

UNIVERSIDADE DE SÃO PAULO FACULDADE DE ECONOMIA,
ADMINISTRAÇÃO E CONTABILIDADE DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Rafael Pereira Oliveira

**Até Onde foi a Lama? Análise dos Efeitos do
Rompimento da Barragem de Rejeitos em
Mariana (MG) em Indicadores Educacionais**

São Paulo

2021

Prof. Dr. Vahan Agopyan
Reitor da Universidade de São Paulo

Prof. Dr. Fábio Frezatti
Diretor da Faculdade de Economia, Administração e Contabilidade

Prof. Dr. José Carlos de Souza Santos
Chefe do Departamento de Economia

Prof. Dr. Wilfredo Fernando Leiva Maldonado
Coordenador do Programa de Pós-Graduação em Economia

UNIVERSIDADE DE SÃO PAULO FACULDADE DE ECONOMIA,
ADMINISTRAÇÃO E CONTABILIDADE DEPARTAMENTO DE ECONOMIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

Rafael Pereira Oliveira

**Até Onde foi a Lama? Análise dos Efeitos do
Rompimento da Barragem de Rejeitos em Mariana (MG)
em Indicadores Educacionais**

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Ciências.

Orientador: Prof. Dr. Ariaster Baumgratz Chimeli

Versão Corrigida

São Paulo

2021

Catálogo na Publicação (CIP)
Ficha Catalográfica com dados inseridos pelo autor

Oliveira, Rafael Pereira

Até Onde foi a Lama? Análise dos Efeitos do Rompimento da Barragem de Rejeitos em Mariana (MG) em Indicadores Educacionais / Rafael Pereira Oliveira. – São Paulo, 2021.

67p.

Dissertação (Mestrado) - Universidade de São Paulo, 2021.
Orientador: Prof. Dr. Ariaster Baumgratz Chimeli.

1. Desastres naturais. 2. Rompimento de barragem de rejeitos.
3. Educação. 4. Diferenças-em-Diferenças. 5. Controle Sintético.
I. Universidade de São Paulo. Faculdade de Economia, Administração e Contabilidade. II. Título.

Rafael Pereira Oliveira

**Até Onde foi a Lama? Análise dos Efeitos do
Rompimento da Barragem de Rejeitos em Mariana (MG)
em Indicadores Educacionais**

Dissertação apresentada ao Programa de Pós-Graduação em Economia do Departamento de Economia da Faculdade de Economia, Administração e Contabilidade da Universidade de São Paulo, como requisito parcial para a obtenção do título de Mestre em Ciências.

Prof. Dr. Ariaster Baumgratz Chimeli
FEA/USP

Prof. Dr. Naércio Menezes
FEA/USP

Prof. Dr. Reynaldo Fernandes
FEA-RP/USP

Profa. Dra. Cristine Pinto
EESP/FGV

São Paulo

2021

Agradecimentos

Agradeço, em especial, aos meus pais, Glória Oliveira e Gesner Oliveira, por todo o apoio incondicional. Este trabalho e todo meu mestrado só foram possíveis graças a vocês. Serei eternamente grato por tudo o que fizeram e fazem por mim. Agradeço também ao meu irmão Gabriel, ao meu tio Mota, à minha tia Carmem e à toda minha família pelo carinho de sempre. O mesmo vale para Iracema, Luciana e Edmilson, além da Karla e do Chico, quem também considero parte da família e que cuidam tanto de mim.

Agradeço ao meu orientador, Ariaster Chimeli, por todas as dicas e correções ao trabalho desde o início, bem como pelo tempo precioso disponibilizado para me orientar. As várias conversas que tivemos foram fundamentais para a estruturação deste trabalho. Espero ter sido capaz de ter reproduzido nas próximas páginas ao menos parte das suas excelentes sugestões. Aproveito para agradecer a todo o grupo de orientandos de Economia do Meio Ambiente pelos ricos encontros e pelas ótimas ideias para minha pesquisa. Em especial, agradeço à Liz Matsunaga pelo apoio com os dados e a definição do tema.

Agradeço também à minha parceira, Laura, pelo amor, carinho e companhia ao longo de todo este trabalho. As longas horas de estudo nestes mais de dois anos de mestrado foram muito mais fáceis e divertidas com você. Quando as piadas ficavam ruins e a gente começava a cantar era sinal de que o estudo já não estava rendendo. Foi importante para saber quando parar. Você foi sem dúvida quem mais acompanhou todos os desafios, angústias e alegrias desta difícil, mas importante etapa que se encerra com este trabalho.

Dedico especial agradecimento aos meus colegas e amigos Rodrigo Pacheco, Lucas Chang, Pedro Sayon, Rafael Misso e José Matheus pela parceria e companhia ao longo de todo o mestrado, pelas risadas e pelas ricas discussões. Esta difícil jornada foi certamente muito mais fácil ao lado de vocês. Ao Rodrigo, um agradecimento especial pela parceria no trabalho de Econometria, que serviu de base para parte desta dissertação, bem como pelas aulas gratuitas de programação. Ao Lucas pela parceria no futebol e por financiar meu tempo livre jogando FIFA e Cartola FC. Ao Pedro pelas corridas na USP e pela importante aula de mapas, sem falar dos vários projetos paralelos que realizamos juntos.

Agradeço a todos meus amigos e demais pessoas que me cercam, incluindo Duda e Ciça. Sem a companhia e apoio de todos nada do que faço seria possível. Por fim, agradeço à Fundação Instituto de Pesquisas Econômicas (FIPE) e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo financiamento a este trabalho.

Resumo

No dia 5 de novembro de 2015, rompeu-se no município de Mariana, Minas Gerais (MG), uma barragem de rejeitos de minério denominada “Fundão”. A barragem armazenava cerca de 56 milhões de metros cúbicos de rejeitos, dos quais cerca de 39 milhões foram liberados. O desastre é considerado o pior da história em termos de volume despejado de rejeitos de mineração. Estima-se que ao menos 40 municípios foram impactados na região da Bacia do Rio Doce. Nota-se na literatura acerca de desastres naturais três canais de influência negativa na Educação: (i) danos à infraestrutura física; (ii) choques negativos de renda; e (iii) danos psicológicos derivados de transtornos de estresse pós-traumático (TEPT). O desastre em Mariana (MG) contempla tais canais de influência, no entanto, há poucos estudos analisando com profundidade os seus impactos na Educação. O objetivo desta dissertação é avaliar se existe uma relação causal entre o rompimento da barragem de Fundão em Mariana (MG) e o desempenho escolar de alunos dos anos finais do ensino fundamental nos municípios afetados. O trabalho baseia-se em indicadores educacionais disponíveis nas bases do Censo Escolar e do Sistema de Avaliação de Educação Básica (Saeb). São analisados dados a nível de município, escola e aluno. O grau de exposição à lama permite uma análise de heterogeneidade geográfica dos impactos. Estima-se tanto um modelo de Diferenças em Diferenças (DD) quanto um de Controle Sintético (CS) específico para Mariana (MG). A desagregação por áreas dos municípios tratados de acordo com a exposição à lama aponta para a importância de as políticas públicas pós-desastre levarem em conta as particularidades das regiões impactadas. Já os resultados a nível de escola sugerem que pode ter havido um aumento da taxa de reprovação nos anos finais do ensino fundamental nos municípios afetados pela lama, mas sem ocupação próxima ao Rio Doce. Nos municípios próximos à barragem de Fundão, observa-se efeito negativo no desempenho médio de Matemática. Em Mariana (MG), de acordo com o método do CS, verifica-se que o desastre elevou a taxa de reprovação nos anos finais do ensino fundamental. Os resultados, no entanto, devem ser vistos com cautela.

Palavras-chave: Desastres naturais; Rompimento de barragem de rejeitos; Educação; Diferenças-em-Diferenças; Controle Sintético.

Abstract

On November 5, 2015, a tailings dam failure occurred at Mariana, in the Brazilian state of Minas Gerais (MG). The so-called “Fundão” dam stored 56 million cubic meters of mud-shaped tailings, of which around 39 million were released. The socio-environmental disaster is considered the worst accident in history in terms of drained volume of mine tailings. At least 40 municipalities were affected in the Doce River watershed region, including both MG and Espírito Santo (ES) states. The available literature regarding natural disasters points to three main channels of negative effects on Education: (i) infrastructure damages; (ii) income negative shocks; and (iii) psychological damages derived from post-traumatic stress disorders (PTSD). It can be seen that Mariana (MG) disaster contemplates such channels, however, there are few studies analyzing in greater depth its impacts on Education. The purpose of this study is to test whether there is a causal relationship between the Fundão dam failure in Mariana (MG) and the academic performance of Middle School students in the affected municipalities. The research is based on educational indicators available at Censo Escolar and Sistema de Avaliação de Educação Básica (Saeb) database. Data is analyzed at the municipality, school and student level. The degree of mud exposure allows a geographical heterogeneity analysis of the disaster impacts. Both a Difference-in-Difference (DD) and a Synthetic Control (SC), focused in Mariana (MG) municipality, models are undertaken. The geographical disaggregation according to mud exposure though points to heterogeneous effects highlighting the importance of post-disaster public policy to take into account the region-specific characteristics. The results at the school level suggest that disapproval rates in Middle Schools may have risen in the municipalities affected by the mud, but without occupation next to the Doce river. In the municipalities next to the Fundão dam we observe negative effects on mathematics average performance. In Mariana (MG), according to the SC approach, we also find evidence that the disaster raised the disapproval rates. The results, however, must be seen with cautiousness.

Keywords: Natural disasters; Tailings dam failure; Education; Difference-in-Difference; Synthetic Control.

Sumário

1	Introdução	1
2	Revisão de Literatura	3
3	Contexto	6
4	Dados	12
5	Estratégia Empírica	18
5.1	Diferenças em Diferenças	18
5.1.1	Escolas e Municípios	23
5.1.2	Alunos	25
5.2	Controle Sintético	26
6	Resultados	31
6.1	Diferenças em Diferenças	31
6.1.1	Escolas e Municípios	34
6.1.2	Alunos	40
6.2	Controle Sintético	41
7	Conclusão	51

Lista de ilustrações

Figura 1 – Trajetória da lama	7
Figura 2 – Categorização COPPETEC	8
Figura 3 – Taxas de Reprovação - Alunos - Categorização COPPETEC	17
Figura 4 – Grupo de Controle Base - Nível Municipal	20
Figura 5 – Grupo de Controle <i>threshold</i> 5% - Nível Municipal	20
Figura 6 – Grupo de Controle <i>threshold</i> 2,5% - Nível Municipal	21
Figura 7 – <i>Event Study</i> - Taxa de Reprovação Anos Finais EF - Escolas	32
Figura 8 – <i>Event Study</i> - Taxa de Reprovação Anos Finais EF - Escolas - Área 3	32
Figura 9 – Mariana vs. Bacia do Rio Doce (MG e ES) - Taxa de Reprovação Anos Finais do EF	41
Figura 10 – Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Todos os <i>lags</i> da <i>outcome variable</i>	44
Figura 11 – Gaps na Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Todos os <i>lags</i> da <i>outcome variable</i>	45
Figura 12 – Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Variação no <i>donor pool: threshold</i> de 2,5% - Todos os <i>lags</i> da <i>outcome variable</i>	46
Figura 13 – Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética (2015 pós) - Todos os <i>lags</i> da <i>outcome variable</i>	47
Figura 14 – Gaps na Taxa de Reprovação Anos Finais EF de Mariana (MG) vs. Gaps Placebo nos 167 municípios da Bacia do Rio Doce - Todos os <i>lags</i> da <i>outcome variable</i>	48
Figura 15 – Gaps na Taxa de Reprovação Anos Finais EF de Mariana (MG) vs. Gaps Placebo em 52 municípios da Bacia do Rio Doce - Todos os <i>lags</i> da <i>outcome variable</i>	49
Figura 16 – Distribuição da razão MSPE pós-tratamento/MSPE pré-tratamento - Todos os <i>lags</i> da <i>outcome variable</i>	50
Figura 17 – Escola Municipal de Bento Rodrigues, subdistrito de Mariana (MG), após o desastre	58
Figura 18 – Taxas de Reprovação - Escolas - Categorização COPPETEC	59
Figura 19 – Movimentação dos alunos entre 2015 e 2016	60
Figura 20 – Movimentação dos alunos entre 2015 e 2017	60
Figura 21 – <i>Event Study</i> - Taxa de Reprovação Anos Finais EF - Escolas - Área 1	61
Figura 22 – <i>Event Study</i> - Probabilidade de Reprovação Anos Finais EF - Alunos	61

Figura 23 – Gaps na Taxa de Reprovação Anos Finais EF - Mariana (MG) vs.
Mariana (MG) sintética - Variação no *donor pool*: *threshold* de 2,5% -
Todos os *lags* da *outcome variable* 67

Lista de tabelas

Tabela 1 – Teste de Médias: Tratados vs. Controles - Escolas e Municípios - Pré-desastre	14
Tabela 2 – Taxa de Reprovação Anos Finais EF - Escolas	35
Tabela 3 – Análise de Heterogeneidade - Taxa de Reprovação Anos Finais EF - Escolas	36
Tabela 4 – Desempenho Matemática 5º Ano do EF - Escolas	37
Tabela 5 – Desempenho Português 5º Ano do EF - Escolas	39
Tabela 6 – Médias das covariadas para Taxa de Reprovação - Anos Finais do EF	42
Tabela 7 – Peso dos municípios na Mariana (MG) sintética	43
Tabela 8 – Categorização COPPETEC - Lista de Municípios	58
Tabela 9 – Teste de Médias: Tratados vs. Controles - Alunos - Pré-desastre . . .	59
Tabela 10 – Taxa de Reprovação Anos Finais EF - Municípios	62
Tabela 11 – Desempenho Matemática 5º Ano do EF - Municípios	63
Tabela 12 – Desempenho Português 5º Ano do EF - Municípios	64
Tabela 13 – Probabilidade de Reprovação - Alunos	65
Tabela 14 – Análise de Heterogeneidade - Probabilidade de Reprovação - Alunos .	66

1 Introdução

No dia 5 de novembro de 2015, rompeu-se no município de Mariana, Minas Gerais, uma barragem de rejeitos de minério denominada “Fundão”. Controlada pela mineradora Samarco Mineração S.A., a barragem armazenava cerca de 56 milhões de metros cúbicos de rejeitos, dos quais 39,2 milhões foram liberados, volume equivalente ao de 15 mil piscinas olímpicas. Trata-se do pior desastre da história em termos de volume de material despejado por barragem de rejeitos de mineração.¹

A lama oriunda da barragem de Fundão engoliu subdistritos vizinhos e destruiu mais de 1.400 hectares ao longo da Bacia do Rio Doce, localizada nos Estados de Minas Gerais e Espírito Santo. Foram 19 mortos e aproximadamente 400 famílias desalojadas em decorrência do desastre. Estima-se que ao menos 40 dos 229 municípios da bacia tenham sido afetados, cerca de 500 mil pessoas.

Além dos conhecidos impactos econômicos e ambientais do referido desastre, há inúmeras evidências de efeitos negativos na área de Educação. De acordo com o Sistema Integrado de Informações sobre Desastres (S2ID), do Ministério de Desenvolvimento Regional, foram identificados danos socioeconômicos em diversos serviços públicos, dentre eles “Ensino” (Ibama, 2015). Conforme Ramboll (2017), o desastre provocou diferentes danos ao sistema educacional, dos quais se destacam: (i) perdas de infraestrutura de educação: escolas total ou parcialmente destruídas; (ii) desagregação de teias sociais existentes nas comunidades escolares; e (iii) interrupção do funcionamento de escolas gerando danos psicopedagógicos às comunidades escolares.

No entanto, não há nenhum estudo analisando com profundidade tais efeitos. Além do mais, a literatura acerca de efeitos de desastres tecnológicos é bastante escassa.² Este trabalho visa a preencher estas lacunas. O objetivo é verificar se há uma relação causal entre o desastre e o desempenho de indicadores educacionais disponíveis nas bases do Censo Escolar e Sistema de Avaliação de Educação Básica (Saeb).³ Para tanto, são testados dois modelos econométricos: um modelo de Diferenças-em-Diferenças (DD) e um modelo de Controle Sintético (CS).

¹Índice desenvolvido pela Bowker Associates, disponível em: <<https://lindsaynewlandbowker.wordpress.com/2015/12/12/samarco-dam-failure-largest-by-far-in-recorded-history/>>. Acesso em: 10/09/2019.

²Desastre tecnológico ou “socioambiental” é entendido neste trabalho como qualquer desastre causado, direta ou indiretamente, pela ação humana. Desastres naturais, por outro lado, são aqueles associados a fenômenos naturais sob os quais os seres humanos não têm controle.

³Base de dados do Censo Escolar database está disponível em: <<http://portal.inep.gov.br/censo-escolar>>; já a base do Saeb está disponível em: <<http://portal.inep.gov.br/educacao-basica/saeb>>.

Os resultados agregados a nível municipal encontrados por meio do método de DD assemelham-se aos da literatura disponível (FGV, 2019). Quando se considera o conjunto de 46 municípios tratados contra os de controle, não parece haver efeitos estatisticamente significativos do desastre na Educação. As variáveis dependentes analisadas foram taxa de reprovação nos anos finais do ensino fundamental, bem como desempenho médio em Matemática e Português no 5º ano do ensino fundamental. Ao nível de escola as estimações agregadas apresentam resultados similares. No caso da taxa de reprovação, no entanto, a análise de heterogeneidade indica que pode ter havido efeitos distintos entre escolas rurais e urbanas. Não se observa efeitos nem nas escolas privadas nem nas públicas. Por sua vez, os resultados desagregados, que discriminam os locais afetados de acordo com sua exposição à lama, apontam no sentido de que o desastre pode ter gerado efeitos regionais heterogêneos.

No caso dos resultados do método de CS, específico para o município de Mariana (MG), são encontradas evidências de aumento da taxa de reprovação associado ao desastre, em especial nos anos de 2016 e 2017. No entanto, tais resultados devem ser vistos com cautela, uma vez que não satisfazem todos os testes de robustez comumente implementados na literatura.

Esta dissertação está organizada da seguinte forma, além desta Introdução. A Seção 2 faz uma breve revisão de literatura sobre o efeito de desastres naturais sobre indicadores educacionais. A Seção 3 explica o contexto do desastre em Mariana (MG) e o relaciona com a literatura analisada. A Seção 4 descreve as bases de dados utilizadas do Censo Escolar e do Saeb. Já a Seção 5 detalha as estratégias empíricas adotadas, quais sejam: método de DD e de CS. A Seção 6 apresenta os resultados de ambas as abordagens, juntamente com testes de robustez. A Seção 7 apresenta as conclusões.

2 Revisão de Literatura

Embora a literatura acerca de desastres tecnológicos seja escassa, diversos estudos têm analisado os efeitos diretos e indiretos de desastres naturais em indicadores educacionais. Incêndios florestais na Austrália no ano de 2009, por exemplo, foram identificados como causa da piora no desempenho em aritmética e interpretação de texto do grupo considerado de alto dano do desastre (Gibbs et al., 2019). Estudantes da Universidade de L'Aquila, na Itália, passaram a ter 6,6 pontos percentuais a menos de probabilidade de se graduar após terremoto ocorrido em 2009, em comparação com estudantes de outras universidades da região, mas mais afastadas do epicentro (Di Pietro, 2018).

Os efeitos negativos dos desastres na performance acadêmica são complexos e podem variar de acordo com o estágio de desenvolvimento do aluno (Pérez-Pereira et al., 2012), condição financeira (Brewer et al., 2017) e de gênero (Di Pietro, 2018). Há estudos que não identificaram efeito algum ou que, inclusive, observaram efeitos positivos no desempenho escolar pós-desastre (Baggerly e Ferreti, 2008; Smilde-van den Doel et al., 2006). Estes autores especulam que os efeitos esperados podem ter sido mitigados por intervenções positivas no ambiente escolar logo após os desastres, em linha com outros estudos (Pane et al., 2008; Sacerdote, 2012).

A despeito das diferentes direções possíveis dos efeitos de desastres em indicadores educacionais, especialmente por conta de políticas compensatórias pós-desastre, nota-se que são três os principais canais para efeitos negativos: (i) danos à infraestrutura física, com efeitos negativos sobre taxas de matrícula e frequência escolar; (ii) choques negativos de renda, ocasionando muitas vezes subnutrição, doenças e aumento da participação de crianças na força de trabalho; e (iii) efeitos derivados de transtornos de estresse pós-traumático (TEPT).

Com relação ao primeiro, a destruição de instalações das escolas, a falta de professores e a realocação dos estudantes em outras localidades mostraram-se fatores relevantes para o desempenho escolar após os furacões Katrina e Rita nos Estados Unidos, em 2005 (Sacerdote, 2012). Por conta da destruição de casas, escolas e outras infraestruturas, terremotos que atingiram o Nepal, em 2016, tiveram efeitos negativos no número de matrículas e frequência escolar em crianças de 5 a 18 anos (Brewer et al., 2017). Fortes tempestades na Costa do Marfim no ano de 1986 reduziram em cerca de 20% o número de matrículas nas regiões afetadas (Jensen, 2000).

Sobre o segundo canal possível, há diversas evidências em crianças de ocorrências de subnutrição após desastres naturais, especialmente em países em desenvolvimento. No caso já mencionado na Costa do Marfim, em uma amostra de 1.600 famílias, as tempestades aumentaram em quase 4% a proporção de crianças desnutridas com até dez anos de idade (Jensen, 2000). Desnutrição, em especial nos períodos críticos pós-desastres, está associada com uma maior incidência de doenças, tanto em crianças como adultos (Victora et al., 2008; Almond e Currie, 2011). Por sua vez, problemas de saúde na infância afetam negativamente o desempenho escolar por meio de fatores como atraso no desenvolvimento motor, redução do Quociente de Inteligência (QI) e menos anos de estudo (Currie, 2009).

Choques negativos de renda também afetam o desempenho escolar via aumento da participação de crianças na força de trabalho (Kousky, 2016). Nas áreas afetadas pelo furacão Mitch, que devastou diversas regiões na América Central, em 1998, em especial Honduras e Nicarágua, houve alta de 58% na participação de crianças na força de trabalho. Crianças frequentando a escola e trabalhando ao mesmo tempo aumentaram de 7,5% para 15,6% devido ao desastre (Baez e Santos, 2007).

Por fim, em meio a tragédias, nota-se que muitas crianças e adolescentes ficam sujeitos a testemunhar destruição em massa, pessoas machucadas ou mortas, evacuações, além de eventuais perdas de parentes próximos ou de migrações para outras localidades. Por conta disso, o canal psicológico, via TEPT, tem se mostrado relevante para o efeito de desastres na performance acadêmica.

Conforme apontam Hoven et al. (2005), diferentes tipos de desastres estão associados com aumento de taxas de doenças psicológicas. No caso citado do terremoto em L'Aquila, a proporção de pessoas da região afetada sofrendo com síndrome metabólica seis meses depois era significativamente mais alta que a de outras regiões no centro da Itália (Di Castelnuovo et al., 2013).

Quando não são devidamente tratados, os TEPT estão diretamente ligados a casos de depressão, ansiedade, transtornos comportamentais e perda de qualidade de vida (Copeland et al., 2007). Destacam-se também efeitos negativos na atenção, memória, velocidade de processamento e habilidades cognitivas como planejamento e resolução de problemas (Barrera-Valencia et al., 2017).

Conforme será detalhado na Seção 3, o desastre no município de Mariana (MG) parece contemplar os três canais de influência negativa sobre indicadores educacionais. Desse modo, é razoável supor que os municípios, escolas e alunos afetados tenham observado piora no desempenho de tais indicadores após o ano de 2015. Por outro lado, também

foram observadas evidências de políticas compensatórias, o que pode ter mitigado ou até compensado os efeitos negativos. Dada a dificuldade em isolar tais políticas, é importante destacar que os efeitos do desastre de Mariana (MG) a serem estimados, assim como observado na literatura citada, incorporam tanto os efeitos negativos quanto os positivos pós-desastre.

Estudo recente divulgado pela FGV foi provavelmente a primeira abordagem para avaliar os efeitos causais do referido desastre (FGV, 2019). Embora diversos indicadores educacionais tenham sido utilizados, bem como diferentes variações de grupos de controle, os resultados não apontam para efeitos estatisticamente significantes na Educação. Como explicado pelos autores, a razão pode ser o nível de agregação das estimações. Foram utilizados apenas dados a nível municipal e não houve avaliação da heterogeneidade dos efeitos na dimensão geográfica. Desse modo, a análise exige maior profundidade e este é um dos objetivos desta pesquisa.

Dentre outros trabalhos existentes sobre o desastre, boa parte concentra-se nos impactos econômicos e ambientais (Simonato et al., 2017; Fernandes et al., 2015; Ramboll, 2017; ANA, 2016). Além disso, na área da Saúde, destacam-se os trabalhos de Nishijima e Rocha (2020), sobre os efeitos em casos de dengue, e o de Matsunaga (2020), sobre os efeitos em doenças mentais e violência doméstica. Ao contrário dos impactos econômicos, ambientais e em indicadores de saúde, destaque-se que impactos na Educação são mais complexos e podem demorar para aparecer.

3 Contexto

Considerado o maior desastre socioambiental da história em volume de material despejado por barragem de rejeitos de mineração, o rompimento da barragem em Mariana (MG) ocorreu na tarde do dia 5 de novembro de 2015. Denominada Fundão, a barragem controlada pela mineradora Samarco Mineração S.A. armazenava cerca de 56 milhões de metros cúbicos de rejeitos, dos quais 39,2 milhões foram liberados para fora da área da mineradora e chegaram ao Rio Gualaxo do Norte, conforme ilustra a Figura 1.⁴

Estima-se que cerca de 10 milhões de metros cúbicos de rejeitos foram depositados no trecho entre Fundão e a Usina hidrelétrica Risoleta Neves, conhecida como “UHE Candonga”, nas proximidades do município de Santa Cruz do Escalvado (MG), em uma extensão de cerca de 100 km. Outros 10 milhões ficaram represados nas imediações da usina e o restante percorreu mais de 600 km até atingir o litoral no município de Linhares (ES). Foram ao menos 40 municípios atingidos na Bacia do Rio Doce, a maioria do Estado de Minas Gerais e o restante do Espírito Santo.⁵

Note-se que a distância de um dado município para a barragem de Fundão, para a foz do Rio Doce ou para um de seus afluentes, constitui fator importante para a heterogeneidade dos efeitos do desastre. A localização das escolas dentro de cada um dos municípios também mostrou-se determinante para a exposição à lama. Houve, portanto, municípios e escolas mais afetados que outros, o que sugere a existência de efeitos de “*dose-response*”. Daí a importância de se definir uma categorização exógena que discrimine as unidades de observação (municípios, escolas e alunos) de maneira adequada.

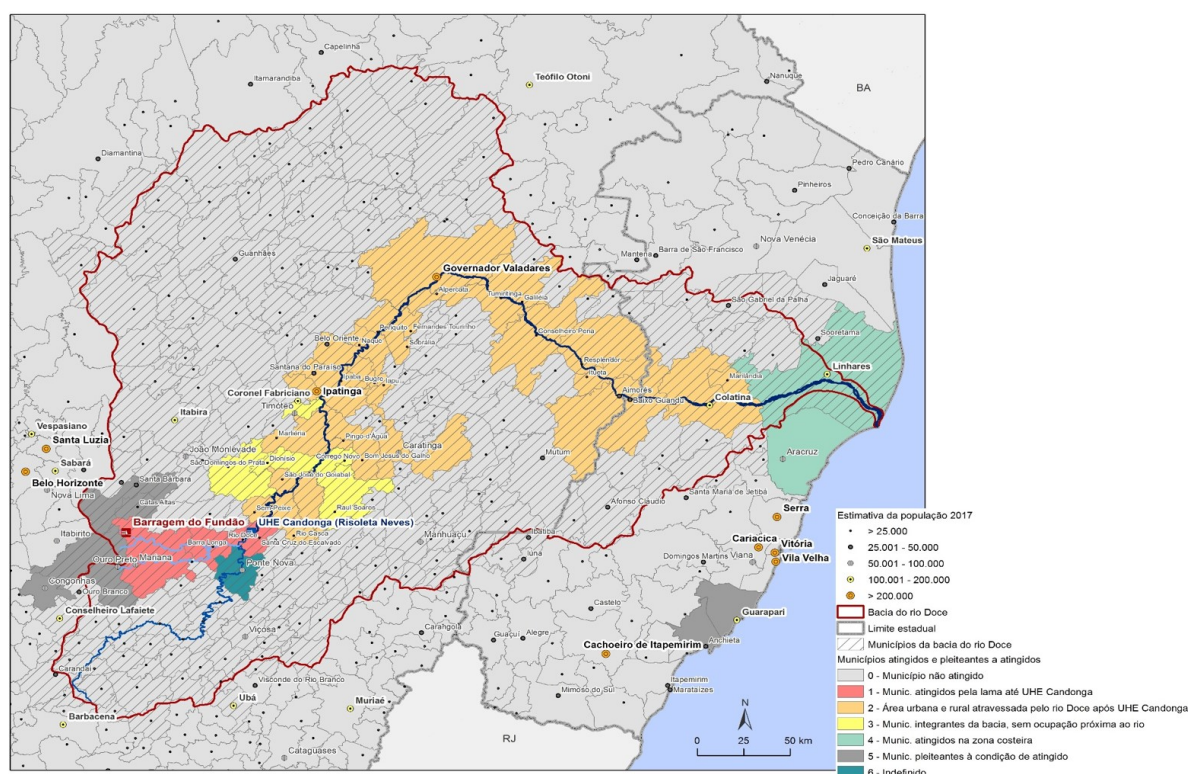
Utiliza-se a categorização a nível municipal, conforme proposta da Fundação Coordenação de Projetos, Pesquisas e Estudos Tecnológicos (COPPETEC) da Universidade Federal do Rio de Janeiro (UFRJ).⁶ Tal categorização também é utilizada por Matsunaga (2020). Os municípios afetados direta e indiretamente pelo desastre são categorizados em 5 áreas, conforme ilustrado pela Figura 2.

⁴Disponível em: <<https://www.caminhodareparacao.org/reparacao-integrada/>>. Acesso em: 02/04/2020.

⁵A Bacia do Rio Doce é uma bacia hidrográfica cuja área pertence ao Estado de Minas Gerais (86%) e Espírito Santo (14%). Disponível em: <<http://www.cbhdoce.org.br/institucional/a-bacia>>. Acesso em: 12/12/2021.

⁶Uma alternativa seria utilizar a categorização proposta pela Fundação Renova, disponível em: <<https://www.fundacaorenova.org/mapa-de-atuacao/>>. Acesso em: 20/04/2020. No entanto, trata-se de categorização possivelmente endógena, baseada nas áreas de atuação da própria entidade responsável pelo programa de reparação.

Figura 2 – Categorização COPPETEC



Fonte: COPPETEC (2016). Elaboração Keyi Ussami.

empresa representavam 1,5% do Produto Interno Bruto (PIB) do primeiro e 6,4% do PIB do último. Os impostos gerados pelas suas atividades correspondiam a 54% da receita de Mariana (MG), 35% da receita de Ouro Preto (MG) e 50% da receita de Anchieta (ES). Apesar de fora da trajetória da lama, os dois últimos sofreram enormes prejuízos com o rompimento da barragem.

O desastre em Mariana (MG) causou a morte de 14 trabalhadores da barragem e de outros cinco moradores de Bento Rodrigues, subdistrito de Mariana (MG) localizado a apenas 5 km de distância da estrutura de Fundão. No total, 225 famílias de Bento Rodrigues, 140 do subdistrito de Paracatu de Baixo e 26 do subdistrito de Gesteira foram reassentadas, boa parte ainda à espera de novas casas.⁸

Por conta dos danos à infraestrutura física como, por exemplo, a destruição de escolas, muitos alunos tiveram de ser realocados. Foi o caso de cerca de 170 alunos dos subdistritos de Bento Rodrigues e Paracatu, com idade entre 4 e 17 anos, da educação infantil e ensino fundamental, que se mudaram para a Escola Dom Luciano Pedro de Almeida, em

⁸Disponível em: <https://www.em.com.br/app/noticia/gerais/2018/11/04/interna_gerais,1002816/3-anos-apos-rompimento-da-barragem-do-fundao-o-pesadelo-continua.shtml>. Acesso em 11/09/2019.

Mariana (MG).⁹ A Figura 17, no Apêndice, mostra como ficou a Escola Municipal de Bento Rodrigues após o desastre. Apesar de terem contado com uma estrutura de apoio no início, há evidências de que os alunos novos foram vítimas de *bullying* na nova escola.¹⁰

Municípios como o de Governador Valadares (MG), com mais de 250 mil habitantes, tiveram seu fornecimento de água interrompido por mais de uma semana.¹¹ Apenas um mês depois do desastre foram encontradas 11 toneladas de peixes mortos, muitos deles em estado de extinção. Atividades como a da pesca ficaram restritas e cerca de 11 mil ribeirinhos foram atingidos, cujo sustento dependia do rio. Tais aspectos constituem parte relevante do choque negativo de renda do desastre.

Mas os danos sociais vão além das mortes, desalojamentos e restrições às atividades econômicas. A sensação de insegurança pós-rompimento afeta também pessoas de áreas adjacentes, usuários do Rio Doce, da área costeira impactada e até habitantes de regiões próximas a outras barragens no Brasil (Ibama, 2015).¹² Em 2018, pesquisa sobre saúde mental realizada com 271 pessoas atingidas mostrou que 12% sofriam de TEPT (PRISMMA, 2018). Sinais de depressão foram relatados por cerca de 30% das vítimas, ou 39% dos jovens. Casos de ansiedade, insônia, problemas cardiovasculares e pensamentos suicidas também foram observados.

Cerca de 16 dias depois do desastre a lama atingiu o Oceano Atlântico e praias no município de Linhares (ES) tiveram de ser interditadas. Ecossistemas marinhos também ficaram comprometidos. Análises da Estação Biologia Marinha Augusto Ruschi, no distrito de Santa Cruz, no município de Aracruz (ES), apontam que os rejeitos só começarão a ser eliminados do mar em 100 anos, no mínimo.¹³

Ao longo do trecho atingido foram constatados ainda: destruição de áreas de preservação permanente e vegetação nativa, morte de animais de produção e impacto na produção rural, prejuízos ao turismo local, dificuldades de geração de energia pelas usinas hidrelétricas afetadas e contaminação da água, além da sensação de perigo e desamparo da população em diversos níveis (Ibama, 2015).

⁹Disponível em: <<https://www.samarco.com/noticia/comunicado-35/>>. Acesso em: 11/09/2019.

¹⁰Disponível em: <<https://brasil.estadao.com.br/noticias/geral,desabrigados-pela-lama-enfrentam-preconceito-e-desconfianca-em-mariana,10000085277>>. Acesso em 11/12/2019.

¹¹Disponível em: <<https://noticias.uol.com.br/ultimas-noticias/bbc/2020/02/24/lama-que-brilha-e-reacende-traumas-do-desastre-de-mariana-para-moradores-da-bacia-do-rio-doce.htm>>. Acesso em: 17/04/2020.

¹²Para fins de simplificação dos modelos estimados, no entanto, municípios em que estão localizadas outras barragens no Brasil não foram considerados nos grupos de tratamento.

¹³Disponível em: <<https://www.otempo.com.br/cmlink/hotsites/mar-de-lama/contamina%C3%A7%C3%A3o-do-rio-doce-amea%C3%A7a-vida-marinha-no-esp%C3%ADrito-santo-1.1161772>>. Acesso em: 11/09/2019.

Note-se, portanto, que o desastre em Mariana (MG) contempla os três canais de influência negativa de desastres naturais sobre indicadores educacionais. De fato, conforme relatório da Ramboll (2017), destinado ao Ministério Público Federal (MPF), o desastre provocou danos que produziram o impedimento, a desarticulação ou o aumento na demanda por processos educativos, como:

- i. Perdas de infraestrutura de educação: escolas total ou parcialmente destruídas;
- ii. Desagregação das teias sociais existentes nas comunidades escolares com prejuízos aos respectivos processos educativos;
- iii. Interrupção do funcionamento de escolas gerando danos psicopedagógicos às comunidades escolares.

Observou-se aumento na demanda por serviços de educação na cidade de Mariana (MG) devido à necessidade de adaptação das escolas receptoras, especialmente no acolhimento aos estudantes e suas famílias por parte das novas comunidades escolares. Além disso, houve interrupção temporária de atividades escolares por conta tanto da interrupção das vias de acesso quanto pela falta de água em muitas escolas dos diversos municípios da Bacia do Rio Doce.

Em março de 2016, foi assinado o Termo de Transação e de Ajustamento de Conduta (TTAC) entre as empresas envolvidas Samarco Mineração S.A, Vale S.A., *BHP Billiton* Brasil Ltda. e os Governos Federal, de Minas Gerais e do Espírito Santo.¹⁴O objetivo foi acelerar a implementação das medidas de recuperação social, ambiental e econômica das regiões atingidas, com a criação da Fundação Renova, que hoje conta com cerca de 7 mil funcionários e ficou responsável pela implantação de 42 programas de reparação na Bacia do Rio Doce.

Os programas da Renova foram divididos em três eixos temáticos: “Pessoas e Comunidades”, “Terra e Água”, e “Reconstrução e Infraestrutura”. O primeiro eixo contempla o programa de “Recuperação das escolas e reintegração da comunidade escolar”, com o objetivo de reconstruir as escolas impactadas de Fundão até a UHE Candonga, além de providenciar todos os meios para reintegração dos alunos e profissionais nas rotinas escolares. Já foram construídas escolas temporárias para Bento Rodrigues, Paracatu de Baixo e Gesteira e um total de 58 escolas municipais de Mariana e Barra Longa, na área 1 proposta pela

¹⁴Disponível em: <<https://www.fundacaorenova.org/wp-content/uploads/2016/07/ttac-final-assinado-para-encaminhamento-e-uso-geral.pdf>>. Acesso em: 28/03/2021.

COPPETEC, seriam envolvidas em atividades de apoio pedagógico e psicopedagógico. Tal programa tem o potencial de mitigar os efeitos negativos do desastre no desempenho acadêmico e na experiência escolar dos alunos afetados. Dessa forma, a relação causal a ser estimada neste trabalho contempla tais potenciais efeitos positivos da Renova.

O orçamento inicialmente previsto para as ações de reparação e compensação da Renova foi de R\$ 11,6 bilhões. Estima-se que até o momento tenham sido desembolsados pela empresa mais de R\$ 8 bilhões para ações de reparação.¹⁵ Foram distribuídos ao menos R\$ 1 bilhão de indenizações para “Dano Água e Dano Geral” e R\$ 1,2 bilhão de “Auxílio Financeiro Emergencial”. Apesar dos avanços, há evidências de atrasos nas obras, não pagamento de indenizações a muitas famílias e críticas às ações da fundação.¹⁶

Em junho de 2018 foi assinado um Termo de Ajustamento de Conduta (TAC) entre Ministério Público de Minas Gerais (MPMG), Ministério Público Federal (MPF), Ministério Público do Espírito Santo (MPES), as Defensorias Públicas dos dois estados e mais nove órgãos públicos, além das empresas Samarco Mineração S.A., Vale S.A. e *BHP Billiton* Brasil Ltda., para alterar a governança da Fundação Renova.¹⁷ O acordo modificou o TTAC, de 2016, com o objetivo de garantir a efetiva participação dos atingidos nas decisões referentes à reparação dos danos causados pelo desastre. Determinou-se o estabelecimento de um processo único de repactuação dos referidos programas, de forma a aprimorar os programas futuros, mas sem prejudicar a manutenção, a continuidade e a eficácia das ações, medidas e projetos já em curso.

Além da dificuldade de monitorar a efetiva implementação dos programas educacionais nos municípios e escolas atingidas, há outro desafio para a estratégia empírica a ser adotada neste trabalho: o descontentamento da população com as empresas diretamente envolvidas pode intensificar o estresse pós-trauma, terceiro canal de influência negativa de desastres naturais em indicadores educacionais. De fato, trata-se de fator relevante na diferenciação entre desastres naturais e tecnológicos (não naturais) e um elemento extra - para além dos potenciais efeitos positivos dos programas da Renova - a ser considerado na interpretação dos efeitos causais do desastre em Mariana (MG).

¹⁵Disponível em: <<https://www.fundacaorenova.org/a-fundacao/>>. Acesso em: 30/04/2020.

¹⁶Disponível em: <<http://agenciabrasil.ebc.com.br/geral/noticia/2019-01/criada-para-reparar-danos-em-mariana-renova-nao-atuara-em-brumadinho>>. Acesso em: 12/09/2019

¹⁷Disponível em: <<https://www.mpmg.mp.br/comunicacao/noticias/acordo-altera-governanca-da-fundacao-renova-para-garantir-participacao-dos-atingidos-pelo-rompimento-da-barragem-de-fundao.htm>>. Acesso em: 28/03/2021.

4 Dados

As duas bases de dados utilizadas neste trabalho são as do Censo Escolar e do Saeb. O Censo Escolar constitui pesquisa estatística coordenada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) que é realizada anualmente com todas as escolas públicas e privadas do Brasil.¹⁸ Desta base, a nível de escola e município, a variável de interesse é a taxa de reprovação no ensino fundamental.¹⁹ A nível de aluno avalia-se o efeito do desastre sobre a probabilidade de reprovação dos alunos que estavam matriculados do 5º ao 9º ano do ensino fundamental em 2015. O período avaliado é de 2010 até 2018 para as análises no nível escola e municípios e de 2010 até 2017 para as análises no nível aluno, dada a disponibilidade de dados.

A coleta de dados do Censo divide-se em duas etapas. A primeira delas é a de “Matrícula Inicial”, com dados de matrículas a nível individual. A data de referência é a última quarta-feira do mês de maio do ano de referência. A segunda etapa refere-se à “Situação do Aluno” e traz informações sobre o rendimento dos alunos ao final do ano letivo. O período de coleta desta última ocorre entre março e maio do ano seguinte ao ano de referência.

Os dados do Censo são disponibilizados de diferentes formas: Sistema Educacenso, Consulta Matrícula, Sinopses Estatísticas, Microdados, Inepdata e Indicadores Educacionais. É desta última base que são coletadas as variáveis de interesse a nível de escola e município. A Figura 18, no Apêndice, apresenta a evolução das taxas de reprovação desde 2010, para tratados e controles, a nível de escola. A princípio, apenas a área 1 parece não apresentar trajetória similar ao grupo de controle.

Dos Microdados são coletadas diferentes variáveis de controle. Há quatro bases de microdados disponíveis: Escolas, Matrículas, Turmas e Docentes.²⁰ Tais bases contemplam quase todo o acervo de dados primários coletados pelo Inep e permitem a reprodução de todos os indicadores calculados pelo órgão.

A base de Escolas reúne informações como localização, dependência administrativa e condições de infraestrutura das escolas. A fim de avaliar o canal de influência negativa na Educação via “Danos à infraestrutura física”, conforme mencionado na Seção 2, são criadas duas variáveis de controle. Uma primeira variável denominada “Infraestrutura” representa uma média de oito *dummies* para existência ou não na escola de: coleta de esgoto pela rede pública; fornecimento de energia elétrica pela rede pública; auditório; área

¹⁸Disponível em: <<http://portal.inep.gov.br/web/guest/censo-escolar>>. Acesso em: 30/04/2020.

¹⁹Embora também tenham sido realizados exercícios com as taxas de aprovação e abandono.

²⁰Disponível em: <<http://portal.inep.gov.br/web/guest/microdados#>>. Acesso em: 25/04/2020.

verde; biblioteca; pátio descoberto; quadra de esportes e refeitório. A segunda variável “Tecnologia” representa uma média de sete *dummies* para a existência ou não na escola de: laboratório de informática; computador para uso dos alunos; impressora; televisão; equipamento multimídia; internet e banda larga.

A base de Matrículas traz informações dos alunos matriculados em todas as escolas do Brasil. Variáveis como idade, sexo, raça e etapa de ensino serão utilizadas ou como controle nas estimações a nível de aluno ou nas análises de heterogeneidade. Outra covariada utilizada é o tamanho da turma, coletada da base de Turmas.

A Tabela 1 apresenta as estatísticas descritivas, para tratados e controles para o período pré-desastre (período de 2010 a 2015). São reportados dados a nível de escola e município e os testes de média. Nota-se diferenças estatisticamente significantes na maioria das variáveis. O grupo de tratamento possuía, antes do desastre, em média, maior rendimento e maior arrecadação de Imposto sobre Circulação de Mercadorias e prestação de Serviços (ICMS) e Compensação Financeira pela Exploração de Recursos Minerais (CFEM) per capita. As condições de infraestrutura também eram melhores, o que está correlacionado com a maior proporção de escolas privadas. Por outro lado, em média, as escolas tratadas reprovavam mais alunos do ensino fundamental e também observavam menor taxa de abandono.²¹

A partir dos dados do Censo foi possível identificar se um aluno matriculado em uma escola afetada pelo desastre, em novembro de 2015, transferiu-se para uma escola não afetada nos anos seguintes. Isso permite verificar uma eventual contaminação do grupo de controle e excluir da análise as escolas não afetadas pelo desastre que, após 2015, receberam alunos considerados afetados. As alternativas de grupo de controle para o método de DD são apresentadas na Seção 5.²²

Para as estimações a nível municipal e de escola, também foram avaliados os desempenhos médios em Português e Matemática dos alunos do 5º ano do ensino fundamental. Tais dados foram obtidos da base do Saeb, conjunto de avaliações externas realizadas a cada dois anos pelo Inep. Participam das provas do Saeb todas as escolas públicas do Brasil que possuem dez ou mais estudantes matriculados no 2º, 5º e 9º ano do ensino fundamental e na 3ª e 4ª série do ensino médio. Devido à falta de dados para escolas privadas, as médias agregadas a nível municipal consideram apenas as escolas públicas.

²¹Conforme será abordado na Seção 5, o método de Diferenças em Diferenças (DD) permite que haja diferenças em nível da média entre os grupos de tratamento e controle do experimento natural.

²²Como será explicado mais adiante, o critério de exclusão do controle por contaminação foi o de a escola receber uma quantidade de alunos tratados maior ou igual a 5 ou 10% do total de seus alunos; e de 2,5 e 5% para o caso dos municípios.

Desde 1990, quando foi criado, o Saeb passou por algumas reestruturações metodológicas.²³ Dentre elas, vale destacar, na sua composição, a incorporação da Avaliação Nacional da Educação Básica (Aneb) e da Avaliação Nacional do Rendimento Escolar (Anresc), ou Prova Brasil, em 2005, e a Avaliação Nacional da Alfabetização (ANA), em 2013. Em 2017, não só as escolas públicas do ensino fundamental, mas também as de ensino médio, públicas e privadas, passaram a ter resultados no Saeb. Em 2019, as siglas ANA, Aneb e Anresc (Prova Brasil) deixaram de existir e todas as avaliações foram unificadas e identificadas pelo nome Saeb. Ressalte-se que, mesmo com as alterações, o sistema não perde em termos de comparabilidade entre as edições. De todo modo, analisam-se os dados desde a prova de 2011 para evitar vieses por conta de tais mudanças.

Tabela 1 – Teste de Médias: Tratados vs. Controles - Escolas e Municípios - Pré-desastre

	Controles		Tratados		p-valor
	Média	SD	Média	SD	
Aprovação nos Anos Finais do Ensino Fundamental (EF)	87.06	9.58	86.89	10.57	0.401
Reprovação nos Anos Finais do EF	9.75	8.20	10.78	9.19	0.000
Abandono nos Anos Finais do EF	3.19	3.80	2.33	4.03	0.000
Matemática EF 9º Ano	259.33	17.80	258.96	18.09	0.554
Português EF 9º Ano	252.53	16.45	253.99	17.78	0.015
Proporção de Escolas Privadas	0.05	0.22	0.14	0.35	0.000
Proporção de Escolas Urbanas	0.43	0.50	0.69	0.46	0.000
Nível Socioeconomico	3.42	0.81	3.84	0.86	0.000
Nível de Complexidade de Gestão das Escolas	2.51	1.59	2.79	1.51	0.000
Média diária de horas de aula nos Anos Finais do EF	4.56	0.68	4.68	0.84	0.000
Adequação do Corpo Docente nos Anos Finais do EF	29.44	30.32	43.70	29.61	0.000
Esforço empreendido pelo docente nos Anos Finais do EF	36.36	42.84	17.16	32.89	0.000
Docentes com curso superior nos Anos Finais do EF	81.77	29.99	88.66	21.09	0.000
Proporção de Escolas em MG	0.84	0.36	0.64	0.48	0.000
Área dos Municípios na Bacia do Rio Doce	96.17	12.72	74.98	39.61	0.000
População Branca em 2010	0.37	0.13	0.37	0.07	0.001
Taxa de Alfabetização em 2010	86.37	4.58	91.71	3.45	0.000
Rendimento Médio em 2010	695.96	219.55	1044.29	228.58	0.000
ICMS per capita - Tendência linear	410.90	790.56	836.93	1385.13	0.000
CFEM per capita - Tendência linear	85.24	683.47	186.53	533.76	0.000
Condições de infraestrutura física	0.42	0.23	0.51	0.23	0.000
Condições de infraestrutura tecnológica	0.55	0.38	0.72	0.32	0.000
PIB Capita	13113.68	18720.33	30841.28	31969.87	0.000

Fonte: Censo Escolar. Elaboração própria

Os resultados do Saeb são apresentados em uma escala de proficiência que varia de 0 a 500 e que é capaz de descrever, em cada nível, as competências e as habilidades que os estudantes demonstram ter desenvolvido. Dentro de cada uma das disciplinas a escala é única e acumulativa para todas as séries avaliadas. Caso a curva de aprendizado dos alunos fosse uma variável relevante, lidar com valores absolutos dos desempenhos médios das

²³Disponível em: <<http://portal.inep.gov.br/web/guest/educacao-basica/saeb/historico>> Acesso em: 30/06/2020.

disciplinas poderia ser problemático. Municípios ou escolas com médias elevadas poderiam apresentar menor variação de um ano para outro relativamente aos demais simplesmente porque seus alunos já estão em uma etapa mais avançada do aprendizado, em que os ganhos marginais são naturalmente menores. Este não parece ser um problema para os modelos a serem apresentados, pois no 5º ano os alunos dos municípios sob análise ainda estão bem distantes da escala máxima de proficiência.

Embora haja informações disponíveis a nível individual, a metodologia do Saeb não permite o cálculo de resultados por aluno, uma vez que cada estudante responde a itens diferentes na prova. Não há, por exemplo, um boletim individual de desempenho no Saeb.²⁴ Daí o porquê dos níveis de agregação escolhidos. Para avaliar a heterogeneidade dos efeitos sobre os indicadores do Saeb também será feita a discriminação dos tratados com base na categorização COPPETEC.

As aplicações das provas são feitas em anos ímpares, geralmente entre outubro e novembro, e os resultados divulgados nos anos pares. Em 2015, a prova do Saeb foi aplicada no período entre 3 e 11 de novembro, enquanto o rompimento da barragem de Fundão ocorreu no dia 5. Como não foi possível identificar quais escolas tiveram suas provas e aulas suspensas, o fato de o desastre ter ocorrido durante a aplicação das provas do Saeb pode comprometer a consistência dos estimadores dos efeitos do desastre por conta de atrito na amostra. Dessa forma, os resultados das estimações que se baseiam nas variáveis de interesse do Saeb devem ser analisados com cautela. De todo modo, para fins de estimação, dado que o conhecimento exigido pela avaliação foi adquirido ao longo de todo o ano, supõe-se neste trabalho que o ano de 2015 é um ano pré-tratamento. O mesmo vale para as estimações com dados do Censo Escolar.

Variáveis socioeconômicas dos municípios foram utilizadas como controle em todas as especificações. O PIB per capita dos municípios, por exemplo, foi obtido do Sistema de Recuperação Automática (SIDRA) do Instituto Brasileiro de Geografia e Estatística (IBGE). Tal variável é utilizada em algumas especificações a fim de captar o canal de influência negativa do desastre na Educação via “Choques negativos de renda”. Os repasses municipais do ICMS foram obtidos junto ao Portal de Transparência de Minas Gerais e do Espírito Santo. Os dados das CFEM foram obtidos da atual Agência Nacional de Mineração (ANM).

Note-se que, assim como as variáveis Infraestrutura e Tecnologia, o PIB per capita também foi afetado pelo rompimento da barragem de Fundão. Isto é, a inclusão destas variáveis

²⁴Disponível em: <<http://portal.inep.gov.br/educacao-basica/saeb/perguntas-frequentes>> Acesso em: 09/06/2020.

nas estimações tendem a absorver parte do efeito do tratamento. Assim, ao controlar pelos canais de danos à infraestrutura e choques negativos de renda, temos que o efeito restante do desastre passa a representar, ao menos em parte, o canal psicológico dos TEPT. Embora seja útil isolar o efeito dos diferentes canais de influência sobre a Educação, o foco deste trabalho é avaliar o efeito global do desastre.²⁵ Dessa forma, nas especificações principais são adicionadas apenas covariadas não afetadas pelo tratamento.

A fim de controlar por variáveis socioeconômicas, mas sem que estas absorvam parte do efeito do tratamento, são criadas tendências lineares para PIB per capita, ICMS per capita e CFEM per capita. As tendências são criadas com base nos dados de 2010 até 2014 e, então, são extrapoladas para o período de 2015 em diante. Tal procedimento também é adotado por Matsunaga (2020).

Vale ressaltar que os dados de rendimento escolar no nível aluno não são disponibilizados para consulta nos Microdados do Censo, mas foram obtidos via Lei de Acesso à Informação (LAI). Com base nestes dados, é possível acompanhar a trajetória escolar de cada aluno, tanto do grupo de controle quanto de tratamento, entre 2010 e 2017.²⁶

Com relação à análise no nível individual, o foco deste trabalho é analisar o efeito do desastre na probabilidade de reprovação dos alunos do 5º ao 9º ano do ensino fundamental matriculados no ensino regular em 2015.²⁷ Para tanto, identificou-se quais alunos pertenciam ao grupo de tratamento e quais pertenciam ao grupo de controle, com base nas categorias COPPETEC, de acordo com o município onde cada aluno estudava no ano de 2015. Feito isso, estes alunos são rastreados desde 2010 até 2017. Isto é, são coletadas informações da escola em que estudaram, etapa de ensino, número de alunos na turma e situação da matrícula (aprovado, reprovado, abandono, transferência ou falecimento) ao final de cada ano. Neste processo, 90.029 alunos do grupo de tratamento e 110.170 alunos do grupo de controle são rastreados neste período totalizando pouco mais de 1,6 milhão de observações.

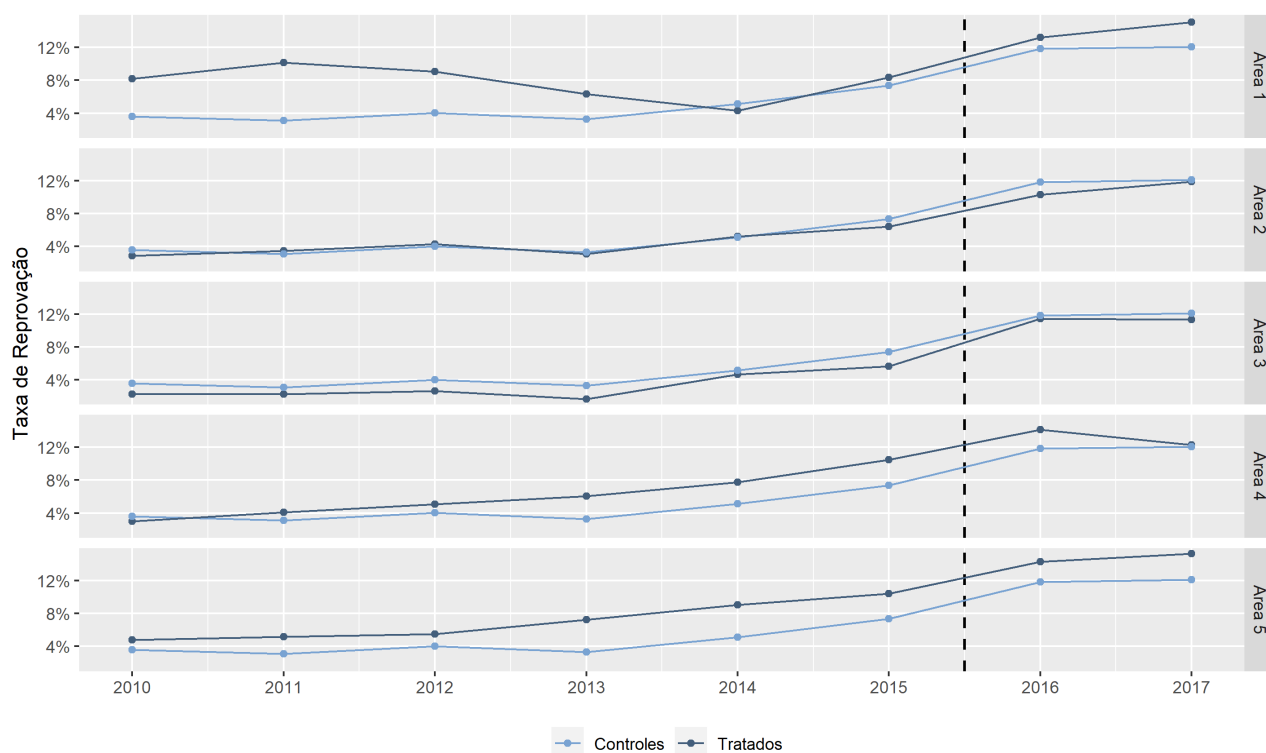
A Figura 3 apresenta a evolução das taxas de reprovação, por região, dos alunos rastreados ao longo do tempo. As curvas de cada área COPPETEC são comparadas à curva de um mesmo grupo de controle, que contém todos os municípios considerados não tratados e que possuem sede na Bacia do Rio Doce. Nota-se um aumento da taxa para a área 1 após o desastre em comparação com as demais áreas.

²⁵Que, vale lembrar, contempla tanto os efeitos dos canais negativos de influência na Educação quanto o canal positivo de políticas compensatórias implementadas pela Fundação Renova.

²⁶Há mudança no código do aluno entre 2017 e 2018. Com essa mudança não é possível rastrear os alunos na base de dados de 2018. Os dados relativos a este último ano foram excluídos da análise.

²⁷Exclui-se desta análise matrículas em cursos técnicos e EJA (Educação de Jovens e Adultos).

Figura 3 – Taxas de Reprovação - Alunos - Categorização COPPETEC



Fonte: Censo Escolar. Elaboração própria.

Já a Tabela 9, no Apêndice, apresenta a comparação de médias entre tratados e controles para algumas variáveis de interesse no período pré-desastre.²⁸ Em 2015, estavam matriculados 310.925 alunos no grupo de tratamento, sendo 12.521 na área 1, a mais afetada pelo desastre. Destes, 84% estudavam em escolas do município de Mariana (MG).

²⁸ Vale mencionar que a variável de número de alunos por turma apresenta alguns valores extremamente elevados para determinadas turmas tanto do grupo de tratamento quanto do grupo de controle. Uma análise dos *outliers* sugere que algumas escolas podem lançar seus alunos em apenas uma ou poucas turmas, não refletindo de fato a turma que cada aluno estuda.

5 Estratégia Empírica

São utilizadas duas abordagens neste trabalho. A primeira delas é o modelo de Diferenças-em-Diferenças (DD), detalhado na Subseção 5.1. A segunda abordagem é o modelo de Controle Sintético (CS), que é apresentado na Subseção 5.2.

5.1 Diferenças em Diferenças

Assim como foi feito em FGV (2019), a primeira abordagem utilizada é o modelo de DD. O rompimento da barragem de Fundão representa um tratamento exógeno e os grupos de tratamento e controle são divididos a nível municipal, conforme categorização proposta pela COPPETEC. As estimações são realizadas a nível de município, escola e aluno.

O grupo de tratamento contempla 46 municípios afetados direta e indiretamente pelo desastre, subdivididos em 5 áreas. O grupo de controle tomado como base inclui todos os demais municípios não afetados e que possuem sede na Bacia do Rio Doce. Tal grupo contempla tanto municípios de Minas Gerais como Espírito Santo.

Conforme mencionado na Seção 4, a partir dos dados do Censo foi possível identificar se um aluno matriculado em uma escola localizada em um município afetado pelo desastre, em 2015, transferiu-se para uma escola localizada em um município não afetado, nos anos seguintes. Embora pequeno, é natural que haja um fluxo de alunos entre escolas e municípios do tratado e controle, todos os anos, conforme ilustram as Figuras 19 e 20, no Apêndice.²⁹

Ocorre que se tal movimentação foi atípica pós-desastre, trata-se de uma primeira evidência de que possivelmente alunos tratados transferiram-se para outras escolas e municípios por conta dos efeitos do desastre. Tal movimentação adicional poderia representar efeitos de “*spill-over*” entre grupo de tratamento e controle, contaminando este último e viesando as estimativas do método de DD.

²⁹Nota-se que mais de 80% dos alunos permaneceram na mesma área após o desastre. A categoria “Não Encontrados” leva em conta também os alunos do último ano do ensino médio que se formaram em 2015 e saíram da amostra.

Assim, como teste de robustez, serão adotadas outras variações possíveis para o grupo de controle, quais sejam:

- i. Municípios não afetados, com sede na Bacia do Rio Doce, exceto aqueles que receberam uma quantidade de alunos tratados maior do que 5% do total de seus alunos em um dado ano pós-tratamento;³⁰
- ii. Municípios não afetados, com sede na Bacia do Rio Doce, exceto aqueles que receberam uma quantidade de alunos tratados maior do que 2,5% do total de seus alunos em um dado ano pós-tratamento;³¹
- iii. Municípios não afetados, com sede na Bacia do Rio Doce e pareados via *propensity score*, que supõe um tratamento não aleatório.³²

As duas primeiras variações do grupo de controle, testadas para as estimações a nível de escola e município, visam a corrigir tais eventuais efeitos de *spillover* decorrentes do desastre. Já a terceira se aproveita do mecanismo de pareamento do *propensity score* para garantir que as unidades do grupo de controle e tratamento tenham a mesma probabilidade de serem tratadas, antes do desastre.

As Figuras 4, 5 e 6 ilustram quais os municípios considerados como grupo de controle de acordo com a especificação base, a com *threshold* de 5% na movimentação de alunos inter-municípios³³ e a com *threshold* de 2,5%³⁴, respectivamente. As estimações do DD a nível de escola e aluno seguem classificação similar.

³⁰Maior que 10% na estimaco a nvel de escola.

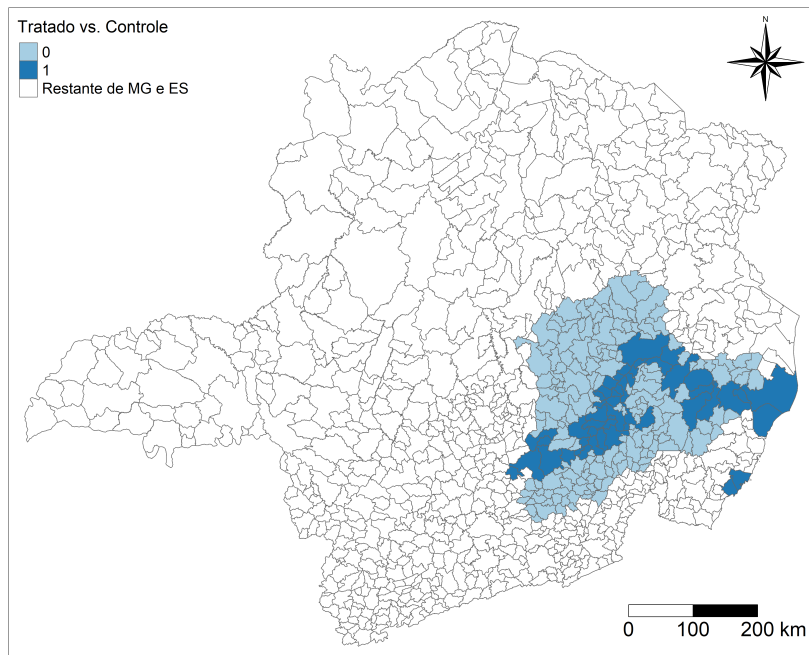
³¹Maior que 5% na estimaco a nvel de escola.

³²Esta foi aplicada apenas para o caso das estimaces a nvel de aluno.

³³Os municpios considerados contaminados ps-tratamento sob este critrio so, de MG: Entre Folhas, Jaguarau, Piedade de Caratinga, Santa Rita do Itueto, So Geraldo da Piedade, So Joo do Oriente; e do ES: Pancas.

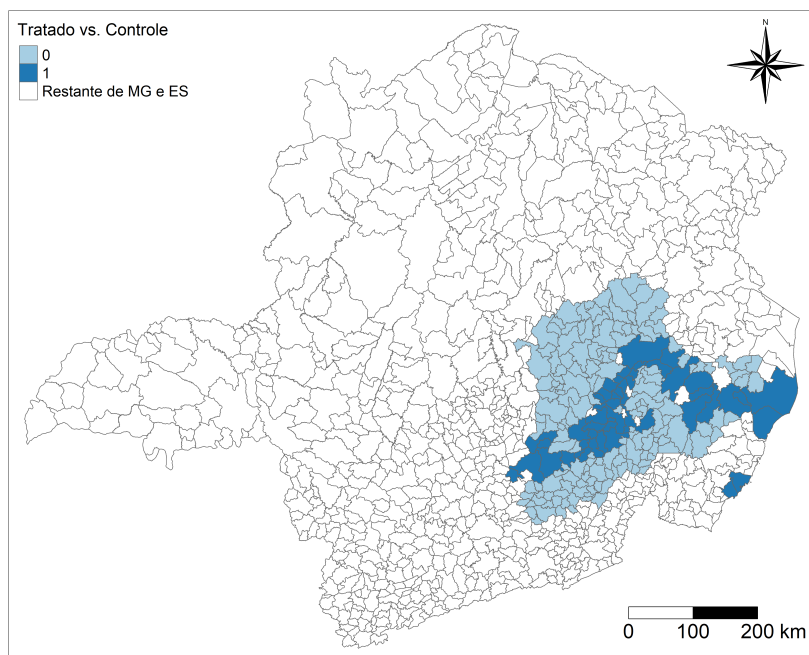
³⁴Alm dos municpios considerados contaminados sob o *threshold* de 5%, so considerados contaminados ps-tratamento sob este critrio, de MG: Acaiaca, Aucena, Alvarenga, Antnio Dias, Capito Andrade, Coroaci, Coronel Fabriciano, Cuparaque, Diogo de Vasconcelos, Divino das Laranjeiras, Dom Cavati, Engenheiro Caldas, Frei Inocncio, Goiabeira, Imb de Minas, Itanhomi, Jampruca, Joansia, Marilac, Mesquita, Santa Brbara do Leste, Santa Rita de Minas, So Geraldo do Baixo, So Jos da Safira, Sardo, Ubaporanga, Vargem Alegre, Vermelho Novo, Mathis Lobato; e do ES: Governador Lindenberg, Laranja da Terra, Rio Bananal e So Roque do Cana.

Figura 4 – Grupo de Controle Base - Nível Municipal



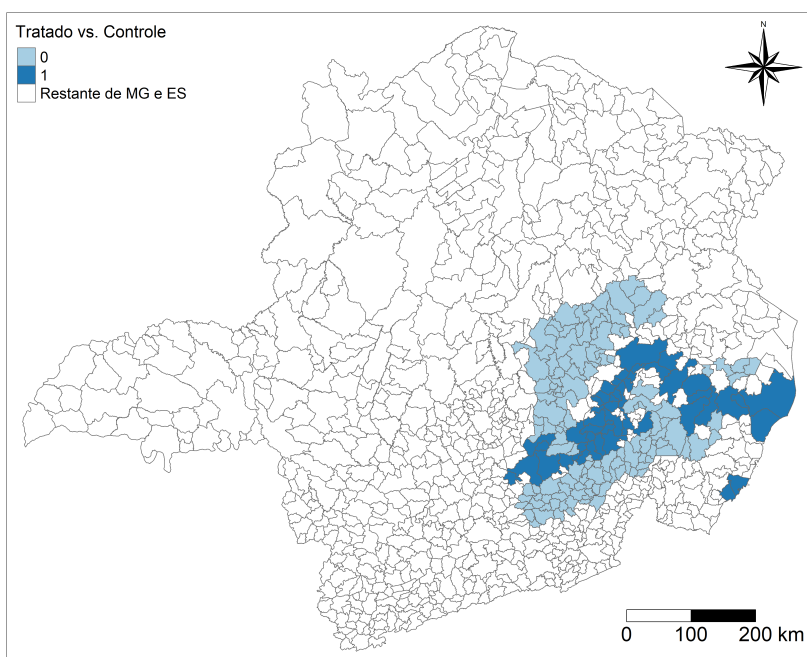
Fonte: COPPETEC (2016). Elaboração própria.

Figura 5 – Grupo de Controle *threshold* 5% - Nível Municipal



Fonte: COPPETEC (2016). Elaboração própria.

Figura 6 – Grupo de Controle *threshold* 2,5% - Nível Municipal



Fonte: COPPETEC (2016). Elaboração própria.

De acordo com Khandker (2010), a hipótese fundamental de identificação do método DD é a de que, na ausência do tratamento, a diferença dos valores médios das taxas de rendimento, desempenhos ou status de aprovação entre os grupos de tratamento e controle deve ser a mesma ao longo dos anos. Tal hipótese é conhecida como a de “tendências paralelas” e será testada com base na estimação da seguinte equação:

$$y_{it} = \sum_{\tau=-5, \tau \neq 0}^3 D_{i\tau} \beta_{\tau} + \mu_i + \gamma_t + X_{it} \delta + \epsilon_{it} \quad (1)$$

em que y_{it} representa algum dos indicadores educacionais mencionados para o aluno, escola ou município i no período t ; μ_i representa efeitos fixos no tempo específicos a alunos, municípios ou escolas i ; γ_t representa *dummies* de ano; X_{it} contempla variáveis explicativas observáveis não afetadas pelo tratamento, que variam no tempo e que serão detalhadas mais adiante; ϵ_{it} representa o termo de erro aleatório e não observável; e D_{it} é uma *dummy* = 1 para as observações tratadas i em um dado ano que está τ anos distante do ano do tratamento.

Ou seja, tem-se um coeficiente β_{τ} para cada ano antes e depois do desastre, relativos ao ano de 2015, tomado como ano de referência. A justificativa para 2015 ser considerado um ano pré-tratamento é de que os alunos passam o ano todo estudando, de modo que o

fato de o último mês letivo ter sido comprometido pelo desastre deveria afetar pouco suas notas anuais ou a decisão do professor a respeito da reprovação ou não dos alunos.³⁵

Embora o caso de Mariana (MG), ao contrário de outros experimentos naturais - como de estimações de efeitos da implementação de legislações específicas, por exemplo - esteja menos sujeito ao fenômeno de “*Ashenfelter’s dip*”, a literatura sugere que a hipótese de tendências paralelas não deve ser tomada como certa.³⁶ Tal fato reforça a importância do teste de tendências paralelas proposto.

Apesar de não ser possível garantir a tendência paralela entre as trajetórias dos grupos de tratado e controle, o teste acima nos dá evidências importantes. Controlando por demais covariadas que variam no tempo e efeitos fixos no tempo, desejamos que não haja diferenças entre os dois grupos, no período pré-tratamento, devido unicamente ao tratamento. Isto é, idealmente, o coeficiente β_τ é não significativo para $\tau \in \{-1, -2, -3, -4, -5\}$. A variação dos grupos de controle tem como objetivo encontrar o contrafactual adequado ao grupo de tratamento.

Por fim, vale destacar que, por incluir demais variáveis de controle, a equação proposta para a estimação do teste de tendências paralelas também pressupõe, em linha com Sant’Anna and Zhao (2020): (i) efeitos homogêneos do tratamento em X_{it} . Isto é, que o rompimento da barragem afeta variáveis incluídas em X_{it} de maneira idêntica, no sentido de que tais variáveis não passaram a afetar o desempenho acadêmico pós-2015 de uma forma distinta entre tratados e controles daquela pré-2015; e (ii) ausência de tendências específicas nas variáveis de controle incluídas em X_{it} entre tratados e controles.³⁷

Como há diferenças no método de estimação e na interpretação dos resultados entre as estimações a nível municipal, escola e de aluno, a estratégia empírica a ser adotada em cada caso é apresentada separadamente. A Subseção 5.1.1 trata da estimação a nível de escolas e municípios. A Subseção 5.1.2 trata do nível de alunos.

³⁵Tal racional é menos direto para o caso das notas de Português e Matemática no Saeb, dada a coincidência nas datas da aplicação das provas e do desastre.

³⁶Ver, por exemplo, trabalhos e notas de aula de Clément de Chaisamartin, disponíveis em: <<https://sites.google.com/site/clementdechaisemartin/>>. Acesso em: 12/12/2021.

³⁷Uma descrição didática da necessidade das duas hipóteses está disponível em: <<https://causalinf.substack.com/p/a-tale-of-time-varying-covariates>>. Acesso em: 12/12/2021.

5.1.1 Escolas e Municípios

A nível de escolas e municípios, o método de DD será estimado via modelo linear em painel com efeitos fixos, a fim de controlar características não observáveis e invariantes no tempo. As variáveis dependentes serão: (i) taxa média de reprovação nos anos finais do ensino fundamental; (ii) desempenho médio do 5º ano do ensino fundamental em Matemática; e (iii) desempenho médio do 5º ano do ensino fundamental em Português.

A princípio, considera-se duas especificações, com erros padrão agrupados (“*clustered*”) por município. A especificação geral, com todos os tratados, é apresentada a seguir:

$$y_{it} = W_{it}\beta + \mu_i + \gamma_t + X_{it}\delta + \epsilon_{it} \quad (2)$$

em que y_{it} representa uma das variáveis dependentes mencionadas no município ou escola i no período t ; W_{it} representa uma variável *dummy* igual a 1 para as observações no grupo de tratamento no período pós-desastre. Isto é, $W_{it} = d_t \times T_i$, com d_t sendo uma variável *dummy* igual a 1 para os períodos após o desastre (2016 em diante) e T_i uma variável *dummy* igual a 1 para os municípios ou escolas i que foram afetados de alguma forma pelo desastre; μ_i representa efeitos fixos no tempo específicos aos municípios e/ou escolas i e γ_t representa *dummies* de ano, que permitem controlar efeitos fixos específicos de cada período; X_{it} contempla variáveis explicativas observáveis que variam no tempo e que não foram afetadas pelo tratamento: tendência linear estimada para PIB per capita, ICMS per capita, CFEM per capita; e por fim, ϵ_{it} representa o termo de erro aleatório e não observável.

O efeito do desastre nas taxas de rendimento ou no desempenho médio das disciplinas de Matemática e Português de um dado município ou escola será medido pelo coeficiente da interação entre o tratamento e o período posterior ao rompimento da barragem: β . Implicitamente, assume-se que o efeito é o mesmo para todos os anos pós-desastre, mas tal hipótese pode ser facilmente relaxada interagindo o indicador da intervenção com *dummies* de ano, conforme é feito na especificação usada para testar a hipótese das tendências paralelas.

Em algumas especificações são incluídos os chamados “*modifiers*”, covariadas afetadas pelo tratamento e que, portanto, absorvem parte do efeito de β : “Infraestrutura”, “Tecnologia” e “PIB per capita”. A ideia é verificar a composição do efeito global do desastre nos indicadores educacionais, a fim de mapear quais os canais de influência negativa tiveram papel mais relevante. A nível municipal apenas o canal de renda via PIB per capita é

incluído nas estimações.³⁸

Na medida em que os municípios ou escolas não tiveram o mesmo grau de exposição ao desastre, uma análise de heterogeneidade exige uma especificação que discrimine as unidades tratadas de acordo com sua categoria COPPETEC. Este trabalho foca nas especificações a nível de escola na medida em que determinadas áreas de tratamento possuem poucos municípios. A especificação por categorias é apresentada a seguir:

$$y_{it} = \sum_{j=1}^5 W_{ijt} \beta_j + \mu_i + \gamma_t + X_{it} \delta + \epsilon_{it} \quad (3)$$

em que W_{ijt} é igual a 1 para as observações no grupo de tratamento j nos períodos pós-desastre. Temos que j varia de 1 a 5 uma vez que as categorizações propostas pela COPPETEC abrangem 5 áreas distintas de tratamento. O efeito do desastre por categorias é medido pelo coeficiente β_j .

³⁸Vale lembrar que as estimações a nível municipal excluem as escolas privadas para as variáveis de interesse da base do Saeb. O desempenho médio nas disciplinas refere-se a uma média das escolas públicas federais, estaduais e municipais da amostra.

5.1.2 Alunos

A nível aluno, o método de DD será estimado via Modelo de Probabilidade Linear (em inglês, “LPM”) com dados em painel, considerando efeitos fixos de tempo, indivíduo e de escola. Opta-se por tal modelo por conta da sua vantagem computacional e facilidade na interpretação dos resultados, assim como em Nishijima e Rocha (2020). A variável dependente será o *status* de Reprovação, isto é, analisa-se o efeito do desastre na probabilidade de reprovação do aluno.

De forma análoga ao realizado para as análises a nível de escola e município, a primeira especificação considera a estimação com todos os tratados, conforme equação a seguir:

$$y_{ijt} = W_{it}\beta + \mu_i + \gamma_t + \theta_j + X_{ijt}\delta + \epsilon_{it} \quad (4)$$

em que y_{ijt} é uma *dummy* igual a 1 se o indivíduo i , na escola j , foi reprovado no instante t ; W_{it} é uma *dummy* igual a 1 para indivíduos do grupo de tratamento em períodos posteriores ao desastre; μ_i é o efeito fixo de indivíduo; γ_t é o efeito fixo de tempo; θ_j corresponde ao efeito fixo da escola j ; X_{ijt} contempla uma série de variáveis observáveis que variam no tempo como, por exemplo: idade do aluno, etapa de ensino cursada em cada período, tendências lineares de PIB, ICMS e CFEM per capita; e ϵ_{it} é o termo de erro aleatório. Conforme descrito anteriormente, o objetivo de utilizar tendências lineares para essas variáveis e não os valores de fato realizados é que tais variáveis são afetadas pelo desastre e poderiam absorver o efeito do tratamento. Estima-se também uma versão modificada da equação acima em que se controla pelo PIB per capita realizado e variáveis de infraestrutura e tecnologia das escolas, os *modifiers*, com o objetivo de estimar os efeitos do tratamento via canais psicológicos. Além disso, o número de alunos na turma do indivíduo i no instante t também foi adicionado como controle.

Uma segunda especificação considera a possibilidade de efeitos heterogêneos decorrentes do desastre entre os alunos das cinco diferentes áreas COPPETEC, com base na seguinte equação:

$$y_{ijt} = \sum_{k=1}^5 W_{ikt}\beta_k + \mu_i + \gamma_t + \theta_j + X_{ijt}\delta + \epsilon_{it} \quad (5)$$

em que W_{ikt} é uma *dummy* que assume valor 1 se o indivíduo i pertence à área k e $t > 2015$.

5.2 Controle Sintético

A segunda abordagem utilizada é o método de CS, desenvolvido por Abadie e Gardeazabal (2003) e expandido por Abadie, Diamond e Hainmueller (2010, 2014). Desde então o método tem sido cada vez mais utilizado para a análise de políticas agregadas e estudos de intervenção (estudos de caso) em uma ampla variedade de tópicos.³⁹Vale destacar o trabalho de Adhikari e Alm (2016), em que os autores avaliam os efeitos de reformas tributárias implementadas, entre 1994 e 2005, no PIB per capita de um conjunto de países do Leste Europeu e Ásia Central. Outro trabalho que merece destaque é o de Cavallo et al. (2013), que avalia os efeitos no crescimento econômico de desastres naturais. Estes dois estudos servirão de base para o arcabouço teórico do CS a ser resumido a seguir para o caso em questão.

O objetivo do método de CS é usar uma combinação convexa das observações do grupo de controle (“*donor pool*”) para criar um grupo de controle “sintético”. A diferença para o método de DD é que tal combinação é “*data-driven*” na medida em que o algoritmo do CS atribui pesos às unidades do *donor pool* de forma a aproximá-lo ao máximo da unidade tratada no período pré-tratamento. Ademais, conforme será detalhado adiante, o CS permite que heterogeneidades individuais variem no tempo.

Como será demonstrado na Seção 6, as variações do grupo de controle para o método de DD não se mostraram suficientes para encontrar o contrafactual adequado ao grupo de tratamento. Nesse sentido, Abadie, Diamond e Hainmueller (2010) argumentam que tal procedimento de escolha acaba se baseando em medidas subjetivas de afinidade entre unidades tratadas e não tratadas, na linha de Angrist e Pischke (2009) e Montalvo (2011).

A fim de garantir a identificação dos efeitos das reformas tributárias no PIB per capita, Adhikari e Alm (2016) adotam uma série de hipóteses comuns na literatura de CS, quais sejam:

³⁸Abadie e Gardeazabal (2003) introduzem o método de CS para avaliar o impacto no crescimento econômico dos conflitos terroristas no País Basco na década de 1960. Os autores estimam que o terrorismo na região gerou um declínio de 10% do PIB per capita em comparação ao grupo de controle não afetado pelos ataques. Abadie, Diamond e Hainmueller (2010) usam a mesma abordagem para avaliar o impacto no consumo de cigarros da Legislação Antitabagista introduzida, em janeiro de 1989, no Estado da Califórnia, nos Estados Unidos. Os resultados indicam redução de 25% no consumo no período entre a implementação da medida e o ano de 2000. Já o trabalho de 2014, com dados de 1960 até 2003, avaliou o impacto econômico na Alemanha Ocidental da reunificação em outubro de 1990, após 45 anos de separação. Os resultados indicam redução de 8% do PIB.

³⁹Além de descrever o método de CS, Adhikari e Alm (2016) fazem breve revisão da literatura recente enumerando uma série de artigos que lançam mão desta abordagem.

- i. Independência condicional: a ocorrência do tratamento e seus resultados potenciais são independentes, condicionais a um conjunto de características observadas;
- ii. Suporte comum: exige que existam unidades não tratadas com características similares à unidade tratada tal que a combinação ponderada das unidades não tratadas reproduza a evolução pré-tratamento da unidade tratada;
- iii. Ausência de choques exógenos: assume que as unidades de controle não passaram por outros choques exógenos que não o tratamento sob análise;
- iv. Ausência de *spillover-effects*: exige que não tenha havido interferência entre unidades tratadas e não tratadas após a ocorrência do tratamento de modo a viesar os resultados;
- v. Ausência de efeito antecipação: implica que as variáveis de interesse não variaram de forma a se antecipar do tratamento esperado.

A natureza extraordinária e, em grande medida, imprevisível - isto é, exógena ao modelo, do desastre em Mariana (MG) simplifica o leque de hipóteses a serem adotadas, especialmente em relação à adoção de reformas tributárias como é o caso dos autores. É razoável supor que o desastre é independente dos resultados potenciais das variáveis educacionais de interesse, bem como não faria sentido esperar nenhum efeito de antecipação no caso de Mariana (MG). Também não parece ter havido nenhum choque exógeno em determinados municípios da Bacia do Rio Doce que poderiam gerar algum tipo de ruído no *donor pool*. Os efeitos da crise econômica do biênio 2015/2016, vale ressaltar, são incorporados nas especificações por meio do PIB per capita dos municípios.

A fim de garantir a não interferência entre tratados e não tratados, assim como foi feito no método de DD, são retirados do *donor pool* aqueles municípios que receberam alunos tratados no período pós-desastre. Naturalmente, municípios que estiveram na trajetória da lama, categorizados em uma das áreas COPPETEC, também foram excluídos do *donor pool*. Conforme Abadie, Diamond e Hainmueller (2010), a hipótese de suporte comum é testada graficamente e por meio da raiz quadrada do erro médio de previsão ao quadrado (*Root Mean Square Prediction Error* - “RMSPE”).

Conforme mencionado, o método de CS tem sido muito utilizado para estudos de caso e análises de políticas agregadas. Neste trabalho optou-se por estudar especificamente o caso de Mariana (MG), sendo o *donor pool* composto de municípios não atingidos pelo rompimento da barragem e as variáveis de controle todas a nível municipal. Assim, sejam $I + 1$ municípios em que o município $i = 1$ representa Mariana (MG), fortemente atingida

pelo rompimento da barragem de Fundão no ano de 2015, e os I municípios restantes constituem potenciais controles (*donor pool*). Seja T_0 o número de anos pré-desastre, com $1 \leq T_0 < T$.

Dada a possibilidade de atrito com os dados do Saeb (desastre ocorreu no período de aplicação das provas), conforme discutido na Seção 4, bem como a menor disponibilidade de dados (provas do Saeb são bianuais), a aplicação do CS foca nos dados do Censo Escolar. Definiu-se como variável dependente para Mariana (MG) a taxa de reprovação nos anos finais do ensino fundamental.

Seja y_{it}^{SR} a variável educacional de interesse observada para o município i do *donor pool*, sem rompimento da barragem (SR), e seja y_{it}^R a variável educacional de interesse observada para Mariana (MG), com rompimento da barragem (R). O resultado observado pode ser escrito genericamente como:

$$y_{it} = \begin{cases} y_{it}^{SR}, & \text{na ausência de rompimento da barragem de Fundão; ou} \\ y_{it}^R \equiv y_{it}^{SR} + \tau_{it}D_{it}, & \text{na presença de rompimento da barragem de Fundão} \end{cases}$$

em que $\tau_{it} = y_{it}^R - y_{it}^{SR}$ representa o efeito do rompimento da barragem de Fundão no município i e ano t quando $D_{it} = 1$, o que ocorre se $t > T_0$ e $i = 1$. Tem-se que $D_{it} = 0$ para o período pré-tratamento ou para $i \neq 1$, o que faz com que $\tau_{it} = 0$. Como o desastre ocorreu em novembro de 2015, são estimadas especificações considerando 2015 tanto como pré ($T_0 = 6$) quanto como pós-tratamento ($T_0 = 5$).

Ocorre que o único resultado observado nos dados é y_{it}^R . É preciso estimar y_{it}^{SR} , que representa o resultado da variável educacional de interesse do município i , no ano t , na hipótese de não ter ocorrido o rompimento da barragem de Fundão. Para tanto, o método de CS assume um modelo linear da forma (CS1):

$$y_{it}^{SR} = \alpha_t + \theta_t Z_i + \lambda_t \mu_i + \epsilon_{it} \quad (6)$$

em que α_t representa um fator não observável comum entre os municípios; Z_i é um vetor de variáveis observadas com coeficientes θ_t - quais sejam: PIB; PIB per capita; valor adicionado bruto (VAB) da indústria no município; impostos, líquidos de subsídios, sobre produtos no município; além de diversas variáveis socioeconômicas do Censo de 2010 ("*baseline*") do IBGE, como taxa de alfabetização; rendimento médio da população;

percentual da população na População Economicamente Ativa (PEA); percentual da população urbana; percentual da população com acesso à energia elétrica; percentual da população com emprego com carteira assinada; percentual de ocupação em setores como o de mineração, construção, comércio, serviços, agropecuário, indústria da transformação e no setor de serviços industriais de utilidade pública; percentual da população branca, preta e parda; PIB e CFEM no *baseline*; além do percentual do município localizado na Bacia do Rio Doce, altitude do município e sua microrregião dentro da bacia. Também são consideradas como variáveis explicativas a taxa de reprovação nos anos finais do ensino fundamental nos anos de 2010, 2011 e 2012.

Já λ_t é um vetor de variáveis comuns não observadas, cujos parâmetros são dados pelo vetor μ_i ; e ϵ_{it} são os termos de erro aleatórios com média zero. Note-se que, ao contrário do método de DD, o método de CS permite que heterogeneidades individuais variem no tempo ($\lambda_t \mu_i$) - o método de DD permite a presença de fatores não observáveis, mas fixos no tempo.

Outra hipótese importante do método de CS (CS2) é a de que existe um vetor de pesos de dimensão $I + 1$, dado por $W^* = (w_2^*, w_3^*, \dots, w_{I+1}^*)'$, tal que $w_i^* \geq 0$ para $i = 2, \dots, I + 1$ e $w_2^* + w_3^* + \dots + w_{I+1}^* = 1$, que faz com que as seguintes equações sejam válidas:

$$\sum_{i=2}^{I+1} w_i^* y_{it} = y_{1t}, \forall t \leq T_0 \quad (7)$$

$$\sum_{i=2}^{I+1} w_i^* Z_i = Z_1 \quad (8)$$

Sob CS1 e CS2 o estimador $\hat{\tau}_{1t} = y_{1t} - \sum_{i=2}^{I+1} w_i^* y_{it}, \forall t > T_0$ é consistente e não viesado para τ_{1t} , conforme provam Abadie, Diamond e Hainmueller (2010). Para isso, no entanto, é importante que o período pré-tratamento seja longo o suficiente, isto é, quanto maior T_0 , melhor. Esta é, precisamente, a principal limitação da abordagem de CS neste trabalho uma vez que o período pré-tratamento começa em 2010 e termina em 2015, isto é, com no máximo seis anos de período pré-tratamento.

³⁹Definidos como SIUP: geração, transmissão e distribuição de eletricidade e gás, captação, tratamento e distribuição de água, esgoto e atividades relacionadas; coleta, tratamento e disposição de resíduos e recuperação de materiais; descontaminação e outros serviços de gestão de resíduos. Considera-se apenas a população acima de 18 anos.

Conforme descreve Adhikari e Alm (2016), o algoritmo do CS é tal que o vetor de pesos ótimos \hat{W} surge da minimização do RMSPE da variável educacional de interesse no período pré-tratamento. Recebem pesos maiores as covariadas e as unidades do *donor pool* que têm maior poder preditivo da variável dependente.

6 Resultados

6.1 Diferenças em Diferenças

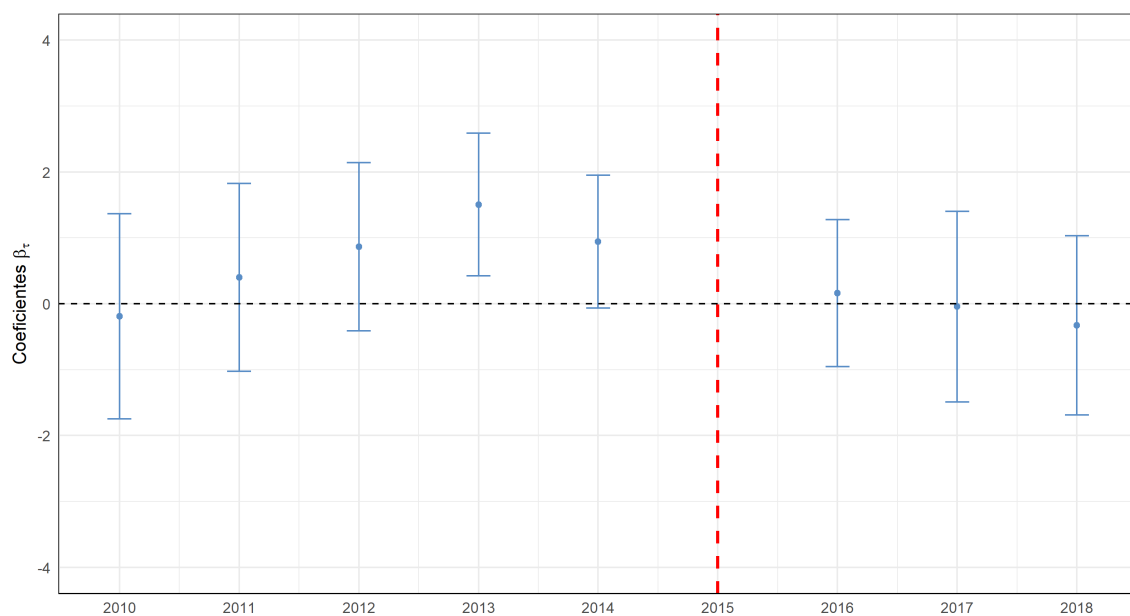
A fim de testar a hipótese fundamental de tendências paralelas, a Figura 7 apresenta o exercício de “*Event-Study*” a nível de escola para a taxa de reprovação nos anos finais do ensino fundamental. Nota-se que apenas o coeficiente $\beta_{\tau=-2}$, para o ano de 2013, mostrou-se significativo. Embora não reportado nesta dissertação, o resultado é similar a nível municipal. Tampouco houve diferenças significativas quando adotadas as demais variações do grupo de controle sugeridas na Seção 5.

Ao discriminar o tratamento de acordo com a categorização COPPETEC, observou-se que o grupo de controle considerado pode não representar o contrafactual adequado para algumas categorias. Para a área 1, com exceção de 2014, todos os anos pré-tratamento apresentaram coeficientes significativos, conforme mostra a Figura 21, no Apêndice. De fato, a área 1, composta pelos municípios de Barra Longa, Mariana, Rio Doce e Santa Cruz do Escalvado, parece possuir características bem distintas das demais. Por conta disso, adota-se o método do Controle Síntético para avaliar os efeitos do desastre no município de Mariana (MG), a ser apresentado na Subseção 6.2.

Para a área 4 também há evidências contrárias à hipótese de tendências paralelas, apenas o ano de 2012 é que apresenta coeficiente estatisticamente não significativo. A área é composta por apenas dois municípios do Espírito Santo, Aracruz e Linhares. Uma vez que o grupo de controle não contempla os municípios não afetados sem sede na Bacia do Rio Doce, acaba por excluir a maioria dos municípios capixabas.

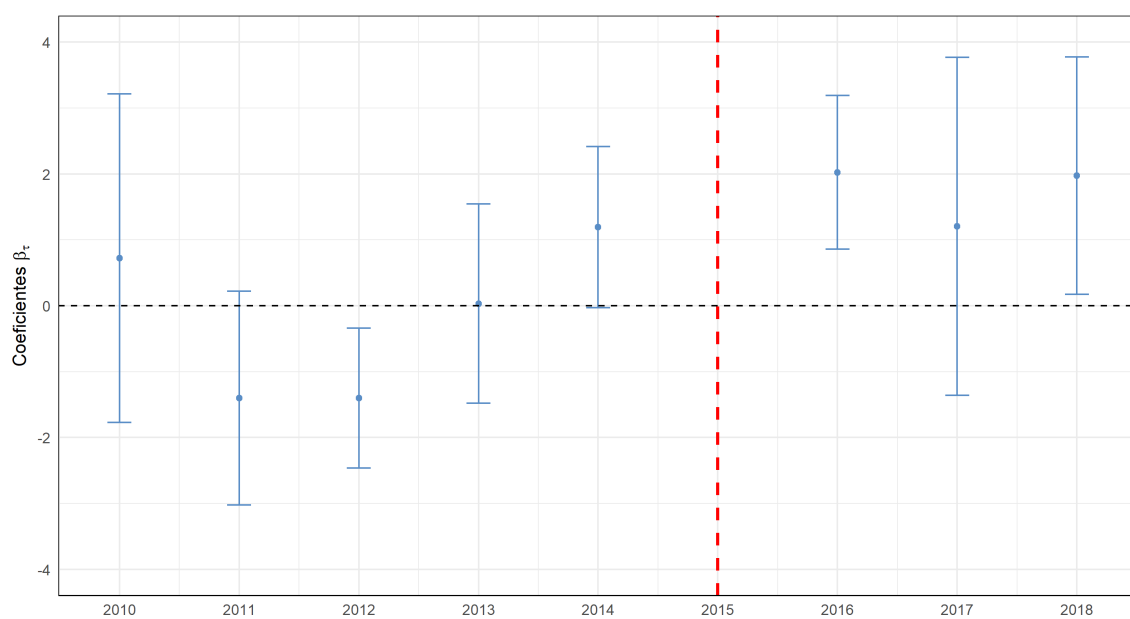
Já as áreas 2, 3 e 5 assemelham-se ao grupo de tratado agregado e praticamente todos os coeficientes pré-tratamento mostraram-se não significativos. A Figura 8 ilustra os resultados do teste de tendências paralelas para a área 3. Nota-se que apenas o ano de 2012 é que apresentou coeficiente estatisticamente significativo no período pré-tratamento.

Figura 7 – *Event Study* - Taxa de Reprovação Anos Finais EF - Escolas



Fonte: Censo Escolar. Especificação agregada: considera o conjunto de escolas tratadas contra o conjunto de escolas de controle. Elaboração própria.

Figura 8 – *Event Study* - Taxa de Reprovação Anos Finais EF - Escolas - Área 3



Fonte: Censo Escolar. Especificação desagregada: considera o conjunto de escolas tratadas da área 3 COPPETEC contra o conjunto de escolas de controle. Elaboração própria.

No caso do desempenho médio em Português e Matemática, o recorte temporal é menor, já que foram utilizados dados bianuais de 2011 a 2017. Embora menos robustos, os exercícios de *Event-Study* apresentam evidência a favor da hipótese de tendências paralelas, tanto a nível de escola quanto de município. Ao discriminar o grupo de tratamento, tais evidências permanecem válidas apenas para as áreas 1, 3 e 5.⁴⁰

Já a Figura 22, no Apêndice, reproduz os resultados encontrados no exercício para o nível aluno. Conforme mencionado anteriormente, os coeficientes de tratamento β_τ para $\tau \in \{-1, -2, -3, -4, -5\}$ deveriam ser não significativos para corroborar com a hipótese de tendências paralelas. No entanto, foram encontrados coeficientes significativos para três dos cinco períodos anteriores ao desastre, um forte indício de que a hipótese de tendências paralelas foi violada.

Assim como no caso das especificações a nível de escola e município, foram testadas diferentes combinações entre os grupos de tratamento e controle para o nível aluno, a saber: (i) utilização do *propensity score* para parear alunos do controle aos alunos tratados, na razão de 1 para 1, definindo dessa forma um novo grupo de controle;⁴¹ (ii) discriminação do grupo de tratamento por áreas COPPETEC, sendo o grupo de controle o mesmo da especificação original; e (iii) discriminação do grupo de tratamento por áreas COPPETEC, mas com um novo grupo de controle criado via *propensity score*, também na razão de 1 para 1. Em todas as variações testadas o resultado dos testes de tendências paralelas não foi satisfatório, isto é, foram encontrados coeficientes significativos relativos a momentos anteriores ao desastre.

Ressalte-se que os resultados do *Event-Study* observados para o nível aluno não contradizem, necessariamente, àqueles observados para os níveis de escola e município. Isso porque o conjunto de alunos tratados não representa exatamente o conjunto de alunos pertencentes às escolas consideradas tratadas. Alunos afetados pelo desastre em 2015, considerados tratados na especificação no nível do aluno e seguidos ao longo do tempo, podem ter mudado de escola ou município no período pré-tratamento, passando de uma escola ou município do grupo de tratamento para uma escola ou município do grupo de controle, e vice-versa.

⁴⁰Como os resultados são similares, optou-se por não reportar os exercícios de *Event Study* a nível municipal, nem aqueles que consideram as variáveis de desempenho médio do Saeb.

⁴¹Para estimar o *propensity score*, utilizou-se um modelo Logit com as seguintes variáveis explicativas: Idade, Sexo, PIB per capita do município entre 2010 e 2015, tempo estudado em escola privada, unidade federativa, etapa de ensino cursada em 2015, número de reprovações entre 2010 e 2015, além de raça.

⁴¹De todo modo, a Subseção 6.1.2 apresenta alguns dos resultados encontrados para as análises a nível de aluno.

A Subseção 6.1.1 apresenta os resultados das estimações a nível de escola e de município. Já a Subseção 6.2.1 apresenta os resultados para o nível do aluno.

6.1.1 Escolas e Municípios

A Tabela 2 apresenta os resultados a nível de escola para a taxa de reprovação nos anos finais do ensino fundamental. Todas as especificações contêm cluster por município, além de efeitos fixos de escola e tempo. Em linha com FGV (2019), não se observa efeito significativo do desastre na estimação agregada (coluna 1). Nota-se efeito positivo na variável Infraestrutura (coluna 2), o que indica que o modelo pode ter capturado efeitos de medidas compensatórias nas condições de infraestrutura física das escolas afetadas. Já a variável PIB per capita não parece ter absorvido efeitos na renda.

A discriminação por áreas (coluna 3) revela efeito significativo e positivo na área 3, que é composta pelas escolas localizadas em municípios afetados pela lama, mas sem ocupação próxima ao Rio Doce. Para a variável dependente em questão, tal área apresentou evidências favoráveis para a hipótese das tendências paralelas, de modo que os resultados indicam que o desastre pode ter elevado a taxa de reprovação nestas escolas. Não se verifica efeitos significativos nas demais escolas das outras áreas COPPETEC.

Quando se incluem os *modifiers* (coluna 4), nota-se que os efeitos positivos para a taxa de reprovação na área 3 continuam estatisticamente significantes. Isto é, mesmo quando se controla pela possível melhora nas condições de infraestrutura das escolas, a taxa de reprovação continua maior pós-desastre, embora em menor magnitude. Observa-se ainda efeitos na área 1, a qual, no entanto, viola a hipótese fundamental de tendências paralelas.

A Tabela 10, no Apêndice, traz os mesmos resultados, mas a nível de município. Todas as especificações contêm cluster por município, além de efeitos fixos de município e tempo. Em linha com FGV (2019), não se observa efeito significativo do desastre na estimação agregada (coluna 1). Não se verifica, neste caso, efeito significativo para a área 3, apenas para a área 2, mas negativo. Esta área também apresentou evidências favoráveis para a hipótese de tendências paralelas, de modo que os resultados a nível municipal indicam que o desastre pode ter reduzido a taxa de reprovação para estes municípios.

⁴¹O fato de a variável PIB per capita não se mostrar relevante para explicar a taxa de reprovação (coluna 2) indica que o efeito surge da ausência das tendências lineares, omitidas nesta especificação.

Tabela 2 – Taxa de Reprovação Anos Finais EF - Escolas

	(1)	(2)	(3)	(4)
Pós × Tratado	-0.21 (0.31)	-0.43 (0.34)		
Infraestrutura		2.09*** (0.48)		2.10*** (0.48)
Tecnologia		-0.87 (0.45)		-0.80 (0.44)
PIB per capita		-0.00 (0.00)		-0.00 (0.00)
Pós × Tratado 1			-1.14 (0.79)	-1.77** (0.65)
Pós × Tratado 2			-0.25 (0.34)	-0.47 (0.38)
Pós × Tratado 3			1.43*** (0.25)	1.07** (0.38)
Pós × Tratado 4			0.29 (0.66)	0.31 (0.62)
Pós × Tratado 5			-0.91 (0.81)	-1.13 (0.80)
Constante	7.09*** (0.37)	7.10*** (0.35)	6.98*** (0.37)	7.13*** (0.36)
Observações	26250	23663	26250	23663
R^2 Ajustado	0.045	0.049	0.047	0.050
Modifiers		✓		✓

Erros padrão robusto em parênteses, com cluster por município.
 Colunas (1) e (3) contêm como controle tendências lineares de PIB, ICMS e CFEM per capita.

Todas as regressões contêm efeitos fixos de escola e de tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A Tabela 3 apresenta uma análise acerca da heterogeneidade dos efeitos a nível de escola. Assim como no resultado agregado, não se observa efeitos específicos do desastre nem para as escolas privadas nem públicas (coluna 1). Para escolas rurais o efeito é negativo e marginalmente significativo (coluna 2). Isto é, o desastre pode ter reduzido a taxa de reprovação nas escolas afetadas pela lama e localizadas na zona rural.

Tabela 3 – Análise de Heterogeneidade - Taxa de Reprovação Anos Finais EF - Escolas

	(1)	(2)
Pós × Tratado × Privada	-0.06 (0.28)	
Pós × Tratado × Pública	-0.23 (0.34)	
Pós × Tratado × Urbana		0.07 (0.34)
Pós × Tratado × Rural		-0.85 (0.45)
Constante	7.10*** (0.38)	7.08*** (0.38)
Observações	26250	26250
R^2 Ajustado	0.045	0.046

Erros padrão robusto em parênteses, com cluster por município.

Todas as regressões contêm efeitos fixos de escola e de tempo,

bem como a linear tendências lineares de PIB, ICMS e CFEM per capita

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A Tabela 4 apresenta os efeitos do desastre no desempenho médio dos alunos em Matemática, a nível de escola. Não se observa efeito significativo para o conjunto de escolas tratadas (coluna 1). O mesmo ocorre quando se considera a adição de variáveis que podem ter sido afetadas pelo desastre: Infraestrutura, Tecnologia e PIB per capita (coluna 2).

Quando se discrimina pelas áreas COPPETEC (coluna 3), nota-se efeito positivo e significativo na área 4, que no entanto não parece ser comparável ao grupo de controle de acordo com os testes de tendências paralelas. Já na área 1 o efeito é negativo e marginalmente significativo. Torna-se ainda mais significativo com a inclusão dos *modifiers* (coluna 4). Como nesse caso há evidências a favor da hipótese de tendências paralelas, o resultado sugere que o desastre pode ter impactado negativamente o desempenho em Matemática dos alunos do 5º ano do ensino fundamental.

Tabela 4 – Desempenho Matemática 5º Ano do EF - Escolas

	(1)	(2)	(3)	(4)
Pós × Tratado	1.67 (1.15)	0.94 (1.19)		
Infraestrutura		1.03 (2.41)		0.96 (2.43)
Tecnologia		6.72** (2.55)		6.79** (2.56)
PIB per capita		-0.00 (0.00)		-0.00 (0.00)
Pós × Tratado 1			-2.63 (1.46)	-5.38*** (1.57)
Pós × Tratado 2			0.87 (1.26)	0.30 (1.24)
Pós × Tratado 3			5.51 (3.72)	5.41 (3.68)
Pós × Tratado 4			3.49** (1.31)	2.87* (1.33)
Pós × Tratado 5			1.90 (2.31)	1.01 (2.92)
Constante	222.98*** (1.01)	215.85*** (2.11)	222.66*** (1.07)	216.03*** (2.10)
Observações	3528	3538	3528	3538
R^2 Ajustado	0.078	0.079	0.079	0.081
Modifiers		✓		✓

Erros padrão robusto em parênteses, com cluster por município.

Colunas (1) e (3) contêm como controle tendências lineares de PIB, ICMS e CFEM per capita.

Todas as regressões contêm efeitos fixos de escola e de tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Conforme mostra a Tabela 11, no Apêndice, os resultados agregados são similares a nível de município (coluna 1). Neste caso, considera-se apenas as escolas públicas dos municípios analisados. Na discriminação por áreas só há efeito significativo e positivo para a área 3 (coluna 3). Como foi mencionado, tal área parece comparável ao grupo de controle em questão, de modo que os resultados indicam que pode ter havido uma melhora no desempenho médio em Matemática dos alunos de escolas públicas em decorrência do

desastre. Vale lembrar que foge do escopo desta pesquisa isolar os efeitos positivos de eventuais medidas compensatórias e de reparação, e que portanto se misturam com os efeitos analisados.

Para o desempenho médio em Português os resultados são similares. A Tabela 5 apresenta os resultados a nível de escola, enquanto a Tabela 12, no Apêndice, apresenta os resultados a nível de município. Não se observam efeitos nem na estimação agregada de escolas tratadas ante escolas do controle (coluna 1) nem quando se discriminam as áreas COPPETEC (coluna 3).

Já ao nível de município, considerando apenas escolas públicas, nota-se efeito positivo e significativo na área 3 (coluna 3). O resultado é ainda robusto à inclusão do PIB per capita (coluna 4). Para a variável dependente em questão, embora menos robusto, o teste de tendências paralelas mostrou-se favorável.

Tabela 5 – Desempenho Português 5^o Ano do EF - Escolas

	(1)	(2)	(3)	(4)
Pós × Tratado	1.28 (1.19)	0.64 (1.30)		
Infraestrutura		0.55 (2.58)		0.44 (2.57)
Tecnologia		4.79 (2.74)		4.79 (2.74)
PIB per capita		0.00 (0.00)		0.00 (0.00)
Pós × Tratado 1			1.53 (2.58)	-2.24 (3.30)
Pós × Tratado 2			0.83 (1.55)	0.44 (1.59)
Pós × Tratado 3			4.67 (3.27)	4.77 (3.24)
Pós × Tratado 4			0.74 (1.65)	0.36 (1.74)
Pós × Tratado 5			1.69 (2.20)	0.41 (3.08)
Constant	201.79*** (0.96)	195.47*** (2.16)	201.65*** (1.04)	195.65*** (2.21)
Observações	3528	3538	3528	3538
R ² Ajustado	0.420	0.420	0.420	0.420
Modifiers		✓		✓

Erros padrão robusto em parênteses, com cluster por município.

Colunas (1) e (3) contêm como controle tendências lineares de PIB, ICMS e CFEM per capita.

Todas as regressões contêm efeitos fixos de escola e de tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6.1.2 Alunos

Conforme discutido no início desta seção, os testes empregados para a verificação da hipótese de tendências paralelas não foram satisfatórios ao nível aluno. Desta forma, os resultados são apresentados apenas no Apêndice. A Tabela 13 apresenta os resultados dos efeitos do desastre na probabilidade de reprovação dos alunos tratados. Todas as regressões consideram efeitos fixos de aluno, escola e tempo, além de controles de idade do aluno e etapa de ensino cursada.⁴² Quando se considera o conjunto de alunos tratados (coluna 1), nota-se um efeito significativo, embora pequeno, de redução na probabilidade de reprovação.

Ao discriminar o grupo de tratamento entre as áreas COPPETEC (coluna 3), percebe-se que o efeito do desastre é heterogêneo, tendo as áreas 2 e 5 efeitos negativos. Nota-se ainda efeito positivo e marginalmente significativo na área 1, mas que desaparece ao se controlar pelos *modifiers* (coluna 4).

A Tabela 14 contempla as análises de heterogeneidade dos efeitos do desastre.⁴³ Percebe-se que o efeito de redução de reprovações é dominado pelos alunos de escola privada, ao passo que alunos de escola pública têm um aumento marginalmente significativo na probabilidade de reprovação. Entre homens e mulheres, os efeitos têm sentidos opostos, com aumento de cerca de 2% na probabilidade de reprovação entre tratados homens e redução de aproximadamente 3% entre tratados do sexo feminino. Algo semelhante é encontrado entre tratados brancos, que têm redução de probabilidade de aproximadamente 2,6%, e demais raças, com aumento marginal na probabilidade de reprovação.

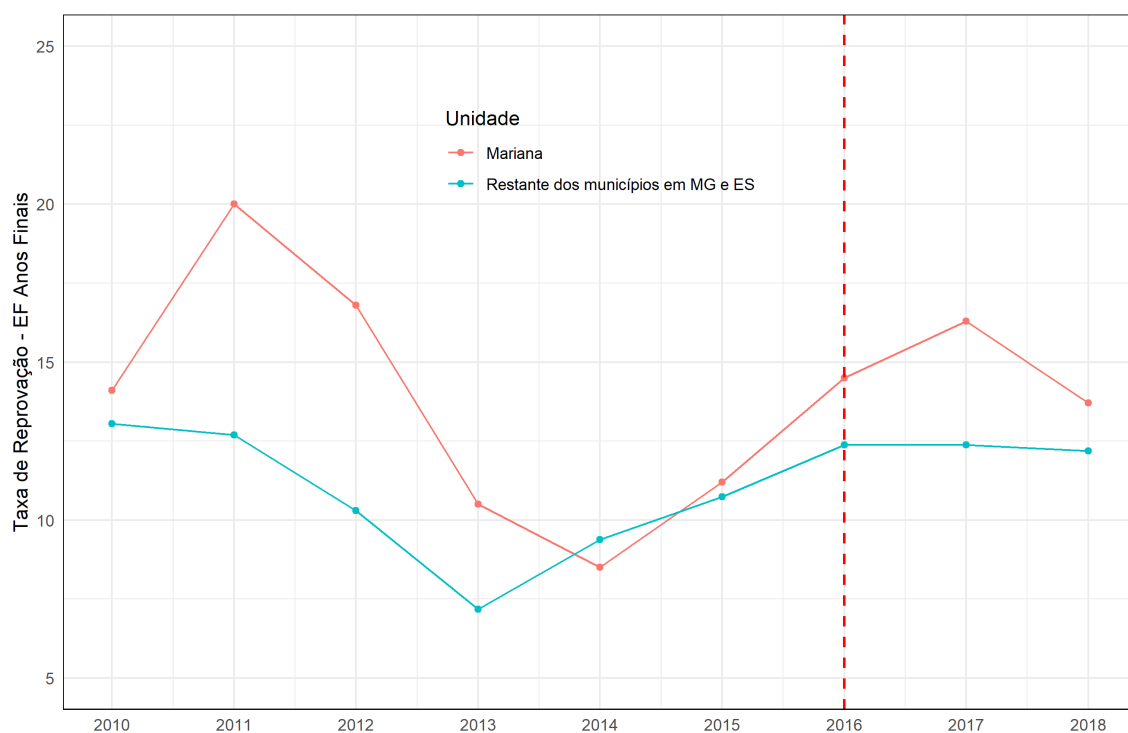
⁴²Assim como nas estimações ao nível de escola e município, as colunas 1 e 3 utilizam como controles adicionais as tendências lineares de ICMS, CFEM e PIB per capita.

⁴³Os controles destas regressões são os mesmos empregados na regressão da coluna 1 da Tabela 13.

6.2 Controle Sintético

A Figura 9 compara a evolução da taxa de reprovação nos anos finais do ensino fundamental em Mariana (MG) com a média simples do restante dos 212 municípios de MG e ES com sede na Bacia do Rio Doce, de 2010 até 2018. Considera-se a taxa média de reprovação para o total de todas as escolas federais, estaduais, municipais e privadas dos municípios da amostra. A linha vermelha tracejada vertical indica que o período de tratamento se inicia em 2016, sendo 2015 o último ano pré-tratamento. Nota-se que a taxa de reprovação em Mariana (MG) costuma ser mais elevada - com exceção do ano de 2014 -, mas não parece acompanhar a média das taxas de reprovação dos restantes dos municípios ao longo do período pré-rompimento. Isto é, a análise gráfica da evolução das taxas pré-tratamento sugere que o conjunto de municípios da bacia não representa um bom contrafactual para Mariana (MG).

Figura 9 – Mariana vs. Bacia do Rio Doce (MG e ES) - Taxa de Reprovação Anos Finais do EF



Fonte: Censo Escolar. O ano de 2015 é considerado como o último ano pré-tratamento. Elaboração própria.

A Tabela 6 compara as médias das variáveis explicativas em Mariana (MG) pré-desastre do controle sintético de Mariana (MG) e do restante da Bacia do Rio Doce. Nota-se que, com exceção do percentual da população parda do Censo de 2010, todas as outras médias

das covariadas do controle sintético de Mariana (MG) estão mais próximas da Mariana (MG) real. Isto é, o controle sintético reproduz as variáveis preditoras da unidade tratada muito melhor do que a média simples do restante da bacia.

Conforme Abadie, Diamond e Hainmueller (2010), o controle sintético força o pesquisador a demonstrar a afinidade entre a região exposta ao tratamento e sua contrapartida contrafactual dada pela combinação convexa das regiões do *donor pool*. Os pesos reportados na Tabela 7 indicam que a taxa de reprovação nos anos finais do ensino fundamental em Mariana (MG) antes do rompimento da barragem de Fundão pode, dadas estas covariadas, ser reproduzida pela combinação dos dados de Barão de Cocais (MG), Ferros (MG), Ponte Nova (MG), São Gonçalo do Rio Abaixo (MG), Virginópolis (MG) e Laranja da Terra (ES). Todos os outros municípios do *donor pool* receberam peso zero.

Tabela 6 – Médias das covariadas para Taxa de Reprovação - Anos Finais do EF

Variáveis	Mariana (MG)		Bacia do Rio Doce
	Real	Controle Sintético	
PIB per capita	87.667,14	52.002,54	13.228,16
PIB	4.946.433,00	1.036.710,00	371.158,41
Taxa de Alfabetização (2010)	94%	93,43%	86,25%
Rendimento Médio (2010)	1.098,64	889,41	670,95
Pop. PEA (2010)	56,14	54,05	53,05%
Pop. Urbana (2010)	88%	80,69%	60,4%
Pop. Ocupada na Mineração (2010)	14,64%	7,34%	1,05%
Pop. Branca (2010)	29,8%	29,44%	35%
Pop. Preta (2010)	18,5%	10,78%	9%
Pop. Parda (2010)	48,9%	58,5%	54%
PIB (2010)	3.690.160,56	885.351,5	289.912,36
CFEM (2010)	65.446.556,18	18.400.000,00	1.098.024,13
Área na Bacia do Rio Doce	100%	99,8%	95,8%
Altitude da sede do(s) município(s) (m)	705,0	709,6	496,7
Código Migrorregião	33	33,06	40,2
Pop. ocupada nos Serviços Industriais de Utilidade Pública (2010)	1,02%	1,12%	0,68%
Pop. Ocupada na Construção (2010)	11,88%	13,5%	7,88%
Pop. Ocupada na Comércio (2010)	13,47%	12,36%	9,64%
Pop. Ocupada na Serviços (2010)	42,76%	38,76%	32,08%
Pop. Ocupada no Agropecuário (2010)	6,92%	11,28%	39,6%
Pop. Ocupada na Indústria de Transformação (2010)	4,93%	6,49%	5,58%
VAB industrial	3.415.395,00	573.564,8	142.618,28
Impostos	136.502,8	67.960,69	30.991,34
Pop. com acesso à energia elétrica (2010)	99,34%	99,46%	98,69%
Taxa de emprego formal (2010)	55,93%	48,45%	23,54%
Taxa de reprovação Anos Finais EF (2010)	14,1%	14,2%	13,1%
Taxa de reprovação Anos Finais EF (2011)	20,0%	19,4%	12,7%
Taxa de reprovação Anos Finais EF (2012)	16,8%	16,2%	10,3%

Fonte: Censo Escolar. Elaboração própria.

Tabela 7 – Peso dos municípios na Mariana (MG) sintética

Município	Peso (%)
Barão de Cocais (MG)	69
Ferros (MG)	4,9
Ponte Nova (MG)	5,8
São Gonçalo do Rio Abaixo (MG)	10
Virginópolis (MG)	10,2
Laranja da Terra (ES)	0,1

Elaboração própria.

Entretanto, conforme ressaltam Ferman, Pinto e Possebom (2020), uma limitação importante do método de CS é o fato de que não há um consenso na literatura a respeito da escolha das covariadas a serem utilizadas na estimação dos pesos no *donor pool*. Se diferentes especificações geram diferentes escolhas das unidades que compõem o grupo contrafactual, então o pesquisador tem a oportunidade de selecionar apenas aquelas especificações que geram resultados estatisticamente significantes, mesmo quando não há, de fato, efeito algum. Tal arbitrariedade no processo de escolha das variáveis é conhecida na literatura econômica como “*cherry picking*”.

De modo a evitar vieses desta natureza, os autores recomendam que o foco seja colocado na especificação que contenha todos os períodos defasados (“*lags*”) da chamada “*outcome variable*” - aqui, no caso, a taxa de reprovação nos anos finais do ensino fundamental -, a não ser que haja uma crença clara de que seja necessário utilizar um dado conjunto de covariadas. Como este não é o caso da pesquisa em questão, considera-se, assim, a especificação do CS com todos os *lags* do período pré-tratamento, isto é, de 2010 até 2015.

A Figura 10 apresenta os resultados do controle sintético para Mariana (MG) considerando 2016 como o primeiro ano pós-rompimento ($T_0 = 6$).⁴⁴ Consideram-se todas as covariadas mencionadas na Subseção 5.2⁴⁵ e um total de 168 municípios não tratados e com sede na Bacia do Rio Doce no *donor pool*.

Nota-se que, ao contrário da taxa de reprovação média nos anos finais do ensino fundamental dos municípios da Bacia do Rio Doce (mostrada na Figura 9), a taxa de reprovação na Mariana (MG) sintética é idêntica à trajetória de Mariana (MG) antes do rompimento da barragem de Fundão. A média dos erros de previsão ao quadrado (em inglês, *Mean Squared Prediction Errors* - “MSPE”) de Mariana (MG) - dada pela média, entre 2010 e 2015, do quadrado das diferenças anuais entre a taxa de reprovação em Mariana (MG) e a

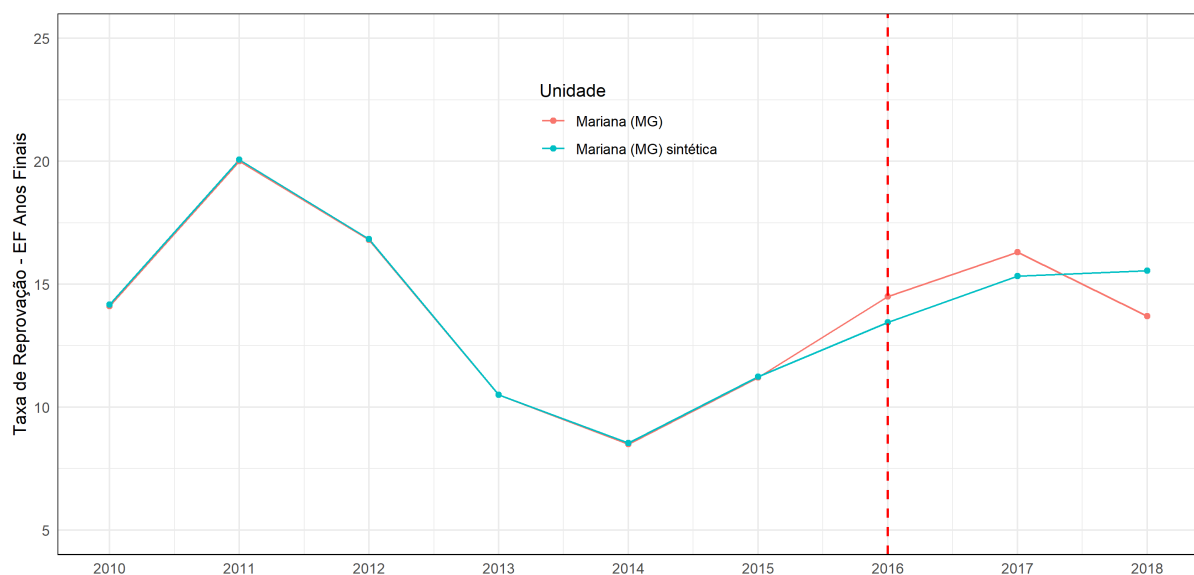
⁴⁴O RMSPE foi de 0,0024.

⁴⁵O método de CS também pondera as covariadas de modo a melhor reproduzir a unidade de tratamento no período pré-desastre.

da Mariana (MG) sintética - é de **0,359**. Combinado ao alto grau de balanceamento entre os preditores, isso sugere que a Mariana (MG) sintética oferece uma boa aproximação da taxa de reprovação nos anos finais do ensino fundamental que teria sido observada nas escolas de Mariana (MG) entre 2016 e 2018, na ausência do desastre.

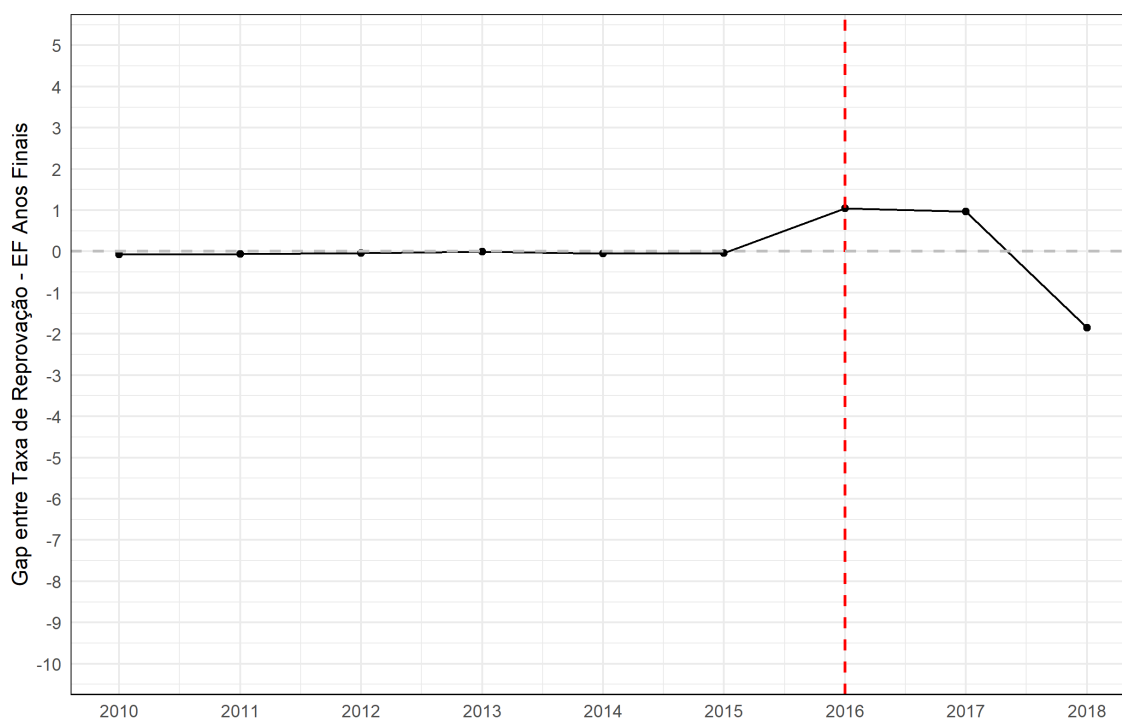
A estimativa do efeito do tratamento é dada pela diferença entre a taxa de reprovação nos anos finais do ensino fundamental observada em Mariana (MG) e aquela estimada para a Mariana (MG) sintética (“*gap*”). A Figura 11 apresenta os efeitos anuais do tratamento ao longo do período analisado, isto é, os *gaps* entre Mariana (MG) e Mariana (MG) sintética entre 2010 e 2018. Os resultados indicam que, na ausência do desastre, as taxas de reprovação teriam sido em até 1 ponto percentual menores do que as de fato observadas em Mariana (MG) em 2016 e 2017. Isto é, tais resultados indicam que o desastre em Mariana (MG) teria elevado as taxas de reprovação nas escolas no município.

Figura 10 – Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Todos os *lags* da *outcome variable*



O ano de 2015 é considerado como o último ano pré-tratamento. Elaboração própria.

Figura 11 – Gaps na Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Todos os *lags* da *outcome variable*



O ano de 2015 é considerado como o último ano pré-tratamento. Elaboração própria.

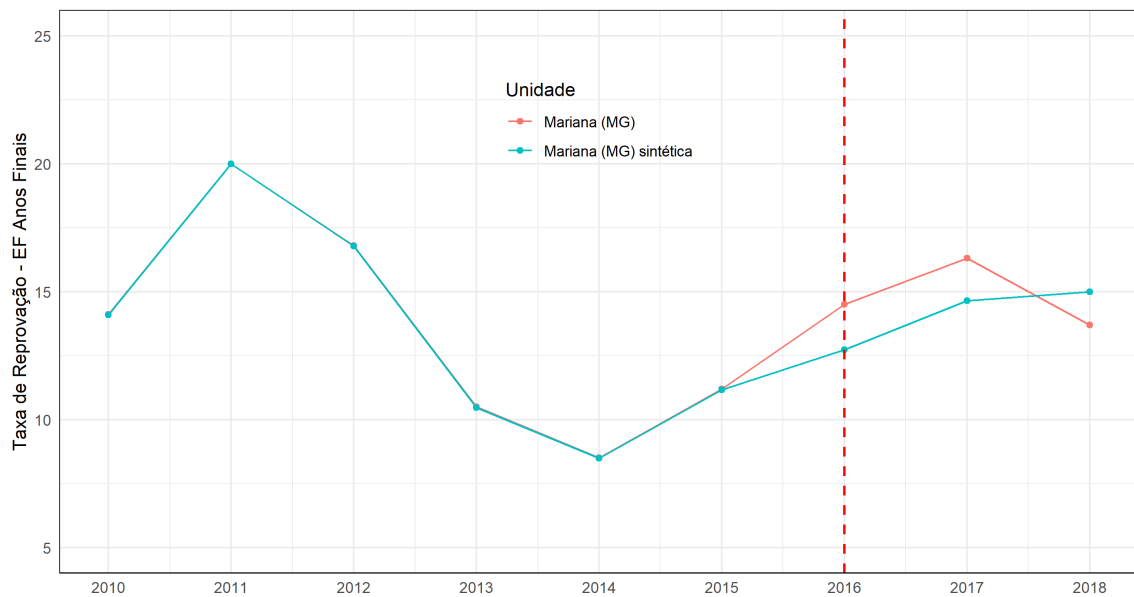
Conforme discutido na Seção 5, Abadie, Diamond e Hainmueller (2010) provam que, sob CS1 e CS2, o estimador do efeito do rompimento da barragem de Fundão é consistente e não viesado. Para tanto, é necessário que T_0 seja grande o suficiente, o que pode não ser o caso do exercício em questão. De todo modo, são realizados alguns testes de robustez comumente discutidos na literatura.

O primeiro teste realizado foi a avaliação da hipótese de ausência de *spillover-effects*. Assim como no método de DD, ajusta-se a composição dos municípios do *donor pool* de modo a evitar qualquer tipo de contaminação e interferência entre o município tratado, Mariana (MG), e os demais municípios de MG e ES da Bacia do Rio Doce. Desse modo, o controle sintético é novamente aplicado para um *donor pool* que agora não contém os 40 municípios em que houve, no período pós-tratamento, uma movimentação de alunos superior a 2,5% do total de alunos matriculados.⁴⁶Neste teste, portanto, o *donor pool* contém apenas 128 municípios.

⁴⁶Nota-se que tal teste é conservador na medida em que são excluídos não só municípios que receberam um fluxo anormal de alunos oriundos de Mariana (MG) no período pós-tratamento, mas também outros que podem ter recebido alunos oriundos de outras regiões localizadas nas áreas COPPETEC.

A Figura 12 mostram que os resultados são similares àqueles apresentados na Figura 10. Isso indica que os resultados apresentados anteriormente, com os 168 municípios, são robustos à variação do *donor pool*. A Figura 23, no Apêndice, apresenta os efeitos anuais do tratamento ao longo do período analisado.

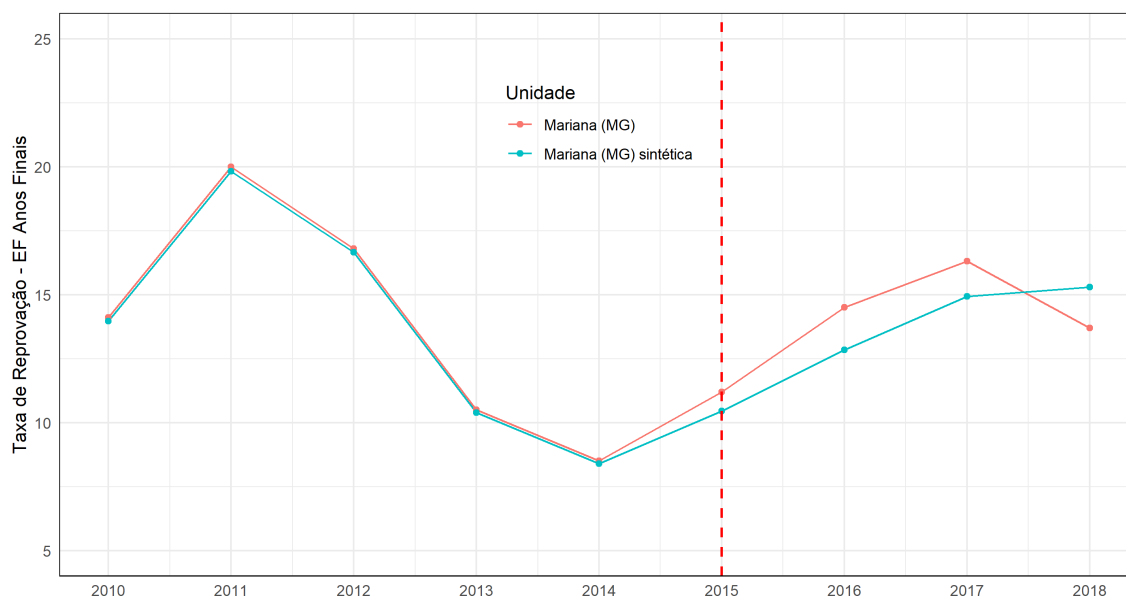
Figura 12 – Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Variação no *donor pool*: *threshold* de 2,5% - Todos os *lags* da *outcome variable*



O ano de 2015 é considerado como o último ano pré-tratamento. Elaboração própria.

Outra hipótese testada foi a de ausência de efeito antecipação. Toma-se, agora, 2015 como um ano pós-tratamento, isto é, $T_0 = 5$. Dada a natureza exógena do desastre ocorrido em Mariana (MG), não faria sentido esperar algum tipo de variação na sua taxa de reprovação frente ao controle sintético que fosse anterior ao tratamento. No entanto, conforme ilustra a Figura 13, o ano de 2015 já indica uma elevação da taxa de reprovação em Mariana (MG) maior do que a da Mariana (MG) sintética. Dessa forma, o resultado levanta dúvidas sobre a escolha do ano de 2015 como um ano pré-desastre: embora o rompimento tenha ocorrido apenas em novembro de 2015, pode ser que tenha afetado a nota anual dos alunos ou a decisão dos professores pela reprovação.

Figura 13 – Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética (2015 pós) - Todos os *lags* da *outcome variable*



O ano de 2014 é considerado como o último ano pré-tratamento. Elaboração própria.

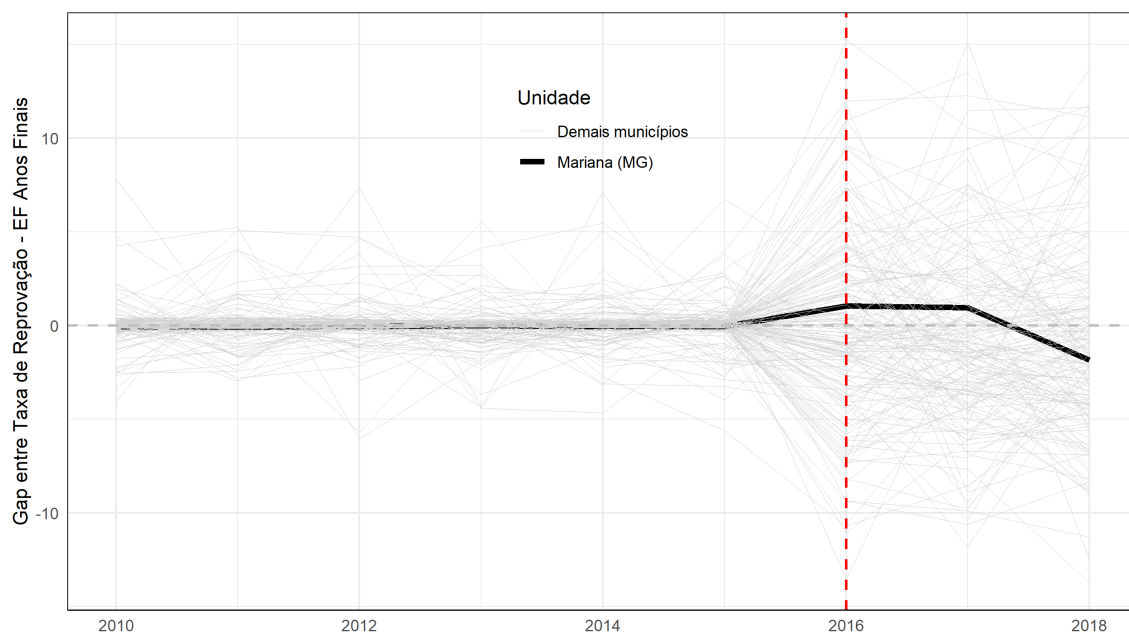
Para avaliar a significância das estimativas encontradas, verifica-se ainda se os resultados foram observados totalmente por acaso. Faz-se a seguinte pergunta: “Com que frequência seriam observados resultados da magnitude estimada se o município escolhido para tratamento tivesse sido escolhido de maneira aleatória ao invés de Mariana (MG)?”. Para responder a esta pergunta são realizados testes de placebo, em linha com Adhikari e Alm (2016); Abadie, Diamond e Hainmueller (2010); entre outros trabalhos envolvendo o método de controle sintético.

Assim, o mesmo exercício realizado para Mariana (MG) é realizado iterativamente para cada um dos 167 demais municípios de MG e ES com sede na Bacia do Rio Doce e que não foram afetados pelo tratamento. Naturalmente, Mariana (MG) é excluída dos *donor pools* para evitar uma possível interferência nos grupos de controle. Conforme destaca Abadie, Diamond e Hainmueller (2010), se os exercícios de placebo geram *gaps* de magnitude similar àquele estimado para Mariana (MG), então a interpretação é de que os resultados do modelo não oferecem evidência significativa de um efeito positivo do rompimento na taxa de reprovação dos anos finais do ensino fundamental. Se, por outro lado, os exercícios de placebo mostram que o *gap* estimado para Mariana (MG) é excepcionalmente elevado relativamente aos demais municípios não afetados pelo tratamento, então os resultados encontrados para Mariana (MG) são considerados estatisticamente significantes.

A Figura 14 apresenta os resultados do teste de placebo. As linhas em cinza claro representam os *gaps* estimados para cada um dos 167 placebos. Isto é, a diferença entre as taxas de reprovação nos anos finais do ensino fundamental dos municípios e dos seus respectivos controles sintéticos - os quais não contêm Mariana (MG) nem qualquer outro município das áreas consideradas afetadas pela COPPETEC (2016).⁴⁷ A linha preta mostra novamente os *gaps* de Mariana (MG).

Tem-se que o MSPE de Mariana (MG) é de apenas 0,0024, enquanto a média dos demais municípios da amostra é 1,383. De fato, a taxa de reprovação nos anos finais do ensino fundamental antes do rompimento da barragem de Fundão, em 2015, não é bem reproduzida a partir de uma combinação convexa dos demais municípios para muitos dos municípios da Bacia do Rio Doce. Como os municípios com baixo *fit* pré-tratamento não fornecem informação adequada para a mensuração dos efeitos relativos sobre Mariana (MG), faz-se necessário aplicar recortes na amostra de 167 municípios excluindo aqueles que apresentam um MPSE acima de determinado *threshold* pré-tratamento.

Figura 14 – Gaps na Taxa de Reprovação Anos Finais EF de Mariana (MG) vs. Gaps Placebo nos 167 municípios da Bacia do Rio Doce - Todos os *lags* da *outcome variable*

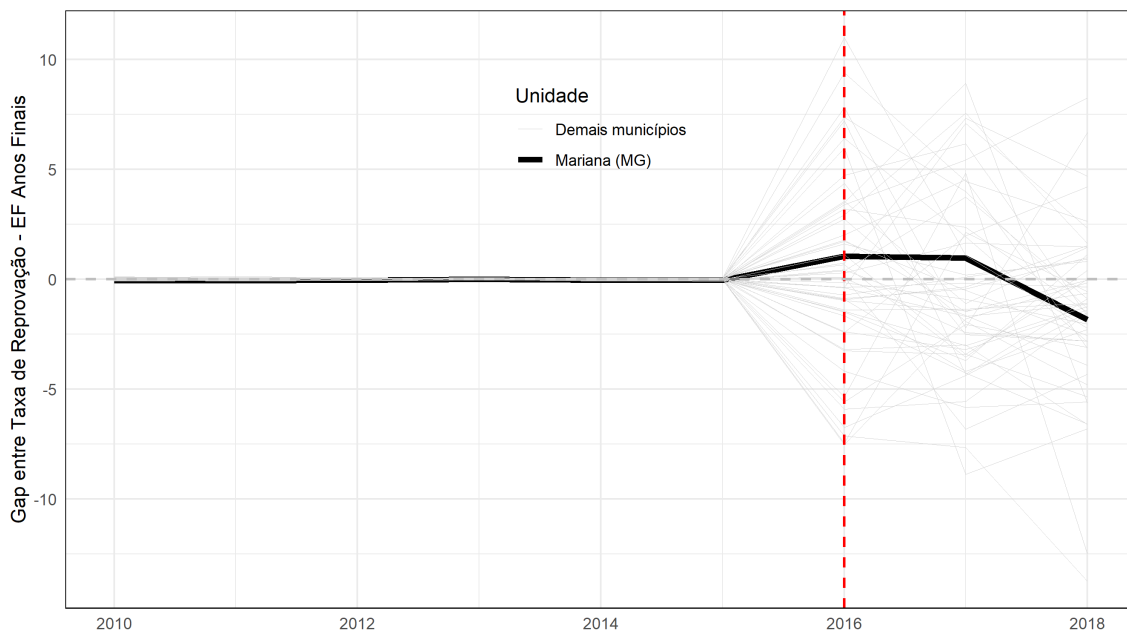


O ano de 2015 é considerado como o último ano pré-tratamento. Elaboração própria.

⁴⁷Incluindo a área 5 dos municípios indiretamente afetados pelo desastre.

A Figura 15 apresenta os mesmos resultados, mas excluindo os 115 municípios cujos testes apresentaram MSPE maior que 2 vezes o MSPE de Mariana (MG). Nota-se que o *gap* de Mariana (MG) não constitui uma das linhas menos usuais da amostra no período pós-tratamento.

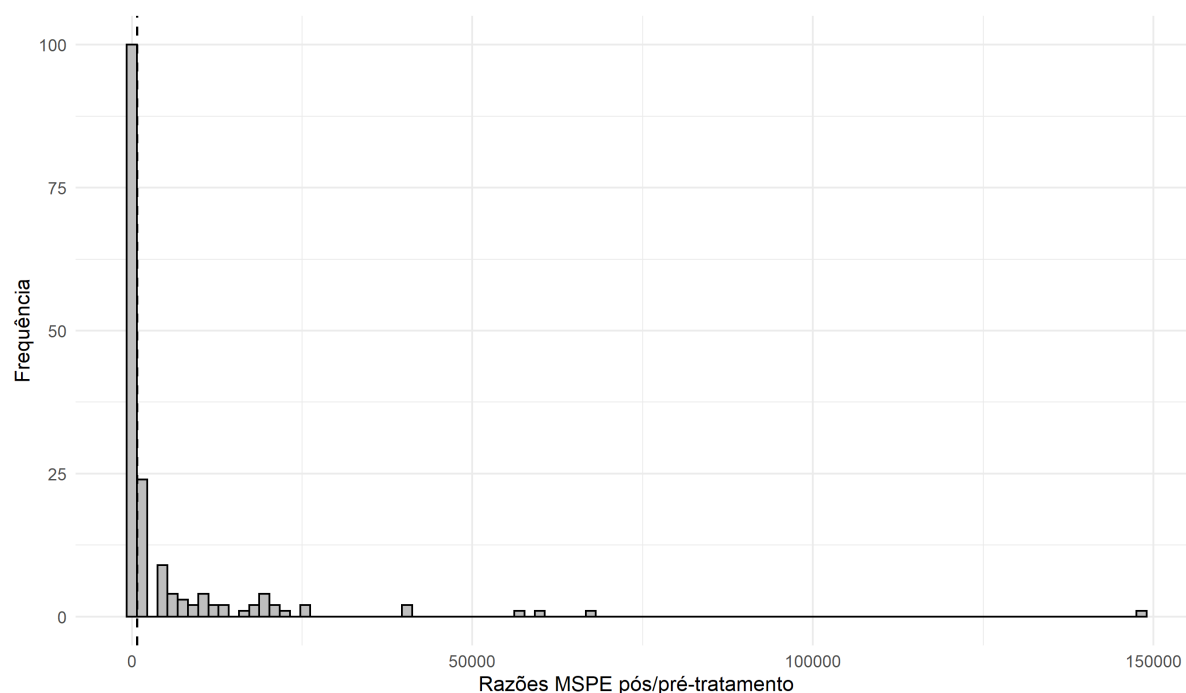
Figura 15 – Gaps na Taxa de Reprovação Anos Finais EF de Mariana (MG) vs. Gaps Placebo em 52 municípios da Bacia do Rio Doce - Todos os *lags* da *outcome variable*



O ano de 2015 é considerado como o último ano pré-tratamento. Considera-se o *threshold* para exclusão dos municípios MSPE pré-tratamento acima de 2 vezes o de Mariana (MG). Elaboração própria.

A Figura 16 reforça estes resultados ao indicar que parte importante dos municípios utilizados como placebos apresentam uma razão entre MSPE pós-tratamento e MSPE pré-tratamento maior que Mariana (MG). Ou seja, se o tratamento fosse atribuído aleatoriamente na amostra, a probabilidade de se obter uma razão maior que de Mariana (MG) seria de $67/168 = 39,8\%$, o que pode comprometer a robustez dos resultados anteriores.

Figura 16 – Distribuição da razão MSPE pós-tratamento/MSPE pré-tratamento - Todos os lags da *outcome variable*



Elaboração própria.

Em suma, os testes de robustez implementados para o método de CS indicam que não há mudanças relevantes nos resultados quando se restringe o *donor pool* para evitar eventuais efeitos de *spillover*. Com relação ao efeito antecipação, os resultados levantam dúvidas sobre o uso de 2015 como um ano pré-tratamento. Por fim, com relação à significância dos resultados avaliada pelo teste de placebo, nota-se que, com uma frequência relativamente alta se observa resultados de magnitude maior em outros municípios escolhidos aleatoriamente em relação àquela estimada para Mariana (MG).

Dessa forma, conclui-se que os resultados obtidos via método de CS para o município de Mariana (MG) sugerem uma alta na taxa média de reprovação nos anos finais do ensino fundamental, em 2016 e 2017, mas tais resultados devem ser vistos com cautela.

7 Conclusão

Este trabalho teve como objetivo verificar, por meio da abordagem de Diferenças-em-Diferenças (DD) e de Controle Sintético (CS), se há uma relação causal entre o rompimento da barragem de Fundão em Mariana (MG), ocorrido em novembro de 2015, e o desempenho de indicadores educacionais coletados das bases do Censo Escolar e do Saeb.

A pesquisa serve de contribuição para uma literatura ainda escassa no que diz respeito a efeitos de desastres tecnológicos, em particular na área de Educação, uma vez que estende a abordagem dos estudos existentes em diferentes sentidos. Primeiro, no âmbito do método de DD, são utilizados microdados da base do Censo Escolar, possibilitando o rastreamento dos alunos ao longo do tempo. Tais informações inéditas a nível individual permitem não só um aprofundamento da análise dos efeitos do desastre em Mariana (MG) como também variações nos grupos de controle a nível de escola e município. Escolas e municípios que receberam, após o desastre, alunos considerados tratados, foram excluídos da amostra de modo a minimizar eventuais efeitos de *spillover*.

Segundo, ainda com relação ao método de DD, este trabalho estende a análise realizada em FGV (2019) ao analisar a heterogeneidade regional dos efeitos do desastre em Mariana (MG). Tal análise é realizada a partir da categorização proposta pela COPPETEC (2016), também utilizada por Matsunaga (2020), que discrimina os municípios da Bacia do Rio Doce de acordo com sua exposição à lama liberada pela barragem de Fundão em Mariana (MG).

Terceiro, a pesquisa é também inovadora na medida em que procura mapear os canais de influência negativa de desastres na Educação, quais sejam, de acordo com a literatura acadêmica apresentada: (i) danos à infraestrutura física, com efeitos negativos sobre taxas de matrícula e frequência escolar; (ii) choques negativos de renda, ocasionando muitas vezes subnutrição, doenças e aumento da participação de crianças na força de trabalho; e (iii) efeitos derivados de transtornos de estresse pós-traumático (TEPT). Os resultados do método de DD indicam que o canal de condições de infraestrutura das escolas, seja física ou tecnológica, merece mais atenção em futuros estudos. O mapeamento do efeito de intervenções positivas e medidas compensatórias, embora mais difícil de ser identificado isoladamente, também constitui possível extensão a ser implementada nesse sentido.

Por fim, este trabalho também contribui com a literatura crescente de CS. Aplica-se o método desenvolvido por Abadie e Gardeazabal (2003) a nível municipal considerando como unidade tratada apenas o município de Mariana (MG). Conforme identificado no

método de DD, tal análise é particularmente importante na medida em que os municípios da área 1 da COPPETEC - municípios atingidos pela lama até UHE Candonga, em especial Mariana (MG) -, apresentaram comportamento muito distinto dos demais municípios da Bacia do Rio Doce em termos de desempenho dos indicadores educacionais.

De maneira geral, os resultados agregados a nível municipal encontrados por meio do método DD assemelham-se aos da literatura disponível (FGV, 2019). De fato, quando se considera o conjunto de 46 municípios tratados contra os de controle, não parece haver efeitos estatisticamente significativos do desastre na Educação. As variáveis dependentes analisadas foram taxa de reprovação nos anos finais do ensino fundamental, bem como desempenho médio em Matemática e Português no 5º ano do ensino fundamental. Ao nível de escola as estimações agregadas apresentam resultados similares. No caso da taxa de reprovação, no entanto, a análise de heterogeneidade indica que pode ter havido efeitos distintos entre escolas rurais e urbanas. Não se observa efeitos nem nas escolas privadas nem nas públicas.

Por sua vez, os resultados desagregados a partir da categorização COPPETEC apontam no sentido de que o desastre de Mariana (MG) pode ter gerado efeitos regionais heterogêneos. As análises ao nível de município e escola, no entanto, apresentam alguns resultados distintos. Para a taxa de reprovação nos anos finais do ensino fundamental, ao nível de escola, verifica-se efeito positivo na área 3, que é composta pelos municípios afetados pela lama, mas sem ocupação próxima ao Rio Doce. Tal resultado é robusto para a adição dos *modifiers*, covariadas afetadas pelo desastre. Já ao nível de município os resultados para esta variável dependente apontam para efeitos negativos na área 2.

Com relação ao desempenho em Matemática, a análise desagregada ao nível de escola indica possível efeito negativo na área 1. Já ao nível de município, nota-se efeito positivo na área 3. Para o caso do desempenho em Português, não se observa efeitos em nenhuma das áreas nas especificações a nível de escola. Já ao nível de município os resultados apontam novamente um efeito positivo para a área 3.

Com relação às análises dos efeitos do desastre no nível aluno, nota-se que os resultados encontrados nos exercícios de *Event Study* não corroboram com a hipótese de tendências paralelas, fundamental para a validade do método de DD. Nenhuma das variações de grupo de controle testadas obteve resultado satisfatório. Por conta disso, os resultados foram apresentados apenas no Apêndice deste trabalho.

Na ausência de resultados robustos via método de DD, em particular para a área 1 da COPPETEC, que contempla Mariana (MG), adotou-se a abordagem alternativa via

CS. Optou-se por analisar especificamente o caso de Mariana (MG), onde ocorreu o rompimento da barragem de Fundão. Dadas as restrições dos dados do Saeb, a aplicação do CS foca nos dados do Censo Escolar, em particular, na taxa de reprovação dos anos finais do ensino fundamental.

Os resultados sugerem que pode ter havido um aumento da taxa de reprovação em Mariana (MG) associado ao desastre, em especial nos anos de 2016 e 2017. Mas tais resultados devem ser vistos com cautela, uma vez que não satisfazem todos os testes de robustez comumente implementados na literatura. Considera-se a especificação sugerida por Ferman, Pinto e Possebom (2020), de modo a evitar “*cherry picking*”. Os resultados indicam que o efeito não ultrapassa 1 ponto percentual nos dois primeiros anos e é negativo em 2018. O teste de placebo para esta especificação, no entanto, apresentou resultados desfavoráveis à significância das estimações.

Referências

- ABADIE, A.; DIAMOND A.; HAINMUELLER, J. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105 (490): 493-505, 2010.
- ABADIE, A.; DIAMOND A.; HAINMUELLER, J. Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, forthcoming, 2014.
- ABADIE, A.; GARDEAZABAL, J. The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1), 113-132, 2003.
- ADHIKARI, B.; ALM, T. Evaluating the Economic Effects of Flat Tax Reforms Using Synthetic Control Methods. *Southern Economic Journal*. DOI: 10.1002/soej.12152, 2016.
- ALMOND, D.; CURRIE, J. Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3), 153-72, 2011.
- ANA. Encarte Especial sobre a Bacia do Rio Doce - Rompimento da Barragem em Mariana-MG. *Conjuntura dos Recursos Hídricos no Brasil. Informe 2015*. Agência Nacional de Águas - ANA, 2016.
- ANGRIST, J. D.; PISCHKE, J-S. *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton: Princeton University Press, 2009.
- BAEZ, J. E.; SANTOS, I. V. Children's vulnerability to weather shocks: A natural disaster as a natural experiment. *Social Science Research Network*, New York, 2007.
- BAGGERLY, J.; FERRETTI, L. K. The impact of the 2004 hurricanes on Florida Comprehensive Assessment Test Scores: Implications for school counselors. *Professional School Counseling*, 12(1), 2156759X0801200101, 2008.
- BARRERA-VALENCIA, M. et al. Cognitive profiles of Post-traumatic Stress Disorder and depression in children and adolescents. *International Journal of Clinical and Health Psychology*, 17(3), 242-250, 2017.
- BREWER, L. et al. Conflict, Earthquakes, and School Outcomes: Two Studies on Nepal, 2017. Disponível em: <<http://hdl.handle.net/1969.1/159135>>
- CAVALLO, E.; GALIANI S.; NOY I.; PANTANO, J. Catastrophic Natural Disasters and

- Economic Growth. *The Review of Economics and Statistics*, 95 (5): 1549-1561, 2013.
- COPELAND, W. E. et al. Traumatic events and posttraumatic stress in childhood. *Archives of General Psychiatry*, 64(5), 577-584, 2007.
- COPPETEC. Avaliação dos impactos da ruptura da barragem de rejeitos de Fundão em Mariana nove meses após o desastre. Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro-RJ, 2016.
- CURRIE, J. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, 47(1), 87-122, 2009.
- DI CASTELNUOVO, A. et al. Metabolic syndrome in survivors from the 2009 earthquake in Italy. *Nutrition, Metabolism and Cardiovascular Diseases*, 23(1), e5-e8, 2013.
- DI PIETRO, G. The academic impact of natural disasters: evidence from L'Aquila earthquake. *Education Economics*, 26(1), 62-77, 2018.
- FERMAN, B.; PINTO, C.; e POSSEBOM, V.. Cherry Picking with Synthetic Controls. *Methods for Policy Analysis*. Association for Public Policy Analysis and Management, 2020.
- FERNANDES, G. W. et al. Deep into the mud: ecological and socio-economic impacts of the dam breach in Mariana, Brazil. *Natureza & Conservação*, 14(2), 35-45, 2016.
- FGV. Avaliação dos impactos e valoração dos danos socioeconômicos causados para as comunidades atingidas pelo rompimento da barragem de fundão: Impactos sobre Educação Básica a partir de Dados Secundários. Fundação Getúlio Vargas. Rio de Janeiro/São Paulo. 165 p, 2019.
- GIBBS, L. et al. Delayed disaster impacts on academic performance of primary school children. *Child Development*, 90(4), 1402-1412, 2019.
- HOVEN, C. W., et al. Psychopathology among New York City public school children 6 months after September 11. *Archives of General Psychiatry*, 62(5), 545-551, 2005.
- IBAMA. Laudo Técnico Preliminar - Impactos ambientais decorrentes do desastre envolvendo o rompimento da barragem de Fundão, em Mariana, Minas Gerais. Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis - Ibama. Diretoria de Proteção Ambiental - DIPRO, 2015.

JENSEN, R. Agricultural volatility and investments in children. *American Economic Review*, 90(2), 399-404, 2000.

KHANDKER, S. et al. *Handbook on impact evaluation: quantitative methods and practices*. The World Bank, 2010.

KOUSKY, C. Impacts of natural disasters on children. *The Future of children*, 73-92, 2016.

MATSUNAGA, L. *Da Lama ao Caos: Impactos do Desastre de Mariana (MG) na Saúde Mental de Municípios da Bacia do Rio Doce*. Dissertação de Mestrado. Faculdade de Economia, Administração e Contabilidade (FEA) da Universidade de São Paulo (USP), 2020.

MONTALVO, Jose G. Voting After Bombings: A Natural Experiment on the Effect of Terrorist Attacks on Democratic Elections. *Review of Economics and Statistics*, 2011, 93 (4): 1146-1154, 2011.

NISHIJIMA, M.; ROCHA, F. F. An economic investigation of the dengue incidence as a result of a tailings dam accident in Brazil. *Journal of Environmental Management*, 253, 109748, 2020.

PANE, J. F. et al. Effects of student displacement in Louisiana during the first academic year after the hurricanes of 2005. *Journal of Education for Students Placed at Risk*, 13(2-3), 168-211, 2008.

PÉREZ-PEREIRA, M. et al. Academic effects of the Prestige oil spill disaster. *The Spanish Journal of Psychology*, 15(3), 1055-1068, 2012.

PRISMMA (2018). *Pesquisa sobre a saúde mental das famílias atingidas pelo rompimento da barragem de Fundão em Mariana*. Maila de Castro Lourenço das Neves et al. organizadores. – Belo Horizonte: Corpus.

RAMBOLL. *Relatório consolidado referente aos trabalhos dos primeiros nove meses de avaliação dos programas socioeconômicos e socioambientais*. Ministério Público Federal - MPF. Dezembro, 2017.

SACERDOTE, B. When the saints go marching out: Long-term outcomes for student evacuees from Hurricanes Katrina and Rita. *American Economic Journal: Applied Economics*, 4(1), 109-35, 2012.

SANT'ANNA, P. H. C.; ZHAO, J. Doubly robust difference-in-differences estimators. *Journal of Econometrics*. Volume 219, Issue 1, November, Pages 101-122, 2020.

SIMONATO, T. C. et al. Projeção dos impactos econômicos regionais do desastre de Mariana-MG (No. 586). Cedeplar, Universidade Federal de Minas Gerais (UFMG), 2018.

SMILDE-VAN, D. A. den-D.; SMIT, C.; WOLLESWINKEL-VAN den Bosch, J. H. School performance and social-emotional behavior of primary school children before and after a disaster. *Pediatrics*, 118(5), e1311-e1320, 2006.

VICTORA, C. G. et al. Maternal and child undernutrition: consequences for adult health and human capital. *The Lancet*, 371(9609), 340-357, 2008.

Apêndice

Tabela 8 – Categorização COPPETEC - Lista de Municípios

Área 1	Área 2	Área 3	Área 4	Área 5
Barra Longa	Alpercata	Marilândia	Dionísio	Anchieta
Mariana	Aimorés	Itueta	Córrego Novo	Catas Altas
Rio Doce	Baixo Guandu	Marliéria	Raul Soares	Congonhas
Santa Cruz do Escalvado	Belo Oriente	Naque	São Domingos da Prata	Guarapari
	Bom Jesus do Galho	Periquito	Timóteo	Ouro Preto
	Bugre	Pingo-d'Água		Santa Bárbara
	Caratinga	Resplendor		
	Colatina	Rio Casca		
	Conselheiro Pena	Santana do Paraíso		
	Fernandes Tourinho	São José do Goiabal		
	Galileia	São Pedro dos Ferros		
	Governador Valadares	Sem-Peixe		
	Iapu	Sobralia		
	Ipaba	Tumiritinga		
	Ipatinga			
4 municípios	29 municípios	5 municípios	2 municípios	6 municípios

Figura 17 – Escola Municipal de Bento Rodrigues, subdistrito de Mariana (MG), após o desastre

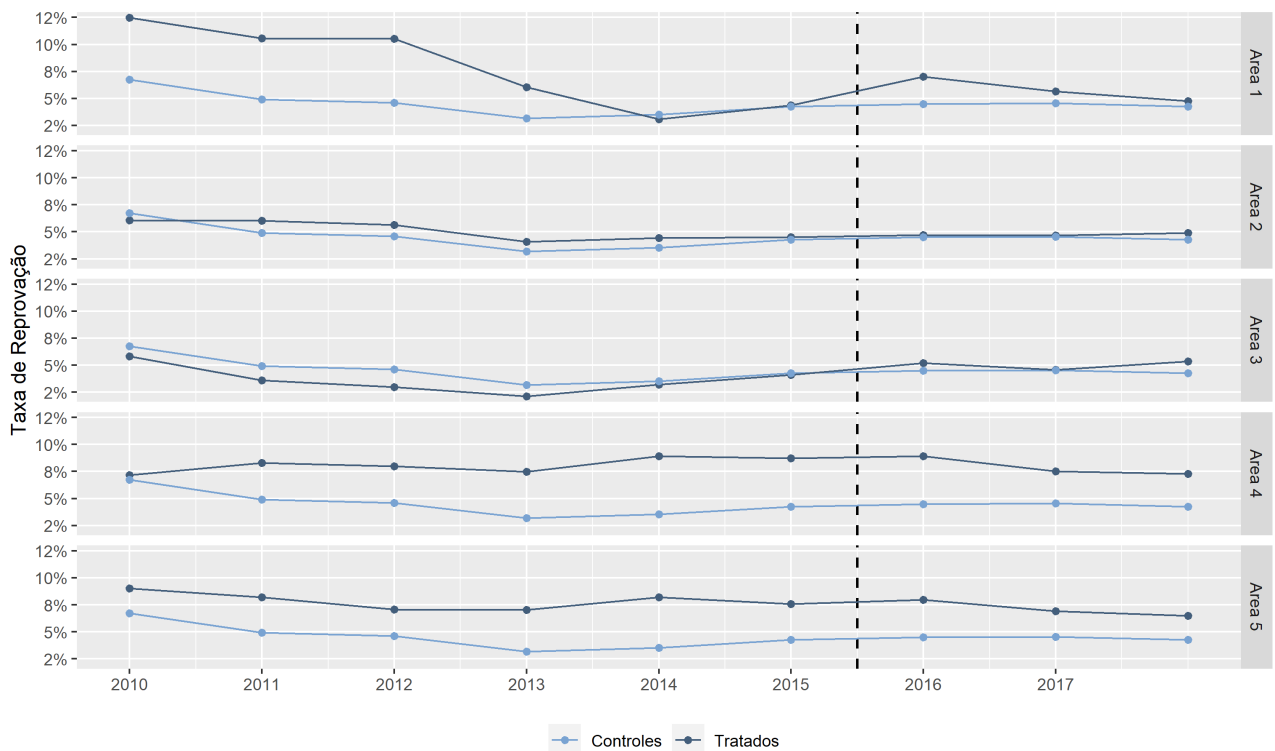


Fonte: Rogério Alvez - TV Senado.

Tabela 9 – Teste de Médias: Tratados vs. Controles - Alunos - Pré-desastre

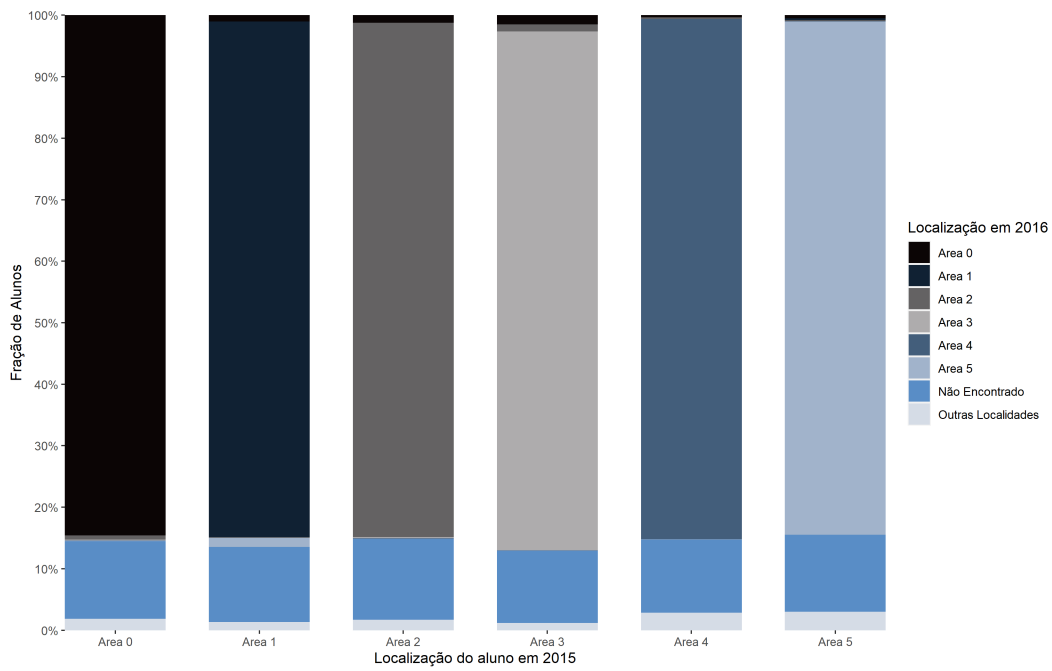
	Controles		Tratados		P-Valor
	Média	SD	Média	SD	
Aprovação	0.90	0.30	0.89	0.31	< 0.001
Reprovação	0.06	0.24	0.07	0.25	< 0.001
Abandono	0.01	0.09	0.01	0.08	< 0.001
Idade	11.36	2.82	11.29	2.79	< 0.001
Fração de Alunos Homens	0.51	0.50	0.50	0.50	0.285
Número de Alunos na Turma	225.66	2866.40	218.84	2793.18	0.129
Tecnologia	0.83	0.24	0.89	0.17	< 0.001
Infraestrutura	0.56	0.22	0.60	0.22	< 0.001
PIB per capita	17601.14	30442.74	33895.54	33072.06	< 0.001
CFEM per capita	151.16	859.51	162.68	497.89	< 0.001
ICMS per capita	432.27	963.63	671.13	972.15	< 0.001
Fração de Alunos em Escola Privada	0.05	0.22	0.11	0.32	< 0.001
Fração de Alunos em MG	0.92	0.28	0.71	0.45	< 0.001
Fração de Alunos Brancos	0.29	0.45	0.26	0.44	< 0.001
Fração de Alunos Pretos	0.06	0.23	0.04	0.20	< 0.001
Fração de Alunos Pardos	0.51	0.50	0.48	0.50	< 0.001

Figura 18 – Taxas de Reprovação - Escolas - Categorização COPPETEC



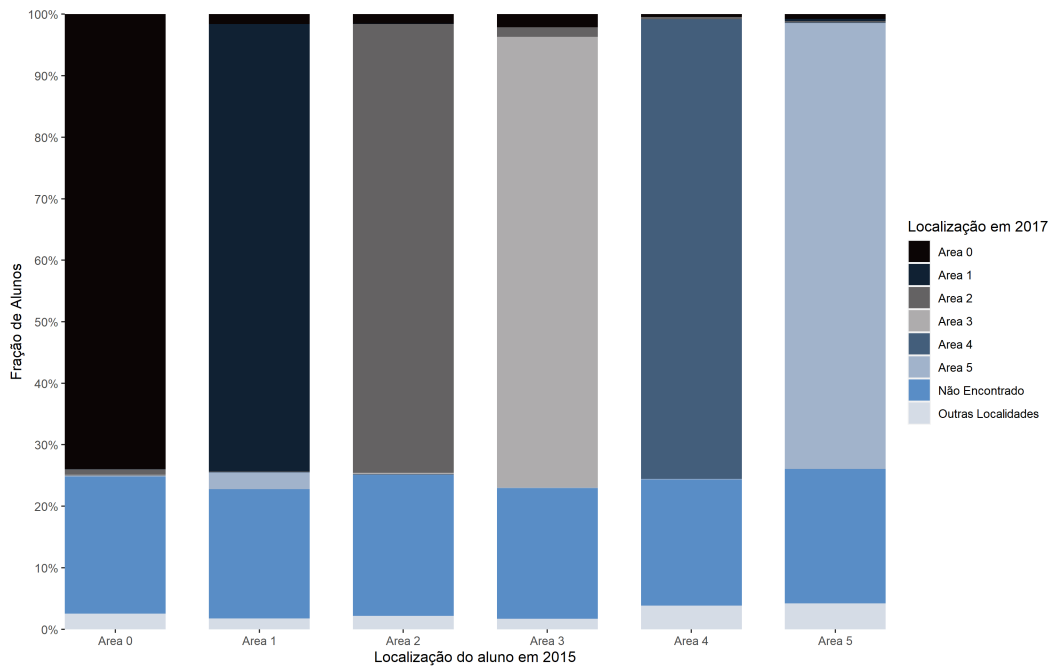
Fonte: Censo Escolar. Elaboração Própria.

Figura 19 – Movimentação dos alunos entre 2015 e 2016



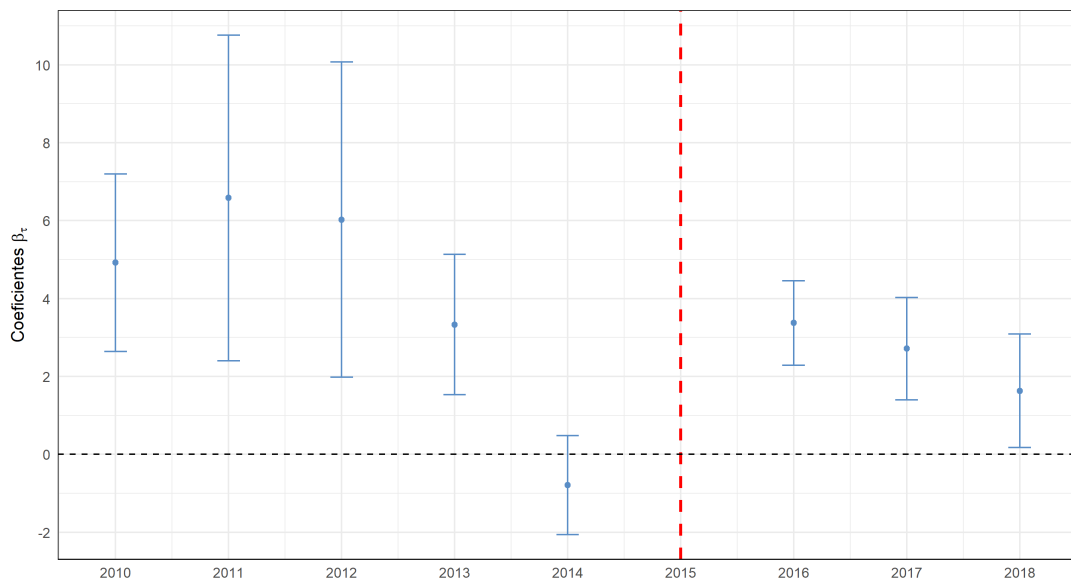
Fonte: Censo Escolar. Elaboração Própria.

Figura 20 – Movimentação dos alunos entre 2015 e 2017



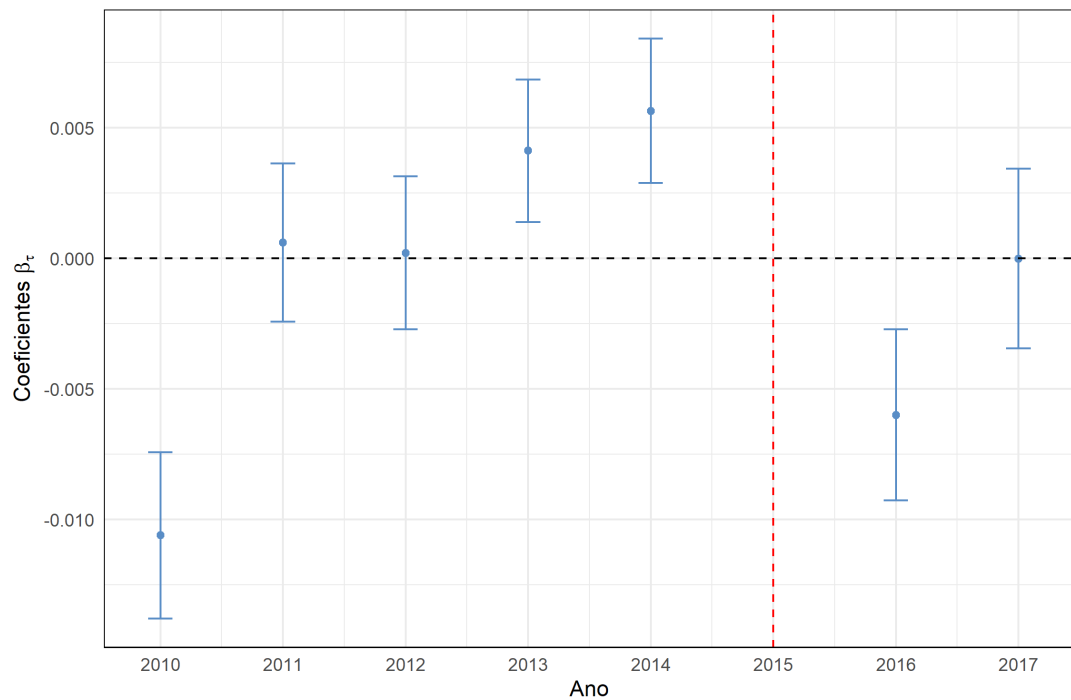
Fonte: Censo Escolar. Elaboração Própria.

Figura 21 – *Event Study* - Taxa de Reprovação Anos Finais EF - Escolas - Área 1



Fonte: Censo Escolar. Especificação desagregada: considera o conjunto de escolas tratadas da área COPPETEC 1 contra o conjunto de escolas de controle. Elaboração Própria.

Figura 22 – *Event Study* - Probabilidade de Reprovação Anos Finais EF - Alunos



Fonte: Censo Escolar. Dados obtidos via LAI junto ao Inep. Elaboração Própria.

Tabela 10 – Taxa de Reprovação Anos Finais EF - Municípios

	(1)	(2)	(3)	(4)
Pós × Tratado	-0.58 (0.32)	-0.76* (0.35)		
PIB per capita		-0.00 (0.00)		-0.00 (0.00)
Pós × Tratado 1			0.13 (1.44)	-0.07 (1.51)
Pós × Tratado 2			-0.80* (0.35)	-1.09** (0.39)
Pós × Tratado 3			0.54 (0.42)	0.60 (0.51)
Pós × Tratado 4			-0.99 (0.81)	-0.93 (0.68)
Pós × Tratado 5			-0.84 (0.98)	-0.68 (0.99)
Constante	9.21*** (0.25)	9.00*** (0.26)	9.18*** (0.25)	8.99*** (0.27)
Observações	1899	1704	1899	1704
R^2 Ajustado	0.204	0.220	0.204	0.220
Modifiers		✓		✓

Erros padrão robusto em parênteses, com cluster por município.

Colunas (1) e (3) contêm como controle tendências lineares de PIB, ICMS e CFEM per capita.

Todas as regressões contêm efeitos fixos de município e de tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tabela 11 – Desempenho Matemática 5º Ano do EF - Municípios

	(1)	(2)	(3)	(4)
Pós × Tratado	-0.13 (1.35)	-0.41 (1.39)		
PIB per capita		-0.00 (0.00)		-0.00 (0.00)
Pós × Tratado 1			-4.89 (2.50)	-5.42* (2.40)
Pós × Tratado 2			-0.92 (1.62)	-1.25 (1.59)
Pós × Tratado 3			3.10* (1.23)	2.82* (1.28)
Pós × Tratado 4			1.80 (1.03)	0.85 (1.00)
Pós × Tratado 5			3.44 (4.09)	3.94 (4.36)
Constante	221.55*** (0.73)	221.44*** (0.76)	221.64*** (0.73)	221.36*** (0.76)
Observações	823	831	823	831
R^2 Ajustado	0.151	0.150	0.150	0.150
Modifiers		✓		✓

Erros padrão robusto em parênteses, com cluster por município.

Colunas (1) e (3) contêm como controle tendências lineares de PIB, ICMS e CFEM per capita.

Todas as regressões contêm efeitos fixos de município e de tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tabela 12 – Desempenho Português 5^o Ano do EF - Municípios

	(1)	(2)	(3)	(4)
Pós × Tratado	1.14 (1.62)	1.14 (1.63)		
PIB per capita		0.00 (0.00)		0.00 (0.00)
Pós × Tratado 1			6.26 (3.54)	5.72 (4.12)
Pós × Tratado 2			-1.64 (1.90)	-1.54 (1.90)
Pós × Tratado 3			8.64** (3.17)	8.34* (3.26)
Pós × Tratado 4			1.11 (2.12)	0.77 (2.16)
Pós × Tratado 5			4.78 (4.30)	5.61 (4.15)
Constante	198.61*** (0.65)	198.24*** (0.77)	198.64*** (0.65)	198.06*** (0.78)
Observações	823	831	823	831
R^2 Ajustado	0.555	0.556	0.558	0.559
Modifiers		✓		✓

Erros padrão robusto em parênteses, com cluster por município.

Colunas (1) e (3) contêm como controle tendências lineares de PIB, ICMS e CFEM per capita.

Todas as regressões contêm efeitos fixos de município e de tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tabela 13 – Probabilidade de Reprovação - Alunos

	(1)	(2)	(3)	(4)
Idade	-0.0047** (0.0017)	-0.0047** (0.0017)	-0.0047** (0.0017)	-0.0048** (0.0017)
Infraestrutura		0.0065** (0.0023)		0.0065** (0.0023)
Tecnologia		0.0007 (0.0023)		0.0008 (0.0023)
Tratado × Pós	-0.0050*** (0.0013)	-0.0062*** (0.0012)		
Tratado 1 × Pós			0.0087 (0.0050)	0.0001 (0.0050)
Tratado 2 × Pós			-0.0070*** (0.0014)	-0.0079*** (0.0014)
Tratado 3 × Pós			0.0012 (0.0033)	0.0008 (0.0033)
Tratado 4 × Pós			0.0007 (0.0029)	0.0005 (0.0029)
Tratado 5 × Pós			-0.0081** (0.0027)	-0.0102*** (0.0027)
Observações	1596428	1585772	1596428	1585772
R^2 Ajustado	0.072	0.072	0.072	0.072
Modifiers		✓		✓

Erros padrão robustos entre parênteses, com cluster de indivíduo.

Incluídos em (1) e (3) tendência linear de PIB, ICMS e CFEM per capita.

Todas as regressões contem efeitos fixos de individuo, escola e tempo.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Tabela 14 – Análise de Heterogeneidade - Probabilidade de Reprovação - Alunos

	(1)	(2)	(3)
Tratado × Pública × Pós	0.0023 (0.0013)		
Tratado × Privada × Pós	-0.0623*** (0.0018)		
Tratado × Homem × Pós		0.0232*** (0.0016)	
Tratado × Mulher × Pós		-0.0339*** (0.0014)	
Tratado × Brancos × Pós			-0.0265*** (0.0017)
Tratado × Demais × Pós			0.0028* (0.0014)
Observações	1596428	1596428	1596428
R^2 Ajustado	0.072	0.073	0.072
Modifiers			

Erros padrão robustos entre parênteses, com cluster de indivíduo.

Todas as regressões contêm efeitos fixos de indivíduo, escola e tempo além de tendências lineares de PIB, ICMS e CFEM per capita.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figura 23 – Gaps na Taxa de Reprovação Anos Finais EF - Mariana (MG) vs. Mariana (MG) sintética - Variação no *donor pool*: *threshold* de 2,5% - Todos os *lags* da *outcome variable*



O ano de 2015 é considerado como o último ano pré-tratamento. Elaboração Própria.