sant'Anna, pode-se avaliar a robustez dos resultados alterando o grupo de controle de unidades “nunca tratadas” para as unidades “ainda não tratadas”. O objetivo é identificar se os resultados obtidos não são decorrentes de um viés de seleção do grupo de controle. Com essa análise, espera-se que os coeficientes obtidos sejam semelhantes aos encontrados quando se utiliza na estimação o grupo “nunca tratado”.

O Apêndice B apresenta os resultados das estimativas mantendo-se constantes todas as especificações utilizadas no tópico 4.1, porém alterando o grupo de controle e utilizando unidades escolares “ainda não tratadas” ao invés de escolas “nunca tratadas”.

Os resultados mostram que não houve mudanças nos sinais dos coeficientes, nem alterações relevantes em seus valores ou mudanças nas suas significâncias estatísticas. Ademais, também não houve grandes alterações nos desvios-padrão, corroborando os resultados encontrados anteriormente.

### 5.3 COMPARAÇÃO DE ESTIMADORES: CALLAWAY & SANT’ANNA X TWFE

Para compreender a dinâmica das diferentes formas de estimação, foram calculados os estimadores utilizando a metodologia TWFE, mantendo todas as covariáveis utilizadas no tópico 4.1. Os resultados encontrados são apresentados na Tabela 7.

Tabela 7: Estimador TWFE

Variável Dependente	Nota Padronizada 5º Ano	Nota Padronizada 9º Ano	Língua Portuguesa 5º Ano	Língua Portuguesa 9º Ano	Matemática 5º Ano	Matemática 9º Ano
Estimador TWFE	-0,13159 ** (0,021)	-0,555 ** (0,0162)	-3,7129 ** (0,5748)	-2,6089 ** (0,8725)	-3,338 ** (0,6482)	-1,5884 ** (0,5119)

Fonte: Elaboração própria.

Notas: \*\* Significativo a 1%. Entre parênteses está apresentado o desvio padrão.

Todos os resultados foram estatisticamente significativos e negativos, sinalizando que a municipalização potencialmente afeta negativamente o desempenho dos alunos.

A título de comparação, a Tabela 8 apresenta os estimadores agregados calculados de acordo com a metodologia Callaway e Sant’Anna e o estimador TWFE.

Tabela 8: Resumo dos efeitos médios agregados calculados de acordo com a metodologia Callaway e Sant’Anna e o estimador TWFE.

Variável Dependente	Nota Padronizada 5º Ano	Nota Padronizada 9º Ano	Língua Portuguesa 5º Ano	Língua Portuguesa 9º Ano	Matemática 5º Ano	Matemática 9º Ano
ATT Simples	-0,1176 * (0,0353)	-0,0136 (0,0392)	-4,8308 * (1,1611)	-0,6526 (1,1231)	-4,8661 * (1,2681)	-0,1624 (1,3236)
ATT Grupos Agregado	-0,0743 * (0,034)	0,0288 (0,0356)	-2,9932 * (1,0882)	0,6872 (1,2121)	-3,4365 * (1,1518)	1,0413 (1,3065)
ATT Dinâmico Agregado	-0,0285 (0,0438)	0,1205 * (0,053)	-2,3671 (1,4182)	3,1484 * (1,4883)	-2,4204 * (1,3981)	4,0804 * (1,9202)
ATT	-0,3105 * (0,0353)	-0,2282 * (0,0392)	-10,5546 * (1,1611)	-6,798 * (1,1231)	-10,9331 * (1,2681)	-6,8936 * (1,3236)



<b>Dinâmico</b>	(0,0433)	(0,0407)	(1,2411)	(1,3508)	(1,3952)	(1,2971)
<b>Agregado</b>						
<b>Balanceado</b>						
<b>ATT</b>	-0,13159 **	-0,555 **	-3,7129 **	-2,6089 **	-3,338 **	-1,5884 **
<b>TWFE</b>	(0,021)	(0,0162)	(0,5748)	(0,8725)	(0,6482)	(0,5119)

Fonte: Elaboração própria.

Notas: \* Significativo a 5%. \*\* Significativo a 1%. Entre parênteses está apresentado o desvio-padrão.

Em virtude das especificidades atribuídas ao estimador por grupos e dinâmico, optou-se por comparar o ATT Simples de Callaway e Sant'Anna com o ATT calculado pela metodologia TWFE.

Ambos sinalizam impactos negativos da municipalização para todas as séries/anos e disciplinas avaliadas. Contudo, percebe-se que os desvios-padrão obtido através do TWFE são inferiores e o nível de significância atribuído a este estimador (1%) é mais atrativo do que o encontrado através do método Calaway e Sant'Anna (5%). Dessa forma, o estimador TWFE informa com maior grau de certeza que o impacto da municipalização no desempenho dos alunos é negativo, ainda que modesto. Em contrapartida, a metodologia de Callaway e Sant'Anna sugere maior cautela ao apresentar muitos estimadores com ausência de significância estatística.

Não obstante, deve-se destacar que embora o estimador TWFE tenha corroborado o estimador simples e identificado resultados negativos, todos os resultados identificados nesta pesquisa não apresentaram magnitude substancial, o que ratifica o identificado pela literatura de não haver impactos diretos e imediatos da municipalização no desempenho escolar.

## 6 CONCLUSÃO

Este trabalho buscou evidenciar os efeitos da política pública da municipalização sobre o desempenho dos alunos. Para tanto, optou-se por utilizar a metodologia mais atual e robusta proposta por Callaway e Sant'Anna (2021), que considera a existência de múltiplos períodos de tempo e prevê diferentes formas de agregação para o efeito médio do tratamento, o que permite detalhar sob diversas óticas o problema analisado destacando a heterogeneidade dos efeitos do tratamento entre grupos e em decorrência do tempo de exposição.

Dessa forma, foram apresentadas novas estimativas sobre a relação entre a política pública de municipalização e a qualidade educacional ofertada - mensurada através do resultado dos estudantes em avaliações Saeb.

O estimador simples sugeriu um efeito geral da municipalização negativo próximo a zero, com maiores efeitos sobre o 5º ano. O estimador por grupos indicou que os grupos experimentaram efeitos ambíguos da municipalização: resultados negativos para o 5º ano e positivos para o 9º ano, todos em pequena magnitude, o que foi corroborado pelos resultados do estimador dinâmico.

A partir da análise do estimador dinâmico, também foi possível perceber que os primeiros anos da municipalização apresentaram efeitos negativos, com impactos positivos posteriores.

Essa percepção foi ratificada através da ótica do estimador dinâmico balanceado para cinco anos, em que todos os resultados foram negativos e estatisticamente significativos, do que se pode concluir que os primeiros anos após a municipalização geram impactos negativos no desempenho dos alunos.

Adicionalmente, buscou-se analisar essa problemática através do prisma da metodologia mais comumente utilizada em econometria para casos de adoção escalonada de tratamento: *Two-Way Fixed Effects*. Os resultados indicaram impactos negativos da municipalização, com significância estatística, porém em pouca magnitude.

Nesse sentido, o estimador simples de Callaway e Sant'Anna e o TWFE não divergiram em demasia em relação aos coeficientes e ambos sinalizaram a existência de impactos negativos da municipalização sobre o desempenho dos alunos.

Não obstante, todos os resultados identificados nesta pesquisa não apresentaram magnitude substancial, o que vai ao encontro do pressuposto de não haver impactos significativos da municipalização sobre desempenho escolar dos alunos identificado por Leme, Paredes e Souza (2009) e Ceneviva (2012).

Por estar amparada tanto na Constituição Federal de 1988 quanto na LDB, bem como em mecanismos de financiamento para auxiliar os municípios a suportarem o aumento de gastos decorrentes da expansão de ensino, é importante compreender o impacto da política de municipalização na proficiência dos alunos, uma vez que, a tendência é que seja aplicada em larga escala no Brasil. Ademais, a despeito da garantia do acesso e permanência dos estudantes nas unidades escolares, as políticas educacionais devem considerar em seu escopo a análise da qualidade do ensino ofertada.

Também deve-se considerar que não são todos os municípios que possuem capacidade para ampliar sua rede e expandir o número de matrículas mantendo a qualidade de ensino. Ao contrário, devido às disparidades evidentes existentes no Brasil, corre-se o risco de uma política

de municipalização resultar em uma queda da qualidade do ensino, como foi sinalizado por alguns estimadores obtidos neste trabalho.

De todo modo, é fundamental monitorar as unidades que passaram pela transferência de gestão para estimular que apresentem melhoria de desempenho.

Assim como a maioria das avaliações de impacto, esta pesquisa tem suas limitações. Além da análise do impacto da alteração de gestão sobre a educação ofertada aos estudantes ser complexa, sobretudo em virtude do fator de maior influência sobre o aprendizado dos estudantes ser o *background* familiar, destaca-se que é marcante a heterogeneidade entre os Estados e Municípios brasileiros, o que dificulta uma análise conjunta e sugere que as unidades escolares sejam analisadas a partir de um menor nível de agregação.

De toda forma, essa pesquisa contribui com a literatura dos impactos da municipalização sobre o desempenho escolar e os resultados aqui apresentados podem ser úteis aos formuladores de políticas auxiliando a definir estratégias capazes de impactar positivamente na aprendizagem dos estudantes.

Destaca-se que pesquisas futuras podem focalizar em diferentes aspectos da municipalização, por exemplo, seu impacto sobre taxas de evasão e distorção idade/série além de buscar compreender quais são os aspectos da gestão que possuem maiores impactos sobre o ensino dos alunos.

## REFERÊNCIAS

BOTH, Ivo, José. **Municipalização da Educação: uma contribuição para um novo paradigma de gestão do ensino fundamental**. Campinas, SP: Papirus, 1997. (Coleção Magistério Formação e trabalho pedagógico).

BRASIL. Congresso Nacional. **Lei nº 9.394, 20 de dezembro de 1996**. Estabelece as diretrizes de bases da educação nacional. Brasília: Câmara dos Deputados, 1996. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/l9394.htm](https://www.planalto.gov.br/ccivil_03/leis/l9394.htm)

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil**. Brasília, DF: Senado Federal, 2016. 496 p. Disponível em: [https://www2.senado.leg.br/bdsf/bitstream/handle/id/518231/CF88\\_Livro\\_EC91\\_2016.pdf](https://www2.senado.leg.br/bdsf/bitstream/handle/id/518231/CF88_Livro_EC91_2016.pdf)

BRASIL. Ministério da Educação. **Lei nº 9.424, de 24 de dezembro de 1996**. Dispõe sobre o Fundo de Manutenção e Desenvolvimento do Ensino Fundamental e de Valorização do Magistério, na forma prevista no artigo 60, § 7º, do Ato das Disposições Constitucionais Transitórias, e dá outras providências. Brasília, 1996a. Disponível em: <http://www.mec.gov.br/sef/fundef/Legisla>.

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

CAVALCANTI, Mônica Maria de Arruda. Avaliação de políticas públicas e programas governamentais: uma abordagem conceitual. **Interfaces de Saberes**, v. 6, n. 1, 2006.

CENEVIVA, RICARDO. O nível de governo importa para a qualidade da política pública? A municipalização da educação fundamental no Brasil. **8º Encontro da ABCP**, 2012.

CUNNINGHAM, Scott. **Causal inference: The mixtape**. Yale university press, 2021.

D'ATRI, Fabiana. **Municipalização do Ensino Fundamental da rede pública: os impactos sobre o desempenho escolar**. Tese de Doutorado. Fundação Getúlio Vargas, São Paulo, 2007.

BAPTISTA, Patrícia Ferreira A política de municipalização do ensino fundamental: limites ao controle de políticas públicas de educação. **Revista Eletrônica da PGE-RJ**, v. 5, n. 1, 2022.

FERREIRA, C. S.; SANTOS, E. N. DOS. Políticas Públicas Educacionais: apontamentos sobre o direito social da qualidade na educação. **Revista Labor**, v. 1, n. 11, p. 143-155, 2014.

FRANCO, C. M. G. A Política Pública do Fundeb em Ação. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 9, n. 7, p. 364–377, 2023.

GOODMAN-BACON, Andrew. Difference-in-differences with variation in treatment timing. **Journal of Econometrics**, v. 225, n. 2, p. 254-277, 2021.

GUIMARÃES, José Luiz. **Desigualdades Regionais na Educação: A municipalização do ensino em São Paulo**. São Paulo, Editora da Universidade Estadual Paulista, 1995.

WOOLDRIDGE, Jeffrey M. **Introdução à econometria: uma abordagem moderna**. Cengage Learning, 2023.

ROTH, Jonathan et al. What's trending in difference-in-differences? A synthesis of the recent econometrics literature. **Journal of Econometrics**, 2023.

LEME, Maria Carolina; PAREDES, Ricardo; SOUZA, André Portela. A municipalização do ensino fundamental e seu impacto sobre a proficiência no Brasil. **Educação básica no Brasil: construindo o país do futuro**. Rio de Janeiro: Campus-Elsevier, p. 261-280, 2009.

MENDES SEGUNDO, Maria das Dores; RABELO, Josefa Jackline A avaliação do Banco Mundial sobre a educação municipal no Brasil: incursões críticas. In: CONGRESSO INTERNACIONAL EM AVALIAÇÃO EDUCACIONAL, 2., 17 a 19 nov. 2005, Fortaleza (CE). **Anais [...]** Fortaleza (CE): UFC/FACED/NAVE, 2005.

MENDES, Cláudia Maria Sales. **A educação e as políticas de municipalização no Ceará**. Tese de doutoramento em Ciências da Educação (especialidade em Sociologia da Educação e Política Educativa) – Universidade do Minho, Portugal, 2022.

NASCIMENTO, E. A. e S.; SANTOS, G. V. dos; CRUVINEL, S. P.; SILVA, J. M. da. Políticas Públicas Educacionais: Educação Básica e de Qualidade. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 9, n. 7, p. 15–30, 2023.

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

RUBIN, Donald B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of educational Psychology**, v. 66, n. 5, p. 688, 1974.

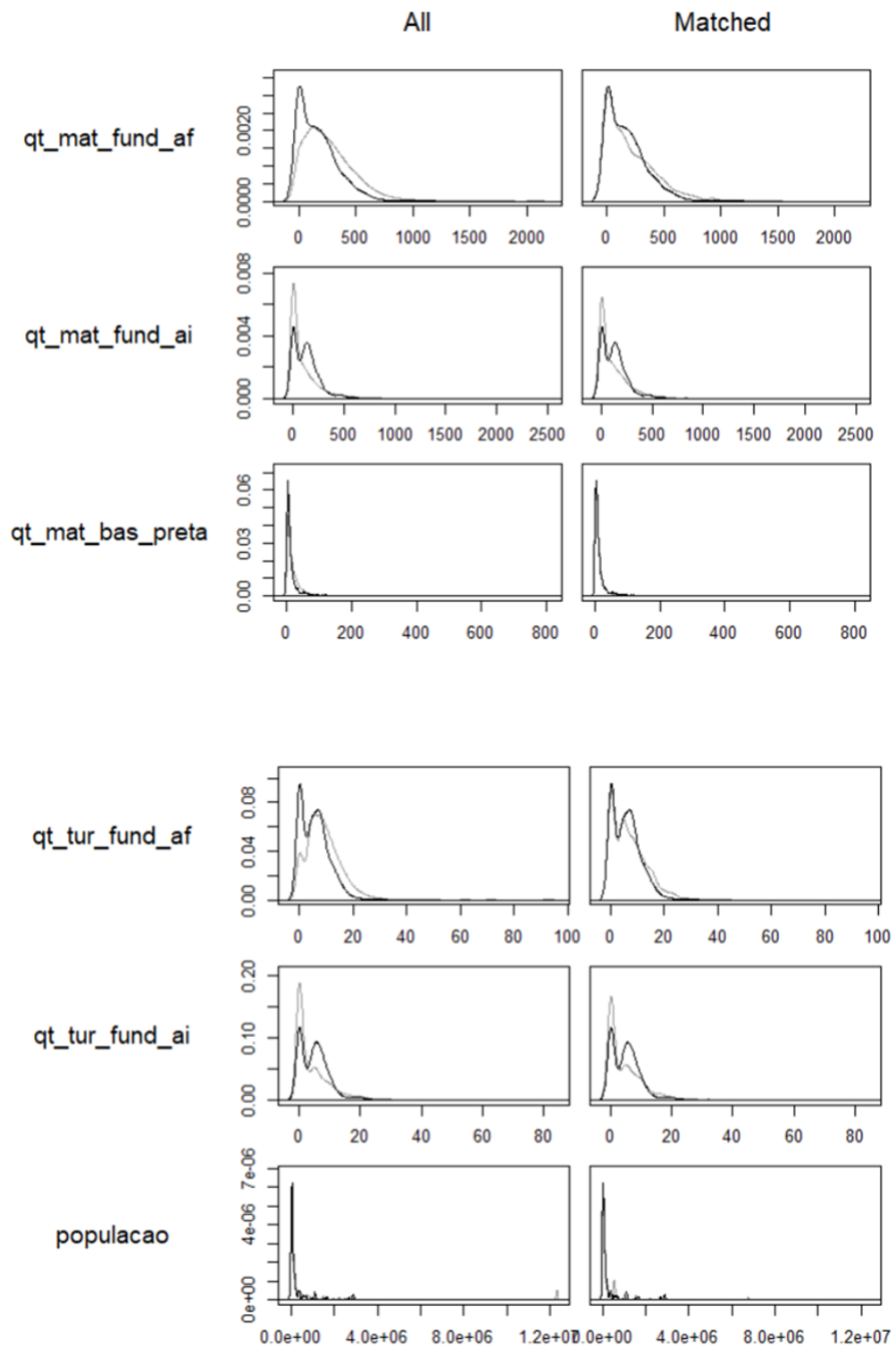
SANTOS, H. C. de S.; VASCONCELOS, J. Políticas Públicas Educacionais e a Responsabilidade Estatal. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, v. 9, n. 5, p. 2649–2658, 2023.

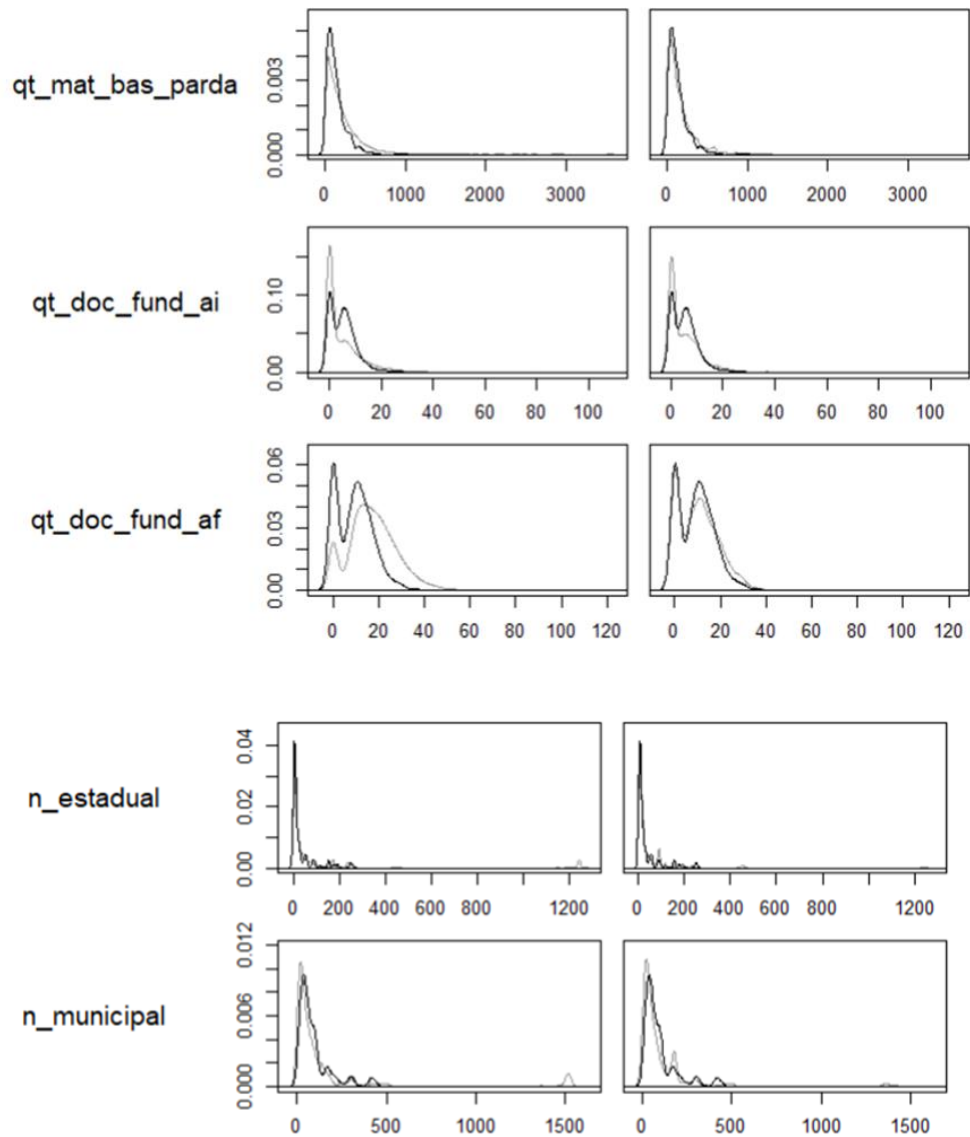
SOARES, ERINEUDA; AMARAL, D. O.; VIDAL, ELOISA. Reordenamento de escolas e municipalização do ensino: o que dizem as pesquisas. **Revista Diálogo Educacional**, v. 22, n. 75, p. 1995-2022, 2022.

SOUZA, Celina. Políticas públicas: uma revisão da literatura. **Sociologias**, p. 20-45, 2006.

TEIXEIRA, Anísio. A municipalização do ensino primário. **Revista do Serviço Público**, v. 75, n. 3, p. 347-364, 1957.

APÊNDICE A – Gráficos de Distribuição de Densidade das variáveis utilizadas no *Propensity Score Matching* (antes e depois do pareamento)





Notas: Os gráficos à esquerda (All) demonstram a distribuição de densidade das variáveis selecionadas para o *Propensity Score Matching*, sendo que a linha mais escura representa o grupo tratado e a linha mais clara, o grupo de controle. Em geral, nota-se distribuições com formatos distintos para os dois grupos no primeiro momento. Já os gráficos à direita (Matched) apresentam a distribuição de densidade dos dois grupos após o pareamento, cujo intuito essencial é tornar os grupos de tratamento e controle mais semelhantes. Nota-se que, na maioria dos casos, o PSM resultou em uma aproximação das distribuições da população tratada e de controle, o que se concluir ser um resultado benéfico.

**APÊNDICE B – Análise de Robustez: efeitos médios do tempo de tratamento para as diferentes variáveis analisadas**

<b>Variável Dependente</b>	<b>Nota Padronizada 5º Ano</b>	<b>Nota Padronizada 9º Ano</b>	<b>Língua Portuguesa 5º Ano</b>	<b>Língua Portuguesa 9º Ano</b>	<b>Matemática 5º Ano</b>	<b>Matemática 9º Ano</b>
<b>ATT simples</b>	-0,1121 * (0,0356)	-0,0129 (0,0377)	-4,8358 * (1,1612)	-0,6287 (1,23)	-4,8637 * (1,2366)	-0,1459 (1,2634)
<b>ATT Dinâmico Agregado</b>	-0,0206 (0,045)	0,121 * (0,0546)	-2,3698 (1,3137)	3,1664 * (1,5188)	-2,418 (1,4226)	4,0931 * (2,0792)
<b>ATT Dinâmico Balanceado</b>	-0,3056 * (0,0382)	-0,227 * (0,0396)	-10,569 * (1,2505)	-6,7542 * (1,261)	-10,9336 * (1,3495)	-6,8629 * (1,3111)
<b>Grupos</b>						
<b>2010</b>	-0,1509 (0,0922)	0,0692 (0,1210)	-5,6977 (2,6108)	0,9696 (3,7022)	-6,2186 (2,6977)	3,1843 (4,0085)
<b>2011</b>	0,0508 (0,2097)	0,2306 (0,1569)	-1,1124 (5,0813)	5,6082 (4,2549)	-1,1972 (6,6132)	8,2300 (6,1249)
<b>2012</b>	-0,2552 * (0,0749)	-0,0849 (0,0815)	-10,4775 * (2,3573)	-3,3017 (2,3288)	-8,7901 * (2,1788)	-1,7929 (2,5255)
<b>2013</b>	-0,2300 * (0,0587)	-0,1949 * (0,0615)	-9,8686 * (1,8230)	-5,6164 (2,0958)	-9,1517 * (2,0534)	-6,0794 (2,0678)
<b>2014</b>	0,1247 (0,1461)	0,1202 (0,1091)	5,1072 (4,0116)	4,4624 (3,6457)	3,3110 (3,9904)	2,7508 (3,8027)
<b>2015</b>	0,0583 (0,1738)	-0,0394 (0,1013)	2,8575 (4,8815)	-0,7607 (3,0448)	1,5397 (4,7694)	-1,6003 (3,6372)
<b>2016</b>	0,3073 (0,1586)	0,1461 (0,1690)	7,5896 (4,4166)	5,2634 (5,1083)	8,1216 (4,1838)	3,5056 (5,6770)
<b>2017</b>	-0,1733 (0,1259)	0,4319 (0,2043)	-5,0728 (3,8208)	9,4238 (6,8419)	-5,1609 (3,8744)	16,4925 (5,4397)
<b>2018</b>	0,3814 (0,1920)	0,4418 * (0,1158)	14,8017 (6,4117)	11,7511 (3,9176)	10,4680 (4,9057)	14,7543 (3,7051)
<b>2019</b>	0,1528 (0,1148)	0,4320 * (0,1017)	10,5298 * (3,9002)	12,3123 * (3,0068)	3,9983 (4,4531)	13,6064 (3,2601)
<b>2020</b>	-0,1597 (0,1653)	-0,1059 (0,1443)	-5,9426 (4,4514)	-3,2344 (3,8327)	-3,9046 (3,8698)	-3,1180 (4,3948)
<b>2021</b>	0,1644 (0,2247)	-0,3174 (1,1836)	6,6996 (6,9313)	4,7886 (19,6298)	1,4368 (7,0401)	-23,8343 (6,8065)
<b>ATT Grupos Agregado</b>	-0,0714 * (0,0342)	0,0293 (0,038)	-2,9965 * (1,2035)	0,7057 (1,1473)	-3,4338 * (1,2077)	1,054 (1,2069)

Fonte: Elaboração própria.

Notas: \* Significativo a 5%. Entre parênteses está apresentado o desvio padrão. Os dados apresentados na tabela foram calculados através do modelo Callaway & Sant’Anna (2021) com as mesmas especificações apresentadas nos resultados desta pesquisa, com exceção de que o grupo de controle é composto por unidades “ainda não tratadas” ao invés de “nunca tratadas”.