

**AVALIAÇÃO DOS IMPACTOS E VALORAÇÃO DOS
DANOS SOCIOECONÔMICOS CAUSADOS PARA
AS COMUNIDADES ATINGIDAS PELO ROMPIMENTO
DA BARRAGEM DE FUNDÃO**

**Avaliação de Impactos do Rompimento
da Barragem de Fundão sobre
o Mercado de Trabalho Formal
dos Municípios Atingidos**



DEZEMBRO DE 2020

Dados internacionais de Catalogação na Publicação
Ficha catalográfica elaborada pelo Sistema de Bibliotecas/FGV

Fundação Getulio Vargas

Avaliação de Impactos do Rompimento da Barragem de Fundão sobre o Mercado de Trabalho Formal dos Municípios Atingidos / Fundação Getulio Vargas – Rio de Janeiro; São Paulo: FGV, 2020.

151 p.

Em colaboração com: Amanda Cappellazzo Arabage, Ana Carolina Marinato Resende, André Portela Fernandes de Souza, Bernardo Ostrovski, Bruno Guimarães de Melo, Giovana Gonçalves Pereira, Marcos Paulo Cambrinha da Costa, Patrícia Franco Ravaioli, Reynaldo Fernandes, Victor Simões Dornelas.

Acima do título: Projeto Rio Doce.

Inclui bibliografia.

1. Projeto Rio Doce. 2. Fundão, Barragem de (MG), 3. Mercado de trabalho. 4. Desastres ambientais – Doce, Rio, Bacia (MG e ES). 5. Equilíbrio econômico. I. Título.

CDD – 627.8

EQUIPE TÉCNICA

Amanda Cappellazzo Arabage

Ana Carolina Marinato Resende

André Portela Fernandes de Souza

Bernardo Ostrovski

Bruno Guimarães de Melo

Giovana Gonçalves Pereira

Marcos Paulo Cambrainha da Costa

Patrícia Franco Ravaioli

Reynaldo Fernandes

Victor Simões Dornelas

LISTA DE FIGURAS

Figura 1 — Fluxograma ilustrando os potenciais canais e possíveis impactos do rompimento da Barragem de Fundão no mercado de trabalho formal dos municípios atingidos	21
Figura 2 — Ilustração do método de diferença-em-diferenças	35
Figura 3 — Mapa ilustrando o grupo atingido e o grupo de comparação 1 (GC 1).....	47
Figura 4 — Mapa ilustrando o grupo atingido e o grupo de comparação 2 (GC 2).....	49
Figura 5 — Distribuição de pessoas no grupo atingido por setores censitários	75

LISTA DE GRÁFICOS

Gráfico 1 — Estimativas de impacto médio sobre o logaritmo natural do número de vínculos ativos de municípios no grupo atingido (comparando com GC 1)	57
Gráfico 2 — Estimativas de impacto médio sobre o logaritmo natural do número de vínculos ativos de municípios no grupo atingido (comparando com GC 2)	59
Gráfico 3 — Estimativas de impacto médio sobre o logaritmo natural da massa salarial de municípios no grupo atingido (comparando com GC 1).....	63
Gráfico 4 — Estimativas de impacto médio sobre o logaritmo natural da massa salarial de municípios no grupo atingido (comparando com GC 2).....	64
Gráfico 5 — Pirâmide etária do grupo atingido.....	71
Gráfico 6 — Distribuição de pessoas no grupo atingido por raça/cor	72
Gráfico 7 — Distribuição de pessoas no grupo atingido por nível de escolaridade	73
Gráfico 8 — Distribuição de pessoas no grupo atingido por setor de atividade	74
Gráfico 9 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo (comparando com GC 1).....	81
Gráfico 10 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo (comparando com GC 2).....	83
Gráfico 11 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo (comparando com GC 1)	87
Gráfico 12 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo (comparando com GC 2)	88
Gráfico 13 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo com o mesmo estabelecimento de 2014 (comparando com GC 1)	93
Gráfico 14 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo com o mesmo estabelecimento de 2014 (comparando com GC 2)	94
Gráfico 15 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 1)	96
Gráfico 16 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 2)	98
Gráfico 17 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 1)	102

Gráfico 18 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 2) 103

LISTA DE TABELAS

Tabela 1 — Resumo dos resultados da avaliação de impactos do rompimento sobre o tamanho do mercado de trabalho formal dos municípios atingidos	10
Tabela 2 — Resumo dos resultados da avaliação de impactos do rompimento sobre a população que antes do desastre estava empregada no mercado de trabalho formal nos municípios atingidos	12
Tabela 3 — Número de trabalhadores ocupados segundo dados da PNAD Contínua e de vínculos de trabalho formal a partir de dados da RAIS	18
Tabela 4 — Salário médio de trabalhadores ocupados segundo dados da PNAD Contínua e de vínculos de trabalho formal a partir de dados da RAIS	20
Tabela 5 — Testes de diferença de médias entre grupos atingido e de comparação ..	52
Tabela 6 — Média e desvio-padrão do número de vínculos ativos por grupo	55
Tabela 7 — Média e desvio-padrão da massa salarial por grupo	61
Tabela 8 — Número de pessoas da coorte com vínculo de trabalho ativo por grupo..	70
Tabela 9 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano por grupo	80
Tabela 10 — Média e desvio-padrão do salário em dezembro do ano de pessoas com vínculo ativo por grupo.....	86
Tabela 11 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano com o mesmo estabelecimento de 2014 por grupo	92
Tabela 12 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano no mesmo setor de atividade econômica de 2014 por grupo	95
Tabela 13 — Média e desvio-padrão do salário em dezembro do ano de pessoas com vínculo ativo no mesmo setor de 2014 por grupo.....	100

SUMÁRIO

SUMÁRIO EXECUTIVO	9
1 INTRODUÇÃO	16
1.1 Características gerais do mercado de trabalho formal de Minas Gerais e Espírito Santo	17
1.2 Possíveis impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal	20
1.3 Evidências de impactos sobre emprego e salários em contexto de desastres...	24
2 FONTES DE DADOS UTILIZADAS	27
2.1 Relação Anual de Informações Sociais (RAIS – Ministério da Economia – ME) ..	27
2.2 Censo Demográfico de 2010 (IBGE)	29
2.3 Estimativa da população residente nos municípios brasileiros (IBGE)	30
2.4 Cadastro Nacional de Endereços para Fins Estatísticos – CNEFE (IBGE) ...	30
2.5 Bases cartográficas	31
3 AVALIAÇÃO DE IMPACTOS SOBRE O MERCADO DE TRABALHO FORMAL DOS MUNICÍPIOS ATINGIDOS	32
3.1 Metodologia	32
3.2 Resultados	55
4 AVALIAÇÃO DE IMPACTOS SOBRE A POPULAÇÃO FORMALMENTE EMPREGADA NOS MUNICÍPIOS ATINGIDOS ANTES DO DESASTRE	66
4.1 Metodologia	67
4.2 Resultados	78
5 CONSIDERAÇÕES FINAIS	106
REFERÊNCIAS	112
APÊNDICE A.....	115
APÊNDICE B.....	123
APÊNDICE C.....	130
APÊNDICE D.....	134

SUMÁRIO EXECUTIVO

Neste relatório são apresentados os resultados da avaliação de impactos causados pelo rompimento da Barragem de Fundão sobre o mercado de trabalho formal da região atingida. Foram utilizados os dados da Relação Anual de Informações Sociais – RAIS (Ministério da Economia – ME), tendo a análise compreendido o período de 2010 a 2017 e abordado dois enfoques distintos: (i) o tamanho do mercado de trabalho formal dos municípios atingidos; e (ii) trajetória no mercado de trabalho formal das pessoas que, antes do desastre, estavam empregadas nos municípios atingidos. O primeiro enfoque visa estimar o impacto sobre o número total de vínculos de empregos formais de trabalho e sua massa salarial na região atingida. Isto porque o impacto pode se dar através de destruição de vínculos formais previamente existentes bem como através da interrupção de abertura de novos vínculos. Foi estimado o efeito líquido destes dois movimentos. O segundo enfoque visa estimar o impacto sobre a manutenção do vínculo formal e salários daqueles trabalhadores e daquelas trabalhadoras que tinham vínculo formal de emprego na região atingida no período imediatamente anterior ao rompimento. Enquanto o primeiro enfoque analisa o impacto agregado no mercado de trabalho formal, o segundo enfoque analisa o impacto sobre um trabalhador formal típico da região atingida.

Para avaliar os impactos do rompimento sobre o tamanho do mercado de trabalho formal dos municípios atingidos, usou-se dados agregados por municípios, sendo os indicadores de impacto de interesse o número de postos formais de trabalho ocupados e a massa salarial. Já para avaliar os impactos do rompimento sobre a população que antes do desastre estava empregada no mercado de trabalho formal nos municípios atingidos, foram usados dados individuais que permitiram analisar a trajetória no mercado de trabalho formal dessas pessoas, utilizando-se indicadores como variável indicadora de vínculo empregatício ativo neste mercado, salário condicional a ter vínculo ativo, variável indicadora de vínculo ativo com o mesmo estabelecimento que tinha em 2014, variável indicadora de vínculo ativo no mesmo setor de atividade econômica em que trabalhava em 2014 e salário condicional a ter vínculo ativo no mesmo setor de atividade econômica em que trabalhava em 2014.

Este relatório adota métodos econométricos de avaliação de impactos, que têm como objetivo investigar relações de causa e efeito entre o rompimento da Barragem de Fundão e o comportamento demonstrado por indicadores quantitativos em momentos posteriores ao do rompimento. Especificamente, utilizou-se um modelo de diferença-em-diferenças que permite a investigação de impactos heterogêneos entre os anos de

2015 a 2017. Neste método, em essência, a evolução ao longo do tempo do indicador de impacto em análise é comparada entre o grupo atingido (formado pelos 45 municípios atingidos ou indivíduos neles) e grupos de comparação (formados por municípios não atingidos ou indivíduos neles) com o objetivo de produzir evidências quantitativas da existência, direção e magnitude do impacto causal do rompimento. Foram utilizados dois diferentes grupos de comparação com o intuito de representar a situação contrafactual do grupo atingido (isto é, como ele estaria caso não tivesse ocorrido o rompimento). São eles: (i) municípios não atingidos de Minas Gerais, Espírito Santo, sul da Bahia e norte do Rio de Janeiro pareados aos atingidos; e (ii) municípios não atingidos da Bahia, Goiás, Mato Grosso do Sul, Rio de Janeiro e São Paulo pareados aos atingidos. A seleção dos municípios que fazem parte de cada grupo de comparação se deu a partir de um processo de pareamento baseado em características observáveis pré-rompimento dos municípios.

A Tabela 1 e a Tabela 2 apresentam, respectivamente, os resumos dos resultados para a avaliação de impactos sobre o tamanho do mercado de trabalho formal dos municípios atingidos e para a avaliação de impactos sobre a população que antes do desastre estava empregada no mercado de trabalho formal nos municípios atingidos. Em ambas as tabelas é possível verificar o impacto estimado para os anos pós-rompimento quando se utiliza na avaliação cada um dos grupos de comparação considerados. Nelas, os símbolos “*”, “**” e “***” referem-se aos níveis de significância de 10%, 5% e 1% respectivamente, indicando a qual nível a variável que mede o impacto é estatisticamente diferente de 0, caso se aplique.

Tabela 1 — Resumo dos resultados da avaliação de impactos do rompimento sobre o tamanho do mercado de trabalho formal dos municípios atingidos

Indicador de impacto	Grupo de comparação	2015	2016	2017
Logaritmo natural do número de vínculos ativos	GC 1	-0,01	-0,01	-0,02
	GC 2	-0,03*	-0,04**	-0,04**
Logaritmo natural da massa salarial	GC 1	-0,02	-0,01	-0,01
	GC 2	-0,03	-0,05*	-0,05

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Considerando a avaliação de impactos sobre o tamanho do mercado de trabalho formal (Tabela 1), nota-se que, no caso do indicador “logaritmo natural do número de vínculos ativos” e considerando o GC 2 como grupo de comparação, o impacto estimado é robusto em termos de magnitude e persistente ao longo dos anos analisados. Tal impacto significa que em 2015 o grupo atingido perdeu, em média, cerca de 300 vínculos

a mais por município, desde 2014, do que perderia se tivesse seguido trajetória similar à dos municípios do GC 2, que representa sua situação contrafactual. Nos anos de 2016 e 2017, estimou-se uma perda média de vínculos ativos desde 2014 devido ao rompimento de 420 e 455 por município, respectivamente. Considerando a totalidade dos 45 municípios atingidos, estima-se que a perda de postos de trabalho devido ao rompimento da Barragem de Fundão foi de aproximadamente 13.500 em 2015, 18.900 em 2016 e 20.475 em 2017.

Ao analisar os resultados para o indicador “logaritmo natural da massa salarial”, é possível verificar que as estimativas de impacto foram sempre negativas, porém estatisticamente significantes ao nível de 10% somente em 2016 na comparação com o GC 2. Para esse ano, estimou-se que a massa salarial foi, em média, R\$ 1,3 milhões (em R\$ de dezembro de 2017) menor por município atingido do que seria a trajetória destes se tivesse seguido desde 2014 a da massa salarial média do GC 2. Tomando como base esta estimativa, a perda da massa salarial para o conjunto dos 45 municípios atingidos seria da ordem de R\$ 58,5 milhões, também em R\$ de dezembro de 2017.

Os resultados referentes à análise no nível município, resumidos na Tabela 1, não mostram evidências de impacto do rompimento sobre os indicadores de impacto analisados do grupo atingido sob os níveis de significância estatística convencionalmente adotados quando se utiliza na análise o grupo de comparação 1 (GC 1). Porém, ao analisar os resultados quando se utiliza o grupo de comparação 2 (GC 2), verifica-se a existência de impactos negativos do rompimento e estatisticamente diferentes de 0 sobre o grupo atingido em ambos os indicadores. A diferença entre os resultados obtidos ao usar um ou outro grupo de comparação levanta a possibilidade de que o impacto do rompimento não se restringiu aos municípios atingidos e afetou também o mercado formal de trabalho de outras localidades, podendo-se conjecturar que os municípios fora do grupo atingido mais suscetíveis a repercutir o impacto do rompimento seriam aqueles que possuem relações de interdependência econômica mais fortes com os atingidos, os quais provavelmente se localizam também em MG e ES. Essa possibilidade de “transbordamento” do impacto do rompimento foi justamente a motivação por trás da construção do GC 2. De fato, a análise em FGV (2020a) sugere que o impacto do rompimento sobre o PIB transcende os municípios da sub-bacia do Rio Doce e afeta também outras regiões de MG e ES. Não surpreende, portanto, que os resultados apresentados aqui não demonstrem vestígios de efeito sobre o mercado de trabalho formal do grupo atingido frente ao GC 1, pois há motivos para se acreditar que também os mercados dos municípios no GC 1 tenham, em média, sido prejudicados pelo rompimento.

Tabela 2 — Resumo dos resultados da avaliação de impactos do rompimento sobre a população que antes do desastre estava empregada no mercado de trabalho formal nos municípios atingidos

Indicador de impacto	Grupo de comparação	2015	2016	2017
Vínculo ativo no ano	GC 1	-0,01	-0,01	-0,01
	GC 2	-0,01	-0,01*	-0,01
Salário condicional a vínculo ativo no ano	GC 1	-31,61	-26,58	-16,01
	GC 2	-72,47*	-90,82*	-107,13*
Vínculo ativo no ano com o mesmo estabelecimento de 2014	GC 1	-0,01	0	-0,01
	GC 2	-0,01	0	0
Vínculo ativo no ano no mesmo setor de atividade econômica de 2014	GC 1	-0,01	-0,01	-0,01
	GC 2	-0,01	-0,01	-0,01
Salário condicional a vínculo ativo no ano no mesmo setor de atividade econômica de 2014	GC 1	-34,9	-37,07	-24,78
	GC 2	-72,8	-95,48**	-105,91*

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Os impactos estimados referentes à análise no nível indivíduo são exibidos na Tabela 2. Quanto à probabilidade de pessoas do grupo atingido terem vínculo de trabalho ativo no mercado de trabalho formal nos anos seguintes ao rompimento, nota-se que esta não se alterou, em média, frente à das pessoas do grupo de comparação seja ele o GC 1 ou o GC 2. Isto porque as estimativas do modelo, apesar de todas negativas, não são estatisticamente diferentes de 0 ao nível de 10% ou, quando o são, mostram-se demasiadamente sensíveis à maneira como o grupo de comparação é construído. Uma exceção relevante, no entanto, diz respeito às pessoas da coorte que em 2014 trabalhavam no setor de mineração (não exibido na tabela): nesse subgrupo, os modelos estimam de forma robusta que a probabilidade de se ter vínculo de trabalho ativo no mercado formal diminuiu após o rompimento. Particularmente em 2017, 1.506 das 3.872 pessoas do grupo atingido que em 2014 atuavam no setor de mineração não tinham vínculo ativo, número que seria menor se o rompimento não tivesse ocorrido: 1.052 usando o GC 1 como representação do contrafactual dos atingidos e 1.001 usando o GC 2.

Levando em consideração o resultado geral para a probabilidade de vínculo ativo no ano (Tabela 2) conjuntamente com o que se estimou para o número de vínculos ativos do mercado de trabalho formal dos municípios atingidos (Tabela 1), tem-se alguma evidência de que o impacto do rompimento parece ter reduzido o número de vínculos não tanto pela extinção de postos de trabalho, mas sim por meio da inviabilização da abertura de novas vagas. Afinal, as estimativas para o grupo atingido na coorte como um todo não apontam para uma mudança na probabilidade de se ter vínculo ativo nos anos posteriores ao rompimento que divergisse de maneira estatisticamente significativa

da dos grupos de comparação. Em outras palavras, a taxa a que as pessoas do grupo atingido deixam de ter vínculo de trabalho ativo no mercado formal parece ter permanecido similar após 2014 à dos grupos de comparação, os quais se supõe não terem sido afetados pelo rompimento. Logo, se os municípios atingidos têm menos vínculos ativos em relação ao GC 2 após o rompimento, as pessoas que antes lá estavam empregadas não saíram do mercado formal de trabalho em proporção maior do que as que antes trabalhavam nos GC 1 ou 2, então o fenômeno que o rompimento parece ter desencadeado nos municípios atingidos é o de impedir que novos postos formais de trabalho se abrissem. É um aspecto importante do impacto sobre o mercado de trabalho formal, porque se pode especular que, devido a ele, o rompimento tenha afetado principalmente os mais jovens, em especial os que só viriam a ingressar no mercado de trabalho após 2015, os quais teriam se deparado com menos oportunidades de emprego formal do que haveria se o rompimento não tivesse acontecido.

Com relação ao indicador de impacto de salário para os indivíduos que têm vínculo ativo ao longo dos anos, as estimativas se revelam negativas para os anos posteriores ao rompimento, porém estatisticamente significantes ao nível de 10% somente quando da comparação com o GC 2. Estima-se que a remuneração dos indivíduos atingidos que tiveram vínculo formal ativo em 2015 se reduziu R\$ 72 em média, agravando-se para uma queda de R\$ 91 em 2016 e de R\$ 107 em 2017, frente ao que se observava em 2014 (valores em R\$ de dezembro de 2017). Ainda considerando o GC 2 como representação da situação contrafactual do grupo atingido, os modelos estimados em subgrupos de pessoas do grupo atingido revelam efeitos heterogêneos do rompimento. Dentre eles, as estimativas sugerem que o impacto do rompimento foi mais intenso sobre o salário de pessoas que antes do desastre trabalhavam em estabelecimentos mais próximos ao Rio Gualaxo do Norte, Rio do Carmo e Rio Doce. Em 2017, por exemplo, o subgrupo de indivíduos da coorte que em 2014 estava empregada em estabelecimentos localizados em setores censitários a até 10 km do rio teve a perda salarial média estimada em R\$ 131.

No caso dos impactos sobre a probabilidade de ter vínculo ativo no ano com o mesmo estabelecimento de 2014 e de probabilidade de ter vínculo ativo no ano no mesmo setor de atividade econômica de 2014, é possível verificar que nenhuma das variáveis que mede os impactos do rompimento pode ser considerada estatisticamente diferente de 0 aos níveis de significância convencionalmente adotados, independentemente do grupo de comparação utilizado na avaliação (GC 1 ou GC 2). Isto é, não se pode rejeitar a hipótese de que o impacto do rompimento sobre tais indicadores tenha sido nulo quando se utilizam os grupos de comparação propostos. No entanto, ao analisar a presença de

possível heterogeneidade nos impactos do rompimento nesse indicador entre subgrupos específicos, a comparação entre o grupo atingido e o GC 2 mostra que houve uma queda na probabilidade de possuir vínculo ativo no mesmo setor de atividade econômica para os indivíduos que trabalhavam em 2014 no setor de agropecuária e pesca. Estima-se que o rompimento da Barragem de Fundão levou 1.593 pessoas que antes trabalhavam em agropecuária e pesca a não conseguir se manter nesse mesmo setor com um vínculo de trabalho formal ativo em 2017.

Para o último indicador de impacto analisado, o salário condicional a possuir vínculo ativo no mesmo setor de atividade econômica de 2014, a comparação entre grupo atingido e GC 1 não aponta a existência de impacto do rompimento que, em média, seja estatisticamente significativa ao nível de 10%. Entretanto, quando se utiliza o GC 2 como representação da situação contrafactual de atingidos, observa-se que o salário foi R\$ 95 e R\$ 106 menor em média para os indivíduos do grupo atingido do que seria sem o rompimento em 2016 e 2017, respectivamente, efeitos que se mostram estatisticamente diferentes de 0 ao nível de significância de 10%. Na comparação com o GC 2, foram encontrados ainda indícios de existência de impactos heterogêneos do rompimento em subgrupos dos indivíduos atingidos da coorte. Dentre eles, estimou-se que a redução de salário dos homens que tiveram vínculo ativo no mesmo setor de atividade de 2014 foi mais intensa que a estimada para todos os atingidos, sendo estatisticamente significativa ao nível de 10% em 2015 e 2016, embora não mais em 2017. Para as mulheres o impacto também foi negativo, no entanto persistiu até 2017, ano em que foi estatisticamente significativa e alcançou magnitude maior em valor absoluto que o estimado para todo o grupo atingido. Outro efeito heterogêneo encontrado refere-se à localização do setor censitário do estabelecimento em que as pessoas da coorte estavam empregadas em 2014, a qual parece ser um fator associado à intensidade dos impactos do rompimento. Estimou-se que, em média, o salário das pessoas com vínculo ativo no mesmo setor de atividade econômica de 2014 teve perdas maiores para aquelas que trabalhavam relativamente próximo aos rios antes do rompimento.

Por fim, é importante ressaltar que os resultados apresentados neste relatório se referem apenas aos impactos do rompimento no mercado de trabalho formal e em um horizonte temporal pós-rompimento de apenas três períodos (31 de dezembro de 2015, 2016 e 2017). Estas duas restrições são devido à inexistência de dados que permitam a realização de uma avaliação de impactos similar para o mercado de trabalho informal e à indisponibilidade de dados mais recentes da RAIS que permitam ampliar o período de análise. Dessa forma, é possível que os impactos do rompimento da Barragem de Fundão sobre emprego e renda do trabalho sejam ainda maiores caso se considere os

efeitos no mercado de trabalho informal e em períodos mais longos (até 2020, por exemplo). Com efeito, os resultados obtidos em FGV (2020b) mostram indícios de que o impacto do rompimento da Barragem de Fundão sobre a renda agregada de Minas Gerais e Espírito Santo considerados conjuntamente se encontrava em trajetória de crescimento durante período de 2015 a 2017, o que vai ao encontro dessa conjectura.

1 INTRODUÇÃO

Este relatório apresenta os resultados de uma avaliação de impactos do rompimento da Barragem de Fundão, ocorrido em 5 de novembro de 2015, sobre o mercado de trabalho formal da região atingida até o ano de 2017.

São utilizadas duas abordagens distintas, que fazem uso de dados agregados no nível de municípios e de dados individuais. A primeira tem como objetivo avaliar os impactos do rompimento sobre o tamanho do mercado de trabalho formal dos municípios atingidos a partir do número de postos formais de trabalho ocupados e da massa salarial. Já a segunda abordagem procura avaliar os impactos do rompimento sobre a população que antes do desastre estava empregada no mercado de trabalho formal nos municípios atingidos. Nesse caso, avaliam-se os impactos do rompimento a partir da análise da trajetória no mercado de trabalho formal dessas pessoas através de indicadores como: variável indicadora de vínculo ativo neste mercado, salário condicional a ter vínculo ativo, variável indicadora de vínculo ativo com o mesmo estabelecimento que tinha em 2014, variável indicadora de vínculo ativo no mesmo setor de atividade econômica em que trabalhava em 2014 e salário condicional a ter vínculo ativo no mesmo setor de atividade econômica em que trabalhava em 2014.

Nas duas abordagens, são utilizadas informações provenientes da Relação Anual de Informações Sociais (RAIS – Ministério da Economia – ME)¹ e a metodologia empregada é a de avaliação econométrica de impactos, especificamente, um modelo de diferença-em-diferenças que permite a investigação de impactos heterogêneos entre os anos de 2015 a 2017. Neste método, em essência, a evolução ao longo do tempo do indicador de impacto em análise é comparada entre regiões atingidas (grupo atingido²) e não atingidas (grupos de comparação) com o objetivo de produzir evidências quantitativas

¹ A Relação Anual de Informações Sociais (RAIS) é um instrumento que visa auxiliar no controle da atividade trabalhista no país, fornecendo informações sobre o mercado de trabalho e dados para a elaboração de estatísticas do trabalho a entidades governamentais.

² O grupo de municípios atingidos é composto por 45 municípios, sendo eles: (i) Em Minas Gerais: Aimorés, Alpercata, Barra Longa, Belo Oriente, Bom Jesus do Galho, Bugre, Caratinga, Conselheiro Pena, Córrego Novo, Dionísio, Fernandes Tourinho, Galileia, Governador Valadares, Iapu, Ipaba, Ipatinga, Itueta, Mariana, Marliéria, Naque, Periquito, Pingo d'Água, Ponte Nova, Raul Soares, Resplendor, Rio Casca, Rio Doce, Santa Cruz do Escalvado, Santana do Paraíso, São Domingos do Prata, São José do Goiabal, São Pedro dos Ferros, Sem-Peixe, Sobrália, Timóteo e Tumiritinga; (ii) No Espírito Santo: Aracruz, Baixo Guandu, Colatina, Conceição da Barra, Fundão, Linhares, Marilândia, Serra e São Mateus. Ressalta-se que essa lista de municípios atingidos foi utilizada para a produção do presente relatório, não configurando uma lista final de municípios atingidos que serão utilizados nos demais estudos produzidos pela FGV. Ressalta-se ainda que essa lista de municípios atingidos não configura um posicionamento da FGV quanto à delimitação de áreas atingidas para qualquer finalidade.

da existência, direção e magnitude do impacto causal do rompimento – isto é, investiga-se a existência de uma relação de causa e efeito. Foram utilizados dois grupos de comparação com o intuito de representar a situação contrafactual do grupo atingido (isto é, como ele estaria caso não tivesse ocorrido o rompimento). São eles: (i) municípios não atingidos de Minas Gerais, Espírito Santo, sul da Bahia e norte do Rio de Janeiro pareados aos atingidos; e (ii) municípios não atingidos da Bahia, Goiás, Mato Grosso do Sul, Rio de Janeiro e São Paulo pareados aos atingidos.³

Este capítulo de introdução é formado por três seções. Na seção 1.1, são expostas informações sobre características gerais do mercado de trabalho formal de Minas Gerais e Espírito Santo. Na seção 1.2, discutem-se os possíveis impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal e os potenciais canais de causalidade envolvidos. Ao final deste capítulo, na seção 1.3, são apresentadas evidências de impactos sobre emprego e renda em contexto de desastres naturais e tecnológicos. O relatório é composto ainda por mais quatro capítulos. No Capítulo 2, são apresentados os dados utilizados, detalhando as respectivas fontes e propósitos dessas informações na análise aqui realizada. No Capítulo 3, apresenta-se a metodologia e os resultados da avaliação de impactos sobre o mercado de trabalho formal dos municípios atingidos. O Capítulo 4, por sua vez, trata da avaliação de impactos sobre a população formalmente empregada nos municípios atingidos antes do desastre. Por fim, no Capítulo 5 são feitas as considerações finais.

1.1 Características gerais do mercado de trabalho formal de Minas Gerais e Espírito Santo

Com o objetivo de ilustrar a representatividade do trabalho formal frente ao informal, esta seção apresenta uma breve análise sobre o mercado de trabalho de Minas Gerais e Espírito Santo, assim como de todo o Brasil. Para tanto, faz-se uso dos dados da Relação Nacional de Informações Sociais (RAIS – Ministério da Economia – ME), os quais são descritos em mais detalhes ao longo deste relatório, e da Pesquisa Nacional por Amostra Domiciliar Contínua (PNAD Contínua), realizada pelo IBGE (IBGE, 2020). As duas fontes de dados registram rendimentos e ocupação dos trabalhadores, porém o fazem a partir de óticas diferentes. A PNAD Contínua é uma pesquisa amostral baseada em visitas em domicílio e realizada diretamente com seus residentes, enquanto

³ A seleção dos municípios que fazem parte de cada grupo de comparação se deu a partir de um processo de pareamento baseado em características observáveis pré-rompimento dos municípios. Foi utilizada a distância de Mahalanobis, tendo sido selecionados três “vizinhos mais próximos” para cada município atingido, sem reposição.

a RAIS, um registro administrativo, é declarada pelos estabelecimentos empregadores formalizados, motivo pelo qual sua base de dados é limitada aos indivíduos que possuem vínculo de trabalho formal.

A Tabela 3 apresenta o total de trabalhadores ocupados no 4º trimestre de 2014 e no 4º trimestre de 2017, tanto no geral quanto apenas no mercado formal, segundo dados da PNAD Contínua.⁴ Além disso, contém o total de vínculos formais de trabalho com remuneração não nula vigentes em 31 de dezembro dos mesmos anos segundo dados da RAIS.⁵ Note que os números referentes a trabalho formal são similares entre PNAD Contínua e RAIS, o que aumenta a credibilidade das duas pesquisas no que diz respeito a retratar o mercado de trabalho formal, uma vez que partem de perspectivas distintas sobre ele. Ainda, a PNAD Contínua permite verificar que mais da metade dos trabalhadores ocupados em 2014 possuía um emprego formal – aproximadamente 53% em ES e 52,4% em MG e no Brasil. Em 2017, no entanto, nota-se uma queda no número absoluto de trabalhadores formais nos três recortes territoriais em comparação com 2014, e as proporções citadas caem de 52,4% e 53% para 48,9% e 48,2% em MG e ES, respectivamente, e de 52,4% para 49% no Brasil como um todo. Tal redução no número de vínculos formais é captada também pela RAIS, segundo a qual há queda na ordem de 350 mil vínculos em MG, de 90 mil em ES e de 3,2 milhões no Brasil.

Tabela 3 — Número de trabalhadores ocupados segundo dados da PNAD Contínua e de vínculos de trabalho formal a partir de dados da RAIS

Ano	Base	Grupo	MG	ES	Brasil
2014	PNAD Contínua	Todos	9.647.603	1.721.227	90.651.430
		Formais	5.059.117	912.640	47.515.658
	RAIS	Formais	4.882.199	931.551	47.614.527
2017	PNAD Contínua	Todos	9.707.635	1.750.113	89.722.611
		Formais	4.748.058	843.853	44.058.068
	RAIS	Formais	4.536.052	840.966	44.369.819

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME) e da PNAD Contínua (IBGE).

Observação: Os vínculos de trabalho formal da RAIS considerados para construção das variáveis são apenas os de pessoas que constam como empregadas em 31 de dezembro do ano indicado com remuneração não nula. A categoria “todos” representada com dados da PNAD Contínua refere-se a trabalhadores formais (com carteira de trabalho assinada, militares e estatutários) e informais (sem carteira de trabalho assinada e conta própria).

⁴ Em rigor, trata-se do número de pessoas ocupadas na semana de referência da pesquisa, a qual é compreendida como a semana de sábado a sábado que precede imediatamente a semana em que foi realizada a entrevista (IBGE, 2020). A divulgação da PNAD Contínua usada, por sua vez, é a referente ao 4º trimestre dos anos indicados.

⁵ Não necessariamente o número de vínculos de trabalho na RAIS corresponde ao número de trabalhadores, uma vez que uma pessoa pode ter ao mesmo tempo mais de um emprego formal.

A Tabela 3 deixa claro que o mercado de trabalho formal encolheu entre 2014, ano que precede o rompimento da Barragem de Fundão, e 2017, que é posterior ao rompimento. Ainda, ter o número de vínculos formais reduzido em mais de três milhões no Brasil como um todo, valor que supera as quedas em MG e ES somadas, sugere que fatores além do rompimento agiram no sentido de desaquecer o mercado formal de trabalho.⁶ Isso explicita que a avaliação de impactos do rompimento da barragem não pode se limitar a constatar se o nível de emprego formal se reduziu ou não em municípios atingidos, mas sim investigar se o rompimento teve influência sobre o mercado de trabalho além daquela provocada por fatores não relacionados ao desastre.

A Tabela 4 segue o modelo da tabela anterior, mas com foco no salário médio conforme captado pela PNAD Contínua e pela RAIS. Os valores apresentados foram todos convertidos para R\$ de dezembro de 2017 por meio do Índice de Preços ao Consumidor Amplo (IPCA), produzido pelo IBGE (IBGE, 2019a). Note que, diferentemente da Tabela 3, a representação do emprego formal pela PNAD Contínua e pela RAIS apresenta discrepâncias mais perceptíveis, as quais potencialmente se devem à origem distinta dos dados de cada base. Por ser declarado pelos estabelecimentos empregadores, é possível que o salário na RAIS consista em valores brutos (sem deduções de imposto de renda ou contribuição ao INSS)⁷, enquanto a remuneração na PNAD Contínua tenda a refletir valores líquidos (declarados pelos próprios indivíduos).⁸

⁶ De fato, a desaceleração do crescimento do PIB brasileiro se inicia em 2014, estando o Brasil em recessão econômica de 2015 até metade de 2017. A série histórica da taxa de crescimento do PIB está disponível em: <https://www.ibge.gov.br/estatisticas/economicas/contas-nacionais/9300-contas-nacionais-trimestrais.html?=&t=series-historicas&utm_source=landing&utm_medium=explica&utm_campaign=pib#evolucao-taxa> Acesso em: 1 dez 2020.

⁷ Os valores que compõem a remuneração declarada na RAIS são discriminados em BRASIL (2020b).

⁸ As informações captadas pela PNAD Contínua são ainda sujeitas a viés de subdeclaração, algo comum em pesquisas baseadas em questionários (*surveys*). Souza (2016, p. 154-155) discute alguns dos principais problemas de representação da renda por esse tipo de pesquisas.

Tabela 4 — Salário médio de trabalhadores ocupados segundo dados da PNAD Contínua e de vínculos de trabalho formal a partir de dados da RAIS

Ano	Base	Grupo	MG	ES	Brasil
2014	PNAD	Todos	1.881,75	2.066,65	2.100,28
	Contínua	Formais	1.970,20	2.102,29	2.316,41
	RAIS	Formais	2.529,09	2.687,16	2.965,85
2017	PNAD	Todos	1.804,00	1.923,41	2.064,59
	Contínua	Formais	2.003,47	2.066,67	2.387,91
	RAIS	Formais	2.534,10	2.624,35	2.973,23

Fonte: Elaboração própria (2020) a partir de dados da RAIS (ME), da PNAD Contínua (IBGE) e da série histórica do IPCA (IBGE).

Observação: Os vínculos de trabalho formal da RAIS considerados para a construção das variáveis são apenas os de pessoas que constam como empregadas em 31 de dezembro do ano indicado com remuneração não nula. O salário na RAIS corresponde ao de dezembro do ano indicado. A categoria “todos” representada com dados da PNAD Contínua refere-se a trabalhadores formais (com carteira de trabalho assinada, militares e estatutários) e informais (sem carteira de trabalho assinada e conta própria). Valores são apresentados em R\$ de dezembro de 2017.

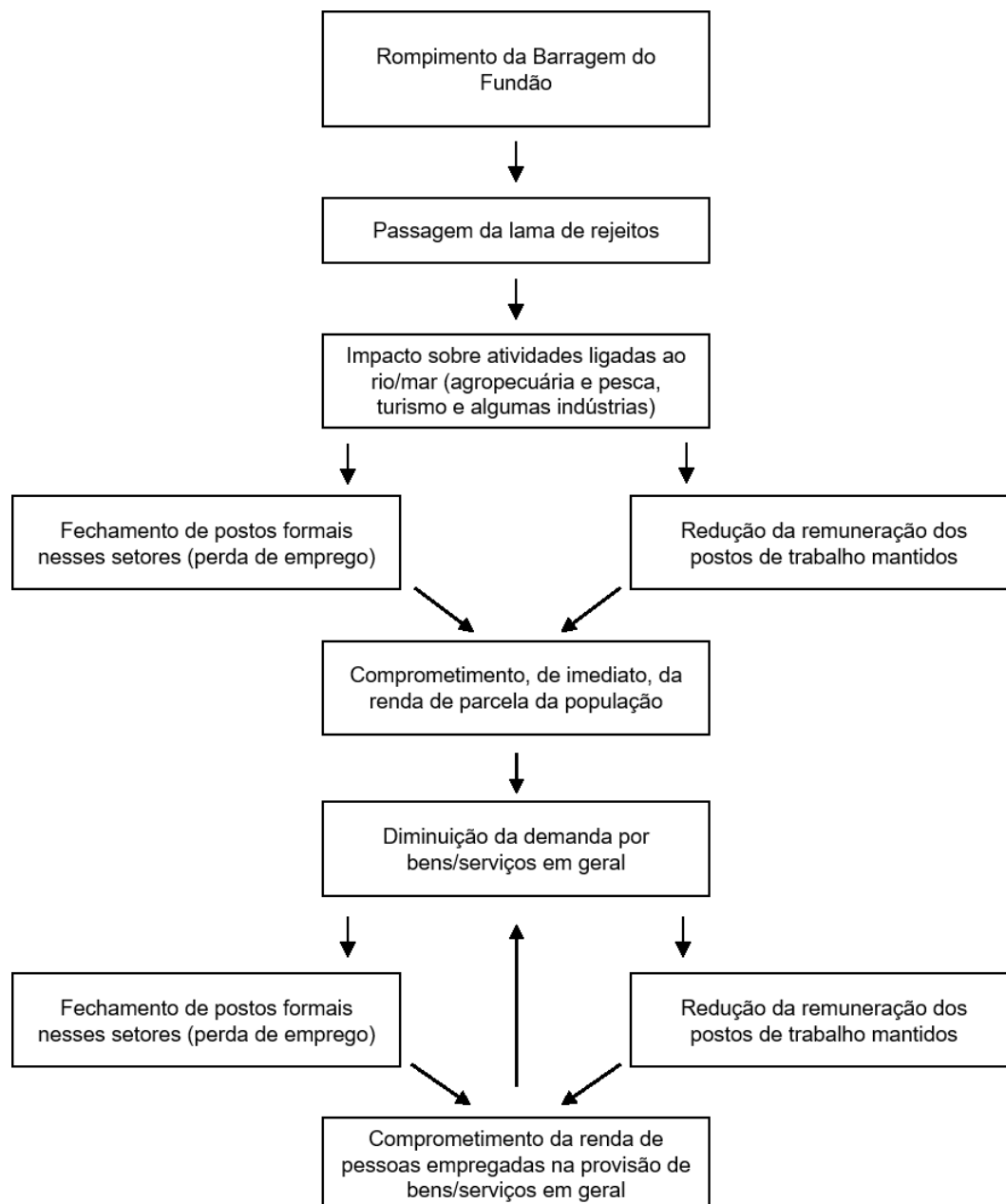
Não obstante a diferença entre PNAD Contínua e RAIS, a Tabela 4 mostra que o salário tanto no geral quanto apenas no setor formal de MG e ES é inferior à média nacional, independentemente da fonte de dados ou do ano analisado. Além disso, o salário médio de todos os trabalhadores se reduziu em termos reais entre 2014 e 2017 nos três recortes regionais, sendo a queda mais acentuada observada no ES, onde ela supera R\$ 140 (redução de 6,9%). Por outro lado, o mesmo não pode ser dito sobre o salário médio no setor formal, os quais tiveram pequeno crescimento em MG e no Brasil e uma redução mais branda no ES, movimentos captados pela PNAD Contínua e pela RAIS. A tabela sugere, então, que as remunerações do mercado formal não apenas conseguiram se manter em patamar superior à média geral, mas também foram menos afetadas pelos fatores que levaram à redução do poder de compra dos trabalhadores em geral. Novamente, fica evidente que diversos fenômenos exerceram influência sobre o mercado de trabalho como um todo desde o rompimento da Barragem de Fundão. Isso enfatiza a necessidade de que a avaliação de impactos seja capaz de enxergar além de fatores não relacionados ao desastre para produzir um diagnóstico acertado dos efeitos exclusivos do rompimento sobre o mercado de trabalho formal dos municípios atingidos.

1.2 Possíveis impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal

Para auxiliar a compreensão das hipóteses investigadas neste relatório, é importante examinar como o rompimento da Barragem de Fundão pode ter causado alterações no

mercado de trabalho formal nos municípios atingidos. Para esse propósito, a Figura 1 ilustra os canais de causalidade através dos quais eventuais impactos do rompimento sobre o emprego formal podem plausivelmente ter se dado.

Figura 1 — Fluxograma ilustrando os potenciais canais e possíveis impactos do rompimento da Barragem de Fundão no mercado de trabalho formal dos municípios atingidos



Fonte: Elaboração própria (2020).

O rompimento da Barragem de Fundão e a consequente passagem da lama de rejeitos podem ter levado à contaminação do Rio Doce e corpos d'água adjacentes, bem como

do solo da área diretamente atingida por ela. Uma eventual contaminação teria potencial para afetar diversas atividades econômicas ao longo da Bacia do Rio Doce que eram dependentes ou ligadas ao rio e, no caso do Espírito Santo, o mar, tais como as de agropecuária, pesca, turismo e mesmo determinados tipos de indústrias, como hidroeletricidade e processamento de celulose.⁹

Sob o cenário de afetação das atividades desses setores pelo rompimento, uma consequência possível seria o fechamento de postos formais de trabalho existentes neles, o que se faria refletir em aumento do desemprego e/ou do emprego informal na população, por exemplo. Além disso, mesmo os postos formais de trabalho que não acabassem fechados em decorrência do rompimento poderiam ser afetados através da redução da remuneração a eles associada. Esses dois canais convergem no sentido de significar uma potencial perda de poder de compra por parte dos trabalhadores formais desses setores.

Com uma parcela da população passando por um processo de redução de renda, poderia se fazer sentir uma queda no consumo por bens e serviços em geral, consequentemente afetando outros setores de atividade, como comércio varejista e indústrias diversas. Frente à redução da demanda pelos bens e serviços que oferecem, os estabelecimentos formais desses setores se veriam forçados a agir de forma similar à descrita acima com relação aos setores afetados imediatamente pelo rompimento: fechando postos de trabalho e/ou reduzindo a remuneração dos postos que mantêm abertos. Isso significaria o comprometimento da renda de pessoas empregadas nesses demais setores, agravando a queda na demanda por bens e serviços produzidos nos municípios atingidos. Note que há um potencial de retroalimentação do impacto do rompimento entre os diversos setores de atividade econômica, onde o desaquecimento progressivo da economia afeta negativamente o mercado formal de trabalho, o que, por sua vez, deprime ainda mais a demanda por bens e serviços.¹⁰

⁹ A interrupção das atividades da Samarco após o rompimento da Barragem de Fundão representou um importante impacto sobre o setor minerador nas localidades em que a empresa atuava, como o município de Mariana. Embora isso se reflita no mercado de trabalho formal, entende-se que é um efeito de abrangência territorial restrita, uma vez que é relativamente pequena a participação do setor de mineração entre postos de trabalho formais na região atingida. Considerando que a Figura 1 pretende ilustrar um panorama geral o suficiente para contemplar efeitos que possam ocorrer em quaisquer pontos da região atingida, optou-se por não incluir nela o impacto do rompimento sobre a atividade de mineração. Não obstante, esse setor é considerado nas análises desenvolvidas no relatório, em especial nas do Capítulo 4.

¹⁰ Evidentemente, todos os impactos sobre o mercado de trabalho formal discutidos podem se aplicar também a pessoas que atuavam como profissionais autônomos ou em empregos sem carteira de trabalho assinada. Destaque foi dado ao trabalho formal por este ser o foco da avaliação de impactos deste relatório, o que se justifica pelo fato de que os dados necessários para a análise estão disponíveis apenas para o mercado formal.

Cabe ressaltar que as análises realizadas neste relatório servem ao propósito de fornecer insumos para responder se o rompimento da Barragem de Fundão causou impactos sobre o mercado de trabalho formal dos municípios atingidos, sendo que nem todos os potenciais canais de causalidade mencionados na Figura 1 são analisados de maneira específica. Como mencionado anteriormente, no Capítulo 3.1 é dado foco à avaliação dos impactos sobre o tamanho do mercado de trabalho formal a partir da análise de indicadores de número de postos formais de trabalho ocupados e da soma das remunerações pagas aos trabalhadores nesses postos. Já no Capítulo 4, é feito um estudo de coorte para investigar se as pessoas que possuíam um vínculo de trabalho formal nos municípios atingidos em 31 de dezembro de 2014 foram impactadas de formas ou com intensidades diferentes das estimadas para o mercado formal como um todo dos municípios atingidos.¹¹

Por fim, um aspecto importante a ser abordado é a existência de programas e ações reparatórias desenvolvidas pela Fundação Renova relacionadas aos impactos do rompimento sobre as economias locais, em particular, os programas que são monitorados pela Câmara Técnica de Economia e Inovação (CT-EI).¹² Neste relatório não foram feitas análises específicas sobre eles nem foram excluídos os possíveis efeitos que eles possam ter exercido sobre os indicadores de impacto aqui abordados, uma vez que tais análises não fazem parte do escopo deste estudo. Cabe mencionar, no entanto, que a maioria dos programas que já estão em andamento atingem setores/atividades econômicas e/ou localidades atingidas específicas, havendo ainda restrições adicionais sobre quais atingidos são elegíveis a participar em alguns casos. Assim, é razoável considerar que eventuais efeitos desses programas e ações de reparação em curso provavelmente são restritos a alguns subgrupos populacionais, provavelmente não sendo verificáveis quando se consideram os municípios atingidos e suas respectivas populações empregadas no mercado de trabalho formal como um todo, que é o que se analisa neste estudo.

¹¹ O termo “coorte” descreve um conjunto específico de indivíduos cuja trajetória registrada em dados é acompanhada ao longo do tempo. Um estudo de coorte retrospectivo seleciona indivíduos expostos e não expostos a um evento de interesse com o objetivo de avaliar como tal evento pode ter influenciado a trajetória daqueles por ele afetados (BRASIL, 2014). Este é o tipo de análise desenvolvida no Capítulo 4: acompanha-se ao longo dos anos a trajetória tanto de pessoas que tinham emprego formal em 31 de dezembro de 2014 em municípios atingidos pelo rompimento da Barragem de Fundão quanto daquelas que também tinham emprego formal em 31 de dezembro de 2014, mas em municípios não atingidos. A avaliação de impactos do rompimento é feita a partir de comparações entre esses dois grupos.

¹² São monitorados pela CT-EI os programas: 15 (Promoção da inovação), 16 (Retomada das atividades aquícolas e pesqueiras), 17 (Retomada das atividades agropecuárias), 18 (Desenvolvimento e diversificação econômica), 19 (Recuperação de micro e pequenos negócios), 20 (Estímulo à contratação local) e 42 (Ressarcimento de gastos públicos extraordinários).

1.3 Evidências de impactos sobre emprego e salários em contexto de desastres

A seguir, comenta-se brevemente alguns resultados obtidos por estudos que investigam impactos causados por desastres, sejam eles naturais ou tecnológicos, sobre emprego e salários nas regiões por eles atingidas. Tais evidências enfatizam a importância da mensuração de possíveis efeitos nessa dimensão para o caso do rompimento da Barragem de Fundão.

A avaliação de impactos feita no presente relatório relaciona-se de forma próxima ao trabalho de Groen et al. (2020), no qual os autores avaliam o impacto do desastre provocado pelos furacões Katrina e Rita, que atingiram a costa do Golfo do México (EUA) em 2005, sobre os mercados locais de trabalho. Um dos pontos em que o artigo se assemelha a este relatório é no uso de uma base de dados que permite acompanhar os indivíduos ao longo do tempo. Groen et al. (2020) definem uma coorte de indivíduos empregados no período anterior ao desastre e acompanham a evolução de seus rendimentos e status de emprego ao longo do tempo. A avaliação dos impactos dos furacões se dá a partir da comparação desse grupo de pessoas em áreas atingidas com um grupo não atingido, construído pelos autores a partir da seleção de condados não atingidos que são mais parecidos com os condados atingidos em características demográficas, composição industrial e rendimentos anteriores ao desastre. As estimativas do impacto foram obtidas através do uso de um modelo econométrico que compara a evolução da variável de interesse, antes e depois do desastre, entre o grupo atingido e o grupo não atingido. Os resultados do trabalho indicam, para o primeiro ano após o desastre, uma queda dos rendimentos (média de US\$ 298, por trimestre) daqueles afetados pelos furacões. No médio e longo prazo, a tendência dos rendimentos, no entanto, é de crescimento (média de US\$ 343 e US\$ 792, por trimestre), de modo que os trabalhadores afetados ganham mais do que os trabalhadores não afetados pelo desastre. Os autores concluem que a queda imediata nos rendimentos ocorre primariamente entre aqueles indivíduos da coorte que, por algum motivo, não mantiveram seus empregos após o desastre. Já o aumento agregado dos rendimentos, observado nos médio e longo prazos, ocorre pelo aumento dos salários daqueles que se mantiveram empregados, e não pelo aumento da proporção de indivíduos empregados após as tempestades. Os autores também apresentam análises agregadas em nível de condados, relatando uma forte queda da taxa de emprego nas áreas afetadas em relação as áreas de comparação no curto e médio prazo, enquanto no longo prazo as taxas de emprego tendem a apresentar um

crescimento similar. Os autores creditam esta recuperação da taxa de emprego a um aumento da demanda por trabalho após o desastre, impulsionado pelo setor de construção.

Em linha com Groen et al. (2020), Deryugina et al. (2018) estima os impactos do furacão Katrina sobre os trabalhadores que residiam em Nova Orleans (EUA) antes do desastre. Como no artigo descrito anteriormente, é aplicado um modelo econométrico que compara evolução ao longo do tempo do grupo atingido com a de grupo de comparação, permitindo analisar como os impactos variam ao longo do tempo. Para a construção do grupo de comparação, os autores escolhem 10 cidades não atingidas pelo desastre, sendo a seleção feita com base na similaridade delas com as atingidas em termos de características demográficas e de perfil de rendimentos nos anos anteriores ao desastre. Os resultados obtidos pelos autores apresentam semelhanças aos resultados de Groen et al. (2020): o impacto visto sobre a renda é transitório; estima-se uma perda média na renda do trabalho de US\$ 2.000 em um ano após o furacão, efeito que não é mais observado dois anos após o desastre. No longo prazo, os trabalhadores afetados pelo desastre apresentam rendimentos maiores que aqueles do grupo de comparação (US\$ 2.500, em média). O mesmo padrão é visto sobre a chance de não ter renda de trabalho (uma aproximação para desemprego): em 2006 e 2007, aqueles afetados pelo desastre possuíam chances maiores (4,2 e 2,1 pontos percentuais, respectivamente) do que aqueles não afetados, diferença que desaparece ao longo do tempo.

Em Martínez, Martínez e Romero-Jarén (2020), também são vistas estimativas que fazem distinção dos efeitos de um desastre nos curto e longo prazos. Os autores encontram impactos de curto prazo do terremoto Bío-Bío que ocorreu em 2010 no Chile sobre a probabilidade de estar empregado: quedas entre 2,5 e 4,2 pontos percentuais, a depender da abordagem econométrica utilizada. Além disso, estimam impactos negativos sobre o salário mensal de 3,9% a 6,4% (mais uma vez, utilizando abordagens diferentes de estimação). Em contrapartida, as estimativas de longo prazo (2011-2017) indicam que esses impactos adversos sobre emprego duraram menos de um ano após o terremoto, já que em 2011 observam-se impactos positivos na taxa de emprego. Assim como em Groen et al. (2020), os autores mencionam um aumento da demanda por trabalho, que estava reprimida nos períodos imediatamente posteriores ao desastre, e um aumento da migração em busca de trabalho como fatores explicativos para a recuperação da taxa de emprego. Adicionalmente, também é verificado pelos autores que a retomada da taxa de emprego ocorre em conjunto com um aumento de tipos de trabalhos temporários e com menor carga horária.

Outros autores tratam do impacto de desastres sobre o mercado de trabalho, porém com foco apenas em efeitos de curto prazo. Ribeiro et al. (2018) concluem que as chuvas e enchentes de 2008 no estado de Santa Catarina causaram perda de 33 mil postos de trabalho por ano entre 2008 e 2010, o que foi causado pelo efeito negativo de 5,13% na produção industrial mensal do estado. De modo similar, Lenzen et al. (2019) examinam os desdobramentos imediatos de impactos do ciclone tropical Debbie ocorrido em 2017 na Austrália. Eles estimam que, devido aos impactos negativos do ciclone sobre os setores mais afetados na região (turismo, carvão, agricultura e comércio), houve perda de 8.487 postos de trabalho em 2017, ao comparar com 2016.

Por outro lado, analisando somente um horizonte de tempo maior, Coffman e Noy (2011) avaliam o impacto de longo prazo do furacão Iniki ocorrido em 1992 em Kauai, Havaí (EUA), sobre a ilha. Os autores estimam, relativo a um mercado de trabalho em Kauai na ausência do furacão, uma perda permanente de 3.400 postos de trabalho no setor privado entre 1992 e 2009 (aproximadamente 15% dos empregos da ilha), além de uma perda de US\$ 650 milhões de dólares em renda pessoal por ano.

Este relatório, conforme mencionado previamente, avalia os impactos do rompimento da Barragem de Fundão exclusivamente sobre o mercado de trabalho formal dos municípios atingidos. Na avaliação, considera-se um horizonte temporal pós-rompimento de apenas três períodos – 31 de dezembro de 2015, 2016 e 2017 –, tratando-se de uma estimação de efeitos de curto prazo, portanto.

2 FONTES DE DADOS UTILIZADAS

A avaliação de impactos sobre empregos e salários no mercado formal de trabalho aqui executada se dá com base nos microdados identificados da Relação Anual de Informações Sociais (RAIS – Ministério da Economia – ME). Esta seção apresenta uma breve contextualização dessa fonte de dados e suas principais características, assim como de outras bases também usadas para construir as análises desenvolvidas neste relatório.

2.1 Relação Anual de Informações Sociais (RAIS – Ministério da Economia – ME)

A Relação Anual de Informações Sociais (RAIS) é um registro administrativo criado pelo Decreto nº 76.900/75 de 23 de dezembro de 1975 para fins como o controle de registros do FGTS e de arrecadação e concessão de benefícios da Previdência. No entanto, é também uma importante ferramenta estatística governamental para “acompanhamento e caracterização do mercado de trabalho formal” (BRASIL, 2020b). Todo estabelecimento deve declarar anualmente a RAIS para o Ministério da Economia, obrigação que se aplica mesmo se em determinado ano ele não teve nenhum trabalhador vinculado ou sequer realizou quaisquer operações. As informações prestadas abrangem o próprio estabelecimento e os vínculos empregatícios a ele ligados, incluindo características pessoais dos empregados.¹³

A divulgação pública dos microdados da RAIS é feita de modo a não permitir identificar os estabelecimentos nem as pessoas vinculadas a eles (“dados não identificados”) (BRASIL, 2020c). Para este relatório, contudo, foram utilizados dados em que constam tais informações (“dados identificados”), obtidos mediante acordo de cooperação técnica com o Ministério da Economia.¹⁴ Neles, constam todos os estabelecimentos e vínculos empregatícios declarados entre os anos de 2010 a 2017, sobre os quais as seguintes informações foram usadas para desenvolver a avaliação de impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal:

¹³ A relação completa de estabelecimentos que devem declarar a RAIS (que inclui empresas privadas e órgãos públicos) e de vínculos empregatícios que devem ser declarados (que inclui celetistas e estatutários) é apresentada em BRASIL (2020a).

¹⁴ Os dados identificados da RAIS foram concedidos pelo então Ministério do Trabalho e Emprego à Escola de Economia de São Paulo da Fundação Getúlio Vargas (FGV EESP) para fins de pesquisa.

- Estabelecimentos: número do CNPJ, número do CEI vinculado, município de localização e endereço completo;
- Vínculos de trabalho: nome da pessoa, número do CPF, número do PIS, números do CNPJ e CEI do estabelecimento a que estava vinculada, data de nascimento, sexo, raça/cor, nível de escolaridade, setor de atividade (conforme CNAE 2.0 – IBGE, 2007) do estabelecimento a que estava vinculada, status de vigência do vínculo em 31 de dezembro de cada ano e remuneração em dezembro de cada ano.¹⁵

O papel desempenhado por cada informação e os procedimentos a elas aplicados são detalhados ao longo do relatório.

Os microdados identificados da RAIS são a principal fonte de informações da análise porque retratam a parcela formal do mercado de trabalho em recortes regionais e temporais que são propícios a uma avaliação de impacto econométrica rigorosa:

- Recorte regional: cada vínculo de trabalho está associado a um estabelecimento cujo município de localização está registrado na RAIS. É possível ainda verificar onde os vínculos se localizam dentro dos municípios a partir da informação de endereço do estabelecimento, o que efetivamente é feito em análises de heterogeneidade no Capítulo 4;
- Recorte temporal: há informações referentes tanto a um período anterior ao rompimento da Barragem de Fundão (2010 a 2014) quanto a um posterior (2015 a 2017).¹⁶

Ademais, a existência de informações particulares a cada trabalhador, como o nome da pessoa e seus números do PIS e CPF, entre outras destacadas anteriormente, faz com que seja possível ainda acompanhar a trajetória de cada indivíduo no mercado de trabalho formal ao longo do tempo.

Devido a essas três características, a RAIS viabiliza uma investigação mais focada e detalhada da existência de impactos do rompimento sobre o mercado de trabalho formal. Ela permite restringir com precisão o conjunto de trabalhadores de interesse ao daqueles nos 45 municípios atingidos e observar o comportamento de seus indicadores por vários períodos antes e após o rompimento, tornando mais rigorosa a diferenciação

¹⁵ A remuneração em 31 de dezembro não inclui 13º salário (BRASIL, 2020a).

¹⁶ O ano de 2015 é considerado parte do período posterior ao rompimento porque as observações referentes a ele usadas na análise são tomadas em 31 de dezembro de 2015. Esse aspecto metodológico é discutido nos Capítulos 3 e 4.

entre impactos do rompimento e oscilações de indicadores do mercado de trabalho formal não relacionadas ao desastre.

Não há nenhuma outra base de dados (pública ou de acesso restrito) que retrate o mercado de trabalho brasileiro com os mesmos recortes regionais e temporais que os da RAIS, apesar de esta ser restrita ao emprego formal. Uma possível alternativa seria a PNAD Contínua do IBGE, cuja divulgação trimestral fornece um panorama do mercado de trabalho formal e informal, conforme discutido na seção 1.1 do produto. Entretanto, a PNAD Contínua é mais limitada em termos de recorte regional, uma vez que sua amostragem não é representativa para municípios brasileiros, apenas Unidades Federativas e regiões metropolitanas. Optar pela PNAD Contínua em detrimento da RAIS prejudicaria a focalização da análise desenvolvida, já que não seria possível restringi-la aos 45 municípios atingidos.

Apesar de permitir observar vínculos formais de trabalho no nível de municípios e em qualquer momento dos anos considerados na análise, a RAIS apresenta algumas limitações. Além da restrição dos dados ao mercado formal, a limitação mais importante é o fato de que a informação de remuneração é referente a dezembro de cada ano, implicando que pessoas cujo vínculo de trabalho é encerrado antes de dezembro não têm informação de salário. Além disso, há casos de vínculos que existiam em dezembro cuja remuneração constava como 0. As decisões metodológicas tomadas para lidar com essas limitações são detalhadas nos Capítulos 3 e 4.

2.2 Censo Demográfico de 2010 (IBGE)

O Censo Demográfico, realizado pelo IBGE, é a principal fonte de referência sobre dados de condições de vida da população brasileira. Ele consiste em um questionário básico, aplicado a todos os domicílios do Brasil com o intuito de coletar informações sobre características do domicílio e dos moradores, e um questionário amostral mais extenso e detalhado que inclui itens sobre educação, migração, trabalho e rendimento, trabalho infantil, entre outros. A pesquisa tem abrangência geográfica nacional com representatividade de diversos recortes regionais, dentre eles o de municípios. Na edição de 2010, a data de referência adotada foi 31 de julho daquele ano (IBGE, 2011).

Este estudo faz uso de três informações obtidas a partir dos dados do Censo Demográfico de 2010:

- Percentual de pessoas de cada município analisado cujos domicílios se localizavam em zonas rurais;

- Percentual de pessoas de cada município analisado cuja raça/cor declarada foi registrada como “branca”;
- Renda domiciliar per capita de todos os indivíduos amostrados nos municípios analisados.

Essas informações têm papel importante na construção de um grupo de municípios não atingidos pelo rompimento da Barragem de Fundão que sirva de comparação com os atingidos no que diz respeito ao mercado formal de trabalho. É sobre comparações feitas entre os dois grupos de municípios (atingidos e não atingidos) que a avaliação de impactos desenvolvida aqui depende crucialmente, conforme exposto ao longo da seção 3.1.

2.3 Estimativa da população residente nos municípios brasileiros (IBGE)

Em todos os anos que não realiza uma contagem da população brasileira ou um Censo Demográfico, o IBGE divulga estimativas do tamanho populacional de todos os municípios do Brasil. A avaliação de impactos do rompimento da Barragem de Fundão sobre o mercado formal de trabalho faz uso especificamente da estimativa feita para 2014, cuja data de referência é 1º de julho daquele ano (IBGE, 2014b). Assim como as informações do Censo Demográfico de 2010, a estimativa de tamanho da população municipal em 2014 é utilizada para a seleção de municípios não atingidos pelo rompimento que servem de comparação com os atingidos.

2.4 Cadastro Nacional de Endereços para Fins Estatísticos – CNEFE (IBGE)

O Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE), em particular a versão utilizada neste relatório, é um registro com os endereços de todos os domicílios e unidades não residenciais visitados pelos pesquisadores do IBGE quando da elaboração do Censo Demográfico de 2010 (IBGE, 2011). Para cada unidade recenseada, o CNEFE apresenta uma caracterização detalhada do endereço, contendo o nome e número do logradouro, complementos, o distrito e o município em que se localizava, entre outros. Uma informação importante fornecida pelo CNEFE é o setor censitário da unidade recenseada em 2010, que se trata do menor recorte territorial investigado pelo Censo daquele ano. Isso viabiliza a construção de uma correspondência entre endereços brasileiros de 2010 e setores censitários, a qual tem

papel central no procedimento de atribuição de estabelecimentos listados na RAIS a esses setores, descrito no APÊNDICE B — Atribuição de estabelecimentos a setores censitários. A informação de setor censitário é utilizada em determinadas análises de efeitos heterogêneos do rompimento da Barragem de Fundão sobre o mercado formal de trabalho, conforme detalhado na seção 4.2.

2.5 Bases cartográficas

Informações geoespaciais de diferentes fontes são usadas neste relatório. Uma delas é a “Base Cartográfica Contínua do Brasil, ao Milionésimo” (BCIM), conjunto de dados geoespaciais produzido pelo IBGE na escala de 1:1.000.000 (IBGE, 2014a). Da versão lançada originalmente em 2014 são retirados os *shapefiles* de limites políticos de municípios e estados usados para produzir os mapas apresentados ao longo deste relatório. A BCIM permite ainda a construção de uma matriz de vizinhança geográfica entre municípios, a qual, juntamente com informações georreferenciadas acerca de aglomerados rurais, territórios indígenas, rodovias e unidades de conservação ambiental presentes na BCIM, é utilizada no procedimento de identificação de setor censitário dos estabelecimentos listados na RAIS.¹⁷

A segunda fonte de dados geoespaciais é a malha territorial do Censo Demográfico de 2010 no nível de setores censitários (IBGE, 2019b). Além de utilizada na produção de mapas, essa informação é crucial para classificar os setores dos municípios atingidos conforme a distância em relação ao Rio do Carmo, ao Rio Gualaxo do Norte e ao Rio Doce. Tal classificação é foco de uma das análises de efeitos heterogêneos do rompimento da Barragem de Fundão sobre aspectos do mercado formal de trabalho, desenvolvida na seção 4.2. Os dados geoespaciais dos três rios, por sua vez, são obtidos do *shapefile* “Massas d’Água” disponibilizado pela Agência Nacional de Águas (ANA, 2016).

¹⁷ Vide APÊNDICE B — Atribuição de estabelecimentos a setores censitários por meio de cruzamento de endereços.

3 AVALIAÇÃO DE IMPACTOS SOBRE O MERCADO DE TRABALHO FORMAL DOS MUNICÍPIOS ATINGIDOS

Como discutido na seção 1.2, há embasamento para se investigar a existência de impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal dos municípios atingidos. Este capítulo é destinado a avaliar se o rompimento alterou o tamanho desses mercados, o que seria percebido através de mudança no número de postos formais de trabalho ocupados e na soma das remunerações pagas aos trabalhadores.

Para produzir as análises apresentadas a seguir, os microdados da RAIS foram agregados de forma a construir um banco de dados no nível de municípios organizado na estrutura de “painel”, na qual os dados referentes a cada município são observados com periodicidade regular ao longo do tempo. Em particular, as informações dos municípios contidas no painel dizem respeito a 31 de dezembro de cada ano entre 2010 e 2017, período a que está voltada a análise desenvolvida.

Este capítulo se inicia com a descrição da metodologia adotada e dos objetos específicos de análise, avançando em seguida para a apresentação dos resultados obtidos em conjunto com estatísticas descritivas calculadas a partir do painel de municípios.

3.1 Metodologia

Este relatório adota métodos econométricos para estudar os impactos do rompimento da Barragem de Fundão sobre o mercado formal de trabalho, tendo como objetivo específico investigar relações de causa e efeito entre o desastre e o comportamento demonstrado por indicadores quantitativos de interesse em momentos posteriores ao do rompimento. Tais indicadores, referidos como “indicadores de impacto” ao longo do texto, consistem em representações de aspectos específicos de um objeto ou fenômeno na forma de medidas quantificáveis, as quais são geralmente obtidas através de processos estatísticos e de contagem. Os indicadores de impacto provêm, assim, uma métrica objetiva passível de ser explorada por análises quantitativas, como as produzidas por métodos econométricos.

Idealmente, a identificação de relações de causa e efeito entre o rompimento e alterações observadas no mercado formal de trabalho seria feita pela comparação do indicador de impacto dos municípios atingidos em dois cenários: o real e o contrafactual. No contexto deste relatório, o cenário real corresponde à situação em que os atingidos

efetivamente se encontram e que está registrada em dados, enquanto o cenário contrafactual seria a situação em que os atingidos estariam se o rompimento da Barragem de Fundão não tivesse acontecido. Se fosse possível observar essas duas situações, então tudo em que o cenário real diferisse do contrafactual poderia ser considerado consequência do rompimento. No entanto, o cenário contrafactual é apenas uma abstração – é impossível saber como a trajetória dos atingidos teria evoluído desde novembro de 2015 se o rompimento não tivesse ocorrido. Não obstante, essa construção teórica fornece as bases para a estratégia empírica que é adotada neste produto.¹⁸

Os efeitos do rompimento sobre indicadores relacionados a empregos e salários no mercado formal de trabalho são estimados a partir da comparação de um grupo de municípios atingidos pelo desastre com um de municípios não atingidos. Este segundo, denominado “grupo de comparação”, é construído sob a premissa de prover uma representação aproximada do que teria acontecido com o grupo atingido caso o rompimento não tivesse ocorrido – isto é, o grupo de comparação deve fornecer uma aproximação razoável de como os indicadores de impacto do grupo atingido teriam se comportado no cenário contrafactual.

Para que um grupo de comparação possa ser considerado uma representação satisfatória da situação contrafactual dos atingidos, há algumas condições que precisam ser atendidas. Quando isso ocorre, as diferenças entre o indicador de impacto do grupo atingido e o grupo de comparação podem ser interpretadas exclusivamente como consequências do rompimento e não de outros eventuais fenômenos concomitantes. Por este motivo, tais condições são frequentemente denominadas “hipóteses de identificação” – elas tornam possível discernir (“identificar”) o impacto do rompimento em meio aos diversos outros fatores captados nos dados.

A literatura econométrica de avaliação de impactos oferece métodos que exploram os dados de diferentes maneiras com o objetivo de estabelecer uma relação causal entre um evento e a trajetória daqueles por ele afetados. Cada método possui um conjunto específico de hipóteses de identificação, e decidir qual a melhor abordagem para cada avaliação é atribuição do pesquisador, que deve ponderar as características do problema em questão em conjunto com a estrutura e qualidade dos dados à disposição. Apesar disso, esses diferentes métodos têm em comum o fato de se basear na

¹⁸ É pertinente destacar que, embora a metodologia guarde semelhanças com aquela já implementada nos relatórios de avaliação de impactos nas áreas de Assistência Social, Educação e Segurança Pública elaborados pela FGV em 2019 (FGV, 2019a; FGV, 2019b; FGV, 2019c), há diferenças importantes causadas pelas características dos dados utilizados. Tais particularidades são apresentadas nesta seção e na seção 4.1.

comparação de indicadores de impacto entre os dois grupos já mencionados, os quais serão detalhados nas seções 3.1.3 e 3.1.4:

- **Grupo atingido**¹⁹: corresponde, no contexto deste relatório, ao grupo de municípios que foram atingidos pelo rompimento da Barragem de Fundão; e
- **Grupo de comparação**²⁰: corresponde, no contexto deste relatório, a um grupo de municípios que não foram atingidos pelo rompimento da Barragem de Fundão, mas que se assemelham àqueles que foram atingidos e, supõe-se, provêm de uma representação aproximada da situação contrafactual destes, isto é, do que teria acontecido com os atingidos caso não tivesse ocorrido o rompimento.

Quando satisfeitas as hipóteses de identificação do método econométrico escolhido, a comparação entre o grupo atingido e de comparação possui interpretação causal com respeito à relação estimada entre o rompimento e a trajetória do indicador de impacto analisado. Torna-se possível verificar se efetivamente existe impacto e, em caso positivo, estimar sua magnitude.

Neste relatório, o método econométrico adotado é o do modelo de diferença-em-diferenças. As seções seguintes apresentam os fundamentos e hipóteses de identificação do método, assim como a maneira como é aplicado à avaliação de impactos sobre o mercado formal de trabalho dos municípios atingidos pelo rompimento da Barragem de Fundão.

3.1.1 Método de diferença-em-diferenças

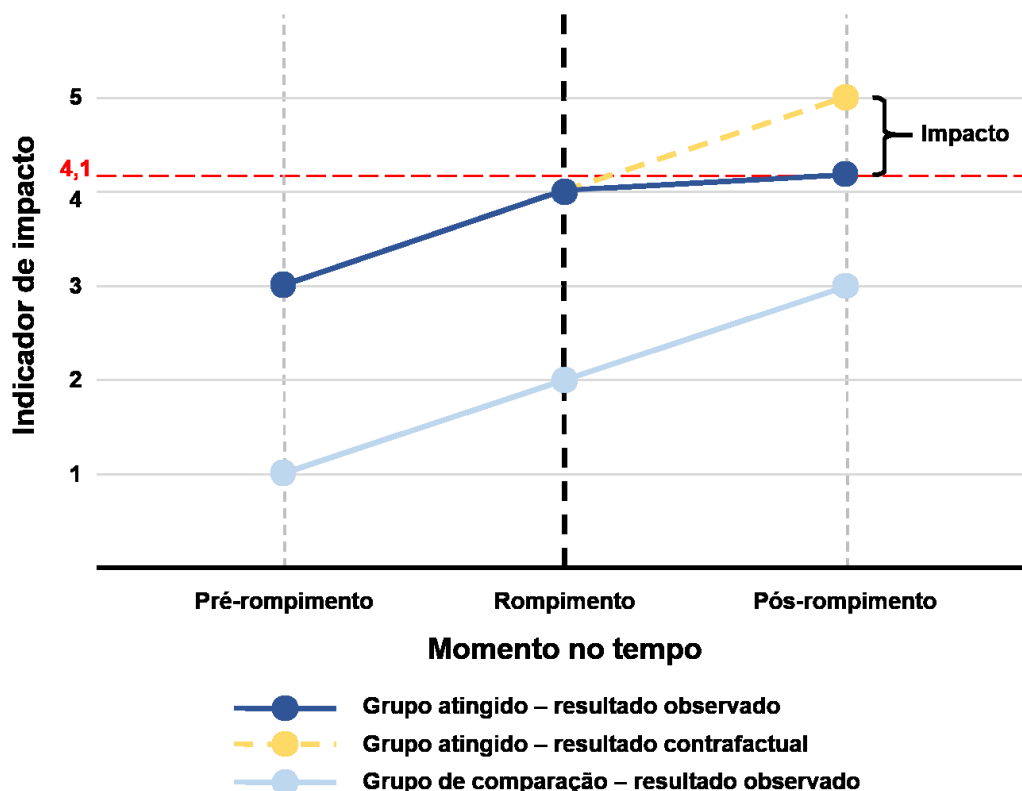
A estimativa de impactos através do método de diferença-em-diferenças é feita a partir da comparação dos indicadores de impacto entre os grupos atingido e de comparação em dois momentos do tempo: antes e depois do rompimento, conforme ilustrado na Figura 2.²¹ Nela, estão representadas a evolução do indicador de impacto dos dois grupos (pontos conectados por linhas em tons de azul) assim como a evolução que o grupo atingido demonstraria no cenário hipotético em que o rompimento não ocorreu (linha tracejada e ponto em amarelo), o qual corresponde à sua situação contrafactual.

¹⁹ O grupo que sofre a intervenção em análise é tipicamente chamado de “grupo de tratamento” na literatura de avaliação econométrica de impactos. No presente estudo, como já mencionado, este grupo é identificado por “grupo atingido”.

²⁰ Na literatura de avaliação econométrica de impactos, este grupo frequentemente é denominado “grupo de controle”.

²¹ Os valores e trajetórias de evolução representados na Figura 2 são genéricos, apresentados para fins de ilustração no método apenas.

Figura 2 — Ilustração do método de diferença-em-diferenças



Se, ao longo do tempo, não acontecesse nada que tivesse influência sobre o indicador de impacto apenas de municípios atingidos, seria de se esperar que a diferença entre o indicador dos dois grupos no momento pós-rompimento (ponto amarelo subtraído do azul-claro: $5 - 3 = 2$) fosse igual à observada antes do rompimento (ponto azul-escuro subtraído do azul-claro: $3 - 1 = 2$). Entretanto, a figura mostra que o indicador do grupo atingido no pós-rompimento é, na realidade, menor do que seria em seu cenário contrafactual (4,1 ao invés de 5), de maneira que a diferença entre o grupo de comparação e atingido é menor depois do que antes do rompimento. Posto de outra forma, caso se subtraia a diferença de indicador observada entre os grupos antes do rompimento da diferença observada entre eles depois (isto é, calcular a “diferença entre as diferenças”: $(4,1 - 3) - (3 - 1)$), obtém-se um valor negativo ($1,1 - 2 = -0,9$), indicando que a discrepância entre os dois grupos diminuiu após o rompimento. A magnitude dessa redução corresponde, na figura, ao ponto azul-escuro subtraído do amarelo ($4,1 - 5 = -0,9$), que é justamente o quão menor o indicador de impacto dos atingidos é em comparação ao que seria se o desastre não tivesse ocorrido – ou seja, trata-se do impacto do rompimento.

A “diferença entre as diferenças” poderá ser considerada o impacto do rompimento se as hipóteses de identificação do modelo de diferença-em-diferenças forem satisfeitas. São elas:

Na ausência do rompimento, os indicadores de impacto dos grupos atingido e de comparação teriam que apresentar evolução semelhante ao longo do tempo – isto é, apresentar tendências “paralelas”. É importante enfatizar que isso não significa que o indicador de impacto dos dois grupos precisaria ter magnitude similar em todos os momentos de tempo na ausência do rompimento, mas apenas que as trajetórias dos dois teriam que ser semelhantes, conforme ilustrado na Figura 2.

- Os grupos atingido e de comparação não podem ter sido afetados de maneira particular e distinta um do outro por um fator externo concomitante ao rompimento e variável no tempo;
- A composição dos grupos atingido e de comparação não pode se alterar de forma significativa ao longo do tempo.

De modo geral, quando alguma das hipóteses parece ser inválida, a causa da “diferença entre as diferenças” não pode seguramente ser atribuída ao rompimento da Barragem de Fundão. Como consequência, torna-se necessário escolher outro grupo de comparação que se acredite ser uma boa representação do contrafactual dos atingidos ou, ainda, repensar a aplicação do método de diferença-em-diferenças ao estudo dos impactos do rompimento.

É importante notar que é difícil, quando não impossível, verificar se essas hipóteses são válidas na prática. Por exemplo, não há como saber se a trajetória contrafactual do grupo atingido seria de fato semelhante à do grupo de comparação escolhido porque, por definição, o cenário contrafactual não é o que se observa na realidade. Não obstante, é possível realizar análises específicas que servem como indicativos de credibilidade de alguma das hipóteses no contexto estudado, as quais são apresentadas conjuntamente com os resultados na seção 3.2.

Uma das propriedades mais atraentes do método de diferença-em-diferenças é o fato de que ele consegue evitar que a estimativa de impacto seja distorcida, ou “enviesada”, por fatores externos invariantes no tempo, independentemente de estes serem ou não observados em dados. Em outras palavras, se as diferenças entre o grupo atingido e o grupo de comparação forem apenas características que, após o rompimento, permanecem iguais ao que eram antes, então ainda assim a estimativa produzida pelo modelo de diferença-em-diferenças consegue identificar o impacto do rompimento. Dessa forma, é desnecessária a preocupação com, por exemplo, as maneiras como

diferenças em leis municipais podem tornar a evolução dos indicadores do grupo de comparação uma representação ruim do contrafactual da dos atingidos, desde que essas leis existam antes do rompimento e não tenham sido alteradas após ele ter ocorrido.²²

A estimativa de impacto obtida através do método de diferença-em-diferenças corresponde ao impacto médio sobre o grupo atingido.²³ Em outras palavras, supondo satisfeitas as hipóteses de identificação, o parâmetro que o modelo estima a partir dos dados é a variação média do indicador de impacto dos municípios atingidos causada pelo desastre. Ele não informa, portanto, o impacto sobre um município em particular, mas sim uma estimativa agregada referente ao conjunto dos atingidos.²⁴

A estimação do impacto médio sobre o grupo atingido pode ser representada em notação algébrica da seguinte forma, onde se traduz a mecânica ilustrada na Figura 2:

- Para os grupos atingido e de comparação, calcula-se separadamente a variação da média do indicador de impacto ao longo do tempo (operação por vezes denominada “primeiras diferenças”):

$$\Delta Y^{\text{atingido}} = E[Y_m | d_m = 1, t = 1] - E[Y_m | d_m = 1, t = 0]$$

$$\Delta Y^{\text{comparação}} = E[Y_m | d_m = 0, t = 1] - E[Y_m | d_m = 0, t = 0]$$

em que

- E indica cálculo do valor esperado (isto é, a média);
- m indica a unidade de observação (município);
- Y_m é o valor do indicador de impacto do município m ;
- $d_m = 1$ indica que o município m é atingido pelo rompimento e $d_m = 0$, que não é atingido;
- $t = 1$ indica o período posterior ao rompimento e $t = 0$, o anterior;
- Subtrai-se a variação da média do grupo de comparação da variação da média do grupo atingido (a “diferença entre as diferenças”):

$$\text{impacto médio sobre atingidos} = \Delta Y^{\text{atingido}} - \Delta Y^{\text{comparação}}$$

²² O método de diferença-em-diferenças não é capaz de identificar o efeito do rompimento quando o indicador de impacto dos municípios é influenciado por características não observadas que variam no tempo e estas diferem entre atingidos e não atingidos. É exatamente sobre isso a que se refere a segunda hipótese de identificação apresentada.

²³ Na terminologia da literatura de avaliação econométrica de impactos, o método de diferença-em-diferenças estima o chamado “efeito médio do tratamento sobre os tratados” (em inglês, *average treatment effect on the treated* – ATT).

²⁴ No Capítulo 4, o método de diferença-em-diferenças é utilizado para estimar efeitos específicos sobre subconjuntos de indivíduos atingidos. Isso é detalhado no referido capítulo.

Embora seja possível operacionalizar o método de diferença-em-diferenças com as operações aritméticas apresentadas acima, as avaliações de impacto produzidas para este relatório utilizam o modelo de regressão linear para fazê-lo. Os detalhes da implementação e as vantagens de estimar o diferença-em-diferenças por meio de uma regressão são discutidos a seguir.

3.1.2 Estimação do diferença-em-diferenças com regressão linear

O modelo de regressão linear é uma ferramenta estatística que, em linhas gerais, estabelece uma relação quantitativa entre um conjunto de atributos de determinado objeto (denominados “variáveis independentes”) e uma característica sua de particular interesse (“variável dependente”). Cada variável independente é associada a um parâmetro, os quais são estimados a partir dos dados ao se solucionar o modelo. Tais parâmetros podem ser informalmente vistos como pesos que refletem a “importância” de cada atributo para se determinar o valor da variável dependente. Ademais, a regressão linear, como indica o nome, presume que a variável dependente é resultado da soma da influência de cada variável independente – ou seja, ela pressupõe que a relação existente entre as variáveis tem a forma de uma equação linear.

Para os propósitos da análise desenvolvida aqui, o modelo de regressão consiste em uma maneira conveniente de se operacionalizar o método de diferença-em-diferenças para estimar os impactos do rompimento da Barragem de Fundão. Dois são os motivos para tal:

- O modelo de regressão acomoda facilmente versões mais flexíveis do procedimento descrito na seção 3.1.1, como se verá a seguir;
- A solução do modelo produz automaticamente as informações necessárias para se testar estatisticamente a “relevância” das variáveis independentes na determinação da dependente.

De fato, o modelo de regressão adotado implementa uma versão modificada do diferença-em-diferenças, na qual não se trabalha com a comparação entre apenas um período posterior e um anterior ao rompimento, como ilustrado na Figura 2.

O modelo, de maneira distinta, toma um ano como referência e então calcula as diferenças dos demais anos com relação a ele. Apesar dessa mudança, a mecânica do método é essencialmente a mesma da descrita na seção 3.1.1. Em notação algébrica, o modelo de regressão estimado é:

Equação 1 – Modelo de regressão para o método de diferença-em-diferenças na análise de municípios

$$Y_{m,t} = \alpha + \sum_{\substack{j=2010, \\ j \neq 2014}}^{2017} \beta_j \cdot \mathbb{I}_{\{m \text{ é atingido}\}} \cdot \mathbb{I}_{\{j=t\}} + \gamma_m + \lambda_t + \varepsilon_{m,t}$$

em que:

- m é o indexador de municípios;
- t é o indexador de tempo, referente a 31 de dezembro de cada ano analisado;
- $Y_{m,t}$ corresponde ao valor do indicador de impacto do município m observado no ano t ;
- Σ é o somatório com indexador j tomado sobre os valores 2010 a 2017 exceto 2014;
- $\mathbb{I}_{\{m \text{ é atingido}\}}$ é uma variável indicadora (*dummy*) que assume valor 1 se o município m fizer parte do grupo atingido pelo desastre e 0 se integrar o grupo de comparação;
- $\mathbb{I}_{\{j=t\}}$ é uma *dummy* que assume valor 1 se o indexador j do somatório for igual ao ano t da observação e 0 em caso contrário;
- α é o parâmetro correspondente ao intercepto do modelo, que capta fatores relevantes para determinar Y , cuja influência é comum a todas as observações – ou seja, comum a todos os municípios em todos os anos considerados;
- β_j , com j indo de 2010 a 2017 exceto 2014, são os parâmetros que representam os estimadores de diferença-em-diferenças para cada ano. O ano de 2014 é tomado como referência, de modo que os β_j expressam o quanto a diferença entre atingidos e não atingidos se modificou entre o ano j e 2014;²⁵
- γ_m é o parâmetro associado ao “efeito fixo” do município m , que capta o efeito de todas as características fixas no tempo de m que são relevantes para determinar Y ;

²⁵ Por este motivo, o modelo não contém um parâmetro β_{2014} , já que a diferença de 2014 com relação a si próprio é nula por construção.

- λ_t é o parâmetro associado ao “efeito fixo” do ano t , que capta o efeito de todos os fatores relevantes para determinar Y que são comuns a todos os municípios e aconteceram especificamente naquele ano;
- $\varepsilon_{m,t}$ é o termo de erro do modelo, que contém a parte de $Y_{m,t}$ não explicada pelos demais elementos. Consideradas as demais variáveis independentes, supõe-se que este termo é distribuído de forma aleatória sobre as observações incluídas no modelo.

O modelo contém então sete parâmetros que correspondem a estimadores de diferença-em-diferenças: β_{2010} , β_{2011} , β_{2012} , β_{2013} , β_{2015} , β_{2016} e β_{2017} . De principal interesse para a avaliação de impactos, no entanto, são apenas os que se referem a momentos posteriores ao rompimento da Barragem de Fundão: 2015, 2016 e 2017.²⁶ São esses parâmetros que, satisfeitas as hipóteses de identificação do método, informam o impacto médio do rompimento sobre os municípios atingidos frente aos do grupo de comparação em relação à situação de 2014. Com três períodos posteriores ao rompimento ao invés de somente um, como na exposição realizada na seção 3.1.1, torna-se possível avaliar, por exemplo, se há efeitos imediatos que se dissipam ao longo do tempo ou se estes são persistentes, ou ainda se a resposta do mercado de trabalho é lenta e começa a se manifestar apenas nos períodos finais da análise. Em outras palavras, eles podem ser utilizados para analisar se há heterogeneidade nos impactos entre os diferentes anos pós-rompimento.

Os parâmetros de diferença-em-diferenças referentes a anos anteriores ao desastre são também importantes, pois servem como indicativos de credibilidade das estimativas obtidas para os anos pós-rompimento. Isto porque, como mencionado anteriormente, o método de diferença-em-diferenças exige que a trajetória do indicador de impacto do grupo de comparação seja uma representação aproximada da trajetória contrafactual do grupo atingido. Assim, se sequer um dos parâmetros pré-desastre do modelo captar que a diferença entre os indicadores dos grupos é distinta da que se observa em 2014, torna-se muito menos convincente o argumento de que os grupos atingido e de comparação evoluiriam de maneira semelhante se o rompimento não tivesse ocorrido. Como ilustração desse ponto, suponha que o modelo estime que a relação entre os grupos era diferente em 2012 quando comparada à de 2014 – dois momentos em que o desastre não existia. Se fosse esse o caso, o que garantiria que uma eventual diferença estimada em 2016 com respeito a 2014 se deveria ao rompimento da

²⁶ Note que as observações de municípios a partir de dados da RAIS são tomadas em 31 de dezembro de cada ano, o que justifica considerar as de 2015 como referentes a um momento pós-rompimento.

Barragem de Fundão e não a outro motivo, como o que teria influenciado os indicadores de impacto entre 2012 e 2014?

É importante manter em mente que, mesmo se nenhum dos parâmetros pré-rompimento do modelo captar diferenças com relação a 2014, não é possível garantir que a hipótese de tendências paralelas seja válida no período pós-rompimento. Trata-se apenas de um indício favorável a ela, que concede maior credibilidade à interpretação de relação causal das estimativas de impacto. Ele não impede que um fenômeno concomitante sem relação com o rompimento afete apenas um dos grupos de municípios após 2015, o que tornaria o grupo de comparação uma representação ruim da situação contrafactual do atingidos. Portanto, permanece a ressalva de que eventuais mudanças observadas nos indicadores de impacto a partir do método de diferença-em-diferenças admitem interpretação de impactos causados pelo rompimento da Barragem de Fundão somente se as hipóteses de identificação do método forem satisfeitas, conforme exposto anteriormente.

Um aspecto importante do modelo de regressão ao qual se fez alusão anteriormente é o fato de as estimativas dos parâmetros serem acompanhadas de intervalos de confiança. Estes intervalos descrevem um conjunto de valores ao redor da estimativa que busca levar em conta as incertezas no processo de estimação oriundas da variabilidade de Y entre os municípios ou mesmo de problemas no registro dos dados que alimentam a análise (“erros de medida”).

Através dos intervalos de confiança, é possível avaliar quão precisas são as estimativas obtidas: um intervalo pequeno significa que, se a magnitude verdadeira do parâmetro estiver dentro dele, então ela estará relativamente próxima ao valor da estimativa. Em contraposição, um intervalo demasiadamente amplo permite que, mesmo se estiver no intervalo, o parâmetro verdadeiro possa ser bastante diferente do valor encontrado para sua estimativa, tornando esta última pouco informativa. Ainda, quando o intervalo de confiança para a estimativa de impacto abrange o valor 0, não se pode afirmar que o rompimento teve, em média, qualquer impacto sobre os municípios atingidos. Em termos técnicos, diz-se que o valor estimado não é considerado estatisticamente significativo dado o nível de significância adotado para a construção do intervalo, que usualmente é de 1%, 5% ou 10%.²⁷

Os intervalos de confiança associados a cada estimativa são calculados de forma a acomodar a autocorrelação municipal, o que significa que a capacidade de avaliar a

²⁷ Mais detalhes sobre a interpretação dos intervalos de confiança podem ser encontrados em FGV (2019c, p. 41).

precisão e significância estatística dos parâmetros estimados é preservada mesmo que o indicador de impacto de cada município seja parcialmente determinado pelo seu valor nos anos anteriores (isto é, correlacionado ao longo do tempo).²⁸

Por fim, considerando se tratar de uma análise com os dados agregados a nível de município, o modelo de regressão é estimado ponderando as observações de cada município pelo número de vínculos de trabalho com remuneração não nula que nele vigiam em 31 de dezembro de 2014. Isso é feito para que as estimativas dos parâmetros de interesse captem o efeito médio do rompimento sobre toda a região atingida, uma vez que, na ausência de ponderação, elas seriam demasiadamente influenciadas pelo comportamento dos indicadores de impacto em municípios relativamente pequenos.²⁹

3.1.3 Grupo atingido

A avaliação de impactos do rompimento da Barragem de Fundão sobre o mercado formal de trabalho considera como grupo atingido o conjunto de 45 municípios citados no TTAC (BRASIL, 2016) e na Deliberação nº 58 do CIF (BRASIL, 2017) além de Ponte Nova, devido à comunidade de Chopotó nele localizada ter sido atingida.³⁰ São eles:

- Em Minas Gerais (MG): Aimorés, Alpercata, Barra Longa, Belo Oriente, Bom Jesus do Galho, Bugre, Caratinga, Conselheiro Pena, Córrego Novo, Dionísio, Fernandes Tourinho, Galileia, Governador Valadares, Iapu, Ipaba, Ipatinga, Itueta, Mariana, Marliéria, Naque, Periquito, Pingo d'Água, Ponte Nova, Raul Soares, Resplendor, Rio Casca, Rio Doce, Santa Cruz do Escalvado, Santana do Paraíso, São Domingos do Prata, São José do Goiabal, São Pedro dos Ferros, Sem-Peixe, Sobrália, Timóteo e Tumiritinga;
- Em Espírito Santo (ES): Aracruz, Baixo Guandu, Colatina, Conceição da Barra, Fundão, Linhares, Marilândia, Serra e São Mateus.³¹

²⁸ Em outras palavras, os erros-padrão calculados, a partir dos quais se constrói os intervalos de confiança, são robustos à heteroscedasticidade *intracluster*, sendo municípios os *clusters* da análise (cujo nível de observação é município-ano).

²⁹ Uma discussão mais extensa sobre o uso de pesos em modelos de regressão com dados em nível de municípios é apresentada em FGV (2019c, p. 45).

³⁰ Ressalta-se que essa lista de municípios atingidos foi utilizada para a produção do presente documento, não configurando uma lista final de municípios atingidos que serão utilizados nos demais estudos produzidos pela Fundação Getúlio Vargas (FGV). Ressalta-se ainda que essa lista de municípios atingidos não configura um posicionamento da FGV quanto à delimitação de áreas atingidas para qualquer finalidade.

³¹ O município de Anchieta, localizado no Espírito Santo, abriga uma usina de pelotização de minério de ferro pertencente à Samarco Mineração S.A. Tal usina tinha participação relevante no PIB municipal de Anchieta, de forma que é possível levantar argumentos favoráveis à inclusão desse município no grupo atingido. Todavia, Anchieta não é reconhecido como

3.1.4 Grupos de comparação

As seções anteriores enfatizam a importância que a escolha de um grupo de comparação tem para determinar a credibilidade de uma avaliação econométrica de impactos. A construção dos grupos de comparação usados nas análises apresentadas neste relatório leva em conta informações socioeconômicas gerais e também específicas ao mercado de trabalho formal dos municípios brasileiros para se tentar determinar quais dos não atingidos pelo rompimento da Barragem de Fundão eram, antes do desastre, similares aos municípios que efetivamente o foram. A estratégia adotada consiste na elaboração de uma métrica para o grau de similaridade entre os municípios, de modo que os municípios não atingidos que, segundo tal métrica, apresentarem as menores diferenças em relação aos atingidos são os que se decide incluir no grupo de comparação.

A motivação por trás dessa estratégia é a hipótese de que seriam boas aproximações à situação contrafactual dos municípios atingidos aqueles não atingidos cujas características observadas antes do rompimento eram próximas das dos atingidos. Sendo eles parecidos antes do desastre, espera-se que tivessem continuado assim na ausência do rompimento, podendo então a evolução dos indicadores de impacto do grupo de comparação ser considerada uma representação aproximada daquela que os atingidos teriam demonstrado em seu cenário contrafactual.

A “métrica de similaridade” mencionada trata-se da distância de Mahalanobis, dada pela fórmula:

$$D_{m,n} = [(X_m - X_n)' \Sigma^{-1} (X_m - X_n)]^{\frac{1}{2}}$$

onde m e n são indexadores de municípios, D é o valor da distância de Mahalanobis, X é o vetor contendo os valores das características municipais que se considera importante serem similares entre atingidos e não atingidos, e Σ é a matriz de covariância das características consideradas.³² A distância de Mahalanobis sintetiza em um único

município atingido nos documentos oficiais relacionados ao Caso Samarco, de modo que passar a fazê-lo aqui criaria inconsistência não somente com as demais análises já produzidas pela FGV, mas também com o processo como um todo. A decisão metodológica adotada foi então não permitir que o município de Anchieta integrasse o grupo atingido nem nenhum grupo de comparação, uma vez que a premissa por trás desse segundo é justamente a inexistência de impacto do rompimento da Barragem de Fundão. Não obstante, para fins de análise de sensibilidade, todos os resultados apresentados neste relatório foram reproduzidos considerando Anchieta um município atingido, fazendo com que esse grupo totalizasse 46 municípios. Nenhuma alteração relevante foi constatada.

³² Seguindo Imbens e Rubin (2015, p. 342), Σ é calculada como a média ponderada entre a matriz de covariância sobre os municípios atingidos e a sobre os não atingidos, onde os pesos correspondem ao número de municípios existente em cada um dos dois grupos.

número o quão diferentes são dois municípios m e n em termos das características consideradas em X , levando ainda em consideração como essas características variam entre todos os municípios atingidos e não atingidos. Se dois municípios comparados forem semelhantes, a distância de Mahalanobis terá um valor baixo; quanto mais diferentes forem, maior seu valor será.

A distância de Mahalanobis é central para o procedimento de *pareamento*, através do qual efetivamente se constrói os grupos de comparação. Para determinado município atingido, calcula-se a distância de Mahalanobis com relação a cada um dos não atingidos. Os três municípios com as menores distâncias com relação ao atingido em questão são incluídos no grupo de comparação, e passa-se então ao próximo município atingido, com o qual se calcula as distâncias com relação aos não atingidos restantes (isto é, excluídos os três que já foram escolhidos para o grupo de comparação). O processo se repete até que todos os 45 municípios atingidos tenham feito três pares com não atingidos, resultando em um grupo de comparação de 135 municípios. Esse procedimento pode ser descrito como um de pareamento de municípios atingidos aos três “vizinhos mais próximos” sem reposição, onde “vizinhança” é dada em termos da distância de Mahalanobis e “sem reposição” reflete o fato de que um município atingido não pode ser pareado a um não atingido que já tenha sido escolhido para outro atingido em etapa anterior. Com relação a esse aspecto, o ordenamento de municípios atingidos no procedimento de pareamento é definido de maneira aleatória – isto é, realiza-se um sorteio para definir qual município atingido será o primeiro a formar pares com três não atingidos, qual será o segundo e assim por diante até que todos os municípios atingidos tenham sido pareados. Procura-se evitar dessa maneira que os resultados das estimações sejam dependentes de um arranjo particular de municípios quando da execução do pareamento.

As variáveis consideradas para cálculo da distância de Mahalanobis, incluídas no vetor X da fórmula apresentada acima, são:

- Provenientes de informações do Censo Demográfico de 2010 (IBGE):
 - Proporção da população do município cujo domicílio se localizava em zonas rurais;
 - Proporção de população do município cuja raça/cor declarada foi registrada como “branca”;
 - Índice de Gini do município, calculado a partir da renda domiciliar per capita de todos os indivíduos lá amostrados.

- Provenientes de informações da RAIS 2014 (ME), considerando apenas vínculos de trabalho com remuneração em dezembro não nula e vigentes em 31 de dezembro daquele ano:
 - Número de vínculos no município como proporção do tamanho da população municipal estimado para 2014;
 - Número de vínculos de mulheres como proporção do número total de vínculos;
 - Número de vínculos de pessoas com idade entre 20 e 39 anos como proporção do número total de vínculos;
 - Número de vínculos de pessoas cujo setor de atividade se enquadrava em agropecuária e pesca como proporção do número total de vínculos;
 - Número de vínculos de pessoas cujo setor de atividade se enquadrava em mineração (exclusive de carvão) como proporção do número total de vínculos;
 - Número de vínculos de pessoas cujo setor de atividade se enquadrava no de indústrias em geral como proporção do número total de vínculos;
 - Número de vínculos de pessoas com ao menos ensino médio completo como proporção do número total de vínculos;
 - Mediana da distribuição do logaritmo natural do salário;
 - Diferença interquartil da distribuição do logaritmo natural do salário.³³

Além das características listadas, o tamanho da população do município estimado para 2014 pelo IBGE também foi usado na execução do pareamento, porém não através de sua inclusão no vetor de variáveis sobre o qual a distância de Mahalanobis é calculada. O procedimento consiste em definir cinco faixas de tamanho populacional distintas, a saber: (i) até 5 mil habitantes; (ii) de 5.001 a 10 mil hab.; (iii) de 10.001 a 20 mil hab.; (iv) de 20.001 a 50 mil hab.; e (v) acima de 50 mil hab. Os municípios, tanto atingidos quanto não atingidos candidatos a pares, são agrupados nessas cinco faixas conforme sua estimativa de tamanho populacional para 2014. O pareamento com distância de Mahalanobis é então aplicado “dentro” de cada um dos grupos, de maneira que um

³³ As variáveis de mediana e diferença interquartil do logaritmo natural dos salários exigem que se aplique a transformação de logaritmo natural à informação de salários em dezembro associada aos vínculos formais do município que vigiam em 31 de dezembro de 2014 e não tinham remuneração nula. Dessa distribuição de salários em logaritmo natural extrai-se o segundo quartil, que corresponde à variável de mediana mencionada, e se subtrai o primeiro do terceiro quartil, obtendo assim a variável de diferença interquartil.

município atingido só possa ser pareado a não atingidos que estejam na mesma faixa de tamanho populacional que ele.³⁴

A avaliação de impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal faz uso de dois grupos de comparação distintos. Ambos têm em comum o fato de serem construídos através do procedimento de pareamento com distância de Mahalanobis sem reposição aos três vizinhos mais próximos. Fundamentalmente, eles diferem apenas no conjunto de municípios não atingidos disponíveis para pareamento com os atingidos. A descrição e motivação por trás de cada um desses conjuntos de “candidatos” a pares são apresentadas a seguir.

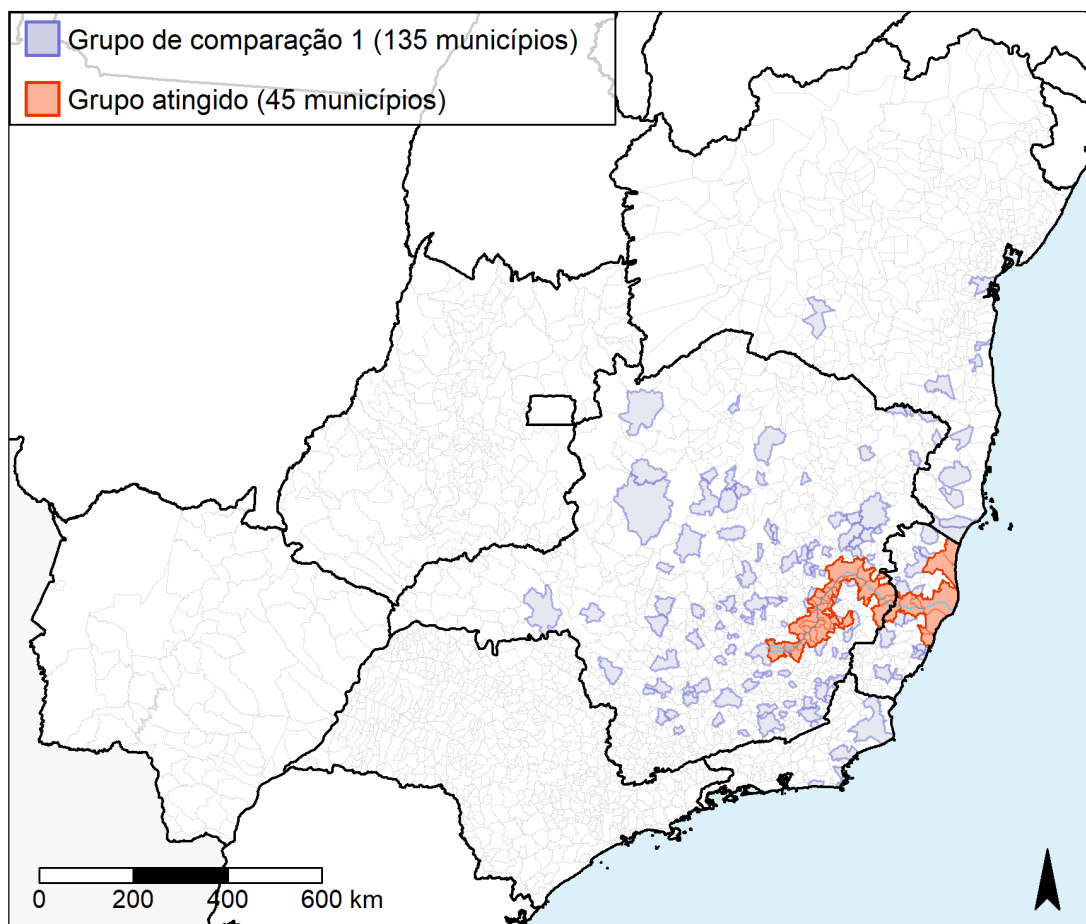
No primeiro deles, denominado “grupo de comparação 1” (GC 1), a localização geográfica de cada município atingido também influencia a execução do pareamento sem figurar entre as variáveis sobre as quais se calcula a distância de Mahalanobis, assim como ocorre com o tamanho populacional. Em particular, determina-se que municípios atingidos de Minas Gerais (MG) só possam ser pareados a não atingidos localizados também em MG, enquanto os atingidos do Espírito Santo (ES) devem ser pareados a não atingidos também do próprio ES ou com municípios do sul da Bahia (BA) ou do norte do Rio de Janeiro (RJ).³⁵ Sendo assim, o uso tanto do tamanho populacional quanto da localização do município faz com que o pareamento sem reposição a 3 vizinhos mais próximos com distância de Mahalanobis seja executado em até dez subgrupos de municípios.³⁶ Os municípios não atingidos selecionados para compor o GC 1 são ilustrados na Figura 3.

³⁴ Quando ocorre de determinada faixa de tamanho populacional não conter municípios não atingidos suficientes para pareamento com todos os atingidos nela existentes, aglutina-se a faixa com a de tamanho imediatamente inferior. Quando este problema ocorre na primeira faixa de tamanho populacional, aglutina-se com a de tamanho imediatamente superior.

³⁵ Os municípios da Bahia considerados candidatos a pares de atingidos do Espírito Santo são os pertencentes às mesorregiões denominadas “Centro Sul Baiano” e “Sul Baiano”. Os do Rio de Janeiro, por sua vez, são aqueles localizados nas mesorregiões “Norte Fluminense” e “Baixadas”. Mais detalhes sobre o conceito de mesorregião são apresentados em IBGE (1990).

³⁶ Da nota de rodapé 34 se extrai o motivo de não ser necessariamente dez grupos.

Figura 3 — Mapa ilustrando o grupo atingido e o grupo de comparação 1 (GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME), do Censo Demográfico (IBGE), de estimativa de tamanho populacional dos municípios (IBGE) e de bases cartográficas (ANA e IBGE).

Observação: O grupo de comparação 1 é composto por municípios não atingidos de Minas Gerais, Espírito Santo, sul da Bahia e norte do Rio de Janeiro pareados aos atingidos.

O uso tanto de MG e ES quanto sul da BA e norte do RJ como área na qual podem ser selecionados municípios não atingidos para integrar o GC 1 é motivado pela relevância que características não observadas em dados (como aspectos culturais, institucionais e governamentais) têm para a determinação do perfil do município tanto em termos socioeconômicos gerais como específicos ao mercado formal de trabalho. Espera-se que as características de municípios não atingidos em MG e ES, assim como o contexto no qual estão inseridos, sejam mais similares aos dos atingidos do que os de municípios localizados em outros estados, em particular os mais distantes. Por sua vez, a inclusão de municípios não atingidos no sul da BA e no norte do RJ como candidatos a pares de atingidos no ES se deve ao fato de que, por serem geograficamente próximos ao ES, potencialmente guardam semelhanças importantes com os municípios desse estado. Além disso, no sul da BA e norte do RJ há rios de portes semelhantes ao do Rio Doce,

cujas fozes estão também localizadas neles³⁷, o que torna ainda mais razoável supor que a dinâmica socioeconômica dessas regiões se assemelha às dos municípios atingidos na região da foz do Rio Doce. Os municípios do sul da BA e norte do RJ podem, portanto, ser também uma alternativa para a construção do GC 1 com respeito aos atingidos do ES. Isso é particularmente relevante porque a quantidade de municípios existentes no Espírito Santo é relativamente pequena (78 no total), além do fato de que os municípios atingidos nele localizados cobrem grande parte de sua faixa litorânea. Restam assim poucos municípios não atingidos no próprio ES que estejam também no litoral, o que poderia levar à inclusão no GC 1 de pares consideravelmente diferentes dos atingidos por falta de alternativas mais adequadas. Para evitar comprometer o potencial do GC 1 como aproximação à situação contrafactual dos atingidos, permite-se então que municípios no sul da BA e norte do RJ estendam o número de possibilidades para o pareamento de municípios atingidos do ES.

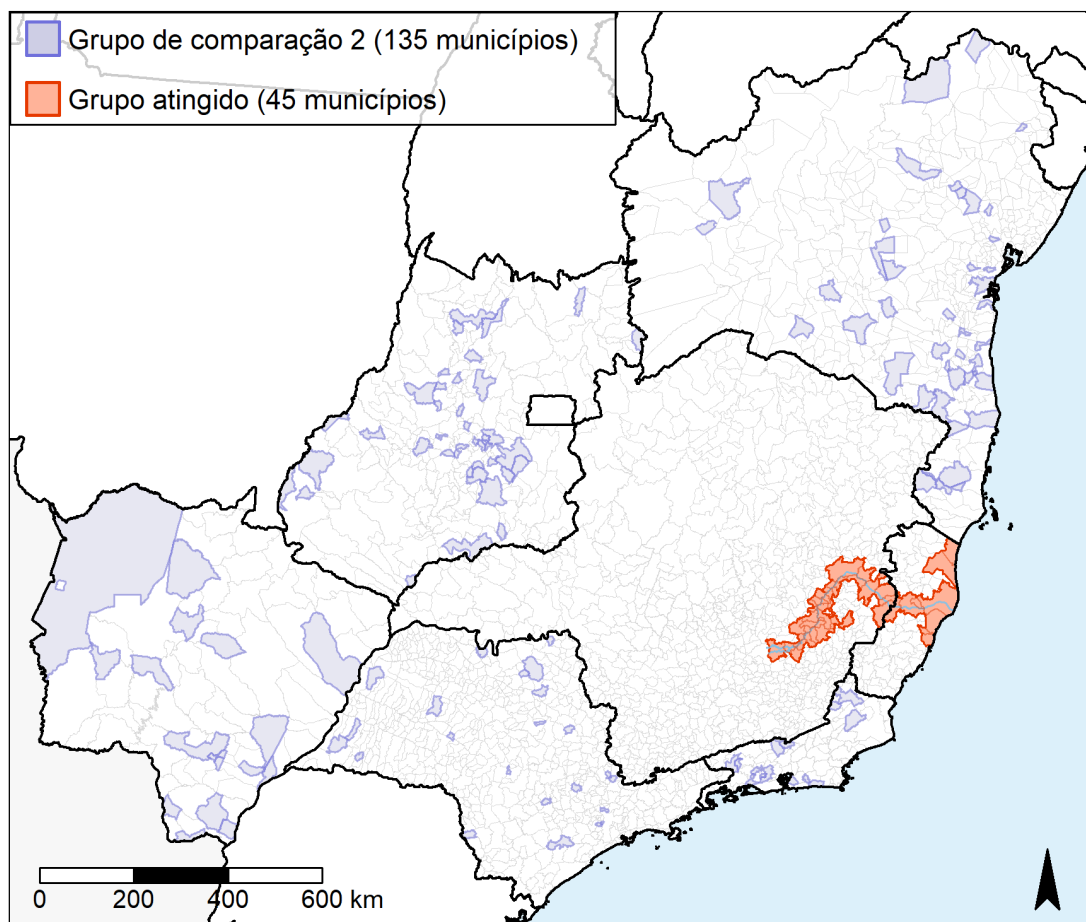
Entretanto, é possível que o impacto do rompimento da Barragem de Fundão sobre o mercado de trabalho formal dos municípios atingidos tenha repercussões para outros municípios, em particular aqueles cujas economias têm maior interdependência com as dos atingidos.³⁸ Considerando então a possibilidade de que o efeito do rompimento possa ter se estendido para todos os municípios de MG e ES, motiva-se a construção do “grupo de comparação 2” (GC 2), para o qual a seleção de municípios não atingidos é restrita àqueles localizados nos estados vizinhos a MG e ES – especificamente, Bahia, Goiás, Mato Grosso do Sul, Rio de Janeiro e São Paulo. Em contraste com o GC 1, um município atingido, independentemente de pertencer a MG ou a ES, pode ser pareado a qualquer não atingido desses cinco estados. Assim, o pareamento com distância de Mahalanobis é executado em até cinco subgrupos de municípios³⁹, definidos apenas pelas faixas de tamanho populacional. A Figura 4 ilustra a composição obtida para o GC 2.

³⁷ São eles o Rio Jequitinhonha e Rio Pardo, ambos na Bahia, e Rio Paraíba do Sul, no Rio de Janeiro.

³⁸ Em outras palavras, pode ser que haja efeitos de “transbordamento” dos impactos do rompimento (em inglês, *spillover effects*).

³⁹ Análogo à nota de rodapé 36.

Figura 4 — Mapa ilustrando o grupo atingido e o grupo de comparação 2 (GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME), do Censo Demográfico (IBGE), de estimativa de tamanho populacional dos municípios (IBGE) e de bases cartográficas (ANA e IBGE).

Observação: O grupo de comparação 2 é composto por municípios não atingidos da Bahia, Goiás, Mato Grosso do Sul, Rio de Janeiro e São Paulo pareados aos atingidos.

Em princípio, não é possível afirmar que um grupo de comparação é seguramente melhor que o outro no sentido de aproximar a situação contrafactual dos atingidos, uma vez que ambos possuem vantagens e desvantagens em potencial relacionadas a vies nas estimações do modelo de diferença-e-diferenças. Como discutido, o GC 1 tende a conter municípios mais parecidos com os atingidos por compartilhar características regionais e institucionais mais próximas, mas pode causar vies de subestimação de eventuais impactos do rompimento sobre os atingidos caso ocorram efeitos de transbordamento. O GC 2, por sua vez, tem propensão menor a ser afetado por tais efeitos de transbordamento, mas pode ser menos comparável com o grupo atingido por ser exclusivamente composto de municípios de fora de MG e ES, o que pode provocar vies de variável omitida sobre as estimativas. Realizar a avaliação de impacto com os dois grupos de comparação é então justificável porque permite utilizar o resultado obtido

com um como base de comparação com o outro, de onde se torna mais fácil interpretar as estimativas e analisar sua sensibilidade à escolha de grupo de comparação.

Ainda no tópico de sensibilidade das estimativas ao grupo de comparação adotado, duas variações do pareamento descrito acima são executadas para fins de análise de robustez dos resultados das estimações de impacto: (i) pareamento a três vizinhos mais próximos com reposição; e (ii) pareamento a cinco vizinhos mais próximos sem reposição. A primeira variação altera o procedimento ao permitir que um município atingido seja pareado a um não atingido que já foi incluído no grupo de comparação. O efeito prático dessa mudança é que, por exemplo, se um dos três pares escolhidos de qualquer município atingido já tiver sido incluído no grupo de comparação, esse município agregará ao grupo de comparação apenas dois novos municípios, e não três como no pareamento sem reposição. Ao fazer isso, espera-se reduzir a possibilidade de introdução de viés nas estimativas, já que se evita a inclusão de municípios menos parecidos com os atingidos. Por outro lado, pode provocar também menor precisão na estimação, uma vez que a variabilidade diminui ao considerar menos municípios diferentes. A segunda variação, por sua vez, não permite reposição, mas aumenta o número de vizinhos de um município atingido que deve integrar o grupo de comparação. Dessa maneira, abre-se margem para maior viés nas estimativas, já que municípios menos parecidos com os atingidos passam a integrar o grupo de comparação, embora tenda a levar o aumento na precisão, decorrente de maior variabilidade advinda do uso de mais municípios diferentes na estimação. Em suma, a motivação por trás desses pareamentos adicionais é verificar se os resultados das estimações na seção 3.2 não são demasiadamente sensíveis à composição dos grupos de comparação. Os resultados das estimações para fins de análise de robustez são apresentados no APÊNDICE C — Tabelas adicionais do capítulo 3.

Por fim, o resultado do pareamento no que se refere à semelhança entre os municípios atingidos e os dos grupos de comparação pode ser analisado na Tabela 5. Nela, as linhas contêm média e desvio-padrão de todas as variáveis consideradas no cálculo da distância de Mahalanobis, e as três colunas à direita permitem comparar esses valores entre os grupos atingido e de comparação 1 e 2 (GC 1 e GC 2, respectivamente). Além dessas estatísticas, há ainda uma terceira linha por bloco de variável, “valor-p”, que sintetiza o resultado de um teste estatístico de comparação de médias entre os grupos: o valor-p na coluna do GC 1 indica a probabilidade de que a diferença entre a média do grupo atingido e a do grupo de comparação 1 seja equivocadamente considerada diferente de 0. Analogamente, o valor-p de GC 2 mostra essa probabilidade para a comparação entre a média do grupo atingido e a do grupo de comparação 2. Quando o

valor-p é inferior a 0,10, por exemplo, diz-se que as médias dos grupos são estatisticamente diferentes ao nível de significância de 10% – isto é, a probabilidade de se estar cometendo um erro ao tratá-las como valores distintos quando na verdade a diferença é irrelevante é menor que 10%.⁴⁰

⁴⁰ É usual tratar como “efetivamente” diferentes valores de média com diferença estatisticamente significativa a um nível igual ou inferior a 10%. Outros valores convencionais para o nível de significância são 5% e 1%.

Tabela 5 — Testes de diferença de médias entre grupos atingido e de comparação

Variáveis	Estatística	Atingidos	GC1	GC2
Proporção da população do município cujo domicílio se localizava em zonas rurais em 2010	Média	0,199	0,205	0,189
	Desvio-padrão	0,149	0,173	0,169
	Valor-p		0,829	0,727
Proporção de população do município cuja raça/cor declarada foi registrada como "branca" em 2010	Média	0,340	0,404	0,432
	Desvio-padrão	0,078	0,140	0,146
	Valor-p		0,003	0,000
Índice de Gini do município em 2010	Média	0,492	0,500	0,485
	Desvio-padrão	0,044	0,048	0,047
	Valor-p		0,345	0,374
Número total de vínculos no município como proporção do tamanho da população municipal estimado para 2014	Média	0,175	0,178	0,170
	Desvio-padrão	0,073	0,104	0,073
	Valor-p		0,861	0,665
Proporção de vínculos de mulheres	Média	0,426	0,434	0,425
	Desvio-padrão	0,068	0,059	0,060
	Valor-p		0,408	0,941
Proporção de vínculos de pessoas com idade entre 20 e 39 anos	Média	0,574	0,574	0,565
	Desvio-padrão	0,050	0,049	0,056
	Valor-p		0,985	0,358
Proporção de vínculos cujo setor de atividade se enquadrava em agropecuária e pesca	Média	0,097	0,090	0,102
	Desvio-padrão	0,088	0,082	0,103
	Valor-p		0,619	0,763
Proporção de vínculos cujo setor de atividade se enquadrava em mineração (exclusive de carvão)	Média	0,012	0,009	0,007
	Desvio-padrão	0,026	0,025	0,017
	Valor-p		0,540	0,122
Proporção de vínculos cujo setor de atividade se enquadrava no de indústrias em geral	Média	0,137	0,141	0,172
	Desvio-padrão	0,101	0,111	0,130
	Valor-p		0,828	0,099
Proporção de vínculos de pessoas com ao menos ensino médio completo	Média	0,591	0,601	0,592
	Desvio-padrão	0,082	0,092	0,090
	Valor-p		0,508	0,922
Mediana da distribuição do logaritmo natural do salário	Média	7,206	7,205	7,268
	Desvio-padrão	0,159	0,224	0,193
	Valor-p		0,978	0,051
Diferença interquartil da distribuição do logaritmo natural do salário	Média	0,661	0,669	0,638
	Desvio-padrão	0,180	0,207	0,130
	Valor-p		0,820	0,360
Teste conjunto de significância estatística ⁴¹	Estatística F		1,133	3,172
	Valor-p		0,336	0,000

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME), do Censo Demográfico de 2010 (IBGE) e da estimativa de tamanho populacional dos municípios (IBGE).

Observação: Os vínculos de trabalho considerados para a construção das variáveis são apenas os de pessoas que constam na RAIS como empregadas em 31 de dezembro de 2014 com remuneração não nula. Todas as variáveis de proporção de vínculos são calculadas com relação ao número total de vínculos no município, com exceção da primeira: "Número total de vínculos no município como proporção do tamanho da população municipal estimado para 2014." Observações de municípios foram ponderadas pelo número de vínculos vigentes em 31 de dezembro de 2014 com remuneração não nula.

⁴¹ O teste conjunto de significância estatística é realizado através da estimação de um modelo de regressão no qual a variável dependente é uma variável indicadora que assume valor 1 se o município for do grupo atingido e 0 caso pertença ao grupo de comparação identificado no cabeçalho da Tabela 5. As variáveis independentes do modelo são as listadas na tabela, que correspondem às que são consideradas no cálculo da distância de Mahalanobis. A estatística de teste, denotada por "estatística F", é comparada aos valores da distribuição probabilística Fisher-Snedecor com 12 e 167 graus de liberdade para avaliar a hipótese nula de que as variáveis são conjuntamente irrelevantes para explicar que o município faz parte do grupo atingido. A rejeição desta hipótese indica que ao menos uma das variáveis independentes é relevante para explicar se o município faz parte do grupo atingido ou do de comparação em questão.

Em geral, as médias das variáveis do grupo de comparação 1 resultaram mais similares às dos atingidos, o que não surpreende quando se considera que municípios geograficamente mais próximos tendem a compartilhar maiores semelhanças em vários aspectos. Pode se considerar até mesmo que isso se aplique a características não captadas por dados (características “não observadas”), uma vez que um conjunto delas pode, em algum nível, estar correlacionado às variáveis incluídas no cálculo da distância de Mahalanobis. A única diferença com o GC 1 que se mostra estatisticamente significativa a 10% é a de proporção da população declarada como branca no Censo Demográfico de 2010 (IBGE), sendo que os municípios do grupo atingido tinham, em média, uma parcela menor de seus habitantes que se reconhecia como branca.

A diferença de médias entre grupos na proporção de brancos na população de 2010 existe também quando se avalia o GC 2, com o qual os atingidos têm ainda diferenças estatisticamente significantes ao nível de 10% na proporção de vínculos de trabalho em atividades industriais e na mediana da distribuição do logaritmo natural dos salários. Além disso, note que o grupo atingido e o GC 2 mostram-se estatisticamente diferentes quando todas as características são consideradas conjuntamente, conforme mostra o valor-p do teste de significância conjunta nas últimas duas linhas da Tabela 5, o que não é o caso para o GC 1.⁴²

Apesar das discrepâncias observadas, é importante lembrar que o método do diferença-em-diferenças não exige que os grupos atingido e de comparação sejam necessariamente iguais em todas as características relevantes para determinação de indicadores de impacto relacionados ao mercado formal de trabalho. O importante para que a principal hipótese de identificação do método seja satisfeita é que a trajetória dos indicadores do grupo de comparação seja igual ao longo do tempo à trajetória contrafactual (não observada) dos atingidos. Cabe destacar que a especificação do modelo de regressão na Equação 1 prevê a inclusão de variáveis que levam em conta “efeitos fixos” dos municípios, os quais neutralizam diferenças entre eles em características invariantes no tempo quando da estimação do modelo, independentemente de estas serem ou não observáveis.

⁴² Ressalte-se que os resultados do pareamento na Tabela 5 foram analisados tendo como referência o nível de significância de 10%, o que consiste em uma escolha arbitrária. Caso se opte por adotar o nível de 5%, um valor também usual para se avaliar significância estatística, a única diferença significativa entre o grupo atingido e os de comparação seria a de proporção da população municipal declarada branca em 2010. Não obstante, a conclusão obtida dos testes de significância conjunta permanece inalterada mesmo com a troca de nível de significância – isto é, as características do GC 2 são em média conjuntamente diferentes das do grupo atingido também ao nível de significância de 5%.

3.1.5 Indicadores de impacto analisados

Os indicadores de impacto descritos a seguir foram escolhidos para a avaliação de impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal dos municípios atingidos por expressarem informações importantes a respeito do tamanho desses mercados, tanto em termos de postos de trabalho como de nível de remuneração.

3.1.5.1 Logaritmo natural do número de vínculos ativos

Neste relatório, entende-se por “vínculo ativo” em determinado ano o vínculo de trabalho formal que consta na RAIS como vigente em 31 de dezembro de tal ano e cuja remuneração de dezembro é maior que 0.⁴³ Então, para um município m observado no ano t , este indicador de impacto corresponde ao logaritmo natural do número total de vínculos com remuneração não nula vigentes em 31 de dezembro do ano t de pessoas empregadas em estabelecimentos que à época se localizavam no município m .

3.1.5.2 Logaritmo natural da massa salarial

Para um município m observado no ano t , este indicador de impacto corresponde ao logaritmo natural da soma das remunerações de vínculos ativos em 31 de dezembro do ano t de pessoas empregadas em estabelecimentos que à época se localizavam no município m . Os valores das remunerações são convertidos para R\$ de dezembro de 2017 utilizando o IPCA.

O uso do logaritmo natural na definição dos dois indicadores torna possível interpretar os parâmetros estimados pelo modelo de regressão como uma aproximação à variação percentual dos indicadores de impacto em resposta às variáveis independentes. Além disso, o logaritmo natural tem como consequência a redução da variância dos indicadores de impacto, algo desejável porque, embora a maioria dos municípios considerados seja relativamente pequena no que se refere a mercado de trabalho, há alguns muito populosos cujo número de vínculos e massa salarial são bem maiores que os dos demais.⁴⁴ Esse tipo de discrepância prejudica a capacidade do modelo de

⁴³ Note que a definição de “vínculo ativo” utilizada neste relatório é diferente daquela que aparece no dicionário de microdados da RAIS, em que “vínculo ativo” refere-se a estar empregado em 31 de dezembro do ano de referência, sem que seja feita qualquer menção à remuneração em dezembro do vínculo.

⁴⁴ Em outras palavras, as distribuições do número de vínculos ativos e da massa salarial são assimétricas à direita.

regressão linear de encontrar um bom ajuste aos dados, levando ainda a estimativas imprecisas. O logaritmo natural atua no sentido de atenuar essa influência.

3.2 Resultados

Esta seção apresenta os resultados da avaliação de impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho dos municípios atingidos. Para fins de exposição, as estimativas da Equação 1 para os dois indicadores de impacto definidos na seção anterior são apresentadas separadamente. As análises são baseadas na comparação do grupo atingido com os grupos de comparação 1 e 2 (GC 1 e GC 2, respectivamente) construídos por meio de pareamento sem reposição aos três vizinhos mais próximos. Tabelas com resultados usando os grupos de controle alternativos – versões de pareamento usando três vizinhos com reposição e cinco vizinhos sem reposição – figuram no APÊNDICE C — Tabelas adicionais do capítulo 3, servindo como análises de robustez das estimativas apresentadas a seguir.

3.2.1 Logaritmo natural do número de vínculos ativos

A Tabela 6 apresenta o número médio de vínculos ativos dos grupos atingido e de comparação 1 e 2 em cada ano do período de análise. A estatística de média é acompanhada de seu desvio-padrão, que permite avaliar o quão próximo a ela tende a ser o número de vínculos ativos de cada um dos municípios dentro dos grupos analisados. Note que a tabela não descreve a trajetória ao longo do tempo do indicador de impacto conforme ele foi definido na seção 3.1.5.1: os valores apresentados nela correspondem à variável em nível – por exemplo, o valor médio do grupo atingido em 2014 é lido como 11.263 vínculos ativos, sendo seu desvio-padrão de 24.533 vínculos ativos. As estatísticas referentes à variável em logaritmo natural constam na Tabela 18 no APÊNDICE C — Tabelas adicionais do capítulo 3.

Tabela 6 — Média e desvio-padrão do número de vínculos ativos por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	10.257	9.752	8.884
	Desvio-padrão	21.788	20.968	18.670
2011	Média	10.685	10.330	9.433
	Desvio-padrão	22.838	22.402	19.927
2012	Média	10.750	10.487	9.760
	Desvio-padrão	23.172	23.390	20.506
2013	Média	11.140	10.850	10.103
	Desvio-padrão	23.891	24.008	21.377
2014	Média	11.263	10.958	10.252

Ano	Estatística	Atingido	GC 1	GC 2
	Desvio-padrão	24.533	24.188	21.659
2015	Média	10.725	10.556	10.037
	Desvio-padrão	23.038	23.023	21.134
2016	Média	10.104	9.961	9.579
	Desvio-padrão	21.679	21.737	20.181
2017	Média	10.143	10.025	9.642
	Desvio-padrão	21.412	21.318	20.063

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Os valores apresentados correspondem à variável em nível e não em logaritmo natural.

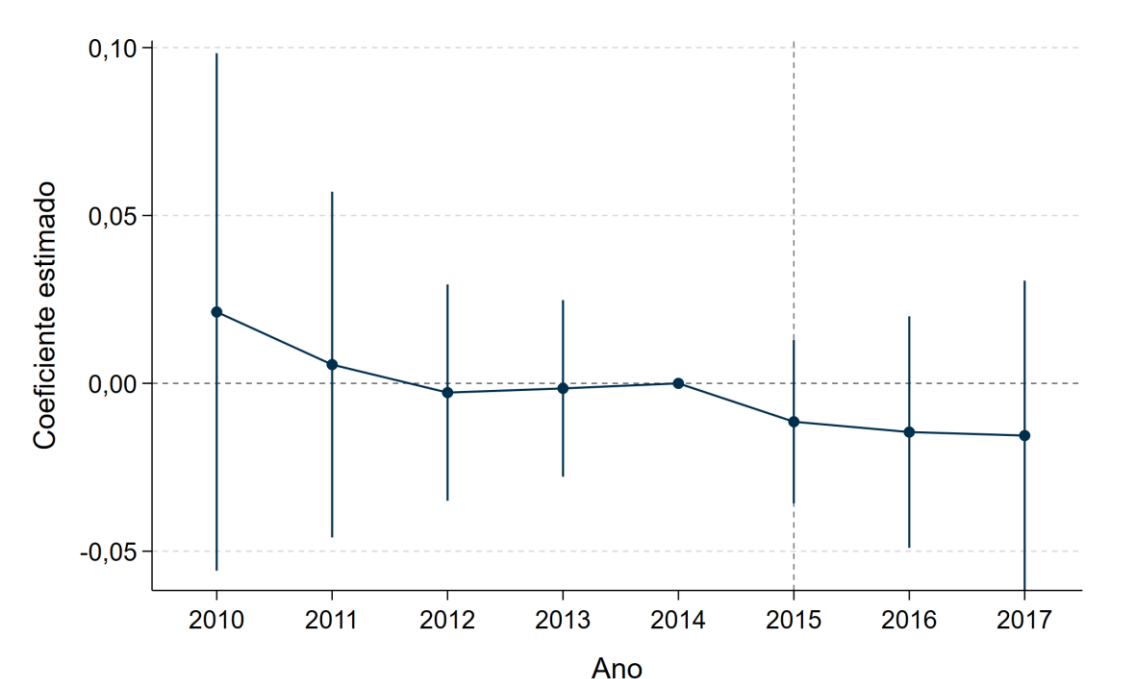
Nota-se de imediato na tabela que o desvio-padrão é sempre maior que o dobro da média, independentemente do grupo ou ano analisado. Isso sugere que a composição de municípios dos dois grupos demonstra grande discrepância no que se refere ao tamanho do mercado formal de trabalho. Em outras palavras, os três grupos contêm municípios relativamente grandes e pequenos, ao invés de vários com um número similar de vínculos ativos. É por causa da tamanha variabilidade que se adota o logaritmo natural do número de vínculos ativos como indicador de impacto, conforme argumentado no final da seção 3.1.5. É motivado por ela também que se estima a Equação 1 com a ponderação de observações, abordagem descrita na seção 3.1.2.

Além dos elevados desvios-padrão das médias, a Tabela 6 mostra que o número médio de vínculos ativos cresceu nos três grupos analisados até 2014, comportamento que se inverte em 2015 e 2016. Ao longo de todo o período, o grupo atingido permaneceu sendo o que, em média, concentrava mais vínculos, seguido pelo GC 1 e pelo GC 2. Apesar disso, a diferença entre eles se reduziu, especialmente depois que se inicia o movimento de queda generalizada no número de vínculos ativos. Em 2014, o grupo atingido tinha em média 11.263 vínculos, 305 a mais que o GC 1 e 1.011 a mais que o GC 2. Já em 2017, a média dos atingidos foi de 10.143, 118 e 501 vínculos a mais que os GC 1 e 2, respectivamente. Ou seja, a diferença com relação aos grupos de comparação se reduziu a menos da metade no período de três anos.

Da Tabela 6 apenas não é possível afirmar se o rompimento da Barragem de Fundão está ou não relacionado à queda relativamente acelerada no número médio de vínculos ativos do grupo atingido frente ao GC 1 e GC 2. Como discutido na seção 3.1, a avaliação econométrica de impactos do rompimento neste relatório é realizada através do modelo de diferença-em-diferenças, cujos resultados são apresentados no Gráfico 1. Nele, o grupo atingido é comparado ao GC 1 para estimar os coeficientes β_{2010} a β_{2017} da Equação 1. Cada ponto no gráfico representa um dos parâmetros, cuja magnitude estimada é indicada no eixo vertical. As linhas verticais que atravessam os

pontos são os intervalos de confiança, construídos sob nível de significância de 10%. Isso significa que quando um intervalo não atravessa a reta horizontal tracejada do valor 0, a variável à qual o intervalo está associado é estatisticamente significativa ao nível de 10%. Posto de outra forma, a probabilidade de se estar cometendo um erro ao tratar tal estimativa como diferente de 0 quando na verdade o impacto médio é nulo é igual ou inferior a 10%.⁴⁵ Os mesmos resultados são apresentados de forma mais completa na Tabela 19 do APÊNDICE C — Tabelas adicionais do capítulo 3.

Gráfico 1 — Estimativas de impacto médio sobre o logaritmo natural do número de vínculos ativos de municípios no grupo atingido (comparando com GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014.

Os coeficientes estimados para os três anos posteriores ao rompimento (2015 a 2017) são negativos com magnitude entre -0,011 e -0,016, tornando-se maiores em valores absolutos conforme mais distante de 2014 é o ano considerado. No ano de 2017, então,

⁴⁵ Expresso ainda de outra maneira, adotando terminologia mais técnica, a probabilidade de se rejeitar a hipótese nula do teste de que o parâmetro é igual a 0 quando ela é verdadeira é igual ou inferior a 10%.

a estimativa de -0,016 significa que o número de vínculos ativos no grupo atingido foi, em média, aproximadamente 1,6% menor do que seria se a trajetória de seu indicador de impacto desde 2014, o ano de referência da análise⁴⁶, fosse semelhante à do GC 1. No entanto, não se pode rejeitar a hipótese de que esse impacto seja nulo, pois o intervalo de confiança dessa estimativa, assim como os de todas as outras, inclui o valor 0. Em outras palavras, nenhum dos parâmetros estimados pode ser considerado estatisticamente diferente de 0 ao nível de significância de 10%.

Os modelos estimados que comparam o grupo atingido ao GC 1 em suas versões de pareamento a três vizinhos mais próximos com reposição e a cinco vizinhos mais próximos sem reposição têm também as variáveis que medem os impactos do rompimento em geral estatisticamente iguais a 0 ao nível de 10% de significância, como mostra a Tabela 19 no APÊNDICE C — Tabelas adicionais do capítulo 3. Todavia, percebe-se que a magnitude dos coeficientes estimados, sempre negativos, apresenta alguma oscilação entre os modelos, um indicativo de que o resultado pode ser sensível à estratégia de construção do grupo de comparação.

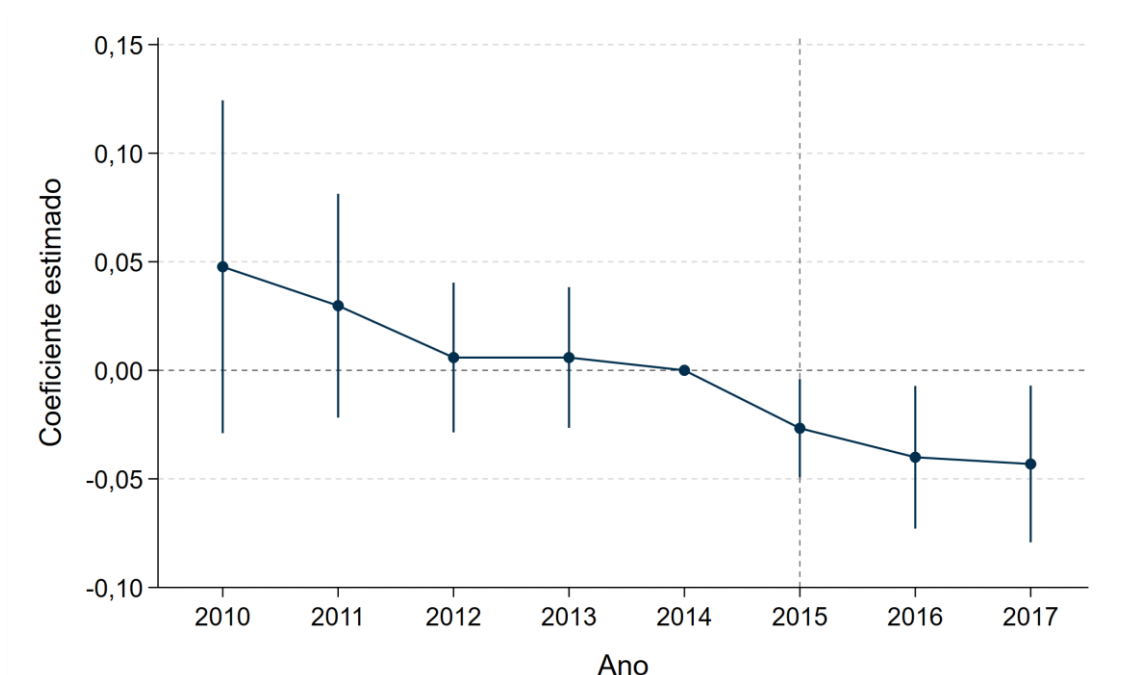
A comparação do grupo atingido com o GC 2 produz resultados diferentes, como mostra o Gráfico 2. Como no Gráfico 1, os três parâmetros referentes ao período pós-rompimento apresentam estimativas cuja magnitude aumenta em valor absoluto à medida que o ano avaliado se distancia de 2014, o ano de referência. Diferentemente do resultado da comparação dos atingidos com o GC 1, no entanto, é o fato de que essas três variáveis são estatisticamente significantes ao nível de 10%⁴⁷, uma vez que seus respectivos intervalos de confiança estimados não contêm o valor 0. Em 2015 a estimativa do diferença-em-diferenças foi de -0,027, indicando que o número médio de vínculos ativos no grupo atingido foi aproximadamente 2,7% menor do que seria se a evolução de seu indicador de impacto tivesse se dado às mesmas taxas de crescimento que a do GC 2, que representa sua situação contrafactual. Levando em consideração que, segundo a Tabela 6, o número médio de vínculos ativos nos municípios atingidos foi de 10.725 em 2015, a redução estimada implica que o grupo atingido perdeu em média cerca de 300 vínculos a mais por município desde 2014 do que perderia se tivesse seguido trajetória similar à dos municípios do GC 2.⁴⁸

⁴⁶ Como notado na seção 3.1.2, a omissão do parâmetro β_{2014} do modelo de regressão reflete a escolha de 2014 como período de referência da análise. Os demais coeficientes devem ser interpretados como diferenças em relação ao que se observava em 2014.

⁴⁷ As variáveis referentes aos impactos do rompimento em 2016 e em 2017 (cujos coeficientes associados na Equação 1 são β_{2016} e β_{2017} , respectivamente) também são estatisticamente significantes ao nível mais rigoroso de 5%.

⁴⁸ Os 300 vínculos mencionados resultam do cálculo $10.725 \cdot \left[1 - \frac{1}{1-0,027}\right] = -297,6 \approx -300$.

Gráfico 2 – Estimativas de impacto médio sobre o logaritmo natural do número de vínculos ativos de municípios no grupo atingido (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014.

Em 2016, o coeficiente estimado é de -0,04, indicando que o número médio de vínculos ativos do grupo atingido diminuiu cerca de 4% a mais do que teria sob a trajetória do GC 2 (perda de aproximadamente 420 vínculos por município⁴⁹). Em 2017, apesar de a Tabela 6 mostrar que o número de vínculos ativos dos atingidos voltou a crescer frente a 2016, a perda acumulada desde 2014 frente ao GC 2 ainda assim aumenta, próxima a 4,3% (cerca de 455 vínculos a menos por município). Isso significa que, mesmo após reverter a tendência de queda no número médio de vínculos ativos observada em 2015 e 2016 na Tabela 6, o crescimento demonstrado pelo grupo atingido foi aquém do que seria sob a trajetória do GC 2, o qual se supõe representar sua situação contrafactual. Considerando a totalidade dos 45 municípios atingidos, estima-se que a perda de postos de trabalho devido ao rompimento da Barragem de Fundão foi de aproximadamente

⁴⁹ O cálculo é análogo ao descrito na nota de rodapé 48, porém usando como referência o número médio de vínculos ativos do grupo atingido em 2016 informado pela Tabela 6.

13.500 em 2015, 18.900 em 2016 e 20.475 em 2017, sempre em comparação ao que se observava em 2014.⁵⁰

É importante também notar no Gráfico 2 que nenhuma das variáveis do diferença-em-diferenças que se referem a momentos anteriores ao rompimento da Barragem de Fundão são estatisticamente diferentes de 0 ao nível de significância de 10%. Sendo assim, o modelo não apresenta evidência de que as trajetórias do indicador de impacto do grupo atingido e do GC 2 seriam divergentes antes do desastre, fato que, por sua vez, torna razoável a hipótese de que elas permaneceriam evoluindo de maneira similar após 2014 caso o rompimento não tivesse ocorrido.⁵¹ A redução no número médio de vínculos ativos do grupo atingido a partir de 2015 destacada anteriormente pode, portanto, ser de fato consequência do rompimento.

Indo ao encontro dessa conclusão, as análises de robustez do resultado apresentado no Gráfico 2 mostram que as estimativas não são sensíveis ao tipo de pareamento utilizado para construir o GC 2. Nota-se na Tabela 20 no APÊNDICE C — Tabelas adicionais do capítulo 3 que os coeficientes estimados de 2015 a 2017 dos modelos com o GC 2 em suas versões alternativas (pareamento a três vizinhos com reposição e a cinco vizinhos sem reposição) são sempre negativos e têm magnitude semelhante à dos resultados principais, apresentados na primeira coluna da tabela, e são também estatisticamente diferentes de 0 ao nível de significância de 10%.⁵²

Os resultados do modelo de diferença-em-diferenças apresentados nesta seção apontam para a existência de impacto do rompimento da Barragem de Fundão sobre os municípios do grupo atingido, no sentido de ter provocado, em média, uma redução no número de vínculos ativos além da que seria observada se o rompimento não tivesse ocorrido. A evidência que leva a tal conclusão, no entanto, advém somente da comparação dos municípios atingidos com o grupo de comparação 2, o qual é formado exclusivamente por municípios de estados vizinhos a MG e ES. Por sua vez, as estimativas do modelo que compara o grupo atingido ao GC 1, cuja composição é restrita a municípios não atingidos de MG, ES, sul da BA e norte do RJ, não mostram

⁵⁰ Os valores de perda de vínculos ativos apresentados para o conjunto dos 45 municípios atingidos foram calculados a partir da perda média por município multiplicada pela quantidade de municípios atingidos. Ou seja, para o ano de 2015 (em que foi estimado uma perda de 300 postos de trabalho), o cálculo foi: $300 \cdot 45 = 13.500$.

⁵¹ Ressalte-se que, conforme discutido na seção 3.1.2, a inexistência de evidência de tendências prévias divergentes não significa a confirmação da hipótese de que o GC 2 representaria uma boa aproximação à situação contrafactual do grupo atingido. Trata-se apenas de um indício favorável à sua credibilidade.

⁵² As variáveis referentes aos impactos do rompimento em 2015, 2016 e 2017 são ainda estatisticamente diferentes de 0 ao nível de 5% de significância, exceto pela que se refere a 2015, quando se utiliza o GC 2 construído a partir do pareamento a três vizinhos com reposição.

diferenças estatisticamente significantes entre eles em nenhum dos anos analisados. Tal divergência entre os resultados dos dois modelos levanta a possibilidade de que o impacto do rompimento sobre número de vínculos ativos não se restringiu aos municípios atingidos e afetou também o mercado formal de trabalho de outras localidades. Pode-se conjecturar que os municípios fora do grupo atingido mais suscetíveis a repercutir o impacto do rompimento seriam aqueles que possuem relações de interdependência econômica mais fortes com os atingidos, os quais provavelmente se localizam também em MG e ES. Essa possibilidade de “transbordamento” do impacto do rompimento é justamente a motivação por trás da construção do GC 2, conforme discutido na seção 3.1.4, a qual parte da premissa de que municípios geograficamente mais distantes da região impactada estariam menos expostos aos impactos do rompimento por ser mais remota a possibilidade de que guardariam relações econômicas relevantes diretamente com os municípios atingidos.

3.2.2 Logaritmo natural da massa salarial

O segundo indicador de impacto analisado, logaritmo natural da massa salarial, é descrito na Tabela 7 sem a efetiva aplicação do logaritmo, assim como feito para o indicador de impacto anterior na Tabela 6. A seguir, portanto, a estatística de média e seu desvio-padrão são apresentados em R\$ de dezembro de 2017 para os três grupos de municípios ao longo do período analisado. A Tabela 21 no APÊNDICE C — Tabelas adicionais do capítulo 3 contém as estatísticas referentes à massa salarial dos municípios transformada com o logaritmo natural.

Tabela 7 — Média e desvio-padrão da massa salarial por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	21.622.742	21.907.766	17.409.278
	Desvio-padrão	51.473.504	70.474.918	37.414.000
2011	Média	22.720.356	24.227.627	19.047.184
	Desvio-padrão	53.693.039	79.385.617	40.696.324
2012	Média	23.744.413	25.693.551	20.679.877
	Desvio-padrão	56.251.936	88.694.067	44.067.328
2013	Média	24.931.531	27.710.284	22.287.533
	Desvio-padrão	58.187.484	94.571.038	47.948.660
2014	Média	25.553.375	28.839.627	23.045.447
	Desvio-padrão	61.835.506	99.130.975	49.478.303
2015	Média	23.347.401	26.546.739	21.794.815
	Desvio-padrão	54.175.621	87.255.355	46.675.586
2016	Média	22.316.948	24.769.795	21.213.438
	Desvio-padrão	51.704.579	77.988.913	45.466.893
2017	Média	22.819.995	24.822.348	21.711.484

Ano	Estatística	Atingido	GC 1	GC 2
	Desvio-padrão	51.766.420	72.081.421	46.108.737

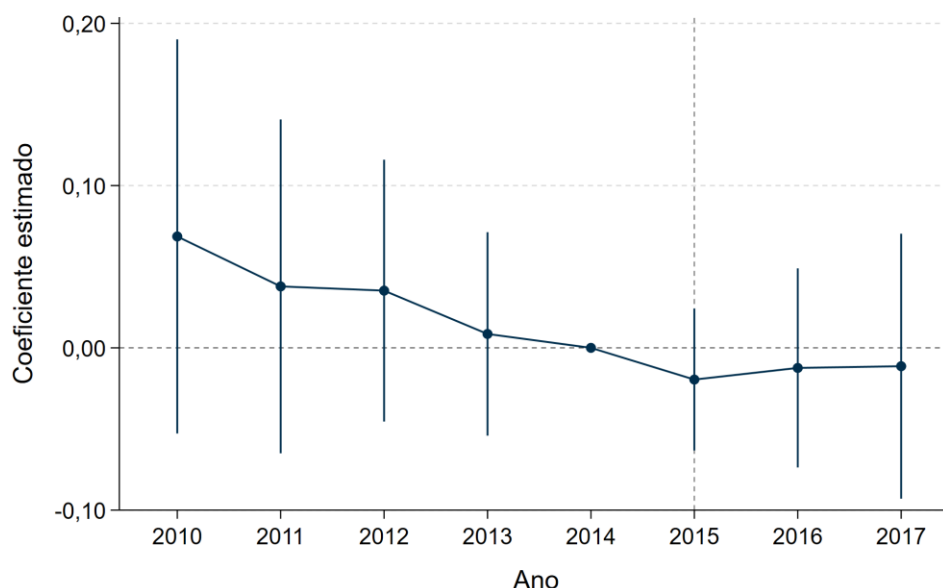
Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Os valores apresentados correspondem à variável em nível, medida em R\$ de dezembro de 2017 e não em logaritmo natural.

Por se tratar da soma das remunerações em dezembro dos vínculos ativos nos municípios em cada ano, a Tabela 7 possui características que remetem à Tabela 6. Os desvios-padrão aqui são como lá: sempre superiores ao dobro das respectivas médias, refletindo a considerável variabilidade entre os municípios dentro de cada grupo no que diz respeito ao tamanho do mercado de trabalho formal. Também os três grupos demonstram tendência de aumento da massa salarial média até 2014, com movimento de queda iniciado em 2015 interrompido por retomada do crescimento em 2017. É importante notar, no entanto, que as taxas de variação ao longo do tempo divergem das observadas para o número médio de vínculos ativos. Enquanto na Tabela 6 o grupo atingido e o GC 1 se mantêm relativamente próximos ao longo de todo o período, aqui se parte de uma diferença de R\$ 285 mil em 2010 para uma que supera R\$ 2 milhões em 2017, sendo 2014 o ano em que a massa salarial de ambos apresentou a maior discrepância entre eles (R\$ 3,28 milhões). Ou seja, a massa salarial do grupo atingido cresceu menos que a do GC 1 – especificamente, 5,5% de 2010 a 2017 contra 13,3% do GC 1 no mesmo período. Este é o caso também quando se analisa o GC 2, uma vez que os R\$ 4,2 milhões em média que os atingidos tinham a mais em 2010 caem para R\$ 1,1 milhões em 2017, tendo havido um crescimento de 24,71% da massa salarial média do GC 2 no período.

Em primeira análise, portanto, os grupos de comparação parecem apresentar trajetórias distintas da do grupo atingido, o que poderia comprometer a estratégia de identificação de impactos do rompimento por meio do modelo de diferença-em-diferenças. O Gráfico 3 mostra, no entanto, que em nenhum dos anos anteriores ao rompimento o logaritmo da massa salarial dos atingidos se altera frente ao do GC 1 de modo a gerar uma discrepância entre eles que seja estatisticamente diferente daquela observada em 2014. Isso significa que o modelo estima as diferenças entre grupo atingido e GC 1 em taxa de crescimento da massa salarial como estatisticamente nulas ao nível de significância de 10%. Em outras palavras, não são encontradas evidências de que as trajetórias de evolução ao longo do tempo do indicador em questão fossem diferentes entre os grupos antes do desastre.

Gráfico 3 — Estimativas de impacto médio sobre o logaritmo natural da massa salarial de municípios no grupo atingido (comparando com GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

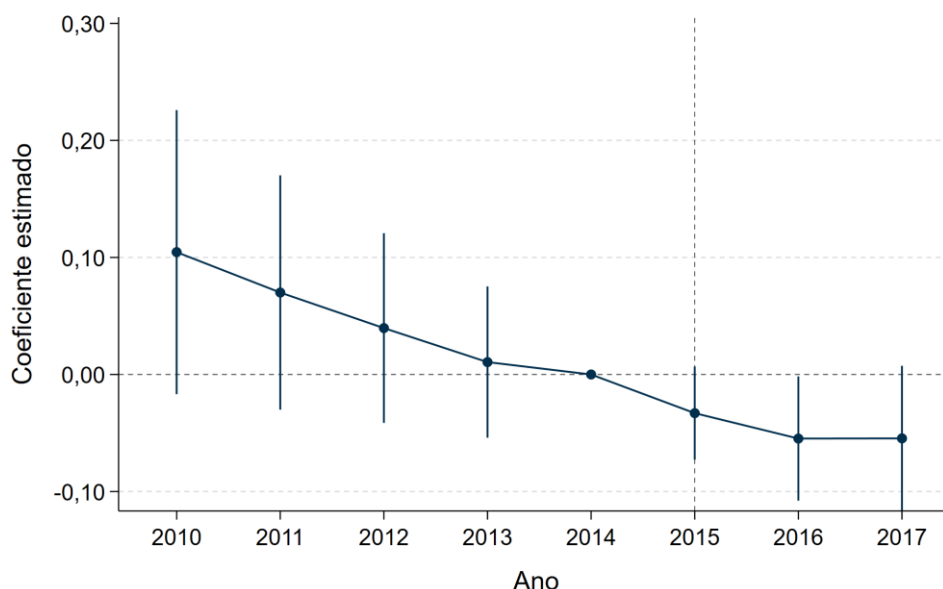
Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014.

As estimativas no Gráfico 3 não permitem afirmar a existência de impacto, em média, do rompimento da Barragem de Fundão sobre a massa salarial dos municípios atingidos quando estes são comparados ao GC 1. Isso porque, embora os coeficientes estimados referentes às variáveis do diferença-em-diferenças após 2014 sejam sempre negativos, elas não são estatisticamente significantes ao nível de 10%. Os modelos com versões alternativas do GC 1 na Tabela 22 vão no mesmo sentido, uma vez que neles também os coeficientes estimados referentes aos anos de 2015 a 2017 são sempre negativos e estatisticamente iguais a 0 ao nível de 10%. Entretanto, nota-se que a magnitude dos coeficientes oscila consideravelmente entre eles: a estimativa para 2017 varia de -0,011 no resultado principal (GC 1 com pareamento a três vizinhos sem reposição) até -0,0577 no modelo com o GC 1 formado por pareamento de três vizinhos com reposição. A sensibilidade das estimativas do diferença-em-diferenças à estratégia de construção do GC 1 levanta a possibilidade de que os resultados no Gráfico 3 não sejam conclusivos a respeito do impacto médio do rompimento sobre a massa salarial do grupo atingido, analogamente ao que se discutiu acima sobre o Gráfico 1.

O Gráfico 4 ilustra as estimativas do modelo quando o grupo atingido é comparado ao GC 2. Os pontos estimados de anos anteriores a 2014 são positivos e demonstram um

padrão decrescente acentuado que em alguns momentos quase se mostra estatisticamente significativa ao nível de 10%. Apesar de não o ser, provoca maior cautela quando da interpretação dos coeficientes estimados referentes aos anos de 2015 a 2017, os quais são todos negativos na ordem de -0,033 a -0,055. O intervalo de confiança da estimativa de 2016 não inclui o valor 0, o que em rigor faz com que a respectiva variável seja estatisticamente significativa ao nível de 10%. Entretanto, é perceptível o quão pequena é a margem que permite essa interpretação: o limite superior do intervalo de confiança é apenas -0,0016 – isto é, a massa salarial do grupo atingido seria em média aproximadamente 0,16% menor do que se tivesse acompanhado a trajetória da do GC 2.⁵³

Gráfico 4 — Estimativas de impacto médio sobre o logaritmo natural da massa salarial de municípios no grupo atingido (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014.

⁵³ O limite superior do intervalo de confiança de nível de significância de 10% para estimativa de 2016 é dado pelo cálculo $-0,0547 + 0,0321 \cdot 1,65341 = -0,00163$, onde $-0,0547$ é o ponto estimado e $0,0321$, seu erro-padrão, ambos informados pela Tabela 23. O número $1,65341$ corresponde ao valor da distribuição *t* de Student com 179 graus de liberdade que deixa 95% da massa probabilística à sua esquerda. A saber, o limite inferior desse intervalo de confiança é $-0,0547 - 0,0321 \cdot 1,65341 = -0,1077$. Note que a variável referente a 2016 é estatisticamente significativa ao nível de 10%, mas não a níveis menores, como 1% ou 5%.

Apesar dessas ressalvas, a Tabela 23 no APÊNDICE C — Tabelas adicionais do capítulo 3 mostra que os coeficientes estimados de 2016 e 2017 têm magnitudes similares entre o modelo principal e os que utilizam o GC 2 em suas versões alternativas de pareamento. Em particular, a variável que mede o impacto em 2016 é sempre estatisticamente diferente de 0 ao nível de significância de 10%, ao passo que nenhuma das referentes aos anos de 2010 a 2013 o é, independentemente da versão de pareamento utilizada para compor o GC 2. Portanto, há evidência favorável à hipótese de que, comparada à do GC 2, a massa salarial média do grupo atingido foi impactada em 2016 pelo rompimento da Barragem de Fundão. O efeito médio estimado de -0,055 indica que, desde 2014, ela foi cerca de 5,5% menor do que seria em seu cenário contrafactual, o qual, aproximado pela trajetória do GC 2, corresponderia a aproximadamente R\$ 23,6 milhões por município em R\$ de dezembro de 2017.⁵⁴ Comparando com o valor de R\$ 22,3 milhões para a massa salarial média do grupo atingido por município observado em 2016, conforme consta na Tabela 7, o rompimento teria levado a uma redução R\$ 1,3 milhões maior do que a que se observaria se ele não tivesse ocorrido. Tomando como base esta estimativa, a perda da massa salarial para o conjunto dos 45 municípios atingidos seria da ordem de R\$ 58,5 milhões, também em R\$ de dezembro de 2017.⁵⁵

De maneira semelhante aos resultados na seção 3.2.1, a redução em massa salarial como consequência do rompimento da Barragem de Fundão é constatada somente quando se compara o grupo atingido ao GC 2, constituído de municípios não atingidos nos estados vizinhos a MG e ES. O comentário feito na seção anterior se aplica aqui: é possível que o efeito do rompimento tenha repercutido nas localidades cujas economias são mais ligadas às dos municípios atingidos, as quais é razoável supor estão também em MG e ES. Não é surpreendente o fato de que os resultados para os dois indicadores de impacto analisados se repliquem nesse sentido, uma vez que ambos representam medidas do tamanho do mercado de trabalho formal dos municípios. É de se esperar que a magnitude da massa salarial de um município seja reflexo do total de vínculos ativos nele existente, uma vez que consiste na soma das remunerações destes.

⁵⁴ O cálculo que obtém R\$ 23,6 milhões é análogo ao demonstrado na nota de rodapé 48: $22.316.948 \cdot \left[1 - \frac{1}{1-0,055}\right] = 23.615.818$, sendo que o valor 22.316.948 corresponde à massa salarial média do grupo atingido observada em 2016, conforme consta na Tabela 7.

⁵⁵ O cálculo realizado de perda da massa salarial para o conjunto dos 45 municípios atingidos é análogo àquele realizado na nota de rodapé 50.

4 AVALIAÇÃO DE IMPACTOS SOBRE A POPULAÇÃO FORMALMENTE EMPREGADA NOS MUNICÍPIOS ATINGIDOS ANTES DO DESASTRE

A análise apresentada neste capítulo é focada nas pessoas que possuíam um vínculo de trabalho formal nos municípios atingidos em 31 de dezembro de 2014, um momento anterior ao desastre. O interesse nesse conjunto particular de indivíduos é motivado pela possibilidade de que as pessoas formalmente empregadas antes do rompimento da Barragem de Fundão possam ter sido impactadas de formas ou com intensidades diferentes das estimadas no capítulo anterior para o mercado formal como um todo dos municípios atingidos.⁵⁶

Não é possível avaliar tal hipótese apenas com os resultados do Capítulo 3.1, porque, para qualquer um dos anos após o rompimento, eles agregam sob um único coeficiente os efeitos não só sobre quem trabalhava com vínculo formal, mas também sobre indivíduos que apenas viriam a ingressar no mercado formal de trabalho após o rompimento, porém deixaram de fazê-lo devido à redução do número de postos de trabalho abertos. Consequentemente, não é claro se o efeito predominante do rompimento sobre os municípios atingidos é no sentido de extinguir vínculos formais antes existentes ou no de inviabilizar a criação de novos postos de trabalho. Em contraposição, um estudo voltado apenas ao conjunto de indivíduos empregados antes do rompimento permite uma investigação mais profunda desse aspecto, já que acompanhar a trajetória das pessoas ao longo do tempo torna possível avaliar se, em média, esse grupo especificamente teve maior dificuldade em se manter no mercado de trabalho formal após o rompimento ou, ainda, se passou a receber remunerações menores.

A próxima seção descreve o banco de dados construído para viabilizar a avaliação de impactos em nível individual, além das modificações realizadas na estratégia empírica para acomodar a mudança na ótica de análise. As seções seguintes avançam à definição dos indicadores de impacto analisados e, por fim, à apresentação dos resultados obtidos.

⁵⁶ Como seria de se esperar, há casos de pessoas registradas na RAIS com um vínculo de trabalho formal vigente em 31 de dezembro de 2014 que não mais existia em 5 de novembro de 2015. A escolha dessa primeira data, no entanto, é baseada em motivos técnicos, conforme se discute na seção 4.1.

4.1 Metodologia

O conjunto de indivíduos com vínculo de trabalho ativo em 2014 – isto é, um vínculo formal de trabalho vigente em 31 de dezembro de 2014, cuja remuneração de dezembro registrada era maior que 0, seguindo a definição apresentada na seção 3.1.5.1 – constitui a coorte de interesse do estudo desenvolvido neste capítulo.

Note que, com o fim de acompanhar ao longo do tempo as pessoas que efetivamente tinham um vínculo ativo quando ocorreu o rompimento da Barragem de Fundão, o ideal seria que a coorte fosse definida com data de referência no dia 5 de novembro de 2015. Em rigor, é de fato possível com os microdados da RAIS adotar a data do rompimento como momento de referência ao invés de 31 de dezembro de 2014. Todavia, o dado de remuneração que está disponível para os anos em análise e que é utilizado neste relatório diz respeito ao mês de dezembro de cada ano, de forma que pessoas cujos vínculos existissem em 5 de novembro de 2015 e não mais em 31 de dezembro de 2015 estariam sem informação de salário. Assim, o dado de salário anterior ao desastre que mais se aproxima da data em que o rompimento efetivamente ocorreu é de dezembro de 2014, motivo pelo qual o dia 31 de dezembro daquele ano foi escolhido para a definição da coorte.

Para realizar a avaliação econométrica de impactos do rompimento sobre a coorte de indivíduos, os microdados da RAIS são organizados de maneira a produzir um banco de dados no nível de pessoas em estrutura de painel, no qual a periodicidade é anual e referente a 31 de dezembro de cada ano. A construção do painel é apresentada em detalhes no APÊNDICE A — Construção do painel de indivíduos. Tem-se então que cada pessoa que possuía ao menos um vínculo ativo em 2014 é acompanhada no banco de dados em todos os anos de 2010 a 2017. No entanto, diferentemente do painel de municípios construído para o Capítulo 3 (no qual todos os municípios são observados em todos os períodos), a pessoa pode ou não aparecer com um vínculo de trabalho ativo nas observações dos demais anos. Quando ocorre o fato de um indivíduo da coorte não ter vínculo ativo em determinado ano, é impossível saber sua situação laboral naquele momento, uma vez que diversos cenários são plausíveis: pode se encontrar desempregado, em um emprego informal, atuando como profissional autônomo ou envolvido em qualquer outra situação que não constitua um vínculo formal de trabalho ativo registrado na RAIS. É importante ter em mente, portanto, que a inexistência de vínculo ativo no mercado formal não implica necessariamente em condição de desemprego.

Uma pessoa da coorte é considerada atingida se o seu vínculo de trabalho ativo em 2014 era com um estabelecimento localizado em um dos 45 municípios do grupo atingido, conforme definido na seção 3.1.3. Por sua vez, uma pessoa faz parte do grupo de comparação da análise se o seu vínculo de trabalho ativo em 2014 era com um estabelecimento localizado em municípios incluídos em um dos grupos de comparação construídos por meio de pareamento com distância de Mahalanobis. Ou seja, a definição de pessoas no grupo atingido ou de comparação deriva da classificação que os municípios nos quais trabalhavam em 2014 receberam na seção 3.1.4⁵⁷, obtida por meio do procedimento de pareamento com distância de Mahalanobis a três vizinhos mais próximos sem reposição. Portanto, há na análise de coorte dois grupos de comparação:

- Grupo de comparação 1 (GC 1): pessoas com vínculo formal de trabalho ativo em 2014 – isto é, empregadas em 31 de dezembro e com remuneração de dezembro não nula – em estabelecimentos localizados em municípios não atingidos de Minas Gerais, Espírito Santo, sul da Bahia e norte do Rio de Janeiro que foram pareados aos atingidos;
- Grupo de comparação 2 (GC 2): pessoas com vínculo formal de trabalho ativo em 2014 – isto é, empregadas em 31 de dezembro e com remuneração de dezembro não nula – em estabelecimentos localizados em municípios da Bahia, Goiás, Mato Grosso do Sul, Rio de Janeiro e São Paulo que foram pareados aos atingidos.

Assim como para as estimativas do Capítulo 33.1, os resultados na seção 4.2 são submetidos a análises de robustez em que os grupos de comparação são construídos com base em municípios não atingidos selecionados por pareamento de cinco vizinhos sem reposição e de três vizinhos com reposição.⁵⁸

Também como no capítulo anterior, o modelo de diferença-em-diferenças é o método econométrico escolhido para avaliar os impactos do rompimento da Barragem de Fundão sobre as pessoas na coorte cujo vínculo ativo em 2014 era em municípios atingidos. O modelo é estimado via regressão linear e tem a seguinte especificação:

⁵⁷ A abordagem de determinar se uma pessoa integra o grupo atingido ou de comparação com base em definições construídas em nível de municípios é executada também em Deryugina et al. (2018) e Groen et al. (2020) para avaliar os impactos das tempestades de 2005 nos EUA (dentre elas, o furacão Katrina) sobre o mercado de trabalho das regiões atingidas, conforme apresentado na seção 1.3.

⁵⁸ Os resultados da análise de robustez são apresentados no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

Equação 2 — Modelo de regressão para o método de diferença-em-diferenças na análise de coorte

$$Y_{i,t} = \alpha + \sum_{\substack{j=2010, \\ j \neq 2014}}^{2017} \beta_j \cdot \mathbb{I}_{\{i \text{ é atingido}\}} \cdot \mathbb{I}_{\{j=t\}} + \gamma_i + \lambda_t + \varepsilon_{i,t}$$

Trata-se do mesmo modelo apresentado na seção 3.1.2, com a exceção de que aqui o nível de observação é o indivíduo i em 31 de dezembro do ano t , o que se reflete nos indexadores dos elementos da Equação 2. Consequentemente, como na Equação 1, o modelo contém sete parâmetros que correspondem a estimadores de diferença-em-diferenças – β_{2010} , β_{2011} , β_{2012} , β_{2013} , β_{2015} , β_{2016} e β_{2017} –, sendo que os de interesse para a avaliação de impactos do rompimento da Barragem de Fundão são os de 2015, 2016 e 2017. Satisfeitas as hipóteses de identificação do método, esses três parâmetros têm a interpretação de impacto médio do rompimento sobre as pessoas que em 2014 tinham vínculo de trabalho ativo em municípios atingidos, informando o quanto o rompimento alterou seus indicadores de impacto frente aos do grupo de comparação com relação à situação que se observava em 31 de dezembro de 2014.

Os parâmetros de diferença-em-diferenças referentes a anos anteriores ao desastre são novamente relevantes por permitir verificar se a trajetória dos indicadores de impacto das pessoas no grupo atingido e das do grupo de comparação era similar antes do rompimento. Isso serve de critério para avaliar a plausibilidade da hipótese de que os dois grupos teriam evolução semelhante caso o rompimento não tivesse ocorrido. Por fim, os intervalos de confiança das estimativas são robustos à autocorrelação municipal⁵⁹, evitando assim que sejam distorcidos no caso de os indicadores de impacto de pessoas de um mesmo município serem correlacionados entre si e/ou parcialmente determinados por seus valores em anos anteriores.

4.1.1 Caracterização dos indivíduos atingidos na coorte

Antes de proceder à definição dos indicadores de impacto analisados e aos resultados dos modelos estimados, esta seção traz um breve panorama sobre os indivíduos na coorte, em particular aqueles que integram o grupo atingido. A Tabela 8 mostra quantas pessoas possuíam um vínculo de trabalho ativo – isto é, vigente em 31 de dezembro e com remuneração não nula, conforme definição na seção 3.1.5.1 – em cada ano entre

⁵⁹ Os *clusters* são formados pelos municípios nos quais as pessoas da coorte trabalhavam em 31 de dezembro de 2014.

2010 e 2017. Os números apresentados ilustram não apenas o grupo atingido, mas também os de comparação 1 e 2.

Tabela 8 — Número de pessoas da coorte com vínculo de trabalho ativo por grupo

Ano	Atingido	GC 1	GC 2
2010	295.024	859.754	781.496
2011	318.848	933.073	851.547
2012	340.363	994.934	917.117
2013	373.764	1.093.966	1.013.983
2014	495.392	1.441.898	1.359.404
2015	372.237	1.099.245	1.036.278
2016	333.020	979.673	932.091
2017	317.894	940.906	886.507

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Os vínculos de trabalho considerados “ativos” são apenas os que constam na RAIS como vigentes em 31 de dezembro de cada ano com remuneração não nula.

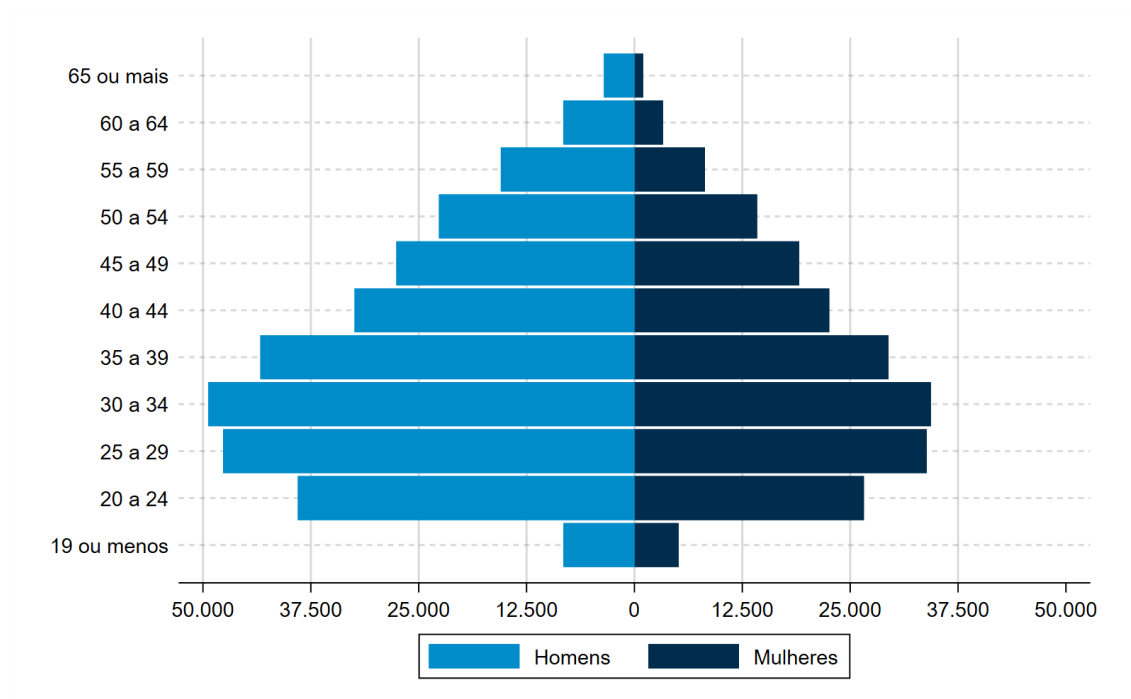
Para os três grupos, o ano de 2014 é aquele em que o maior número de pessoas se encontrava com vínculo de trabalho ativo, reflexo direto do fato de que a data de referência da coorte é 31 de dezembro desse ano. O número referente ao grupo atingido mostra que a análise deste capítulo diz respeito aos impactos do rompimento da Barragem de Fundão sobre a trajetória no mercado de trabalho formal de 495.392 pessoas.

Conforme se afasta de 2014, o número de pessoas com vínculos ativos diminui, ficando tão menor quão mais distante o momento de tempo considerado. Em 2010, por exemplo, mais de 40% dos indivíduos dos três grupos não foram localizados nos microdados da RAIS ou foram, porém com um vínculo de trabalho encerrado antes de 31 de dezembro daquele ano. Com efeito, a taxa de queda do número de pessoas com vínculos ativos em relação a 2014 é sempre semelhante entre os três grupos. De 2014 para 2013, a redução é de 24,6% no grupo atingido, de 24,1% no GC 1 e de 25,4% no GC 2; de 2013 para 2012 é de 8,9%, 9% e 9,6%, respectivamente. Mesmo após o rompimento, a redução de 2014 para 2015 é de 24,9% no grupo atingido, sendo de 23,8% para os dois grupos de comparação; de 2016 para 2017 é de 4,5%, 4% e 4,9%, respectivamente. Tal tendência de queda não é surpreendente considerando a maneira como foi construído o painel, uma vez que o número de pessoas com vínculos ativos nunca pode superar o de 2014.⁶⁰

⁶⁰ Mais detalhes no APÊNDICE A — Construção do painel de indivíduos.

Com foco específico nas pessoas do grupo atingido, os gráficos e o mapa a seguir ilustram algumas de suas características conforme registradas pelos empregadores. Os dados utilizados se referem ao ano de 2014, no qual todas tinham vínculo de trabalho ativo. O Gráfico 5 mostra a distribuição das pessoas atingidas por idade e sexo, deixando claro que homens são maioria em todas as faixas etárias. No total, as mulheres representam 39,9% dos atingidos (197.795 pessoas), percentual menor do que quando se compara com o mercado formal de MG ou de ES como um todo, apresentado na Tabela 24⁶¹, no qual as mulheres ocupavam, respectivamente, 43,6% e 42,9% dos vínculos ativos segundo dados da RAIS de 2014.

Gráfico 5 — Pirâmide etária do grupo atingido



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

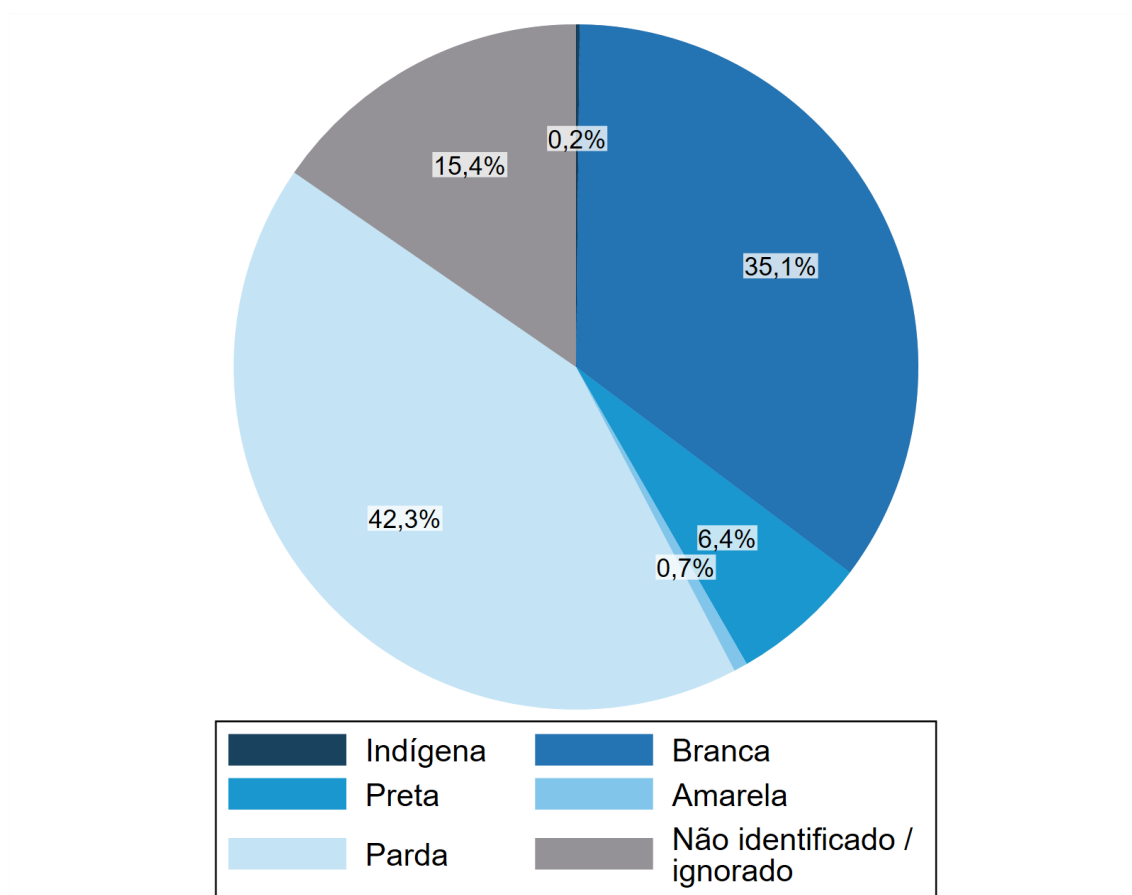
Observação: Informações se referem a pessoas com vínculos de trabalho vigentes em 31 de dezembro de 2014 e com remuneração não nula.

Ainda, a pirâmide etária mostra que a distribuição de trabalhadores por idade é concentrada nas faixas que abrangem 20 a 39 anos, contendo 61,3% (303.685 indivíduos) do total de pessoas da coorte no grupo atingido. Adicionalmente, há proporcionalmente menos mulheres que homens com mais de 60 anos registrados com vínculos ativos em 2014 (5,8% e 8%, respectivamente).

⁶¹ A Tabela 24 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 apresenta estatísticas sobre vínculos ativos em 2014 para os estados de MG e ES.

O Gráfico 6 ilustra a distribuição de pessoas por raça/cor registrada pelo empregador. Uma parcela considerável dos atingidos que compõem a coorte foi registrada com raça/cor “não identificada/ignorada”, embora esta parcela seja inferior à que se observa para os estados de MG e ES, respectivamente de 23,2% e 22,1%.⁶² Esta, porém, não é a única distinção relevante dos atingidos que compõem a coorte em relação ao que se vê para os estados, uma vez que a proporção de pessoas do grupo atingido registradas como pardas em 2014 era de 42,3% (209.305 indivíduos), enquanto em MG foi de 28,3% e em ES, 38,8%. A proporção pequena de indígenas entre os atingidos que compõem a coorte, por outro lado, é similar à registrada para MG e ES.

Gráfico 6 — Distribuição de pessoas no grupo atingido por raça/cor



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

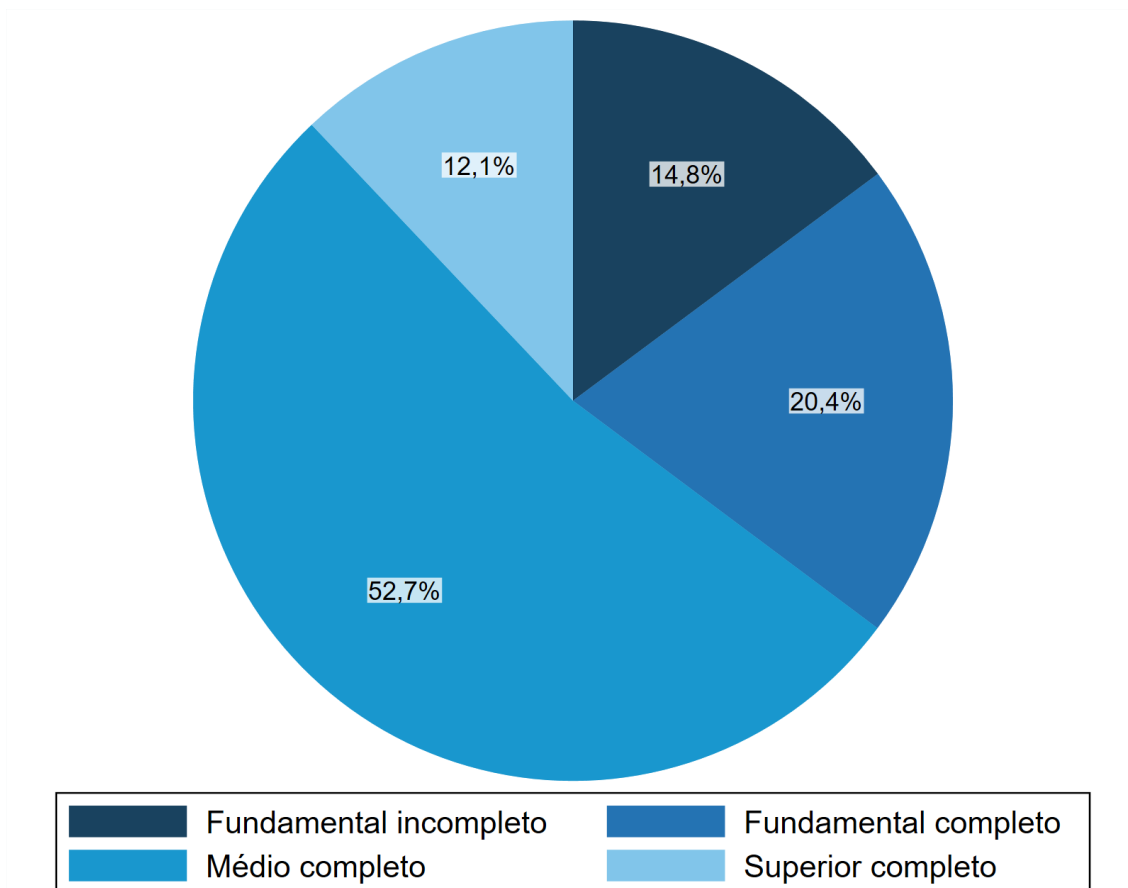
Observação: Informações se referem a pessoas com vínculos de trabalho vigentes em 31 de dezembro de 2014 e com remuneração não nula.

O Gráfico 7 se volta ao nível de escolaridade dos atingidos da coorte em 2014 e mostra que mais da metade deles possuía ensino médio completo (261.166 indivíduos),

⁶² Informações disponíveis na Tabela 24 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

percentual superior ao observado para MG e ES na Tabela 24. A parcela de pessoas com ensino fundamental completo também é maior entre as pessoas do grupo atingido, assim como a parcela sem fundamental completo quando comparado ao ES. A proporção de pessoas com ensino superior completo, por sua vez, é mais de cinco pontos percentuais menor que as dos estados.

Gráfico 7 — Distribuição de pessoas no grupo atingido por nível de escolaridade



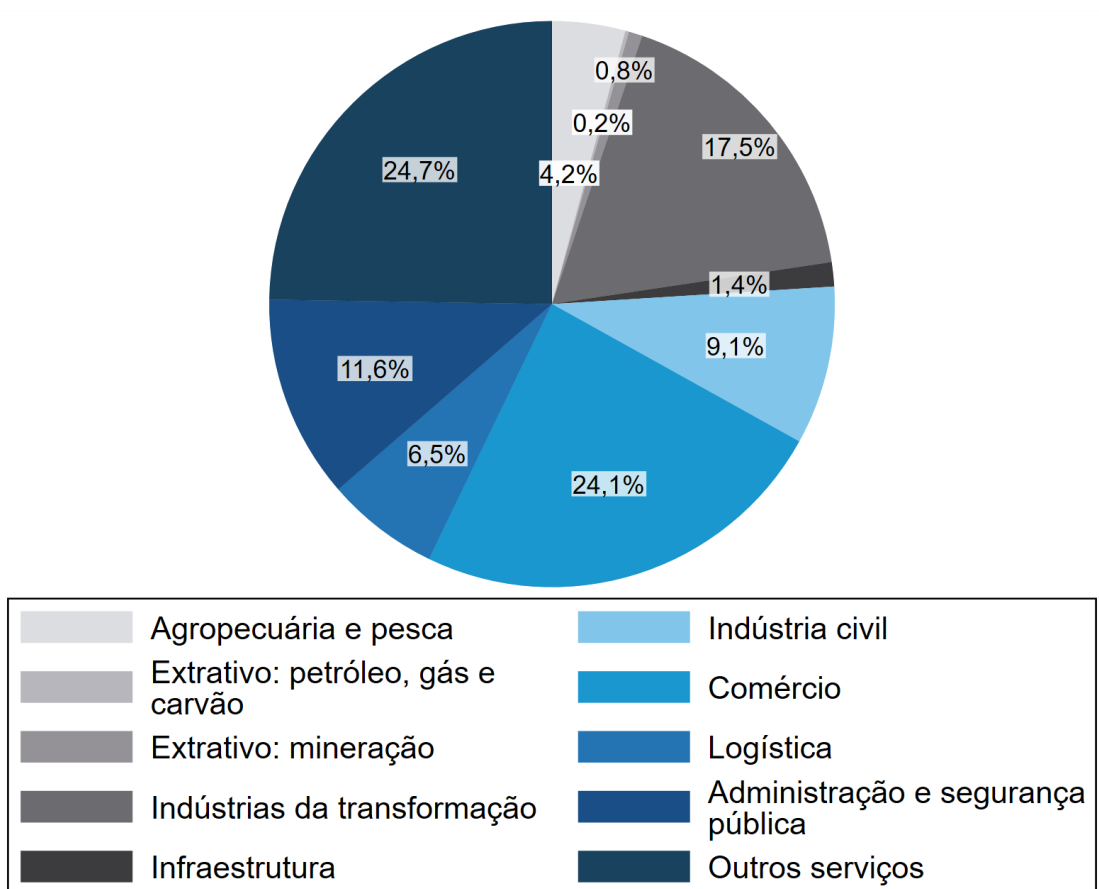
Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).
Observação: Informações se referem a pessoas com vínculos de trabalho vigentes em 31 de dezembro de 2014 e com remuneração não nula.

O Gráfico 8 agrupa as pessoas pelo setor de atividade em que trabalhavam em 31 de dezembro de 2014 com base no código CNAE 2.0 informado pelo empregador.⁶³ Quase um quarto dos atingidos da coorte tinha emprego ligado ao comércio (119.371 indivíduos), enquanto outros 24,7% (122.598 indivíduos) trabalhavam em “outros serviços”, o que inclui áreas como restaurantes, bancos, educação e saúde. Os outros

⁶³ A Tabela 25 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 contém a correspondência entre os códigos da CNAE 2.0 e a classificação de atividades utilizada no Gráfico 8 e na Tabela 24, esta última no mesmo apêndice.

dois setores que concentravam fatias superiores a 10% do grupo atingido são o de indústrias da transformação, com 17,5% das pessoas (86.499 indivíduos), e administração e segurança pública, com 11,6% (57.541 indivíduos). São também esses quatro setores que em 2014 mais concentraram vínculos de trabalho ativos nos estados de MG e ES, com a diferença de que neles as áreas de administração e segurança pública e indústrias da transformação apareceram em posições inversas.⁶⁴

Gráfico 8 — Distribuição de pessoas no grupo atingido por setor de atividade



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

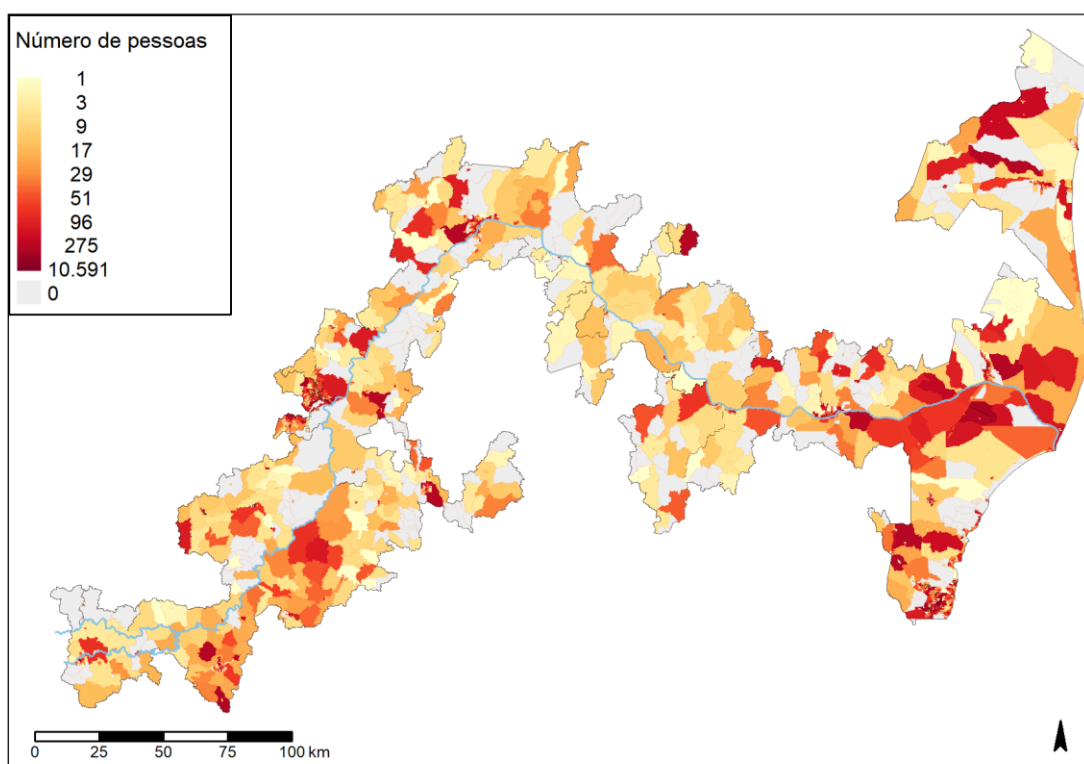
Observação: Informações se referem a pessoas com vínculos de trabalho vigentes em 31 de dezembro de 2014 e com remuneração não nula.

Por fim, a Figura 5 apresenta um mapa dos municípios atingidos com as divisões de setores censitários usadas no Censo Demográfico de 2010. Cada setor sinaliza por meio de sua cor o número de pessoas do grupo atingido da coorte que nele se localizavam em 2014. Note que a localização não se refere ao endereço domiciliar da pessoa, mas

⁶⁴ Informações disponíveis na Tabela 24 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

sim ao endereço do estabelecimento com o qual a pessoa tinha vínculo ativo em 31 de dezembro de 2014.⁶⁵ A distribuição de pessoas entre os 3.441 setores (2.838 urbanos e 603 rurais) é claramente desigual, sendo que metade deles está associado a menos de 30 trabalhadores. Especificamente entre setores rurais, o setor mediano possui apenas 11 trabalhadores. Essas estatísticas, no entanto, desconsideram os 573 setores (16,6% dos 3.441) sem nenhum trabalhador associado. Embora 361 dos setores sem trabalhadores encontrados sejam urbanos, os 212 restantes representam 35,2% do total de setores rurais.

Figura 5 — Distribuição de pessoas no grupo atingido por setores censitários



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME), CNEFE (IBGE) e de bases cartográficas (ANA e IBGE).

Observação: Informações se referem a pessoas com vínculos de trabalho vigentes em 31 de dezembro de 2014 e com remuneração não nula. A cada pessoa se atribuiu a localização do estabelecimento com o qual tinha vínculo de trabalho em 31 de dezembro de 2014. As linhas em azul-claro ilustram o Rio Gualaxo do Norte, Rio do Carmo e Rio Doce.

⁶⁵ Os microdados da RAIS apresentam o endereço completo de todos os estabelecimentos declarantes. A partir dessa informação, utilizou-se um algoritmo que a cruzou com o CNEFE e bases georreferenciadas do IBGE para atribuir cada estabelecimento a um setor censitário. O procedimento está descrito no APÊNDICE B — Atribuição de estabelecimentos a setores censitários. Dos 495.392 indivíduos da coorte no grupo atingido, 15.621 (3,2%) não puderam ser alocados em um setor censitário devido à impossibilidade de localizar nas bases do IBGE os endereços dos estabelecimentos a que estavam associados.

Com relação aos setores censitários que margeiam o Rio Gualaxo do Norte, Rio do Carmo e Rio Doce (ilustrados pelas linhas em azul-claro na Figura 5), não se tem clara a tendência a concentrar mais ou menos pessoas da coorte. Considerando a região como um todo, proximidade aos rios não parece ser um fator particularmente importante para o desenvolvimento de empreendimentos formais, seja em quantidade ou tamanho. Entretanto, à medida que se avança no território do ES em direção ao litoral, nota-se que os setores censitários parecem se tornar sistematicamente mais densos em torno do Rio Doce.

4.1.2 Indicadores de impacto analisados

Os indicadores de impacto escolhidos para investigar os efeitos do rompimento sobre a coorte de trabalhadores descrita anteriormente contemplam, assim como a análise desenvolvida no Capítulo 3.13, questões relacionadas a postos de trabalho e remuneração no mercado formal. No entanto, a ótica adotada aqui diz respeito ao indivíduo, não mais a municípios, de modo que os indicadores de impacto apresentados a seguir descrevem a trajetória particular de cada pessoa enquanto formalmente empregada.

Em acordo com os indicadores de impacto definidos da seção 3.1.5, o termo “vínculo ativo” em determinado ano se refere a um vínculo de trabalho formal com remuneração de dezembro não nula que consta na RAIS como vigente em 31 de dezembro daquele ano, e os indicadores de impacto que tratam de salários se baseiam em valores convertidos para R\$ de dezembro de 2017 usando o IPCA (IBGE). Aos salários foi aplicado ainda um procedimento de censura de dados conhecido como *winsorizing*: todas as observações de salários superiores ao 99º percentil da distribuição de remunerações não nulas em 31 de dezembro foram igualadas ao valor desse percentil. Simetricamente, todas as observações de salários inferiores ao 1º percentil da mesma distribuição foram igualadas ao valor desse percentil.

4.1.2.1 Vínculo ativo no ano

Para um indivíduo i observado no ano t , este indicador de impacto corresponde a uma variável indicadora (*dummy*) que assume valor 1 se o indivíduo teve um vínculo de trabalho ativo em t , independentemente de qual setor de atividade ou município de localização, e 0 em caso contrário. Note que, pela maneira como é definida a coorte de interesse, este indicador vale 1 em 2014 para qualquer indivíduo, já que todos necessariamente tinham um vínculo ativo na data de referência da coorte.

A estimativa de impacto obtida pelo modelo de regressão aplicado a esse indicador tem a interpretação de efeito médio sobre a probabilidade de um indivíduo atingido da coorte ter vínculo de trabalho formal ativo no ano t , o qual é medido em pontos percentuais.⁶⁶

4.1.2.2 Salário condicional a vínculo ativo no ano

Para um indivíduo i observado no ano t , este indicador de impacto corresponde ao valor da remuneração em dezembro de t associada ao vínculo de trabalho ativo de i naquele ano. Note que esta definição implica que, se no ano t uma pessoa não teve vínculo ativo, ela não é considerada na análise de impacto do rompimento sobre este indicador.

4.1.2.3 Vínculo ativo no ano com o mesmo estabelecimento de 2014

Para um indivíduo i observado no ano t , este indicador de impacto corresponde a uma variável indicadora (*dummy*) que vale 1 se ele teve vínculo ativo em t junto ao estabelecimento no qual estava empregado também em 31 de dezembro de 2014. O indicador assume o valor 0 se o indivíduo i não teve vínculo formal ativo no ano t ou se teve com um estabelecimento diferente do de 2014.

Como na seção 4.1.2.1, a estimativa de impacto obtida pelo modelo de regressão tem a interpretação de efeito médio sobre uma probabilidade, medido em pontos percentuais.

4.1.2.4 Vínculo ativo no ano no mesmo setor de atividade econômica de 2014

Para um indivíduo i observado no ano t , este indicador de impacto corresponde a uma variável indicadora (*dummy*) que vale 1 se ele teve vínculo ativo em t junto a um estabelecimento no mesmo setor de atividade econômica daquele em que estava empregado em 31 de dezembro de 2014.⁶⁷ O indicador assume o valor 0 se o indivíduo i não teve vínculo formal ativo no ano t ou se teve junto a estabelecimento em um setor de atividade econômica diferente daquele de 2014.

⁶⁶ Por se tratar de uma variável binária, o modelo expresso na Equação 2 quando aplicado a esse indicador de impacto tem a denominação específica de “modelo de probabilidade linear”.

⁶⁷ Para fins de construção desse indicador de impacto, considerou-se o “setor de atividade” de um indivíduo os primeiros três dígitos do código da CNAE 2.0 associado ao seu vínculo de trabalho, o que corresponde ao que o IBGE denota por “grupo de atividade” (IBGE, 2007).

Como na seção 4.1.2.1, a estimativa de impacto obtida pelo modelo de regressão tem a interpretação de efeito médio sobre uma probabilidade, medido em pontos percentuais.

4.1.2.5 Salário condicional a vínculo ativo no ano no mesmo setor de atividade econômica de 2014

Para um indivíduo i observado no ano t , este indicador de impacto corresponde ao valor da remuneração em dezembro de t associada ao vínculo de trabalho ativo de i naquele ano se este tiver sido junto a estabelecimento no mesmo setor de atividade econômica daquele em que estava empregado em 31 de dezembro de 2014. Ou seja, se no ano t uma pessoa não teve vínculo formal ativo ou teve com estabelecimento em outro setor de atividade econômica, ela não é considerada na análise de impacto sobre este indicador.

4.2 Resultados

Esta seção apresenta os resultados da avaliação de impactos do rompimento da Barragem de Fundão sobre os indivíduos da coorte que integram o grupo atingido. De maneira similar à seção 3.2, as estimativas da Equação 2 para os cinco indicadores de impacto definidos na seção anterior são apresentadas separadamente. As análises são baseadas na comparação do grupo atingido com os grupos de comparação 1 e 2 (GC 1 e GC 2, respectivamente) construídos por meio de pareamento de municípios aos três vizinhos mais próximos sem reposição. Para permitir avaliar a sensibilidade dos resultados a mudanças na construção do grupo de comparação, tabelas com resultados usando as versões de três vizinhos com reposição e cinco vizinhos sem reposição do GC 1 e do GC 2 são apresentadas no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

Em complemento aos resultados principais, a discussão a seguir traz ainda, quando pertinente, menções a modelos estimados em subgrupos de indivíduos da coorte. A motivação por trás de tais modelos é investigar a possibilidade de efeitos heterogêneos do rompimento da Barragem de Fundão entre os atingidos. Por exemplo, avaliar se existe uma faixa etária específica na qual o impacto médio sobre salário é diferente das demais, ou se a trajetória de mulheres e de homens no mercado de trabalho formal é afetada de maneira diferente pelo rompimento. Os subgrupos são definidos com base nas características das pessoas registradas em seus vínculos de trabalho de 2014. São eles:

- Sexo;
- Raça/cor;
- Setor de atividade (conforme Tabela 25 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4);
- Faixa etária;
- Nível de escolaridade;
- Pessoas do grupo atingido empregadas em setores censitários a no máximo 1, 5 ou 10 km do Rio Gualaxo do Norte, Rio do Carmo e Rio Doce (conforme Figura 6 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4);
- Pessoas do grupo atingido empregadas em municípios divididos por recorte territorial (Alto, Médio ou Baixo Rio Doce, conforme Figura 7 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4).

As tabelas referentes aos modelos por subgrupos explicitamente citados ao longo desta seção constam no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4. Embora não apresentados no relatório, os demais resultados serão prontamente disponibilizados mediante solicitação aos autores.

4.2.1 Vínculo ativo no ano

A Tabela 9 permite observar o comportamento do indicador de impacto “vínculo ativo no ano” no grupo atingido, GC 1 e GC 2 ao longo do período de análise. Por se tratar de uma variável *dummy* (isto é, binária que assume apenas os valores 0 ou 1), a estatística de média referente a ela revela a fração de pessoas de cada grupo que possuía vínculo de trabalho ativo no ano indicado. Em 2010, por exemplo, 59,6% das pessoas da coorte no grupo atingido tinham vínculo ativo, parcela similar à do GC 1 e mais de dois pontos percentuais (p. p.) acima da do GC 2. O desvio-padrão da média, por sua vez, é tão mais alto quão mais próximo a média for de 0,5, consequência novamente do fato de o indicador de impacto ser uma variável *dummy*.⁶⁸

⁶⁸ O menor valor que o desvio-padrão da média de uma variável *dummy* pode atingir é 0, que ocorre quando todas as pessoas do grupo (independentemente de ser atingido ou de comparação) têm o mesmo valor no indicador de impacto em determinado ano – no caso, nenhuma com vínculo ativo (média = 0) ou todas com vínculo ativo (média = 1). Seu maior valor possível é de 0,5, que ocorre quando o número de pessoas no grupo sem vínculo ativo no ano é igual ao das que têm (média = 0,5) – ou seja, 50% das pessoas com indicador de impacto 0,5 maior que a média (indicador = 1) e 50% delas com indicador 0,5 menor que a média (indicador = 0).

Tabela 9 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	0,596	0,596	0,575
	Desvio-padrão	0,491	0,491	0,494
2011	Média	0,644	0,647	0,626
	Desvio-padrão	0,479	0,478	0,484
2012	Média	0,687	0,690	0,675
	Desvio-padrão	0,464	0,462	0,469
2013	Média	0,754	0,759	0,746
	Desvio-padrão	0,430	0,428	0,435
2014	Média	1	1	1
	Desvio-padrão	0	0	0
2015	Média	0,751	0,762	0,762
	Desvio-padrão	0,432	0,426	0,426
2016	Média	0,672	0,679	0,686
	Desvio-padrão	0,469	0,467	0,464
2017	Média	0,642	0,653	0,652
	Desvio-padrão	0,480	0,476	0,476

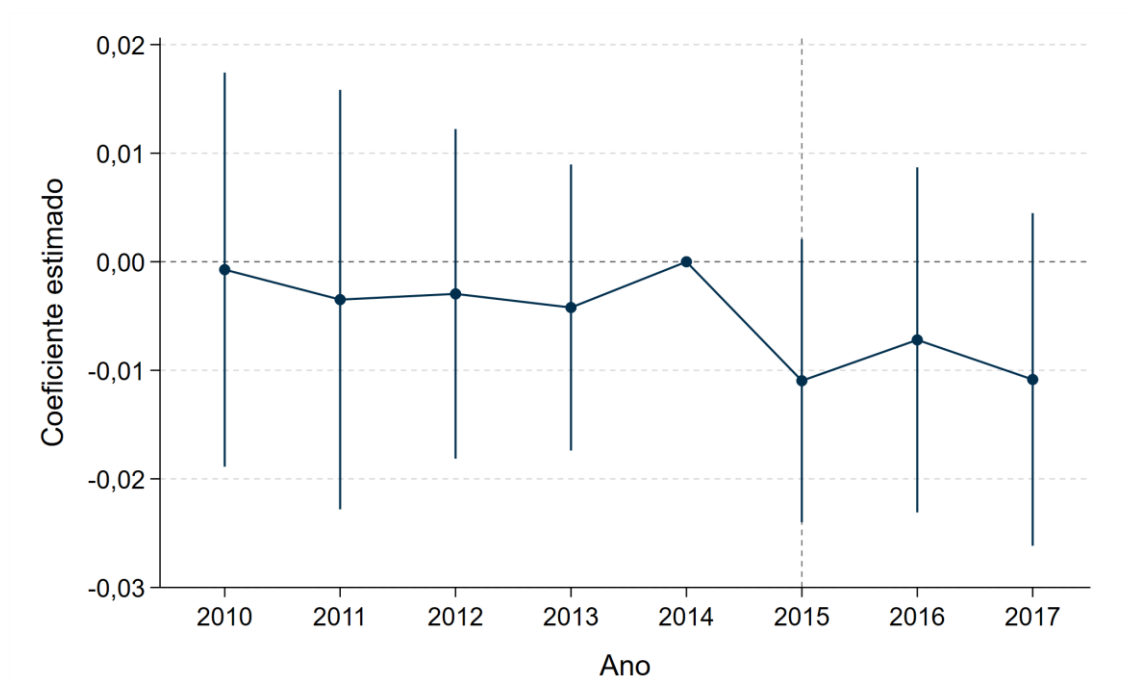
Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

No ano de 2014, os três grupos apresentam média igual a 1 e desvio-padrão nulo, reflexo direto do fato de que todas as pessoas da coorte, por definição, tiveram vínculo de trabalho ativo naquele ano. Nos demais anos, percebe-se que a parcela de pessoas com vínculo ativo se reduz conforme mais distante o momento em relação ao ano de referência (2014), comportamento já ilustrado na Tabela 8. Nos três momentos após o rompimento (2015 a 2017), nota-se que a proporção de pessoas no grupo atingido com vínculo ativo se mantém por volta de 1 p. p. abaixo das observadas nos grupos de comparação. Tal padrão é ausente nos anos que precedem o desastre, nos quais a média do indicador de impacto do grupo atingido se situava entre a do GC 1 e a do GC 2.

O Gráfico 9 ilustra as estimativas dos parâmetros β_{2010} a β_{2017} da Equação 2 quando o GC 1 é usado para avaliar os impactos do rompimento sobre o grupo atingido. Como na seção 3.2, os coeficientes estimados para os parâmetros do modelo correspondem aos pontos no gráfico e as linhas verticais aos respectivos intervalos de confiança, construídos sob o nível de significância de 10%. Novamente, quando um intervalo não atravessa a reta horizontal do valor 0, a variável à qual o intervalo está associado é estatisticamente significativo ao nível de 10%. Em outras palavras, a probabilidade de se estar cometendo um erro ao tratar tal estimativa como diferente de 0 quando na

verdade o impacto médio é nulo é igual ou inferior a 10%.⁶⁹ Os mesmos resultados são apresentados de forma mais completa na Tabela 26 do APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

Gráfico 9 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo (comparando com GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

Nota-se que as estimativas referentes aos anos anteriores ao rompimento da Barragem de Fundão, em especial as de 2011 a 2013, oscilam em torno de -0,003, indicando ser em média 0,3 p. p. mais baixa a probabilidade de indivíduos atingidos terem vínculo ativo do que a de pessoas do GC1 em comparação ao que era em 2014. Nos anos pós-rompimento (2015 a 2017), as estimativas passam a oscilar em um patamar inferior, ao redor de -0,01, cuja interpretação é de redução média de 1 p. p. na probabilidade de pessoas do grupo atingido terem vínculo de trabalho ativo em comparação à do GC 1 com relação a 2014. Apesar disso, os intervalos de confiança ilustram que nenhuma

⁶⁹ Expresso ainda de outra maneira, adotando terminologia mais técnica, a probabilidade de se rejeitar a hipótese nula do teste de que o parâmetro é igual a 0 quando ela é verdadeira é igual ou inferior a 10%.

destas variáveis pode ser considerada estatisticamente diferente de 0 ao nível de significância de 10%, uma vez que seus intervalos abrangem o valor 0.⁷⁰

O comportamento dos coeficientes estimados replica o observado para as estatísticas do grupo atingido e do GC 1 na Tabela 9: a diferença entre as médias lá apresentadas resulta em valores iguais aos dos coeficientes ilustrados no Gráfico 9. Isto se deve ao fato de que todas as pessoas da coorte tinham o mesmo valor de indicador de impacto no ano de referência, 2014. Nesse sentido, portanto, as estimativas do modelo de diferença-em-diferenças não trazem informação nova, mas tendo em consideração as características do método apresentadas na seção 3.1, fica claro que as evidências trazidas pelo modelo permitem uma avaliação mais criteriosa da relação que se pode estabelecer entre o rompimento da Barragem de Fundão e a trajetória do indicador de impacto do grupo atingido. Como nenhuma das variáveis que mede os impactos do rompimento se mostrou estatisticamente significativa, não se pode rejeitar a hipótese de que o rompimento tenha impacto nulo sobre a proporção de atingidos da coorte com vínculo ativo a partir de 2015 quando comparada à de indivíduos do GC 1 e em relação a 2014. Por outro lado, o resultado obtido também não pode ser usado como evidência de que o rompimento não teve, em média, impacto sobre a probabilidade de indivíduos atingidos terem um vínculo de trabalho ativo. O motivo para tal é que a Tabela 26 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 mostra que, quando se utiliza o GC 1 construído através de pareamento com reposição a três vizinhos mais próximos, são obtidas estimativas de redução entre 1,5 e 2,2 p. p. em 2015, 2016 e 2017, sendo elas efeitos estatisticamente significantes ao nível de 5%. Portanto, a estimação da Equação 2 para esse indicador de impacto se mostra sensível à maneira como é construído o GC 1, o que indica que os resultados ilustrados no Gráfico 9 não podem ser tomados como conclusivos sobre a inexistência de impacto nessa dimensão sobre os indivíduos atingidos da coorte.

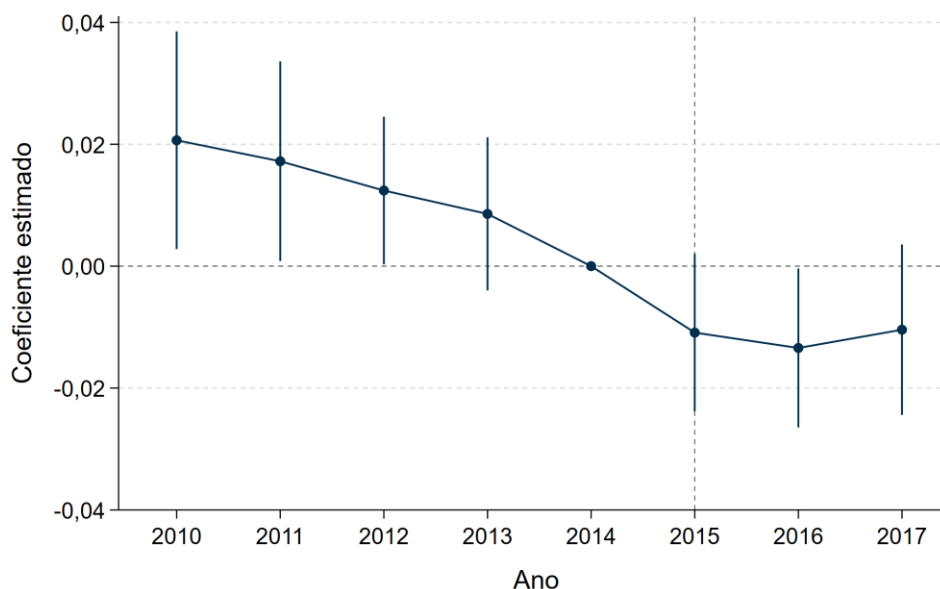
O Gráfico 10 mostra os resultados de análise similar à anterior, porém comparando as pessoas do grupo atingido às do GC 2. Quatro das sete variáveis de interesse do modelo⁷¹ mostram-se estatisticamente diferentes de 0 ao nível de significância de 10%, três das quais referem-se a anos que precedem o rompimento da Barragem de Fundão. Os coeficientes de 2010 a 2012 mostram que a diferença entre grupo atingido e o GC 2 apresentava tendência de queda, o que implica que as trajetórias do indicador de impacto dos dois grupos não eram paralelas antes do desastre. Isso sugere que o

⁷⁰ Como na seção 3.2, o coeficiente de 2014 é igualado a 0 por este ser o ano de referência da análise. Na estimação do modelo, β_{2014} é omitido e por consequência não possui intervalo de confiança associado.

⁷¹ O parâmetro referente a 2014 é omitido. Veja a nota de rodapé 70.

comportamento do indicador do GC 2 no período após o rompimento pode não ser, em média, uma representação satisfatória da situação contrafactual do grupo atingido.

Gráfico 10 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

Como consequência, o coeficiente referente a 2016, de magnitude estimada -0,013 e intervalo de confiança que não contém o 0, não pode razoavelmente ser interpretado como um efeito causal do rompimento. As análises de robustez na Tabela 27 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 corroboram esse entendimento. Embora a comparação dos atingidos com o GC 2 de cinco vizinhos sem reposição exiba impactos estatisticamente significantes em 2015 e 2016 sem diferenças entre os grupos nos anos anteriores ao desastre, o de três vizinhos com reposição mostra o contrário: as tendências prévias são distintas e nenhuma variável que mede o impacto do rompimento é significativa ao nível de 10%. Em outras palavras, a evidência produzida pelo modelo de diferença-em-diferenças quando o GC 2 é usado como grupo de comparação não se mostra adequada para avaliar a existência de impactos do rompimento sobre a probabilidade média de indivíduos do grupo atingido apresentarem vínculo formal ativo após 2014.

Quanto à investigação de efeitos heterogêneos em subgrupos, é digno de destaque o fato de que, diferentemente do discutido acima para o grupo atingido como um todo, a comparação tanto com o GC 1 quanto com o GC 2 de três vizinhos sem reposição mostra que a probabilidade de pessoas do grupo atingido terem um vínculo ativo se reduziu após o rompimento, em média, para aquelas que em 2014 estavam empregadas no setor de mineração. Como ilustra a Tabela 29 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4, a probabilidade média para os atingidos foi em 2017 11,7 p. p. menor que a do GC 1 com relação ao que a diferença era em 2014; comparados ao GC 2, a redução foi de 13,1 p. p.. A referida variável é estatisticamente significativa não somente ao nível de 10%, mas mesmo ao de 1% nos dois casos (quando se utiliza o GC 1 ou o GC 2). É importante também notar que os coeficientes estimados para 2016 são também de magnitudes comparáveis ao usar o GC 1 ou o GC 2 na análise, sendo que a respectiva variável é estatisticamente diferente de 0 ao nível de 5% nos dois casos. Ao mesmo tempo, nenhum dos parâmetros de 2010 a 2013 sugere que as trajetórias do indicador de impacto dos atingidos e dos GCs se comportaram de maneira divergente no período pré-rompimento. Ademais, esses resultados se sustentam mesmo nas análises de robustez com os GCs por meio de uso de pareamentos distintos⁷², sugerindo que as pessoas da coorte que em 2014 trabalhavam no setor de atividade econômica de mineração em municípios atingidos foram, em média, prejudicadas pelo rompimento da Barragem de Fundão no que se refere à permanência no mercado formal de trabalho nos anos subsequentes analisados.⁷³

A Tabela 28 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 mostra que 61,1% das pessoas do grupo atingido que estavam no setor de mineração em 2014 tiveram vínculo de trabalho ativo em 2017. Levando em consideração o Gráfico 8, isso significa que, de 3.872 pessoas (0,7816% dos 495.392 indivíduos atingidos na coorte, mostrado no gráfico como 0,8%), 1.506 não tiveram vínculo formal de trabalho com remuneração não nula registrado em 31 de dezembro de 2017. O percentual

⁷² Estes resultados não são apresentados no relatório, porém estão disponíveis mediante solicitação aos autores.

⁷³ Segundo a Tabela 29 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4, a comparação do grupo atingido com o GC 1 no ano de 2015 revela um coeficiente positivo, sendo o impacto estatisticamente diferente de 0 ao nível de significância de 10%. A Tabela 28 no mesmo apêndice mostra que esse resultado se deve à acentuada redução em 2015 na parcela de trabalhadores com vínculos ativos dentre aqueles que em 2014 estavam empregados no setor de mineração em municípios que compõem o GC 1. Tal queda não se observa no GC 2, cuja redução em 2015 em relação a 2014 é semelhante à observada no grupo atingido. Sendo assim, o impacto positivo e estatisticamente significativo de 2015 parece ser antes consequência de algum fator relacionado aos municípios do GC 1 do que do rompimento da Barragem de Fundão.

correspondente foi de 72,8% no GC 1 e de 74,2% no GC 2, refletindo as quedas relativas do grupo atingido mencionadas anteriormente, estimadas em 11,7 p. p e 13,1 p. p. em comparação a 2014 quando se utilizam os grupos de comparação 1 e 2, respectivamente. É importante destacar que as pessoas com vínculos ativos nos demais anos pós-rompimento não necessariamente os têm no mesmo estabelecimento, setor de atividade econômica e/ou no mesmo município de 2014. Ou seja, os resultados para esse indicador de impacto não captam casos de pessoas que tiveram o vínculo de trabalho de 2014 terminado após o rompimento, mas conseguiram se manter no mercado formal em outro emprego nos anos subsequentes.⁷⁴

4.2.2 Salário condicional a vínculo ativo no ano

A Tabela 10 exibe média e desvio-padrão do salário em dezembro das pessoas do grupo atingido, do GC 1 e do GC 2 que, em cada ano, foram registradas com um vínculo de trabalho ativo. O número de indivíduos cuja informação é usada para o cálculo das estatísticas é dado na Tabela 8, de onde segue que no ano de 2014, o qual marca a definição da coorte, a média e seu desvio-padrão consideram todos os indivíduos analisados.⁷⁵

⁷⁴ Os resultados nas seções 4.2.3 e 4.2.4 permitem considerações sobre permanência de emprego junto ao mesmo estabelecimento ou no mesmo setor de atividade.

⁷⁵ O fato de que a avaliação de impacto sobre salários não considera todos os indivíduos da coorte em todos os anos considerados pode ser interpretado como uma limitação da análise desenvolvida em termos de representatividade dos resultados. Uma forma alternativa de proceder, que contorna tal restrição, é por meio da metodologia de Groen et al. (2020), uma vez que com ela os autores conseguem analisar os impactos dos furacões Katrina e Rita sobre salários sem descartar nenhum indivíduo membro da coorte de seu estudo. Isto é feito através da imputação de remuneração nula às observações de pessoas que não têm emprego registrado, procedimento que torna possível decompor o efeito do furacão sobre salários em dois canais: (i) *within employment*, referente a pessoas que são observadas com empregos ao longo do tempo; e (ii) *shifts between employment and nonemployment*, referente às pessoas que ao longo do tempo deixam de aparecer como empregadas no banco de dados dos autores. Apesar de não constar neste relatório, a abordagem metodológica de Groen et al. (2020) foi aplicada ao contexto de avaliação de impactos do rompimento da Barragem de Fundão e os resultados da decomposição proposta pelos autores se mostraram qualitativamente similares aos apresentados na seção 4.2.1 e aos que se apresenta nesta seção 4.2.2. A saber, o componente *within employment* em grande parte replica a magnitude e significância estatística dos coeficientes no Gráfico 11 e no Gráfico 12, enquanto o componente *shifts between employment and nonemployment* reflete em grande parte o que se observa no Gráfico 9 e no Gráfico 10. Tais resultados vão em linha com o esperado quando se considera a maneira como os dois componentes do efeito são definidos por Groen et al. (2020). As estimativas obtidas para o impacto do rompimento seguindo essa metodologia estão disponíveis mediante solicitação aos autores deste relatório.

Tabela 10 — Média e desvio-padrão do salário em dezembro do ano de pessoas com vínculo ativo por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	2.160,14	2.339,66	2.108,60
	Desvio-padrão	2.244,28	2.821,57	2.143,14
2011	Média	2.208,79	2.441,01	2.174,83
	Desvio-padrão	2.261,16	2.890,26	2.173,02
2012	Média	2.330,45	2.560,72	2.279,02
	Desvio-padrão	2.331,54	2.977,65	2.234,49
2013	Média	2.322,72	2.638,85	2.336,19
	Desvio-padrão	2.253,37	3.011,69	2.244,39
2014	Média	2.227,21	2.460,98	2.210,34
	Desvio-padrão	2.144,42	2.827,77	2.096,97
2015	Média	2.328,32	2.618,56	2.358,71
	Desvio-padrão	2.217,65	2.955,84	2.228,62
2016	Média	2.400,90	2.674,58	2.453,98
	Desvio-padrão	2.273,73	2.961,43	2.282,95
2017	Média	2.469,54	2.729,92	2.538,04
	Desvio-padrão	2.306,37	2.960,75	2.329,62

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

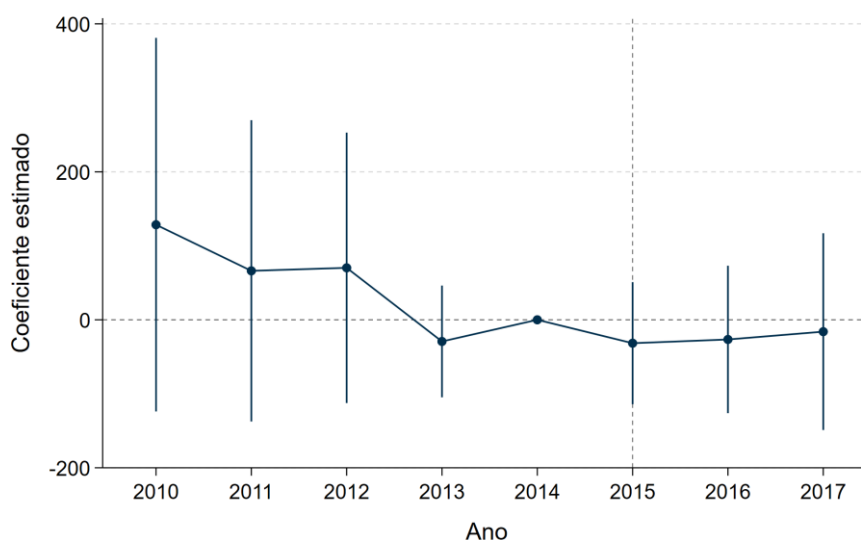
Observação: São consideradas apenas pessoas que possuíam vínculo de trabalho com remuneração não nula vigente em 31 de dezembro do ano considerado. Valores são apresentados em R\$ de dezembro de 2017.

Em média, a remuneração das pessoas com vínculo ativo do GC 1 é sempre superior à dos atingidos por uma diferença que varia de R\$ 179,52 (em 2010) a R\$ 316,13 (em 2013). Apesar disso, tal discrepância fica bem aquém do desvio-padrão da média dos dois grupos em todos os anos analisados, indicando que em ambos a distribuição de salários apresenta um número relevante de pessoas com remunerações consideravelmente acima e abaixo da média.⁷⁶ O mesmo se aplica ao GC 2, embora a remuneração média nele seja mais próxima à observada para o grupo atingido. De 2010 a 2012, nota-se que o salário médio, então inferior ao dos atingidos, aumentou a ponto de superá-lo em 2013, momento a partir do qual os dois grupos se mantiveram equiparáveis até 2015, anos em que a diferença entre eles não passou de R\$ 30. No ano de 2014, os três grupos demonstraram queda do salário médio, aparentemente revertida em 2015 e ponto de retomada do movimento de alta para 2016 e 2017. No entanto, enquanto a média dos atingidos nesses anos cresceu às taxas de 3,1% e 2,9% frente aos anos imediatamente anteriores, respectivamente, as do GC 1 foram mais modestas (2,1% nos dois anos) e as do GC 2, mais aceleradas (4% e 3,4%).

⁷⁶ A influência de indivíduos com remunerações muito destoantes das demais (*outliers*) sobre as estatísticas é limitada devido ao procedimento de *winsorizing* aplicado à variável de salário, conforme mencionado na seção 4.1.2.

Somente a Tabela 10 não permite, no entanto, analisar se o diferencial no crescimento do salário médio observado entre os três grupos de pessoas após o rompimento pode estar de alguma forma vinculado a ele. Para esse fim, o Gráfico 11 apresenta os coeficientes do diferença-em-diferenças estimados através da regressão da Equação 2, utilizando como grupo de comparação o GC 1 e tomando a remuneração das pessoas com vínculos ativos como indicador de impacto. O eixo vertical do gráfico informa a magnitude da estimativa em R\$ de dezembro de 2017.

Gráfico 11 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo (comparando com GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

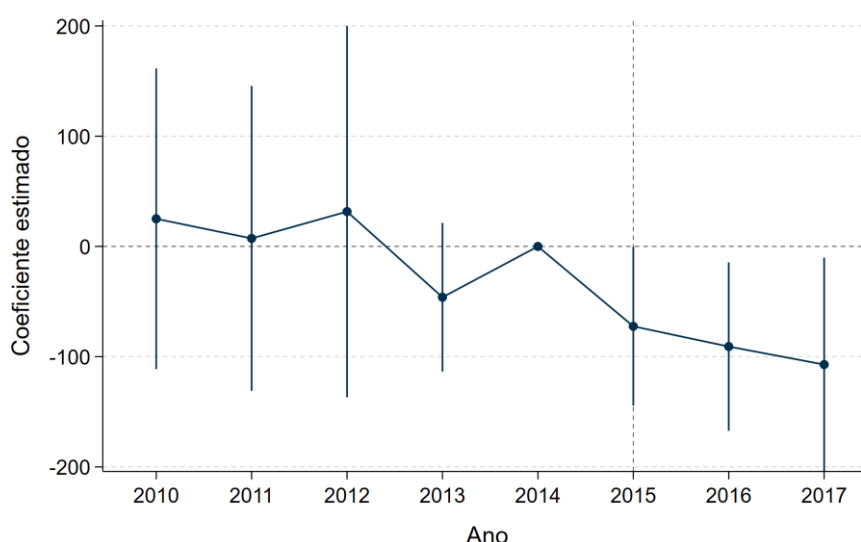
Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

Nota-se de imediato que todos os intervalos de confiança contêm o valor 0, indicando que nenhuma das variáveis pode ser considerada estatisticamente diferente de 0 ao nível de significância de 10%. Este resultado é similar ao que se observa quando o modelo é estimado usando o GC 1 em suas versões de três vizinhos com reposição e cinco vizinhos sem reposição, conforme apresentado na Tabela 30 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4. Sendo assim, o modelo não traz evidências claras de mudança na remuneração média de pessoas do grupo atingido que conseguiram se manter no mercado de trabalho formal nos anos após o rompimento em comparação ao GC 1.

O Gráfico 12, por sua vez, compara os atingidos da coorte aos indivíduos do GC 2. Diferentemente do observado no resultado anterior, as variáveis que medem o impacto do rompimento nos três anos que o sucedem (2015 a 2017) se mostram estatisticamente significantes ao nível de 10%, embora não ao nível mais rigoroso de 1% ou 5%. Apesar disso, elas revelam uma queda persistente na remuneração dos atingidos frente à do GC 2: comparado ao que essa diferença era em 2014, os atingidos perdem em média R\$ 72 em 2015, R\$ 91 em 2016 e R\$ 107 em 2017 – isto é, uma discrepância que se elevou em aproximadamente R\$ 20 por ano. Nota-se ainda que as variáveis referentes aos anos anteriores ao desastre se mostram estatisticamente iguais a 0, o que não contesta a hipótese de que a evolução do indicador de impacto do GC 2 seria uma boa aproximação à situação contrafactual dos atingidos.

Ademais, a Tabela 31 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 com o GC 2 em suas versões alternativas mostra que o resultado apresentado no Gráfico 12 é robusto a mudanças na forma de construção do grupo de comparação via pareamento. Nela, vê-se que as estimativas das variáveis que medem os impactos têm magnitude próxima às observadas no Gráfico 12, sendo elas também estatisticamente significantes ao nível de 10% quando se utiliza o GC 2 em suas versões de pareamento usando cinco vizinhos sem reposição ou de três vizinhos com reposição.

Gráfico 12 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

A evidência produzida pelo modelo de diferença-em-diferenças utilizando o grupo de comparação construído com pessoas de fora de MG e ES (GC 2) sugere então que o rompimento da Barragem de Fundão levou a uma queda na remuneração média das pessoas do grupo atingido que tiveram vínculos de trabalho formal ativos nos anos posteriores a 2014. Em particular, o modelo aponta que o salário médio dos atingidos da coorte empregados no mercado formal em 2017 seria de R\$ 2.576 ao invés de R\$ 2.469,54 se o rompimento não tivesse ocorrido.⁷⁷ Além disso, o fato de a comparação com o GC 1 não revelar efeito semelhante sugere que o impacto do rompimento pode ter repercutido sobre os demais municípios de MG e ES, de modo que neles também a remuneração no mercado formal tenha sido prejudicada. Tal resultado estaria em linha com a motivação por trás da criação do GC 2, conforme discutido na seção 3.1.4.

As estimativas do impacto médio do rompimento sobre os indivíduos atingidos com vínculos ativos são ainda heterogêneas em alguns aspectos relevantes, conforme mostram as tabelas no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 citadas a seguir. A discussão sobre efeitos heterogêneos é limitada nesta seção à comparação de subgrupos de indivíduos atingidos da coorte a equivalentes apenas do GC 2, uma vez que, em geral, os resultados para esses subgrupos quando se utiliza o GC 1 na análise replicam aquele observado no Gráfico 11.⁷⁸

Segundo a Tabela 32, o impacto do rompimento sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo no mercado formal de trabalho parece ter-se dissipado mais rapidamente para os homens. Isso porque, embora as estimativas de impacto sobre homens sejam similares às do resultado principal (Gráfico 12) em termos de sinal (sempre negativo) e magnitude dos coeficientes estimados⁷⁹, os efeitos não são estatisticamente diferentes de 0 ao nível de 10% em 2016 nem 2017. No caso de mulheres, também os coeficientes estimados são semelhantes em sinal (sempre negativo) e magnitude aos do resultado principal no Gráfico 12⁸⁰, porém os efeitos são estatisticamente significantes ao nível de 10% nos três anos após o rompimento, sendo que os referentes aos anos de 2016 e 2017 também o são ao nível de 5%.

⁷⁷ O valor de R\$ 2.576 resulta da soma do valor absoluto da estimativa de 2017 (R\$ 107) com o salário médio do grupo atingido observado naquele ano (R\$ 2.469,54, segundo a Tabela 10), ignorados os centavos.

⁷⁸ Embora tais resultados quando se utiliza o GC 1 não tenham sido incluídos neste relatório por motivo de concisão, estão disponíveis mediante solicitação aos autores.

⁷⁹ Exceto para 2015, quando o coeficiente é de aproximadamente -R\$ 93 para os homens contra -R\$ 72 quando não se fazem recortes por sexo.

⁸⁰ Exceto para 2015, quando o coeficiente é de aproximadamente -R\$ 41 para as mulheres contra -R\$ 72 quando não se fazem recortes por sexo.

Enfocando subgrupos por raça/cor, a Tabela 33 evidencia que pessoas sob apenas duas das cinco classificações tiveram estimativas de impacto estatisticamente diferentes de 0 ao nível de 10% de significância. Os impactos estimados sobre indivíduos que em 2014 foram registrados como pardos pelos empregadores têm magnitude similar aos do resultado principal (Gráfico 12), enquanto aqueles classificados como de raça/cor amarela mostram coeficientes sempre maiores em valor absoluto, chegando à redução média de R\$ 241 em 2017, mais que o dobro da estimativa correspondente para todos os atingidos (-R\$ 107). Merece destaque o fato de que os coeficientes estimados para os anos de 2015 a 2017 dos demais subgrupos (indígena, branca e preta) têm magnitude sempre inferior em valor absoluto aos do resultado principal, sendo os impactos ainda estatisticamente nulos ao nível de 10% de significância. Considerando que, segundo o Gráfico 6, as pessoas registradas como pardas representam 42,3% do grupo atingido, o resultado principal apresentado no Gráfico 12 parece então estar refletindo em grande parte o impacto estimado sobre indivíduos atingidos da coorte que são registrados como pardos e que tiveram vínculo de trabalho ativo nos anos seguintes ao rompimento.

No que se refere ao nível de escolaridade, nota-se na Tabela 34 que a magnitude do impacto sobre pessoas atingidas da coorte que em 2014 foram registradas na RAIS como tendo ensino superior completo é mais que o dobro da do resultado principal (Gráfico 12), sendo esses impactos estatisticamente significantes ao nível de 5% nos três anos pós-rompimento analisado. Em 2017, por exemplo, o coeficiente estimado para o impacto do rompimento sobre esse grupo foi de -R\$ 258, frente à estimativa de -R\$ 107 quando não se fazem recortes por nível de escolaridade. Por outro lado, para pessoas atingidas da coorte que em 2014 foram registradas na RAIS como tendo ensino fundamental incompleto e completo os coeficientes estimados são de magnitude consideravelmente mais baixa e, em sua maioria, esses impactos são estatisticamente nulos ao nível de 10% de significância. Tal padrão sugere que o impacto do rompimento da Barragem de Fundão sobre o mercado formal foi maior em postos de trabalho mais qualificado, os quais têm, em geral, maiores remunerações.

Por fim, a Tabela 35 mostra que a magnitude do impacto parece ter sido mais intensa sobre as pessoas do grupo atingido que em 2014 trabalhavam em setores censitários entre 1 e 10 km do Rio Gualaxo do Norte, Rio do Carmo e Rio Doce. Tomando as estimativas de 2017 como ilustração, a tabela mostra que elas foram -R\$ 102 para pessoas que estavam a até 1 km dos rios, -R\$ 121 para aquelas a até 5 km e -R\$ 131 para aquelas a até 10 km. Uma vez que a mudança de setores a até 1 km para até 5 km fez com que o impacto estimado aumentasse R\$ 19 em valor absoluto, as pessoas

que em 2014 trabalhavam em setores entre 1 e 5 km dos rios devem ter sofrido, em média, perdas maiores na remuneração do que aquelas de setores a até 1 km. A mudança de setores a até 5 km para até 10 km também elevou a estimativa de impacto, implicando de maneira similar que o efeito do sobre as pessoas em setores na faixa de 5 a até 10 km foi mais intenso do que sobre aquelas a até 5 km dos rios. Considerando que a estimativa de 2017 sob o resultado principal foi de -R\$ 107 (Gráfico 12), a magnitude do impacto estimado sobre os indivíduos que em 2014 trabalhavam em estabelecimentos localizados em setores a até 5 km e a até 10 km foi, em média, superior ao de todo o grupo atingido. É importante notar que, segundo a Figura 6, a faixa de setores entre 1 e 10 km dos rios continha 236.758 pessoas do grupo atingido em 2014, representando aproximadamente 49% do total de 479.771 indivíduos que possuem informação de setor censitário.⁸¹ Avançando até 2017, mesmo somente entre os indivíduos atingidos com vínculo ativo nesse ano, a proporção de pessoas cujo emprego formal em 2014 era em setores na faixa de 1 a 10 km do rio ainda era cerca de 48% do total.⁸² Isso mostra que quase metade dos indivíduos atingidos teve impacto negativo sobre salários em média mais intenso que o do grupo como um todo.

4.2.3 Vínculo ativo no ano com o mesmo estabelecimento de 2014

A Tabela 11 ilustra o comportamento do indicador de impacto “vínculo ativo no ano com o mesmo estabelecimento de 2014” ao longo dos anos de 2010 até 2017 para o grupo atingido e os GC 1 e 2. É possível verificar um padrão semelhante ao da Tabela 9, já que aqui a proporção de pessoas de cada grupo que possuía vínculo de trabalho ativo com o mesmo estabelecimento de 2014 cai à medida que o ano considerado se afasta de 2014. Ao observar o ano de 2017, cerca de 40% dos trabalhadores que possuíam vínculos ativos em 2014 continuaram exercendo suas atividades no mesmo estabelecimento, independentemente do grupo analisado. Cabe destacar que, uma vez que o indicador em questão é mais restritivo que o de “vínculo ativo no ano” na seção 4.2.1, as médias apresentadas a seguir são sempre menores ou iguais às aquelas apresentadas na Tabela 9.

⁸¹ Há 15.621 indivíduos atingidos na coorte aos quais não foi possível atribuir um setor censitário, conforme apontado na nota de rodapé 65.

⁸² O número exato de pessoas do grupo atingido com vínculo ativo em 2017 cujo emprego em 2014 foi em setores censitários que distavam entre 1 e 10 km dos rios é 149.305. Por sua vez, o número total de pessoas atingidas com vínculo ativo em 2017 e informação de setor censitário é 308.639.

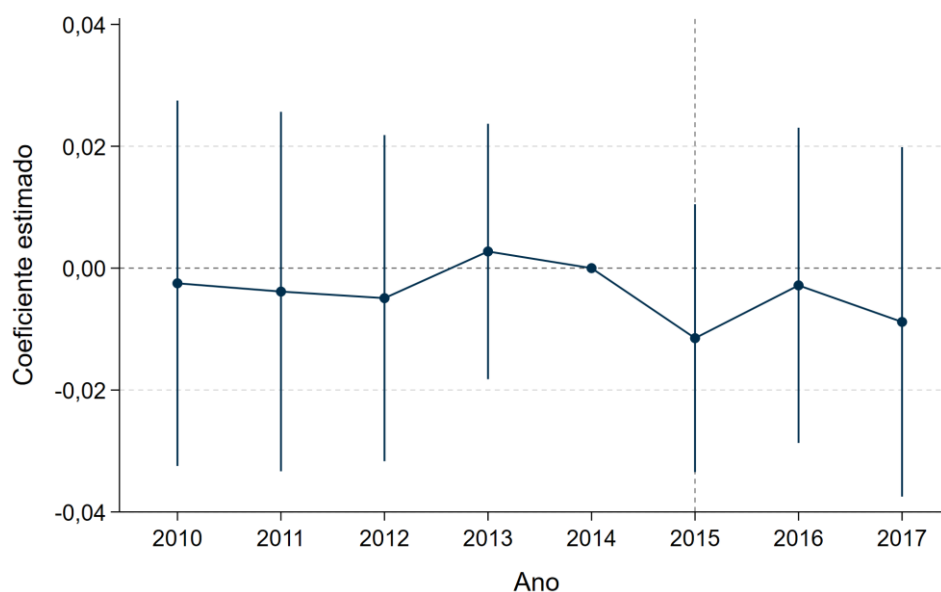
Tabela 11 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano com o mesmo estabelecimento de 2014 por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	0,295	0,297	0,280
	Desvio-padrão	0,456	0,457	0,449
2011	Média	0,363	0,366	0,350
	Desvio-padrão	0,481	0,482	0,477
2012	Média	0,455	0,460	0,446
	Desvio-padrão	0,498	0,498	0,497
2013	Média	0,622	0,619	0,605
	Desvio-padrão	0,485	0,486	0,489
2014	Média	1	1	1
	Desvio-padrão	0	0	0
2015	Média	0,639	0,650	0,645
	Desvio-padrão	0,480	0,477	0,478
2016	Média	0,493	0,496	0,494
	Desvio-padrão	0,500	0,500	0,500
2017	Média	0,406	0,415	0,409
	Desvio-padrão	0,491	0,493	0,492

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

O Gráfico 13 representa as estimativas de impacto médio sobre a probabilidade de um indivíduo no grupo atingido possuir vínculo de trabalho ativo com o mesmo estabelecimento de 2014, utilizando o GC 1 como comparativo. É possível verificar que todos os coeficientes estimados a partir da Equação 2, excetuando-se β_{2013} , são negativos, variando entre -0,003 e 0,003 nos anos anteriores ao rompimento e entre -0,012 e -0,003 nos posteriores. No entanto, os intervalos de confiança mostram que nenhuma destas variáveis pode ser considerada estatisticamente diferente de 0 ao nível de significância de 10%, uma vez que todos eles contêm o valor 0. Assim, não se pode rejeitar a hipótese de que o impacto do rompimento sobre o grupo atingido na dimensão captada pelo indicador analisado tenha sido, em média, nulo quando comparado ao GC 1 e considerando o referido nível de significância estatística.

Gráfico 13 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo com o mesmo estabelecimento de 2014 (comparando com GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

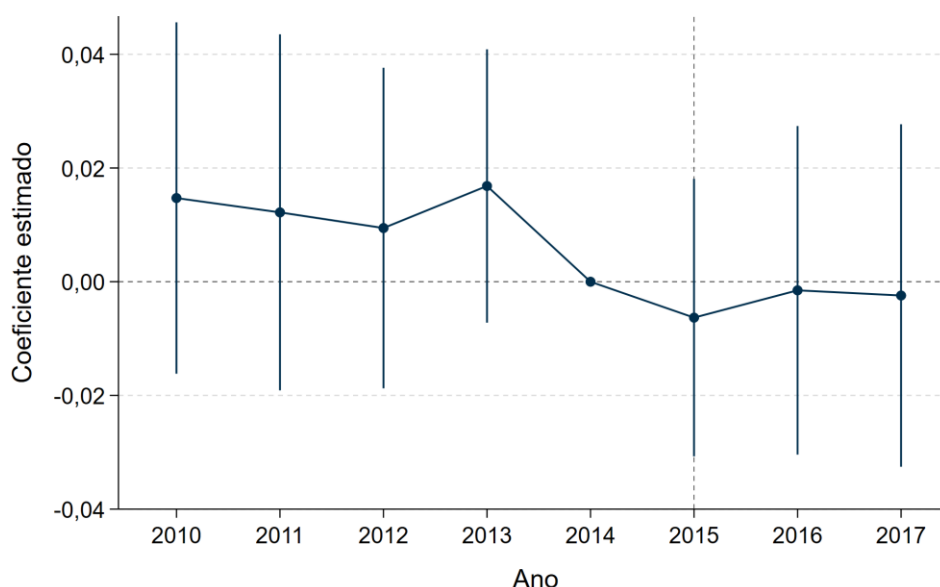
Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

Com o intuito de avaliar a robustez dos resultados apresentados no Gráfico 13, a Tabela 36 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 mostra que há sensibilidade dos coeficientes à forma de construção do GC 1. Em particular, as estimativas para os anos de 2015 até 2017, apesar de todas negativas, variam em magnitude e também em significância estatística das variáveis correspondentes entre as diferentes versões do GC 1. Especificamente para a versão de pareamento com reposição aos três vizinhos mais próximos, as variáveis do diferença-em-diferenças após o rompimento são estatisticamente significantes a 10%. Assim, é possível que os resultados apresentados no Gráfico 13 não sejam efetivamente conclusivos com relação à inexistência de efeito médio do rompimento sobre o grupo atingido quando se considera o GC 1 como representação de sua situação contrafactual.

O Gráfico 14 ilustra os resultados da estimação quando se compara o grupo atingido ao GC 2. Novamente, nenhuma das variáveis do diferença-em-diferenças ilustradas se mostra estatisticamente significativa ao nível de 10%, de forma que não se pode rejeitar a hipótese de que o impacto do rompimento sobre este indicador tenha sido, em média, nulo. Por sua vez, os resultados das estimativas de modelos com versões alternativas do GC 2 apresentados na Tabela 37 do APÊNDICE D — Tabelas e figuras adicionais

do capítulo 4 deixam aparente que, especialmente em 2016 e 2017, os coeficientes estimados são pequenos em valores absolutos e oscilam de valores negativos a positivos a depender da versão do pareamento utilizada para construir o GC 2. Apesar dessa sensibilidade, nenhuma das variáveis que medem o impacto pós-rompimento nesses modelos é estatisticamente significativa ao nível de significância de 10%. Converte-se então para a conclusão de que as evidências produzidas pelo método de diferença-em-diferenças não apontam para mudanças, em média, na probabilidade de indivíduos atingidos terem vínculo ativo com o mesmo estabelecimento de 2014 frente à das pessoas no GC 2 após o rompimento.

Gráfico 14 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo com o mesmo estabelecimento de 2014 (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

A investigação de impactos do rompimento em subgrupos de pessoas atingidas da coorte sobre o indicador de vínculo ativo no ano com o mesmo estabelecimento de 2014 também não encontrou resultados robustos apontando a existência de tais efeitos heterogêneos. Apesar de não incluídos neste relatório por motivos de concisão, estas análises serão disponibilizadas mediante solicitação aos autores.

4.2.4 Vínculo ativo no ano no mesmo setor de atividade econômica de 2014

A Tabela 12 exibe as estatísticas de média e desvio-padrão do indicador de impacto “vínculo ativo no ano no mesmo setor de atividade econômica de 2014” ao longo dos anos de 2010 a 2017 para o grupo atingido e os grupos de comparação 1 e 2 (GC 1 e GC 2, respectivamente). Como reflexo da definição menos restritiva do indicador quando comparada à da seção anterior, a parcela de pessoas que possuem vínculos ativos no mesmo setor de 2014 a cada ano é maior que a de pessoas que possuem vínculo com o mesmo estabelecimento de 2014. Apesar dessa diferença, as médias decrescem quanto mais distante o ano considerado estiver de 2014, como nas seções 4.2.1 e 4.2.3. Nota-se que os valores apresentados para as médias são sempre bastante similares entre os três grupos a cada ano. Em 2017, por exemplo, 46,3% dos atingidos na coorte possuíam vínculo de trabalho ativo no mesmo setor de atividade econômica de 2014, enquanto no GC 1 e GC 2 esta parcela foi 47,6% e 47,3%, respectivamente. Outro ponto de destaque é a queda da média ao longo dos anos pós-rompimento: entre 2014 e 2015 houve uma queda de 33,6 pontos percentuais (p. p.) na parcela de indivíduos do grupo atingido com vínculo ativo no mesmo setor que 2014 (de 100% em 2014, por construção, para 66,4% em 2015); e de 2016 a 2017, por sua vez, o percentual caiu de 53,6% para 46,3% (variação de 7,3 p. p.) para este mesmo grupo.

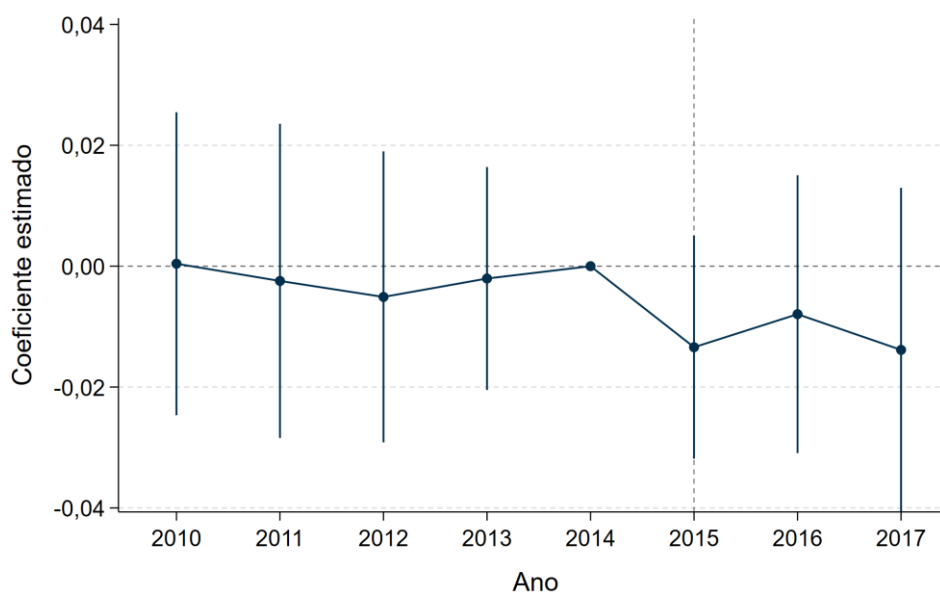
Tabela 12 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano no mesmo setor de atividade econômica de 2014 por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	0,364	0,364	0,346
	Desvio-padrão	0,481	0,481	0,476
2011	Média	0,428	0,430	0,418
	Desvio-padrão	0,495	0,495	0,493
2012	Média	0,509	0,514	0,502
	Desvio-padrão	0,500	0,500	0,500
2013	Média	0,654	0,656	0,645
	Desvio-padrão	0,476	0,475	0,479
2014	Média	1	1	1
	Desvio-padrão	0	0	0
2015	Média	0,664	0,678	0,675
	Desvio-padrão	0,472	0,467	0,468
2016	Média	0,536	0,544	0,547
	Desvio-padrão	0,499	0,498	0,498
2017	Média	0,463	0,476	0,473
	Desvio-padrão	0,499	0,499	0,499

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

O Gráfico 15 ilustra os resultados da estimativa do modelo definido na Equação 2 para o indicador de impacto em questão quando se compara os indivíduos do grupo atingido aos do GC 1. Os coeficientes associados a todos os períodos analisados são negativos, excetuando-se o referente ao ano 2010, estimado marginalmente acima de 0 (0,0004). Entretanto, aqueles referentes aos anos posteriores ao rompimento da Barragem de Fundão se mostram em patamar inferior aos de 2011 a 2013, possível indicativo de uma mudança na tendência da diferença entre os grupos existente até então. Apesar disso, os intervalos de confiança mostram que, ao nível de significância de 10%, não se pode rejeitar a hipótese de que as variáveis do diferença-em-diferenças sejam estatisticamente nulas. Em outras palavras, não se pode afirmar a tal nível de significância que a diferença entre as médias do indicador do grupo atingido e do GC 1 se alterou em qualquer um dos anos analisados com relação à que se observava em 2014.

Gráfico 15 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 1)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

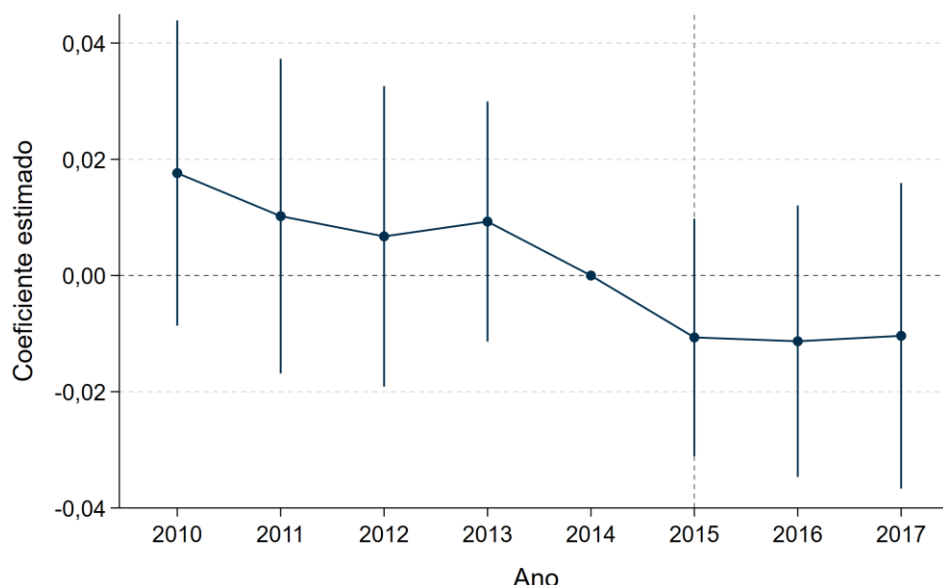
Ao analisar a robustez dos resultados ao uso de versões alternativas do pareamento empregado na construção do GC 1, verifica-se na Tabela 38 no APÊNDICE D —

Tabelas e figuras adicionais do capítulo 4 que tanto a magnitude dos coeficientes estimados como a significância estatística das variáveis correspondentes são sensíveis à estratégia de pareamento de municípios utilizada. Em particular, a comparação do grupo atingido com o GC 1 de três vizinhos mais próximos com reposição mostra efeitos estatisticamente diferentes de 0 ao nível de 10% e também de 5% de significância. Os coeficientes estimados são sempre negativos e de magnitude superior aos observados no Gráfico 15 nos três anos pós-rompimento analisados. Assim, frente à considerável variabilidade dos resultados relacionada à estratégia de construção do GC 1, as estimativas apresentadas no Gráfico 15 devem ser tratadas antes como inconclusivas do que como evidência da inexistência de efeito médio do rompimento da Barragem de Fundão sobre o grupo atingido no que se refere indicador de impacto analisado.

Com relação aos resultados obtidos a partir da comparação do grupo atingido com o GC 2, o Gráfico 16 mostra estimativas de impacto negativas de 2015 a 2017 sobre a probabilidade de indivíduos atingidos da coorte possuírem vínculo ativo no mesmo setor de atividade econômica em que estavam empregados em 2014. Como no Gráfico 15, também os intervalos de confiança associados a todos os pontos estimados atravessam a linha horizontal do valor 0, implicando que as variáveis do diferença-em-diferenças ilustradas não podem ser consideradas estatisticamente diferentes de 0 sob o nível de significância de 10%. Ou seja, os resultados não provêm base para afirmar que a relação entre as médias do indicador de impacto do grupo atingido e do GC 2 se alterou após o rompimento com relação a 2014.

Os modelos estimados na Tabela 39 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 apontam alguma flutuação na magnitude das estimativas referentes aos anos pós-rompimento quando se utilizam as versões alternativas de pareamento para a construção do GC 2, porém não em termos de sinal dos coeficientes estimados ou de significância estatística. De fato, nenhum dos modelos alternativos mostra as variáveis que medem os impactos do rompimento estatisticamente significantes ao nível de 10%. Tendo isso em vista, as análises de robustez não contrariam a evidência dada no Gráfico 16: a hipótese de que o indicador de impacto dos indivíduos do grupo atingido não teria se alterado desde 2014, em média, frente ao das pessoas do GC 2 não pode ser rejeitada ao nível de significância de 10%.

Gráfico 16 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

No entanto, a investigação de possível heterogeneidade nos impactos do rompimento sobre esse indicador revela indícios de que pessoas que em 2014 trabalhavam no setor de atividade econômica de agropecuária e pesca podem ter sido impactadas de maneira diferente da estimada para o grupo atingido como um todo. A Tabela 40 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 indica que, em média, a probabilidade de pessoas nesse subgrupo de possuir vínculo ativo em agropecuária e pesca se reduziu após o rompimento frente ao GC 2, sendo o maior impacto observado no ano de 2016. Em 2015, o coeficiente estimado é de -0,077 (redução de redução de 7,7 p. p.), sendo a variável correspondente estatisticamente significativa a 10%. Já em 2016, a queda frente ao GC 2 com relação a 2014 se agrava para 9,7 p. p., retornando ao patamar de redução estimada de 7,7 p. p. em 2017. Estas duas últimas variáveis, referentes a 2016 e 2017, são estatisticamente diferentes de 0 ao nível de significância de 10% e de 5%. Tais estimativas se mostram ainda robustas a mudanças na forma de construir o GC 2 – cinco vizinhos sem reposição ou três vizinhos com reposição no processo de pareamento.⁸³ Além disso, uma vez que os impactos do rompimento não

⁸³ Resultados dos modelos com versões alternativas do GC 2 não são apresentados no relatório, mas estão disponíveis mediante solicitação aos autores.

se mostraram estatisticamente significantes ao nível de 10% quando se faz esta mesma análise usando o GC 1, não se pode descartar a possibilidade de que este impacto tenha transcendido o grupo atingido e afetado também outros municípios de MG e ES; ou seja, de que tenha havido efeitos de transbordamento para municípios que integram o GC 1.

Como ilustração do resultado discutido acima em número de pessoas, a Tabela 41 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 informa que a parcela de indivíduos do grupo atingido que trabalhava em agropecuária e pesca em 2014 e teve vínculo ativo em 2017 nesse mesmo setor era de 42,117%, equivalente a 8.733 do total de 20.735 pessoas nesse subgrupo.⁸⁴ A estimativa de impacto -0,077 associada a 2017, no entanto, implica que nesse ano seriam 10.326 pessoas com vínculo ativo em agropecuária e pesca se o indicador de impacto do grupo atingido tivesse seguido a trajetória do GC 2, que representa o que teria acontecido na ausência do rompimento.⁸⁵ Supondo que esse grupo de comparação de fato representa uma boa aproximação à situação contrafactual das pessoas da coorte em municípios atingidos, pode-se considerar que o rompimento da Barragem de Fundão levou 1.593 pessoas que antes trabalhavam em agropecuária e pesca a não conseguir se manter nesse mesmo setor com um vínculo de trabalho formal ativo em 2017.⁸⁶

4.2.5 Salário condicional a vínculo ativo no ano no mesmo setor de atividade econômica de 2014

A Tabela 13 ilustra o comportamento das estatísticas de média e desvio-padrão do salário em dezembro das pessoas do grupo atingido, do GC 1 e do GC 2 que, em cada ano, foram registradas com um vínculo de trabalho ativo no mesmo setor de atividade econômica no qual atuavam em 2014. A informação de número de indivíduos que é usada para o cálculo das estatísticas a seguir é apresentada na Tabela 42 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4. Por construção, a média e seu desvio-padrão no ano de 2014, o qual marca a definição da coorte, consideram

⁸⁴ O total de 20.735 pessoas é obtido do Gráfico 8: 4,1856% de 495.392 indivíduos da coorte no grupo atingido (mostrado no gráfico como 4,2%). Por sua vez, o percentual de 42,117% é exibido na Tabela 41 como 0,421.

⁸⁵ O número de 10.326 pessoas resulta do cálculo: $20.735 \cdot [0,421 + 0,077] = 10.326,03 \approx 10.326$, onde 0,421 é a parcela de pessoas do subgrupo que teve em 2017 vínculo ativo no mesmo setor de 2014, conforme a Tabela 41, e 0,077 é o valor absoluto da estimativa de impacto para 2017.

⁸⁶ Este número resulta da subtração: $10.326 - 8.733 = 1.593$.

todos os indivíduos analisados: 495.392 no grupo atingido, 1.441.898 e 1.359.404 no GC 1 e GC 2, respectivamente.

Tabela 13 — Média e desvio-padrão do salário em dezembro do ano de pessoas com vínculo ativo no mesmo setor de 2014 por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	2.452,49	2.557,07	2.274,67
	Desvio-padrão	2.541,91	3.026,91	2.265,46
2011	Média	2.447,42	2.649,06	2.326,13
	Desvio-padrão	2.489,88	3.099,40	2.280,93
2012	Média	2.496,74	2.703,54	2.372,35
	Desvio-padrão	2.502,12	3.117,01	2.286,41
2013	Média	2.388,67	2.697,56	2.375,13
	Desvio-padrão	2.318,69	3.067,51	2.265,36
2014	Média	2.227,21	2.460,98	2.210,34
	Desvio-padrão	2.144,42	2.827,77	2.096,97
2015	Média	2.364,97	2.660,69	2.381,07
	Desvio-padrão	2.238,42	3.008,27	2.232,98
2016	Média	2.498,33	2.787,49	2.534,00
	Desvio-padrão	2.341,85	3.076,73	2.337,22
2017	Média	2.659,47	2.912,73	2.665,33
	Desvio-padrão	2.496,61	3.151,26	2.426,93

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: São consideradas apenas pessoas que possuíam vínculo de trabalho com remuneração não nula vigente em 31 de dezembro do ano considerado e no mesmo setor de atividade econômica em que trabalhavam em 2014. Valores são apresentados em R\$ de dezembro de 2017.

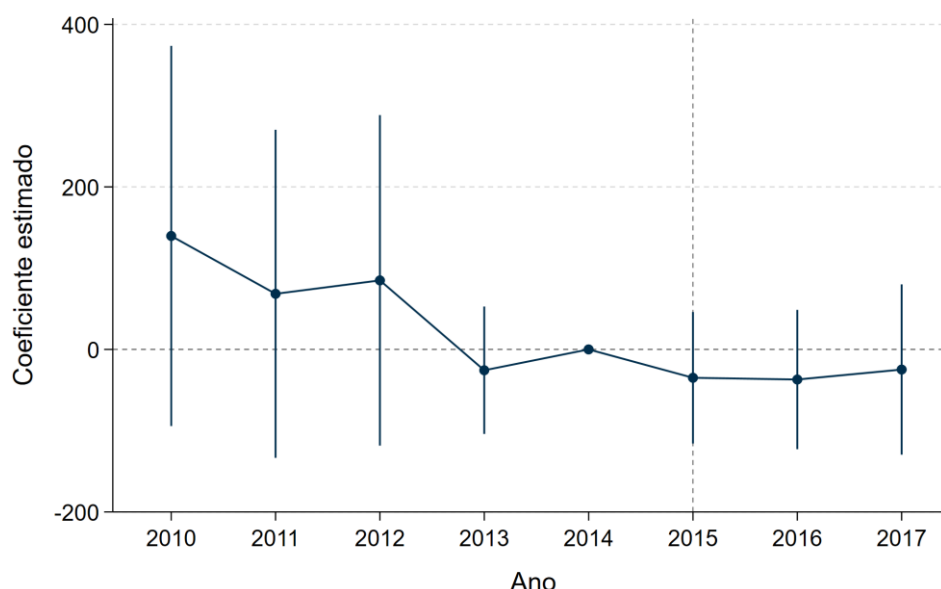
De maneira geral, ao se observar a Tabela 13, a média da remuneração dos indivíduos que têm um vínculo ativo no ano no mesmo setor de atividade econômica de 2014 no GC 1 é sempre superior à observada para o grupo atingido por valores que variam entre R\$ 100 e R\$ 300, aproximadamente. O desvio-padrão para cada um dos grupos é bastante elevado em todos os anos, sugerindo heterogeneidade considerável de remuneração entre os indivíduos que compõem cada um deles. O GC 2 possui remunerações médias mais similares àquelas apresentadas pelo grupo atingido, especialmente quando se consideram anos mais próximos ao do rompimento, porém apresenta também desvios-padrão bastante elevados. Merece destaque o fato de que, ao se comparar os valores apresentados na Tabela 10 com os da Tabela 13, a média dos salários daqueles indivíduos que no ano em análise atuavam no mesmo setor de atividade econômica que em 2014 (Tabela 13) é sempre maior do que as observadas quando não se faz restrição sobre o setor de atividade econômica de atuação (Tabela

10).⁸⁷ Uma remuneração média maior daqueles que não migram de setor de atividade econômica ao longo do tempo é esperada e pode ser explicada por eventual recompensa pela maior experiência nas atividades desempenhadas e/ou pela ausência do custo de aprendizado de especificidades de um novo setor.

Os coeficientes estimados a partir do método de diferença-em-diferenças (Equação 2) são apresentados no Gráfico 17, tendo a remuneração dos indivíduos com vínculos ativos no mesmo setor de atividade econômica que em 2014 como indicador de impacto e o GC 1 como grupo de comparação. Como nos gráficos da seção 4.2.2, o eixo vertical informa a magnitude da estimativa em R\$ de dezembro de 2017. Ao analisar os resultados, nota-se que todos os intervalos de confiança estimados contêm o valor 0, indicando que nenhuma das variáveis pode ser considerada estatisticamente diferente de 0 ao nível de significância de 10%, independentemente de se referirem a anos anteriores ou posteriores ao rompimento. Este resultado é similar ao que se observa quando se estima o modelo usando o GC 1 em suas versões de três vizinhos com reposição ou cinco vizinhos sem reposição, conforme apresentado na Tabela 43 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4, onde também os coeficientes estimados para 2015 a 2017 são negativos. Assim, os resultados do modelo de diferença-em-diferenças comparado o grupo atingido às diferentes versões do GC 1 parecem convergir ao não permitir rejeitar ao nível de significância de 10% a hipótese de que a remuneração média de pessoas do grupo atingido não se alterou frente à daquelas do GC 1 que conseguiram se manter no mercado de trabalho formal no mesmo setor de atividade econômica.

⁸⁷ A única exceção é o ano de 2014, quando a diferença é sempre igual a 0 por construção.

Gráfico 17 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 1)



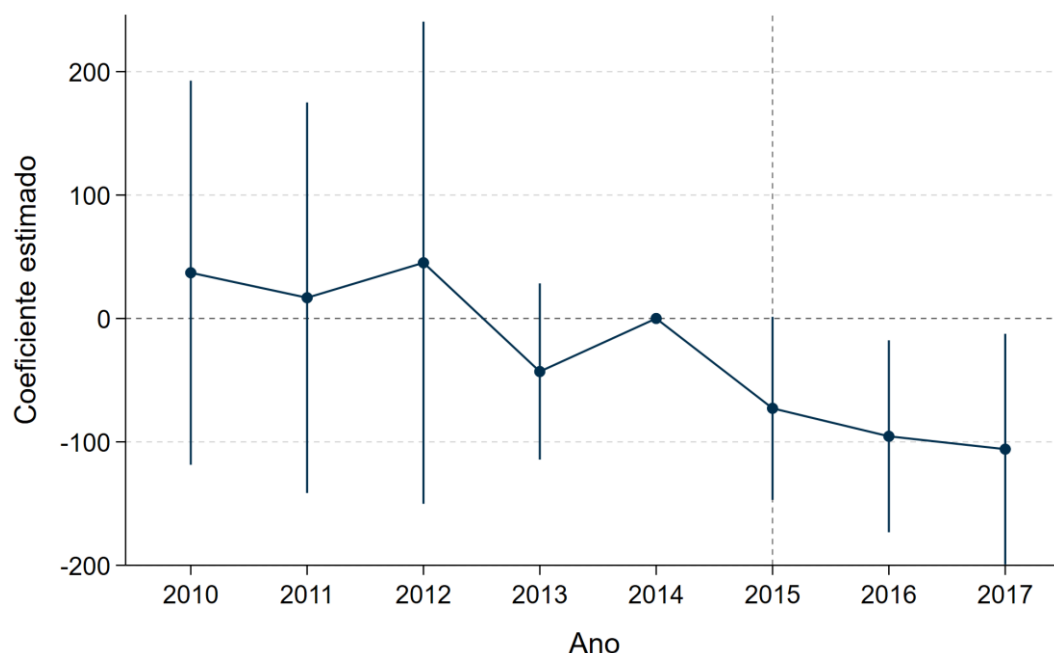
Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 1 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

O Gráfico 18 ilustra os resultados de análise similar, mas estimados a partir da comparação do grupo atingido com o GC 2. Os coeficientes referentes ao período pós-rompimento são todos negativos, sendo que os efeitos se mostram estatisticamente significantes ao nível de 10% em 2016 e 2017.⁸⁸ As estimativas sugerem uma queda na remuneração dos atingidos que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 frente àqueles do GC 2, a qual aumenta ao longo dos dois últimos anos analisados: comparado ao que essa diferença entre grupos era em 2014, os atingidos perderam, em média, R\$ 95,48 em 2016 e R\$ 105,91 em 2017. Note que as variáveis referentes aos anos pré-rompimento não são estatisticamente significantes ao nível de 10%, indicando que não existiram mudanças relevantes antes do desastre na diferença entre as médias do indicador de impacto do grupo atingido e do GC 2 com relação a 2014. Trata-se, portanto, de evidência favorável à credibilidade da hipótese de tendências paralelas requerida pelo modelo de diferença-em-diferenças.

⁸⁸ A variável referente ao impacto no ano de 2016 também é estatisticamente significativa ao nível de 5%, conforme ilustra a Tabela 44 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

Gráfico 18 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 2)



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2 na versão de pareamento sem reposição a três vizinhos mais próximos. Os intervalos de confiança foram construídos com nível de significância de 10% a partir de erros-padrão *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

As análises de robustez na Tabela 44 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4 indicam que o resultado apresentado no Gráfico 18 não se altera de maneira relevante em resposta a mudanças no tipo de pareamento utilizado para a construção do GC 2. Em linha com a seção 4.2.2, as estimativas produzidas pelo modelo de diferença-em-diferenças ganham assim credibilidade enquanto evidências de que o rompimento da Barragem de Fundão teve, em média, impacto negativo sobre a remuneração das pessoas do grupo atingido que apresentaram vínculo ativo no mesmo setor de atividade de 2014 frente às do GC 2.

Também remetem aos resultados da seção 4.2.2 os modelos estimados em subgrupos de indivíduos atingidos, voltados à investigação de efeitos heterogêneos do rompimento. A Tabela 45 mostra que o impacto do rompimento sobre homens que tiveram vínculo ativo no mesmo setor de atividade de 2014 é mais intenso que o estimado para todos os atingidos (Gráfico 18) e estatisticamente significativo ao nível de 10% em 2015 e 2016, porém não mais em 2017, apontando possível recuperação mais célere do que a observada para mulheres. Como no impacto sobre salário de mulheres

com vínculo ativo independentemente do setor de atividade (Tabela 32), o impacto para elas persiste até 2017, alcançando neste ano magnitude maior em valor absoluto que o estimado para todo o grupo atingido. Verifica-se que o impacto para as mulheres, nesse caso, é estatisticamente significativo aos níveis de 10% e mesmo 5% nos anos de 2016 e 2017, sendo o sinal do coeficiente estimado relacionado ao impacto sempre negativo nos anos pós-rompimento analisados (2015, 2016 e 2017).

A Tabela 46 retoma os subgrupos de raça/cor e mostra que, como na Tabela 33, os impactos são estatisticamente significantes ao nível de 10% somente para as pessoas registradas pelos empregadores em 2014 como amarelas e pardas. Especificamente para os pardos, o impacto estimado é bastante similar ao do resultado principal com exceção da variável que mede o impacto em 2015, cujo coeficiente se mostra R\$ 10 mais negativo que o correspondente no modelo principal e também estatisticamente diferente de 0 ao nível de significância de 10%. Com relação às pessoas classificadas como de raça/cor amarela, repete-se o comportamento observado para impacto sobre salário de pessoas com vínculo ativo independentemente do setor de atividade em 2014 (Tabela 33): as variáveis referentes aos anos posteriores ao rompimento se mostram estatisticamente significantes ao nível de 10% e com coeficientes estimados maiores em valor absoluto que os do resultado principal.⁸⁹ Todavia, a magnitude dos coeficientes estimados quando se considera o subgrupo de raça/cor amarela é menor que a observada na Tabela 33, indicando que o impacto do rompimento sobre esse subgrupo foi, em média, menos intenso sobre as pessoas que conseguiram se manter com vínculo ativo no mesmo setor de atividade econômica em que trabalhavam em 2014.

Com relação ao nível de escolaridade, o subgrupo com ensino superior completo novamente se destaca na Tabela 47 com todas as variáveis pós-rompimento estatisticamente significantes ao nível de 10% (e de 5%, no caso de 2015 e 2016) e com coeficientes estimados de magnitude superior ao dobro das do modelo principal. Em comparação com o resultado sobre salários desse subgrupo considerando vínculos ativos independentemente de serem ou não no mesmo setor de atividade econômica de 2014 (Tabela 34), os parâmetros estimados apresentam somente pequenas diferenças, indicando que o impacto não parece ter variado, em média, a depender do setor de atividade econômica que um indivíduo com ensino superior completo passou a ter vínculo ativo depois do rompimento. O mesmo se aplica aos subgrupos dos demais níveis de escolaridade, inclusive nos casos em que as variáveis que medem o impacto

⁸⁹ As variáveis que medem o impacto em 2016 e em 2017 são estatisticamente significantes mesmo considerando o rigoroso nível de significância de 1%.

não se mostram estatisticamente significantes, como o de ensino fundamental completo ou o de fundamental incompleto em alguns dos anos analisados.

Por fim, a Tabela 48 apresenta a avaliação de impactos do rompimento por subgrupo de pessoas da coorte atingidas determinado pela localização do setor censitário em que elas estavam empregadas em 2014. Diferentemente do observado para o indicador de impacto da seção 4.2.2, o efeito estimado mais intenso do rompimento para 2016 e 2017, em média, é observado para os indivíduos que em 2014 trabalhavam em setores censitários a até 5 km do Rio Gualaxo do Norte, Rio do Carmo e Rio Doce, e não para aqueles até 10 km como na Tabela 35. Não obstante, os três subgrupos destacados na Tabela 48 exibem estimativas de impacto cuja magnitude supera a do resultado principal (Gráfico 18) em todos os anos posteriores ao rompimento, sendo eles sempre negativos.⁹⁰ Portanto, a localização do setor censitário em que estavam empregadas em 2014 parece ser um fator associado à intensidade média dos impactos do rompimento sobre o salário das pessoas com vínculo ativo no mesmo setor de atividade econômica de 2014, sendo as perdas maiores para aquelas relativamente próximas aos rios.

⁹⁰ Em rigor, o coeficiente de 2015 estimado para pessoas em setores censitários a até 1 km dos rios é menor em valor absoluto que o do resultado principal. No entanto, o efeito sobre esse subgrupo se mostra estatisticamente significativo ao nível de 10%, enquanto o Gráfico 18 mostra que a variável de 2015 não pode ser considerada estatisticamente diferente de 0 ao mesmo nível de significância para o caso do resultado principal.

5 CONSIDERAÇÕES FINAIS

Este relatório investigou a existência de impactos do rompimento da Barragem de Fundão sobre o mercado de trabalho formal. As análises realizadas contemplaram o período de 2010 a 2017 e abordaram dois enfoques distintos: (i) impactos sobre o tamanho do mercado formal dos municípios atingidos; e (ii) impactos sobre a trajetória no mercado formal de pessoas que, antes do rompimento, possuíam vínculo formal de trabalho junto a estabelecimentos localizados nos municípios atingidos.

As duas análises consistiram em avaliações econométricas de impacto e fizeram uso de uma mesma metodologia: o modelo de diferença-em-diferenças modificado para permitir a estimação do impacto em cada ano de 2015 a 2017 com relação a 2014, o qual foi definido como ano pré-rompimento de referência. Os resultados foram obtidos ao se comparar ao longo do tempo o indicador de impacto do grupo atingido com o do grupo de comparação, sendo que para este segundo se usou duas definições distintas: (i) municípios não atingidos de Minas Gerais, Espírito Santo, sul da Bahia e norte do Rio de Janeiro pareados aos atingidos (GC 1); e (ii) municípios não atingidos da Bahia, Goiás, Mato Grosso do Sul, Rio de Janeiro e São Paulo pareados aos atingidos (GC 2). O procedimento de pareamento se baseou na métrica de distância de Mahalanobis e levou em consideração aspectos geográficos, características do mercado de trabalho formal e informações socioeconômicas gerais para tentar determinar quais dos municípios não atingidos pelo rompimento eram, antes do desastre, similares aos municípios que foram atingidos, o que é importante para que o grupo de comparação seja uma boa representação da situação contrafactual do grupo atingido (isto é, a situação deles caso o rompimento não tivesse ocorrido).

A análise que enfoca o tamanho como um todo do mercado de trabalho formal dos municípios atingidos investigou dois indicadores de impacto: (i) o logaritmo natural do número de vínculos ativos; e (ii) o logaritmo natural da massa salarial. Com relação ao primeiro deles, as estimativas do modelo para o impacto do rompimento se mostraram negativas nos anos de 2015 a 2017, indicando redução do número de vínculos ativos nos municípios atingidos frente aos do grupo de comparação com relação a 2014. Todavia, estes efeitos só são estatisticamente diferentes de 0 ao nível de significância de 10% quando se compara os municípios atingidos aos do GC 2: nesse caso, estima-se uma perda média no número de vínculos ativos com relação a 2014 de aproximadamente 300 no ano de 2015, 420 em 2016 e 455 vínculos por município atingido em 2017. Se os 45 municípios atingidos tivessem seguido a trajetória dos do GC 2 após 2014, portanto, teriam 20.475 vínculos ativos a mais em 2017 do que o que

efetivamente se observou.⁹¹ Com relação ao segundo indicador de impacto, o logaritmo natural da massa salarial, as estimativas de impacto foram também negativas, porém estatisticamente significantes ao nível de 10% somente em 2016 na comparação com o GC 2. Para esse ano, estimou-se que a massa salarial foi, em média, R\$ 1,3 milhões menor por município atingido do que seria a trajetória destes se tivesse seguido desde 2014 a da massa salarial média do GC 2.⁹² Sob a hipótese de que esse grupo de comparação é uma boa aproximação à situação contrafactual do grupo atingido, a qual não é contrariada pelas estimativas de diferença-em-diferenças para os anos anteriores a 2014, pode-se considerar que o rompimento da Barragem de Fundão teve, em média, impacto negativo sobre o mercado de trabalho formal dos municípios atingidos.

Um aspecto interessante dessa análise é que evidências de impacto do rompimento emergem somente quando da comparação do grupo atingido com o grupo de comparação composto por municípios de UFs vizinhas a MG e ES. Ao longo do relatório, levantou-se a possibilidade de que o GC 1, composto em sua maioria de municípios pertencentes a MG e ES, poderia estar particularmente sujeito a repercutir os impactos do rompimento, o que o tornaria uma aproximação ruim à situação contrafactual dos atingidos. Com efeito, o fato de que as comparações com o GC 1 produzem estimativas que não são estatisticamente significantes vai ao encontro dos resultados apresentados nos relatórios da FGV que realizam o diagnóstico do impacto do rompimento da Barragem de Fundão com enfoque macroeconômico. Em particular, a análise em FGV (2020a) sugere que o impacto do rompimento sobre o PIB transcende os municípios da sub-bacia do Rio Doce e afeta também outras regiões de MG e ES. Não surpreende, portanto, que os resultados apresentados aqui não demonstrem vestígios de efeito sobre o mercado de trabalho formal do grupo atingido frente ao GC 1, já que há motivos para se acreditar que também os mercados dos municípios no GC 1 tenham, em média, sido prejudicados pelo rompimento.

Os resultados do estudo realizado com a coorte de indivíduos, o qual se voltou à trajetória no mercado de trabalho formal das pessoas que antes do rompimento possuíam vínculo formal junto a estabelecimentos em municípios atingidos, mostrou-se também elucidativa com respeito aos impactos do rompimento da Barragem de Fundão. Nela, cinco indicadores de impacto foram investigados: (i) vínculo ativo no ano; (ii) salário condicional a vínculo ativo no ano; (iii) vínculo ativo no ano junto ao mesmo estabelecimento de 2014; (iv) vínculo ativo no ano no mesmo setor de atividade

⁹¹ O número 20.475 resulta do cálculo $45 \cdot 455 = 20.475$, em que 45 é o número de municípios atingidos e 455 é a quantidade de vínculos ativos que, em média, os municípios atingidos teriam a mais em 2017 do que o que de fato se observou.

⁹² Valores em R\$ de dezembro de 2017.

econômica de 2014; e (v) salário condicional a vínculo ativo no ano no mesmo setor de atividade econômica de 2014.

A avaliação de impactos sobre o primeiro indicador sugere que a probabilidade de pessoas do grupo atingido terem vínculo de trabalho ativo no mercado de trabalho formal nos anos seguintes ao rompimento não se alterou, em média, frente à das pessoas do GC 1 ou do GC 2. Isto porque as estimativas do modelo, apesar de todas negativas, não são estatisticamente significantes ao nível de 10% ou, quando o são, mostram-se demasiadamente sensíveis à maneira como o grupo de comparação é construído. Uma exceção relevante, no entanto, diz respeito às pessoas da coorte que em 2014 trabalhavam no setor de mineração: nesse subgrupo, os modelos estimam de forma robusta que a probabilidade de se ter vínculo de trabalho ativo no mercado formal diminuiu após o rompimento. Particularmente em 2017, 1.506 das 3.872 pessoas do grupo atingido que em 2014 atuavam no setor de mineração não têm vínculo ativo, número que seria menor tivesse o rompimento não ocorrido: 1.052 usando o GC 1 como representação do contrafactual dos atingidos e 1.001 usando o GC 2.⁹³

Levando em consideração o resultado geral para a probabilidade de vínculo ativo no ano conjuntamente com o que se estimou para o número de vínculos ativos do mercado de trabalho formal dos municípios atingidos, tem-se alguma evidência de que o impacto do rompimento parece ter reduzido o número de vínculos não tanto pela extinção de postos de trabalho, mas sim por meio da inviabilização da abertura de novas vagas. Afinal, as estimativas para o grupo atingido na coorte como um todo não apontam uma mudança na probabilidade de se ter vínculo ativo nos anos posteriores ao rompimento que divergissem de maneira estatisticamente significativa da dos grupos de comparação. Em outras palavras, a taxa a que as pessoas da coorte no grupo atingido deixam de ter vínculo de trabalho ativo no mercado formal parece ter permanecido similar após 2014 à dos grupos de comparação, os quais se supõe não terem sido afetados pelo rompimento. Logo, se os municípios atingidos têm menos vínculos ativos em relação ao GC 2 após o rompimento, porém as pessoas que antes lá estavam empregadas não saíram do mercado formal de trabalho em proporção maior do que as que antes trabalhavam nos GC 1 ou 2, o fenômeno que o rompimento parece ter desencadeado nos municípios atingidos é o de impedir que novos postos formais de trabalho se

⁹³ Estes números resultam dos cálculos $3.872 \cdot [1 - (0,611 + 0,1173)] = 1.052,022 \approx 1.052$ e $3.872 \cdot [1 - (0,611 + 0,1306)] = 1.000,525 \approx 1.001$, onde 3.872 é o número de pessoas do grupo atingido que em 2014 trabalhavam no setor de mineração, 0,611 é a parcela das pessoas do subgrupo atingido que em 2017 trabalhava no setor de mineração e 0,1173 e 0,1306 são o valor absoluto das estimativas de impacto sobre o indicador quando se usa, respectivamente, o GC 1 e 2 como grupos de comparação, conforme consta na Tabela 29 no APÊNDICE D — Tabelas e figuras adicionais do capítulo 4.

abrissem. É um aspecto importante do impacto sobre o mercado de trabalho formal, porque pode-se especular que, devido a ele, o rompimento tenha afetado principalmente os mais jovens, em especial os que só viriam a ingressar no mercado de trabalho após 2015, uma vez que se deparariam com menos oportunidades de emprego formal do que haveria se o rompimento não tivesse acontecido.

Com relação ao segundo indicador de impacto analisado, o de salário para os indivíduos que têm vínculo ativo ao longo dos anos, as estimativas se revelam negativas para os anos posteriores ao rompimento, porém estatisticamente significantes ao nível de 10% somente quando da comparação com o GC 2. Mais uma vez, em linha com as análises em FGV (2020a), as evidências de impacto produzidas pelo modelo de diferença-em-diferenças apontam para a existência de “transbordamento” do efeito do rompimento, uma vez que os salários das pessoas no grupo atingido não se alteram desde 2014, em média, frente aos das pessoas no GC 1. Quando comparadas às pessoas do GC 2, no entanto, estima-se que a remuneração dos indivíduos atingidos na coorte que tiveram vínculo formal ativo em 2015 se reduziu R\$ 72 em média, agravando-se para queda de R\$ 91 em 2016 e de R\$ 107 em 2017 frente ao que se observava em 2014.⁹⁴ Ainda considerando o GC 2 como representação da situação contrafactual do grupo atingido, os modelos estimados em subgrupos de pessoas do grupo atingido revelam efeitos heterogêneos do rompimento. Dentre eles, as estimativas sugerem que o impacto do rompimento foi mais intenso sobre o salário de pessoas que antes do desastre trabalhavam em estabelecimentos mais próximos ao Rio Gualaxo do Norte, Rio do Carmo e Rio Doce. Em 2017, por exemplo, o subgrupo de indivíduos da coorte que em 2014 estava empregada em estabelecimentos localizados em setores censitários a até 10 km do rio teve perda salarial média estimada em R\$ 131.

Quanto à avaliação de impactos do rompimento sobre a probabilidade de ter vínculo ativo no ano junto ao mesmo estabelecimento de 2014, as estimativas são sempre negativas, independentemente do grupo de comparação utilizado, mas os efeitos não podem ser considerados estatisticamente diferentes de 0 ao nível de 10% de significância. Com relação aos impactos estimados sobre a probabilidade de possuir vínculo ativo no mesmo setor de atividade econômica em que o indivíduo trabalhava em 2014, também não se pode rejeitar ao nível de 10% a hipótese de que os impactos do rompimento neste indicador sejam iguais a 0 independentemente do grupo de comparação utilizado na avaliação (GC 1 ou GC 2). No entanto, a análise de possível heterogeneidade nos impactos do rompimento sobre esse indicador revelou por meio

⁹⁴ Valores em R\$ de dezembro de 2017.

da comparação entre grupo atingido e GC 2 que houve uma queda na probabilidade de possuir vínculo ativo no mesmo setor de atividade econômica para os indivíduos que trabalhavam em 2014 no setor de agropecuária e pesca. Estima-se que o rompimento da Barragem de Fundão levou 1.593 pessoas que antes trabalhavam nesse setor de atividade a não conseguir se manter nele com um vínculo de trabalho formal ativo em 2017.⁹⁵

Para o último indicador de impacto analisado, o salário condicional a possuir vínculo ativo no mesmo setor de atividade econômica de 2014, a comparação entre grupo atingido e GC 1 não aponta a existência de impacto do rompimento que, em média, seja estatisticamente significativa ao nível de 10%. Entretanto, quando se utiliza o GC 2 como representação da situação contrafactual do atingidos, observa-se que o salário foi R\$ 95 e R\$ 106 menor em média para os indivíduos do grupo atingido do que seria sem o rompimento em 2016 e 2017, respectivamente, efeitos que se mostram estatisticamente diferentes de 0 ao nível de significância de 10%. Na comparação com o GC 2, foram encontrados ainda indícios de existência de impactos heterogêneos do rompimento em subgrupos dos indivíduos atingidos da coorte. Dentre eles, estimou-se que a redução de salário dos homens que tiveram vínculo ativo no mesmo setor de atividade de 2014 foi mais intensa que a estimada para todos os atingidos, sendo estatisticamente significativa ao nível de 10% em 2015 e 2016, embora não mais em 2017. Para as mulheres o impacto também foi negativo, no entanto persistiu até 2017, ano em que foi estatisticamente significativo e alcançou magnitude maior em valor absoluto que o estimado para todo o grupo atingido. Outro efeito heterogêneo encontrado refere-se à localização do setor censitário do estabelecimento em que as pessoas da coorte estavam empregadas em 2014, a qual parece ser um fator associado à intensidade dos impactos do rompimento. Estimou-se que, em média, o salário das pessoas com vínculo ativo no mesmo setor de atividade econômica de 2014 teve perdas maiores para aquelas que trabalhavam relativamente próximo aos rios antes do rompimento.

Por fim, é importante ressaltar que os resultados apresentados neste relatório se referem apenas aos impactos do rompimento no mercado de trabalho formal e em um horizonte temporal pós-rompimento de apenas três períodos (31 de dezembro de 2015, 2016 e 2017). Estas duas restrições são causadas pela inexistência de dados que

⁹⁵ Este número resulta da subtração: $10.326 - 8.733 = 1.593$. O número de 10.326 pessoas resulta do cálculo: $20.735 \cdot [0,421 + 0,077] = 10.326,03 \approx 10.326$, onde: (i) 20.735 é a quantidade de pessoas que trabalhavam em agropecuária e pesca em 2014; (ii) 0,421 é a parcela de pessoas do subgrupo que teve em 2017 vínculo ativo no mesmo setor de 2014; e (iii) 0,077 é o valor absoluto da estimativa de impacto para 2017. Por sua vez, 8.733 é a parcela de indivíduos do grupo atingido que trabalhavam em agropecuária e pesca em 2014 e tiveram vínculo ativo em 2017 nesse mesmo setor.

permitam a realização de uma avaliação de impactos similar para o mercado de trabalho informal e para a indisponibilidade de dados mais recentes da RAIS que permitam ampliar o período de análise. Dessa forma, é possível que os impactos do rompimento da Barragem de Fundão sobre emprego e renda do trabalho sejam ainda maiores caso se considere os efeitos no mercado de trabalho informal e em períodos mais longos (até 2020, por exemplo). Com efeito, os resultados obtidos em FGV (2020b) mostram indícios de que o impacto do rompimento da Barragem de Fundão sobre a renda agregada de Minas Gerais e Espírito Santo considerados conjuntamente se encontrava em trajetória de crescimento durante período de 2015 a 2017, o que vai ao encontro dessa conjectura.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS (ANA). **Massas d'Água** (Espelhos d'Água). Brasília: ANA, 2016. Disponível em: <<http://www.metadados.inde.gov.br/geonetwork/srv/br/metadata.show.embedded?uuid=7d054e5a-8cc9-403c-9f1a-085fd933610c>>. Acesso em: 5 jun. 2019.

BRASIL. **Solicitar consulta às estatísticas da Rais e do Caged**. 2020c. Disponível em: <<https://www.gov.br/pt-br/servicos/solicitar-consulta-as-estatisticas-da-raais-e-do-caged>>. Acesso em: 1 dez. 2020.

_____. Comitê Interfederativo (CIF). **Deliberação nº 58**. Belo Horizonte: CIF, 2017. Disponível em: <http://www.ibama.gov.br/phocadownload/cif/deliberacoes/2017/cif-2017-03-31-deliberacao_58.PDF>. Acesso em: 8 dez. 2020.

_____. Ministério da Economia (ME). **Manual de orientação da Relação Anual de Informações Sociais (RAIS)**: Ano-base 2019. Brasília: ME, 2020a. Disponível em: <http://www.rais.gov.br/sitio/rais_ftp/ManualRAIS2019.pdf>. Acesso em: 1 dez. 2020.

_____. Ministério da Economia (ME). **Programa de Disseminação das Estatísticas do Trabalho**: O que é RAIS? Brasília: ME, 2020b. Disponível em: <<http://pdet.mte.gov.br/o-que-e-rais>>. Acesso em: 1 dez. 2020.

_____. Ministério da Saúde (MS). **Diretrizes metodológicas**: Elaboração de revisão sistemática e metanálise de estudos observacionais comparativos sobre fatores de risco e prognóstico. Brasília: MS, 2014. Disponível em: <http://bvsmms.saude.gov.br/bvs/publicacoes/diretrizes_metodologicas_fatores_risco_prognostico.pdf>. Acesso em: 14 jan. 2021.

_____. Tribunal Federal Regional da 1ª Região. Ação Civil Pública nº 006975861.2015.4.01.3400. **Termo de Transação e Ajustamento de Conduta (TTAC)**. Brasília: 2016. Disponível em: <www.samarco.com/wp-content/uploads/2016/07/TTACFINAL.pdf>. Acesso em: 24 nov. 2020.

COFFMAN, M.; NOY, I. Hurricane Iniki: Measuring the long-term economic impact of a natural disaster using synthetic control. **Environment and Development Economics**, v. 17, n. 2, p. 187-205, 2011.

DER LOO, M. **String Metrics in Stringdist**. RDocumentation, 2019. Disponível em: <<https://www.rdocumentation.org/packages/stringdist/versions/0.9.5.5/topics/stringdist-metrics>>. Acesso em: 16 dez. 2020.

DERYUGINA, T.; KAWANO, L.; LEVITT, S. The economic impact of hurricane Katrina on its victims: Evidence from individual tax returns. **American Economic Journal: Applied Economics**, v. 10, n. 2, p. 202-233, 2018.

FUNDAÇÃO GETULIO VARGAS (FGV). **Impacto do Rompimento da Barragem de Fundão sobre a Renda Agregada nos Estados e Sub-Regiões de Minas Gerais e Espírito Santo**. Rio de Janeiro; São Paulo: FGV, 2020a

_____. **Impactos do Rompimento da Barragem de Fundão sobre a Renda Agregada de Minas Gerais e Espírito Santo**: Análise de Determinantes e Canais a partir do Modelo de Equilíbrio Geral. Rio de Janeiro; São Paulo: FGV, 2020b.

_____. **Impactos sobre Assistência Social a partir de Dados Secundários.** Rio de Janeiro; São Paulo: FGV, 2019a. Disponível em: <http://www.mpf.mp.br/grandes-casos/caso-samarco/documentos/fgv/fgv_impactos-sobre-assistencia-social-a-partir-de-dados-secundarios.pdf>. Acesso em: 10 nov. 2020.

_____. **Impactos sobre Educação Básica a partir de Dados Secundários.** Rio de Janeiro; São Paulo: FGV, 2019b. Disponível em: <http://www.mpf.mp.br/grandes-casos/caso-samarco/documentos/fgv/fgv_impactos-sobre-educacao-basica-a-partir-de-dados-secundarios>. Acesso em: 10 nov. 2020.

_____. **Impactos sobre Segurança Pública a partir de Dados Secundários.** Rio de Janeiro; São Paulo: FGV, 2019c. Disponível em: <http://www.mpf.mp.br/grandes-casos/caso-samarco/documentos/fgv/fgv_impactos-sobre-seguranca-publica-a-partir-de-dados-secundarios.pdf>. Acesso em: 10 nov. 2020.

GROEN, J.; KUTZBACH, M.; POLIVKA, A. Storms and jobs: The effect of hurricanes on individuals' employment and earnings over the long term. **Journal of Labor Economics**, v. 38, n. 3, p. 653-685, 2020.

IMBENS, G.; RUBIN, D. **Causal inference for statistics, social, and biomedical sciences:** An introduction. Nova York: Cambridge University Press, 2015.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Base Cartográfica Contínua do Brasil, ao Milionésimo** – BCIM: 4ª versão. Rio de Janeiro: IBGE, 2014a. v. 1. Disponível em: <https://geofp.ibge.gov.br/cartas_e_mapas/bases_cartograficas_continuas/bcim/versao2014/informacoes_tecnicas/bcim_v4_doc_tecnica_vol_i.pdf>. Acesso em: 8 mai. 2020.

_____. **Classificação Nacional de Atividades Econômicas Versão 2.0:** Subclasses para uso da administração pública. Rio de Janeiro: IBGE, 2007. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv36477.pdf>>. Acesso em: 9 nov. 2020.

_____. **Divisão regional do Brasil em mesorregiões e microrregiões geográficas.** Rio de Janeiro: IBGE, 1990. v. 1. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/monografias/GEBIS-RJ/DRB/Divisaoregional_v01.pdf>. Acesso em: 19 nov. 2019.

_____. **Estimativas da população residente nos municípios brasileiros com data de referência em 1º de julho de 2014.** Rio de Janeiro: IBGE, 2014b. Disponível em: <<https://biblioteca.ibge.gov.br/index.php/biblioteca-catalogo?view=detalhes&id=297745>>. Acesso em: 27 out. 2020.

_____. **Índice Nacional de Preços ao Consumidor Amplo – IPCA:** Séries históricas. Rio de Janeiro: IBGE, 2019a. Disponível em: <<https://www.ibge.gov.br/estatisticas/economicas/precos-e-custos/9256-indice-nacional-de-precos-ao-consumidor-amplo.html?=&t=series-historicas>>. Acesso em: 16 set. 2020.

_____. **Malha de setores censitários.** Rio de Janeiro: IBGE, 2019b. Disponível em: <<https://www.ibge.gov.br/geociencias/organizacao-do-territorio/malhas-territoriais/26565-malhas-de-setores-censitarios-divisoes-intramunicipais.html?edicao=26589&t=o-que-e>>. Acesso em: 10 nov. 2020.

_____. **Pesquisa Nacional por Amostra de Domicílios Contínua**: Notas técnicas versão 1.8. Rio de Janeiro: IBGE, 2020. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv101733_notas_tecnicas.pdf>. Acesso em: 25 nov. 2020.

_____. **Sinopse do Censo Demográfico 2010**. Rio de Janeiro: IBGE, 2011. Disponível em: <<https://biblioteca.ibge.gov.br/visualizacao/livros/liv49230.pdf>>. Acesso em: 2 mai. 2019.

JOY OF DATA. **Comparison of String Distance Algorithms**. Alemanha, 2013. Disponível em: <<http://www.joyofdata.de/blog/comparison-of-string-distance-algorithms/>>. Acesso em: 16 dez. 2020.

LENZEN, M.; ET AL. Economic damage and spillovers from a tropical cyclone. **Natural Hazards & Earth System Sciences**, v. 19, n. 1, 2019.

MARTÍNEZ, M. J.; MARTÍNEZ, M. J.; ROMERO-JARÉN, R. How resilient is the labour market against natural disaster? Evaluating the effects from the 2010 earthquake in Chile. **Natural Hazards**, v. 104, n. 2, p. 1.481-1.533, 2020.

RIBEIRO, F. G. ET AL. O impacto econômico dos desastres naturais: o caso das chuvas de 2008 em Santa Catarina. **Planejamento e Políticas Públicas**, n. 43, 2014.

SKABA, D. A.; **Contrato no AS – 338/2017** – Produto 03. Brasil: Ministério do Desenvolvimento Social (MDS), 2017.

SOUZA, P. H. G. F. de. **A desigualdade vista do topo**: a concentração de renda entre os ricos no Brasil, 1926-2013. 2016. 377 f. Tese (Doutorado em Sociologia) – Universidade de Brasília, Brasília, 2016.

WIKIPEDIA. **Jaro-Winkler distance**. 2020. Disponível em: <https://en.wikipedia.org/wiki/Jaro%E2%80%93Winkler_distance>. Acesso em: 16 dez 2020.

APÊNDICE A — Construção do painel de indivíduos

Este apêndice descreve os procedimentos executados para construir o painel de indivíduos a partir do qual é desenvolvida a análise do Capítulo 4. Em essência, o esforço se deu em dois sentidos:

- Transformar o que originalmente é uma base de dados em nível de vínculos de trabalho (os microdados brutos da RAIS – ME) em uma em nível de pessoas, já que um indivíduo pode ter tido mais de um vínculo de trabalho em um mesmo ano;
- Rastrear cada pessoa encontrada em 2014, ano de definição da coorte, nos microdados da RAIS referentes aos demais anos (cada ano consiste em um arquivo separado), obtendo assim uma base de dados com estrutura de painel.

Para tanto, foi desenvolvido um algoritmo que basicamente explora três informações dos vínculos nos microdados da RAIS: o nome e os números do PIS e CPF das pessoas que os possuíam. Com isso, foi possível realizar comparações entre vínculos de trabalho a fim de decidir se eram ou não pertencentes a uma mesma pessoa, procedimento que viabilizou a construção do painel.

Para fins de exposição, o algoritmo foi dividido em quatro etapas, três das quais dependem das mencionadas comparações entre vínculos com base no nome e números do PIS e CPF dos indivíduos. Elas são realizadas por uma regra que determina se um conjunto de nomes é ou não suficientemente similar para ser considerado pertencente a uma única pessoa. A regra foi definida arbitrariamente e é descrita ao longo da apresentação da Etapa 1 do algoritmo.

Etapa 1: Seleção de pessoas que compõem a coorte

O primeiro passo consiste na determinação de quais indivíduos integram a coorte do estudo. Como discutido no Capítulo 4, este conjunto se refere às pessoas que em 31 de dezembro de 2014 possuíam vínculo de trabalho vigente com remuneração não nula (“vínculo ativo”⁹⁶) junto a estabelecimentos localizados em municípios pertencentes ao grupo atingido ou a um dos grupos de comparação. Portanto, a primeira etapa de construção do painel se volta aos microdados da RAIS de 2014. O objetivo é obter uma lista de indivíduos, porém a base da RAIS está originalmente em nível de vínculos e, como mencionado, uma pessoa pode estar associada a vários vínculos de trabalho em um mesmo ano.

⁹⁶ A definição de “vínculo ativo” é dada originalmente na seção 3.1.5.1.

Uma maneira de proceder é agrupar todos os vínculos cujas informações remetem a uma mesma pessoa e selecionar apenas um deles para representá-la no painel. De fato, cada observação da base bruta da RAIS apresenta dados do trabalhador que poderiam servir a esse fim, como o nome, o número do CPF e o número de inscrição no PIS declarados pelo empregador. Entretanto, erros no preenchimento destas informações podem ocorrer, de maneira que um mesmo número do PIS ou do CPF, conforme consta na base, pode agregar vínculos de pessoas diferentes. Por sua vez, o uso somente dos nomes para localizar pessoas entre os milhões de vínculos registrados é excessivamente exigente em termos computacionais e também suscetível a falhas, já que também podem acontecer erros em preenchimento de nomes, assim como casos de homônimos.

Para lidar com esses problemas, construiu-se uma regra decisória que, dentro do conjunto de vínculos ligados a cada PIS ou CPF, compara os nomes das pessoas e determina por semelhança textual se:

- Trata-se de múltiplos vínculos de um mesmo indivíduo; ou
- Há múltiplos indivíduos registrados sob um mesmo PIS ou CPF.

A regra decisória é apresentada a seguir.

Regra decisória sobre similaridade de nomes⁹⁷

A regra decisória é usada para julgar se os nomes a ela apresentados são ou não pertencentes a um mesmo indivíduo. Os critérios através dos quais se produz tal decisão são arbitrários, tendo sido calibrados ao longo de seu desenvolvimento. Em sua versão final, a regra atingiu desempenho superior a 94% em dois testes de *benchmark* realizados com, respectivamente, 500 e 200 grupos de PIS ou CPF que possuem mais de um nome associado selecionados aleatoriamente – isto é, a regra julgou corretamente se os nomes pertenciam a pessoas iguais ou diferentes em 94% dos casos. Seus critérios, cujos parâmetros são explicados logo a seguir, são:

- I Se $q_{\text{nome completo}} \geq 15$ e $jw_{\text{sigla}} \geq 0,15$, então são pessoas diferentes;
- II Caso $jw_{\text{sigla}} < 0,15$, são a mesma pessoa;

⁹⁷ As informações sobre as métricas de distância textual usadas foram retiradas das seguintes referências: Joy of data (2013), der Loo (2019), Wikipedia (2020).

III Caso $q_{\text{nome completo}} < 15$ e $jw_{\text{sigla}} \geq 0,15$, se $q_{\text{nome completo}} < 5$ ou $jw_{\text{primeiro nome}} < 0,15$ ou $jw_{\text{primeiro nome modificado}} < 0,15$ ⁹⁸, então são a mesma pessoa;

IV Caso $q_{\text{nome completo}} \geq 5$ e $jw_{\text{sigla}} \geq 0,15$ e $jw_{\text{primeiro nome}} \geq 0,15$ e $jw_{\text{primeiro nome modificado}} \geq 0,15$, então:

a. Se $jw_{\text{sobrenome}} < 0,15$ e $jw_{\text{primeiro nome}} < 0,25$, são a mesma pessoa;

b. Caso contrário, são pessoas diferentes.

Aqui, q se refere à distância em “q-grams”. Um “q-gram” é toda sequência de comprimento q de caracteres consecutivos em um pedaço de texto. Ao comparar dois pedaços de texto, essa métrica de distância conta quantas sequências de um pedaço não estão no outro (e vice-versa). Adotou-se $q = 1$ para a construção do painel, de forma que a métrica conta quantos caracteres de um texto não são comuns ao outro.

Por sua vez, jw se refere à distância Jaro-Winkler, dada pela fórmula:

$$jw(A, B) = \begin{cases} \left[1 - \frac{1}{3} \cdot \left(\frac{m}{T_A} + \frac{m}{T_B} + \frac{m-t}{m} \right) \right] \cdot (1 - l \cdot p), & \text{se } m > 0, \\ 1, & \text{se } m = 0 \end{cases}$$

em que:

- T_A é o tamanho do texto A e T_B é o tamanho do texto B ;
- m é o número de caracteres que A e B têm em comum⁹⁹;
- t é o número de transposições necessárias entre caracteres em comum para que fiquem na mesma ordem;
- l é o mínimo entre o número de caracteres até a primeira diferença e o número 4;
- $p \in [0, 0,25]$ é o parâmetro de Winkler, que reduz a distância entre textos que começam de maneira igual. Adota-se $p = 0,1$, seguindo o trabalho do próprio Winkler.

Etapa 1 (continuação)

⁹⁸ “Primeiro nome modificado” é a primeira palavra que aparece no nome da pessoa que não é “José” nem “Maria”. A opção metodológica de não considerar esses dois tipos de primeiros nomes foi feita para aumentar a precisão dos resultados obtidos com a regra decisória, uma vez que se tratam de nomes observados com frequência alta no Brasil.

⁹⁹ Considera-se que um caractere c é comum aos textos A e B se a distância em caracteres entre a posição na qual c ocorre em A e a posição na qual ocorre em B for menor ou igual a $\lfloor \max\{T_A, T_B\}/2 - 1 \rfloor$.

A regra decisória é usada em quatro momentos:

- I Para cada número PIS existente na base, destaca-se todos os nomes a ele associados. Para os números que possuem mais de um nome, aplica-se a regra decisória, gerando uma marcação de quais números de PIS estão associados a múltiplos vínculos de uma mesma pessoa ou de pessoas potencialmente diferentes;
- II Dentre os vínculos cujos PIS foram marcados como associados a pessoas diferentes no passo anterior, lista-se os números de CPF e se destaca os nomes associados a cada um deles. Aplica-se a regra decisória, gerando uma marcação de quais números CPF (dentre os dos vínculos cujos PIS foram sinalizados como problemáticos no passo anterior) estão associados a múltiplos vínculos de uma mesma pessoa ou de pessoas potencialmente diferentes. Gera-se aqui uma lista de vínculos candidatos a descarte, uma vez que o número do PIS nem posteriormente o do CPF foram capazes de identificar uma única pessoa a eles associada;
- III Repete-se o passo I, porém utilizando como referência o número do CPF ao invés do PIS;
- IV Repete-se o passo II, porém utilizando o número do PIS como referência dentre os vínculos cujos CPF foram sinalizados como problemáticos no passo III. Gera-se aqui uma segunda lista de vínculos candidatos a descarte, uma vez que o número do CPF nem posteriormente o do PIS foram capazes de identificar uma única pessoa a eles associada.

Apenas os vínculos que constarem nas duas listas de descarte (geradas nos passos II e IV) são efetivamente excluídos da base. Os demais seguem a um procedimento de consolidação, no qual se cria uma nova variável identificadora de indivíduos que atribui um código único de pessoa a todos os vínculos a ela pertencentes.

Para cada indivíduo com múltiplos vínculos, resta determinar qual deles será o que o representa no painel, constituindo assim sua observação em 2014. A abordagem escolhida foi selecionar o “vínculo principal” usando critérios semelhantes aos adotados na PNAD Contínua para determinar o trabalho principal de um respondente (IBGE, 2020, p. 39). Tais critérios consistem em uma hierarquia, sendo que os passos sucessores só são executados se os anteriores não forem capazes de selecionar um único vínculo por pessoa. Cada passo é aplicado somente aos vínculos filtrados pelo passo anterior:

- I Vínculo vigente em 31 de dezembro de 2014 com remuneração não nula (“vínculo ativo”). Se a pessoa não possuía vínculos ativos em 2014, então ela não pode integrar a coorte;
- II Vínculo que especificar o maior número de horas de trabalho contratadas;
- III Vínculo com a maior remuneração em dezembro de 2014;
- IV Vínculo cuja data de admissão for a mais antiga;
- V Se a data de admissão mais antiga for comum a mais de um vínculo, seleciona-se um deles aleatoriamente.

Quando aplicado a uma pessoa que possuiu apenas um vínculo de trabalho em 2014, o procedimento de seleção de vínculo principal se limita ao passo I: verificar se tal vínculo era ou não ativo. Em caso negativo, como especificado anteriormente, a pessoa não pode integrar a coorte. Ao fim da Etapa 1 de processamento, tem-se como resultado uma base de dados no nível de pessoas, em que as informações apresentadas dizem respeito ao principal vínculo de trabalho no mercado formal que tiveram em 2014.

Etapa 2: Busca de indivíduos da coorte nas bases da RAIS dos demais períodos

A segunda etapa é voltada a rastrear as pessoas que integram a coorte nos microdados da RAIS dos sete demais anos (2010 a 2013 e 2015 a 2017) referentes ao Brasil inteiro.¹⁰⁰ Em cada uma dessas bases, filtra-se somente os vínculos associados a um número PIS ou CPF que consta na coorte. Aplica-se então o procedimento descrito na Etapa 1 a esse conjunto restrito de vínculos, o que produz também bases com apenas uma observação por pessoa para cada ano.

Há dois detalhes importantes que merecem destaque. O primeiro deles é que, quando da seleção do vínculo de trabalho principal, uma pessoa que não teve vínculo ativo em determinado ano não é excluída da base. Ao invés disso, sua observação no ano em questão recebe uma marcação indicando que naquele momento do tempo ela não teve vínculo de trabalho ativo.

¹⁰⁰ Note que, na Etapa 1, foi feita a determinação de quais pessoas integram a coorte do estudo, havendo nela uma restrição geográfica: integram a coorte as pessoas que possuíam vínculo ativo em 2014 junto a estabelecimentos localizados em municípios pertencentes ao grupo atingido ou a um dos grupos de comparação. A análise realizada no Capítulo 4 trata-se de um estudo desta coorte, em particular, da trajetória das pessoas que a compõem no mercado formal de trabalho ao longo do tempo, sem que haja restrição geográfica ao local de atuação nos demais períodos anteriores e posteriores ao rompimento analisados (2010 a 2013 e 2015 a 2017, respectivamente). Por isso, para esses sete anos, na Etapa 2 se realiza a busca pelas pessoas que integram a coorte nos microdados da RAIS referentes ao Brasil inteiro.

Já o segundo detalhe chama atenção para o fato de que as pessoas que integram as bases construídas para os demais anos são indivíduos que naquele ano foram registrados na RAIS com números do PIS e/ou do CPF iguais ao de pessoas na base de 2014 que define a coorte. Como pode haver erros de preenchimento, é importante confirmar se as observações de indivíduos com os mesmos números de PIS e/ou CPF ao longo dos anos são de fato referentes a uma mesma pessoa, o que é o foco da Etapa 3 de processamento.

Etapa 3: Comparação entre vínculos da coorte e dos demais anos

É preciso que cada pessoa esteja consistentemente identificada por um mesmo código em todos anos. Para atribuir o código identificador de pessoa criado na Etapa 1 às observações das pessoas dos demais anos, recorre-se mais uma vez aos números do PIS e CPF como maneira de localizar os indivíduos nas várias bases. Porém, como destacado acima, é possível que haja erros de preenchimento nessas informações, sendo assim importante confirmar se as observações de indivíduos com os mesmos números de PIS e/ou CPF ao longo dos anos são de fato referentes a uma mesma pessoa. Consequentemente, a regra decisória de nomes descrita na Etapa 1 é novamente aplicada.

A base de 2014 é empilhada com cada uma das dos demais anos, formando sete novas bases com pares de anos (2010 e 2014, 2011 e 2014 etc.). A partir disso, o problema se torna similar ao abordado nas Etapas 1 e 2: deve-se verificar através dos nomes se, para cada PIS ou CPF, o conjunto de vínculos associado pertence a uma mesma pessoa ou a pessoas diferentes por meio da regra decisória. No caso em que se julga a partir da regra que observações com mesmo PIS ou CPF em 2014 e em outro ano são de pessoas distintas, opta-se pelo retirar essa pessoa da coorte.¹⁰¹ Executado tal procedimento, atribui-se o código identificador de pessoa criado na base de 2014 às observações dela nos demais anos, resultando em oito bases – 2010 a 2017 – no nível de pessoas com código identificador padronizado.

É importante destacar que, ao agregar as informações de períodos diferentes, é possível que nesta etapa sejam encontrados novos vínculos de uma pessoa em um mesmo ano. Como ilustração, suponha que a situação a seguir exista na base de 2014 empilhada com 2015.

¹⁰¹ Com esse descarte, perde-se 125 pessoas do grupo atingido, 437 do grupo de comparação 1 e 501 do grupo de comparação 2. Note que esses números não chegam a 0,05% do tamanho dos grupos na coorte, conforme apresentado na linha referente a 2014 da Tabela 8.

Tabela 14 — Exemplo de situação possível na construção do painel

PIS	CPF	Ano	Nome
1	2	2014	A
3	4	2014	A
5	6	2014	B
3	2	2015	A
1	6	2015	B

Fonte: Elaboração própria (2020).

Como em 2014 o indivíduo A apareceu em duas entradas tanto PIS como CPF diferentes, em nenhum dos quatro passos descritos na Etapa 1 estes dois vínculos foram comparados. Consequentemente, foram tratados como se pertencessem a pessoas diferentes. Todavia, a Etapa 3 considera conjuntamente a informação de 2014 e a de 2015 que traz o PIS de uma das observações de 2014 e o CPF de outra. Assim, ela promoverá a comparação das observações com PIS = 3 no passo 1 e com CPF = 2 no passo 3. Uma vez determinado pela regra decisória que nos dois casos os vínculos pertencem ao indivíduo A, por transitividade concluirá que os dois vínculos de 2014 antes tratados como pertencentes a pessoas diferentes são também da mesma pessoa. Faz-se novamente necessário, portanto, aplicar o algoritmo de seleção de vínculo principal para que o indivíduo A apareça uma única vez em 2014, o ano de definição da coorte.

Note que a informação de diferentes anos pode também levar à comparação de vínculos de pessoas antes corretamente sinalizadas como diferentes. O indivíduo B em 2015 aparece com o mesmo CPF de 2014, porém registrado com o PIS igual ao da primeira observação de A em 2014. Ao analisar os vínculos ligados ao PIS de valor 1, a regra decisória estará lidando com nomes que não havia comparado na Etapa 1. Caso falhe e diagnostique incorretamente que A e B são a mesma pessoa, suas informações serão tratadas com a de um indivíduo com múltiplos vínculos nos dois anos. Isso reforça o papel crucial que a regra decisória desempenha na construção do painel.

Outra ressalva a se fazer é que, se uma pessoa possuir vínculo ativo em mais de um ano seguido, não significa necessariamente que é o mesmo emprego sendo observado em todos os momentos. Isto é, um indivíduo que tinha vínculo ativo em 31 de dezembro de 2014 pode ter mudado de emprego formal ao longo de 2015 e permanecido nessa posição até ao menos o último dia do ano. O painel permite observar o estabelecimento e setor de atividade de cada vínculo ativo, de modo que é possível identificar se a pessoa muda ou não de emprego, algo que é explorado pelos indicadores de impacto definidos nas seções 4.1.2.3 a 4.1.2.5.

Por outro lado, uma pessoa ter em dado ano um vínculo ativo que é diferente do existente em outros períodos não significa necessariamente que o vínculo observado anteriormente foi encerrado. Pode ocorrer de ela ter acumulado empregos e seu novo vínculo de trabalho ter sido selecionado como o principal, de modo que o vínculo anterior, embora ainda ativo, deixe de ser representado no painel. Este problema tende a ser pouco expressivo, no entanto: em 2015, apenas 9 pessoas do grupo atingido apresentaram mais de um vínculo ativo (0,002%), 51 do GC 1 (0,005%) e 42 do GC 2 (0,004%). Os números são semelhantes nos demais anos.

Etapa 4: Empilhamento e finalização do painel

Nesta última etapa, as oito bases geradas na Etapa 3 são empilhadas. Nos casos em que uma pessoa não foi localizada em determinado ano, inclui-se uma observação “negativa” – isto é, acrescenta-se uma linha à base com a marcação de que aquela pessoa não foi encontrada na RAIS daquele ano, o que implica na marcação de que ela também não possuiu vínculo ativo no período. Como resultado, cada ano no painel possui a mesma quantidade de linhas, que corresponde ao número de pessoas na coorte.

APÊNDICE B — Atribuição de estabelecimentos a setores censitários por meio de cruzamento de endereços

A atribuição da informação de setores censitários aos endereços dos estabelecimentos listados na RAIS é realizada por meio das bases de dados do Cadastro Nacional de Endereços para Fins Estatísticos (CNEFE) e a Base Cartográfica Contínua do Brasil, ao Milionésimo (BCIM), ambas de acesso público disponibilizado pelo IBGE. Como descrito no Capítulo 2, a primeira relaciona todos os endereços do Brasil a setores censitários, enquanto a segunda contém informações georreferenciadas de rodovias nacionais e estaduais, aglomerados rurais, terras indígenas e unidades de proteção ambientais. O procedimento aplicado consiste essencialmente em cruzar as bases de dados da RAIS e as do IBGE (CNEFE e BCIM) por meio das informações textuais de endereços que são comuns entre elas. Após isso, atribui-se a cada estabelecimento na RAIS o dado de setor censitário correspondente ao seu endereço encontrado nas bases do IBGE. A metodologia desenvolvida tem como principal referência o trabalho de Skaba (2017), realizado para o então Ministério de Desenvolvimento Social em parceria com o IBGE, que teve como fim atribuir informação de setores censitários aos domicílios registrados no Cadastro Único.

Este apêndice descreve de maneira abreviada as etapas do procedimento de atribuição de informação de setor censitário aos estabelecimentos na RAIS. Antes de prosseguir a elas, cabe destacar que foram realizados os seguintes procedimentos preliminares:

- I Atribuição de setor censitário às unidades geoespacializadas da BCIM: consistiu na interseção da BCIM, a qual contém pontos, linhas e polígonos que identificam a localização de comunidades, rodovias e unidades de conservação ambiental, com a malha de setores censitários do Censo Demográfico de 2010, também do IBGE, a qual contém os polígonos que determinam os limites geográficos dos setores censitários;
- II Padronização de informações de endereço das bases: buscou-se com esse tratamento aumentar a chance de relacionar as bases por meio das informações textuais de endereço. Todos os acentos, preposições e espaços foram removidos, as letras foram transformadas em maiúsculas e, por fim, foram aplicadas regras de fonetização para padronizar letras que possuem o mesmo som. Esse tratamento foi aplicado às variáveis textuais da RAIS e das bases do IBGE que continham informações de nome do logradouro, nome da localidade, tipo do logradouro e título do logradouro;

III Criação de variável identificadora (ID) com base na semelhança entre nomes dos logradouros das bases relacionadas: esse procedimento consistiu na identificação de endereços no CNEFE cujos nomes do logradouro fossem “suficientemente parecidos” com os de endereços de estabelecimentos na RAIS. O critério de semelhança entre nomes de logradouros se pautou por métricas de distância textual, às quais se impôs limites de tolerância arbitrários. Os endereços na RAIS e no CNEFE cujos nomes de logradouro foram julgados como suficientemente parecidos recebem um código identificador único que é registrado em uma variável auxiliar, referida como ID no restante deste apêndice.¹⁰²

O cruzamento entre as bases de dados foi realizado em 41 etapas de pareamento divididas em oito conjuntos denominados “Partes”. A Tabela 15, apresentada ao fim deste apêndice, mostra de maneira sintética o que compõe cada uma das Partes do processamento e o conjunto de endereços a que são aplicadas. A coluna “Condição de aplicação” destaca que, finda a Parte 1, só avançam às seguintes partes os endereços de estabelecimentos na RAIS para os quais não se encontrou nenhum correspondente nas bases do IBGE ou, ainda, aqueles para os quais os endereços correspondentes encontrados estavam distribuídos por mais de 10 setores censitários distintos. Caso contrário, o processamento acaba para os endereços na RAIS que são encontrados em um único setor censitário e avança para a realização de um sorteio se forem encontrados entre 2 e 10 setores, o qual consiste em atribuir aleatoriamente ao endereço RAIS o dado de um dos setores a que foi pareado ao longo das etapas da Parte em questão.¹⁰³

A Tabela 16 provê uma visão mais detalhada das etapas que compõem cada Parte. Na Parte 1, são aplicadas as etapas 1 a 27 considerando somente os endereços do CNEFE localizados no mesmo município dos endereços da RAIS e utilizando o nome do logradouro como pivô do pareamento – ou seja, todas as etapas utilizam essa

¹⁰² As métricas de distância textual utilizadas guardam semelhanças com as descritas no APÊNDICE A — Construção do painel de indivíduos. No entanto, as etapas envolvidas e o processo de calibragem para a determinação dos limites de tolerância foram feitos de maneira específica para a criação da referida variável identificadora (ID), justamente para permitir maior adequação aos dados desta aplicação. O detalhamento da metodologia utilizada neste procedimento preliminar pode ser solicitado aos autores.

¹⁰³ Ademais, o sorteio é ponderado pela quantidade de domicílios que existia em cada setor censitário em 2010 e também leva em consideração o número de vezes que cada setor censitário foi pareado ao endereço da RAIS. Por exemplo, suponha que um endereço da RAIS foi pareado a nove endereços no CNEFE, os quais estão distribuídos por oito setores censitários. Isso significa que dois dos endereços encontrados no CNEFE estão em um mesmo setor. Portanto, esse setor terá chance maior de ser sorteado como o dado de setor censitário a se atribuir ao endereço da RAIS do que teria se tivesse aparecido uma única vez.

informação no cruzamento das bases. A Parte 1 é aplicada a endereços da RAIS cujo nome do logradouro não é ausente (“nome!=miss”) e a sucessão de etapas depende da quantidade de setores censitários encontrados na etapa anterior – isto é, um endereço da RAIS só é levado para a próxima etapa se na atual não tiver sido encontrado nenhum correspondente para ele no CNEFE ou se os endereços que foram encontrados se localizarem em mais de um setor censitário. Ao fim da Parte 1, como descrito no parágrafo anterior, os endereços da RAIS para os quais foi encontrado apenas um setor têm seu processo de atribuição de setor censitário concluído. Os endereços para os quais foram encontrados entre 2 e 10 setores censitários correspondentes são levados ao sorteio, enquanto os endereços para os quais foram encontrados 0 ou mais de 10 setores censitários são levados para a Parte 2 de pareamento.

A Parte 2 é similar à primeira, porém com a diferença de que nela se utiliza o ID como pivô do pareamento. Dessa maneira, apesar de se conceder maior tolerância a erros no registro do nome do logradouro, espera-se aumentar a chance de encontrar pares de endereços no CNEFE para os endereços da RAIS não pareados na Parte 1. As regras que definem a atribuição do setor censitário aos endereços da RAIS e os endereços que deverão ser levados à Parte 3 são as mesmas da Parte 1.

A Parte 3 inclui somente duas etapas de pareamento, a 28 e 29, que utilizam informações do endereço referentes a rodovias. Tanto o CNEFE como a BCIM são usados para a busca de endereços da RAIS, sendo considerados das primeiras somente os trechos das rodovias que cortam o município do endereço da RAIS sendo analisado. Nesta parte, diferentemente das anteriores, não há sorteio: a atribuição de setor censitário acontece apenas para os endereços para os quais se encontra no processo de pareamento um único setor censitário. Os pares de endereços da RAIS para os quais foram encontrados de 2 a 10 setores censitários diferentes são armazenados e agregados aos setores que entram para o sorteio na Parte 4, apenas se nesta não for encontrado um único setor censitário correspondente para o endereço da RAIS em análise, como explicado a seguir. Os endereços para os quais não se encontra nenhum setor censitário ou mais de 10 setores são levados para a Parte 4.

A Parte 4 do processo é uma tentativa de encontrar o setor censitário do endereço da RAIS por meio do nome da localidade do endereço, que pode coincidir com o de uma comunidade rural ou indígena ou mesmo com o nome de uma unidade de conservação. Portanto, aqui também se faz uso das bases da BCIM além da do CNEFE. Nesta parte, o sorteio para casos de endereços para os quais foram encontrados de 2 a 10 setores censitários considera também os setores da Parte 3, caso estes tenham sido armazenados como descrito anteriormente. Se para o endereço da RAIS não for

encontrado nenhum ou mais de 10 setores censitários diferentes nesta Parte, o sorteio se limita aos setores encontrados na Parte 3, caso tenham sido armazenados. Somente os endereços da RAIS que chegam ao término da Parte 4 sem dado de setor censitário atribuído são levados para a Parte 5.

As Partes 5 a 8 são iguais às quatro primeiras em termos das etapas de pareamento envolvidas. A principal diferença é que as informações usadas para executá-las, antes referentes ao mesmo município do endereço da RAIS em análise, agora se referem aos municípios vizinhos. Além disso, não há regra de sorteio, de forma que um endereço da RAIS só recebe informação de setor censitário se aquele setor for o único ao qual ele foi pareado.

A Tabela 17 resume os resultados do processamento, apresentando quantos endereços de estabelecimentos na RAIS registrados em municípios atingidos receberam dado de setor censitário em cada Parte. Do total de 88.060 endereços que foram buscados nas bases do IBGE, 81.573 terminaram o processo com código de setor censitário atribuído (92,6%), sendo que 66.642 não necessitaram de sorteio (equivalente a 81,7% dos endereços encontrados e 75,7% de todos os buscados). A maioria dos endereços da RAIS é localizada já na Parte 1 – 60.372 deles (74% do total encontrado), dos quais 8.502 (14,1%) passaram pela fase de sorteio. Nas Partes 3 e 4, os sorteios têm participação mais relevante no resultado, o que se justifica pelo fato de que as informações usadas para pareamento nessas etapas se referem a áreas potencialmente amplas, propensas a abranger mais de um setor censitário.¹⁰⁴

¹⁰⁴ É em decorrência disso que casos de endereços para os quais se encontra mais de 10 setores censitários diferentes não são levados para a fase de sorteio, mas sim tratados como não encontrados. Entende-se que sortear entre mais que 10 setores retornaria uma informação excessivamente imprecisa da localização do estabelecimento em questão.

Tabela 15 — Partes do processo de atribuição de setores censitários aos endereços da RAIS

Parte	Descrição das etapas percorridas	Município	Condição de aplicação	Regra de sorteio
Parte 1	Etapas 1 a 27, com nome lograd.	Mesmo município	Todos os endereços da RAIS	Entre 2 e 10 setores
Parte 2	Etapas 1 a 27, com ID		Endereços da RAIS que a Parte 1 retornou 0 ou mais de 10 setores	Entre 2 e 10 setores
Parte 3	Etapas 28 e 29		Endereços da RAIS que a Parte 2 retornou 0 ou mais de 10 setores	Entre 2 e 10 setores, acumula com Parte 4
Parte 4	Etapas 30 a 41		Endereços da RAIS que a Parte 3 retornou 0 ou mais de 1 setor	Entre 2 a 10 setores, sorteio conjunto com setores da Parte 3
Parte 5	Etapas 1 a 27, com nome lograd.	Municípios vizinhos	Endereços da RAIS que a Parte 4 retornou 0 ou mais de 1 setor	-
Parte 6	Etapas 1 a 27, com ID		Endereços da RAIS que a Parte 5 retornou 0 ou mais de 1 setor	-
Parte 7	Etapas 28 e 29		Endereços da RAIS que a Parte 6 retornou 0 ou mais de 1 setor	-
Parte 8	Etapas 30 a 41		Endereços da RAIS que a Parte 7 retornou 0 ou mais de 1 setor	-

Fonte: Elaboração própria (2020).

Tabela 16 — Etapas de pareamento entre as bases da RAIS e do IBGE

Etapa	Informações utilizadas para pareamento						Regra para aplicação
	RAIS	CNEFE	Rodovia	Aglomerados rurais	Terras indígenas	Unidade de proteção ambiental	
1	(nome lograd. ou ID)	-	-	-	-	-	nome!=miss
2	(nome lograd. ou ID) + intervalo	-	-	-	-	-	nome!=miss & etapa1>1
3	(nome lograd. ou ID) + número	-	-	-	-	-	nome!=miss & etapa2>1
4	(nome lograd. ou ID) + tipo	-	-	-	-	-	nome!=miss & npar=0 & etapa1>1
5	(nome lograd. ou ID) + tipo + intervalo	-	-	-	-	-	nome!=miss & etapa4>1
6	(nome lograd. ou ID) + tipo + número	-	-	-	-	-	nome!=miss & etapa5>1
7	(nome lograd. ou ID) + título	-	-	-	-	-	nome!=miss & npar=0 & etapa1>1
8	(nome lograd. ou ID) + título + intervalo	-	-	-	-	-	nome!=miss & etapa7>1
9	(nome lograd. ou ID) + título + número	-	-	-	-	-	nome!=miss & etapa8>1
10	(nome lograd. ou ID) + tipo + título	-	-	-	-	-	nome!=miss & npar=0 & etapa4>1 & etapa7>1
11	(nome lograd. ou ID) + tipo+ título + intervalo	-	-	-	-	-	nome!=miss & etapa10>1
12	(nome lograd. ou ID) + tipo+ título + número	-	-	-	-	-	nome!=miss & etapa11>1

Etapa	Informações utilizadas para pareamento						Regra para aplicação
	RAIS	CNEFE	Rodovia	Aglomerados rurais	Terras indígenas	Unidade de proteção ambiental	
13	(nome lograd. ou ID) + CEP	-	-	-	-	-	nome!=miss & npar=0 & etapa1>1
14	(nome lograd. ou ID) + CEP + intervalo	-	-	-	-	-	nome!=miss & etapa13>1
15	(nome lograd. ou ID) + CEP + número	-	-	-	-	-	nome!=miss & etapa14>1
16	(nome lograd. ou ID) + localidade	-	-	-	-	-	nome!=miss & npar=0 & etapa1>1
17	(nome lograd. ou ID) + localidade + intervalo	-	-	-	-	-	nome!=miss & etapa16>1
18	(nome lograd. ou ID) + localidade + número	-	-	-	-	-	nome!=miss & etapa17>1
19	(nome lograd. ou ID) + localidade + tipo	-	-	-	-	-	nome!=miss & npar=0 & etapa4>1 & etapa16>1
20	(nome lograd. ou ID) + localidade + tipo + intervalo	-	-	-	-	-	nome!=miss & etapa19>1
21	(nome lograd. ou ID) + localidade + tipo + número	-	-	-	-	-	nome!=miss & etapa20>1
22	(nome lograd. ou ID) + localidade + título	-	-	-	-	-	nome!=miss & npar=0 & etapa7>1 & etapa16>1
23	(nome lograd. ou ID) + localidade + título + intervalo	-	-	-	-	-	nome!=miss & etapa22>1
24	(nome lograd. ou ID) + localidade + título + número	-	-	-	-	-	nome!=miss & etapa23>1
25	(nome lograd. ou ID) + localidade + tipo + título	-	-	-	-	-	nome!=miss & npar=0 & etapa10>1 & etapa16>1
26	(nome lograd. ou ID) + localidade + tipo+ título + intervalo	-	-	-	-	-	nome!=miss & etapa25>1
27	(nome lograd. ou ID) + localidade + tipo+ título + número	-	-	-	-	-	nome!=miss & etapa26>1
28a	UF + código da estrada			-	-	-	(nome==miss OU etapa1==0) & código da estrada!=miss
29a	UF + código + km da estrada			-	-	-	(nome==miss OU etapa1==0) & código da estrada!=miss & etapa28a != 1
28b	UF "alternativa" + código da estrada	UF + código da estrada		-	-	-	(nome==miss OU etapa1==0) & código da estrada!=miss & etapa29a != 1
29b	UF "alternativa" + código + km da estrada	UF + código + km da estrada		-	-	-	(nome==miss OU etapa1==0) & código da estrada!=miss & etapa28b != 1
30	localidade	-	-	Nome do aglomerado	-	-	etapa 29b == 0 OU etapa 29b > 10
31	nome lograd.	-	-	Nome do aglomerado	-	-	etapa 30 != 1
32	localidade	-	-	-	Nome da terra indígena	-	etapa 31 != 1
33	nome lograd.	-	-	-	Nome da terra indígena	-	etapa 32 != 1

Etapa	Informações utilizadas para pareamento						Regra para aplicação
	RAIS	CNEFE	Rodovia	Aglomerados rurais	Terras indígenas	Unidade de proteção ambiental	
34	localidade	-	-	-	-	Nome da unidade de conservação ambiental	etapa 33 != 1
35	nome lograd.	-	-	-	-	Nome da unidade de conservação ambiental	etapa 34 != 1
36	Localidade		-	-	-	-	etapa 35 != 1
37	localidade + intervalo		-	-	-	-	etapa 36 != 1
38	localidade + número		-	-	-	-	etapa 37 != 1
39	nome lograd.	localidade	-	-	-	-	etapa 38 != 1
40	nome lograd. + intervalo	localidade + intervalo	-	-	-	-	etapa 39 != 1
41	nome lograd. + número	localidade + número	-	-	-	-	etapa 40 != 1

Fonte: Elaboração própria (2020) com base em Skaba (2017).

Observação: Na última coluna, “etapa #” significa o número de setores únicos (distintos) resultante do pareamento na etapa # e “npar” que representa a variável binária (*dummy*) que identifica se o endereço encontrou um único setor censitário em alguma das etapas anteriores. Além disso, “nome” se refere ao nome do logradouro e “miss”, à ausência de informação. Na segunda coluna, “UF alternativa” se refere ao uso de “BR” ao invés da sigla do estado em que o município está, porque alguns endereços apresentam apenas o código da rodovia, sem identificação de jurisdição.

Tabela 17 — Resultados do processamento para atribuição de dado de setor censitário aos endereços da RAIS

Parte	Endereços encontrados		Total
	Sem sorteio	Com sorteio	
1	51.870	8.502	60.372
2	5.542	948	6.490
3	217	490	707
4	6.044	4.991	11.035
5	1.292	-	1.292
6	254	-	254
7	4	-	4
8	1.419	-	1.419
Total encontrado	66.642	14.931	81.573
Não encontrado			6.487
Total			88.060

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME), CNEFE (IBGE) e BCIM (IBGE).

APÊNDICE C — Tabelas adicionais do capítulo 3

Tabela 18 — Média e desvio-padrão do logaritmo natural do número de vínculos ativos por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	7,66	7,58	7,64
	Desvio-padrão	1,76	1,79	1,61
2011	Média	7,74	7,65	7,68
	Desvio-padrão	1,73	1,75	1,63
2012	Média	7,66	7,61	7,66
	Desvio-padrão	1,79	1,79	1,67
2013	Média	7,73	7,68	7,70
	Desvio-padrão	1,75	1,77	1,68
2014	Média	7,75	7,70	7,74
	Desvio-padrão	1,75	1,76	1,64
2015	Média	7,73	7,69	7,73
	Desvio-padrão	1,73	1,74	1,66
2016	Média	7,64	7,62	7,68
	Desvio-padrão	1,74	1,76	1,66
2017	Média	7,73	7,69	7,73
	Desvio-padrão	1,69	1,71	1,62

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Tabela 19 — Estimativas de impacto médio sobre o logaritmo natural do número de vínculos ativos de municípios no grupo atingido (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0213 [0,0466]	0,0276 [0,0456]	-0,0044 [0,0466]
2011	0,0056 [0,0311]	0,0135 [0,0305]	-0,0075 [0,0330]
2012	-0,0028 [0,0195]	0,0038 [0,0186]	-0,0109 [0,0219]
2013	-0,0015 [0,0159]	-0,0035 [0,0154]	-0,0052 [0,0168]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,0114 [0,0147]	-0,0137 [0,0134]	-0,0189 [0,0139]
2016	-0,0145 [0,0208]	-0,0216 [0,0185]	-0,0372* [0,0198]
2017	-0,0156 [0,0279]	-0,0291 [0,0219]	-0,0423 [0,0257]
Intercepto	10,5157*** [0,0042]	10,7437*** [0,0023]	10,3811*** [0,0060]
Nº de observações	1.440	2.160	1.192
Nº de municípios	180	270	149

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
Coeficiente de determinação (R2) ajustado	0,998	0,998	0,998

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 20 — Estimativas de impacto médio sobre o logaritmo natural do número de vínculos ativos de municípios no grupo atingido (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0477 [0,0464]	0,028 [0,0461]	0,0557 [0,0465]
2011	0,0298 [0,0312]	0,0156 [0,0309]	0,0362 [0,0315]
2012	0,0059 [0,0209]	0,0017 [0,0190]	0,0138 [0,0204]
2013	0,0059 [0,0196]	0,0014 [0,0168]	0,0124 [0,0199]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,0267* [0,0137]	-0,0283** [0,0133]	-0,0267* [0,0138]
2016	-0,0400** [0,0199]	-0,0390** [0,0182]	-0,0408** [0,0204]
2017	-0,0431** [0,0218]	-0,0426** [0,0206]	-0,0454** [0,0226]
Intercepto	10,4424*** [0,0041]	10,4442*** [0,0027]	10,4986*** [0,0045]
Nº de observações	1.440	2.160	1.240
Nº de municípios	180	270	155
Coeficiente de determinação (R2) ajustado	0,997	0,997	0,997

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 21 — Média e desvio-padrão do logaritmo natural da massa salarial por grupo

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	14,99	14,87	14,99
	Desvio-padrão	1,94	1,96	1,76
2011	Média	15,09	14,98	15,06
	Desvio-padrão	1,91	1,93	1,79
2012	Média	15,03	14,98	15,08
	Desvio-padrão	1,99	1,96	1,83
2013	Média	15,15	15,08	15,16
	Desvio-padrão	1,93	1,94	1,83
2014	Média	15,17	15,11	15,23
	Desvio-padrão	1,94	1,94	1,80
2015	Média	15,12	15,08	15,21
	Desvio-padrão	1,92	1,93	1,80
2016	Média	15,07	15,03	15,18
	Desvio-padrão	1,91	1,93	1,80
2017	Média	15,19	15,13	15,25
	Desvio-padrão	1,86	1,87	1,75

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Tabela 22 — Estimativas de impacto médio sobre o logaritmo natural da massa salarial de municípios no grupo atingido (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0687	0,0656	0,0428
	[0,0735]	[0,0727]	[0,0764]
2011	0,0379	0,0406	0,0219
	[0,0622]	[0,0596]	[0,0677]
2012	0,0353	0,0392	0,0326
	[0,0488]	[0,0472]	[0,0526]
2013	0,0086	0,0065	0,0145
	[0,0379]	[0,0369]	[0,0395]
2014 (ano-base)	0	0	0
	[.]	[.]	[.]
2015	-0,0195	-0,0216	-0,0315
	[0,0264]	[0,0240]	[0,0256]
2016	-0,0124	-0,0244	-0,0524
	[0,0371]	[0,0324]	[0,0320]
2017	-0,0113	-0,0303	-0,0577
	[0,0494]	[0,0390]	[0,0401]
Intercepto	18,2244***	18,4841***	18,0142***
	[0,0083]	[0,0046]	[0,0123]
Nº de observações	1.440	2.160	1.192
Nº de municípios	180	270	149
Coefficiente de determinação (R2) ajustado	0,996	0,997	0,996

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 23 — Estimativas de impacto médio sobre o logaritmo natural da massa salarial de municípios no grupo atingido (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,1046 [0,0734]	0,0688 [0,0730]	0,1199 [0,0733]
2011	0,07 [0,0606]	0,0461 [0,0599]	0,0808 [0,0608]
2012	0,0397 [0,0490]	0,0296 [0,0475]	0,0514 [0,0490]
2013	0,0106 [0,0391]	0,0025 [0,0373]	0,0184 [0,0394]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,033 [0,0240]	-0,0420* [0,0234]	-0,0329 [0,0242]
2016	-0,0547* [0,0321]	-0,0551* [0,0303]	-0,0555* [0,0327]
2017	-0,0546 [0,0375]	-0,057 [0,0347]	-0,0567 [0,0385]
Intercepto	18,0989*** [0,0081]	18,1293*** [0,0054]	18,1494*** [0,0088]
Nº de observações	1.440	2.160	1.240
Nº de municípios	180	270	155
Coeficiente de determinação (R2) ajustado	0,995	0,995	0,995

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 1, a qual inclui efeitos fixos de ano e município. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município. Observações de municípios são ponderadas pelo número de vínculos ativos que lá existia em 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

APÊNDICE D — Tabelas e figuras adicionais do capítulo 4

Tabela 24 — Distribuição de vínculos ativos em 2014 de MG e ES por características socioeconômicas

Característica	Categorias	MG		ES	
		Total	%	Total	%
Total		4.882.199		931.551	
Sexo	Mulher	2.127.958	43,6%	399.519	42,9%
	Homem	2.754.241	56,4%	532.032	57,1%
Raça/cor	Indígena	7.334	0,2%	1.328	0,1%
	Branca	2.045.402	41,9%	305.238	32,8%
	Preta	284.160	5,8%	51.566	5,5%
	Amarela	29.058	0,6%	5.241	0,6%
	Parda	1.381.562	28,3%	361.865	38,8%
	Não identificado	1.134.683	23,2%	206.313	22,1%
Setor de atividade	Agropecuária e pesca	253.508	5,2%	32.456	3,5%
	Extrativo: petróleo, gás e carvão	1.177	0,0%	3.215	0,3%
	Extrativo: mineração	61.774	1,3%	10.047	1,1%
	Indústrias da transformação	776.698	15,9%	116.200	12,5%
	Infraestrutura	45.478	0,9%	10.286	1,1%
	Indústria civil	327.446	6,7%	56.671	6,1%
	Comércio	982.856	20,1%	200.306	21,5%
	Logística	243.490	5,0%	67.864	7,3%
	Administração e segurança pública	868.339	17,8%	172.536	18,5%
	Outros serviços	1.321.433	27,1%	261.970	28,1%
Faixa etária	19 anos ou menos	232.177	4,8%	42.217	4,5%
	20 a 24 anos	626.051	12,8%	118.503	12,7%
	25 a 29 anos	756.193	15,5%	153.097	16,4%
	30 a 34 anos	779.208	16,0%	154.823	16,6%
	35 a 39 anos	673.028	13,8%	130.268	14,0%
	40 a 44 anos	550.485	11,3%	101.682	10,9%
	45 a 49 anos	489.681	10,0%	89.331	9,6%
	50 a 54 anos	384.247	7,9%	69.680	7,5%
	55 a 59 anos	242.588	5,0%	44.469	4,8%
	60 a 64 anos	105.604	2,2%	20.169	2,2%
Nível educacional	65 anos ou mais	42.937	0,9%	7.312	0,8%
	Ensino Fundamental incompleto	788.020	16,1%	116.998	12,6%
	Ensino Fundamental completo	958.910	19,6%	177.813	19,1%
	Ensino Médio completo	2.281.269	46,7%	470.433	50,5%
	Ensino Superior completo	854.000	17,5%	166.307	17,9%

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

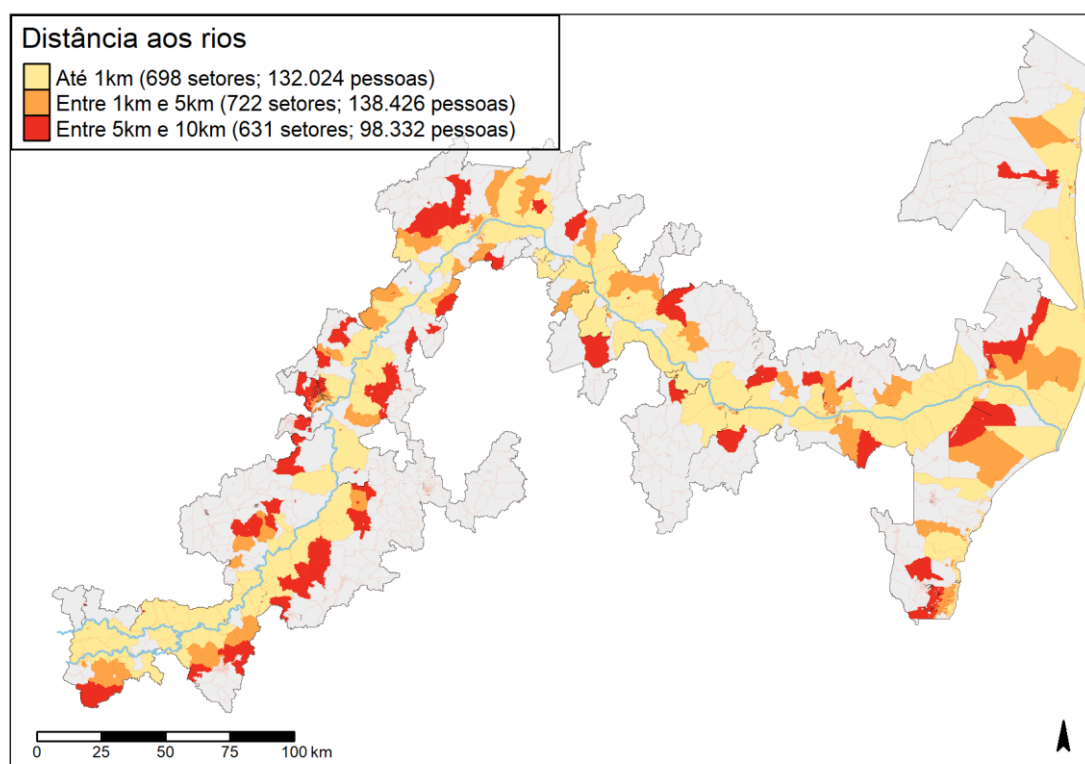
Observação: Os vínculos de trabalho considerados “ativos” são apenas os que constam na RAIS como vigentes em 31 de dezembro com remuneração não nula.

Tabela 25 — Classificação de setores de atividade criada a partir da CNAE 2.0

Setor	Divisão de atividade – CNAE 2.0 (2 dígitos)
Agropecuária e pesca	01, 02 e 03
Extrativo: petróleo, gás e carvão	05, 06 e 09
Extrativo: mineração	07 e 08
Indústrias da transformação	10 a 33
Infraestrutura	35 a 39
Indústria civil	41 a 43
Comércio	45 a 47
Logística	49 a 53
Administração e segurança pública	84
Outros serviços	maior ou igual a 55, exceto 84

Fonte: Elaboração própria (2020) a partir da CNAE 2.0 (IBGE).

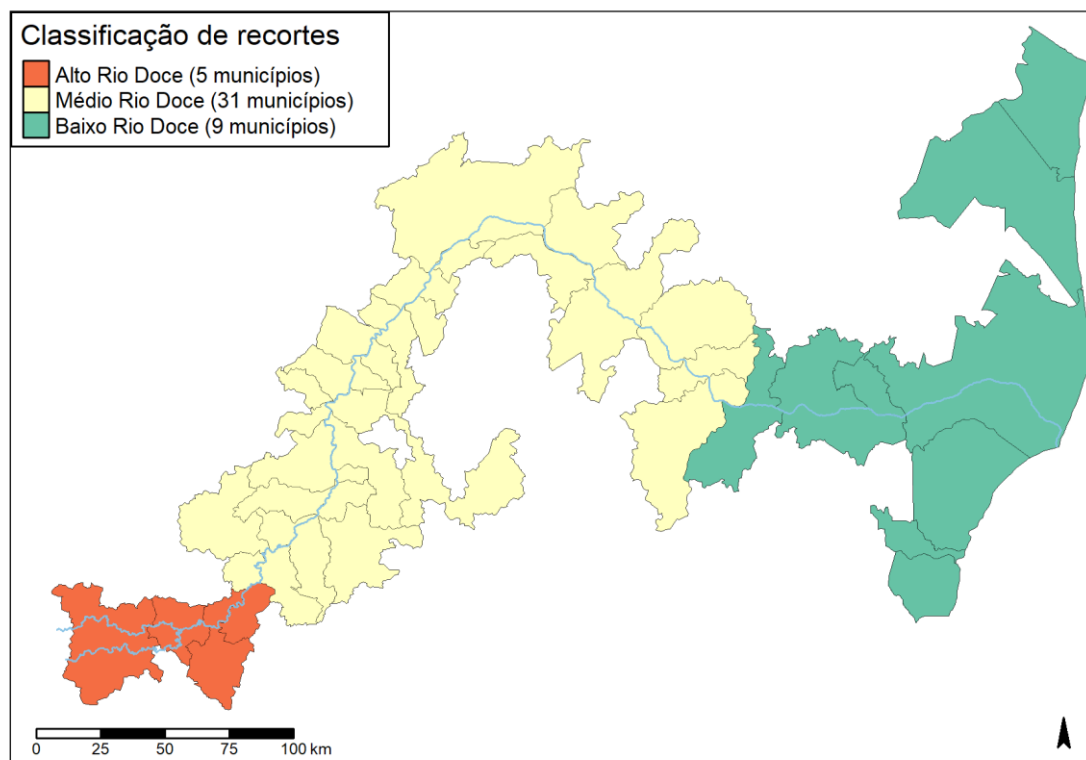
Figura 6 — Classificação de setores censitários conforme distância ao Rio Gualaxo do Norte, Rio do Carmo e Rio Doce



Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME) e de bases cartográficas (ANA e IBGE).

Observação: Considera-se que um setor censitário está a uma dada quantidade de km dos rios se ao menos um ponto dele estiver dentro de um *buffer* ao redor dos rios com raio igual àquela quantidade de km. Setores censitários que em 2014 não apresentavam nenhuma pessoa da coorte não são destacados no mapa (isto é, aparecem em cinza).

Figura 7 — Classificação de municípios do grupo atingido conforme recorte territorial



Fonte: Elaboração própria (2020) a partir de bases cartográficas (ANA e IBGE).

Tabela 26 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	-0,0007 [0,0110]	-0,0037 [0,0125]	-0,0032 [0,0101]
2011	-0,0035 [0,0117]	-0,0064 [0,0120]	-0,0045 [0,0099]
2012	-0,003 [0,0092]	-0,005 [0,0099]	-0,0039 [0,0082]
2013	-0,0042 [0,0080]	-0,0059 [0,0079]	-0,0056 [0,0065]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,011 [0,0079]	-0,0132 [0,0080]	-0,0152** [0,0074]
2016	-0,0072 [0,0096]	-0,0136 [0,0089]	-0,0190** [0,0093]
2017	-0,0108 [0,0093]	-0,0173** [0,0087]	-0,0216** [0,0103]
Intercepto	0,7233*** [0,0017]	0,7264*** [0,0011]	0,7274*** [0,0022]
Nº de observações	15.498.320	25.538.664	10.822.472
Nº de municípios	180	270	149

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
Coefficiente de determinação (R ²) ajustado	0,300	0,302	0,303

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 27 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0207* [0,0108]	0,0126 [0,0104]	0,0250** [0,0106]
2011	0,0172* [0,0099]	0,0103 [0,0096]	0,0214** [0,0097]
2012	0,0124* [0,0073]	0,0061 [0,0069]	0,0159** [0,0072]
2013	0,0086 [0,0076]	0,0019 [0,0065]	0,0121 [0,0077]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,0109 [0,0078]	-0,0129* [0,0074]	-0,0102 [0,0081]
2016	-0,0134* [0,0079]	-0,0137* [0,0076]	-0,0119 [0,0081]
2017	-0,0104 [0,0085]	-0,0112 [0,0089]	-0,0083 [0,0087]
Intercepto	0,7152*** [0,0015]	0,7191*** [0,0010]	0,7127*** [0,0016]
Nº de observações	14.838.368	21.607.728	13.742.480
Nº de municípios	180	270	155
Coefficiente de determinação (R ²) ajustado	0,303	0,302	0,303

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 28 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano por grupo, apenas para aqueles que em 2014 trabalhavam no setor de atividade econômica de mineração

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	0,717	0,694	0,694
	Desvio-padrão	0,451	0,461	0,461
2011	Média	0,769	0,778	0,756
	Desvio-padrão	0,421	0,416	0,429
2012	Média	0,798	0,822	0,799
	Desvio-padrão	0,401	0,382	0,401
2013	Média	0,844	0,861	0,862
	Desvio-padrão	0,363	0,346	0,345
2014	Média	1	1	1
	Desvio-padrão	0	0	0
2015	Média	0,852	0,719	0,830
	Desvio-padrão	0,355	0,450	0,376
2016	Média	0,695	0,755	0,779
	Desvio-padrão	0,461	0,430	0,415
2017	Média	0,611	0,728	0,742
	Desvio-padrão	0,488	0,445	0,438

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: São consideradas apenas pessoas que tinham vínculo ativo no setor de atividade econômica de mineração em 2014.

Tabela 29 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo – Efeitos heterogêneos para aqueles que em 2014 trabalhavam no setor de atividade econômica de mineração

Ano	GC 1	GC 2
2010	0,0224	0,0222
	[0,0819]	[0,0708]
2011	-0,0082	0,0132
	[0,0668]	[0,0655]
2012	-0,0239	-0,001
	[0,0643]	[0,0611]
2013	-0,0176	-0,0178
	[0,0576]	[0,0564]
2014 (ano-base)	0	0
	[.]	[.]
2015	0,1334*	0,0227
	[0,0791]	[0,0442]
2016	-0,0604**	-0,0844***
	[0,0279]	[0,0291]
2017	-0,1173***	-0,1306***
	[0,0380]	[0,0411]
Intercepto	0,7947***	0,8077***
	[0,0095]	[0,0137]

Ano	GC 1	GC 2
Nº de observações	114.832	82.560
Nº de municípios	110	95
Coefficiente de determinação (R2) ajustado	0,252	0,261

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. São incluídas no modelo apenas pessoas que tinham vínculo ativo no setor de atividade econômica de mineração em 2014. Cada coluna da tabela indica o grupo de comparação com o qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 30 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	128,5325 [152,7044]	103,7517 [107,0287]	-6,3826 [83,8841]
2011	66,2091 [123,1514]	64,2717 [97,1303]	-25,7365 [84,6682]
2012	70,2562 [110,5398]	76,7806 [104,6650]	25,6183 [102,0875]
2013	-29,2566 [45,6577]	-24,9466 [42,2810]	-45,4646 [40,8692]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-31,6085 [49,9091]	-35,903 [46,2485]	-46,7556 [43,6522]
2016	-26,5845 [60,2409]	-38,326 [51,2549]	-64,5896 [44,9151]
2017	-16,0115 [80,4154]	-46,1145 [66,4397]	-67,0633 [58,2077]
Intercepto	2495,8479*** [17,0902]	2578,8262*** [8,5507]	2252,3409*** [17,9550]
Nº de observações	11.128.748	18.419.525	7.793.023
Nº de municípios	180	270	149
Coefficiente de determinação (R2) ajustado	0,889	0,894	0,870

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014.

Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 31 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	25,0967 [82,4847]	27,5224 [80,6823]	30,856 [83,3167]
2011	7,2725 [83,6324]	11,1028 [82,4237]	11,5799 [84,1362]
2012	31,5776 [101,8493]	37,8397 [101,3399]	33,7985 [102,0760]
2013	-46,088 [40,8008]	-44,8313 [40,5637]	-42,9569 [40,8805]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-72,4671* [43,5690]	-72,3051* [43,1210]	-76,4426* [43,6891]
2016	-90,8234* [46,2121]	-81,1689* [45,0984]	-97,0716** [46,5310]
2017	-107,1305* [58,6407]	-98,1056* [57,1798]	-113,8979* [59,2091]
Intercepto	2320,5477*** [13,0984]	2374,9642*** [8,9405]	2309,7763*** [14,2026]
Nº de observações	10563849	15448946	9759567
Nº de municípios	180	270	155
Coeficiente de determinação (R2) ajustado	0,855	0,861	0,855

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 32 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo – Efeitos heterogêneos por sexo

Ano	Mulheres	Homens
2010	32,7271 [52,4284]	21,923 [105,5266]
2011	5,2894 [39,2558]	9,0176 [114,9535]
2012	-1,9349 [44,2249]	52,2571 [143,2280]
2013	-13,8847 [31,3161]	-67,7746 [50,0095]
2014 (ano-base)	0 [.]	0 [.]
2015	-40,8450* [24,1073]	-93,2657* [55,5079]

Ano	Mulheres	Homens
2016	-84,5926*** [30,2896]	-93,4956 [59,2116]
2017	-107,8911** [43,1100]	-104,392 [70,4111]
Intercepto	2048,9042*** [6,4006]	2505,3121*** [17,4829]
Nº de observações	4.279.048	6.284.801
Nº de municípios	180	180
Coefficiente de determinação (R2) ajustado	0,836	0,861

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 33 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo – Efeitos heterogêneos por raça/cor

Ano	Indígena	Branca	Preta	Amarela	Parda
2010	-42,6941 [166,1509]	41,5319 [95,0184]	-6,9857 [95,1510]	99,4827 [99,7368]	8,4132 [90,4954]
2011	-90,6553 [132,4917]	23,5785 [94,6523]	-11,591 [98,1289]	19,4219 [83,7852]	14,7265 [99,5968]
2012	-40,6112 [70,2879]	38,6273 [102,4809]	5,6071 [107,1097]	35,8368 [59,9876]	60,8074 [136,0476]
2013	-18,9099 [76,0626]	-33,5452 [39,9278]	-33,4873 [50,3001]	-54,6751 [46,9025]	-41,1756 [42,6763]
2014 (ano-base)	0 [.]	0 [.]	0 [.]	0 [.]	0 [.]
2015	-32,9983 [36,6996]	-61,4547 [44,4254]	-63,9036 [44,9251]	-108,5924** [53,7315]	-80,2760* [45,1451]
2016	-19,5212 [36,4143]	-47,2074 [44,2242]	-54,364 [46,1851]	-189,8267*** [55,3889]	-90,8723* [47,0243]
2017	-18,6263 [76,3012]	-41,0344 [49,5091]	-58,6493 [57,5838]	-241,2613*** [59,1235]	-104,2278* [54,7467]
Intercepto	1686,2261*** [5,9909]	2450,0609*** [12,3654]	1902,3602*** [15,6582]	2448,4316*** [13,0256]	2065,0745*** [17,4694]
Nº de observações	42.765	4.087.220	629.012	66.000	3.617.382
Nº de municípios	114	180	180	146	180
Coefficiente de determinação (R2) ajustado	0,861	0,875	0,829	0,898	0,839

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 34 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo – Efeitos heterogêneos nível de escolaridade

Ano	Fundamental incompleto	Fundamental completo	Médio completo	Superior completo
2010	25,5411 [30,7335]	33,2609 [40,2796]	0,1508 [102,9074]	34,6411 [168,9996]
2011	3,3781 [26,5845]	17,38 [39,8372]	-5,8453 [105,0444]	-2,4511 [154,8001]
2012	10,2407 [24,6113]	42,3113 [52,2947]	37,1758 [124,9886]	-10,1048 [184,3855]
2013	-12,6002 [12,5216]	-17,0631 [12,4333]	-55,9049 [48,5265]	-96,218 [95,9439]
2014 (ano-base)	0 [.]	0 [.]	0 [.]	0 [.]
2015	-16,5353 [13,6066]	-23,9781 [22,8136]	-78,5201 [52,4877]	-166,9543** [78,6651]
2016	-22,2533 [20,4952]	-30,7592 [26,4528]	-90,8521* [51,9717]	-225,3757** [94,1087]
2017	-39,1782* [22,9224]	-41,0873 [34,0098]	-100,0735 [63,0601]	-257,5502** [125,4933]
Intercepto	1675,8756*** [4,2625]	1720,6455*** [5,7753]	2065,1773*** [16,4889]	4587,0339*** [23,2677]
Nº de observações	1.469.322	2.115.972	5.390.885	1.587.670
Nº de municípios	180	180	180	180
Coefficiente de determinação (R2) ajustado	0,795	0,786	0,807	0,843

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 35 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem um vínculo ativo – Efeitos heterogêneos para pessoas cujo emprego em 2014 era em setores censitários próximos aos rios

Ano	Até 1 km	Até 5 km	Até 10 km
2010	26,3384 [57,8541]	23,9478 [112,5896]	11,8218 [102,6357]
2011	-9,7314 [37,5352]	13,2039 [120,5255]	-0,6422 [107,2723]
2012	-31,2209 [43,0778]	45,52 [145,7661]	22,5903 [123,5666]
2013	-30,251 [30,3801]	-62,771 [67,0105]	-59,7882 [52,0599]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-53,2394* [29,7675]	-95,3989 [59,3633]	-92,7834* [51,5612]

Ano	Até 1 km	Até 5 km	Até 10 km
2016	-88,6017** [40,8288]	-114,0404* [59,1644]	-116,6033** [52,5837]
2017	-101,7093* [61,4346]	-121,2429* [71,3362]	-130,5487* [67,3539]
Intercepto	2319,9475*** [2,4449]	2343,3111*** [11,6351]	2333,4853*** [13,0496]
Nº de observações	8.507.754	9.298.859	9.843.551
Nº de municípios	177	179	180
Coefficiente de determinação (R2) ajustado	0,853	0,855	0,854

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 36 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo com o mesmo estabelecimento de 2014 (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	-0,0025 [0,0181]	-0,0051 [0,0200]	-0,0191 [0,0177]
2011	-0,0038 [0,0178]	-0,0064 [0,0200]	-0,0177 [0,0177]
2012	-0,0049 [0,0162]	-0,0072 [0,0184]	-0,0177 [0,0161]
2013	0,0027 [0,0127]	-0,0006 [0,0138]	-0,0095 [0,0129]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,0115 [0,0133]	-0,0126 [0,0139]	-0,0243* [0,0129]
2016	-0,0028 [0,0156]	-0,0081 [0,0162]	-0,0250* [0,0143]
2017	-0,0088 [0,0173]	-0,0128 [0,0178]	-0,0311* [0,0159]
Intercepto	0,5380*** [0,0033]	0,5406*** [0,0022]	0,5521*** [0,0046]
Nº de observações	15.498.320	25.538.664	10.822.472
Nº de municípios	180	270	149
Coefficiente de determinação (R2) ajustado	0,474	0,476	0,473

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 37 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo com o mesmo estabelecimento de 2014 (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0147 [0,0187]	0,0058 [0,0179]	0,0201 [0,0188]
2011	0,0122 [0,0189]	0,0038 [0,0182]	0,0172 [0,0191]
2012	0,0094 [0,0171]	0,0018 [0,0165]	0,0134 [0,0173]
2013	0,0168 [0,0145]	0,006 [0,0130]	0,0215 [0,0148]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,0063 [0,0148]	-0,0104 [0,0134]	-0,0042 [0,0153]
2016	-0,0015 [0,0175]	-0,0037 [0,0156]	0,0015 [0,0182]
2017	-0,0024 [0,0182]	-0,0036 [0,0172]	0,0017 [0,0187]
Intercepto	0,5287*** [0,0037]	0,5341*** [0,0024]	0,5251*** [0,0041]
Nº de observações	14.838.368	21.607.728	13.742.480
Nº de municípios	180	270	155
Coeficiente de determinação (R2) ajustado	0,473	0,472	0,473

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 38 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0004 [0,0152]	-0,0008 [0,0180]	-0,0136 [0,0152]
2011	-0,0024 [0,0157]	-0,0046 [0,0184]	-0,0147 [0,0164]
2012	-0,0051 [0,0146]	-0,0063 [0,0172]	-0,0153 [0,0156]
2013	-0,002 [0,0111]	-0,0047 [0,0123]	-0,0105 [0,0124]
2014 (ano-base)	0 [.]	0 [.]	0 [.]

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2015	-0,0134 [0,0111]	-0,0153 [0,0118]	-0,0253** [0,0107]
2016	-0,0079 [0,0139]	-0,0134 [0,0145]	-0,0280** [0,0129]
2017	-0,0138 [0,0162]	-0,0176 [0,0171]	-0,0339** [0,0151]
Intercepto	0,5827*** [0,0029]	0,5850*** [0,0020]	0,5948*** [0,0043]
Nº de observações	15.498.320	25.538.664	10.822.472
Nº de municípios	180	270	149
Coefficiente de determinação (R2) ajustado	0,438	0,441	0,438

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 39 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	0,0176 [0,0159]	0,0086 [0,0149]	0,0233 [0,0160]
2011	0,0102 [0,0164]	0,0016 [0,0155]	0,0151 [0,0165]
2012	0,0067 [0,0156]	-0,0007 [0,0146]	0,0112 [0,0158]
2013	0,0093 [0,0125]	0,0006 [0,0109]	0,013 [0,0129]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-0,0107 [0,0124]	-0,0156 [0,0111]	-0,0088 [0,0128]
2016	-0,0113 [0,0141]	-0,0143 [0,0128]	-0,009 [0,0146]
2017	-0,0104 [0,0159]	-0,0132 [0,0151]	-0,0072 [0,0162]
Intercepto	0,5757*** [0,0032]	0,5813*** [0,0021]	0,5725*** [0,0035]
Nº de observações	14.838.368	21.607.728	13.742.480
Nº de municípios	180	270	155
Coefficiente de determinação (R2) ajustado	0,442	0,440	0,443

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 40 — Estimativas de impacto médio sobre a probabilidade de pessoas no grupo atingido terem um vínculo ativo no mesmo setor de 2014 – Efeitos heterogêneos para pessoas cujo emprego em 2014 era no setor de atividade econômica de agropecuária e pesca

Ano	GC 1	GC 2
2010	0,0067 [0,0246]	0,0066 [0,0238]
2011	0,0177 [0,0257]	0,0256 [0,0261]
2012	0,0112 [0,0261]	0,012 [0,0269]
2013	0,0101 [0,0257]	0,0182 [0,0254]
2014 (ano-base)	0 [.]	0 [.]
2015	-0,0348 [0,0480]	-0,0773* [0,0459]
2016	-0,0708 [0,0490]	-0,0968** [0,0475]
2017	-0,0406 [0,0385]	-0,0770** [0,0370]
Intercepto	0,5753*** [0,0052]	0,5863*** [0,0053]
Nº de observações	590800	531.736
Nº de municípios	180	176
Coeficiente de determinação (R2) ajustado	0,401	0,388

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. São incluídas no modelo apenas pessoas que tinham vínculo ativo no setor de atividade econômica de agropecuária e pesca em 2014. Cada coluna da tabela indica o grupo de comparação com o qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 41 — Média e desvio-padrão da variável indicadora de vínculo ativo no ano no mesmo setor de atividade econômica de 2014 por grupo, especificamente para pessoas cujo emprego em 2014 era no setor de agropecuária e pesca

Ano	Estatística	Atingido	GC 1	GC 2
2010	Média	0,387	0,381	0,381
	Desvio-padrão	0,487	0,486	0,486
2011	Média	0,464	0,446	0,439
	Desvio-padrão	0,499	0,497	0,496
2012	Média	0,521	0,510	0,509
	Desvio-padrão	0,500	0,500	0,500
2013	Média	0,651	0,641	0,632
	Desvio-padrão	0,477	0,480	0,482
2014	Média	1	1	1
	Desvio-padrão	0	0	0

Ano	Estatística	Atingido	GC 1	GC 2
2015	Média	0,599	0,634	0,676
	Desvio-padrão	0,490	0,482	0,468
2016	Média	0,459	0,530	0,556
	Desvio-padrão	0,498	0,499	0,497
2017	Média	0,421	0,462	0,498
	Desvio-padrão	0,494	0,499	0,500

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: São consideradas apenas pessoas que tinham vínculo ativo no setor de agropecuária e pesca em 2014.

Tabela 42 — Número de pessoas da coorte com vínculo de trabalho ativo no mesmo setor de atividade econômica de 2014 por grupo

Ano	Atingido	GC 1	GC 2
2010	180.327	524.272	470.867
2011	212.055	620.733	568.011
2012	251.993	740.787	682.340
2013	323.990	945.954	876.442
2014	495.392	1.441.898	1.359.404
2015	329.084	977.166	917.526
2016	265.387	783.890	743.627
2017	229.119	686.845	642.835

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Os vínculos de trabalho considerados “ativos” são apenas os que constam na RAIS como vigentes em 31 de dezembro de cada ano com remuneração não nula e no mesmo setor de atividade econômica em que trabalhavam em 2014.

Tabela 43 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 1)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	139,6157 [141,4785]	113,9886 [108,5019]	28,1294 [96,4542]
2011	68,4122 [122,1078]	73,7322 [103,7999]	-6,123 [96,0179]
2012	84,9842 [122,9949]	92,8245 [119,4232]	50,0659 [118,7085]
2013	-25,667 [47,4123]	-22,3293 [44,2331]	-38,5248 [43,4076]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-34,895 [48,9017]	-37,5827 [47,1282]	-45,1248 [45,0742]
2016	-37,0668 [51,8449]	-42,4336 [49,5626]	-60,5311 [44,4402]
2017	-24,7808 [63,3709]	-46,8734 [59,9601]	-50,7388 [54,9081]
Intercepto	2624,5207*** [14,8367]	2731,0386*** [8,0330]	2352,4354*** [17,6514]

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
Nº de observações	8.755.151	14.490.421	6.195.067
Nº de municípios	180	270	149
Coefficiente de determinação (R2) ajustado	0,917	0,919	0,899

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 44 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 (comparando com GC 2)

Ano	3 vizinhos sem reposição	5 vizinhos sem reposição	3 vizinhos com reposição
2010	37,0418 [94,1210]	50,8354 [92,2008]	42,8146 [95,1654]
2011	16,7791 [95,7025]	25,4438 [94,2349]	19,9783 [96,3878]
2012	45,1421 [118,1567]	55,5471 [117,5303]	46,8168 [118,4672]
2013	-43,0107 [43,1794]	-40,7688 [43,0063]	-40,3456 [43,2976]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-72,8037 [44,8397]	-73,3139* [44,4157]	-75,7748* [45,0715]
2016	-95,4816** [47,0336]	-86,4975* [45,3309]	-99,0067** [47,8208]
2017	-105,9118* [56,5355]	-99,2092* [54,6016]	-110,1775* [57,5042]
Intercepto	2419,7812*** [12,9957]	2481,1466*** [8,8325]	2409,5634*** [14,1168]
Nº de observações	8.296.161	12.187.204	7.650.205
Nº de municípios	180	270	155
Coefficiente de determinação (R2) ajustado	0,886	0,891	0,887

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. Cada coluna da tabela indica a versão do grupo de comparação com a qual os atingidos são comparados. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 45 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 – Efeitos heterogêneos por sexo

Ano	Mulheres	Homens
2010	35,4287 [57,9632]	40,3863 [128,9470]
2011	6,6258 [43,5764]	25,0228 [139,6402]
2012	1,7023 [48,7018]	74,8273 [173,1796]
2013	-7,9097 [34,2177]	-67,4001 [52,6692]
2014 (ano-base)	0 [.]	0 [.]
2015	-35,0332 [25,4877]	-99,2095* [57,1791]
2016	-84,6815** [34,2993]	-102,2153* [60,0576]
2017	-113,3531** [46,1123]	-98,7292 [67,7295]
Intercepto	2122,7292*** [6,2715]	2635,6669*** [17,6634]
Nº de observações	3.493.815	4.802.346
Nº de municípios	180	180
Coeficiente de determinação (R2) ajustado	0,865	0,894

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 46 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 – Efeitos heterogêneos por raça/cor

Ano	Indígena	Branca	Preta	Amarela	Parda
2010	-85,674 [203,7338]	45,7859 [106,1807]	14,6794 [107,7392]	173,07 [116,1999]	35,0282 [119,5914]
2011	-131,2358 [120,4161]	23,6136 [107,0216]	5,7206 [114,4856]	58,119 [84,9792]	39,1905 [129,1691]
2012	-55,5122 [59,0504]	48,7745 [116,3959]	26,4977 [129,6675]	52,5201 [71,0481]	91,5446 [173,0449]
2013	-69,0442 [63,1983]	-31,3625 [41,9406]	-23,222 [53,2626]	-62,1141 [56,5495]	-38,3685 [46,9772]
2014 (ano-base)	0 [.]	0 [.]	0 [.]	0 [.]	0 [.]
2015	-23,9264 [41,3469]	-60,5221 [43,2342]	-67,1276 [47,7146]	-94,7264* [54,5496]	-83,6309* [46,9578]

Ano	Indígena	Branca	Preta	Amarela	Parda
2016	-15,0361 [37,7396]	-48,9766 [42,7306]	-62,4463 [47,5223]	-153,5646*** [56,8820]	-92,3639* [47,7643]
2017	83,9899 [55,4765]	-37,6421 [45,8386]	-74,2541 [53,5777]	-196,1787*** [53,1189]	-102,9077** [50,3372]
Intercepto	1700,6203*** [5,1096]	2556,6924*** [12,0295]	1978,9635*** [15,7411]	2594,1691*** [13,1070]	2155,1638*** [18,2853]
Nº de observações	35.374	3.194.044	475.903	50.185	2.687.797
Nº de municípios	113	180	180	146	180
Coefficiente de determinação (R2) ajustado	0,911	0,907	0,869	0,924	0,882

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 47 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 – Efeitos heterogêneos por nível de escolaridade

Ano	Fundamental incompleto	Fundamental completo	Médio completo	Superior completo
2010	15,9684 [31,3181]	45,0845 [49,8757]	41,0046 [126,3337]	12,3793 [175,5727]
2011	-8,0372 [25,6747]	23,3018 [46,2043]	19,727 [128,4971]	-2,2477 [160,1305]
2012	1,3743 [23,8273]	52,2635 [62,0265]	63,0797 [152,6426]	7,4606 [192,9303]
2013	-14,3671 [12,2460]	-15,7438 [12,4780]	-56,8633 [52,2740]	-69,837 [97,2057]
2014 (ano-base)	0 [.]	0 [.]	0 [.]	0 [.]
2015	-16,7999 [13,4744]	-26,0483 [22,1533]	-81,8254 [54,8007]	-157,8971** [77,3274]
2016	-21,8291 [20,4981]	-35,1441 [24,2274]	-96,4053* [52,9139]	-231,9902** [96,4902]
2017	-40,9220** [20,5978]	-44,3911 [30,7287]	-99,4814* [59,0017]	-248,8222* [127,8196]
Intercepto	1715,1144*** [3,6703]	1772,0892*** [5,6670]	2161,8769*** [16,5617]	4698,7946*** [22,9216]
Nº de observações	1.212.932	1.653.813	4.117.139	1.312.277
Nº de municípios	180	180	180	180
Coefficiente de determinação (R2) ajustado	0,830	0,830	0,849	0,878

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).

Tabela 48 — Estimativas de impacto médio sobre o salário de pessoas no grupo atingido que possuem vínculo ativo no mesmo setor de atividade econômica de 2014 – Efeitos heterogêneos para pessoas cujo emprego em 2014 era em setores censitários próximos aos rios

Ano	Até 1 km	Até 5 km	Até 10 km
2010	29,1776 [61,5052]	54,8754 [135,2790]	36,8336 [122,7327]
2011	-10,4799 [46,0681]	40,6658 [143,3644]	18,793 [127,5965]
2012	-34,2021 [51,0439]	69,1169 [169,9369]	40,4892 [143,9405]
2013	-23,5028 [33,3441]	-59,1841 [71,4795]	-57,8161 [55,6762]
2014 (ano-base)	0 [.]	0 [.]	0 [.]
2015	-62,3322* [35,3160]	-104,5417 [64,9825]	-96,2573* [53,4020]
2016	-118,5568** [51,1603]	-137,0347** [65,4521]	-127,8632** [52,5892]
2017	-129,5815* [71,4562]	-143,6491* [73,8692]	-137,1775** [62,5648]
Intercepto	2411,1127*** [2,7006]	2442,8794*** [12,2534]	2433,5845*** [13,0216]
Nº de observações	6.711.317	7.320.374	7.715.671
Nº de municípios	177	179	180
Coeficiente de determinação (R2) ajustado	0,885	0,887	0,886

Fonte: Elaboração própria (2020) a partir dos dados da RAIS (ME).

Observação: Coeficientes no gráfico correspondem a estimativas da Equação 2, a qual inclui efeitos fixos de ano e indivíduo. O grupo de comparação utilizado é o GC 2. São incluídas no modelo apenas pessoas que integram o subgrupo identificado no cabeçalho das colunas. Erros-padrão (entre colchetes) *clusterizados* no município em que a pessoa trabalhava em 31 de dezembro de 2014. Asteriscos ao lado das estimativas indicam valor-p igual ou inferior a 0,1 (*), 0,05 (**) e 0,01 (***).