



Universidade de Brasília  
Faculdade de Economia, Administração, Contabilidade e  
Gestão de Políticas Públicas  
Programa de Pós Graduação em Economia

**Desastre de Mariana e Resultados Educacionais:  
Um Estudo de Evento**

Vinícius de Oliveira Alves

Brasília - DF

2022

Vinícius de Oliveira Alves

**Desastre de Mariana e Resultados Educacionais:  
Um Estudo de Evento**

Dissertação apresentada ao Programa de  
Pós-Graduação em Economia da  
Universidade de Brasília como requisito  
para obtenção do título de Mestre

Orientador: Dr. Rafael Terra de Menezes

Brasília - DF

2022

Vinícius de Oliveira Alves

**Desastre de Mariana e Resultados Educacionais:  
Um Estudo de Evento**

Dissertação apresentada ao Programa de Pós-Graduação em Economia da Universidade de Brasília como requisito para obtenção do título de Mestre

Orientador: Dr. Rafael Terra de Menezes

Trabalho aprovado em 24 de junho de 2022.

---

Dr. Felipe dos Santos Martins  
(Membro Externo)  
(Ipea)

---

Dr. Vander Mendes Lucas  
(Membro Interno)  
(UnB)

---

Dr. Rafael Terra de Menezes  
(Orientador)  
(UnB)

Brasília - DF  
2022

## Resumo

Este trabalho busca realizar uma análise dos impactos do Desastre de Mariana, onde em 2015 houve rompimento da Barragem de Fundão, sobre resultados educacionais. Para tanto, utilizou-se um método quase experimental de inferência causal, o controle sintético. Os resultados indicam que, quando análise se restringe a Mariana como unidade tratada, o efeito estimado só é estatisticamente significativo em 3 das 15 medidas de desempenho escolar, com destaque para a taxa de distorção-idade-série (TDI) do 1º ano do ensino fundamental, que apresentou um aumento transitório de quase 30 pontos percentuais em Mariana relativamente ao controle sintético. Em contrapartida, quando a análise centrou-se no conjunto de municípios mais afetados, o cenário inverteu-se, pois houve significância na maioria das variáveis estudadas, nas quais os efeitos médios do tratamento, que foram aumentos nas taxas examinadas, oscilaram de 0,72 pontos percentuais (abandono no 1º ano) a 13,58 pontos percentuais (TDI no 9º ano).

**Palavras - chave:** desastre natural, educação, controle sintético

**Classificação JEL:** Q50, I20

## Abstract

This study aims to evaluate the impact of Mariana Disaster, where in 2015 there was a mining dam collapse at Barragem de Fundão, over educational outcomes. To do so, we implement a quasi experimental method of causal inference, the synthetic control. Results suggest that, when we restrict the analysis to the case where Mariana is the single treated unit, the estimated effect is statistically meaningful in only 3 out of 15 educational performance measures, the greatest of which is for age-grade distortion (overage), a transitory increase of nearly 30 percentage points in Mariana compared to the synthetic control. On the other hand, when we consider the set of directly damaged municipalities as the treated unit, the outcome changes, as the estimated effect was statistically meaningful for the majority of studied variables. In this case, the impacts, which were generally an increase in the examined rates, range from 0,72 (absenteeism in first grade) percentage points to 13,58 percentage points (overage in ninth grade)

**Keywords:** natural disaster, education, synthetic control

**JEL classification:** Q50, I20

# 1 Introdução

Desastres Naturais configuram-se como choques adversos que podem afetar o bem estar dos indivíduos em dimensões que se relacionam mutuamente, tais como saúde e acumulação de capital humano.<sup>1</sup> No caso da educação, no período recente, o interesse sobre os efeitos desses choques reaparece em razão da Pandemia de Covid 19. De fato, embora se trate de um choque recente, já existem algumas tentativas de avaliação dos impactos desse evento na formação de capital humano, por exemplo, em países da África.<sup>2</sup> No Brasil, um esforço inicial para o entendimento de como as escolas reagiram a esse evento adverso é a "Pesquisa Resposta Educacional à Pandemia de COVID-19", realizada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP).

Não obstante essa pesquisa do INEP seja importante para a obtenção de um panorama do ensino básico durante a pandemia, o efeito causal desta pode ser de difícil mensuração, visto que ela atingiu, em tese, todo o país. Ou seja, a busca por um grupo de comparação seria desafiadora. Diante desse entrave, uma forma de, pelo menos, obtermos *insights* sobre esse tema consiste em analisar os impactos de choques passados cuja área afetada seja mais restrita. Nesse contexto, neste trabalho, buscamos, por meio de um método quase experimental de inferência causal, o controle sintético, realizar uma análise dos impactos do Desastre de Mariana (MG), onde em 2015 houve rompimento da Barragem Fundão, sobre resultados educacionais.

Em 5 Novembro desse ano, houve um rompimento da referida barragem, uma tragédia em razão da qual a população de Mariana e de outros municípios próximos tem sido afetada, por exemplo, em termos de danos à infraestrutura, meio ambiente, saúde e educação. Com efeito, estudos recentes, tais como os de Mrejen *et al* (2020) e Carrilo *et al* (2020), sugerem a ocorrência de danos a saúde física das crianças nascidas logo após esse desastre. Todavia, a literatura ainda carece de evidência empírica a respeito dos impactos educacionais dessa tragédia, o que reforça a relevância da agenda de pesquisa do presente estudo. De toda forma, a princípio, segundo o relatório Grupo Força Tarefa (2016), o ensino básico foi atingido, seja pela destruição de escolas públicas de ensino fundamental, seja pela impossibilidade de deslocamento a essas instituições devido ao comprometimento da infraestrutura de acesso.<sup>3</sup>

Além disso, conquanto não diretamente impactados por danos na estrutura física das escolas, em tese, estudantes do ensino médio podem ter sido também prejudicados, sobretudo no caso de Mariana, cuja economia, conforme sugere o relatório supracitado do governo de Minas Gerais, é mais dependente da atividade mineradora em comparação aos demais municípios atingidos. Tal atividade foi inclusive suspensa por 5 anos a partir do fim de 2015.<sup>4</sup> Nesse cenário, uma possibilidade seria esse choque adverso comprometer o orçamento das famílias e, destarte, impor pressões adicionais sobre a entrada de adolescentes no mercado de trabalho e, assim, afetar os estudos desses jovens.

A assertiva anterior é, na verdade, apenas uma conjectura para motivar o porquê de se esperar efeitos deletérios desse desastre também nos estudantes do ensino médio. De toda forma, independente do nível de ensino, há razões adicionais para se esperar impactos na educação básica. Afinal, ainda segundo o relatório do governo de Minas Gerais, a saúde mental da população local foi afetada. Assim, caso esse trauma altere, por exemplo, a estabilidade emocional das crianças e adolescentes atingidos, é possível haver comprometimento de habilidades não cogni-

---

<sup>1</sup>Baez *et al* (2010)

<sup>2</sup>Agrist *et al* (2021) e Sabates *et al* (2021)

<sup>3</sup>As informações mencionadas sobre os prejuízos iniciais atrelados ao desastre de Mariana encontram-se em mais detalhes no relatório Grupo Força Tarefa (2016).

<sup>4</sup>Conforme noticiado pelo portal G1 em 2020. Esse texto está disponível neste link: <https://glo.bo/3tYwrTx>

tivas, as quais, segundo literatura recente, podem influenciar o desempenho escolar. <sup>5</sup> Ou seja, os possíveis danos a educação básica podem se manifestar também de forma indireta.

Nesse sentido, além desta introdução e de uma conclusão, este estudo estrutura-se em mais quatro seções. Na segunda, realiza-se uma revisão da literatura aplicada a respeito do impacto de desastres naturais sobre resultados educacionais; na terceira, apresenta-se dados, variáveis utilizadas e se discute o recorte espacial adotado; na quarta, explana-se a metodologia empírica implementada e, na quinta, discorre-se sobre os resultados das estimações dos modelos de controle sintético.

---

<sup>5</sup>Para mais detalhes sobre a relação entre habilidades não cognitivas e desempenho escolar, veja Santos e Primi (2014).

## 2 Revisão de Literatura

A literatura aplicada tem documentado os impactos de uma variedade de desastres naturais, tais como terremotos, ciclones e enchentes, sobre resultados educacionais no ensino básico. Os efeitos de curto prazo normalmente são mensurados sobre a taxa de crianças matriculadas e, em alguns casos, sobre o desempenho estudantil. Os impactos de tais choques sobre a taxa de matrícula são mistos. Por exemplo, na Etiópia, Índia e Peru, Nguyen e Pham (2018) não encontram impacto estatisticamente significativo da exposição a enchentes no percentual de matriculados. Já no Vietnã, esses mesmos autores encontram redução de 5% na probabilidade de uma criança afetada por esses desastres estar matriculada.

Por sua vez, na Indonésia, Rush (2018) encontra também efeitos negativos nesse sentido associados a enchentes, embora, nesse caso, a magnitude do impacto dependa da gravidade do choque em termos de danos na infraestrutura, quantidade de pessoas mortas, feridas ou desaparecidas. Por outro lado, existe evidência segundo a qual os prejuízos escolares de curto prazo atrelados a desastres naturais podem ser atenuados pela provisão de auxílio financeiro aos afetados. É o que sugere Andrabi *et al* (2021), em cujo estudo não se encontra evidência de redução na taxa de matrícula escolar para crianças residentes em regiões acometidas por um terremoto no Paquistão.

Contudo, em se tratando de desempenho escolar, esses autores encontram déficits educacionais. Em particular, independente da idade, alunos mais afetados pelo terremoto apresentaram redução de 0,24 desvios padrão em testes padronizados comparativamente a estudantes não afetados. Em geral, Andrabi *et al* (2021) reconhece que, em função da indisponibilidade de dados, é escassa a evidência empírica a respeito dos efeitos de desastres naturais sobre o desempenho estudantil. Apesar disso, Nguyen e Pham (2018) conseguiram documentar resultados, embora mistos, nesse sentido para o caso de secas e enchentes, consoante discutiremos sobre a seguir.

No Peru e na Etiópia, esses choques adversos não afetaram as habilidades cognitivas de crianças em testes de matemática e vocabulário. Já no caso da Índia e do Vietnã, os efeitos negativos oscilam numa perda de 3% a 6% na pontuação em exames de matemática e vocabulário. Ou seja, a presença ou ausência desses efeitos deletérios de curto prazo pode variar de acordo com o país analisado.

No longo prazo, a literatura sugere o seguinte: os prejuízos associados a desastres naturais podem persistir de tal forma a comprometer a acumulação de capital humano. Nesse caso, normalmente os efeitos mensurados relacionam-se ao nível de escolaridade dos afetados. Por exemplo, 20 anos após um terremoto ocorrido no Nepal, Paudel e Ryu (2018) comparam indivíduos que, na época, eram crianças de até 2 anos as quais foram, ou não, afetadas por esse choque. Como resultado, os impactados pelo terremoto, em média, completaram 0,8 séries a menos relativamente ao grupo de comparação. Mais ainda, os afetados tem probabilidade 13,8% menor de terminar o ensino fundamental e 10% menor de concluir o ensino médio.

Por sua vez, Caruso (2017), ao analisar diversas modalidades de desastres, tais como enchentes, ciclones, deslizamentos de terra e terremotos, ocorridos na América Latina ao longo do século XX, conclui que, em média, essas catástrofes afetam negativamente o nível de escolaridade dos afetados. Esse efeito seria mais severo nos primeiros anos de vida da criança. Ademais, Wang *et al* (2017) e Caruso e Miller (2015) encontram impactos nesse sentido para, respectivamente, China e Peru.

Em essência, a evidência empírica disponível sugere, para diferentes países em distintos continentes, efeitos adversos de desastres naturais sobre a acumulação de capital humano no longo prazo. Esses impactos negativos, segundo estudos recentes, podem se transmitir até mesmo

intergeracionalmente.<sup>6</sup> Ademais, tais prejuízos, na verdade, tem se mostrado heterogêneos a depender de algumas características observadas, tais como sexo e renda familiar, dos afetados pelos choques.<sup>7</sup>

Em suma, discorreremos sobre a literatura aplicada a respeito dos impactos de desastres naturais sobre variáveis educacionais, tais como taxa de matrícula de crianças, desempenho e nível de escolaridade. Nesse processo, percebemos que, no curto prazo, o aparecimento de algum efeito negativo varia de acordo com o país analisado. Já no longo prazo, para diversas localidades em distintos continentes, a literatura documenta prejuízos na acumulação de capital humano. Ademais, normalmente, esses estudos tem identificado esses desastres como um experimento natural e avaliam os possíveis efeitos usando o estimador de diferenças em diferenças. Nessa linha, este estudo se diferencia dessa literatura devido a implementação do método de controle sintético, cujos dados e variáveis utilizadas serão discutidas a seguir.

---

<sup>6</sup>Exemplos de trabalhos aplicados que discutem essa persistência intergeracional são Caruso e Miller (2015) e Caruso (2017)

<sup>7</sup>Para uma discussão desses efeitos heterogêneos, veja, por exemplo, Paudel e Ryu (2018) e Caruso e Miller (2015)

## 3 Dados

### 3.1 Fontes e variáveis utilizadas

A proposta consiste em utilizar dados do INEP referentes ao período de 2007 a 2018. Em particular, coletam-se estatísticas dos Indicadores Educacionais, Sinopses Estatísticas e Censo Escolar, a partir dos quais obtém-se algumas medidas de resultados escolares no nível municipal. Nessa linha, objetiva-se investigar o comportamento das taxas de abandono, reprovação e distorção-idade-série do 1º, 5º e 9º ano do ensino fundamental, bem como da 1ª e 3ª série do ensino médio. Essas serão as variáveis dependentes da estratégia de estimação, explanada na próxima seção. Ademais, coletam-se, para cada uma dessas etapas educacionais no intervalo de tempo supracitado, um conjunto de variáveis de controle, elencadas a seguir:

Tabela 1: Variáveis preditoras utilizadas

---

População
PIB <i>per capita</i> real
Taxa de Distorção Idade Série (TDI)
Taxa de Abandono
Taxa de Reprovação
Média Hora Aula
Média Aluno Turma
Taxa de Evasão
Taxa de Regularidade do Docente
Quantidade de Matrículas
Percentual de Docentes com Ensino Superior
Número de Escolas (por nível de ensino)
Salário Real Médio dos Docentes em 2014

---

### 3.2 Recorte espacial adotado

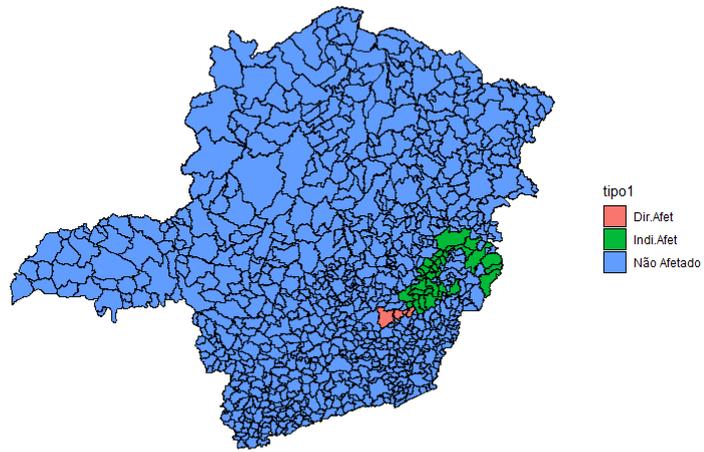
A figura 1 ilustra a geografia do desastre em análise. Segundo relatório do governo de Minas Gerais, dos 853 municípios mineiros, 35 foram de algum modo afetados pelo rompimento da Barragem de Fundão, dos quais 4, destacados em vermelho no mapa abaixo, foram impactados diretamente. Trata-se de locais atingidos efetivamente pela lama. Nesse caso, Mariana foi o município mais atingido. Além disso, outros 31 municípios, destacados em verde, foram indiretamente afetados. Isso significa que, embora não tenham sido atingidos pela lama, houve impacto, por exemplo, em termos de água contaminada. Por outro lado, os demais 818 municípios mineiros, destacados em azul, não foram impactados por essa tragédia.<sup>8</sup>

Como a estratégia de estimação inicial centra-se no impacto do desastre sobre Mariana, remove-se da análise os outros 3 municípios diretamente afetados e os 31 indiretamente atingidos. Ademais, com vistas a buscar unidades de comparação geograficamente mais próximas desse município, restringimos o recorte espacial para cidades circunscritas a um raio de 100km de Mariana.<sup>9</sup> O resultado dessa delimitação, ilustrado na figura 2, consiste em uma unidade exposta à intervenção e 102 unidades de controle potenciais.

<sup>8</sup>O relatório mencionado é o mesmo da seção 1, o Força Tarefa (2016)

<sup>9</sup>Trabalha-se com a distância em linha reta do centro geométrico, denominado centroide, de um município ao de Mariana

Figura 1: Regiões Afetadas em Minas Gerais

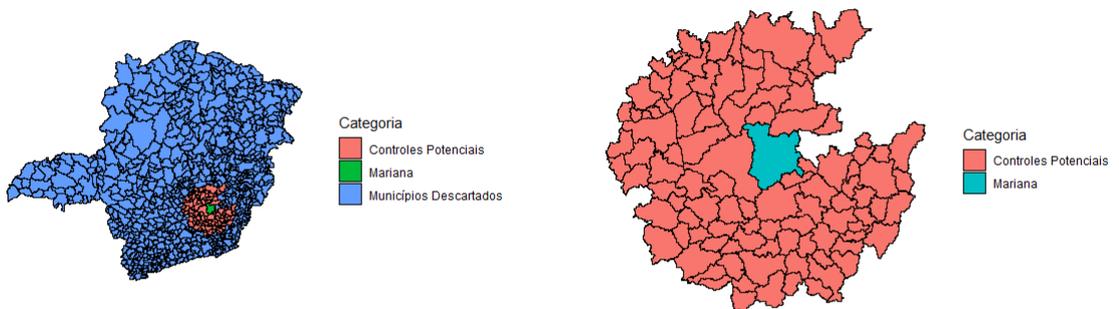


Fonte: Elaboração própria com base em dados territoriais do IBGE sobre a malha municipal de Minas Gerais (2020)

Figura 2: Recorte Espacial adotado: raio de 100km de Mariana

(a)

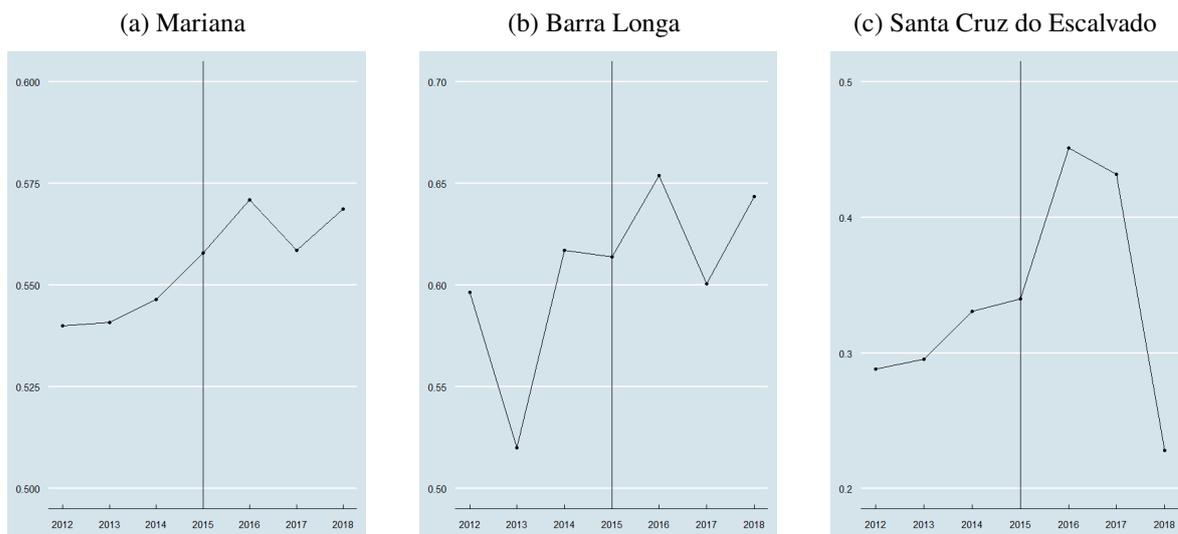
(b)



Fonte: Elaboração própria com base em dados territoriais do IBGE sobre a malha municipal de Minas Gerais (2020)

Embora Mariana tenha sido o município mais afetado, houve, vide figura 1, outros 3 municípios diretamente impactados pela lama que, portanto, são também passíveis de terem sofrido prejuízos educacionais. Nessa linha, com o intuito de motivar a inclusão dessas 3 cidades na análise, apresentamos um indicador de como a concentração da quantidade de alunos matriculados nas escolas públicas evoluiu em 3 desses 4 municípios no período de 2012 a 2018. Em particular, examinar-se-á brevemente o índice de Gini para os anos iniciais do ensino fundamental.<sup>10</sup>

Figura 3: Índice de Gini para a quantidade de matrículas nos anos iniciais do ensino fundamental



Fonte: Elaboração própria com base me dados do Censo Escolar do ensino básico

Observação: dado o número de escolas públicas no município que ofertam esse nível educacional, mede-se o grau de concentração das matrículas nas escolas

Em tese, seria de se esperar um aumento na concentração da quantidade de alunos matriculado particularmente nas escolas de Mariana, pois trata-se do único município no qual houve destruição de escolas, de modo que alguns alunos foram realocados para outras instituições de ensino, um movimento que também ocorreu devido a necessidade de encaminhar famílias desalojadas ou desabrigadas para outros locais. No entanto, entre 2015 e 2016, conforme aponta o painel (a) da figura 3, a concentração da quantidade de matrículas só subiu marginalmente, reflexo do aumento inferior a 0,02 pontos do índice de Gini nesse período. Similarmente, vide painel (b) dessa figura, em Barra Longa, onde o prejuízo educacional inicial mais expressivo foi a suspensão de aulas em algumas escolas, não houve aumento expressivo do indicador em análise no ano posterior ao desastre.<sup>11</sup>

Já em Santa Cruz do Escalvado, entre 2015 e 2016, consoante o painel (c), o índice de Gini subiu mais de 0,1 pontos e, partir do ano seguinte, volta a cair até retornar, em 2018, ao patamar pré desastre. Esse aumento de concentração transitório chama a atenção, pois esse município foi o único no qual, a princípio, segundo relatório Grupo Força Tarefa (2016), não houve comprometimento da infraestrutura física das escolas ou dos canais de acesso a elas. Nesse sentido, uma possível explicação está no fato de, segundo processo do poder judiciário

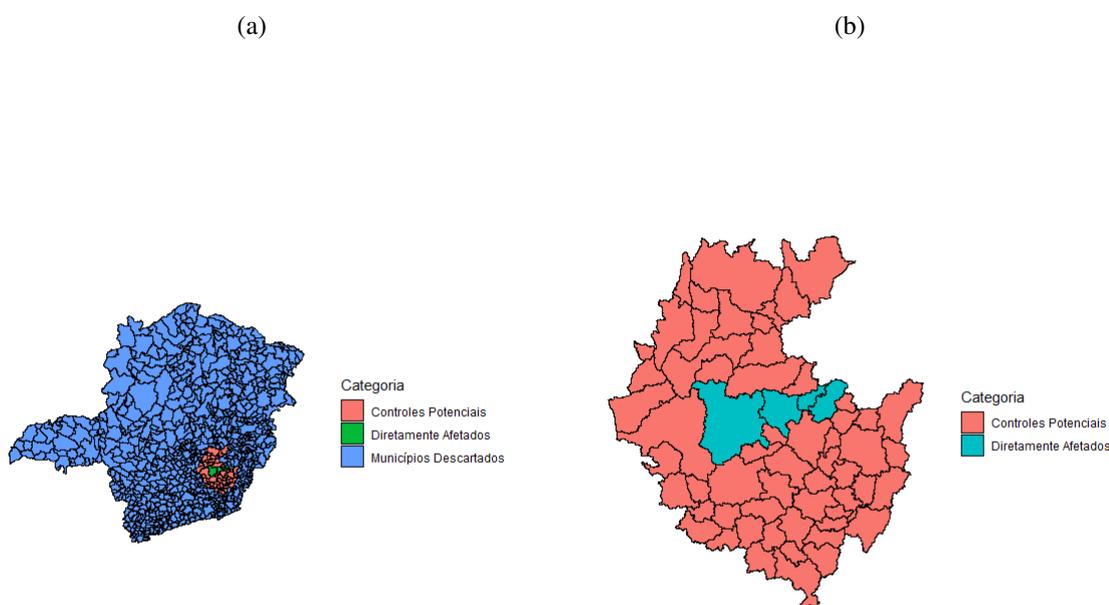
<sup>10</sup>Em Rio Doce, de acordo com dados do Censo do ensino básico, não havia quantidade suficiente de escolas com alunos nesse nível educacional para se analisar a concentração das matrículas. Por isso, só se apresentou o índice de Gini para os outros 3 municípios

<sup>11</sup>Essas informações sobre como o sistema de ensino de cada municípios foi afetado encontra no relatório Grupo Força Tarefa (2016)

de Minas Gerais, esse município ter enfrentado, a partir de 2016, um aumento populacional devido aos funcionários e respectivas famílias que migraram para a cidade para prestação do serviço de remoção de rejeitos do desastre por empresas contratadas.<sup>12</sup> Esse fluxo migratório, ainda segundo esse processo, provocou elevação do número de alunos na rede municipal de ensino. Portanto, em teoria, esse repentino crescimento populacional poderia ter ocasionado um desarranjo temporário na alocação dos alunos nas escolas públicas e, destarte, o incremento transitório na concentração das matrículas.

De toda forma, a análise acima serve apenas para ilustrar que outros municípios, além de Mariana, podem ter sido potencialmente afetados no tocante a educação básica. Nessa medida, conforme será discutido na metodologia, a segunda estratégia de estimação se propõem a avaliar o impacto do desastre no conjunto de municípios diretamente afetados, destacados em vermelho na Figura 1. Nesse caso, há também os municípios de Barra Longa, Rio Doce e Santa Cruz do Escalvado. Desse modo, assim como no caso anterior, remove-se da análise as 31 cidades não atingidas. Além disso, objetiva-se delimitar unidades de comparação geograficamente próximas desse conjunto de locais diretamente afetados, uma tarefa realizada em 2 passos. Primeiramente, para cada um desses 4 municípios, derivamos a região circunscrita a um raio de 100km. Em seguida, calculamos a intersecção entre as 4 regiões obtidas. O resultado desse exercício, ilustrado na figura 4, são 58 municípios que comporão os controles potenciais para o conjunto das 4 cidades impactadas.

Figura 4: Recorte Espacial adotado: intersecção dos raios de 100km de cada município diretamente afetado



Fonte: Elaboração própria com base em dados territoriais do IBGE sobre a malha municipal de Minas Gerais (2020)

<sup>12</sup>Trata-se do Processo n.º 5001772-35.2018.8.13.0521 do Tribunal de Justiça de Minas Gerais, disponível em <https://bitly.com/yMHgsQ>

## 4 Metodologia

### 4.1 Controle Sintético com uma unidade tratada: Mariana

O desastre de Mariana pode ser entendido como um experimento natural para o efeito de um choque adverso nos resultados escolares. Nesse caso, ao delimitarmos o recorte espacial como os municípios não afetados dentro de um raio de 100km de Mariana (veja a subseção 3.2), temos uma unidade agregada exposta ao desastre a partir de 2015 e 102 unidades de comparação potenciais, os municípios mineiros não afetados.

Nesse cenário, uma questão seria como escolher o grupo de controle ideal entre tais candidatos. A escolha de apenas um deles, numa comparação de caso, pode ser arbitrária diante de tantas possibilidades. Por outro lado, a determinação de um grupo de candidatos implica saber qual a combinação que melhor se aproxima do município de Mariana.

Diante desse problema, uma alternativa é utilizar o método de controle sintético. Este deriva uma unidade de comparação artificial, a qual consiste numa combinação linear convexa de todos os candidatos a unidade de controle. Essa ponderação seria feita com base em características observadas preditivas do resultado de interesse no período pré intervenção.

Formalmente, seguindo a abordagem de Abadie *et al* (2010), suponha que observamos  $J + 1$  unidades e denote por  $Y_{it}$  o resultado observado da variável de interesse para a unidade  $i$  no período  $t$ . Se  $i = 1$ , temos a unidade tratada (Mariana). Por outro lado, para  $i \in \{2, \dots, J + 1\}$ , temos os candidatos a unidade de comparação (municípios não afetados)

Nesse contexto, a estimação do efeito  $\hat{\alpha}_t$  dessa intervenção para cada período  $t$  a partir da ocorrência do desastre de Mariana (a partir de 2015) pressupõem a obtenção dos pesos da equação abaixo:

$$\hat{\alpha}_t = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \quad (1)$$

Em que

- $Y_{it}$  : Resultado educacional do município  $i$  no ano  $t$ ;  $i = 1, 2, \dots, J + 1$
- $w_j^* \in [0, 1]$  : peso atribuído a unidade de comparação potencial  $j$

Na prática, o vetor de pesos é obtido por um algoritmo de minimização dupla com base nos controles potenciais disponíveis e em características observadas pré intervenção que sejam preditoras do resultado de interesse antes da ocorrência do desastre natural. Nesse processo, mantendo a notação acima e supondo que se utilize  $k$  covariadas preditoras, a princípio, é necessário realizar uma combinação entre as  $k$  características da unidade exposta à intervenção (numa matriz  $X_0$  de ordem  $(k \times 1)$ ) com as características correspondentes das unidades potenciais de controle, formalizadas numa matriz  $X_1$  de ordem  $(k \times J)$ . Assim, a ideia seria escolher o vetor  $W^*$ , que atribui pesos a cada controle potencial, de modo a minimizar a distância entre  $X_1$  e  $X_0W$ . Isso significa minimizar esta equação com respeito a  $W$ :

$$\|X_1 - X_0W\| = \sqrt{(X_1 - X_0W)'V(X_1 - X_0W)} \quad (2)$$

Em que

- $V$  é definido como uma matriz  $(k \times k)$  simétrica e positiva semidefinida.

Note que, ao minimizar a equação acima,  $W^*$  fica em função de  $V$ , uma matriz introduzida para se atribuir pesos às características observadas contidas em  $X_0$  e  $X_1$ . Na abordagem de Abadie *et al* (2010),  $V^*$  é escolhida dentre todas as matrizes diagonais (e com as características indicadas acima) tais que o Erro Quadrático Médio de Previsão ( $MSPE$ ) da variável dependente  $Y$  seja minimizado ao longo do intervalo de tempo pré intervenção. Essa é a segunda otimização do algoritmo de minimização dupla.

#### 4.1.1 Inferência Estatística

Abadie *et al* (2010) propõem um método de inferência baseado em permutações por meio do qual realiza-se um teste de placebos. Antes de expô-lo, contudo, convém definir mais explicitamente o  $MSPE$ . Para isso, manteremos a notação da subseção 4.1 com estas adições: seja  $T$  o total de períodos nos quais se observam as unidades analisadas e  $T_0$  o período a partir do qual começa a exposição da unidade tratada a intervenção. Nesse caso, define-se o  $MSPE$  tal qual a equação abaixo:

$$MSPE = \sqrt{\frac{1}{T - T_0} \sum_{t=T_0+1}^T (Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt})^2} \quad (3)$$

A ideia do procedimento é atribuir o tratamento a cada um dos controles potenciais, estimando mais  $J$  vezes o controle sintético. Dessa forma, cada unidade de comparação torna-se um placebo. Assim, se o tratamento for real, espera-se observar um efeito da intervenção na unidade tratada e efeito nenhum nos placebos. Para quantificar essa proposta objetivamente, calcula-se  $MSPE$  definido em (3) para a unidade tratada e os placebos em dois subperíodos: o pré intervenção, denominado  $MSPE_{pre}$ , e o pós intervenção, denominado  $MSPE_{pos}$ . Feito isso, calcula-se a razão entre essas duas medidas para cada unidade  $j$ , consoante abaixo:

$$r_j = \frac{MSPE_{pos}}{MSPE_{pre}} \quad (4)$$

A intuição de (4) é a seguinte: se a intervenção produziu algum efeito na unidade  $j$ , o  $MSPE_{pos}$  deveria ser maior do que  $MSPE_{pre}$ , de modo que seria esperado  $r_j > 1$ . Além disso, se a intervenção for real, ela só ocorreu de fato na unidade exposta ao tratamento (Mariana, em nosso estudo). Logo, essa razão da unidade tratada deveria se destacar (ser maior, na maior parte dos casos) em relação às demais. Assim, uma maneira de quantificar essa relação consiste em calcular percentual de unidades  $j$  com  $r_j \geq r_1$ . Esse percentual, entendido como a probabilidade de ocorrência de um evento extremo na distribuição de permutações (um efeito da intervenção num placebo), é interpretado com a medida de  $p$  value da inferência estatística.

## 4.2 Controle sintético com múltiplas unidades tratadas: municípios diretamente afetados

Embora Mariana tenha sido o município mais afetado pelo desastre natural em pauta, recorde, da seção 3, que houve outros 3 municípios diretamente atingidos, os quais foram a princípio excluídos da análise na subseção anterior. Contudo, uma outra estratégia de estimação possível seria mensurar um efeito agregado do desastre para os 4 municípios diretamente atingidos. Nesse caso, segundo Abadie e L'Hour (2021), umas das formas possíveis de se utilizar o método quase experimental em análise seria estimar o controle sintético para cada unidade tratada e, em

seguida, calcular o efeito médio do tratamento como a média aritmética dos efeitos estimados para cada unidade exposta a intervenção. De forma objetiva, se há  $L$  municípios tratados, o efeito médio da intervenção ao longo do período de exposição para cada unidade impactada é, mantendo a notação da subseção 4.1, assim definido:

$$\hat{\alpha}_l = \frac{1}{T - T_0} (Y_{1lt} - \sum_{j=2}^{J+1} w_j^* Y_{jt}) \quad (5)$$

Em que

- $l \in 1, 2, \dots, L$

Nesse sentido, o efeito médio do tratamento  $\hat{\tau}$  é dado por:

$$\hat{\tau} = \frac{1}{L} \sum_{l=1}^L \alpha_l \quad (6)$$

#### 4.2.1 Inferência estatística

Para o caso de múltiplas unidades tratadas, Abadie e L'Hour (2021) propõem metodologia similar a exposta na subsubseção 4.1.1, com a diferença nos insumos utilizados. Em particular, se houve  $L$  unidades tratadas, realiza-se um procedimento de *bootstrap* no qual cada placebo é definido num sorteio de  $L$  unidades selecionadas dentre os controles potenciais (delimitados, neste estudo, na seção 3) para receberem a intervenção. Esse sorteio é feito várias vezes e, em cada uma delas, ao conjunto de unidades pseudotratadas, estima-se o controle sintético tal qual descrito na subseção 4.2.<sup>13</sup> Feito isso, obtemos  $J + 1$  medidas de efeito médio do tratamento  $\hat{\tau}_j$ , a partir das quais calcula-se o percentual de vezes nos quais o  $\hat{\tau}_j$  de um placebo foi superior ao  $\hat{\tau}_j$  do conjunto de unidades efetivamente expostas a intervenção. Esse percentual, por sua vez, é interpretado como o *p value* da inferência estatística.<sup>14</sup>

---

<sup>13</sup>Realizamos 500 sorteios, número default da função `pla.dist()`, utilizada no software estatístico R para a inferência estatística do controle sintético com múltiplas unidades tratadas. Para mais detalhes dessa função e o pacote a ela associado, veja documentação disponível em: <https://bityli.com/LkBoMd>

<sup>14</sup>Em tese, seria possível aplicar esse mesmo procedimento para o caso de uma única unidade tratada, embora exista o risco, nesse caso, de se utilizar municípios repetidos na inferência ou, ainda, de algum controle potencial não ser sorteado. Nesse sentido, aplicamos a metodologia no caso de Mariana como única unidade tratada. Contudo, das 15 variáveis dependentes analisadas, em apenas 3 o cálculo da inferência foi factível (nas demais, o software R apresentou um erro durante a estimação, que, por isso, não se concluiu). Por essa razão, optamos por alocar esses resultados no apêndice e manter os dois procedimentos distintos de inferência

## 5 Resultados

### 5.1 Controle Sintético com uma unidade tratada: Mariana

Começaremos apresentando um panorama geral dos resultados e, em seguida, discorreremos sobre os mais importantes. Nessa medida, a tabela 2 apresenta o *p value* do efeito estimado do desastre de Mariana sobre cada variável dependente examinada. Nota-se que, das 15 medidas de desempenho escolar, em apenas 3 houve impacto estatisticamente significativo. Em particular, no TDI do 1º ano do fundamental, a significância estatística alcança os 5%, enquanto, para TDI no 5º do fundamental e Taxa Reprovação no 1º ano dessa mesma etapa, as estimativas foram marginalmente significativas. Nas demais variáveis, não se observa significância estatística nem a 10%

Tabela 2: P Value dos Resultados Estimados por Variável Dependente

	TDI	Abandono	Reprovação
1º ano	0,02	0,16	0,08
5º ano	0,09	0,21	0,22
9º ano	0,82	0,67	0,79
1ª série EM	0,71	0,60	0,76
3ª série EM	0,24	0,90	0,97

Fonte: Elaboração própria

Nota: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Quanto a direção dos impactos, ressalta-se que, nos casos em que foi possível a rejeição da hipótese nula, encontrou-se, em geral, efeitos deletérios, no sentido de um *gap* positivo entre Mariana e o controle sintético nas taxas de desempenho avaliadas. Contudo, a magnitude dos efeitos é bastante diversa a depender do nível de ensino e da medida de resultado escolar. É o que ilustra a tabela 3, a qual exibe, para cada variável dependente, o assim chamado efeito máximo do tratamento. Para entendê-lo, recorde, da seção 4, que o efeito do tratamento é a diferença, na variável dependente, entre a unidade tratada e a de controle em cada ano pós intervenção. Logo, o efeito máximo é simplesmente o maior *gap* entre Mariana e a unidade artificial de comparação no período posterior ao desastre. <sup>15</sup>

<sup>15</sup>Uma alternativa seria trabalhar com o efeito médio ao longo do período pós desastre. Contudo, como, em muitos casos, o efeito, quando existiu, foi transitório, a média subestimaria os efeitos observados.

Tabela 3: Efeito Máximo do Tratamento por Variável Dependente (em pontos percentuais)

	TDI	Abandono	Reprovação
1º ano	28,00	0,12	2,96
5º ano	8,63	0,78	1,69
9º ano	8,50	2,91	3,28
1ª série EM	6,28	9,18	2,30
3ª série EM	7,69	1,93	3,59

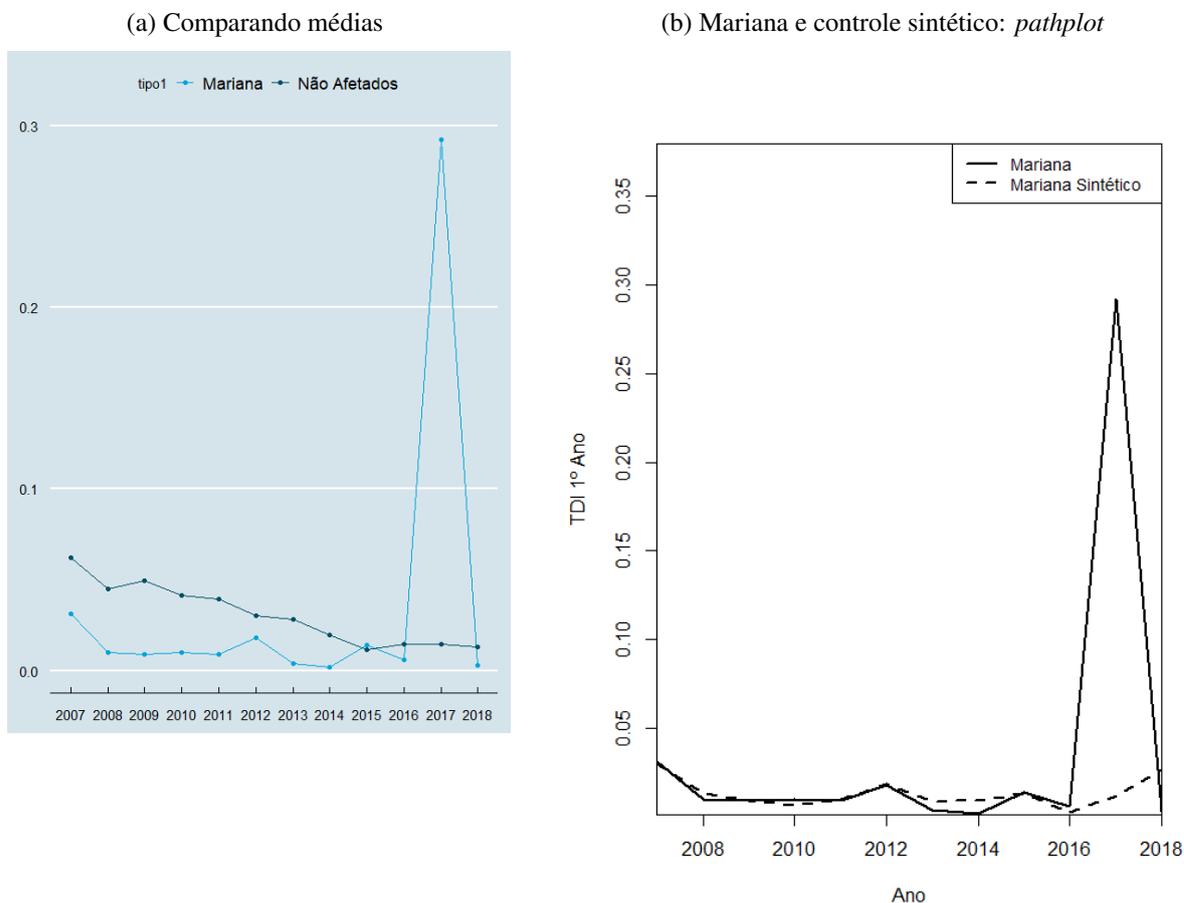
Fonte: Elaboração própria

Nota: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Nesse sentido, destaca-se o impacto na TDI do 1º ano do fundamental, para a qual observou-se um aumento de quase 30 pontos percentuais em Mariana comparativamente ao controle sintético entre 2016 e 2017. Por sua vez, no 5º ano do fundamental o impacto na TDI foi mais modesto, um aumento de 8,6 pontos percentuais entre 2015 e 2016. Já com relação a taxa de Reprovação no 1º ano, houve o menor impacto: um aumento de cerca de 3 pontos percentuais entre 2016 e 2017. Nas demais variáveis, embora a magnitude indique aumento nas taxas, não houve significância estatística. Concluído esse panorama inicial, passamos, a seguir, para uma análise mais minuciosa com respeito às três variáveis em cujo efeito estimado foi possível a rejeição da hipótese nula.

### 5.1.1 TDI 1º ano fundamental

Figura 5: TDI 1º ano fundamental



Fonte: Elaboração própria

No período pré intervenção, até 2014, a TDI no 1º ano, em Mariana e na média dos municípios não afetados, permanece relativamente estável, embora com tendência de queda, conforme aponta o painel (a) da figura 5. Essa tendência pré desastre se confirma no painel (b), pelo qual nota-se, por outro lado, em 2017, um aumento da ordem de 28 pontos percentuais na TDI de Mariana comparativamente ao controle sintético. Esse descolamento é transitório, visto que, já no ano seguinte, a variável de interesse volta ao patamar próximo de zero e da unidade artificial de comparação.

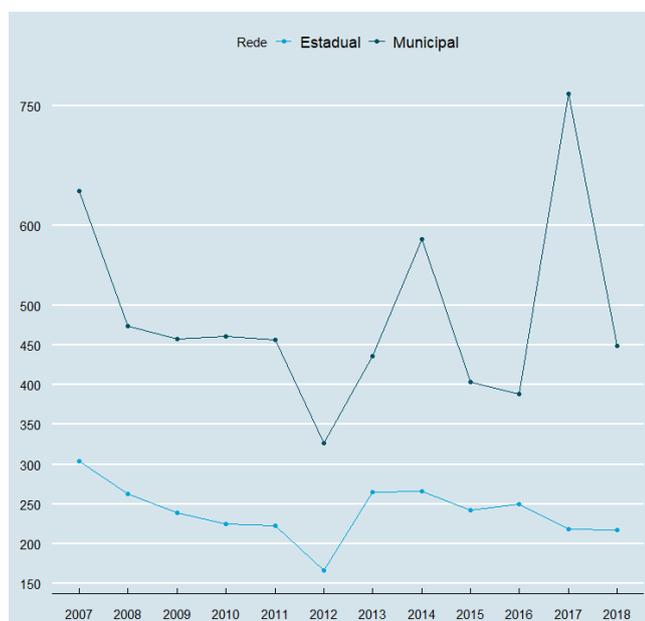
Essa evidência de um possível efeito chama atenção por pelo menos dois motivos. A distorção idade-série, em teoria, pode ser reflexo de reprovação ou abandono excessivo em séries anteriores ou na própria série em análise. Contudo, como o 1º ano marca o início do ensino básico, não há níveis anteriores os quais poderiam ter contribuído para um atraso no ciclo escolar. Além disso, conforme disposto nas tabelas 2 e 3, a comparação das taxas de abandono e reprovação de Mariana e do controle sintético não apresenta indícios convincentes para justificar um aumento tão expressivo e transitório da TDI em Mariana. Nesse sentido, uma possível explicação pode estar relacionada a trajetória de matrículas no 1º ano, conforme discutiremos sobre brevemente a seguir.

No período de 2008 a 2016, vide figura 6, ao comparar a quantidade de matriculados no 1º ano em Mariana na rede municipal e na estadual, esta permanece estável em torno de 300

alunos, enquanto aquela oscila entre 300 e 400 estudantes na maior parte do tempo, variando no máximo cerca de 30% de um ano para o outro. Contudo, em 2017, conquanto o número de matriculados permaneça estável na rede estadual, o número de alunos quase dobra na rede municipal desse município, um aumento em magnitude não observado nos anos anteriores.

Com base nesse cenário, formulamos a seguinte hipótese: um número razoável de alunos deveriam ter entrado no 1º ano em período anterior a 2017, mas não o fizeram, pois, dentre outros motivos, o desastre de Mariana incentivou as famílias a postergarem ainda mais a entrada dessas crianças no ensino básico. Dessa forma, quando enfim iniciaram o primeiro ciclo educacional em 2017, já estavam em idade consideravelmente acima da esperada para esse nível escolar, o que contribuiu para esse salto repentino e transitório na TDI do 1º ano. Afinal, no ano seguinte, em 2018 o fluxo de matrículas volta a ficar próximo do patamar pré 2017, e a TDI volta a ficar próxima de zero. Evidentemente, essa explicação é apenas uma hipótese a qual não foi testada. Trata-se, isso sim, de uma tentativa de explicar esse efeito com uma narrativa consistente com os dados apresentados.

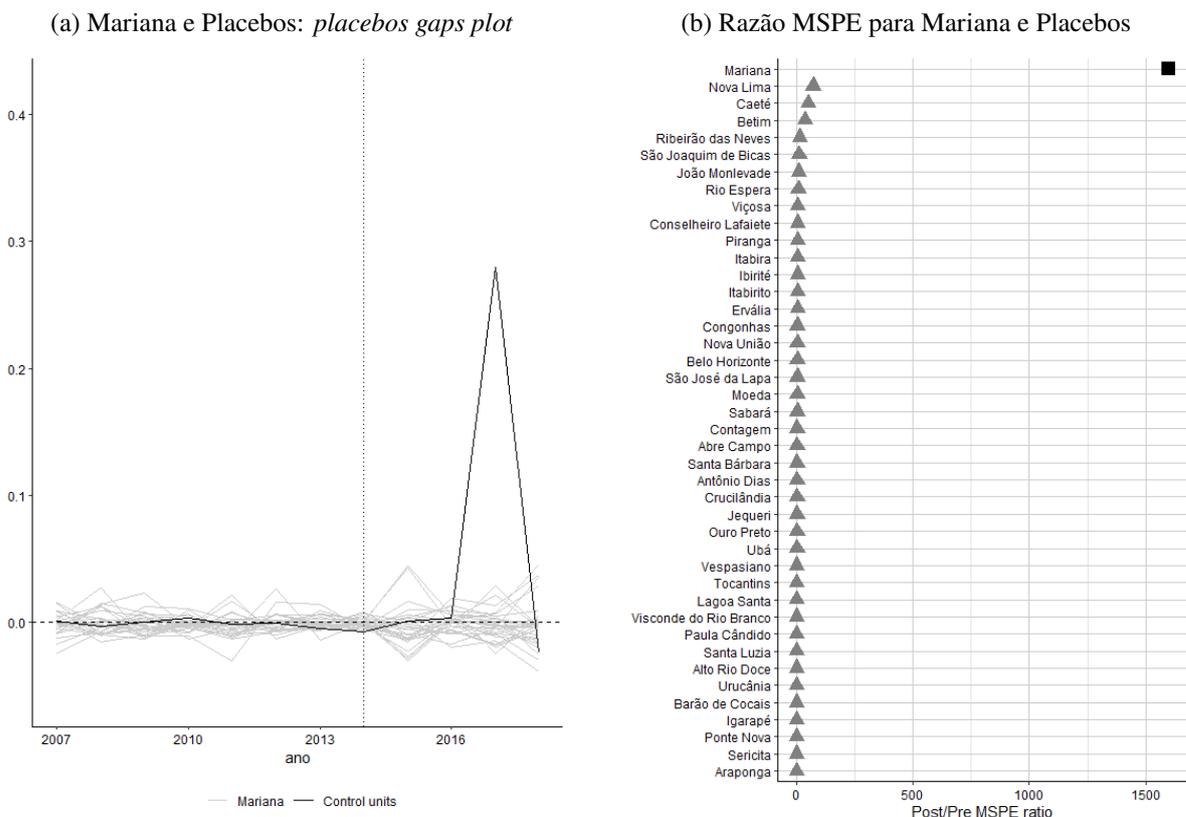
Figura 6: Quantidade de matrículas no 1º ano em Mariana: rede estadual e municipal



Fonte: Elaboração própria com base em dados dos indicadores educacionais do INEP

Prosseguindo com as estimações do controle sintético, abordemos a inferência estatística. Para tanto, seguindo a metodologia da subsubseção 4.1.1, começamos com o painel (a) da figura 7, que ilustra um exercício conhecido na literatura como *placebos gaps plot*. A ideia é comparar o efeito estimado da intervenção em Mariana, unidade tratada, com o correspondente ao dos placebos, para os quais não se deveria observar efeito algum. Na prática, ainda nesse mesmo painel, isso é verificado ao notarmos que, uma vez excluídos os casos extremos, em 2017, período após o desastre no qual ocorre o salto da TDI em Mariana, o *gap plot* desse município está bem acima em relação ao das demais unidades de comparação. Ou seja, em geral, os municípios estão abaixo de Mariana por essa métrica, o que respalda a possibilidade um efeito crível desse desastre natural sobre o resultado educacional em pauta.

Figura 7: Inferência estatística: TDI 1º ano fundamental



Fonte: Elaboração própria

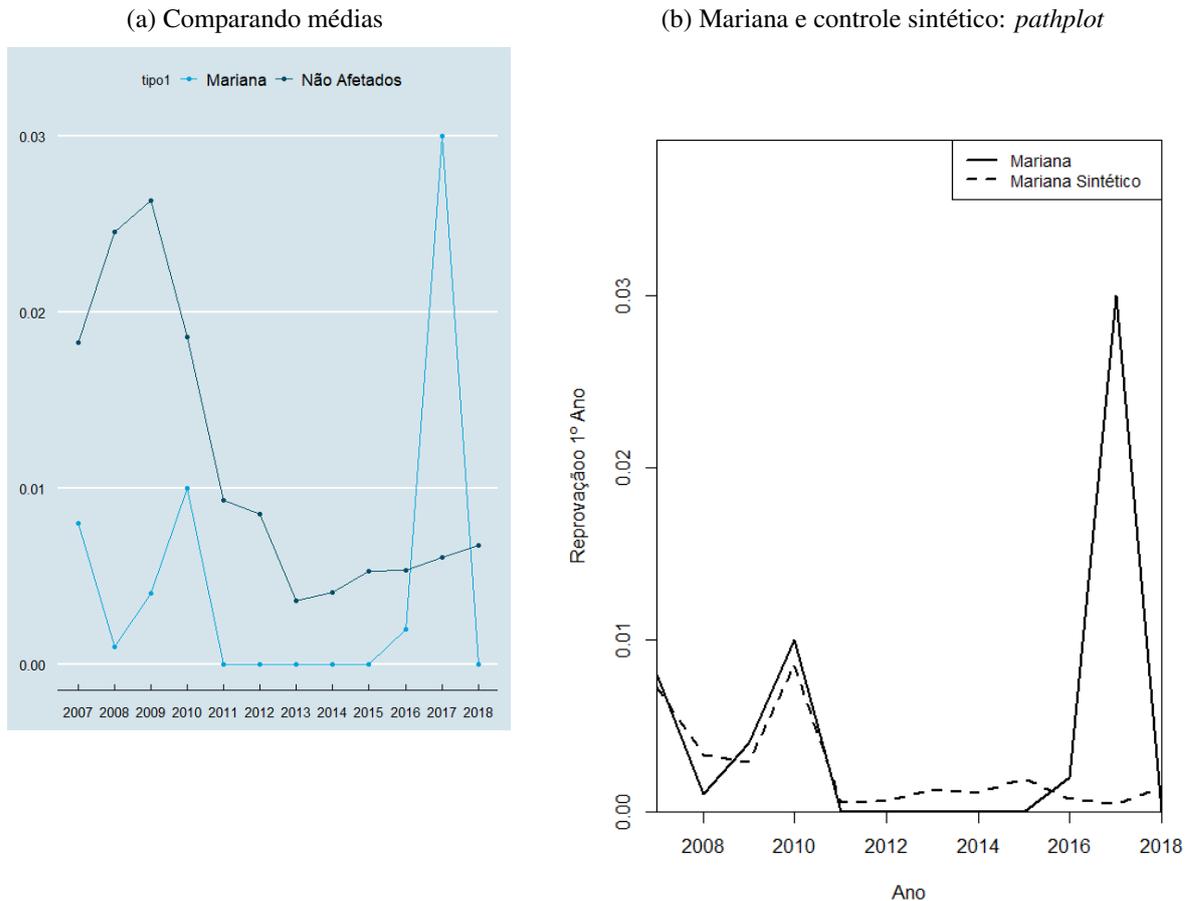
O painel (b) da figura 7 mostra a razão entre o Erro de Previsão Quadrático Médio (MSPE) pós e pré tratamento, doravante denominado razão pós/pré MSPE, para Mariana e Placebos. Note que, com respeito a esse quociente, Mariana encontra-se no topo da distribuição, pois é o primeiro num ranking em ordem decrescente de 42 municípios.<sup>16</sup> Isso significa que, caso a intervenção fosse designada de forma aleatória nos dados, a probabilidade de se obter uma razão pré/pós MSPE maior ou igual a de Mariana é  $1/42 = 0.0238$ , medida interpretada como o *p value* associado. Ou seja, o efeito estimado do desastre de Mariana seria estatisticamente significativo a 5%.

### 5.1.2 Reprovação 1º ano fundamental

A comparação de médias, realizada no painel (a) da figura 8, fornece indícios de uma tendência corroborada no *path plot* (painel (b)): até 2014, as taxas de reprovação de Mariana e do controle sintético eram parecidas e, na maior parte do tempo, próximas de zero. Em contrapartida, no ano de 2017, a taxa de reprovação de Mariana aumenta em quase 3 pontos percentuais e se distancia temporariamente da unidade artificial de comparação, pois, no ano seguinte, a variável de interesse retorna ao nível próximo de zero e do controle sintético. Isto é, a figura 8 sugere que o desastre em análise pode ter contribuído para um aumento pequeno e efêmero na reprovação do 1º ano.

<sup>16</sup>Dos 102 municípios não afetados a 100km de Mariana, 60 foram descartados no algoritmo de minimização dupla, nesse caso em particular, por indisponibilidade de informações da TDI no 1º ano em um ou mais anos do período pré tratamento.

Figura 8: Taxa de reprovação 1º ano fundamental



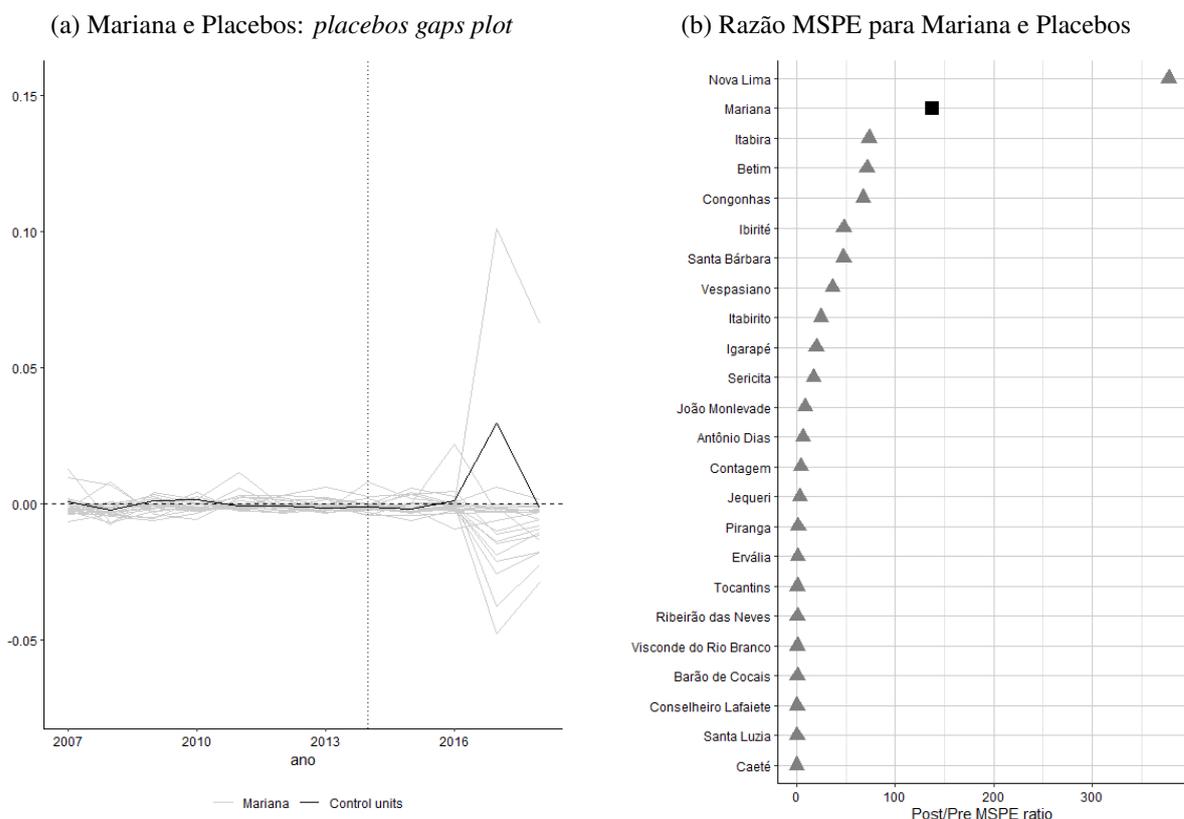
Fonte: Elaboração própria

Em se tratando de inferência estatística, de forma análoga a subsubseção 5.1.1, começamos pelo painel (a) da figura 9, o qual explicita o *placebo gaps plot*, pelo qual compara-se o efeito da intervenção na unidade efetivamente tratada (Mariana) com o efeito nos placebos, os controles potenciais. Nesse sentido, o *gaps plot* de Mariana encontra-se bem acima comparativamente às demais unidades de comparação. Isto é, em geral, os municípios não afetados estão abaixo de Mariana por essa métrica, o que reforça a possibilidade de um efeito real, embora de pequena magnitude e transitório, desse desastre natural sobre o resultado educacional em análise.

Por seu turno, o painel (b) ilustra a razão entre o MSPE pós e pré tratamento para a unidade tratada e placebos. Note que, em relação a esse quociente, Mariana encontra-se no topo da distribuição, pois é o segundo num ranking em ordem decrescente de 24 municípios.<sup>17</sup> Isso significa que, caso a intervenção fosse designada de forma aleatória nos dados, a probabilidade de se obter uma razão pós/pré MSPE maior ou igual a de Mariana é  $2/24 = 0.0833$ , medida interpretada como o *p value* associado. Ou seja, o efeito estimado do desastre de Mariana seria estatisticamente significativo a 10%.

<sup>17</sup>O *default* do software R é descartar municípios com razão MSPE pré pós superior 20 vezes ou mais comparativamente a unidade tratada. São os assim chamados "casos extremos". Por isso, diminuiu o número municípios em relação a análise da distorção idade série, quando havia 42 no total.

Figura 9: Inferência estatística: Reprovação 1º ano fundamental

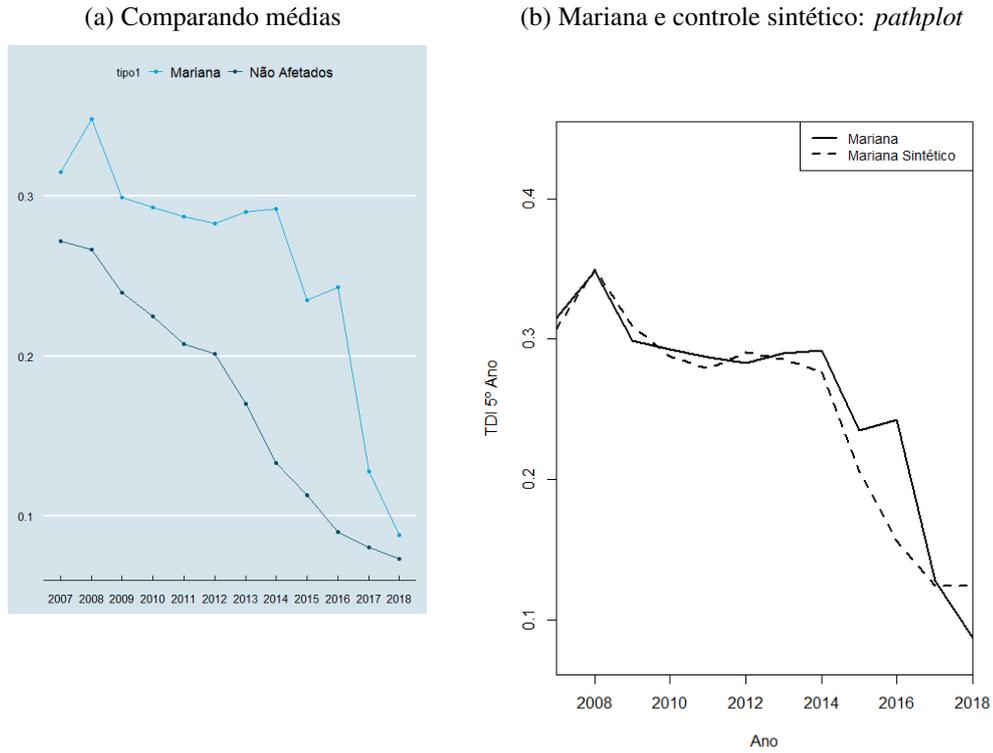


### 5.1.3 TDI 5º ano fundamental

Na comparação entre Mariana e a média dos municípios não afetados, vide painel (a) da figura 10, nota-se que as tendências da TDI do 5º ano são decrescentes e paralelas até 2014, período no qual o controle sintético apresenta trajetória parecida ao do município atingido, conforme sugere *path plot* exibido no painel (b) dessa mesma figura. Por outro lado, em 2016, a TDI em Mariana aumenta em 8,6 pontos percentuais comparativamente a unidade artificial de comparação. Nos anos seguinte até 2018, a variável de interesse volta a cair na unidade tratada até convergir para um patamar similar ao do controle sintético. Portanto, a figura 10 indica que o evento em análise pode ter contribuído para um aumento expressivo, porém transitório, nesse resultado educacional.

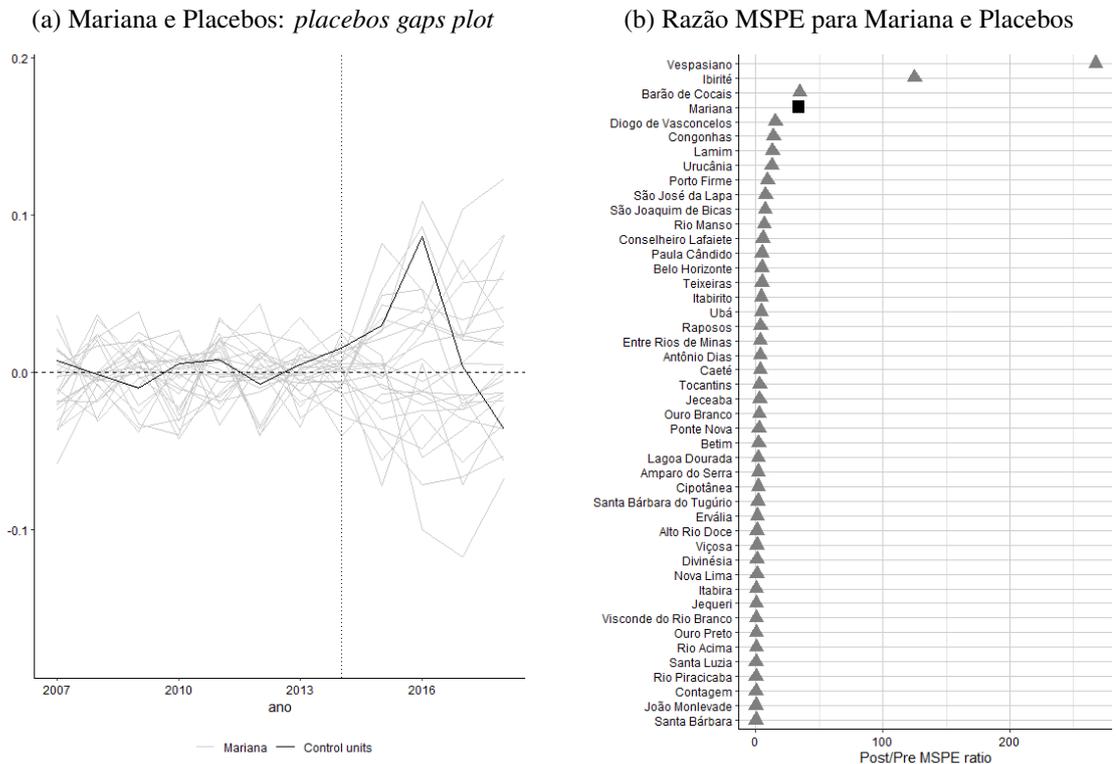
Quanto a inferência estatística, mantendo o padrão das subsubseções 5.1.1 e 5.1.2, começamos pelo painel (a) da figura 11, que exibe o *placebos gaps plot*, a partir do qual compara-se o impacto da intervenção na unidade efetivamente tratada com o efeito nos placebos, os controles potenciais. Nessa linha, o *gaps plot* de Mariana se destaca comparativamente a maioria das unidades de comparação, uma métrica que respalda a possibilidade de um efeito real, não obstante transitório, desse desastre natural.

Figura 10: TDI 5º ano fundamental



Fonte: Elaboração própria

Figura 11: Inferência estatística: TDI 5º ano fundamental



Fonte: Elaboração própria

Por sua vez, o painel (b) da figura 11 apresenta a razão entre o MSPE pós e pré tratamento para a unidade afetada e placebos. Com respeito a esse quociente, percebe-se que Mariana encontra-se próximo do topo da distribuição, pois é o quarto num ranking de 46 municípios. Isso significa o seguinte: caso a intervenção fosse designada de forma aleatória nos dados, a probabilidade de se obter uma razão pós/pré MSPE maior ou igual a de Mariana é  $4/46 = 0.087$ , medida interpretada como o *p value* associado. Ou seja, o efeito estimado do desastre de Mariana seria estatisticamente significativo a 10%.

## 5.2 Controle sintético com múltiplas unidades tratadas: municípios diretamente afetados

De forma análoga a subseção 4.1, começaremos apresentando um panorama geral dos resultados e, em seguida, discorreremos sobre os mais importantes. Nesse sentido, a tabela 4 a seguir apresenta o *p value* do desastre em pauta sobre cada variável dependente avaliada. Percebe-se que, das 14 medidas de desempenho escolar, em apenas 5 houve impacto não significativo estatisticamente. Em contrapartida, para o TDI e reprovação no 5º ano e para TDI na 3ª série do ensino médio, os efeitos foram marginalmente significativos (10%), enquanto nas outras 9 variáveis, a rejeição da hipótese nula ocorreu a 5% .<sup>18</sup>

Tabela 4: P Value dos Resultados Estimados por Variável Dependente

	TDI	Abandono	Reprovação
1º ano	-	0,02	0,37
5º ano	0,10	0,00	0,07
9º ano	0,00	0,00	0,16
1ª série EM	0.194	0.492	0,04
3ª série EM	0,06	0,39	0,02

Fonte: Elaboração própria

Nota1: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Nota2: Valores com “ - ” indicam que a estimação não foi possível devido a excesso de *missing* na variável dependente

Com relação a direção dos impactos, enfatiza-se que, nos casos para os quais foi possível a rejeição da hipótese nula, encontrou-se, em geral, efeitos prejudiciais, no sentido de um *gap* positivo entre o conjunto de municípios diretamente afetados e o controle sintético nas taxas de desempenho examinadas. Todavia, a magnitude dos efeitos é bastante diversa a depender do nível de ensino e da medida de resultado escolar. De fato, é o que sugere a tabela 5, a qual apresenta, para cada variável dependente, o assim chamado efeito médio do tratamento, que abarca os *gaps* entre o conjunto de locais atingidos e a unidade artificial de comparação no período posterior ao desastre.

<sup>18</sup>Como teste de robustez, estimou-se esse modelo definindo o período inicial de exposição ao desastre como 2016, em vez de 2015. Nessa linha, visto que as alterações nos resultados, para a maior parte das variáveis, foi marginal, alocamos esse exercício no apêndice.

Tabela 5: Efeito médio do tratamento (Multiple Synt; em pontos percentuais)

	TDI	Abandono	Reprovação
1º ano	-	0,72	1,07
5º ano	7,73	2,00	6,77
9º ano	13,58	6,43	6,82
1ª série EM	9,85	5,89	12,92
3ª série EM	10,23	3,86	9,96

Fonte: Elaboração própria

Nota1: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

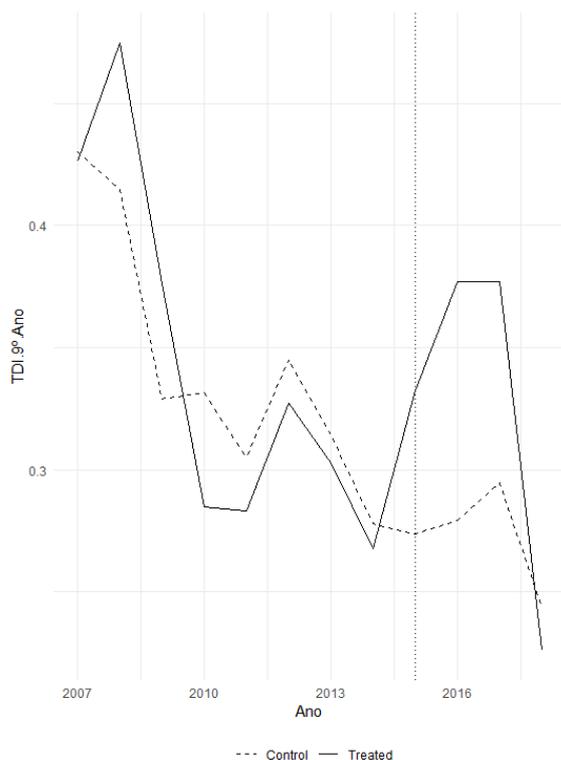
Nota2: Valores com “ - ” indicam que a estimação não foi possível devido a excesso de *missing* na variável dependente

Nessa linha, destacam-se 4 casos nos quais o impacto foi um aumento próximo ou superior a 10 pontos percentuais: TDI 9º ano (13,58 pontos percentuais), Reprovação na 1ª série do Ensino Médio (12,92 pontos percentuais) e TDI e Reprovação na 3ª série do Ensino Médio (10,23 e 9,96 pontos percentuais, respectivamente). Nas demais variáveis, exceto os casos sem significância estatística, o efeito oscila entre 0,72 e 7,73 pontos percentuais. Findado esse panorama inicial, realiza-se, a seguir, uma análise mais aprofundada com respeito a três das nove variáveis em cujo efeito estimado foi possível a rejeição da hipótese nula.

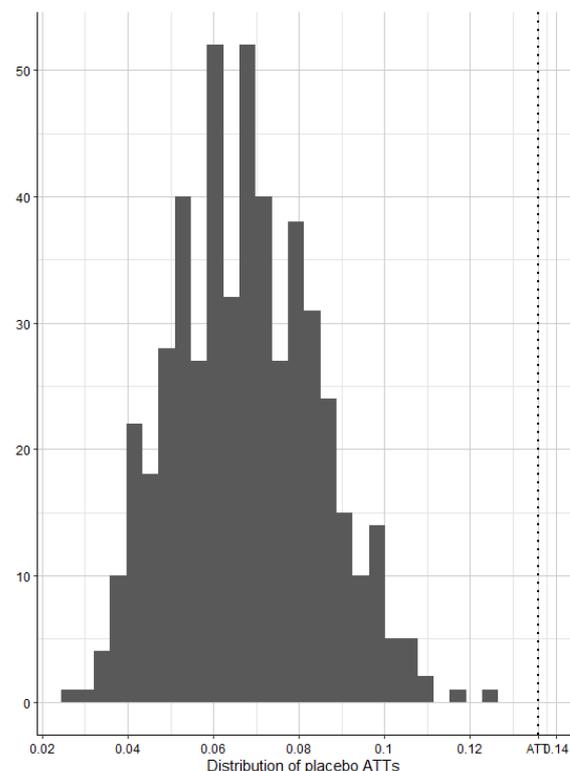
## 5.2.1 TDI 9º ano fundamental

Figura 12: TDI 9º ano fundamental

(a) Conjunto de municípios afetados e controle sintético: *path plot*



(b) Conjunto de municípios afetados e controle sintético: inferência estatística



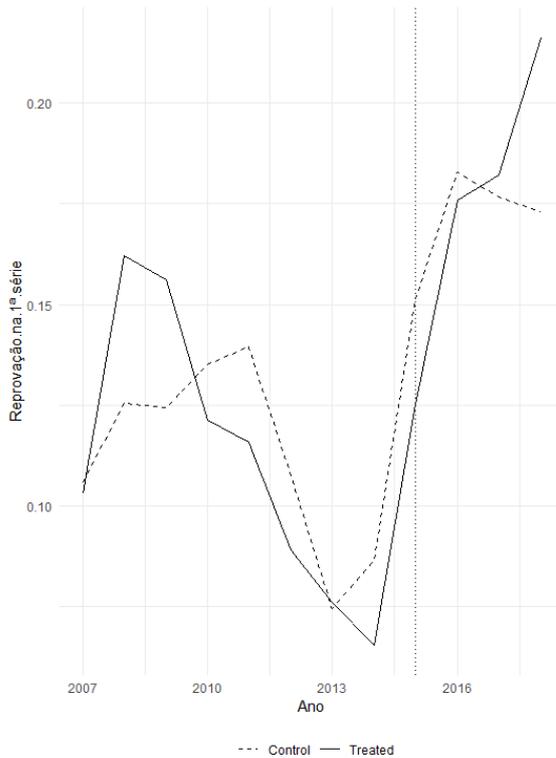
Fonte: Elaboração própria

Consoante evidencia o painel (a) da figura 12, até 2013, a TDI do 9º ano apresentou trajetória, na maior parte do tempo, decrescente e parecida no conjunto de municípios atingidos e no controle sintético. Em seguida, entre 2014 e 2017, essa variável aumenta para a unidade tratada, a qual se distancia da unidade de comparação, o que resulta num efeito médio do tratamento de 13,58 pontos percentuais. Contudo, no ano seguinte, a TDI dos municípios afetados e da unidade artificial voltam a cair e se aproximam num patamar em torno de 20 pontos percentuais. Por seu turno, o painel (b) dessa figura apresenta, seguindo a metodologia da subseção 4.2, a distribuição do efeito médio do desastre nas cidades atingidas e nos placebos. Nesse caso, nota-se que o conjunto de municípios afetados, por apresentar o maior impacto estimado, destaca-se na ponta a direita da distribuição, o que resulta num *p value* de 0,00. Em suma, esse exercício aponta um efeito deletério do evento em análise sobre a TDI no 9º ano, contudo, esse impacto se dissipou ao longo do tempo e foi revertido cerca de 3 anos após a tragédia.

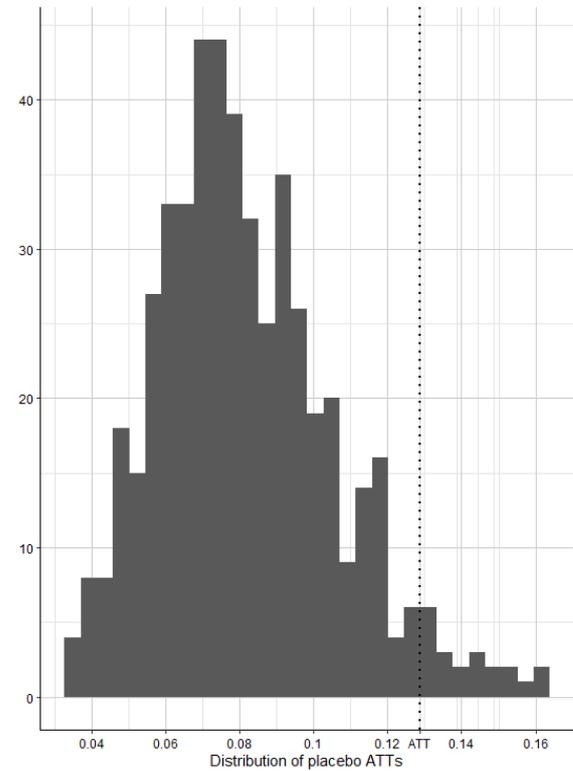
## 5.2.2 Reprovação 1ª série ensino médio

Figura 13: Reprovação 1ª série ensino médio

(a) Conjunto de municípios afetados e controle sintético: *path plot*



(b) Conjunto de municípios afetados e controle sintético: inferência estatística



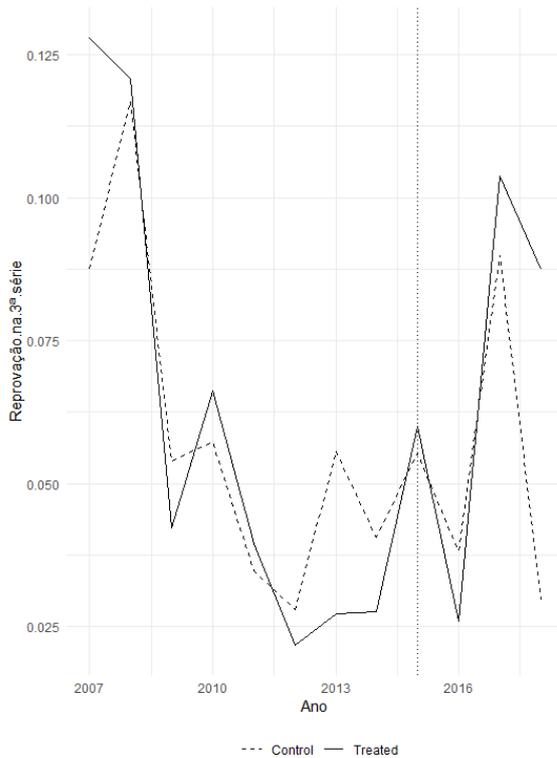
Fonte: Elaboração própria

Entre 2007 e 2014, vide painel (a) da figura 13, as trajetórias da taxa de reprovação dos municípios afetados e do controle sintéticos eram similares e, na maior parte do tempo, decrescentes. Após 2014, ambas começam a subir, contudo, a partir de 2016, enquanto a variável de interesse se estabiliza para o controle sintético, ela continua a aumentar para o conjunto de cidades afetadas até 2018, último ano da série temporal. Por sua vez, no painel (b) dessa figura, ilustra-se a distribuição do efeito médio do desastre nos locais atingidos e nos placebos. Percebe-se, nesse caso, que o conjunto de municípios afetados, por apresentar um dos maiores impactos estimados, situa-se próximo da ponta a direita da distribuição, resultando num *p value* de 0,04. Em síntese, as estimações sugerem um efeito que, embora seja defasado, significou um aumento médio de 12,92 pontos percentuais na taxa de reprovação em pauta no período posterior ao desastre.

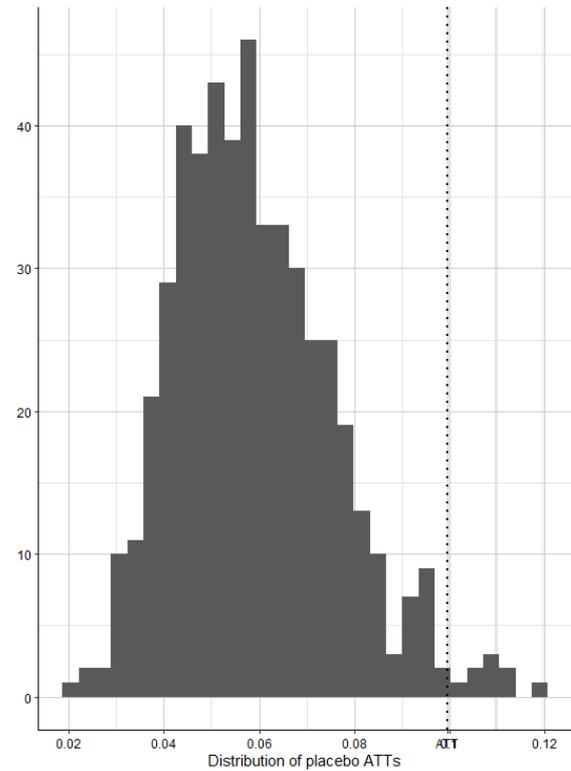
### 5.2.3 Reprovação 3ª série ensino médio

Figura 14: Reprovação 3ª série ensino médio

(a) Conjunto de municípios afetados e controle sintético: *path plot*



(b) Conjunto de municípios afetados e controle sintético: inferência estatística



Fonte: Elaboração própria

Conforme mostra o painel (a) da figura 14, entre 2007 e 2016, nota-se uma trajetória, na maior parte do tempo, decrescente da taxa de reprovação na 3ª série tanto para o controle sintético, como também para o conjunto de cidades atingidas. Em 2017, essa taxa sobe em ambos casos, mas o aumento é mais acentuado para o conjunto de municípios afetados, de modo que estes ficam acima da unidade artificial de comparação mesmo em 2018, ano no qual a tendência passa a ser de queda em ambos casos. Esse descolamento a partir de 2017 se traduziu num efeito médio do tratamento de 9,96 pontos percentuais no período pós intervenção. Por sua parte, o painel (b) dessa figura explicita a distribuição do efeito médio do desastre nos locais atingidos e nos placebos. Nessa medida, como o conjunto de municípios afetados apresenta um dos maiores impactos estimados, esse grupo localiza-se próximo da ponta direita da distribuição, o que resulta num *p value* de 0,02.

## 6 Conclusão

O presente trabalho buscou realizar uma análise dos impactos do Desastre de Mariana, onde em 2015 houve rompimento da Barragem de Fundão, sobre resultados educacionais. Para tanto, utilizou-se um método quase experimental de inferência causal, o controle sintético. Nessa medida, o primeiro passo foi reconhecer a existência de uma literatura dedicada a análise dos efeitos de desastres naturais sobre variáveis educacionais. Esses estudos sugerem que, não obstante os resultados sejam mistos quanto aos efeitos de curto prazo, os impactos de longo prazo costumam ser deletérios na acumulação de capital humano.

Em seguida, para responder especificamente a pergunta de pesquisa escolhida, definimos dois recortes espaciais. No primeiro, Mariana consistiu na unidade tratada; no segundo, o conjunto de cidades diretamente afetadas constituiu a unidade exposta à intervenção. Em ambos casos, os controles potenciais foram derivados de municípios mineiros não atingidos pelo desastre e geograficamente mais próximos da região afetada. Além disso, estudou-se 3 indicadores de desempenho escolar — as taxas de abandono, reprovação e distorção-idade-série —, para o primeiro, quinto e nono ano do ensino fundamental, bem como para a primeira e terceira série do ensino médio.

Nesse sentido, quando a análise se restringiu aos impactos sobre Mariana, não houve significância estatística no efeito estimado para a maioria dos indicadores educacionais avaliados. Contudo, das 3 variáveis para as quais se rejeitou a hipótese nula, destacou-se a TDI no 1º ano do ensino fundamental, que apresentou, dois anos após a tragédia, um aumento transitório de quase 30 pontos percentuais em Mariana relativamente ao controle sintético. Nesse caso, na ausência de um efeito compatível com esse em outros indicadores de desempenho no 1º ano, formulou-se a hipótese de que esse impacto pode estar relacionado a uma postergação da entrada de crianças no 1º ano após esse desastre, uma conjectura a qual, embora não testada rigorosamente, é consistente com a dinâmica da quantidade de alunos matriculados nessa série entre 2007 e 2018.

Por outro lado, quando a análise centrou-se no conjunto de municípios diretamente afetados, o cenário se inverteu, pois houve significância estatística no efeito estimado sobre a maioria dos resultados educacionais examinados. Apesar disso, a magnitude dos impactos apresentou amplitude elevada a depender da medida de desempenho e da série analisada. De fato, os efeitos médios estimados, que foram em geral aumentos nas taxas avaliadas, oscilaram de 0,72 pontos percentuais (abandono no 1º ano) a 13,58 pontos percentuais (TDI no 9º ano).

Em suma, embora os resultados sugiram efeitos de curto prazo deletérios, esses impactos não se manifestaram em todas as dimensões do desempenho escolar e, nos casos detectados, o prejuízo educacional foi heterogêneo em função da série e da medida de desempenho escolar. De toda forma, esse cenário é consistente com a literatura sobre desastres naturais, de acordo com a qual os impactos nos resultados escolares variam de acordo com a circunstância em que se deu a tragédia. Nesse sentido, uma perspectiva para trabalhos futuros consiste em avaliar se haverá efeitos de longo prazo sobre a acumulação de capital humano. Afinal, conforme apontam Mrejen *et al* (2020) e Carrilo *et al* (2020), houve danos a saúde física das crianças nascidas logo após essa tragédia.

## Referências

ABADIE, Alberto; DIAMOND, Alexis; HAINMUELLER, Jens. Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. **Journal of the American statistical Association**, v. 105, n. 490, p. 493-505, 2010.

ABADIE, Alberto; L' HOUR, Jérémy. A penalized synthetic control estimator for disaggregated data. **Journal of the American Statistical Association**, v. 116, n. 536, p. 1817-1834, 2021.

ANDRABI, Tahir; DANIELS, Benjamin; DAS, Jishnu. Human capital accumulation and disasters: Evidence from the Pakistan earthquake of 2005. **Journal of Human Resources**, p. 0520-10887R1, 2021.

ANGRIST, Noam et al. Building back better to avert a learning catastrophe: Estimating learning loss from COVID-19 school shutdowns in Africa and facilitating short-term and long-term learning recovery. **International Journal of Educational Development**, v. 84, p. 102397, 2021.

BAEZ, Javier E.; DE LA FUENTE, Alejandro; SANTOS, Indhira. Do natural disasters affect human capital? An assessment based on existing empirical evidence. **IZA Discussion Papers** 5164. Institute for the Study of Labor (IZA), v. 10, 2010. Disponível em: <<https://docs.iza.org/dp5164.pdf>>. Acesso em: 6 de fevereiro. 2022.

CARRILLO, Bladimir et al. Avoidable environmental disasters and infant health: Evidence from a mining dam collapse in Brazil. **Health economics**, v. 29, n. 12, p. 1786-1794, 2020.

CARUSO, Germán; MILLER, Sebastian. Long run effects and intergenerational transmission of natural disasters: A case study on the 1970 Ancash Earthquake. **Journal of development economics**, v. 117, p. 134-150, 2015.

CARUSO, Germán Daniel. The legacy of natural disasters: The intergenerational impact of 100 years of disasters in Latin America. **Journal of Development Economics**, v. 127, p. 209-233, 2017

MREJEN, Matias; PERELMAN, Julian; MACHADO, Danielle Carusi. Environmental disasters and birth outcomes: Impact of a tailings dam breakage in Brazil. **Social Science & Medicine**, v. 250, p. 112868, 2020.

NGUYEN, Cuong Viet; MINH PHAM, Nguyet. The impact of natural disasters on children's education: Comparative evidence from Ethiopia, India, Peru, and Vietnam. **Review of Development Economics**, v. 22, n. 4, p. 1561-1589, 2018.

PAUDEL, Jayash; RYU, Hanbyul. Natural disasters and human capital: The case of Nepal's earthquake. **World Development**, v. 111, p. 1-12, 2018

RUSH, John V. The impact of natural disasters on education in Indonesia. **Economics of disasters and climate change**, v. 2, n. 2, p. 137-158, 2018.

SABATES, Ricardo; CARTER, Emma; STERN, Jonathan MB. Using educational transitions to estimate learning loss due to COVID-19 school closures: The case of Complementary Basic Education in Ghana. **International Journal of Educational Development**, v. 82, p. 102377, 2021.

SANTOS, Daniel; PRIMI, Ricardo. Desenvolvimento socioemocional e aprendizado escolar: uma proposta de mensuração para apoiar políticas públicas. **Relatório sobre resultados preliminares do projeto de medição de competências socioemocionais no Rio de Janeiro. São Paulo: OCDE, SEEDUC, Instituto Ayrton Senna**, 2014. Disponível em: <<https://bityli.com/ICySkb>>. Acesso em: 6 de fevereiro. 2022.

WANG, Jun; YANG, Juan; LI, Bo. Pain of disasters: The educational cost of exogenous shocks evidence from Tangshan Earthquake in 1976. **China Economic Review**, v. 46, p. 27-49,

2017.

FORÇA-TAREFA, Grupo. Avaliação dos efeitos e desdobramentos do rompimento da Barragem de Fundão em Mariana-MG. **Belo Horizonte: Secretaria de Desenvolvimento Regional, Política Urbana e Gestão Metropolitana, Governo do Estado de Minas Gerais**, 2016. Disponível em:

<<https://bitly.com/rpTQuU>>. Acesso em: 6 de fevereiro. 2022.

## Apêndice

### A1: Controle sintético para o conjunto de municípios diretamente afetados (ano inicial de exposição a intervenção igual a 2016)

Tabela A1.1: P Value dos Resultados Estimados por Variável Dependente (Ano inicial da intervenção igual a 2016)

	TDI	Abandono	Reprovação
1º ano	-	0,00	0,28
5º ano	0,53	0,10	0,42
9º ano	0,00	0,00	0,78
1ª série EM	0,11	-	0,02
3ª série EM	0,01	0,33	0,00

Fonte: Elaboração própria

Nota: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Nota2: Valores com “ - ” indicam que a estimação não foi possível devido a excesso de *missing* na variável dependente

Tabela A1.2: Efeito médio do tratamento (em pontos percentuais; ano inicial da intervenção igual a 2016)

	TDI	Abandono	Reprovação
1º ano	-	0,65	1,27
5º ano	4,69	0,7	2,85
9º ano	14,98	7,46	4,04
1ª série EM	11,72	-	12,92
3ª série EM	12,3	3,63	12,05

Fonte: Elaboração própria

Nota: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Nota2: Valores com “ - ” indicam que a estimação não foi possível devido a excesso de *missing* na variável dependente

## A2 Controle sintético para Mariana usando metodologia de inferência do caso com múltiplas unidades tratadas

Tabela A2.1: P Value dos Resultados Estimados por Variável Dependente (em pontos percentuais)

	TDI	Abandono	Reprovação
1º ano	0,00	0,99	0,39

Fonte: Elaboração própria

Nota1: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Nota2: A estimação só funcionou para TDI do 1º ano; nas demais variáveis, o *software* R apresentou um erro durante a estimação, o que impossibilitou a obtenção dos resultados

Tabela A2.2: Efeito médio do tratamento (em pontos percentuais)

	TDI	Abandono	Reprovação
1º ano	15,66	0,08	1,54

Fonte: Elaboração própria

Nota1: valores em verde indicam significância estatística menor ou igual a 10%; valores em vermelho, significância estatística maior que 10%

Nota2: A estimação só funcionou para TDI do 1º ano; nas demais variáveis, o *software* R apresentou um erro durante a estimação, o que impossibilitou a obtenção dos resultados