



REVISTA
CADERNOS DE
FINANÇAS
PÚBLICAS

Análise Contrafactual dos Impactos Covid-19 na Arrecadação de Entes Subnacionais e Políticas de Isolamento Social e Lockdown.

André Nunes Maranhão

FGV/Banco do Brasil

RESUMO

O objetivo deste estudo é estimar os impactos contrafactuais da pandemia do Covid-19 na arrecadação estadual e taxas de mortalidade das UF do Brasil, considerando a adoção de políticas de lockdown e de isolamento social. Os modelos DiD, indicam que consumo diário de energia elétrica esteve inversamente associado as taxas de mortalidade estaduais. Os resultados ArCo evidenciaram que quanto maior a severidade pandêmica maiores foram os impactos em arrecadação, e quanto maior a taxa de isolamento social menores foram as taxas de mortalidade contrafactuais. Não há evidências significantes da política de lockdown sobre arrecadação ou taxa de mortalidade estadual.

Palavras-chave: Covid-19, Arrecadação, Taxa de Mortalidade, Inferência Contrafactual.

Classificação JEL: I18, C32, H27

SUMÁRIO

Resumo	2
1 Introdução	4
2 Literatura	5
2.1 Literatura sobre Impactos Covid-19 na Política Fiscal.....	6
3 Metodologia	7
3.1 Introdução a Inferência Causal e os Modelos de Dados em Painel de Diferenças-em-Diferenças.....	7
3.2 Modelos de Dados em Painel de Contrafactuais Artificiais de Alta Dimensão.....	9
3.2.1 Estimando Tempo de Intervenção Desconhecido em Modelos ArCo	11
4. Resultados	12
4.1 Descrição, Tratamento e Estatísticas Descritiva dos Dados.....	12
4.2 Resultados dos Modelos de Dados em Painel de Diferenças-em-Diferenças.....	25
4.2.1 Resultados dos Modelos DiD para Arrecadação Estadual	26
4.2.2 Resultados dos Modelos DiD para Taxa de Mortalidade	32
4.2.3 Conclusões Contrafactuais dos Modelos Dif-em-Dif	38
4.3 Resultados dos Modelos de Dados em Painel Contrafactuais Artificiais.....	40
5. Conclusão	51
Referências	54
6. Apêndice	59

1 Introdução

A pandemia do Covid-19 trouxe ao mundo, ao Brasil e as Unidades Federativas novos comportamentos, novas políticas, novos temas de pesquisa e alterou de maneira abrupta as escolhas dos agentes diante de um novo cenário com risco sanitário adicional. A velocidade de propagação de casos combinado a aceleração de letalidade na sequência gerou impacto de políticas públicas que foram desde restrições completas de movimentação, a políticas de socorro sócio-econômico. Os desdobramentos dessas políticas e/ou da evolução pandêmica geraram efeitos em cadeia nos mercados e na economia como um todo. Os entes subnacionais sofreram diretamente esse choque exógeno, os impactos se traduziram como uma drástica elevação da incerteza e, por conseguinte, uma elevação na percepção individual de risco dos agentes. A queda generalizada de arrecadação seria reflexo desse comportamento, fazendo com que governos subnacionais adotassem diferentes políticas no combate as elevações das taxas de mortalidade pandêmica. Dentro da gama de políticas adotadas, duas em particular foram mais difundidas: a recomendação ao isolamento social, deixando a critério individual a responsabilidade, benefícios e custos de adesão a essa política; a adoção de decretos legais de restrição de movimentos, ou como ficou conhecida, política de lockdown.

Várias importantes questões surgem sobre os temas relativos à arrecadação e taxas de mortalidade estaduais. Estudos podem, a depender da complexidade, ser direcionados para análise da trajetória dos fatos como ocorreram e podem chegar a importantes conclusões, contudo, em se tratando de um evento inesperado, abrupto e repentino, como foi o evento pandêmico do Covid-19, acabou por torná-lo um evento quase-experimental, fazendo com que as questões mais relevantes surgam dos questionamentos do que teria ocorrido em um contexto contrafactual, ou seja, o que teria ocorrido com um mesmo ente subnacional, em termos de arrecadação e taxas de mortalidade, caso não tivesse a intervenção pandêmica, ou se caso não tivesse adotado uma política de lockdown na intervenção pandêmica, ou se não houvesse uma grande adesão ao isolamento social. As respostas dessas importantes questões nos conduzem a inferência contrafactual (ou causal), e nesse contexto foram construídos dois painéis: painel de arrecadação, contendo dados mensais des-sazonalizados de arrecadação e atividade comercial de janeiro de 2000 a junho de 2020 perfazendo um total de 246 observações temporais por UF, e um painel pandêmico, contendo dados diários das taxas de mortalidade por 100 mil/hab. e consumo diário de energia elétrica do dia 22/02/2020 a 14/08/2020, perfazendo um total de 172 observações temporais por UF. Foi agregada a cada um desses painéis informações estaduais, no formato de variáveis dummy, para construção dos efeitos contrafactuais a serem estimados: PIBs acima da mediana nacional, alta densidade demográfica, adoção a política de lockdown¹ e alta adesão ao isolamento social (para o painel de arrecadação ainda foi considerado indicadora de gravidade pandêmica).

O primeiro passo foi apresentar os detalhes de construção das variáveis de estudo, bem como um detalhamento exploratório dos dados. Duas abordagens de inferência contrafactual temporal são colocadas a prova para estudar os dois painéis dados com os estados brasileiros: modelos de painel de Diferenças-em-Diferenças (DiD) e os recentes modelos de painel contrafactuais artificiais de al-

1 Apesar de estritamente não terem o mesmo significado, utilizaremos sem perda de generalidade, lockdown como sinônimo de lockout.

ta-dimensão (ArCo). Os resultados dessas duas abordagens convergem ao analisar as políticas de isolamento social e lockdown, as evidências encontradas indicam que alta adesão ao isolamento social teve impacto negativo contrafactual, estatisticamente significativo na arrecadação, enquanto que uma maior taxa de isolamento social está associado a uma menor taxa de mortalidade pandêmica nos estados. A adoção de política de lockdown não apresentou efeitos estatisticamente significantes nem na arrecadação e nem nas taxas de mortalidade, em nenhuma abordagem estimada.

O artigo apresenta na seção 2 uma breve revisão da literatura sobre diferentes impactos da pandemia Covid-19 em diferentes áreas, inclusive no tocante a questão fiscal de arrecadação. Na seção 3 detalhamos todos os modelos utilizados no estudo, sendo que a descrição dos dados e resultados são apresentados na seção 4, as conclusões do estudo na seção 5 e uma seção final dedicada a análise das implicações em políticas públicas dos resultados do estudo e da aplicação dos métodos aqui propostos.

2 Literatura

Nenhum outro evento datado teve o alcance, o impacto, a magnitude de forma tão súbita como a pandemia do Covid-19. A pandemia do Covid-19 marcou a história mundial recente não apenas pela alta velocidade de propagação e intensidade de letalidade, mas também pelos impactos nas mais diferentes áreas de conhecimento. Assim que seus impactos tomaram proporções mundiais, pesquisadores de diferentes áreas da ciência aceleraram esforços em identificar os diferentes efeitos desse novo fenômeno em seus respectivos nichos de pesquisa.

Os números de casos, de óbitos por Covid-19 e recuperados passou a ter seus números atualizados e acompanhados diariamente na maior parte dos países² e publicações surgiram já estimando seus impactos em mortes futuras, como no estudo Roberton et al. (2020), que apresenta estimativas iniciais dos efeitos indiretos da pandemia Covid-19 na mortalidade materna e infantil em países de baixa e média renda. Sumner et al. (2020) estimaram os efeitos sobre a pobreza mundial e chegaram a preocupante conclusão de que, *ceteris paribus*, assumindo uma contração de 5 por cento na renda per capita, o mundo poderia testemunhar um aumento potencial no número de pessoas pobres, em relação aos números de 2018, de mais de 80 milhões para a linha de pobreza de US \$ 1,9/dia, de mais de 130 milhões para o US \$ 3,2/dia padrão, e de quase 124 milhões para a linha superior de US \$ 5,5/dia.

Os efeitos econômicos estudados não se limitaram a temática futura da pobreza, mas também aos efeitos correntes sobre o desemprego, como no estudo de Fairlie et al. (2020), que mostra como a pandemia impactou o desemprego das minorias usando microdados até abril de 2020. Afro-americanos experimentaram um aumento no desemprego para 16,6 por cento, menos do que o antecipado com base em recessões anteriores. Outros estudos como Coates et al. (2020) e Kurmann et al. (2020), mostram outros diferentes impactos sobre o desemprego. As políticas de lockdown também foram alvo de pesquisas, como o estudo de Ozili e Arun (2020), mostravam que o aumento do número de dias de bloqueio, as decisões de política monetária e as restrições às viagens internacionais afetaram

² A universidade JHU criou um painel de informações da pandemia para um grande grupo de países disponível no site: <https://coronavirus.jhu.edu>

severamente o nível das atividades econômicas e o fechamento, a abertura, o menor e o maior preço das ações dos principais índices do mercado de ações.

Apesar da maioria dos impactos da pandemia do Covid-19 terem caráter negativos, alguns estudos relatam benefícios ecológicos e climáticos que a abrupta desaceleração social causou, como por exemplo Wang e Su (2020), Mandaland Pal (2020), Norouzi et al. (2020) e Le Quéré et al. (2020), apenas para citar alguns exemplos. Os impactos nos mercados financeiros, segundo Baker et al. (2020) e Fernandes (2020), foram sem precedentes. Baker et al. (2020) mostram que nenhum surto de doença infecciosa anterior, incluindo a gripe espanhola, impactou a bolsa de valores norte americana tão vigorosamente quanto a pandemia Covid-19. Na verdade, as pandemias anteriores deixaram apenas leves traços no Mercado de ações dos EUA. As evidências desse estudo sugere que as restrições dos governos à atividade comercial e distanciamento social voluntário, operaram como poderosos efeitos em uma economia mundial orientada para serviços, essas seriam as principais razões pelas quais o mercado de ações dos EUA reagiu com tamanha força para Covid-19 do que para pandemias anteriores.

O estudo de De Vito e Gomez (2020) investigaram como a crise de saúde Covid-19 pode afetar a liquidez de empresas listadas em 26 países. Os resultados indicaram que os empréstimos-ponte são mais econômicos para evitar uma crise de caixa massiva. Os estudos de Garcin et al.(2020) utilizaram uma abordagem não paramétrica para estimar as variações no tempo das densidade de retornos de preços diários de vários índices de ações e, usando várias estatísticas de divergência, puderam descrever a cronologia da crise, bem como as disparidades regionais. Por exemplo, um impacto mais limitado do Covid-19 nos mercados financeiros da China, um forte impacto nos EUA e uma lenta recuperação na Europa. Zhang et al. (2020) mostram que a rápida disseminação do coronavírus (Covid-19) teve impactos dramáticos na volatilidade e nos níveis de risco nos mercados financeiros em todo o mundo, fazendo com que os investidores sofressem perdas significativas em um período muito curto de tempo. Al-Awadhi et al. (2020) mostraram que tanto o crescimento diário no total de casos confirmados quanto no total de casos de morte causados por Covid-19 têm efeitos negativos significativos no retorno das ações em todas as empresas nas bolsas chinesas. O estudo de Topcu and Gulal (2020) estimando os impactos do Covid-19 nos mercados de ações de países emergentes, revelam que o impacto negativo da pandemia nos mercados de ações emergentes caiu gradualmente e começou a diminuir em meados de abril. Em termos de classificação regional, o impacto do surto foi maior nos mercados emergentes asiáticos, enquanto os mercados emergentes na Europa tiveram menor intensidade.

2.1 Literatura sobre Impactos Covid-19 na Política Fiscal

Durante as crises econômicas, os orçamentos dos governos estaduais e locais ficam sob pressão. O estresse surge devido a contrações nas receitas e arrecadações e aumentos nas necessidades de despesas. O mesmo é amplamente verdadeiro para os governos federal, estadual e local, um recurso para essa situação ocorre quando o governo federal tem poderes de emitir dívida de curto e longo prazo. Governos estaduais e locais, em contraste, têm restrições de vários graus de severidade em sua autoridade legal para emitir dívida em resposta a gastos inesperados ou choques de receita (Poterba,1994; Clemens e Miran 2012).

Comparamos o provável impacto da pandemia nas receitas dos governos estaduais com as mais comuns contrações econômicas. Ao contrário das contrações típicas, durante as quais a renda diminui mais drasticamente do que o consumo (Canova, 1998) os lockdowns da Covid-19 geraram quedas excepcionalmente grandes em consumo em relação à renda. Isso é em grande parte porque as receitas foram impulsionadas por estímulos, alguns dos quais são tributáveis e outros não. Além disso, a Covid-19 resultou em um declínio dramático nas despesas de consumo pessoal com saúde, restaurantes e alojamentos. No curto prazo as tensões de receita foram, portanto, particularmente graves em estados que dependem em um grau significativo de impostos sobre vendas, e, em particular, nas vendas em setores expostos. Por outro lado, como na maioria das recessões, é improvável que as bases fiscais incidentes sobre as propriedades contraem significativamente durante a própria recessão, porque os valores das propriedades são normalmente reavaliados com atrasos substanciais, conforme Lutz, Molloy, and Shan (2011). No âmbito subnacional brasileiro os estados perderam participação relativa na distribuição das disponibilidades tributárias e os municípios afirmaram a posição obtida no ordenamento constitucional de 1988, conforme Tristao (2003).

Os estudos com a resposta dos governos para amenizar os efeitos perversos também começaram a surgir, como por exemplo o estudo de Casado et al. (2020), que fornece uma maneira de avaliar quase em tempo real medidas fiscais para estímulo atividade econômica local. Por outro lado impactos nas finanças públicas também foi tema de estudo de Clemens e Veuger (2020), que estimaram uma queda de arrecadação de US\$ 106 bilhões projetados para o exercício de 2021. Mesmo os impactos diários na economia foram estimados utilizando dados de consumo de energia elétrica para Itália, como mostra Fezzi e Fanghella (2020).

O objetivo dessa breve revisão não é esgotar todos os inúmeros estudos que tratam dos impactos da pandemia do Covid-19, mas trazer exemplos de diferentes áreas de pesquisa, até a temática de arrecadação estadual que será tratado como o tema principal deste estudo.

3 Metodologia

3.1 Introdução a Inferência Causal e os Modelos de Dados em Painel de Diferenças-em-Diferenças

O estudo experimental aleatorizado, desde surgimento dos testes estatísticos de hipóteses, se consolidou como um dos principais pilares do desenvolvimento científico. No ambiente clínico, a dosimetria controlada, e a identificação de grupos de controle e de tratamento, permitiu identificar a eficácia de medicamentos ou de tratamentos. No contexto ciências sociais, a reprodução desse ferramental metodológico tem se mostrado um grande desafio. Desde o trabalho pioneiro de John Snow (1855), que buscou identificar as causas da epidemia de cólera em Londres, até a primeira versão do modelo de diferenças-em-diferenças apresentado por Ashenfelter e Card (1985) foram mais de cem anos. Diante da impossibilidade de haver um experimento completamente randomizado, a estatística evoluiu para métodos que permitisse uma pseudo randomização por meio de fenômenos quase-experimentais. A ideia era estimar efeitos contrafactuais a partir da reconstrução de uma pseudo-unidade de tratamento que permitisse obter uma estimativa dos efeitos de uma intervenção.

Dessa forma, considere a notação apresentada por Carvalho et al. (2018);

- 1) Unidades: indexados por $i=1, \dots, n$, que pode ser indivíduos, estados, países, etc. O tratamento ocorre em uma dessas unidades e não afeta as outras;
- 2) Variáveis: $z_{it} = (z_{it}^1, \dots, z_{it}^{q_i})$ sendo $q_i > 1$ a quantidade total de variáveis relacionadas as unidades de tratamento;
- 3) Intervenção: A intervenção ocorreu apenas na unidade tratada no momento $T_0 = \lambda_0 T$ com $\lambda_0 \in (0, 1)$.

Suponha, sem perda de generalidade, que a unidade tratada é a primeira ($i=1$). Considere ainda que $z_{1t}^{(0)}$ e $z_{1t}^{(1)}$ sejam os resultados da primeira unidade com tratamento e sem tratamento, respectivamente. Normalmente, não observamos os dois resultados simultaneamente. Em vez disso, observamos:

$$z_{1t} = D_t z_{1t}^{(0)} + (1 + D_t) z_{1t}^{(1)} \quad (5)$$

Onde D_t assume valor 1 se a unidade estiver em tratamento no momento t e 0 caso contrário. O objetivo é testar a hipótese de os efeitos da intervenção serem estatisticamente significativos para $t > T_0$. As intervenções são consideradas no seguinte formato:

$$y_t^{(1)} = \begin{cases} y_t^{(0)}, & t = 1, \dots, T_0 \\ y_t^{(0)} + \delta_t, & t = T_0, \dots, T \end{cases} \quad (6)$$

Onde y_t representa a variável resposta de interesse. Considerando um contexto de dados em painel, o efeito δ_t pode ser estimado a partir dos modelos de efeito fixos, conforme Angrist (2008):

$$y_{it} = \alpha_i + \lambda_t + \theta D_{int} + \phi D_{trat} + \delta(D_{int} \times D_{trat}) + z'_{it} \beta + \varepsilon_{it} \quad (7)$$

Onde, D_{int} representa a dummy de intervenção e D_{trat} a dummy de tratamento. Os efeitos contrafactuais de Diferenças-em-Diferenças (DiD) são obtidos pelo parâmetro δ . Para Card and Krueger (1994) a principal vantagem do método DiD é a de que ele é capaz de controlar para as influências das variáveis de resultado para características não-observáveis que sejam fixas no tempo, essa é uma vantagem importante, porque caso essas características não-observáveis influenciem sobre a decisão de participar do tratamento, ocorrerá o que chamamos de viés de seleção. Ou seja, a ferramenta DiD é capaz de levar em conta a associação entre a variável de resultado, a participação no tratamento e as características não observáveis dos indivíduos que sejam invariantes no tempo, contornando o viés de seleção. As principais hipóteses desse modelo são:

- 1) O modelo de regressão foi corretamente especificado, com inclusão de outras variáveis independentes de controle;
- 2) O termo de erro é em média igual a zero: $E(\varepsilon_{it}) = 0$;
- 3) O termo de erro não é correlacionado com outras variáveis na equação. Esse pressuposto,

- conhecido como pressuposto de tendência paralela;
- 4) As variáveis omitidas são constantes no tempo;
 - 5) Não dependência cross-section;
 - 6) Não há endogeneidade presente no modelo;

Várias dessas hipóteses foram contestadas por Bertrand et al. (2004), em particular o pressuposto-chave de tendências comuns que é chamado de parallel trend assumption. Esse pressuposto não possui teste formal na literatura e dessa forma a validade de suas inferências contrafactuais depende diretamente da crença em sua validade nos modelos estimados.

3.2 Modelos de Dados em Painel de Contrafactuais Artificiais de Alta Dimensão

Uma alternativa recente para enfrentar as hipóteses restritivas dos modelos DiD, foi o surgimento dos modelos contrafactuais sintéticos ou artificiais. O método de controle sintético (SCM) proposto por Abadie et al. (2010), Abadie et al. (2015), Abadie and Gardeazabal (2003) é uma abordagem popular para estimar o impacto de um tratamento em uma única unidade em ambientes com um número grande de unidades de controle e com muitos pré-tratados com resultados para todas as unidades. A ideia é construir uma média ponderada das unidades de controle, conhecido como controle sintético, que corresponda a uma unidade tratada mais estatisticamente próximo de uma unidade pré-tratada. O impacto estimado é então a diferença nos resultados pós-tratamento entre a unidade tratada e o controle sintético. Uma limitação importante desta abordagem é que, enquanto SCM minimiza o desequilíbrio nos resultados de pré-tratamento, geralmente falha em atingir o equilíbrio exato entre tratados e não tratados pós-intervenção devido à maldição da dimensionalidade, o desequilíbrio resultante pode levar a um grave viés.

A primeira solução proposta para eliminação desse viés foi proposto por Hsiao, Ching, and Wan (2012), que considera um método de duas etapas em que, em sua primeira etapa, o contrafactual para uma única variável de interesse tratada é construído como uma combinação linear de um conjunto de baixa dimensão de covariáveis observadas a partir de elementos pré-selecionados de um grupo de pares. O modelo é estimado por mínimos quadrados ordinários usando dados do período pré-intervenção. Seus resultados teóricos foram derivados sob a hipótese de especificação correta de um modelo linear de dados em painel com fatores comuns e sem covariáveis. A seleção dos pares incluídos na combinação linear é feita por critérios de informação.

O método proposto por Carvalho et al. (2018), de painéis contrafactuais artificiais em alta dimensão (ArCo), generaliza o trabalho de Hsiao, Ching, and Wan (2012) em direções importantes. Primeiro, considerando estimativa LASSO na primeira etapa, permitimos que um grande número de covariáveis/pares sejam incluídos, não exigindo qualquer seleção de pré-estimativa que pode enviesar as estimativas. Além disso, estimativa LASSO é bastante atraente quando o tamanho da amostra é pequeno em comparação com o número de parâmetros a serem estimados. Em segundo lugar os resultados teóricos são derivados sem suposições estritas sobre o processo gerador da série, que se assume ser desconhecido. Outra grande vantagem desse método é não precisar estimar a verdadeira expectativa condicional. Este é um bom recurso da metodologia ArCo, já que normalmente os modelos são mal

especificados. Terceiro, não há restrição de balanceamento entre tratados e não tratados, ou seja, não há restrição a análise a uma única variável tratada. Ainda é possível, por exemplo, medir o impacto das intervenções em diversas variáveis da unidade tratada simultaneamente. Os modelos ArCo também permitem testes em vários momentos da variável de interesse. Quarto, a metodologia ArCo pode ser aplicada quando o tempo de intervenção é desconhecido, permitindo que seja estimado endogeneamente o tempo de intervenção. Considere as hipóteses até a equação (6).

A metodologia ArCo propõe, $y_t^{(j)} = h(z_{1t}^j) \in R^q, j \in \{0,1\}$, $h(\cdot)$ sendo uma função mensurável de z_{1t} e $\{\delta_t\}_{t=T_0}^T$ sendo uma sequência determinística. A função $h(\cdot)$ é uma função geral que permite intervenções na média, na variância, na covariância e etc. O método ArCo concentra-se na hipótese:

$$H_0: \Delta_T = \frac{1}{T-T_0+1} \sum_{t=T_0}^T \delta_t = 0 \tag{8}$$

Onde Δ_T representa o efeito médio do tratamento durante o período de tratamento. Não observamos $y_t^{(0)}$ para $t \geq T_0$, esse será o valor contrafactual, ou seja, qual seria o y_t na ausência da intervenção. Na primeira etapa do método ArCo, seja $z_{0t} = (z_{2t}', \dots, z_{nt}')'$ e $Z_{0t} = (Z_{2t}', \dots, Z_{nt}')'$ uma coleta de todas as unidades não tratadas até uma defasagem arbitrária $p \geq 0$. A dimensão de Z_{0t} depende do número de unidades, número de defasagens e número de variáveis por unidade. Considere agora o seguinte modelo para $y_t^{(0)}$:

$$y_t^{(0)} = \mathcal{M}_t + v_t \tag{9}$$

Com as únicas hipóteses de $E(v_t) = 0$ e M_t sendo um mapeamento mensurável, não necessariamente uma função explícita, ou seja, podendo assumir qualquer forma funcional, como por exemplo, árvore de decisão, florestas randômicas, funções não lineares, modelos de dados em painel, funções temporais com diversas estruturas de dependência, etc. O primeiro passo do modelo ArCo é estimar a equação (9) utilizando as $T_0 - 1$ observações, haja vista que para $t < T_0$ temos $y_t = y_t^{(0)}$. Com o conjunto $T_1 = T_0 - 1$ e $T_2 = T - T_0 + 1$, é possível estimar $\hat{M}_{t,T_1} = \hat{M}_{t,T_1}(Z_{0t})$ e utilizá-lo para construção do contrafactual:

$$y_t^{(0)} = \begin{cases} y_t^{(0)}, & t = 1, \dots, T_0 \\ \hat{\mathcal{M}}_{t,T_1}, & t = T_0, \dots, T \end{cases} \tag{10}$$

Finalmente o estimador ArCo é definido como:

$$\hat{\Delta}_T = \frac{1}{T - T_0 + 1} \sum_{t=T_0}^T \hat{\delta}_t \tag{11}$$

Onde $\hat{\Delta}_T = y_t - y_t^{(0)}$.

A metodologia ArCo é uma abordagem em dois estágios de estimação, sendo o primeiro estágio a estimação de M_t na amostra pré-intervenção e na segunda etapa estimamos $\hat{\Delta}_T$ que é o impacto médio da intervenção.

Conforme destaca Carvalho et al. (2018), quando comparados aos estimadores DiD, as vantagens da metodologia ArCo são três. Primeiro não precisamos que o número de unidades tratadas cresça. Na verdade, a situação laboriosa é quando há uma única unidade tratada. A segunda e mais importante diferença é que a metodologia ArCo foi desenvolvida para situações em que as $n-1$ unidades não tratadas diferem substancialmente do tratado e não pode formar um grupo de controle mesmo após o condicionamento em um conjunto de observáveis. Por fim, a metodologia ArCo funciona mesmo sem a hipótese de tendências paralelas.

Embora, os métodos ArCo e SCM construam um contrafactual em função de variáveis observadas de um grupo de pares, as duas abordagens têm diferenças importantes. Apenas elencando os pontos principais listados por Carvalho et al. (2018), primeiro, o método SCM depende de uma combinação convexa de pares para construir o contrafactual que, como apontado por Ferman e Pinto (2016), gera viés no estimador. Isso é claramente evidenciado em experimentos de simulação. A solução ArCo é uma função geral, possivelmente não linear. Mesmo no caso de linearidade, o método não impõe nenhuma restrição aos parâmetros. Por exemplo, a restrição de que os pesos nos métodos SCM são todos positivos também é uma forte restrição. Além disso, os pesos no método SCM são geralmente estimados usando médias de tempo das variáveis observadas para cada par. Portanto, toda a dinâmica da série temporal é removida e os pesos são determinados em um cenário de corte transversal puro. Em algumas aplicações do SCM método, o número de observações para estimar os pesos é muito menor do que o número de parâmetros a serem determinados. Por exemplo, Abadie and Gardeazabal (2003), os autores tem 13 observações para estimar 16 parâmetros.

Além disso, como destacado em Carvalho et al. (2018), o método SCM foi projetado para avaliar os efeitos da intervenção em uma única variável. Para avaliar os efeitos em um vetor de variáveis, o método deve ser aplicado várias vezes. A metodologia ArCo pode ser aplicada diretamente a um vetor de variáveis de interesse. Além disso, não há procedimento inferencial formal para teste de hipótese no método SCM, enquanto na metodologia ArCo, um método simples e uniforme teste válido e padrão pode ser aplicado. Uma última desvantagem do SCM, destacado em Ferman, Pinto, e Possebom (2016), o método SCM não oferece nenhum guia sobre a seleção de variáveis utilizadas para construção dos pesos que serão utilizados na criação dos valores contrafactuais.

Para fins de estimação utilizaremos os modelos lineares generalizados com penalização complexa, que combinam diferentes tipos de métodos de regularização l_1 (LASSO), l_2 (Regressão Ridge) e a combinação dessas duas (o elastic net), com algoritmo de estimação cíclico de coordenada descente, conforme apresentado por Friedman et al (2010). Essa formulação garante estimar diferentes estruturas temporais da dinâmica das variáveis e unidades dos painéis utilizados para o estudo.

3.2.1 Estimando Tempo de Intervenção Desconhecido em Modelos ArCo

Em muitos casos, é razoável supor que o tempo de intervenção é desconhecido. Por exemplo, embora alguma nova política foi iniciada em um momento conhecido, seus efeitos podem ter sido antecipados devido a expectativas. Ou ainda, a diferença entre os valores estimados nos modelos pré-intervenção ainda não se configuram estatisticamente diferentes, de nenhum momento observado até certo ponto pós-intervenção. Nesses casos, independentemente da fonte de incerteza, é necessário

estimar λ_0 ajustando o Estimador ArCo para ser uma função de λ , $\Delta_T(\lambda)$, conforme destaca Carvalho et al. (2018).

Basicamente, determina-se $\Lambda = (\underline{\lambda}, \bar{\lambda})$ como um limite para λ_0 para evitar viés de amostra finita perto dos limites e definimos $\|\cdot\|_p$ como a norma l_p , então o tempo de intervenção será estimado como:

$$\hat{\lambda}_{0,p} = \arg \max_{\lambda \in \Lambda} \|\hat{\Delta}_T(\lambda)\|_p$$

Dessa forma, $\hat{\lambda}_{0,p}$ será aquele valor que maximiza a norma l_p do estimador ArCo.

4. Resultados

4.1 Descrição, Tratamento e Estatísticas Descritiva dos Dados

A construção da base de dados utilizada nesse estudo considerou a possibilidade de construção de dois painéis balanceados para os estados brasileiros considerando a dinâmica mensal de arrecadação no primeiro, e a dinâmica pandêmica diária no segundo. Essa formulação permitiu a implementação dos modelos de painel Dif-em-Dif, além da estimação dos modelos contrafactuais artificiais de dados em painel. A construção desses painéis considerou, portanto a disponibilidade em diferentes fontes de informações de dados, de informações estaduais.

Tabela 1 - Descrição dos Dados

Painel de Estudo de Arrecadação Estadual		
Variável	Descrição	Fonte
Arrecadação	Dados mensais de arrecadação estadual	CONFAZ - Conselho Nacional de Política Fazendária
Atividade Comercial	Índice de volume estadual de vendas no comércio varejista	IBGE - Tabela 3416
Dados Estaduais	Dados de população e densidade demográfica	IBGE - Tabela Cidades e Estados
Painel de Estudo Pandêmico Estadual		
Variável	Descrição	Fonte
Mortalidade Estadual	Dados diários de óbitos por 100 mil habitantes	Dados ARPEN – SRAG Associação de Cartórios Brasileiros
Consumo de Energia Elétrica	Consumo diário de energia elétrica estadual	Câmara de Comércio de Energia Elétrica

Fonte: Elaboração própria.

A Tabela 1 descreve o conjunto de informações disponíveis para aberturas em nível dos entes subnacionais. Os dados mensais consideraram o início da série em janeiro de 2000 ao último dado recente disponibilizado pelo CONFAZ referente ao mês de junho de 2020, perfazendo um total de 246 observações temporais por Unidade Federativa. O índice de volume estadual de vendas no comércio

varejista (PMC), foi utilizado na construção do painel de estudo da Arrecadação como proxy de atividade econômica em nível estadual, presente para todos estados brasileiros. Na construção do painel pandêmico estadual o número de óbitos foi ajustado para representar uma taxa por 100 mil habitantes e dessa forma permitir estimação livre dos efeitos de escala populacional. Os dados diários de consumo de energia elétrica foram utilizados como uma proxy do nível de atividade econômica diária em nível estadual. Esses dados diários contemplam o período que começa no dia 22 de fevereiro de 2020 ao dia 14 de agosto, com uma total de 172 observações temporais por UF.

Os dados de arrecadação estadual foram submetidos a ajuste sazonal utilizando X-13ARIMA-SEATS (GÓMEZ; MARAVALL, 1996) considerando as especificações de ajustes sazonal da Pesquisa Mensal do Comércio (PMC). Nessa especificação são considerados todos os efeitos de calendário e também os feriados móveis, como Carnaval, Páscoa e Corpus Christi, além da hipótese de sazonalidade estocástica. Também foram utilizados os valores com ajuste sazonal da própria PMC.

Com os painéis construídos para os estudos contrafactuais pandêmicos e de arrecadação foram construídas as variáveis de estimação dos efeitos Dif-em-Dif, conforme a Tabela 2:

Tabela 2 - Construção dos Efeitos Dif-em-Dif

UF	I_iso	I_grave	I_Lockout	I_PIB	I_dens
RO	1	0	0	0	1
AC	1	1	0	0	0
AM	1	1	1	0	1
RR	0	1	0	0	1
PA	1	1	1	1	0
AP	1	1	1	0	0
TO	0	0	1	0	0
MA	1	0	1	0	0
PI	1	0	1	0	0
CE	1	1	1	1	1
RN	0	1	1	0	1
PB	0	0	0	0	1
PE	1	1	1	1	1
AL	1	1	0	0	1
SE	0	1	0	0	1
BA	0	0	0	1	0
MG	0	0	0	1	0
ES	0	1	0	0	1
RJ	1	1	1	1	1
SP	0	1	0	1	1
PR	0	0	1	1	0

SC	0	0	0	1	1
RS	1	0	0	1	1
MS	0	0	1	0	0
MT	0	0	0	1	0
GO	0	0	0	1	0
DF	1	0	0	1	1
Total (1):	13	13	12	13	15
Total (0):	14	14	15	14	12

Nota: As variáveis I_PIB e I_dens foram construídas utilizando os dados IBGE - Tabela Cidades e Estados:

I_PIB: indicadora de PIB nominal do Estado acima da mediana brasileira;

I_dens: indicadora de densidade demográfica acima da mediana brasileira.

A variável I_grave foi construída considerando o resultados do número de óbitos por 100 mil habitantes no dia 14/08/2020, observado no dia 31/08/2020 após 15 dias das revisões de suas vintages:

I_grave: indicadora de taxa de óbitos por 100 habitantes acima da mediana brasileira.

A variável I_lockout foi construída a partir das publicações oficiais dos governos estaduais ou municipais decretando restrições explícitas de movimentação pública, fechamento de estabelecimentos e áreas em pelo menos uma cidade no Estado:

I_lockout: indicadora de Lockdown na UF.

A variável I_iso foi construída a partir do índice de Isolamento Social desenvolvido pela Inloco, utilizando tecnologia que mapeia os deslocamentos o via geolocalização de 60 M de brasileiros.

O índice representa o percentual da população que está respeitando a recomendação de isolamento:

I_iso: indicadora de alta adesão ao isolamento social, ou seja, UF com índice acima da mediana brasileira, considerando o percentil 90 do índice de abril a agosto por UF. O detalhamento dessa variável está disponibilizada na Tabela 2 do Apêndice.

Fonte: Elaboração própria.

A partir da Tabela 2 podemos observar que treze estados apresentaram alta adesão às recomendações de isolamento social (RO, AC, AM, PA, AP, MA, PI, CE, PE, AL, RJ, RS e DF), outros treze estados brasileiros foram classificados como grave condição pandêmica (AC, AM, RR, PA, AP, CE, RN, PE, AL, SE, ES, RJ e SP). Não obstante da gravidade pandêmica, alguns estados apresentaram baixa adesão ao isolamento social, como são os casos de RR, RN, SE, ES e SP. Em doze estados foram verificadas políticas extremas de lockdown (AM, PA, AP, TO, MA, PI, CE, RN, PE, RJ, PR e MS), dentre estes, sete estados apresentaram grave condição pandêmica (AM, PA, AP, CE, RN, PE e RJ). Dentre os estados que utilizaram de políticas de lockdown, quatro estados apresentaram baixa adesão ao isolamento social (TO, RN, PR e MS), o que evidencia, ainda de maneira exploratória, políticas de lockdown podem não surtir efeitos sobre a pandemia regional.

Para cada painel de estudo (Arrecadação e Pandêmico), dado que individualmente tem dinâmicas temporais distintas, arrecadação com dados mensais e pandêmico com dados diários, a marcação de início da pandemia foi distinta para cada um desses estudos. A marcação do início da pan-

demia dos dados de arrecadação ocorre no mês de fevereiro de 2020 para todos os estados, enquanto para os dados pandêmicos diários, a marcação de início da pandemia regional ocorre a partir da data de registro do primeiro óbito por Covid-19.

A combinação dentro de cada painel de estudo, das dummies de início da pandemia com as variáveis apresentadas na Tabela 2, nos permite estimar os efeitos contrafactuais para cada uma dessas variáveis no contexto de modelo de dados em painel. Apresentamos na Tabela 3 os resultados das estatísticas descritivas considerando os períodos pré-pandêmicos e pandêmicos da arrecadação por UF, considerando na comparação, o mesmo período do ano de 2019.

Os resultados da Tabela 3 evidencia uma queda na arrecadação média para quase totalidade dos estados brasileiros. A arrecadação sofreu uma queda, em média, em termos percentuais³, de 5,9%, com uma mediana de 7,9%. Entre os 27 estados, cinco apresentaram aumento na média de arrecadação no período pandêmico, em termos percentuais: MT (12,12%), MS (8,63%), RR (1,68%), AM (1,57%) e PA (0,96%). Quando consideramos a magnitude da queda de arrecadação média, oito estados se destacam em termos percentuais: CE (20,12%), AP (12,24%), RJ (9,8%), SE (9,79%), BA (9,79%), AL (9,55%), RN (9,3%) e PE (9,22%). Verifica-se também o aumento da variabilidade da arrecadação mensal no período pandêmico em todos os estados brasileiros, indicando uma heterogeneidade dos impactos causados pela pandemia, e dessa forma um aumento significativo dos respectivos coeficientes de variação³, de 0,04 (4%) em média, para 0,31 (31%). O valor de mínima arrecadação observada no período pandêmico foi menor do que a do período de comparação para todas as UF, contudo o valor de máxima de arrecadação no período pandêmico, também foi maior para todas as UF quando comparado ao período pré-pandêmico, indicando um forte processo de recuperação mesmo durante a pandemia.

A Tabela 4 apresenta sub-tabelas contingenciais dos valores médios de arrecadação considerando o horizonte de um ano para os efeitos Dif-em-Dif a serem estimados na seção subsequente. Em todas sub-tabelas, foram considerados os períodos pré-pandêmicos e pandêmicos e os diferentes efeitos. Os resultados exploratórios indicam um valor médio menor de arrecadação para os estados que adotaram políticas de lockdown, contudo com um percentual de redução maior no período pandêmico, dos estados sem adoção de políticas de lockdown. Em termos de magnitude dessa redução, as UF sem adoção desse tipo de política, tem redução média de 12% contra 11% das UF com indicativo de Lockdown.

Tabela 3 - Estatísticas Descritivas da Arrecadação Estadual

³ Os valores percentuais e dos coeficientes de variação e foram suprimidos no corpo do documentos, mas encontraram-se disponíveis na Tabela 1 do Apêndice.

Período Pré-Pandêmico - Fev/19 a Jun/19				
UF	Média	Desvio-Padrão	Mínimo	Máximo
AC	118.718.836	5.158.994	112.183.348	126.379.247
AL	382.763.013	14.715.722	358.533.112	399.889.500
AM	893.765.124	30.258.090	837.213.145	924.164.388
AP	108.525.059	9.524.292	98.254.478	124.227.344
BA	316.898.516	13.159.724	304.762.542	342.029.640
CE	1.268.607.493	195.779.012	1.154.429.405	1.659.341.608
DF	831.397.940	33.375.465	781.465.587	875.027.455
ES	1.100.205.868	38.276.588	1.039.408.797	1.151.339.846
GO	1.618.987.955	30.746.079	1.566.144.244	1.661.465.411
MA	700.339.769	10.423.302	683.168.572	712.678.381
MG	5.065.035.909	207.330.852	4.902.924.089	5.461.236.594
MS	951.107.584	28.960.048	925.896.395	1.007.821.225
MT	1.289.483.003	49.180.447	1.193.876.485	1.333.980.338
PA	1.254.469.014	20.365.520	1.222.777.930	1.283.159.830
PB	531.278.479	16.459.643	520.221.155	563.908.915
PE	1.574.603.331	14.260.261	1.555.223.789	1.595.001.879
PI	403.835.095	58.528.223	370.369.931	520.563.259
PR	3.166.321.838	124.875.423	2.928.379.055	3.283.532.729
RJ	3.778.213.948	170.760.613	3.517.872.384	4.029.731.761
RN	511.513.210	7.536.897	498.232.612	519.102.381
RO	460.699.961	17.908.043	425.930.941	475.537.152
RR	112.417.924	10.905.604	102.417.512	130.392.358
RS	3.243.057.391	44.331.689	3.194.431.122	3.310.832.943
SC	2.259.346.778	30.049.282	2.227.686.105	2.310.512.344
SE	316.898.516	13.159.724	304.762.542	342.029.640
SP	14.285.223.260	113.509.250	14.126.824.600	14.480.172.500
TO	270.006.497	2.866.144	265.598.780	273.510.374

Período Pandêmico - Fev/20 a Jun/20				
UF	Média	Desvio-Padrão	Mínimo	Máximo
AC	109.629.719	35.244.266	2.422.956	236.679.573
AL	346.222.065	115.063.618	39.787.712	487.949.653
AM	907.835.174	253.161.189	107.403.471	1.147.625.532
AP	90.904.959	33.755.759	7.924.209	152.456.200
BA	285.872.278	94.104.732	29.079.511	466.062.373
CE	1.013.387.755	339.473.309	133.101.937	1.659.341.608
DF	790.269.364	229.330.160	111.820.637	927.703.651
ES	1.019.272.995	277.871.976	148.987.517	1.186.503.150
GO	1.527.685.647	521.912.294	159.901.431	3.801.654.704
MA	652.389.353	205.979.476	47.738.867	868.317.392
MG	4.638.169.079	1.409.095.708	627.408.247	5.911.600.362
MS	1.033.274.147	273.585.094	82.353.863	1.123.311.276
MT	1.445.877.552	396.832.788	114.724.064	1.963.284.757
PA	1.266.561.811	374.105.831	93.397.989	1.423.875.035
PB	488.810.145	154.569.576	54.564.453	604.026.441
PE	1.429.409.575	452.471.639	179.944.260	2.244.184.330
PI	373.560.448	128.779.839	32.887.344	577.171.411
PR	2.964.979.362	961.120.964	352.367.410	4.672.946.912
RJ	3.408.064.641	1.054.176.319	658.052.220	5.276.449.316
RN	463.894.723	147.337.786	52.648.467	549.003.521
RO	447.634.715	131.217.988	26.167.426	564.114.745
RR	114.307.254	30.555.323	6.786.233	148.880.307
RS	2.999.974.113	915.341.543	126.992.109	4.536.925.026
SC	2.081.635.333	606.025.609	226.542.180	2.402.363.655
SE	285.872.278	94.104.732	29.079.511	466.062.373
SP	13.043.716.280	3.949.209.732	2.562.458.125	17.502.688.200
TO	266.340.656	81.200.035	21.098.352	369.957.240

Nota: Os valores de arrecadação com ajuste sazonal.

Fonte: Elaboração própria.

Tabela 4 - Tabelas Contingenciais de Valores Médios de Arrecadação dos Efeitos Dif-em-Dif - Jun/19 a Jun/20

	Sem Lockdown	Com Lockdown	Média
Pré-pandêmico	2.236.205.888	1.292.203.812	1.816.649.410
Pandêmico	1.974.996.588	1.155.883.550	1.610.946.349
Média	2.135.740.773	1.239.772.942	1.737.532.848

	Baixo Isolamento	Alto Isolamento	Média
Pré-pandêmico	2.382.823.098	1.206.923.900	1.816.649.410
Pandêmico	2.118.550.552	1.064.295.669	1.610.946.349
Média	2.281.179.811	1.152.066.888	1.737.532.848

	Menor Gravidade	Maior Gravidade	Média
Pré-pandêmico	1.586.154.927	2.064.874.238	1.816.649.410
Pandêmico	1.428.319.442	1.807.621.479	1.610.946.349
Média	1.525.448.971	1.965.930.869	1.737.532.848

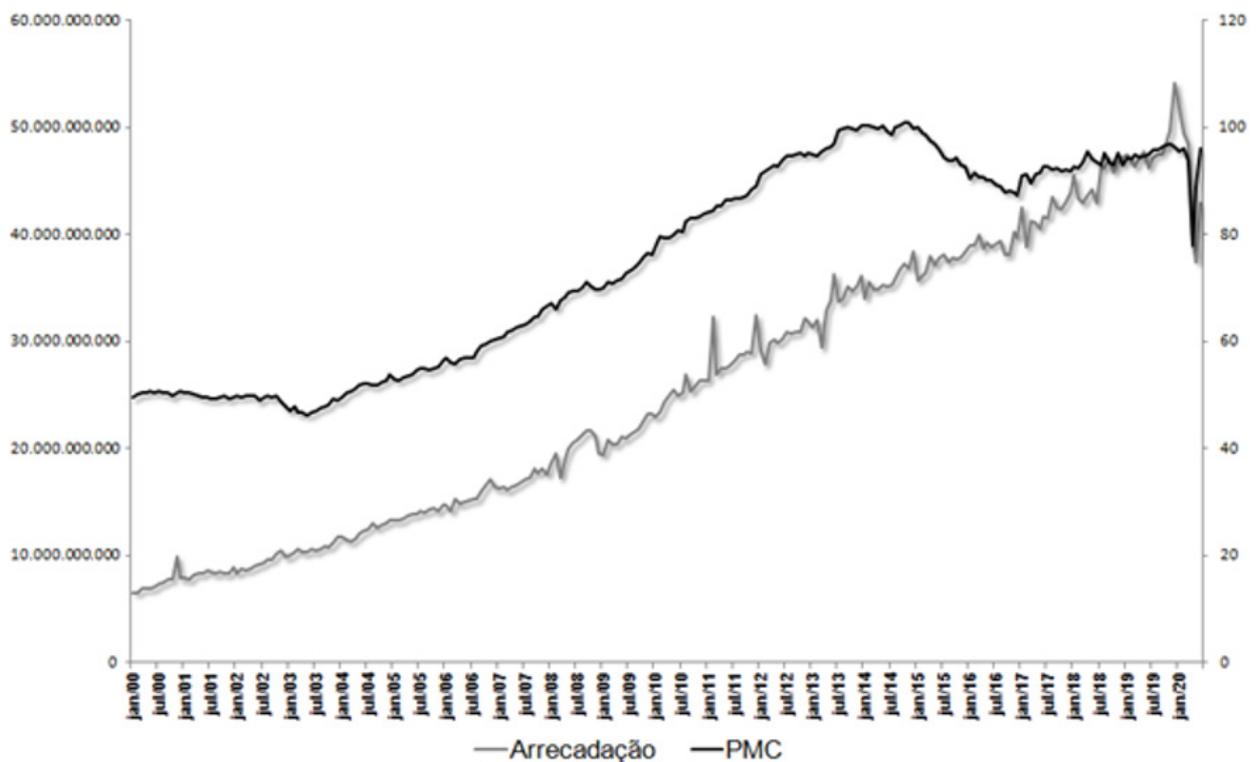
	Menores PIB	Maiores PIB	Média
Pré-pandêmico	518.906.999	3.214.218.160	1.816.649.410
Pandêmico	471.424.902	2.838.123.291	1.610.946.349
Média	500.644.654	3.069.566.288	1.737.532.848

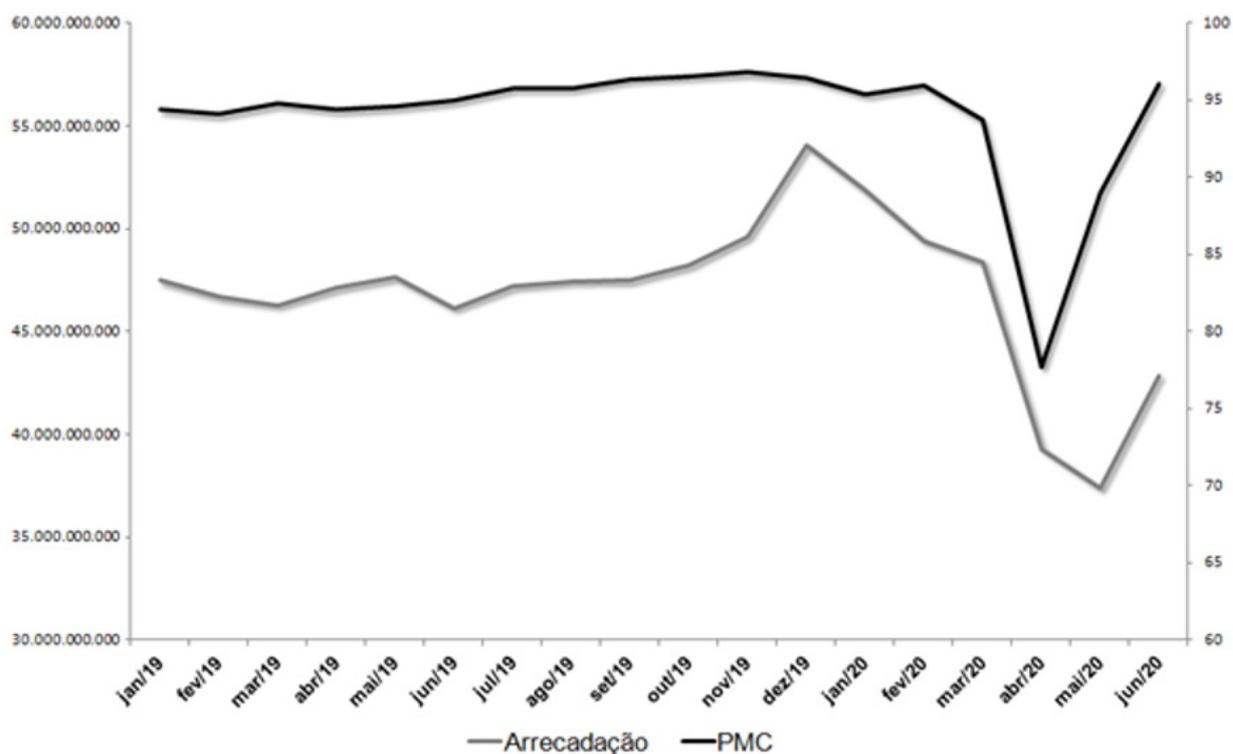
	Menores Densidades	Maiores Densidades	Média
Pré-pandêmico	1.345.633.604	2.193.462.055	1.816.649.410
Pandêmico	1.221.270.418	1.922.687.094	1.610.946.349
Média	1.297.801.609	2.089.317.839	1.737.532.848

Fonte: Elaboração própria.

A alta adesão populacional ao isolamento está associada, de maneira exploratória, a um menor valor médio de arrecadação, contudo com uma maior redução percentual das UF com indicativo de alto índice de isolamento no período pandêmico. Os estados com quadro pandêmico mais grave sofreram, percentualmente, uma maior redução de arrecadação média no período pandêmico, contudo em termos de arrecadação geral, estão associados à UFs com maior arrecadação. Em termos de PIB, os estados acima da mediana nacional, apresentaram uma redução média de arrecadação, maior em termos absolutos e percentuais, em relação aos demais UFs no período pandêmico. Os resultados observados para sub-tabela do PIB, são válidos para a densidade demográfica. Para um melhor entendimento da relação entre arrecadação e atividade comercial, apresetamos na Figura 1 gráficos de toda série histórica dessas variáveis, apresentando uma ênfase para o período recente de janeiro de 2019 em diante, onde é possível identificar com clareza a queda ocasionada pela pandemia do Covid-19.

Figura 1 – Dinâmica de Arrecadação mensal e Atividade do Comércio Varejista



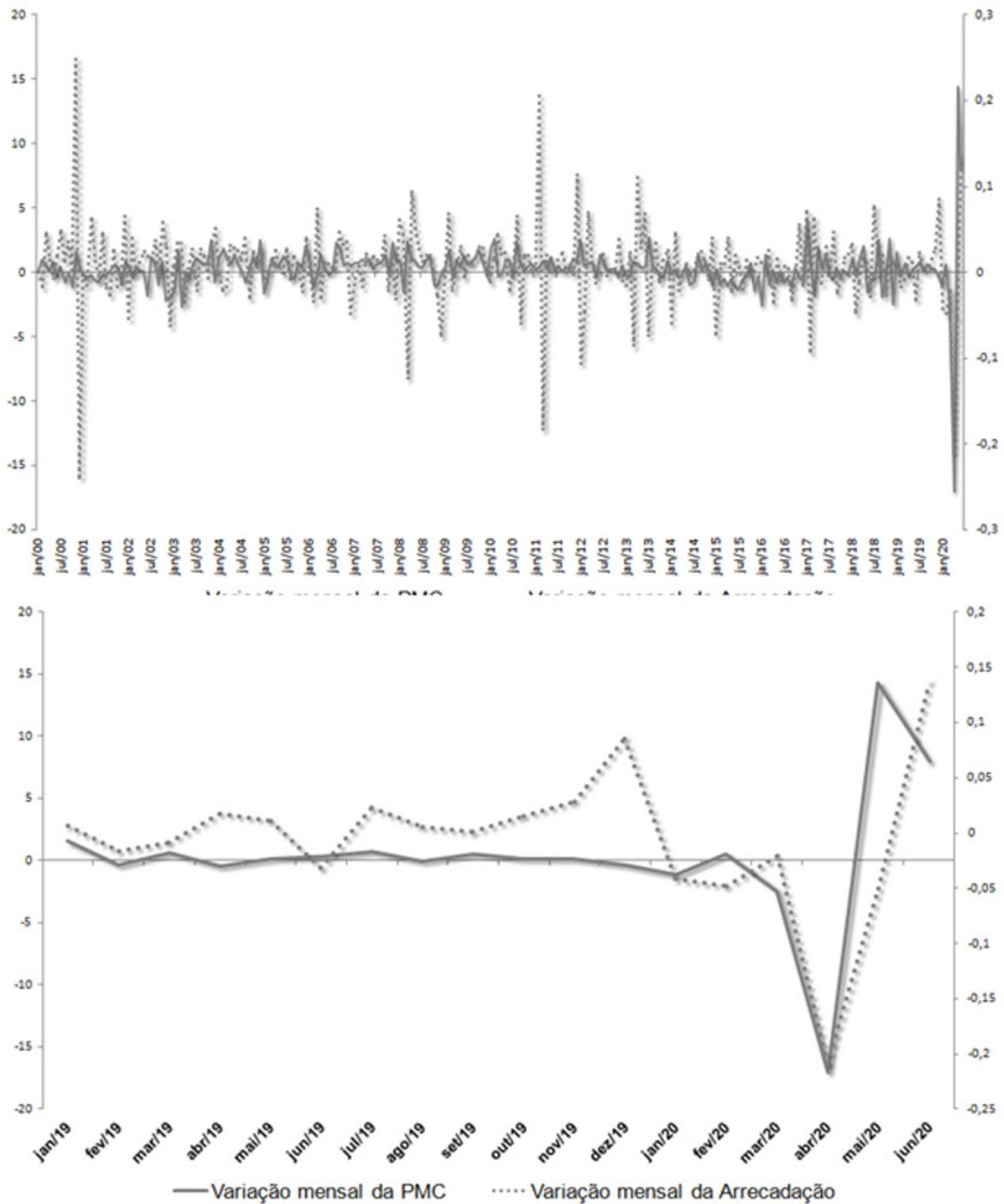


Fonte: CONFAZ, IBGE e Elaboração própria.

Na figura 2 apresentamos a primeira diferença do logaritmo dessas variáveis, também dando ênfase para o período recente pós-janeiro de 2019. Esse gráfico permite identificar mais claramente a direta relação da atividade comercial e a arrecadação. Dessa forma, esses indícios gráficos, reforçam a escolha dessa variável para proxy de atividade econômica estadual⁴.

Figura 2 – Dinâmica da Variação Mensal da Arrecadação e do Comércio Varejista.

⁴ Haja vista não haver, até onde os pesquisadores tem conhecimento, outra variável considerando o horizonte temporal e todas as unidades federativas, para o estudo desse painel.



Fonte: CONFAZ, IBGE e Elaboração própria.

Quando consideramos o painel epidêmico, apresentamos na Tabela 5 as estatísticas descritivas usuais da pandemia e os respectivos efeitos Dif-em-Dif a serem estimados. O número de óbitos por 100 mil/hab. precisa ser ponderado pela evolução epidêmica em cada estado, e, portanto, considerar o número de dias desde registro de primeiro óbito Covid-19. Quando ranqueamos da maior taxa de mortalidade para menor e associamos aos efeitos, podemos traçar, de maneira exploratória, um perfil

de agrupamento das UFs. O primeiro corte traz os estados com taxa acima de 70 mortes por 100 mil/hab, nesse grupo existe uma predominância de estados, com alta adesão ao isolamento social, com adoção de política de Lockdown, alta densidade demográfica e com PIB acima da mediana nacional.

Tabela 5 - Estado Pandêmico por UF

UF	Número de Óbitos Covid-19 por 100 mil/hab.	Número de dias de pandemia	I_iso	I_Lockout	I_grave	I_PIB	I_dens
AM	90	140	1	1	1	0	1
CE	82	142	1	1	1	1	1
PA	71	136	1	1	1	1	0
RR	71	133	0	0	1	0	1
RJ	70	149	1	1	1	1	1
AP	67	133	1	1	1	0	0
PE	66	142	1	1	1	1	1
AC	56	130	1	0	1	0	0
ES	55	135	0	0	1	0	1
SE	44	135	0	0	1	0	1
SP	42	150	0	0	1	1	1
AL	41	137	1	0	1	0	1
RN	40	139	0	1	1	0	1
RO	39	137	1	0	0	0	1
MA	39	138	1	1	0	0	0
PB	36	136	0	0	0	0	1
DF	33	139	1	0	0	1	1
MT	32	134	0	0	0	1	0
PI	29	140	1	1	0	0	0
TO	21	122	0	1	0	0	0
BA	18	139	0	0	0	1	0
GO	14	140	0	0	0	1	0
PR	11	141	0	1	0	1	0
RS	11	141	1	0	0	1	1
SC	10	140	0	0	0	1	1
MG	9	138	0	0	0	1	0
MS	8	137	0	1	0	0	0

Nota: O número de dias de pandemia é contado a partir do primeiro registro de óbito por Covid-19.

Fonte: Elaboração própria.

O segundo corte, apresenta de mortalidade entre 40 e 69 mortes por 100 mil/hab, nesse grupo metade dos estados teve grande adesão ao isolamento social, com apenas dois estados com política de Lockdown, sendo a maioria de alta densidade demográfica e de menores PIBs. O terceiro corte, tem a mortalidade entre 20 e 39 mortes por 100 mil/hab, novamente tem metade dos estados com alta adesão, contudo com poucos estados com política de Lockdown, na maior parte UFs com menores PIB e com metade sendo de alta densidade demográfica. O último grupo tem mortalidade até 19 mortes por 100 mil/hab, na sua maioria estados com PIB acima da mediana nacional, com baixa adesão ao isolamento social, baixa densidade demográfica e com apenas dois estados com política de Lockdown. Todos esses resultados serão submetidos à inferência contrafactual para que tenham tais perspectivas validadas. Na Tabela 6 apresentamos sub-tabelas contingenciais dos efeitos Dif-em-Dif a serem estimados na próxima seção.

Tabela 6 - Tabelas Contingenciais de Número médio de Óbitos Covid-19 por 100 mil/hab. dos Efeitos Dif-em-Dif - Primeiros 70 dias pandêmicos.

	Sem Lockdown	Com Lockdown	Média
Baixo Isolamento	59	36	53
Alto Isolamento	64	110	92
Média	61	85	72

	Menor Gravidade	Maior Gravidade	Média
Baixo Isolamento	32	90	53
Alto Isolamento	55	115	92
Média	40	105	72

	Menores PIB	Maiores PIB	Média
Baixo Isolamento	69	36	53
Alto Isolamento	87	97	92
Média	78	64	72

	Menores Densidades	Maiores Densidades	Média
Baixo Isolamento	29	76	53
Alto Isolamento	89	94	92
Média	54	86	72

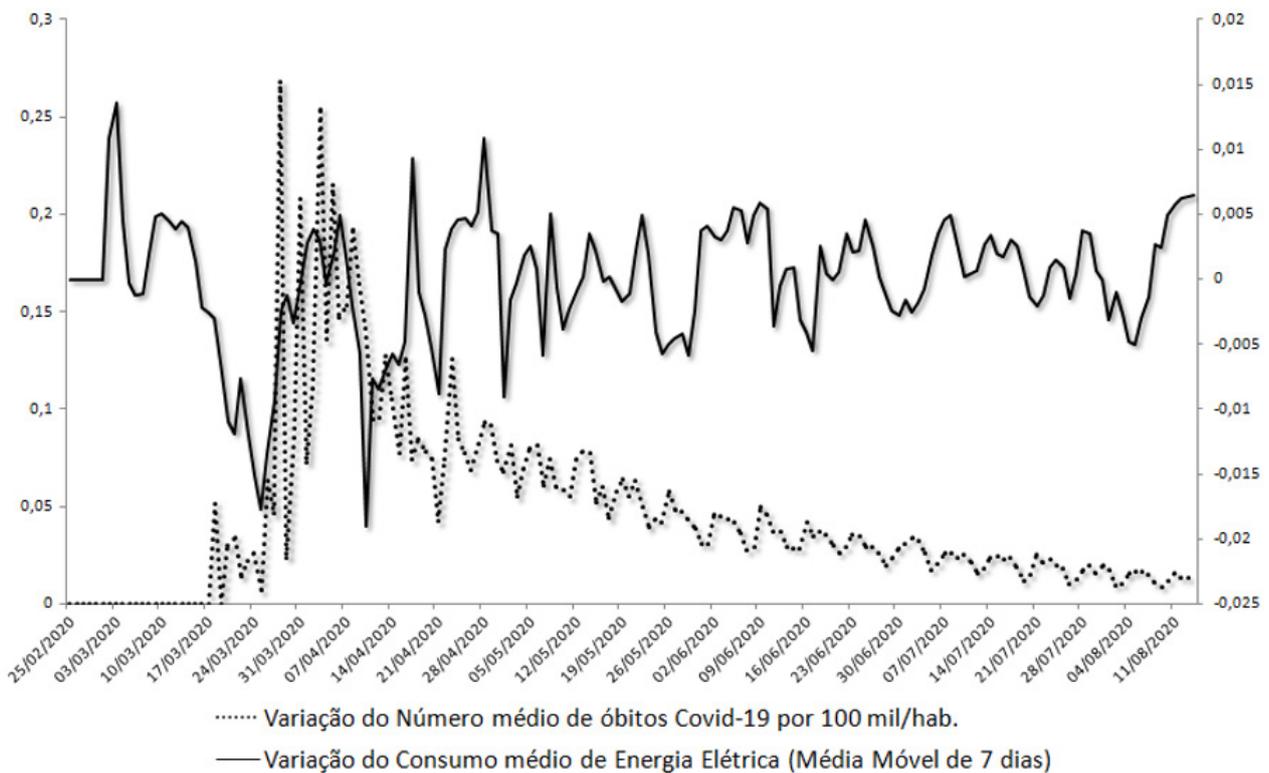
Fonte: Elaboração própria.

Os resultados da Tabela 6 foram considerados até 70º dia pandêmico para que fosse possível a comparação de médias de mortalidade por UF com diferentes evoluções da pandemia. O resultado da primeira sub-tabela indica que a taxa média de mortalidade é maior para UFs com políticas de Lockdown, e ainda maior para UFs com maior adesão ao isolamento social. Quando consideramos a gravidade pandêmica observamos que para os estados com baixa adesão ao isolamento social, a média de mortalidade é quase três vezes maior nos casos graves, e mais que o dobro para os estados de alta adesão. Considerando a distribuições dos indicativos de PIB, observamos que os estados de maior PIB e com baixa adesão tem menores taxas médias de mortalidade em relação a todas outras possibilidades. Com relação ao indicativo de densidade demográfica, estados com menor adesão ao isolamento social e baixa densidade demográfica apresentam menores médias de mortalidade. Os resultados gerais dessas sub-tabelas, indicam, de maneira exploratória, que quanto mais grave a pandemia na UF, maior a adesões ao isolamento social, à adoção de políticas de lockdown. Novamente, esses resultados carecem de inferência causal para elucidação desses indicativos iniciais.

Com a escolha do uso de energia elétrica como proxy para o nível de atividade econômica diária estadual, apresentamos na Figura 3, a relação dessa variável com a evolução pandêmica. Os gráficos para os dados agregados de todos os estados brasileiros precisam ser analisados com transformação de média móvel de 7 dias dos dados originais de consumo de energia elétrica para eliminar sua sazonalidade diária. A indicação gráfica sugere que em muitos momentos, existe uma relação positiva entre o consumo de energia elétrica e a evolução pandêmica.

Tanto no painel de estudo de arrecadação quanto no painel de estudo pandêmico por UF, é fundamental a utilização de covariável de controle para inferência contrafactual, tanto para os modelos de painel Dif-em-Dif quanto nos modelos temporais contrafactuais artificiais. Os indícios das estatísticas descritivas, muito embora, possam contribuir para uma prévia identificação de possíveis traços do fenômeno em estudo, não podem ser utilizados diretamente para avaliação de políticas públicas ou dos eventos quase experimentais proporcionados pelo advento da pandemia. Dessa forma na próxima seção iniciamos nossa inferência contrafactual pelos modelos de painel de diferença-em-diferença.

Figura 3 – Dinâmica Diária da Pandemia e Consumo de Energia Elétrica.



Fonte: CCEE, SRAG ARPEN e Elaboração própria.

4.2 Resultados dos Modelos de Dados em Painel de Diferenças-em-Diferenças

Apresentamos nessa seção os resultados dos modelos de dados em painel Dif-em-Dif considerando dois escopos de estudos:

- a) Estudo contrafactual das arrecadações mensais estaduais estimadas;
- b) Estudo contrafactual da taxa de mortalidade diária estimada.

Uma importante propriedade para utilização de modelos lineares que combinam séries de tempo diz respeito à estacionariedade. Negligenciar tal propriedade faz com que quaisquer resultados inferenciais possam ser espúrios e, portanto não representando a realidade do fenômeno. Essa preocupação evoluiu na literatura econométrica para a extensão dos tradicionais testes de raiz unitária para o contexto de dados em painel. Nesse sentido uma das principais motivações em adequar os testes de raiz unitária para painéis foi a possibilidade de aumento do poder do teste, ponto no qual, os testes para uma única série temporal, sofreram por muito tempo. Dentre algumas possibilidades de testes de raiz unitária para dados de painel, utilizamos nesse estudo o teste do tipo Fisher que combina os p-valores dos testes de raiz unitária unidade-específicos usando os quatro métodos propostos por Choi (2001). A hipótese nula sendo testada por estes métodos é que todas as unidades do painel contêm uma raiz unitária. Para um número finito de unidades, a alternativa é que pelo menos uma unidade seja estacionária. Com a finalidade de inferência contrafactual por meio de modelos de painéis de diferenças-em-diferenças, apresentamos na Tabela 7 os resultados dos testes de raiz unitária para todas as variáveis que serão utilizadas no estudo.

Tabela 7 - Teste do tipo Fisher para estacionariedade de Dados em Painel

Teste	Estatística	P-valor
Arrecadação	Pm = -4.5625	0.01
PMC	Pm = -7.3819	0.01
Var Arrecadação	Pm = -33.787	0.01
Var PMC	Pm = -23.007	0.01
NR_OBITOS100	Pm = -8.3726	0.01
Consumo_EE_MWh	Pm = -4.5062	0.01
Var NR_OBITOS100	Pm = -10.705	0.01
Var Consumo_EE_MWh	Pm = -14.898	0.01

Nota: 1) Var se refere a primeira diferença do logaritmo da variável, ou seja, sua variação na margem.

2) Para as variáveis Arrecadação e PMC, foram testadas diferentes defasagens: 2, 5 e 10. As conclusões não se alteram e reportamos os resultados para 10° defasagem.

3) Para as variáveis Número de óbitos por 100 mil/hab e Consumo de energia elétrica, foram testadas diferentes defasagens: 5, 10, 20 e 30. As conclusões não se alteram e reportamos os resultados para 30° defasagem.

Fonte: Elaboração própria.

As conclusões dos testes, mesmo considerando diferentes defasagens, indicam séries estacionárias em nível, sempre rejeitando a hipótese de presença de raiz unitária. Dessa forma, temos a evidência de estacionariedade que nos permite utilizar os métodos tradicionais de estimação de dados em painel.

4.2.1 Resultados dos Modelos DiD para Arrecadação Estadual

Partindo de uma abordagem de hipóteses mais restritas para modelos mais gerais, estimamos em um primeiro momento um painel Pooling para a arrecadação mensal das UF da amostra, com a forte hipótese de não haver heterogeneidade individual, e de que os parâmetros são constantes no tempo. Os resultados desse primeiro modelo estão detalhados na Tabela 8 a seguir.

Os resultados desse modelo tem significância global validada pelo teste F com uma parcela de 74% da variabilidade da arrecadação dos estados, no período amostral, explicada. Considerando os parâmetros estimados, observamos que, sendo validas as hipóteses desse modelo, constata-se que a atividade comercial tem impacto⁵ médio positivo, estatisticamente significativo, de grande magnitude na arrecadação mensal dos estados. Considerando as variáveis de efeitos contrafactuais, apenas três efeitos Dif.-em-Dif, dentre os cinco, foram estatisticamente significantes. Estados que não estão entre os maiores PIB, teriam tido uma maior arrecadação (média) durante o período pandêmico, caso ti-

5 A presença da população tem objetivo de corrigir distorções que a diferença populacional possa causar nos coeficientes estimados, alternativamente foram estimados painéis com arrecadação per capita, como os resultados não se alteraram significativamente, optamos por reportar os modelos com parâmetro populacional estimado explicitamente descrito.

vessem o nível de PIB elevado. Estados que não tiveram as mais graves pandemias Covid-19, teriam uma arrecadação (em média) maior, caso tivessem enfrentado uma pandemia mais severa, esse é um resultado contraintuitivo, e pode ser resultado das hipóteses assumidas nesse modelo. O último efeito contrafactual estatisticamente significantes está relacionado aos estados de maior adesão ao isolamento social: estados com baixa adesão teriam uma queda (em média) de arrecadação, caso tivessem tido uma alta adesão à política de isolamento social. O efeito contrafactual do lockdown não se mostrou estatisticamente significante. Ressaltamos que as hipótese desse modelo são extremamente fortes e sugerem cautela com essas evidência.

Tabela 8 - Modelo de Dados em Painel Dif-em-Dif para Arrecadação Estadual e os efeitos Covid-19

Modelo Pooling - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
(Intercept)	-1.5650e+09	4.4335e+07	-35.2991	< 2.2e-16 ***
PMC_Vendas	1.8875e+07	5.2359e+05	36.0499	< 2.2e-16 ***
pop	1.8072e+02	1.6256e+00	111.1711	< 2.2e-16 ***
I_grave	-5.3478e+06	2.8437e+07	-0.1881	0.85083
I_PIB	-5.4803e+08	2.9784e+07	-18.4005	< 2.2e-16 ***
I_iso	-6.1535e+07	2.6230e+07	-2.3460	0.01901 *
I_Lockout	-1.1552e+08	2.7133e+07	-4.2574	2.097e-05 ***
I_densidade	1.6022e+08	2.7957e+07	5.7309	1.043e-08 ***
I_pan	-1.6800e+08	1.8886e+08	-0.8895	0.37376
I_pan_lock	-8.2539e+07	1.9030e+08	-0.4337	0.66450
I_pan_grave	3.2059e+08	1.9387e+08	1.6537	0.09824 .
I_pan_pib	1.0737e+09	1.6706e+08	6.4269	1.392e-10 ***
I_pan_iso	-3.0134e+08	1.7902e+08	-1.6833	0.09237 .
I_pan_den	1.3748e+08	1.9557e+08	0.7029	0.48212
OBS				6642
R2				0.74999
R2 ajustado				0.7495
Estatística F				1529.43 com 13 e 6628 GL, p-valor: < 2,22e-16

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Uma maneira de flexibilizar tais restrições é a utilização dos modelos de efeitos fixos, que assume que características individuais das UF (heterogeneidade individual) podem ser estimadas para cada Estado da amostra. Essa abordagem permite modelar diferenças no comportamento individual das UF, os parâmetros ainda constantes no tempo, mas podendo variando de estado para estado. O primeiro modelo de efeito fixo estimado, também conhecido como LSDV (Least Square Dummy Variable), tem seus resultados estimados apresentados na Tabela 9.

Tabela 9 - Modelo de Dados em Painel Dif-em-Dif para Arrecadação Estadual e os efeitos Covid-19

Modelo de Efeitos Fixos (LSDV) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
PMC_Vendas	1.882e+07	4.583e+05	41.073	< 2e-16 ***
pop	7.315e+00	4.309e+01	0.170	0.865201
I_grave	7.367e+09	1.672e+09	4.407	1.06e-05 ***
I_PIB	8.057e+09	1.917e+09	4.203	2.67e-05 ***
I_iso	6.970e+08	1.957e+08	3.562	0.000370 ***
I_Lockout	-1.019e+09	8.876e+07	-11.485	< 2e-16 ***
I_densidade	-8.611e+09	1.611e+09	-5.345	9.36e-08 ***
I_pan	-1.669e+08	1.627e+08	-1.026	0.304937
I_pan_lock	-8.242e+07	1.639e+08	-0.503	0.615047
I_pan_grave	3.205e+08	1.669e+08	1.920	0.054960 .
I_pan_pib	1.073e+09	1.439e+08	7.461	9.67e-14 ***
I_pan_iso	-3.018e+08	1.542e+08	-1.957	0.050342 .
I_pan_den	1.374e+08	1.684e+08	0.816	0.414468
factor(sg_uf)AC	-9.242e+09	1.820e+09	-5.078	3.93e-07 ***
factor(sg_uf)AL	-5.838e+08	1.391e+08	-4.196	2.75e-05 ***
factor(sg_uf)AM	5.588e+08	1.733e+08	3.224	0.001272 **
factor(sg_uf)AP	-8.226e+09	1.889e+09	-4.356	1.35e-05 ***
factor(sg_uf)BA	-9.344e+09	1.279e+09	-7.305	3.11e-13 ***
factor(sg_uf)CE	-7.287e+09	1.837e+09	-3.967	7.36e-05 ***
factor(sg_uf)DF	-1.194e+09	3.676e+08	-3.247	0.001171 **
factor(sg_uf)ES	4.262e+08	1.273e+08	3.348	0.000818 ***
factor(sg_uf)GO	-8.576e+09	1.617e+09	-5.304	1.17e-07 ***
factor(sg_uf)MA	-6.422e+08	1.171e+08	-5.484	4.32e-08 ***

factor(sg_uf)MG	-7.029e+09	1.009e+09	-6.966	3.57e-12 ***
factor(sg_uf)MS	1.848e+08	8.952e+07	2.064	0.039071 *
factor(sg_uf)MT	-8.867e+09	1.769e+09	-5.012	5.54e-07 ***
factor(sg_uf)PA	-1.597e+10	3.469e+09	-4.603	4.24e-06 ***
factor(sg_uf)PB	7.584e+09	1.783e+09	4.253	2.14e-05 ***
factor(sg_uf)PE	-7.129e+09	1.819e+09	-3.919	8.97e-05 ***
factor(sg_uf)PI	-8.777e+08	1.498e+08	-5.859	4.88e-09 ***
factor(sg_uf)PR	-6.965e+09	1.494e+09	-4.661	3.22e-06 ***
factor(sg_uf)RJ	-5.813e+09	1.488e+09	-3.907	9.42e-05 ***
factor(sg_uf)RN	1.162e+09	8.876e+07	13.097	< 2e-16 ***
factor(sg_uf)RO	6.957e+09	1.506e+09	4.621	3.90e-06 ***
factor(sg_uf)RR	6.961e+09	1.616e+09	4.832	3.96e-06 ***
factor(sg_uf)RS	-1.054e+09	3.565e+08	-3.136	0.001682 **
factor(sg_uf)SC	-1.001e+09	3.591e+08	-3.201	0.001191 **
factor(sg_uf)SE	-8.151e+08	1.321e+08	-5.989	4.98e-09 ***
factor(sg_uf)SP	-7.983e+09	1.237e+09	-3.992	9.62e-05 ***
factor(sg_uf)TO	1.925e+08	8.987e+07	2.071	0.039981 *
OBS				6642
R2				0.8526
R2 ajustado				0.8518
Estatística F				1124 com 34 e 6608 GL, p-valor: < 2,22e- 16

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Os resultados do modelo LSDV, assim como o modelo Pooling, tem significância global confirmada pelo teste F, contudo com um percentual maior de variância explicada no conceito de R2, de 85%. Os resultados dos parâmetros estimados no modelo Pooling, continuaram validos no modelo LSDV, com exceção da covariável de população, haja vista que parte dos ajustes proposto pela inclusão dessa variável, acabam capturados nos parâmetros de heterogeneidade individual. Os efeitos causais estimados por Dif.-em-Dif, também se mantiveram em relação ao modelo sem heterogeneidade individual, tendo mesmo sinais e interpretações. A heterogeneidade individual nesse modelo tem por hipótese ser uma característica intrínseca da UF, não sendo dada por um choque aleatório (Modelos de Efeitos Aleatórios), sendo estimada por parâmetros interceptos individuais. Os resultados da Tabela 9

apresentam evidências de que a heterogeneidade individual está presente, de maneira estatisticamente significativa, para todos os estados brasileiros.

Outra abordagem de estimação de efeitos fixos, conhecida como Within, extrai a média temporal das variáveis de seus respectivos países e então re-estima o painel. Os resultados desse método está detalhado na Tabela 10.

Tabela 10 - Modelo de Dados em Painel Dif-em-Dif para Arrecadação Estadual e os efeitos Covid-19

Modelo de Efeitos Fixos (Within) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
PMC_Vendas	18824291	458319	41.0725	< 2.2e-16 ***
I_pan	-166876696	162650639	-1.0260	0.30494
I_pan_lock	-82417546	163882571	-0.5029	0.61505
I_pan_grave	320462375	166948386	1.9195	0.05496 .
I_pan_pib	1073443562	143868757	7.4613	9.669e-14 ***
I_pan_iso	-301767421	144167379	-2.0932	0.03645 *
I_pan_den	137448923	168420210	0.8161	0.41447
OBS				6642
R2				0.21767
R2 ajustado				0.21376
Estatística F				262.65 com 7 e 6608 GL, p-valor: < 2,22e- 16

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Observe que para esse modelo variáveis constantes no tempo, tem parâmetros nulos, haja vista o primeiro passo de extração da média temporal. Os resultados gerais se mantém em relação à versão LSDV, com significância global, os mesmos efeitos contrafactuais estatisticamente significantes. Todos os resultados apresentados são válidos apenas na condição de que as hipóteses dos modelos estejam ativas. Falhas nas hipóteses geram inferências e/ou estimativas incorretas, desse modo submetemos os modelos estimados a uma bateria de vários testes que tornem suas conclusões mais confiáveis, e dessa forma apresentamos os resultados de todos esses testes na Tabela 11.

Tabela 11 - Testes em Modelos de Dados em Painel

Teste	Estatística	Graus de Liberdade	P-valor
Teste F para efeitos individuais	F = 116.48	gl1 = 20, gl2 = 6608	< 2.2e-16
Breusch-Pagan studentizado	BP = 764.38	gl = 12	< 2.2e-16
Teste Breusch-Godfrey/Wooldridge para correlação serial em modelos de painel	chisq = 6315.9	gl = 246	< 2.2e-16
Teste Pesaran para dependência Cross-Sectional em Painéis	z = 75.192	-	< 2.2e-16
Teste do tipo Fisher para estacionariedade do resíduo do Painel	Pm = -6.7966	-	0.01
Teste de Wooldridge para efeitos individuais não-observáveis	z = 1.8955	-	0.05803

Fonte: Elaboração própria.

O primeiro teste coloca a hipótese do modelo Pooling a prova, ou seja, a existência ou não de heterogeneidade individual. O p-valor aponta para a existência de efeitos individuais estatisticamente significantes, rejeitando a hipótese nula do teste. Dessa forma as evidências dos modelos de efeitos fixos devem ser visto com maior segurança e, portanto serão aplicados os demais testes para o modelo Within, dado sua parcimônia em termos de parâmetros estimados, resultando em um maior grau de liberdade, importante característica na inferência em modelos lineares. O teste Breusch-Pagan na sequência, não rejeita a hipótese de heterocedasticidade, ou seja, a variância não é constante e, portanto as inferências utilizando o teste t sobre parâmetros estimados estão comprometidos. O teste Breusch-Godfrey/Wooldridge apresenta evidência estatística de presença de autocorrelação serial, que tem efeitos similares a da heterocedasticidade sobre a inferência de parâmetros, novamente comprometendo a inferência dos testes t. A dependência cross-sectional se apresenta em painéis com longas séries de tempo. O teste de Pesaran é utilizado para esse tipo de dependência, sendo que a hipótese nula é de que os resíduos através dos indivíduos não estão correlacionados, dessa forma temos evidência de existe uma estrutura de dependência entre os resíduos das UF da amostra. O teste do tipo Fisher para estacionariedade dos resíduos garante que bom ajuste do modelo linear de dados em painel. Por fim o teste de Wooldridge para efeitos individuais não observáveis está no limiar (p-valor de 0,058) de rejeição da hipótese nula de ausência desses efeitos individuais não observáveis.

Heteroscedasticidade e autocorrelação serial são problemas bastante recorrentes em modelos de Dif-em-Dif. Uma prática comum nestas circunstâncias é empregar um estimador consistente da matriz de covariâncias a fim de formular teste de hipóteses assintoticamente válidos. Diante das evidências de heterocedasticidade e autocorrelação serial no modelo Within reestimamos esse modelo considerando a correção proposta por Cribari-Neto (2004) ao estimador HC4 amplamente empregado e apresentado por Arellano (2003,1987). Esse estimador é consistente tanto na presença de heteroscedasticidade quando de autocorrelação serial dos erros além de considerar o contexto de pequenas amostras cross-section e possíveis pontos de alavancagem. Os resultados do modelo Within robusto a quebra das hipóteses de heterocedasticidade e autocorrelação serial é apresentado na Tabela 12.

Tabela 12 - Modelo de Dados em Painel Dif-em-Dif para Arrecadação Estadual e os efeitos Covid-19

Modelo de Efeitos Fixos (Within+Corr. Het.Auto-Corr.) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
PMC_Vendas	18824291	6453252	2.9170	0.003546 **
I_pan	-166876696	261506471	-0.6381	0.523407
I_pan_lock	-82417546	239601546	-0.3440	0.730874
I_pan_grave	320462375	420633067	0.7619	0.446172
I_pan_pib	1073443562	358424948	2.9949	0.002756 **
I_pan_iso	-301767421	155002391	-1.9469	0.051681 .
I_pan_den	137448923	229617721	0.5986	0.549461

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

A correção da heterocedasticidade e autocorrelação serial altera significativamente o desvio-padrão dos parâmetros estimados e, portanto também toda inferência dos testes t. Os novos resultados apenas três variáveis são estatisticamente significantes: o volume de vendas no comércio, com parâmetro positivo, indicando que cada um ponto de alteração no índice PMC, gera (em média) dezoito milhões de reais de arrecadação, já livre de efeitos sazonais; Os efeitos contrafactuais de adesão a política de isolamento social na pandemia, com parâmetro negativo, ou seja, estados que não tiveram alta adesão ao isolamento social, teriam tido uma queda (em média) de pouco mais de trezentos milhões (considerando o desvio-padrão estimado, os impactos médios estariam entre 146 milhões a 456 milhões), em relação aos estados que tiveram uma alta adesão ao isolamento social. Os estados que não estão no grupo dos maiores PIB estaduais, teriam tido um aumento médio de arrecadação de pouco mais de um bilhão de reais, caso tivessem um PIB entre os maiores do Brasil.

4.2.2 Resultados dos Modelos DiD para Taxa de Mortalidade

Seguindo as mesmas etapas de modelagem da seção 4.2.1, apresentamos nessa seção os resultados de inferência causal para as taxas de mortalidade das UF. Conforme os resultados da Tabela 8 as variáveis desse painel (taxas de mortalidade e consumo de energia elétrica) também são estacionárias no contexto de dados em painel e portanto estimamos inicialmente o modelo Pooling de hipóteses mais restritas sobre a heterogeneidade individual, sendo seus resultados apresentados na Tabela 13.

Tabela 13 - Modelo de Dados em Painel Dif-em-Dif para Número de Óbitos por 100 mil/hab e Características Estaduais

Modelo Pooling - Parâmetros Estimados

	Estimativa	Desv.Pad	t valor	Pr(> t)
(Intercept)	7.0277e+00	2.6570e+00	2.6449	0.0081980 **
Consumo_EE_MWh	3.5343e-05	9.4917e-06	3.7236	0.0001988 ***
I_PIB	-8.5225e+00	2.7126e+00	-3.1418	0.0016896 **
I_iso	-8.9520e+00	2.6112e+00	-3.4283	0.0006126 ***
I_Lockout	-1.2293e+01	1.8661e+00	-6.5875	4.971e-11 ***
I_densidade	-8.4938e+00	2.6747e+00	-3.1756	0.0015049 **
I_pan	2.5916e+01	2.7642e+00	9.3754	< 2.2e-16 ***
I_pan_lock	3.7456e+01	2.1568e+00	17.3664	< 2.2e-16 ***
I_pan_pib	-7.6100e+00	2.8479e+00	-2.6722	0.0075624 **
I_pan_iso	2.2424e+01	2.8482e+00	7.8732	4.274e-15 ***
I_pan_den	1.4058e+01	2.8517e+00	4.9295	8.531e-07 ***
OBS				4644
R2				0.44589
R2 ajustado				0.44015
Estatística F				77.6167 com 10 e 4633 GL, p-valor: < 2,22e-16

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Os resultados do modelo Pooling para taxa de mortalidade apresenta significância global pelo teste F com 45% da variabilidade da taxa de mortalidade explicada no conceito R2. Todos os parâmetros estimados se mostraram estatisticamente significantes para os mais tradicionais níveis de significância.

Considerando válidas as hipóteses desse modelo, o consumo de energia elétrica, utilizado como proxy para o nível de atividade econômica diária, tem parâmetro positivo, de onde depreende-se que aumentos no nível de atividade estão associados a elevações médias da taxa de mortalidade. Dentre os efeitos causais, os estados sem política de lockdown e baixa adesão a política de isolamento social, teriam (em média) maiores taxas de mortalidade, caso tivessem adotado tal política ou tivessem tido alta adesão ao isolamento social. Essa mesma interpretação é validade para estados com mais baixa densidade demográfica. A presença entre os maiores PIBs do Brasil, teria em média, um efeito contrafactual negativo, ou seja, o grupo de menor PIB, teria tido uma menor mortalidade, caso seu nível de PIB fosse maior. Novamente, os resultados contraintuitivos desse modelo, devem ser lidos com cautela, haja vista a restrição de suas hipóteses. Esses resultados novamente, carecem de flexibilização

de suas hipóteses e validações de testes para dados em painel para assegurar essas inferências. Considerando modelos de efeitos fixos, a abordagem LSDV foi estimada e seus resultados apresentados na Tabela 14.

Tabela 14 - Modelo de Dados em Painel Dif-em-Dif para para Número de Óbitos por 100 mil/hab e Características Estaduais

Modelo de Efeitos Fixos (LSDV) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
Consumo_EE_	4.519e-04	6.315e-05	7.155	9.67e-13 ***
MWh				
I_PIB	-1.598e+02	2.360e+01	-6.774	1.41e-11 ***
I_iso	8.441e+01	1.856e+01	4.548	5.55e-06 ***
I_Lockout	-7.172e+00	3.412e+00	-2.102	0.035592 *
I_densidade	-3.775e+00	3.608e+00	-1.046	0.295508
I_pan	2.641e+01	2.662e+00	9.921	< 2e-16 ***
I_pan_lock	1.052e+00	5.208e+00	0.202	0.839914
I_pan_pib	-4.397e+00	2.796e+00	-1.572	0.115918
I_pan_iso	2.048e+01	2.744e+00	7.462	1.02e-13 ***
I_pan_den	1.658e+01	2.768e+00	5.991	2.25e-09 ***
factor(UF)AC	-7.913e+01	1.874e+01	-4.222	2.47e-05 ***
factor(UF)AL	-1.040e+02	1.883e+01	-5.521	3.55e-08 ***
factor(UF)AM	-6.384e+01	1.922e+01	-3.322	0.000901 ***
factor(UF)AP	-6.520e+01	1.895e+01	-3.440	0.000587 ***
factor(UF)BA	1.224e+02	1.873e+01	6.536	6.99e-11 ***
factor(UF)CE	8.817e+01	6.659e+00	13.241	< 2e-16 ***
factor(UF)DF	4.951e+01	5.878e+00	8.422	< 2e-16 ***
factor(UF)ES	-1.300e+00	4.182e+00	-0.311	0.755952
factor(UF)GO	1.315e+02	2.037e+01	6.459	1.17e-10 ***
factor(UF)MA	-9.477e+01	1.968e+01	-4.815	1.52e-06 ***
factor(UF)MG	7.532e+01	1.350e+01	5.581	2.53e-08 ***
factor(UF)MS	-1.492e+01	4.016e+00	-3.717	0.000204 ***
factor(UF)MT	1.557e+02	2.169e+01	7.177	8.25e-13 ***
factor(UF)PA	7.753e+01	7.867e+00	9.855	< 2e-16 ***
factor(UF)PB	-9.167e+00	3.986e+00	-2.300	0.021513 *

factor(UF)PE	7.038e+01	6.369e+00	11.051	< 2e-16 ***
factor(UF)PI	-9.712e+01	1.895e+01	-5.125	3.10e-07 ***
factor(UF)PR	1.178e+02	1.844e+01	6.388	1.84e-10 ***
factor(UF)RJ	3.838e+01	5.968e+00	6.431	1.40e-10 ***
factor(UF)RN	-6.351e+01	1.055e+01	-3.611	0.000599 ***
factor(UF)RO	-1.052e+02	1.870e+01	-5.627	1.95e-08 ***
factor(UF)RR	2.519e+01	4.027e+00	6.255	4.33e-10 ***
factor(UF)RS	1.099e+02	1.654e+01	6.644	4.75e-10 ***
factor(UF)SC	1.079e+02	1.923e+01	5.613	2.11e-08 ***
factor(UF)SE	-9.433e+01	1.811e+01	-5.208	3.05e-07 ***
factor(UF)SP	3.929e+01	5.657e+00	6.921	1.89e-10 ***
factor(UF)TO	-9.005e+01	1.991e+01	-4.522	1.61e-07 ***
OBS				4644
R2				0.5669
R2 ajustado				0.5639
Estatística F				188.6 com 32 e 4612 GL, p-valor: < 2,22e- 16

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Os resultados do modelo LSDV apresenta significância global pelo teste F com uma elevação do R2, chegando a 57%, o que pode está ocorrendo pela redução de graus de liberdade em razão do aumento do número de parâmetros estimados, característico desse modelo. O consumo de energia elétrica permaneceu estatisticamente significativo e positivo, mantendo sua interpretação do modelo Pooling. Dentre os quatro⁶ efeitos Dif-em-Dif estimados nesse modelo, podemos observar que apenas os efeitos de adesão a política de isolamento social e densidade demográfica foram estatisticamente significantes, ou seja, estados que não tiveram alta adesão ao isolamento social, teriam tido (em média) taxas de mortalidade maiores, caso sua população tivesse aderido a essa política, enquanto que estados com baixa densidade demográfica, teriam tido (em média) maior taxa de mortalidade maior, caso tivessem maior densidade demográfica. Novamente, o resultado contraintuitivo do efeito I_pan_iso, pode está relacionado com as hipóteses desse modelo, e serão melhor descritos nas próximas etapas do estudo.

⁶ Observe que nos modelos de dados em painel com taxa de mortalidade não consideramos o efeito Dif-em-Dif da gravidade pandêmica por UF considerada nos modelos com painel de arrecadação, haja vista que a gravidade, agora como taxa de mortalidade, é a variável resposta do modelo.

Quando analisamos a heterogeneidade individual estimadas pelos parâmetros intercepto do modelo LSDV observamos todos as UF tem parâmetros significativos, exceto o Estado do Espírito Santo. O grande número de parâmetros estatisticamente significativos nesse modelo sugere um elevado custo de graus de liberdade já destacado anteriormente, recomendando a estimação do modelo de efeito fixo Within. Os resultados desse modelo estão detalhados na Tabela 15.

Tabela 15 - Modelo de Dados em Painel Dif-em-Dif para Número de Óbitos por 100 mil/hab e Características Estaduais

Modelo de Efeitos Fixos (Within) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
Consumo_EE_MWh	4.5187e-04	6.3152e-05	7.1552	9.673e-13 ***
I_pan	2.6408e+01	2.6620e+00	9.9206	< 2.2e-16 ***
I_pan_lock	1.052e+00	5.208e+00	0.202	0.839914
I_pan_pib	-4.3971e+00	2.7963e+00	-1.5724	0.1159
I_pan_iso	2.0475e+01	2.7441e+00	7.4616	1.016e-13 ***
I_pan_den	1.6581e+01	2.7679e+00	5.9907	2.249e-09 ***
OBS				4644
R2				0.18656
R2 ajustado				0.1811
Estatística F				211.554 com 5 e 4612 GL, p-valor: < 2,22e-16

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

O modelo Within apresentou significância global pelo teste F e com R2, de 18%. Os parâmetros estatisticamente significativos permanecem os mesmo, com os mesmo sinais e interpretações. Esses resultados inferenciais precisam serem testadas em relação as hipótese sobre variância, efeitos individuais não observados, etc. O resultado desses testes são apresentados na Tabela 16.

Tabela 16 - Testes em Modelos de Dados em Painel

Teste	Estatística	Graus de Liberdade	P-valor
Teste F para efeitos individuais	F = 20.245	gl1 = 21, gl2 = 4612	< 2.2e-16
Breusch-Pagan studentizado	BP = 1025.2	gl = 10	< 2.2e-16
Teste Breusch-Godfrey/Wooldridge para correlação serial em modelos de painel	chisq = 4500.4	gl = 172	< 2.2e-16
Teste Pesaran para dependência Cross-Sectional em Painéis	z = 198.74	-	< 2.2e-16
Teste do tipo Fisher para estacionariedade do resíduo do Painel	Pm = -6.9967	-	0.01
Teste de Wooldridge para efeitos individuais não-observáveis	z = 2.7444	-	0.006063

Fonte: Elaboração própria.

O teste F para efeitos individuais conclui pela presença de heterogeneidade individual, sugerindo o uso dos modelos de efeitos fixos. Assim como nos modelos do painel de arrecadação mensal, os testes Breusch-Pagan e Breusch-Godfrey/Wooldridge, apresentam evidências de heterocedasticidade e autocorrelação serial, comprometendo as inferências de teste t apresentadas nos modelos estimados. Muitos embora os resíduos sejam estacionários pelo teste tipo Fisher, temos a indicação de presença de dependência cross-section pelo teste Pesaran. O teste de Wooldridge, não rejeita a hipótese de efeitos individuais não-observáveis, esses efeitos individuais não-observáveis podem ter várias origens, mas talvez a mais relevante seja a possibilidade de endogeneidade entre taxa de mortalidade e consumo de energia elétrica como proxy para atividade econômica diária, onde, uma maior taxa de mortalidade pode reduzir a atividade econômica, assim como uma maior nível de atividade econômica, pode sugerir uma maior incidência do número de caso e por fim aumentando a taxa de mortalidade, contudo essa questão será melhor detalhada na sequência do estudo. Diante dessas evidências o modelo Within foi estimado com a mesma correção do estimador HC4 comentado na seção anterior. Os resultados estão contidos na Tabela 17.

Tabela 17 - Modelo de Dados em Painel Dif-em-Dif para Número de Óbitos por 100 mil/hab e Características Estaduais

Modelo de Efeitos Fixos (Within+Corr. Het.Auto-Corr.) - Parâmetros Estimados

	Estimativa	Desv.Pad	t valor	Pr(> t)
Consumo_EE_MWh	0.00045187	0.00020202	2.2368	0.02535 *
I_pan	26.40819890	5.94871728	4.4393	9.236e-06 ***
I_pan_lock	1.0520e+00	4.4589e+00	0.2359	0.8134946
I_pan_pib	-4.39706543	7.63732143	-0.5757	0.56482
I_pan_iso	20.47524110	7.25676182	2.8215	0.00480 **
I_pan_den	16.58146673	6.99635416	2.3700	0.01783 *
I_pan_den	137448923	229617721	0.5986	0.549461

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 '.' 1

Fonte: Elaboração própria.

Os resultados da correção para heterocedasticidade e autocorrelação não altera o número de parâmetros estatisticamente significantes. A cada aumento de 10.000 Megawatt-hora⁷, existiu, no período amostral, um aumento médio de 5 novos óbitos por 100 mil habitantes, ou ainda, a cada aumento de 100.000 Megawatt-hora, existiu 45 novos óbitos por 100 mil habitantes. Dentre os efeitos Dif-em-Dif significantes, estados com baixa densidade demográfica, teriam uma taxa de mortalidade maior (em média) de 17 óbitos em 100 mil habitantes, caso tivessem alta densidade demográfica. O efeito Dif-em-Dif associado a política de lockdown não se mostrou estatisticamente significativo em nenhum modelo. O efeito Dif-em-Dif associado a política de isolamento social, com parâmetro positivo, se traduz como: estados com baixa adesão ao isolamento social teriam em média, uma taxa de mortalidade de 20 óbitos a mais por 100 mil habitantes, caso tivessem uma alta adesão. Esse resultado contraintuitivo, tem íntima relação com a hipótese de efeitos individuais não-observáveis verificada no teste de Wooldridge, ou seja a possível presença de endogeneidade entre, taxa de mortalidade e adesão ao isolamento social, onde podemos ter, uma maior adesão a política de isolamento social em razão de uma maior taxa de mortalidade. Outro possível efeito bem conhecido na literatura de modelos de painel em Dif-em-Dif, que pode explicar a presença de efeitos individuais não-observáveis, são os efeitos defasados da variável resposta, ou seja a dinâmica entre a taxa de mortalidade de períodos anteriores no efeito Dif-em-Dif estimado, onde, um aumento da taxa de mortalidade em um período, induziria uma maior adesão no período seguinte.

4.2.3 Conclusões Contrafactuais dos Modelos Dif-em-Dif

Os resultados apresentados nas duas últimas seções apresentam evidências, tanto da dinâmica do nível de atividade econômica, com suas proxies, para arrecadação mensal e para taxa de mortalida-

⁷ Para que tenhamos uma referência do tamanho desse impacto, um megawatt-hora (Mwh) é igual a 1.000 Kilowatt-hora (Kwh). É igual a 1.000 quilowatts de eletricidade usada continuamente por uma hora. É equivalente à quantidade de eletricidade usada por cerca de 330 casas durante uma hora.

de diária do Covid-19 por UF, além do teste de vários efeitos contrafactuais. Os resultados destacam a importância da atividade comercial para arrecadação mensal dos estados brasileiro, no modelo de melhor ajuste, esse impacto médio de 18 milhões a cada unidade do índice do volume de venda, esse impacto ser de 11 a 24 milhões em média, já considerando a sazonalidade. Entre os efeitos contrafactuais estimados no painel de arrecadação, temos evidências de que estados brasileiros de maior PIB, tiveram melhor desempenho durante a pandemia, haja vista que, UF fora do grupo de maiores PIB, teriam um aumento médio de pouco mais de um bilhão de arrecadação, caso tivessem maior nível de PIB. O efeito associado a política de isolamento social, se mostrou estatisticamente significativo com sinal negativo, ou seja, temos evidências de que estados com alta adesão a essa política, tiveram consideráveis perdas de arrecadação, pois, estados que tiveram baixa adesão ao isolamento social, teriam tido redução média de arrecadação entre 146 a 456 milhões de reais, no contexto contrafactual.

Com relação ao painel de mortalidade dos estados brasileiros, temos a evidência de que o nível de atividade econômico, medido pelo consumo de energia elétrica, tem impacto positivo na taxa de mortalidade durante o período amostral, o aumento de 10.000 Megawatt-hora, gerou um aumento médio de 5 novos óbitos por 100 mil habitantes. Outro resultado esses modelos com painel de taxa de mortalidade revela que a pandemia tem forte componente de densidade demográfica, com efeito contrafactual estimado de 17 óbitos (em média) a mais em relação aos estados de baixa densidade demográfica. Um resultado contraintuitivo relacionado ao efeito Dif-em-Dif de política de isolamento social lança alertas sobre as inferências obtidas utilizando os modelos DiD. Os resultados do melhor modelo ajustado indicam que estados com baixa adesão teriam em média, 20 óbitos a mais (em média) na taxa de mortalidade, caso tivessem alta adesão ao isolamento social. Esse resultado, juntamente ao teste de Wooldridge de efeitos individuais não-observáveis, sugerem haver a natureza de outras dinâmicas nesse painel: endogeneidade e/ou efeitos defasados da taxa de mortalidade na adesão a política de isolamento social.

Em resumo, a arrecadação estadual mensal apresentou efeitos estatísticos significantes de redução nos estados com alta adesão ao isolamento social, enquanto a taxa de mortalidade apresentou forte componente demográfico, contudo em nenhum dos painéis o efeito contrafactual dos estados que adotaram a política de lockdown se mostrou estatisticamente significativo, ou seja, não há evidências de impacto sobre arrecadação, ou sobre as taxas de mortalidade. Todos esses resultados inferências contrafactuais, dependem diretamente das hipóteses dos modelos de dados painel de Diferenças-em-Diferenças. Nesse sentido, uma hipótese central é a de que os estados são diferentes entre si, mas de maneira constante, isso se traduz em não haver tendências presente na componente de heterogeneidade individual, ou seja as tendências são necessariamente paralelas, o que em muitas situações é uma hipótese extremamente restritiva. Os testes Pesaran para dependência cross-section e de Wooldridge para efeitos individuais não observáveis, além da persistente presença de heterocedasticidade e autocorrelação serial, são indícios de que as hipóteses dos modelos Dif-em-Dif, e suas inferências podem está falhando conforme Bertrand et al (2004). Outra sensível questão diz respeito aos efeitos defasados de intervenções, como por exemplo, políticas de lockdown ou isolamento social, em seus efeitos podem ocorrer de maneira lenta, criando uma estrutura de autocorrelação e demandando técnicas GMM, como a de Arellano e Bond (1991), para sua estimação, em uma literatura ainda

não consolidada para efeitos Dif-em-Dif. Caso a distribuição da variável resposta esteja mudando no tempo, e mesmo introduzindo dinâmica temporal nos modelos de dados em painel, permanece presente a necessidade de tendências paralelas e com ela todas suas dificuldades inferências. O problema de endogeneidade também pode gerar estimativas viesadas e inconsistentes, nesse caso, agravando e tornando ainda mais restritivas as hipóteses de tendências paralelas.

Uma alternativa a todas essas questões de inferência causal em modelos de dados em painel são os modelos de painel contrafactuais artificiais (ArCo) em Alta-Dimensão apresentado por Carvalho et al. (2018). A abordagem ArCo engloba o método PF proposto por Hsiao et al. (2012) e pode ser visto como uma generalização da abordagem de Controle Sintético nas mesmas linhas discutidas por Doudchenko e Imbens (2016) e Ferman e Pinto (2016). O método ArCo é mais adequado do que os estimadores de Diferenças-em-Diferenças (DiD), mesmo para estudos de caso comparativos quando há uma única unidade tratada e nenhum grupo de controle semelhante está disponível, mesmo após a inclusão de muitas variáveis de controle, ou presença de endogeneidade. Além disso, a abordagem ArCo relaxa a restrição estrita de tendência paralela dos métodos DiD. Recentemente, Gobillon e Magnac (2016) generalizam estimadores DiD considerando uma especificação correta modelo de painel linear encontrado com regressores estritamente exógenos e efeitos fixos interativos representados como uma série de fatores comuns com cargas heterogêneas. Seus resultados teóricos dependem de duplo assintóticos quando T e n vão ao infinito. Os autores permitem que os fatores de confusão comuns têm tendências determinísticas não lineares. O método ArCo difere de Gobillon e Magnac (2016) em várias direções:

- 1) O modelo não é considerado como especificado corretamente e não há necessidade de estimar os fatores comuns. A estimativa consistente de fatores precisa que tanto o a série temporal e as dimensões da seção transversal divergem ao infinito e podem ser severamente viesadas em pequenas amostras;
- 2) A metodologia ArCo requer apenas a divergência da dimensão da série temporal. Além disso, o regressores não precisam ser estritamente exógenos, o que é uma suposição irreal na maioria das aplicações com dados agregados. Tendências não lineares heterogêneas também são permitidas, mas não há necessidade de estimar (explicitamente ou por meio de fatores comuns);
- 3) Finalmente, como no caso DiD, o ArCo também não exigem que o número de unidades tratadas cresça ou tenha um grupo de controle confiável (após o condicionamento covariáveis).

Diante de todas essas questões apresentamos na próxima seção os resultados ArCo para estimar comportamentos contrafactuais nos painéis de arrecadação e de taxas de mortalidade.

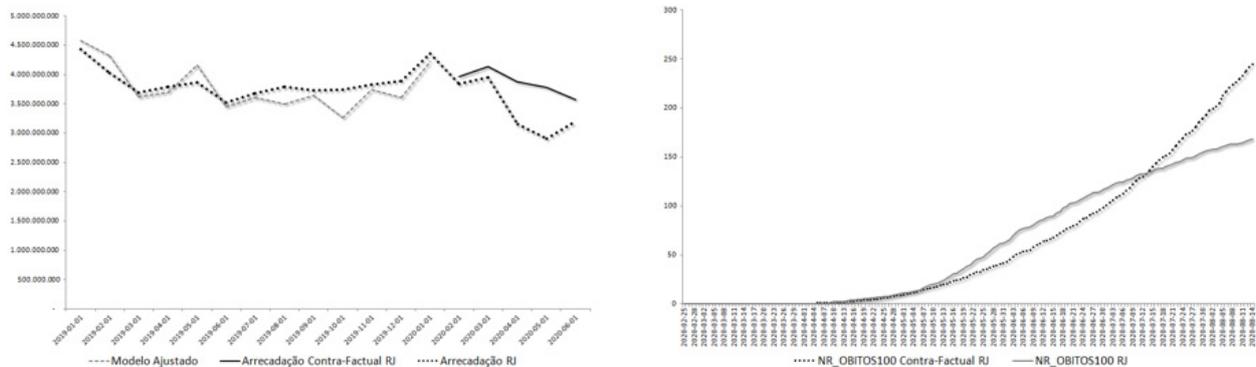
4.3 Resultados dos Modelos de Dados em Painel Contrafactuais Artificiais

Apresentamos nessa seção os resultados dos modelos de dados em painel contrafactual artificial (ArCo) para os painéis de arrecadação e taxas de mortalidade estaduais. O ponto de intervenção para o estudo da arrecadação mensal foi exogenamente estabelecido como o mês de fevereiro de 2020. Para o painel de taxas de mortalidade, a data de intervenção foi estimada endogenamente conforme apresentado na seção 3.2.1, haja vista as diferentes velocidades de evolução pandêmicas por UF, e

representa o momento no qual, existe diferença estatisticamente significativa entre as taxas de mortalidade observadas e contrafactualmente estimadas.

Inicialmente apresentamos estudos de casos para elucidar os resultados obtidos e como serão tratados. Na Figura 4 apresentamos os resultados do modelo ArCo para o RJ.

Figura 4 – Estudo de Caso: RJ

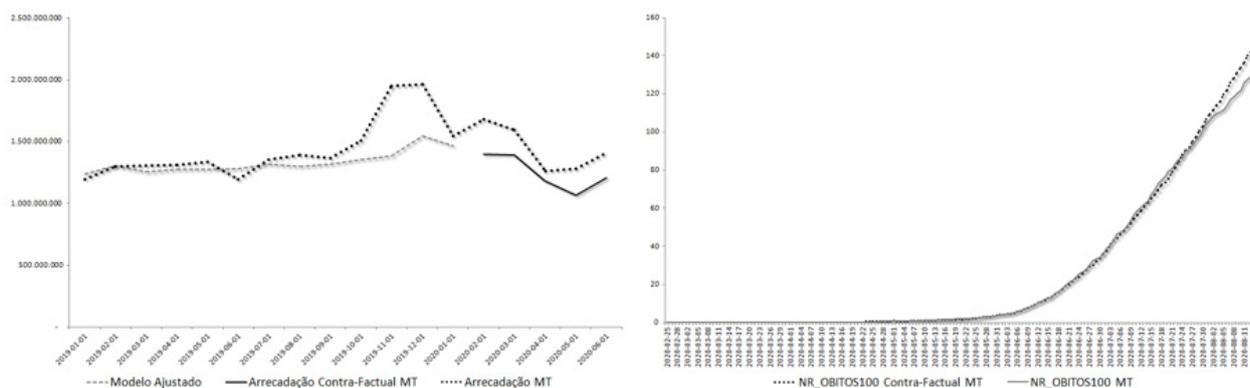


Fonte: Elaboração própria.

O caso do RJ reflete o comportamento contrafactual da arrecadação, a partir da intervenção do mês de fevereiro de 2020, acima da arrecadação observada, ou seja, em um contexto contrafactual sem pandemia, utilizando a melhor estimativa de um RJ-Contrafactual, a arrecadação do RJ teria sido significativamente maior, revelando que para o RJ a pandemia trouxe severas percas de arrecadação. O comportamento da taxa de mortalidade esteve até o dia 12/07 esteve acima da taxa de mortalidade contrafactual, e nesse trecho a pandemia foi mais severa do que seria no contexto contrafactual, contudo a partir dessa data, a taxa de mortalidade se manteve abaixo da taxa do RJ-Contrafactual, ou seja, a partir do dia 12/07, a pandemia passou a ser menos severa para o RJ, caso seguisse seu crescimento contrafactual.

Na sequência apresentamos na Figura 5 os resultados para o caso do MT, que é o contra-ponto do caso do RJ. Podemos observar graficamente que a arrecadação observada no período pandêmico foi acima da arrecadação do MT-Contrafactual a partir da intervenção, e nesse caso, o MT teria a princípio tido menores impactos da pandemia na arrecadação. Em termos de taxa de mortalidade, não observamos para o MT, quase durante todo período amostral, diferença as taxas observadas e contra-factuais, somente a partir do dia 31/07, as taxas do MT-Contrafactual começam a descolar, indicando uma pandemia menos severa para esse Estado.

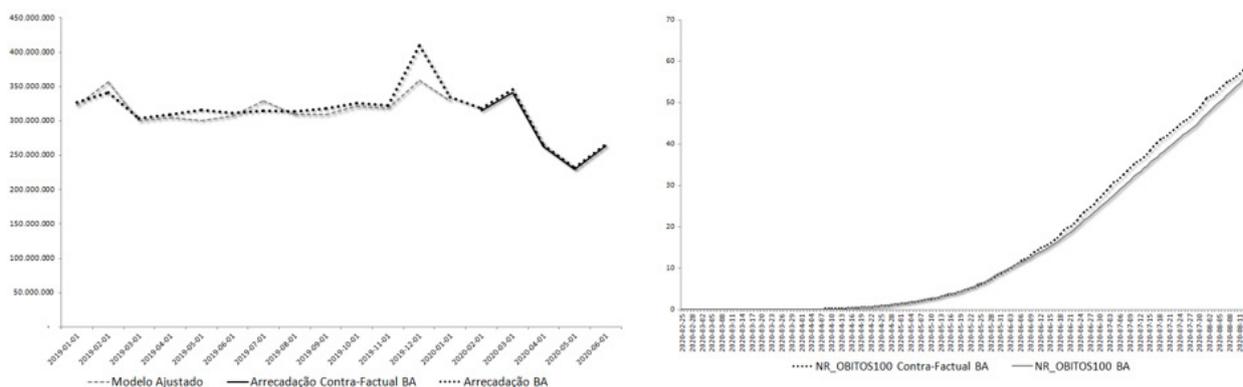
Figura 5 – Estudo de Caso: MT



Fonte: Elaboração própria.

No estudo de caso da Bahia, apresentado na Figura 6, constamos pelas estimativas Arco que a arrecadação não foi, estatisticamente, maior ou menor, do que seria em um contexto contrafactual, e em termos gravidade pandêmica, os valores contrafactuais estiveram ligeiramente acima dos valores observados a partir do dia 11/06, indicando que a pandemia nesse estado, foi também, ligeiramente menos severa.

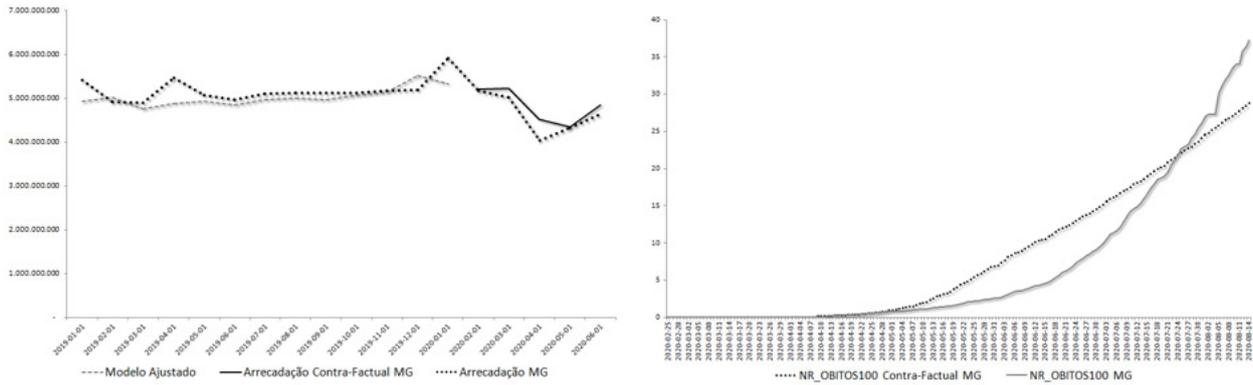
Figura 6 – Estudo de Caso: BA



Fonte: Elaboração própria.

Existe várias possibilidade de combinações de impactos de arrecadação e gravidade pandêmica, sendo que cada Estado, apresentará sua própria peculiaridade durante o período amostral. O detalhamento de cada Estado, por si mesmo, seria objeto de um estudo próprio. Contudo, ressaltamos os estudos de caso de MG e SP, por serem os dois estados brasileiros de maior arrecadação absoluta, conforme detalhado na Tabela 3. Podemos observar na Figura 7, que o Estado de MG, a exemplo do RJ, tem valores de arrecadação, acima da arrecadação de MG-Contrafactual, o que pode se traduzir, em mais um dos estados com grande impacto negativo de arrecadação. Do ponto de vista pandêmico, o estado de MG esteve boa parte do período amostral abaixo do valor contrafactual, contudo a partir do dia 24/07, observamos uma aceleração da taxa de mortalidade acima de seu valor contrafactual, o que indica que a pandemia se tornou mais severa a partir dessa data.

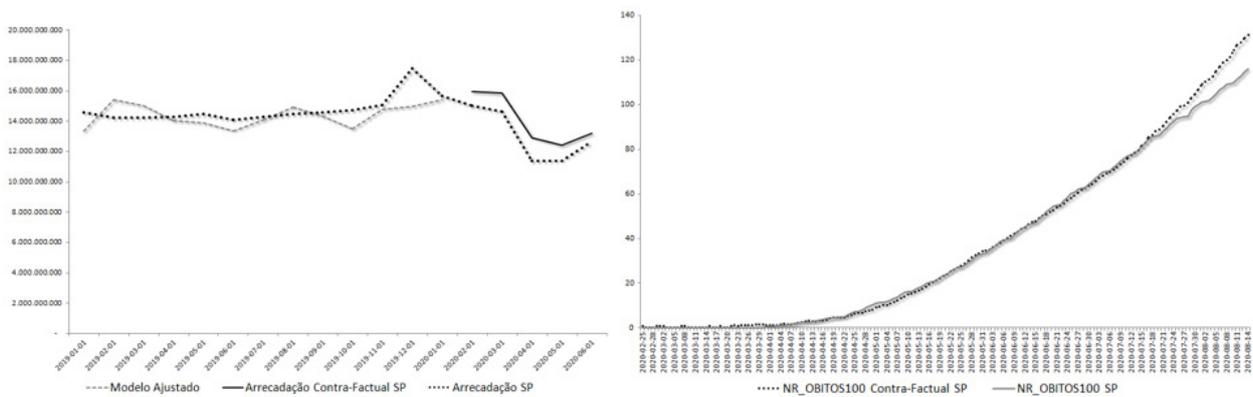
Figura 7 – Estudo de Caso: MG



Fonte: Elaboração própria.

Apresentamos na Figura 8 o estudo de caso de SP, a exemplo de MG e RJ, teve valores de arrecadação abaixo do que seria em um contexto contrafactual, e dado sua representatividade em termos de arrecadação nacional, seu impacto da pandemia em termos de arrecadação, foi o maior. Em termos de taxa de mortalidade, o Estado de SP, esteve na maior parte do período amostral, muito próximo entre taxas de mortalidade observada e contrafactual, o descolamento ocorre após o dia 18/07, quando os valores da taxa de SP-Contrafactual acelera-se, e, portanto, a partir dessa data, verificamos uma menor severidade da pandemia.

Figura 8 – Estudo de Caso: SP



Fonte: Elaboração própria.

Os resultados apresentados nesses estudos de caso mostram a heterogeneidade dos impactos pandêmicos nas arrecadações estaduais e também as diferentes evoluções pandêmicas por UF. Para termos ideia desse impacto em termos contrafactuals utilizaremos duas medidas: Diferença Contrafactual (DCF) e a Diferença Contrafactual Percentual (DCFP). A primeira medida é a diferença entre a variável contrafactual e seu valor observado, enquanto que a segunda considera essa diferença em modulo, divida pelo valor observado, permitindo termos uma ideia percentual da diferença contrafactual. Enquanto a média da DCFP permite termos uma medida de impacto total, sem discriminar se o contrafactual esteve acima ou abaixo do observado, a média do DCF fornece a ideia de predominância de estado contrafactual. Apresentamos na Tabela 18 e 19 os resultados dessas medidas para as 27 UF, ranqueadas da maior média de DCF de arrecadação para menor.

Tabela 18 - Impactos Pandêmicos Contrafactuais Estimados do Covid-19 por UF

UF	I_iso	I_lock-down	Arrecadação Média Contra-factual - período pandêmico	Arrecadação Média - período pandêmico	Arrecadação Média fev/2019 a jun/2019	Número de Óbitos Contrafactual por 100 mil/hab. - fev/20 a jun/20 (Média)	Número de Óbitos por 100 mil/hab. - fev/20 a jun/20 (Média)	Número de Óbitos Contrafactual por 100 mil/hab. (Média)	Número de Óbitos por 100 mil/hab. (Média)	Média de DCF de Arrecadação
SP	0	0	14.083.004.476	13.043.716.280	14.285.223.260	17,69	17,62	37,97	36,66	1.039.288.196
RJ	1	1	3.862.599.233	3.408.064.641	3.778.213.948	23,29	30,54	61,17	60,43	454.534.592
MG	0	0	4.823.427.609	4.638.169.079	5.065.035.909	3,49	1,70	8,24	6,97	185.258.530
CE	1	1	1.075.034.627	1.013.387.755	1.268.607.493	34,73	35,42	77,97	68,00	61.646.872
DF	1	0	846.771.113	790.269.364	831.397.940	7,80	6,99	32,57	26,55	56.501.749
MA	1	1	680.025.842	652.389.353	700.339.769	14,08	14,22	29,08	30,91	27.636.489
ES	0	0	1.045.954.533	1.019.272.995	1.100.205.868	17,00	17,32	42,57	42,85	26.681.538
AL	1	0	360.726.969	346.222.065	382.763.013	14,61	14,21	34,93	33,04	14.504.904
PI	1	1	382.974.294	373.560.448	403.835.095	4,31	6,90	14,50	23,67	9.413.846
AP	1	1	96.880.118	90.904.959	108.525.059	26,63	25,71	62,15	51,69	5.975.159
RN	0	1	468.331.822	463.894.723	511.513.210	8,07	10,63	23,76	32,13	4.437.099
BA	0	0	282.431.402	285.872.278	316.898.516	5,27	5,02	15,42	14,45	-3.440.876
SE	0	0	282.431.402	285.872.278	316.898.516	8,84	9,27	32,17	34,75	-3.440.876
TO	0	1	258.849.905	266.340.656	270.006.497	5,12	5,26	14,18	14,93	-7.490.752
PB	0	0	476.617.518	488.810.145	531.278.479	10,86	10,41	28,28	28,32	-12.192.627
AC	1	0	97.090.969	109.629.719	118.718.836	16,71	18,37	44,20	42,01	-12.538.751
RR	0	0	98.255.633	114.307.254	112.417.924	12,65	20,58	33,27	54,66	-16.051.621
RO	1	0	429.499.992	447.634.715	460.699.961	10,19	11,33	26,07	30,82	-18.134.723
PE	1	1	1.390.580.295	1.429.409.575	1.574.603.331	36,72	28,16	71,75	54,33	-38.829.280
AM	1	1	861.810.572	907.835.174	893.765.124	46,58	45,52	82,06	73,60	-46.024.602
RS	1	0	2.937.456.125	2.999.974.113	3.243.057.391	3,21	2,30	8,64	8,69	-62.517.988
PA	1	1	1.179.434.964	1.266.561.811	1.254.469.014	24,18	30,93	59,59	56,51	-87.126.847
SC	0	0	1.959.003.523	2.081.635.333	2.259.346.778	2,72	2,22	11,03	8,10	-122.631.810
MS	0	1	900.654.874	1.033.274.147	951.107.584	1,01	0,99	4,00	6,28	-132.619.272
MT	0	0	1.252.394.123	1.445.877.552	1.289.483.003	4,14	4,15	25,80	25,18	-193.483.429
GO	0	0	1.282.694.014	1.527.685.647	1.618.987.955	2,08	2,36	10,11	11,50	-244.991.633
PR	0	1	2.576.657.594	2.964.979.362	3.166.321.838	2,07	2,27	9,66	8,85	-388.321.768

Fonte: Elaboração própria.

Tabela 19 - Impactos Pandêmicos Contrafactuais Estimados do Covid-19 por UF- Diferença Contrafactual.

UF	I_iso	I_lockdown	Média de DCF de Arrecadação	Média de DCFP de Arrecadação	Média de DCF de Nr. Óbitos por 100mil/hab.	Média de DCFP de Nr. Óbitos por 100mil/hab.	Arrecadação Média fev/2019 a jun/2019	Número de Óbitos
SP	0	0	1.039.288.196	0,0815	-1,2286	3,7248	14.285.223.260	26.625
RJ	1	1	454.534.592	0,1448	-0,7407	0,1588	3.778.213.948	14.514
MG	0	0	185.258.530	0,0422	-1,2650	0,4668	5.065.035.909	3.945
CE	1	1	61.646.872	0,0948	-9,8995	0,2981	1.268.607.493	8.127
DF	1	0	56.501.749	0,0890	-6,0045	0,1167	831.397.940	1.936
MA	1	1	27.636.489	0,0611	1,8361	0,2348	700.339.769	3.241
ES	0	0	26.681.538	0,1191	0,3041	0,0737	1.100.205.868	2.850
AL	1	0	14.504.904	0,0422	-1,8702	0,1424	382.763.013	1.733
PI	1	1	9.413.846	0,0739	9,1650	0,2131	403.835.095	1.582
AP	1	1	5.975.159	0,0946	-10,4585	0,1212	108.525.059	612
RN	0	1	4.437.099	0,0393	8,3633	0,1524	511.513.210	2.042
BA	0	0	-3.440.876	0,0117	-0,9635	0,0973	316.898.516	4.273
SE	0	0	-3.440.876	0,0117	2,5939	0,1031	316.898.516	1.685
TO	0	1	-7.490.752	0,0524	0,7643	0,0626	270.006.497	500
PB	0	0	-12.192.627	0,0478	0,0372	0,0739	531.278.479	2.114
AC	1	0	-12.538.751	0,1150	-2,1686	0,0893	118.718.836	576
RR	0	0	-16.051.621	0,1395	21,3955	0,3659	112.417.924	565
RO	1	0	-18.134.723	0,0416	4,7485	0,1283	460.699.961	1.001
PE	1	1	-38.829.280	0,0557	-17,4170	0,2321	1.574.603.331	7.114
AM	1	1	-46.024.602	0,0829	-8,4517	0,1274	893.765.124	3.457
RS	1	0	-62.517.988	0,0405	0,0478	0,1923	3.243.057.391	2.632
PA	1	1	-87.126.847	0,0701	-3,0688	0,3355	1.254.469.014	5.927
SC	0	0	-122.631.810	0,0570	-2,9255	0,2449	2.259.346.778	1.743
MS	0	1	-132.619.272	0,1286	2,2859	0,1382	951.107.584	591
MT	0	0	-193.483.429	0,1319	-0,5965	0,1084	1.289.483.003	2.286
GO	0	0	-244.991.633	0,1581	1,3879	0,1148	1.618.987.955	2.287
PR	0	1	-388.321.768	0,1329	-0,8059	0,0838	3.166.321.838	2.613

Nota: O número de óbitos por UF referente a última data do painel de pandemia Covid-19: 14/08/2020 - Total: 106.571 (segunda vintage). | Fonte: Elaboração própria.

Os resultados da Tabela 18 e 19 resumem os impactos contrafactuais da pandemia Covid-19 na arrecadação estadual além das evoluções pandêmicas contrafactuais por UF. Os resultados da Tabela 19 mostram que para 11 estados (SP, RJ, MG, CE, DF, MA, ES, AL, PI, AP e RN), a DCF de arrecadação foi positiva, o que indica que os valores observados de arrecadação estiveram, na maior parte do período amostral, acima de seus respectivos valores contrafactuais. Para esses estados, a perda de arrecadação em razão da pandemia do Covid-19 foi além da simples queda verificada nos valores observados, fato que foi comum a grande maioria de todos os estados brasileiros conforme apresentado na Tabela 18. Para esses estados, o impacto Covid-19 em arrecadação pode ser considerado mais grave no sentido que mesmo os estados, que sem o contexto pandêmico, teriam quedas em suas arrecadações, observaram quedas de arrecadação, além do que seria essa sua queda natural de arrecadação. Os demais estados em que a DCF foi negativa, a interpretação se inverte, ou seja, para esses estados, a eventual queda de arrecadação, esteve aquém do seria sua queda natural de arrecadação nas condições do Covid-19.

Quando considerando as maiores Diferenças Contrafactuais positivas, temos a sequência dos estados de SP, RJ e MG, que são respectivamente, as três maiores arrecadações nacionais. Para esses estados, o impacto da pandemia na arrecadação, pode ser considerado severo. Em contrapartida, em termos de DCF da taxa de mortalidade, apresentam valores negativos (nas posições: 11º, 14º e 10º respectivamente). Ou seja, para esses estados a pandemia, em termos contrafactuais, foi severa, contudo não estiveram entre as mais graves pandemias em relação a outros entes subnacionais, apesar de SP e RJ terem os maiores números absolutos de óbitos nacionais (26.625 e 14.514 na data de 14/08), conforme a Tabela 19. Quando analisamos a média de DCF da taxa de mortalidade dos estados com DFC de arrecadação positiva, na maior parte dos casos, apresentam DFC de mortalidade negativas, o que indica a taxa de mortalidade observada, esteve (ainda que ligeiramente para alguns estados) acima da taxa de mortalidade contrafactual, ou seja, a pandemia Covid-19 foi individualmente severa. A maior DCF de arrecadação positiva refere-se ao Estado de SP. Segundo os dados da Tabela 18, SP observou uma queda média de arrecadação de 1.241 bilhões de reais no período pandêmico, quando contrafactual estimado indica um valor (em média) de 1.039 bilhões de reais maior do que esse valor observado, ou seja, em um contexto não pandêmico, esperávamos um valor significativamente maior de arrecadação em mais de 200 milhões de reais.

A maior DCF de arrecadação negativa foi observada para o Estado do PR, pela Tabela 18, o PR teve uma queda média de pouco mais de duzentos milhões de reais, essa queda deveria ter sido maior, em cerca de 388 milhões de reais, segundo o PR-Contrafactual. Quando consideramos o DCF da taxa de mortalidade do PR, temos um valor negativo e dessa forma, temos uma evidência de uma pandemia Covid-19 individualmente severa, contudo observamos que o PR não esteve entre as mais graves pandemias regionais. Quando comparamos com PE, por exemplo, observamos que esse Estado teve o maior DCF negativa da taxa de mortalidade, o que indica que esse Estado sofreu a mais severa pandemia individual, a taxa de mortalidade, esteve muito acima do número que seria seu PE-Contrafactual, não obstante, esse Estado teve DCF negativa de arrecadação, que por outro lado, indica que sua arrecadação não caiu, como naturalmente cairia, o que segundo a Tabela 18, representou uma queda de 9,22%, equivalente a 145 milhões de reais, quando contrafactualmente, deveria ser uma

queda (em média) 38 milhões de reais acima do valor. Outro destaque da Tabela 19, mostra o Estado do RR, com a maior DCF de mortalidade positiva, o que indica que o número de óbitos por 100 mil/hab, deveria ser muito acima da taxa de mortalidade observada, com o número total de 564 óbitos registrados, contudo, um problema de subnotificação pode ser responsável por esse efeito.

Quando consideramos a DCF positiva de arrecadação na Tabela 19, temos o conjunto de estados que sofreram os maiores impactos pandêmicos, por outro lado quando observamos nessa mesma Tabela, o conjunto de estados que tiveram uma DCF negativa temos o conjunto de estados que sofreram pandemias mais severas. Dessa forma temos um primeiro conjunto formado por estados com DCF de arrecadação positiva e DCF de mortalidade negativas: SP, RJ, MG, CE, DF, AL e AP, esses estados tiveram grande impacto de arrecadação e também sofreram pandemias mais severas do que seus respectivos contrafactuais. Um segundo grupo é formado por estados com DCF de arrecadação positiva, e com DCF de mortalidade também positivas: ES, PI, MA e RN, esses estados tiveram grande impacto de arrecadação, contudo sofreram pandemias menos severas do que seus respectivos contrafactuais. O terceiro grupo é formado por estados que observamos DCF negativas de arrecadação e DCF de mortalidade também negativas: BA, AC, PE, AM, PA, SC, MT e PR, nesses estados, observamos um menor impacto na arrecadação no período pandêmico concomitante com pandemias mais severas considerando seus respectivos contrafactuais. O último grupo é formado por estados com DCF de arrecadação negativas e com DCF de mortalidade positivas: SE, TO, PB, RR, RO, RS, MS e GO, nesses estados tiveram menores impactos sobre arrecadação e pandemias menos severas, considerando seus respectivos valores contrafactuais.

Em termos absolutos os maiores impactos de arrecadação (DCFP) foram observados no GO (15%) e RJ (14%), como o DCF de arrecadação positiva do RJ, temos a evidência de que esse Estado tenha enfrentado o maior impacto negativo de arrecadação contrafactual, enquanto o Estado de GO, com DCF de arrecadação negativa, indica ter sido o Estado de menor impacto negativo de arrecadação contrafactual. Em termos de pandemia, os maiores impactos absolutos foram observados em SP e MG, com DCF negativa, indicando uma grande severidade pandêmica contrafactual nesses estados, tendo como contrapartida o Estado de RR, com impacto de 36%, e com DCF de mortalidade positiva, o que indica que esse Estado tenha vivenciado uma menor severidade pandêmica contrafactual. Considerando o impacto total de arrecadação observamos que a soma total de DCF apresenta valor positivo, indicando que de modo geral houve um grande impacto de arrecadação no país. Considerando da mesma forma toda DCF de mortalidade, observamos um valor negativo, o que indica que a pandemia foi severa de maneira geral no país.

Para obtermos evidências finais dos impactos da pandemia na arrecadação, bem como diagnosticar os efeitos de políticas de isolamento social e lockdown na pandemia estadual, foram estimados dois painéis: Painel de efeitos fixos com a DCF de arrecadação dos estados tendo a DCF de mortalidade como covariável além de testar os efeitos de isolamento social e lockdown considerando apenas o período pandêmico (de fevereiro a junho de 2020); Painel de efeitos fixos com a média de DCF de mortalidade por UF para os meses pandêmicos (de fevereiro a junho de 2020) tendo como covariável a respectiva taxa de isolamento e a indicadora de lockdown. Os resultados desses dois painéis estimados são apresentados respectivamente nas Tabelas 20 e 21.

Tabela 20 - Modelo de Dados em Painel para Arrecadação Estadual e efeitos Covid-19

Modelo de Efeitos Fixos (LSDV) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
DCFObitos	-90343	54091	-1.671	0.09706 .
I_iso	39353870	21423725	1.836	0.06829 .
I_lockdown	-7466091	60046747	-0.124	0.90128
factor(UF)AC	50392044	85174463	0.592	0.55536
factor(UF)AL	77106320	84927612	0.908	0.36599
factor(UF)AM	23937433	104004458	0.230	0.81841
factor(UF)AP	75958276	104002575	0.730	0.46679
factor(UF)BA	-3479525	60048642	-0.058	0.95390
factor(UF)CE	104666982	108343305	0.966	0.33621
factor(UF)DF	119038244	84917603	1.402	0.16389
factor(UF)ES	26736733	60051963	0.445	0.65706
factor(UF)GO	-244946268	60049849	-4.079	8.78e-05 ***
factor(UF)MA	97786680	104026473	0.940	0.34935
factor(UF)MG	184975632	60216406	3.072	0.00270 **
factor(UF)MS	-125156043	84918230	-1.474	0.14349
factor(UF)MT	-193477577	60045520	-3.222	0.00169 **
factor(UF)PA	-15937872	105708160	-0.151	0.88044
factor(UF)PB	-12260144	60055197	-0.204	0.83863
factor(UF)PE	29949093	105854903	0.283	0.77779
factor(UF)PI	79954209	104350239	0.766	0.44525
factor(UF)PR	-380823790	84917164	-4.485	1.86e-05 ***
factor(UF)RJ	525806996	105955626	4.963	2.67e-06 ***
factor(UF)RN	12310671	85138475	0.145	0.88530
factor(UF)RO	44709470	85077892	0.526	0.60032
factor(UF)RR	-14795879	63328951	-0.234	0.81572
factor(UF)RS	-103920411	90016741	-0.115	0.90823
factor(UF)SC	-122710721	60058766	-2.043	0.04352 *
factor(UF)SE	-3370319	60056095	-0.056	0.95535
factor(UF)SP	1039294371	60045528	17.308	< 2e-16 ***

factor(UF)TO	41709561	84076812	0.496	0.62059
OBS				135
R2				0.8095
R2 ajustado				0.7592
Estatística F				16.09 com 28 e 106 GL, p-valor: < 2,22e- 16

Breusch-Pagan studentizado: BP = 3.673, GL = 3, P-valor = 0.299

Teste Breusch-Godfrey/Wooldridge: chisq = 3.992, GL = 4, p-value = 0.407

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Os resultados da tabela 20 apresentam significância estatística global reportada pelo teste F com baixo P-valor, com R2 ajustado de 75%. O parâmetro de DFC de mortalidade se mostrou estatisticamente significativa a 10% com sinal negativo, ou seja, quanto mais severa a pandemia, maior o impacto na arrecadação em termos contrafactuais. A cada aumento da diferença negativa (ou seja, da pandemia mais severa, pois os valores observados estão, nesse caso, acima do contrafactual) contrafactual de um óbito, gera em média um aumento de 90 mil reais na diferença contrafactual de arrecadação. Como a DCF positiva representa que os valores de arrecadação contrafactuais estão acima dos valores observados, temos a evidência de que quanto mais severa a pandemia, maiores foram os impactos na arrecadação. Nesse mesmo sentido, também se mostrou estatisticamente significativa variável I_iso, mostrando o grande impacto do isolamento social na DCF de arrecadação, estados com grande adesão ao isolamento social, tiveram um impacto médio 39 milhões de reais maior em relação aos estados de menor adesão, como DCF positiva, temos a evidência do grande impacto na arrecadação dos estados. Por outro lado, não há evidência estatisticamente significativa de efeitos do lockdown na DCF de arrecadação. Todas as inferências são válidas, em particular pela ausência de heterocedasticidade e autocorrelação serial verificadas pelos testes de Breusch-Pagan e Breusch-Godfrey/Wooldridge respectivamente. Apresentamos na Tabela 21 os resultados relativos à severidade pandêmica.

Tabela 21 - Modelo de Dados em Painel para Mortalidade Contrafactual por UF

Modelo de Efeitos Fixos (LSDV) - Parâmetros Estimados				
	Estimativa	Desv.Pad	t valor	Pr(> t)
Taxa_iso	0.76656	0.42691	1.795	0.074670
I_lockdown	0.13121	1.92691	0.068	0.945839
factor(UF)AC	2.19634	2.72506	0.806	0.422042
factor(UF)AL	0.44388	2.72506	0.163	0.870913
factor(UF)AM	-0.24850	3.33751	-0.074	0.940787

factor(UF)AP	-0.13633	3.33751	-0.041	0.967494
factor(UF)BA	-0.20563	1.92691	-0.107	0.915215
factor(UF)CE	1.50447	3.47379	0.433	0.665820
factor(UF)DF	0.09847	2.72506	0.036	0.971243
factor(UF)ES	0.29367	1.92691	0.152	0.879156
factor(UF)GO	0.24136	1.92691	0.125	0.900553
factor(UF)MA	0.75259	3.33751	0.225	0.822024
factor(UF)MG	-1.50516	1.92691	-0.781	0.436452
factor(UF)MS	-0.14643	2.72506	-0.054	0.957245
factor(UF)MT	0.03114	1.92691	0.016	0.987137
factor(UF)PA	6.27946	3.33751	1.881	0.062624 .
factor(UF)PB	-0.35922	1.92691	-0.186	0.852464
factor(UF)PE	-6.54619	3.33751	-1.961	0.052430 .
factor(UF)PI	2.82851	3.33751	0.847	0.398612
factor(UF)PR	0.03845	2.72506	0.014	0.988769
factor(UF)RJ	6.72335	3.33751	2.014	0.046467 *
factor(UF)RN	2.03680	2.72506	0.747	0.456442
factor(UF)RO	1.73558	2.72506	0.637	0.525554
factor(UF)RR	6.68120	1.92691	3.467	0.000758 ***
factor(UF)RS	-0.40791	1.90581	-0.214	0.830843
factor(UF)SC	-0.41985	1.92691	-0.218	0.827933
factor(UF)SE	0.37540	1.92691	0.195	0.845903
factor(UF)SP	0.03286	1.92691	0.017	0.986428
factor(UF)TO	1.56728	2.58229	0.606	0.544914
OBS				135
R2				0.8095
R2 ajustado				0.7592
Estatística F				16.09 com 28 e 106 GL, p-valor: < 2,22e- 16

Breusch-Pagan studentizado: BP = 3.671, GL = 3, P-valor = 0.299

Teste Breusch-Godfrey/Wooldridge: chisq = 3.991, GL = 4, p-value = 0.407

Nota: 0 '***' 0,001 '**' 0,01 '*' 0,05 '.' 0,1 ' ' 1

Fonte: Elaboração própria.

Os resultados da Tabela 21 tem validade global estatisticamente significativa pelo teste F com baixo P-valor, com R2 ajustado de 25%. A taxa de isolamento se mostrou estatisticamente significativa, com valor positivo, o que se traduz como: a cada aumento unitário na taxa de isolamento social, tem um impacto médio de 0,77 na DCF de mortalidade, como no caso positivo dessa diferença, representa uma menor severidade pandêmica, temos a evidência que uma maior taxa de isolamento, está associado a uma pandemia menos severa em termos contrafactuais, ou seja, o número de óbitos por 100 mil/hab. contrafactual maior do que o número de óbitos por 100 mil/hab. observado. Por outro lado, a política de lockdown não se mostrou estatisticamente significativa, ou seja, a adoção dessa política não tem efeitos na diferença contrafactual, nem positiva, nem negativamente. Todas essas inferências se mostram válidas, em particular, pela ausência de heterocedasticidade e autocorrelação serial verificadas pelos testes de Breusch-Pagan e Breusch-Godfrey/Wooldridge respectivamente.

5. Conclusão

Nesse artigo investigamos como a pandemia do Covid-19 afetou a arrecadação dos entes subnacionais. A proposta desse estudo teve como linha principal de investigação a inferência contrafactual tanto da arrecadação dos estados quanto da sua respectiva dinâmica pandêmica. Para essa finalidade construímos dois painéis de estudo, o primeiro com os dados de arrecadação e atividade comercial, já com ajuste sazonal, com dinâmica mensal por UF e o segundo painel contendo dados de mortalidade por 100 mil/hab. e consumo de energia elétrica, com dinâmica diária por UF. O período amostral compreendeu o período de janeiro de 2000 a junho de 2020 para o painel de arrecadação, perfazendo uma amostra de 246 observações temporais, enquanto o painel pandêmico compreendeu o período do dia 22/02/20 a 14/08/20, perfazendo uma amostra de 172 observações temporais.

O primeiro passo foi à construção de um conjunto de variáveis indicadoras considerando as características das UFs como PIB e densidade demográfica, além da adoção de políticas de lockdown, de adesão ao isolamento social e de gravidade pandêmica. Uma completa análise exploratória detalha a relação entre as variáveis utilizadas no estudo. O estudo utilizou duas abordagens de inferência causal: modelos de dados em painel de Diferenças-em-Diferenças (DiD) e os modelos de painel contrafactuais artificiais (ArCo).

As conclusões gerais dos modelos DiD estimados depreende que aumento da atividade comercial tem grande impacto na arrecadação estadual (o aumento de uma unidade de volume de vendas do comércio, gera um aumento médio de 18 milhões na arrecadação). Dentre os efeitos contrafactuais estimados, a alta adesão a políticas de isolamento social e estados com PIB mais altos, apresentaram efeitos estatisticamente significantes. Dessa forma, estados que apresentaram alta adesão à política de isolamento social tiveram um impacto negativo de arrecadação, em média, de 300 milhões, podendo esse impacto ser de 146 a 456 milhões, em relação aos estados que não tiveram alta adesão, de maneira contrafactual. Os estados com PIB abaixo da mediana nacional teriam um aumento médio de quase um bilhão de reais na arrecadação durante a pandemia, caso fossem do conjunto de UF de maiores PIBs. Os efeitos contrafactuais da adoção de políticas de lockdown não se mostraram estatisticamente significantes no painel Dif-em-Dif de arrecadação. Com relação ao painel com dados pandêmicos diários, foram encontradas evidências de que um aumento de 10.000 MegaWatt-horas, gerou um au-

mento (em média) de 5 óbitos na taxa de mortalidade por 100 mil/hab. Dois efeitos contrafactuais se mostraram estatisticamente significantes, a densidade demográfica e a alta adesão a políticas de isolamento social. Dessa forma estados com grande densidade demográfica, tiveram (em média) 17 óbitos por 100 mil/hab. a mais, caso não tivessem essa densidade demográfica, o que ressalta a importante característica demográfica da pandemia. Novamente os efeitos contrafactuais associados a política de lockdown não foram estatisticamente significantes. Os resultados do melhor modelo ajustado indicam que estados com baixa adesão teriam em média, 20 óbitos a mais (em média) na taxa de mortalidade, caso tivessem alta adesão ao isolamento social. Esse resultado contraintuitivo revela a possibilidade de endogeneidade entre a taxa de mortalidade e adesão ao isolamento social nos estados, uma das fragilidades dos modelos de Painel Dif-em-Dif. Portanto todas essas inferências DiD são sensíveis a um conjunto de hipóteses, dentre essas, destaca-se a hipótese de que os estados são diferentes entre si, mas de maneira constante, isso se traduz em não haver tendências presente na componente de heterogeneidade individual, ou seja, as tendências são necessariamente paralelas, o que em muitas situações é uma hipótese extremamente restritiva, como alternativa a essas (e outras) fragilidades dos modelos DiD foram estimados os modelos para painéis contrafactuais artificiais (ArCo).

Os resultados dos modelos ArCo permitiram, individualmente por Estado, construir sua respectiva arrecadação contrafactual, bem como sua pandemia contrafactual. Com esses resultados foi possível observar o comportamento dos painéis a partir de uma data de intervenção (fixada exogenamente como o mês de fevereiro de 2020 para o painel de arrecadação e endogenamente estimada para o painel pandêmico, dado as diferentes evoluções regionais do Covid-19). Os resultados revelaram que os impactos foram heterogêneos nos estados, arrecadações como as de SP e do RJ, se mostraram abaixo de seus valores contrafactuais, ou seja, para esses estados, a arrecadação teria sido maior em um contexto contrafactual, e portanto sofreram grandes impactos na arrecadação. Outros estados, como o PR e GO, apresentaram arrecadação contrafactual abaixo de seus valores observados, ou seja, a arrecadação desses estados teria sido menor no contexto contrafactual, portanto sofreram impactos menores em termos de arrecadação. Houveram ainda estados como BA e RN, em que quase não há distinção entre os valores observados de arrecadação e seus contrafactuais, e dessa forma a arrecadação seguiu sua trajetória natural. Em uma análise similar, verificamos nesse estudo, estados como o PI e RN, onde o número de óbitos por 100 mil/hab. esteve abaixo de seus respectivos valores contrafactuais, e, portanto, enfrentaram pandemias menos severas, ao passo que estados como PE e AP, observamos valores contrafactuais da taxa de mortalidade muito abaixo de seus respectivos valores observados, e portanto a severidade pandêmica foi alta. Temos ainda casos em que os valores das taxas de mortalidade, contrafactuais e observadas, praticamente não apresentaram diferença, como por exemplo, os estados de SP, RJ, BA e PR, ou seja, para esses estados, a pandemia seguiu seu curso natural.

Para dimensionar os impactos da Covid-19 a partir da data de intervenção, foram construídas duas medidas de diferença entre valores contrafactuais e valores observados de arrecadação e taxa de mortalidade estadual: Diferença Contrafactual (DCF) média, como a diferença simples, e a Diferença Contrafactual Percentual (DCFP) média, que considera o valor absoluto dessa diferença dividida pelo valor observado, medindo impactos absolutos. Quando consideramos o DCF (positiva ou negativa)

para arrecadação e taxa de mortalidade, foi possível identificar quatro grupos de estados. O primeiro grupo formado com estados com DCF de arrecadação positiva e DCF de mortalidade negativas: SP, RJ, MG, CE, DF, AL e AP, esses estados tiveram grande impacto de arrecadação e também sofreram pandemias mais severas do que seus respectivos contrafactuais. O segundo grupo formado com estados com DCF de arrecadação positiva e DCF negativa de taxa de mortalidade: ES, PI, MA e RN, esses estados tiveram grande impacto de arrecadação, contudo sofreram pandemias menos severas do que seus respectivos contrafactuais. O terceiro grupo, formado com estados com DCF de arrecadação negativa e DCF da taxa de mortalidade negativa: BA, AC, PE, AM, PA, SC, MT e PR, nesses estados, observamos um menor impacto na arrecadação no período pandêmico concomitante com pandemias mais severas considerando seus respectivos contrafactuais. E por fim, um grupo de estados formados por DCF de arrecadação negativa e DCF da taxa de mortalidade positiva: SE, TO, PB, RR, RO, RS, MS e GO, nesses estados tiveram menores impactos sobre arrecadação e pandemias menos severas, considerando seus respectivos valores contrafactuais.

A fim de elucidar as questões chaves do estudo, foram estimados dois últimos painéis mensais para o período pandêmico considerando as medidas de impacto pandêmico contrafactual (DCF). O primeiro painel de efeitos fixos apresentou evidências de que quanto maior a severidade pandêmica, maiores foram os impactos em arrecadação, bem como, quanto maior a adesão à política de isolamento social maiores foram os impactos negativos na arrecadação, contudo não houve evidência estatisticamente significativa de impactos na arrecadação da adoção de políticas de lockdown. O segundo painel de efeitos fixos apresentaram evidências de que quanto maior as taxas de isolamento social, menos severas foram às pandemias regionais, contudo a adoção de políticas de lockdown não apresentou significância estatística. Os resultados inferenciais causais dos métodos ArCo e Dif-em-Dif convergem em relação as duas políticas principais de combate ao Covid-19. Uma grande adesão a políticas de isolamento social apresentaram evidências de grandes impactos na arrecadação e também na taxa de mortalidade estadual, por outro lado, não houve nenhuma evidência estatística de efeitos da adoção de políticas de lockdown, muito embora não tenham apresentado efeitos sobre arrecadação estadual, outras variáveis não relacionadas no estudo, como taxa de desemprego, falências e outras patologias não diagnosticadas podem ter tido grande elevação em função da adoção da política de lockdown. O estudo conclui que os impactos Covid-19 na arrecadação foram heterogêneos, assim como também apresentou evidências de que as pandemias regionais foram heterogêneas, com estados sofrendo maiores e menores impactos em arrecadação, bem como, enfrentando pandemias mais severas do que outros estados. A política voluntária de adesão ao isolamento social, foi efetivo para redução da taxa de mortalidade com custo em arrecadação dos entes subnacionais, mas nenhum efeitos foi observado com políticas deliberativas dos governos subnacionais. Dessa forma o estudo contribui com a literatura como um pioneiro estudo dos impactos contrafactuais da pandemia do Covid-19 para arrecadação estadual.

Referências

Abadie, A., Diamond, A. and Hainmueller, J.: 2010, Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program, *Journal of the American statistical Association* 105(490), 493–505.

Abadie, A., Diamond, A. and Hainmueller, J.: 2015, Comparative politics and the synthetic control method, *American Journal of Political Science* 59(2), 495–510.

Abadie, A. and Gardeazabal, J.: 2003, The economic costs of conflict: A case study of the basque country, *American economic review* 93(1), 113–132.

Al-Awadhi, A. M., Al-Saifi, K., Al-Awadhi, A. and Alhamadi, S.: 2020, Death and contagious infectious diseases: Impact of the covid-19 virus on stock market returns, *Journal of Behavioral and Experimental Finance* p. 100326.

Albulescu, C.: 2020a, Coronavirus and financial volatility: 40 days of fasting and fear, arXiv preprint arXiv:2003.04005.

Albulescu, C. T.: 2020b, Covid-19 and the united states financial markets' volatility, *Finance Research Letters* p. 101699.

Alchian, A. A.: 1953, The meaning of utility measurement, *The American Economic Review* 43(1), 26–50.

Angrist, J. D. and Pischke, J.-S.: 2008, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press.

Arellano, M.: 2003, *Panel data econometrics*, Oxford university press.

Arellano, M. and Bond, S.: 1991, Some tests of specification for panel data: Monte carlo evidence and an application to employment equations, *The review of economic studies* 58(2), 277–297.

Arellano, M. et al.: 1987, Computing robust standard errors for within-groups estimators, *Oxford bulletin of Economics and Statistics* 49(4), 431–434.

Ashenfelter, O. and Card, D.: 1985, Using the longitudinal structure of earnings to estimate the effect of training programs, *review of economics and statistics*; 67 (4).

Baker, S. R., Bloom, N., Davis, S. J., Kost, K., Sammon, M. and Viratyosin, T.: 2020, The unprecedented stock market reaction to covid-19, *The Review of Asset Pricing Studies* .

- Baumol, W. J.: 1951, The neumann-morgenstern utility index—an ordinalist view, *Journal of Political economy* 59(1), 61–66.
- Baumol, W. J.: 1958, The cardinal utility which is ordinal, *The Economic Journal* 68(272), 665–672.
- Bertrand, M., Duflo, E. and Mullainathan, S.: 2004, How much should we trust differences-in-differences estimates?, *The Quarterly journal of economics* 119(1), 249–275.
- Canova, F.: 1998, Detrending and business cycle facts, *Journal of monetary economics* 41(3), 475–512.
- Card, D. and Krueger, A. B.: 2000, Minimum wages and employment: a case study of the fast-food industry in new jersey and pennsylvania: reply, *American Economic Review* 90(5), 1397–1420.
- Carvalho, C., Masini, R. and Medeiros, M. C.: 2018, Arco: an artificial counterfactual approach for high-dimensional panel time-series data, *Journal of econometrics* 207(2), 352–380.
- Casado, M. G., Glennon, B., Lane, J., McQuown, D., Rich, D. and Weinberg, B. A.: 2020, The effect of fiscal stimulus: Evidence from covid-19. Working Paper.
- Choi, I.: 2001, Unit root tests for panel data, *Journal of international money and Finance* 20(2), 249–272.
- Clemens, J. and Miran, S.: 2012, Fiscal policy multipliers on subnational government spending, *American Economic Journal: Economic Policy* 4(2), 46–68.
- Clemens, J. and Veuger, S.: 2020, Implications of the covid-19 pandemic for state government tax revenues. Working Paper.
- Coates, B., Cowgill, M., Chen, T. and Mackey, W.: 2020, Shutdown: estimating the covid-19 employment shock, Grattan Institute, Melbourne, Victoria .
- Corbet, S., Hou, G., Hu, Y., Oxley, L. and Xu, D.: 2020, Pandemic-related financial market volatility spillovers: Evidence from the chinese covid-19 epicentre, Available at SSRN 3618736 .
- Cribari-Neto, F.: 2004, Asymptotic inference under heteroskedasticity of unknown form, *Computational Statistics & Data Analysis* 45(2), 215–233.
- De Vito, A. and Gomez, J.-P.: 2020, Estimating the covid-19 cash crunch: Global evidence and policy, *Journal of Accounting and Public Policy* p. 106741.

Debreu, G.: 1954, Representation of a preference ordering by a numerical function, *Decision processes* 3, 159–165.

Demsetz, H.: 1968, Why regulate utilities?, *The Journal of Law and Economics* 11(1), 55–65.

Doudchenko, N. and Imbens, G. W.: 2016, Balancing, regression, difference-in-differences and synthetic control methods: A synthesis, Technical report, National Bureau of Economic Research.

Fairlie, R. W., Couch, K. and Xu, H.: 2020, The impacts of covid-19 on minority unemployment: First evidence from april 2020 cps microdata.

Ferman, B. and Pinto, C.: 2016, Revisiting the synthetic control estimator, Working Paper FGV .

Ferman, B. and Pinto, C.: 2019, Inference in differences-in-differences with few treated groups and heteroskedasticity, *Review of Economics and Statistics* 101(3), 452–467.

Ferman, B., Pinto, C. and Possebom, V.: 2020, Cherry picking with synthetic controls, *Journal of Policy Analysis and Management* 39(2), 510–532.

Fernandes, N.: 2020, Economic effects of coronavirus outbreak (covid-19) on the world economy, Available at SSRN 3557504 .

Fezzi, C. and Fanghella, V.: 2020, Real-time estimation of the short-run impact of covid-19 on economic activity using electricity market data, *Environmental and Resource Economics* 76(4), 885–900.

Friedman, J., Hastie, T. and Tibshirani, R.: 2010, Regularization paths for generalized linear models via coordinate descent, *Journal of statistical software* 33(1), 1.

Garcin, M., Klein, J. and Laaribi, S.: 2020, Estimation of time-varying kernel densities and chronology of the impact of covid-19 on financial markets, arXiv preprint arXiv:2007.09043 .

Gobillon, L. and Magnac, T.: 2016, Regional policy evaluation: Interactive fixed effects and synthetic controls, *Review of Economics and Statistics* 98(3), 535–551.

Gómez, V. and Maravall, A.: 1996, Programs tramo and seats, instruction for user (beta version: September 1996), Technical report, Banco de España.

Hsiao, C., Steve Ching, H. and Ki Wan, S.: 2012, A panel data approach for program evaluation: measuring the benefits of political and economic integration of hong kong with mainland china, *Journal of Applied Econometrics* 27(5), 705–740.

Ibikunle, G. and Rzayev, K.: 2020, Volatility, dark trading and market quality: evidence from the 2020 covid-19 pandemic-driven market volatility, Available at SSRN 3586410 .

Knight, F. H.: 1921, Risk, uncertainty and profit, Vol. 31, Houghton Mifflin.

Kurmann, A., Lale, E. and Ta, L.: 2020, The impact of covid-19 on us employment and hours: Real-time estimates with homebase data, May). http://www.andrekurmann.com/hb_covid .

Laffont, J.-J. and Martimort, D.: 2009, The theory of incentives: the principal-agent model, Princeton university press.

Le Quéré, C., Jackson, R. B., Jones, M. W., Smith, A. J., Abernethy, S., Andrew, R. M., De-Gol, A. J., Willis, D. R., Shan, Y., Canadell, J. G. et al.: 2020, Temporary reduction in daily global co 2 emissions during the covid-19 forced confinement, Nature Climate Change pp. 1–7.

LeRoy, S. F. and Singell Jr, L. D.: 1987, Knight on risk and uncertainty, Journal of political economy 95(2), 394–406.

Lutz, B., Molloy, R. and Shan, H.: 2011, The housing crisis and state and local government tax revenue: Five channels, Regional Science and Urban Economics 41(4), 306–319.

Mandal, I. and Pal, S.: 2020, Covid-19 pandemic persuaded lockdown effects on environment over stone quarrying and crushing areas, Science of The Total Environment 732, 139281.

Milani, B., Bender Filho, R., Ceretta, P. S., Vieira, K. M. and Coronel, D. A.: 2014, Spillover de liquidez no mercado brasileiro, Revista Pensamento Contemporâneo em Administração 8(2), 55–74.

Morgenstern, O. and Von Neumann, J.: 1953, Theory of games and economic behavior, Princeton university press.

Norouzi, N., de Rubens, G. Z., Choubanpishhezafar, S. and Enevoldsen, P.: 2020, When pandemics impact economies and climate change: Exploring the impacts of covid-19 on oil and electricity demand in china, Energy Research & Social Science 68, 101654.

Ozili, P. K. and Arun, T.: 2020, Spillover of covid-19: impact on the global economy, Available at SSRN 3562570 .

Poterba, J. M.: 1994, State responses to fiscal crises: The effects of budgetary institutions and politics, Journal of political Economy 102(4), 799–821.

Roberton, T., Carter, E. D., Chou, V. B., Stegmuller, A. R., Jackson, B. D., Tam, Y., Sawadogo-Lewis, T. and Walker, N.: 2020, Early estimates of the indirect effects of the covid-19 pandemic on maternal and child mortality in low-income and middle-income countries: a modelling study, *The Lancet Global Health* .

Romano, F.: 2020, An estimate of the economic impact of covid-19 on australia, Available at SSRN 3581382 .

Snow, J.: 1855, On the mode of communication of cholera, John Churchill.

Sumner, A., Hoy, C., Ortiz-Juarez, E. et al.: 2020, Estimates of the impact of covid-19 on global poverty, UNU-WIDER, April pp. 800–9.

Topcu, M. and Gulal, O. S.: 2020, The impact of covid-19 on emerging stock markets, *Finance Research Letters* p. 101691.

Tristão, J. A. M.: 2003, A administração tributária dos municípios brasileiros: uma avaliação do desempenho da arrecadação, PhD thesis.

Wang, Q. and Su, M.: 2020, A preliminary assessment of the impact of covid-19 on environment—a case study of china, *Science of the Total Environment* p. 138915.

Zhang, D., Hu, M. and Ji, Q.: 2020, Financial markets under the global pandemic of covid-19, *Finance Research Letters* p. 101528.

6. Apêndice

Tabela 1 - Variações Percentuais de Arrecadação e Coeficientes de Variação na Comparação Pandêmica

UF	Varição Percentual entre a média Pré-Pandêmica e Pandêmica	Coef. Variação Pré-Pandêmico	Coef. Variação Pandêmico
AC	-7,66	0,04	0,32
AL	-9,55	0,04	0,33
AM	1,57	0,03	0,28
AP	-16,24	0,09	0,37
BA	-9,79	0,04	0,33
CE	-20,12	0,15	0,33
DF	-4,95	0,04	0,29
ES	-7,36	0,03	0,27
GO	-5,64	0,02	0,34
MA	-6,85	0,01	0,32
MG	-8,43	0,04	0,30
MS	8,64	0,03	0,26
MT	12,13	0,04	0,27
PA	0,96	0,02	0,30
PB	-7,99	0,03	0,32
PE	-9,22	0,01	0,32
PI	-7,50	0,14	0,34
PR	-6,36	0,04	0,32
RJ	-9,80	0,05	0,31
RN	-9,31	0,01	0,32
RO	-2,84	0,04	0,29
RR	1,68	0,10	0,27
RS	-7,50	0,01	0,31
SC	-7,87	0,01	0,29
SE	-9,79	0,04	0,33
SP	-8,69	0,01	0,30
TO	-1,36	0,01	0,30
Média	-5,92	0,04	0,31
Mediana	-7,50	0,04	0,31

Fonte: Elaboração própria.

Tabela 2 - Construção de Variáveis de Isolamento Social por UF

UF	IIS_3004	IIS_3005	IIS_3006	IIS_3007	IIS_1508	IIS Percentil 90	I_Iso
RO	36,27	51,34	42,11	38,79	41,12	47,65	1
AC	39,78	54,50	42,67	40,57	43,28	50,01	1
AM	44,19	50,74	40,56	39,07	40,67	48,12	1
RR	37,19	49,84	38,72	37,04	38,63	45,39	0
PA	45,26	49,96	38,26	36,87	40,38	48,08	1
AP	42,92	56,05	42,17	38,42	40,69	50,80	1
TO	34,04	46,03	33,88	34,50	38,07	42,85	0
MA	44,20	48,75	37,32	37,12	38,96	46,93	1
PI	42,48	51,81	40,32	39,41	45,60	49,33	1
CE	44,68	52,94	41,06	39,39	40,85	49,64	1
RN	41,00	49,01	39,16	36,88	38,40	45,81	0
PB	41,86	50,27	39,24	37,38	38,56	46,91	0
PE	46,10	53,55	39,66	37,20	39,87	50,57	1
AL	42,35	53,52	39,40	37,34	39,22	49,05	1
SE	40,08	48,99	38,21	39,98	38,74	45,43	0
BA	38,66	50,08	39,92	38,57	40,31	46,17	0
MG	37,40	47,37	38,22	35,79	39,11	44,07	0
ES	38,14	50,81	38,82	39,69	38,28	46,36	0
RJ	44,38	51,00	40,91	39,05	39,25	48,35	1
SP	40,14	49,78	38,82	36,61	39,44	45,92	0
PR	35,58	47,74	39,11	36,02	43,92	46,21	0
SC	36,42	48,55	43,20	37,54	41,41	46,41	0
RS	38,01	52,17	45,87	37,37	42,06	49,65	1
MS	31,93	46,20	37,24	36,72	39,27	43,43	0
MT	33,81	46,76	40,78	37,09	38,96	44,37	0
GO	43,23	44,54	37,10	35,90	37,30	44,02	0
DF	49,77	49,49	41,04	38,55	39,49	49,66	1
Média	40,37	50,07	39,77	37,74	40,07	47,08	
Mediana	40,14	49,96	39,40	37,37	39,44	46,91	
Desvio Padrão	4,28	2,71	2,33	1,46	1,90	2,28	
CV	0,11	0,05	0,06	0,04	0,05	0,05	

Fonte: Elaboração própria.

Estudos de Casos - Modelos de Painéis Contrafactuais Artificiais

