

Resumo

Este trabalho tem o objetivo de verificar os impactos da construção de usinas eólicas nas economias municipais. Mais especificamente, pretende-se aqui analisar a evolução do produto interno bruto (PIB) *per capita* dos municípios que receberam investimentos na construção de usinas eólicas no período de 2007 a 2014, em comparação a um grupo de outros municípios com características semelhantes, mas que não receberam essa intervenção. O estudo utilizou uma base de dados municipais construída por uma consolidação de 247 variáveis provenientes de 12 fontes diferentes. Foi utilizada uma metodologia para a redução da dimensionalidade do espaço paramétrico, chamada adaLASSO, com o objetivo de selecionar as covariadas mais importantes. O método do controle sintético foi aplicado para cada caso, somando 37 avaliações individuais, que foram então compiladas e analisadas de maneira agregada. Os resultados obtidos se mostraram heterogêneos, mas com efeitos medianos positivos, concordando com a bibliografia levantada. Os efeitos medianos compilados nos municípios estão entre +3,3% e +10,1% em comparação com o grupo de controle. Os efeitos são maiores a partir de dois anos após as construções e em parques eólicos com investimentos mais elevados.

Palavras-chave: Usinas eólicas. Avaliação de impacto. Controle sintético.

Abstract

This study examines the impacts of wind farm plant construction on municipal economies. More specifically, it analyzes the evolution of the GDP per capita of municipalities that received investments in wind farm plant construction from 2007 to 2014, compared to a group of municipalities with similar characteristics which underwent no such intervention. Data were obtained from a municipal database constructed by consolidating 247 variables from 12 different sources. The dimensionality of the parametric space was reduced using adaLASSO to select the most important covariates. Synthetic control was applied for each case, resulting in 37 individual assessments which were then compiled and analyzed in an aggregated manner. The heterogeneous results showed positive median effects in line with the reviewed literature. The median effects compiled in municipalities ranged from +3.3 to +10.1% when compared with the control group. Effects were more pronounced starting from two years after construction and in wind parks with higher investments.

Keywords: Wind power plants. Impact evaluation. Synthetic control methods.

JEL Classification: C23, Q42, R58

Introdução

A energia eólica, em relação a outras fontes de energia, apresenta uma série de vantagens. Em primeiro lugar, trata-se de uma fonte de energia renovável e limpa. Além disso, não produz resíduos ao gerar eletricidade. Também se deve destacar que a fonte é considerada inesgotável e não há custos associados à obtenção de uma matéria-prima, assim como há baixos riscos ambientais envolvidos na construção de suas usinas. Há estudos que procuram analisar a importância da implementação de energias renováveis em nível local. Pelo fato de poder coexistir com outras atividades ligadas ao uso da terra, os parques eólicos podem ajudar no desenvolvimento econômico de regiões agrícolas, aumentando a qualidade de vida e reduzindo desigualdades sociais. Ou seja, devido aos arrendamentos de terras para a instalação de torres eólicas, o investimento em energia eólica pode estar associado ao aumento de renda dos pequenos proprietários de terras no interior do Brasil. O Banco Nacional de Desenvolvimento Econômico e Social (BNDES) contribui para o setor de geração de energia eólica tanto pelo apoio a projetos de construção de novas plantas como pelo financiamento à aquisição de máquinas e equipamentos.

Nesse sentido, este trabalho tem o objetivo de verificar possíveis impactos da política de construção de usinas eólicas nas economias municipais. Mais especificamente, pretende-se aqui analisar a evolução do produto interno bruto (PIB) *per capita* dos municípios que receberam investimentos na construção de usinas eólicas em comparação a outro grupo de municípios com características semelhantes, mas que não receberam essa intervenção. O período contemplado pela avaliação vai de 2007 a 2014, devido à disponibilidade de dados do acompanhamento das centrais geradoras

eólicas, da Agência Nacional de Energia Elétrica (Aneel), sobre as características das obras em andamento, assim como da capacidade já instalada por município.

Para que os resultados estimados em um exercício empírico possam ser interpretados para uma análise da intervenção como um todo, é necessário que as estimativas obtidas por essa técnica permitam inferências estatísticas confiáveis. Para isso, é fundamental que a amostra de unidades tratadas tenha um volume apropriado. Esse ponto é sensível para muitas intervenções públicas, como para o caso do impacto local, em nível geográfico, das grandes obras de infraestrutura. Nesse sentido, este texto apresenta uma metodologia adequada para os casos em que as intervenções atingem uma pequena quantidade de unidades. Essa metodologia é a estimação por controle sintético, a qual procura comparar cada unidade tratada com uma unidade artificial de controle, que consiste em uma média ponderada de todas as unidades elegíveis para comparação disponíveis na base de dados.¹

O método do controle sintético foi aplicado para cada município beneficiado pela construção de parques eólicos, somando 37 avaliações individuais, que foram então compiladas e analisadas de maneira agregada. Os resultados obtidos se mostraram heterogêneos, mas com efeitos medianos positivos, concordando com a bibliografia levantada. Os efeitos medianos compilados sobre o PIB *per capita* dos municípios estão entre +3,3% e +10,1% em comparação com o grupo de controle. Os efeitos são maiores a partir de dois anos após as construções e em parques eólicos com investimentos mais elevados.

¹ Ver, por exemplo, Abadie e Gardeazabal (2003) e Abadie, Diamond e Hainmueller (2010, 2011, 2015).

O estudo aqui realizado utilizou uma base de dados municipais que consolida 247 variáveis quantitativas provenientes de 12 fontes diferentes. Essas fontes são o Instituto Brasileiro de Geografia e Estatística (IBGE), o Sistema de Informações sobre Mortalidade (SIM) e o Cadastro Nacional de Estabelecimentos de Saúde (CNES), ambos do Departamento de Informática do Sistema Único de Saúde (Datasus), Informações da Estatística Bancária por Município (Estban), dados da Secretaria de Comércio Exterior (Secex), órgão do Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC), a Relação Anual de Informações Sociais (Rais), do Ministério do Trabalho e Emprego, informações financeiras dos municípios (Finbra), banco de dados criado pela Secretaria do Tesouro Nacional (STN), em convênio com a Caixa Econômica Federal (CEF), o Censo Escolar e o Censo do Ensino Superior, ambos do Instituto Nacional de Estudos e Pesquisas (Inep), o Sistema Nacional de Informações sobre Saneamento (SNIS) e uma série de variáveis derivadas construídas com a análise dessas variáveis brutas. Por fim, o banco de dados da Aneel forneceu informações sobre os valores de potência outorgada para todas as usinas eólicas avaliadas, assim como o ano de construção e o município em que está localizado o parque eólico. Para reduzir a dimensionalidade do modelo utilizado neste exercício empírico, de modo a selecionar as variáveis explicativas mais relevantes em relação à variável de interesse, foi utilizada uma metodologia de seleção de variáveis denominada Adaptive LASSO, ou adaLASSO. Em resumo, essa metodologia introduz uma penalização sobre a soma dos coeficientes estimados, de modo a reduzir a complexidade do modelo e seu custo em termos computacionais.

O trabalho está organizado da seguinte maneira: esta introdução é seguida por uma descrição sobre o panorama da energia eólica

no Brasil, com base em dados da Aneel, que inclui uma revisão da bibliografia sobre seus impactos em nível local. Após isso, é descrita a base de dados consolidada, assim como as principais informações de suas fontes. A metodologia é apresentada a seguir, na qual são expostos os modelos teóricos da estimação por controle sintético e do método adaLASSO de seleção de variáveis de controle. Após isso, são apresentados os resultados das estimações individuais compiladas. Por último, as considerações finais e os possíveis desdobramentos futuros dessa linha de pesquisa encerram o trabalho.

Panorama do setor de energia eólica no Brasil

Conceitualmente, pode-se definir um parque eólico ou uma usina eólica como um conjunto de moinhos de vento ou turbinas usados para gerar energia elétrica por meio de seus aerogeradores, que são empurrados pelo vento. As turbinas são destinadas a transformar energia eólica em elétrica. A principal vantagem da energia eólica em relação a outras fontes é que se trata de uma fonte de energia renovável e limpa, pois não emite os gases de efeito estufa (GEE) que contribuem para o aquecimento global. Além disso, não produz resíduos ao gerar eletricidade. É importante destacar, ainda, que a fonte é considerada inesgotável e não há custos associados à obtenção de uma matéria-prima, diferentemente do que ocorre com combustíveis fósseis, assim como há baixos riscos ambientais em comparação com os que podem existir em usinas hidrelétrica e nucleares, por exemplo (Costa, R. A.; Casotti; Azevedo, 2009).

Atualmente, a maior fonte energética do Brasil advém de usinas hidrelétricas. Apesar de sua produção não ser poluente e ser barata,

essa dependência pode acarretar danos. Isso porque, em períodos de seca, os reservatórios de água podem esvaziar, o que leva à necessidade de colocar em funcionamento usinas termelétricas, caras e poluentes. Para que o gargalo de energia seja solucionado, investimentos em energia alternativa tornam-se de suma importância. O Brasil tem um dos maiores volumes de ventos do mundo, assim como baixa probabilidade de ocorrência de fenômenos climáticos extremos, o que permite ampliar seu uso de energia eólica. Segundo o Atlas do Potencial Eólico (Amarante; Zack; Sá, 2001), o território nacional apresenta ventos com potencial que proporcionariam o equivalente a 272 terawatt-hora por ano (TWh/ano), o que representa aproximadamente 64% do consumo nacional de energia elétrica, que gira em torno de 424 TW/ano.

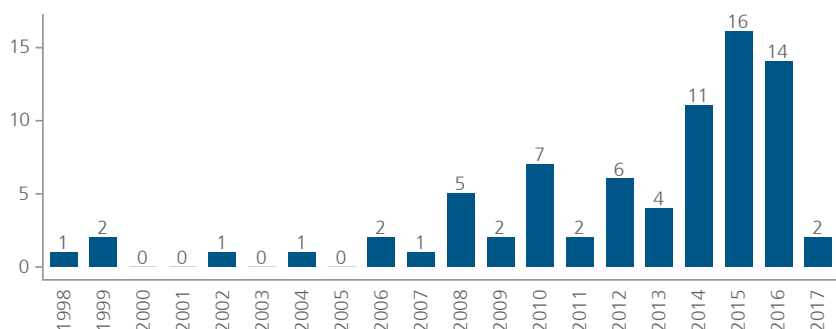
O Brasil, nos últimos anos, vem acelerando a implantação de parques eólicos. A primeira turbina eólica foi instalada em Fernando de Noronha, em 1992 (Gouvêa; Silva, 2018). Atualmente, 14 estados produzem energia elétrica por meio de usinas eólicas. Até 2014, a quantidade de energia eólica produzida era de 4 terawatt-hora por ano. Ao fim de 2022, o Brasil contava com capacidade instalada de 23,76 GW, 11,5% na matriz elétrica total do país (EPE, 2023).

Os incentivos no setor de energia eólica no Brasil tiveram início em 2002, por meio do Programa de Incentivo às Fontes Alternativas de Energia Elétrica (Proinfa). O Proinfa ganhou força a partir de 2004, sob o Novo Modelo do Setor Elétrico (Gouvêa; Silva, 2018). A partir de 2009, houve uma aceleração dos investimentos no setor. Os investimentos ocorrem por meio de leilões especificamente para a geração de eólicas e outras fontes de energias alternativas, por serem pouco competitivas em relação às fontes

tradicionais (hidrelétrica, gás natural, óleo combustível etc.) (Lage; Processi, 2013).

O Gráfico 1 apresenta o número de municípios brasileiros que receberam suas primeiras usinas eólicas por ano. Vale ressaltar, de acordo com o gráfico, que o período de maior inauguração de eólicas por município foi entre 2014 e 2017. Dessa maneira, a maior parte dos parques eólicos não pôde ser avaliada neste trabalho por falta de dados nas variáveis do banco da Aneel.

Gráfico 1 • Número de municípios com parques eólicos pelo ano de sua primeira usina construída

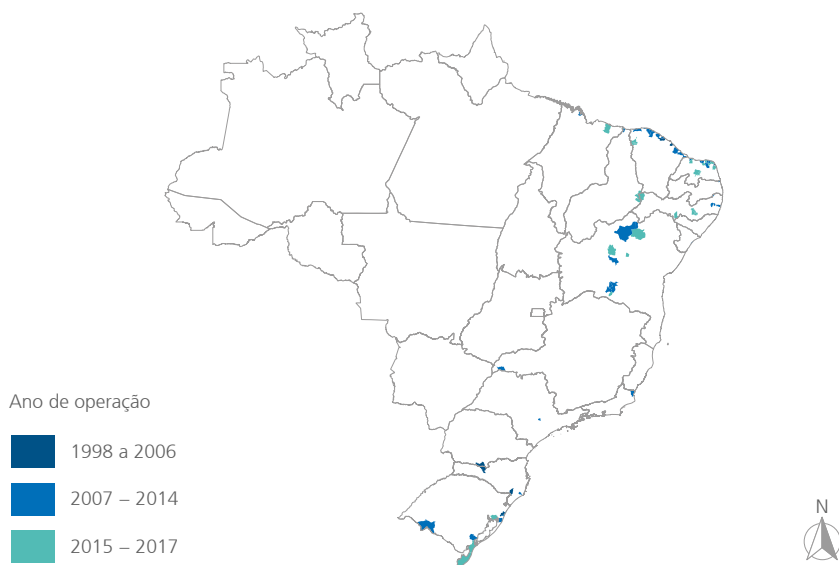


Fonte: Elaboração própria com base em dados da Aneel.

O mapa apresentado na Figura 1 indica como as usinas eólicas estão distribuídas no Brasil, agregando os parques eólicos por período de construção. Os períodos foram divididos em três grupos, isto é, antes (1998-2007), durante (2008-2014) e depois (2015-2017) do período de análise deste trabalho, definidos pela disponibilidade de dados. A região Nordeste apresenta a maioria das usinas eólicas no país, com 78% do total. O estado com a maior quantidade de usinas é o Rio Grande do Norte, concentrando 28% do total entre os 14 estados. A região Sul vem atrás, com 21%, sendo que só o Rio

Grande do Sul representa 17% de eólicas no Brasil. A região Sudeste conta com menos de 1% da matriz energética eólica. O mapa também evidencia uma concentração de usinas no período mais recente.

Figura 1 • Mapa dos municípios com parques eólicos pelo ano da primeira usina construída



Fonte: Elaboração própria com base em dados da Aneel.

Há estudos que procuram analisar a importância da introdução de energias renováveis em nível local. Pelo fato de poder coexistir com outras atividades ligadas ao uso da terra, os parques eólicos podem ajudar no desenvolvimento econômico de regiões agrícolas, com capacidade de melhorar, assim, esse setor, aumentando a qualidade de vida e reduzindo desigualdades sociais (AIE, 2002; Costa, R. C.; Prates, 2005). Mais especificamente, o investimento em energia eólica está associado ao aumento de renda dos pequenos proprietários de

terras em regiões ermas do Brasil. Isso se deve aos arrendamentos de suas terras para a instalação de torres eólicas (Aneel, 2014).

Na literatura, há uma crescente série de estudos empíricos que buscam medir o impacto de usinas eólicas em nível local no Brasil. Simas (2012) procurou verificar a contribuição da energia eólica para a geração de empregos a partir de dados primários levantados em entrevistas com representantes de 18 parques eólicos e com gestores de empresas de componentes para o setor e utilizando uma análise baseada em matriz insumo-produto. O trabalho chegou à conclusão de que o setor de energia eólica tem potencial de gerar 330 mil empregos até 2020, principalmente no setor de construção.

Estudando especificamente os efeitos em âmbito municipal, Resende (2015) avaliou o impacto da construção de usinas eólicas sobre os preços dos aluguéis no entorno das regiões. O trabalho utilizou dados do Censo Demográfico do IBGE de 1991 e 2010 para contemplar 18 municípios que receberam alguma usina eólica entre esses anos. Por meio da metodologia do controle sintético, o estudo concluiu que os aluguéis ficaram 8,3% mais caros nos municípios que receberam usinas eólicas do que foi previsto pelo grupo de controle.

Particularmente para os municípios da região Nordeste, Rodrigues, Gonçalves e Chagas (2016) analisaram o impacto dos parques eólicos no nível de emprego, na massa salarial e no número de firmas. A análise partiu da verificação dos dados da Aneel e das informações de mercado de trabalho disponibilizadas pela Rais, do Ministério do Trabalho. Por meio de um pareamento por escore de propensão, o estudo verificou que os municípios com usinas eólicas apresentaram maiores massas salariais e

número de firmas do que os demais, ao passo que o efeito sobre o emprego não foi significativo.

Já para o caso da região Sul, Rintzel, Alves e Massuquetti (2017) analisaram dados de mercado de trabalho da Rais, dados de valor adicionado desagregado por grandes setores do IBGE e informações de receitas de impostos do Sistema de Coleta de Dados Contábeis e Fiscais dos Entes da Federação (SISTN). O estudo procurou verificar efeitos da instalação de parques eólicos nos municípios por meio do método estrutural diferencial, que é baseado na comparação do desempenho dos municípios tratados com a média da região para cada variável. O trabalho identificou resultados positivos para todas as variáveis, especialmente para o nível de emprego.

Ou seja, os estudos empíricos analisados destacaram potenciais efeitos positivos da construção de usinas eólicas nas suas localidades. Todavia, essas pesquisas observaram casos específicos e não controlaram o problema da micronumerosidade de casos tratados, o que compromete inferências estatísticas sobre a avaliação. Dessa maneira, a contribuição deste trabalho é exatamente utilizar uma metodologia apropriada para avaliação com poucos casos tratados, valendo-se de ampla disponibilidade de dados para melhor verificar o impacto da construção de usinas eólicas nos municípios brasileiros.

Nesse levantamento da literatura empírica, é importante destacar o trabalho de Assunção, Costa e Szerman (2016). Esse trabalho procurou avaliar os efeitos da construção de usinas hidrelétricas em 82 municípios brasileiros entre 2002 e 2011 em uma série de indicadores socioeconômicos municipais. Os autores utilizaram a metodologia do controle sintético, baseada na construção de

82 estudos de caso comparáveis. A pesquisa observou que o impacto da construção das usinas hidrelétricas é estimulado apenas no curto prazo, tendendo a zero no quinto ano após o início das obras. O único impacto de médio prazo foi identificado no nível de emprego formal no município. Por fim, os efeitos nos municípios são muito heterogêneos, em todos os indicadores estimados.

Base de dados

Com o objetivo de verificar possíveis efeitos locais da política de construção de usinas eólicas no Brasil, este trabalho utilizou uma extensa base de dados dos municípios brasileiros. Essa base é uma consolidação de informações de diversas fontes, que serão descritas a seguir. A base municipal conta com um total de 247 variáveis numéricas, que podem servir como indicadores de interesse ou covariadas para as estimações de controle sintético. Essas variáveis são provenientes de 12 fontes diferentes de informações municipais, cuja descrição está no Apêndice A deste artigo.

Devido à disponibilidade de dados da Aneel, a avaliação considera como unidades tratadas os 37 municípios beneficiados de 2007 a 2014 por entrada em operação de sua primeira usina eólica, expostos na Tabela 1. Já os potenciais controles (isto é, o *donor pool*) incluem um total de 5.490 municípios que não têm usinas eólicas. Esse conjunto não inclui o município de Gravatá (PE), pois, apesar de sua primeira usina eólica ter entrado em construção em 2010, suas obras iniciaram apenas em 2005, de modo que faltam dados para a análise. Outros 42 municípios que receberam sua primeira usina eólica antes de 2007 ou após 2014 foram considerados contaminados e excluídos da amostra.

O exercício realizado procurou construir um controle sintético para cada unidade tratada, a partir de combinações de municípios não tratados de todo o Brasil. Considerando que os maiores efeitos sobre as economias locais são provenientes das obras de implantação dos parques, para melhor capturar os efeitos da construção das usinas eólicas, o controle sintético foi aplicado para cada caso com base no ano de início das obras civis das estruturas, de acordo com o acompanhamento das centrais geradoras eólicas, da Aneel. Para evitar distorções, alguns ajustes foram realizados. Para as obras que duraram mais de um ano, se seu mês de início fosse entre julho e dezembro, o ano de tratamento seria deslocado para o ano imediatamente posterior, uma vez que a maior parte da construção aconteceria nesse ano. Para as obras que duraram até um ano, considerou-se que o ano de tratamento seria aquele em que ocorreu o maior número de dias de obras.

Tabela 1 • Lista de municípios avaliados, ano e potência de sua primeira usina eólica

Município	Tratamento	Potência outorgada (kW)	Anos de obras
Acaraú – CE	2008	70.800	2
Amontada – CE	2009	54.600	0
Aracati – CE	2007	10.500	1
Areia Branca – RN	2012	20.000	1
Barra dos Coqueiros – SE	2012	34.500	0
Beberibe – CE	2008	25.600	0
Boituva – SP	2012	2,24	0
Brotas de Macaúbas – BA	2011	95.190	1
Cabo de Santo Agostinho – PE	2012	2.000	0
Caetité – BA	2011	296.820	3
Camocim – CE	2008	105.000	1
Cururupu – MA	2008	22,5	0
Galinhos – RN	2012	118.570	2
Guamaré – RN	2010	51.000	0
Guanambi – BA	2011	167.840	3
Igaporã – BA	2011	143.840	3
Itarema – CE	2012	30.000	2

Continua

Continuação

Município	Tratamento	Potência outorgada (kW)	Anos de obras
Iturama – MG	2012	156	0
João Câmara – RN	2011	39.600	1
Macaparana – PE	2010	4.950	0
Mataraca – PB	2007	10.200	0
Palmares do Sul – RS	2010	9.200	0
Paracuru – CE	2007	25.200	1
Parazinho – RN	2011	466.000	3
Parnaíba – PI	2008	18.000	0
Pedra Grande – RN	2012	118.400	2
Pelotas – RS	2014	1,98	0
Pombos – PE	2009	4.950	1
Sant'Ana do Livramento – RS	2011	60.000	0
São Francisco de Itabapoana - RJ	2010	28.050	0
São Miguel do Gostoso – RN	2013	51.200	1
Sento Sé – BA	2012	90.000	1
Sobradinho – BA	2012	48.000	1
Trairi – CE	2012	55.392	1
Tramandaí – RS	2010	70.000	1
Tubarão – SC	2014	2.099,5	0
Xangri-lá – RS	2014	27.675	0

Fonte: Elaboração própria com base em dados da Aneel.

A Tabela 2 apresenta estatísticas descritivas sobre os municípios tratados em relação ao total de municípios elegíveis como unidades de comparação (*donor pool*), considerando o ano imediatamente anterior aos primeiros tratamentos (2006). Apresentam-se a média e o desvio-padrão de 16 indicadores com o objetivo de exibir um perfil dos dois grupos de municípios. Conforme mostra a tabela, os tratados estavam, em 2006, com uma renda anual média por habitante ligeiramente acima da média nacional. Além disso, apresentam maior porte do que os não tratados em termos de

população, número de estabelecimentos, receitas totais, despesas com saúde e educação. Já os potenciais controles tinham vantagem em relação ao PIB e às receitas tributárias. Pelos dados apresentados, não é possível inferir diferenciais de níveis de desenvolvimento entre os dois grupos.

Tabela 2 • Perfil das unidades (2006)

Variável	Donor pool		Tratados	
	Média	Desvio-padrão	Média	Desvio-padrão
PIB (R\$)	432.458.035,18	4.687.964.973,25	402.910.918,92	592.410.986,73
PIB <i>per capita</i> (R\$)	8.052,04	9.605,62	9.167,03	9.715,34
Adm. pública/ PIB (%)	33,37	17,00	30,94	16,97
Agropecuária/ PIB (%)	22,27	14,72	12,07	9,72
Indústria/PIB (%)	13,77	14,92	25,01	24,25
Serviços/PIB (%)	30,59	12,57	31,99	16,00
População	33.017,24	197.831,82	49.427,19	62.365,64
N. estabelecimentos	445,85	3.596,20	520,84	1.042,37
N. vínculos de emprego	6.250,85	66.662,33	6.087,70	10.492,52
Massa salarial (R\$)	7.343.984,40	109.022.685,90	5.094.955,44	10.576.407,88
Receita total (R\$)	33.571.586,64	280.819.705,38	42.600.486,33	49.834.522,83
Receita tributária (R\$)	6.426.819,96	119.969.415,13	4.137.954,52	6.393.477,11
Desp. educação (R\$)	8.170.058,99	58.382.637,37	11.505.349,30	11.511.485,81
Desp. saúde (R\$)	7.280.861,06	51.270.657,61	9.035.692,83	12.589.765,74
Total de municípios	5.490	–	37	–

Fonte: Elaboração própria com base em dados do IBGE, da Rais e da Finbra.

Metodologia

Controle sintético

A análise de impacto de intervenções públicas em nível de localidade, como da construção de usinas eólicas sobre o desempenho econômico dos municípios contemplados, acarreta uma série de desafios de ordem empírica (Assunção; Costa; Szerman, 2016). Em primeiro lugar, é preciso que alguma técnica seja utilizada para construção de um cenário contrafactual. Ou seja, é necessário que se leve em consideração a mensuração da variável de interesse em um cenário hipotético no qual esses municípios não tenham sofrido intervenção. Dessa maneira, a medida do impacto será a diferença entre a variável de interesse – no caso, o PIB *per capita* – observada nos municípios beneficiados (tratados) e o PIB *per capita* do contrafactual após o tratamento.

Em segundo lugar, a escolha de construir obras de infraestrutura em localidades específicas é baseada em fatores políticos, econômicos e ambientais. Isto é, como a seleção dos municípios contemplados por usinas eólicas não é aleatória, espera-se que as localidades tratadas tenham características diferentes das não tratadas. Dessa maneira, métodos empíricos baseados em simples comparações de médias entre o grupo de unidades afetadas pela intervenção e o grupo das não afetadas levariam a estimativas viesadas sobre supostos impactos.

Em terceiro, o número de unidades afetadas por esse tipo de intervenção geralmente é pequeno. Dessa maneira, os resultados estimados por meio de métodos de análise de impacto baseados em pareamento e em diferenças-em-diferenças podem ter problemas de inferência estatística.

Nesse sentido, a metodologia de controle sintético foi originalmente concebida para estudos de caso, sendo apropriada para quando se tem poucos tratados, como é o caso em análise. Parte-se do pressuposto de que uma combinação de unidades não tratadas compõe melhor contrafactual para a tratada do que qualquer outra individualmente. Para definir as ponderações aplicadas a cada combinação, o método se vale de características mensuráveis de todas as unidades disponíveis para a análise. O trabalho da metodologia de controle sintético é procurar simular, a partir de uma combinação de informações observadas de unidades não tratadas, a mesma trajetória prévia ao tratamento que se observou na unidade beneficiária, conforme foi desenvolvido por Abadie, Diamond e Hainmueller (2003, 2010, 2015).²

Supõe-se uma amostra de $J + 1$ unidades (por exemplo, municípios) indexadas por j . A unidade $j = 1$ é a unidade de interesse, isto é, a unidade tratada por uma intervenção. As demais unidades, de $j = 2$ a $j = J + 1$, constituem o chamado *donor pool*, ou seja, o conjunto de unidades não tratadas pela intervenção e que servem como potenciais unidades de comparação com a unidade tratada.

Pressupõe-se que a amostra seja um painel balanceado, isto é, que seja composta por dados longitudinais e que todas as suas unidades sejam observadas no período de $t = 1, \dots, T$. A amostra inclui um número positivo de períodos pré-intervenção T_0 , assim como de períodos pós-intervenção T_1 , de modo que $T = T_0 + T_1$. A intervenção que será avaliada consiste na exposição da unidade $j = 1$ ao tratamento durante os períodos $t = T_0 + 1, \dots, T$, considerando que essa intervenção não tem efeitos durante o período pré-tratamento $t = 1, \dots, T_0$.

² Mais detalhes sobre a implementação da metodologia de controle sintético podem ser consultados em Martini e outros (2018).

Dessa maneira, o objetivo da análise de impacto nessa amostra é medir o efeito da intervenção sobre a unidade tratada em um indicador de interesse para o período pós-tratamento.

Por hipótese, considera-se que as características pré-tratamento da unidade de interesse são mais bem aproximadas por uma combinação das unidades não tratadas do que por qualquer uma dessas unidades não tratadas isoladamente. Dessa maneira, o controle sintético pode ser entendido como uma média ponderada das unidades do *donor pool* que será comparada com a unidade tratada. O controle sintético é representado por um vetor $(J \times 1)$ de pesos $W = (w_2, \dots, w_{J+1})$ tal que $0 \leq w_j \leq 1$ para $j = 2, \dots, J$ e $w_2 + \dots + w_{J+1} = 1$. Dessa maneira, a escolha de qualquer valor particular de W é equivalente à escolha de um controle sintético. O vetor X_i ($K \times 1$) contém as características pré-tratamento da unidade tratada, as quais pretende-se aproximar o máximo possível. X_0 , por sua vez, é uma matriz ($K \times J$) que tem os valores das mesmas variáveis para o *donor pool*. Observa-se que K equivale ao número de variáveis disponíveis para mensurar as características das unidades no período pré-tratamento, sendo preditoras da variável de interesse e não sendo afetadas pela intervenção nesse período. Ou seja, o conjunto do total das variáveis preditoras disponíveis é representado por $m = 1, \dots, K$. Nesse conjunto de variáveis, pode-se incluir os valores da própria variável de interesse antes do tratamento.

A diferença entre as características da unidade tratada e do controle sintético é dada pelo vetor $X_1 - X_0W$, sendo que o objetivo da metodologia aqui aplicada é escolher o vetor de pesos W^* , que minimiza essa distância. Esse valor é obtido da seguinte maneira: seja X_{1m} o valor da variável preditora m para a unidade tratada,

e X_{0m} , um vetor $1 \times J$ que contém os valores da mesma variável para as unidades do *donor pool*, o W^* ótimo é aquele que minimiza:

$$\min_{W \in W} \sum_{m=1}^K v_m (X_{1m} - X_{0m}W)^2$$

Nessa equação, v_m é um peso que reflete a importância relativa atribuída à variável m quando se mede a discrepância entre X_1 e X_0W .

Y_{jt} é a variável de interesse da unidade j no tempo t . Y_1 é um vetor $(T_1 \times 1)$ dos valores pós-intervenção da variável de interesse para a unidade tratada, de modo que $Y_1 = (Y_{1T_0+1}, \dots, Y_{1T})'$. Y_0 é uma matriz $(T_1 \times J)$ em que a coluna j contém os valores pós-intervenção da variável de interesse para a unidade $j + 1$. Dessa maneira, a variável de interesse do controle sintético é $Y_1^* = Y_0W^*$.

O estimador de controle sintético do impacto do tratamento é dado pela comparação entre os valores da variável de interesse para a unidade tratada e para a unidade de controle sintético no período pós-tratamento:

$$\delta = Y_1 - Y_1^* \\ \delta = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$$

A consistência do estimador de controle sintético é maior quanto maior for o número de períodos pré-tratamento que estiver presente na base de dados (Abadie; Diamond; Hainmueller, 2010). Segundo esse estudo, esse fator contribui para reduzir o papel de variáveis não observadas na determinação da trajetória pré-tratamento da variável de interesse.

A maneira mais comum de mensurar a qualidade do ajuste por meio do método do controle sintético é pelo erro quadrático médio pré-tratamento, ou Root Mean Squared Prediction Error (RMSPE).

Esse indicador equivale à razão entre os desvios quadrados da trajetória da variável de interesse entre a unidade de referência (tratada ou placebo) e seu correspondente controle sintético antes do ponto de tratamento. Quanto menor for esse valor, menor é a distância entre as trajetórias e melhor é o ajuste.

$$RMSPE = \left(\frac{1}{T_0} \sum_{t=1}^{T_0} \left(Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt} \right)^2 \right)^{1/2}$$

Em casos de intervenções com mais de uma unidade tratada, o controle sintético deve ser aplicado para cada caso individual. Após isso, para se obter medidas do efeito da intervenção como um todo, as estimativas individuais devem ser compiladas (Assunção; Costa; Szerman, 2016). Ou seja, em vez de se considerar uma única unidade tratada $j = 1$, considera-se um conjunto de G unidades tratadas por uma intervenção, as quais são indexadas por $g = 1, 2, \dots, G$. T_{0g} é o ano em que houve o tratamento em cada unidade tratada. Para melhor comparar o tratamento em diferentes anos, eles são normalizados em $\tau = t - T_{0g}$, em que $\tau = 0$ é o ano do tratamento de cada unidade em G .

Após a normalização temporal, $\delta_{j\tau g} = y_{j\tau g} - \hat{y}_{j\tau g}^*$ representa o efeito estimado da intervenção na unidade $g \in G$, pertencente ao conjunto J de unidades tratadas e não tratadas no período τ . Por exemplo, pode se referir ao impacto da construção da usina eólica no município g , integrante do total J de municípios brasileiros, no ano τ . Dessa maneira, os resultados das unidades tratadas são compilados para cada ponto no tempo de modo a se obter uma distribuição empírica. Portanto, o impacto médio dos G tratamentos em cada unidade g será:

$$\bar{\delta}_t = \frac{\sum_{g=1}^G \delta_{g1\tau}^*}{G} = \frac{\sum_{g=1}^G (y_{g1\tau} - y_{g1\tau}^*)}{G}$$

Em caso de elevada heterogeneidade de efeitos entre as unidades tratadas, pode ser mais vantajoso utilizar a mediana em vez da média para compilar os resultados individuais, assim como os percentis 0,25 e 0,75 para cada caso:

$$Px(\delta_{\tau}) = Px(\delta_{g1\tau}^*)$$

Nessa fórmula, Px equivale ao percentil escolhido do efeito para cada τ de tratamento.

Observa-se que, nos casos em que a intervenção não foi aleatoriamente atribuída entre as unidades, dois problemas podem ameaçar as conclusões em termos de inferência a respeito dos efeitos individuais compilados. Em primeiro lugar, pode haver viés, isto é, a trajetória da variável de interesse de uma unidade tratada pode divergir do seu controle sintético desde antes da intervenção. Em segundo lugar, pode existir heterocedasticidade de choques idiosincráticos. Esse problema acontece quando uma unidade tratada recebe choques aleatórios sobre sua variável de interesse com uma variância diferente do seu controle sintético. Se os choques atingem a unidade tratada com maior variância do que o controle sintético, pode haver superestimação do efeito do tratamento; se atingem o controle sintético com maior variância do que a unidade tratada, pode existir subestimação do tratamento.

Os dois problemas – viés e heterocedasticidade de choques – são detectados no exercício empírico por meio de um mau ajuste de pré-tratamento da variável de interesse, e isso pode ser testado sob a forma de um elevado RMSPE pré-tratamento. Por isso, uma forma simples de controlar a influência desses casos na estimação do efeito compilado das intervenções é eliminar as unidades tratadas com um nível de RMSPE pré-tratamento acima do patamar escolhido.

AdaLASSO – método de seleção de covariadas

Os modelos de alta dimensionalidade estão cada vez mais presentes na literatura, já que a inclusão de muitas variáveis pode contribuir para ganhos de capacidade preditiva (Konzen, 2014). Porém, quando a dimensionalidade do modelo é grande em relação ao tamanho da amostra, os métodos tradicionais de regressão podem apresentar problemas. Primeiro, aumenta a dificuldade de tornar os modelos interpretáveis. Segundo, os modelos perdem robustez. Terceiro, há comprometimento da eficiência computacional. Quarto, há perda de eficiência em termos de inferência estatística. Quinto, há problemas com correlação espúria entre as covariadas do modelo, a qual pode ser elevada mesmo quando elas forem independentes e identicamente distribuídas.

No caso de modelos de controle sintético, esses problemas tendem a ser mais graves, já que o método tem sua eficiência computacional muito sensível à extensão das bases de dados. Não obstante, nesses modelos a estimação do contrafactual é dependente do conjunto de covariadas presente na base de dados de análise. Por isso, há a necessidade de filtrar as variáveis mais importantes para explicar a trajetória da variável de interesse sobre a qual será calculado o efeito da intervenção realizada.

Uma solução para problemas referentes à alta dimensionalidade dos modelos é a suposição de esparsidade do vetor de parâmetros. Isto é, a suposição de que muitos de seus componentes são iguais a zero. Essa hipótese pode produzir estimativas viesadas, mas contribui para a identificação das covariadas mais importantes, assim como para a obtenção de um modelo mais parcimonioso. Da mesma maneira, reduz a complexidade do modelo e seu custo em termos computacionais.

O Least Absolute Shrinkage and Selection Operator (LASSO) é um método de encolhimento do conjunto de coeficientes estimados de um modelo desenvolvido por Tibshirani (1996). Esse método consiste, simplificadamente, na introdução de uma penalização ao conjunto de norma L1 dos coeficientes, isto é, uma penalização na soma dos valores absolutos dos coeficientes. Seu objetivo é permitir a estimação de um modelo que produza previsões com pequena variância e que determine um conjunto de preditores que melhor explicam a variável de interesse. A penalização introduzida tende a zerar alguns dos coeficientes estimados, o que não apenas reduz a dimensionalidade do espaço paramétrico, como também seleciona as covariadas mais relevantes.

O LASSO pode ser entendido como uma técnica de regularização. Isto é, considera-se uma função de erro do tipo $E = medida\ de\ erro + \lambda * complexidade\ do\ modelo$. No caso de uma regressão, a medida de erro equivale à soma dos quadrados dos resíduos estimados. O segundo termo representa a penalização dos modelos com maior complexidade e variância dos estimadores, sendo que λ representa a severidade dessa penalidade. Quanto maior for λ , mais simplificado será o modelo estimado, ainda que isso leve a um maior viés. Em termos formais, as estimativas LASSO são obtidas por meio da minimização dos quadrados dos resíduos sujeita a uma penalização de norma L1 dos coeficientes:

$$\hat{\beta}^{LASSO} = argmin_{\beta_0, \beta_1, \dots, \beta_k} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ji} \right)^2 + \lambda \sum_{j=1}^k |\beta_j| \right\}$$

Aqui, $\lambda \geq 0$ é uma função de um parâmetro de ajuste t , de modo que $\sum_{j=1}^k |\beta_j| \leq t$. Quanto maior o valor de λ , maior será a penalidade

imposta ao somatório dos coeficientes. Por outro lado, se $\lambda = 0$, as estimativas LASSO serão iguais às estimativas de MQO. Esse parâmetro é escolhido por um procedimento de *k-fold cross-validation*, no qual a amostra é particionada aleatoriamente em K subamostras de tamanhos iguais. O modelo é iterativamente estimado eliminando-se uma subamostra. Assim, são feitas previsões com base em cada estimação e comparações com a subamostra removida, calculando-se o erro quadrático médio de previsão (EQMP) para avaliar a qualidade do ajuste naquela subamostra. Portanto, o λ ótimo deve minimizar o EQMP médio nas K subamostras.

O LASSO apresenta menor variabilidade entre outras opções de modelos de redução de dimensionalidade (Konzen, 2014). Além disso, por reduzir alguns coeficientes para zero, destaca as covariadas mais relevantes para explicar uma variável de interesse. Por fim, é capaz de realizar a escolha das variáveis e a estimação dos coeficientes simultaneamente. Contudo, deve-se observar que nem sempre o LASSO é consistente na escolha de variáveis (Zhao; Yu, 2006). Isso significa que a solução esparsa – isto é, de dimensionalidade reduzida, com alguns coeficientes reduzidos a zero – pode não representar o modelo verdadeiro quando o tamanho da amostra tende ao infinito.

Nesse sentido, o LASSO Adaptativo, ou *adaLASSO*, é um método que pretende dar consistência às estimativas LASSO por meio da atribuição de diferentes pesos para diferentes coeficientes (Zou, 2006). Ou seja:

$$\hat{\beta}^{adaLASSO} = \underset{\beta_0, \beta_1, \dots, \beta_k}{argmin} \left\{ \sum_{i=1}^n \left(y_i - \beta_0 - \sum_{j=1}^k \beta_j x_{ji} \right)^2 + \lambda \sum_{j=1}^k w_j |\beta_j| \right\}$$

$$w_j = \left| \hat{\beta}_j^{ridge} \right|^{-\tau}, \tau > 0$$

Nesse caso, W_j é um vetor de pesos individuais para auxiliar na seleção das variáveis relevantes para o modelo. Esse vetor é função de $\hat{\beta}_j^{ridge}$, isto é, dos coeficientes estimados por meio de uma regressão que controla possíveis efeitos de multicolinearidade (regressão *ridge*). Uma variável relevante x_j terá um valor comparativamente $\hat{\beta}_j^{ridge}$ elevado, o que reduz o peso W_j atribuído ao seu coeficiente. Por outro lado, uma variável irrelevante x_j terá um valor $\hat{\beta}_j^{ridge}$ comparativamente baixo, o que aumenta o peso W_j atribuído ao seu coeficiente. Assim, o modelo atribuirá maior penalização dos coeficientes das variáveis que aparentam ser pouco importantes para explicar a variável de interesse.

Sob um conjunto de pesos W_j apropriados, o adaLASSO tem propriedades ORACLE, isto é, é consistente na seleção de variáveis e apresenta normalidade assintótica. Ou seja, o método seleciona corretamente as variáveis relevantes quando o modelo aumenta. Além disso, as estimativas dos coeficientes não nulos seguem assintoticamente a mesma distribuição dos estimadores de MQO quando este for estimado apenas com as variáveis relevantes.

Resultados

O método do controle sintético foi aplicado individualmente para comparar cada município tratado com o seu contrafactual. Foi realizado um total de 37 avaliações individuais. O intuito foi fazer uma análise agregada de todos os municípios brasileiros que tiveram usinas eólicas que entraram em operação de 2008 a 2014 a partir de cada avaliação individual. Para melhor capturar o efeito da construção das usinas eólicas, o controle sintético foi aplicado considerando-se como o momento do tratamento o início das obras civis das

estruturas da primeira usina de cada município, incluindo os ajustes explicados na seção anterior. Para construir a trajetória do PIB *per capita* em cada cenário contrafactual, a estimação por meio do método adaLASSO de seleção de covariadas considerou relevantes seis variáveis em nível do município que a explicam: a receita corrente líquida *per capita* (variável derivada a partir de dados do Finbra e do IBGE), o salário médio (Rais), a proporção da indústria no PIB (IBGE), a despesa *per capita* em cultura (variável derivada a partir de dados do Finbra e do IBGE), o PIB da agropecuária (IBGE) e a proporção da administração pública no PIB (IBGE).

Para fins de compilação dos resultados individuais, observou-se que em quatro municípios foram implantados projetos de microgeração, cujos efeitos são menos representativos e contam com processo diferente no aparato regulatório (registro, em vez de autorização). Esses municípios são Boituva (SP), Cururupu (MA), Iturama (MG) e Pelotas (RS). Como seus resultados individuais poderiam gerar viés para o resultado para o grupo como um todo, esses municípios foram excluídos da análise. Além disso, foram desconsiderados do trabalho outros três municípios: Guamaré (RN), Sobradinho (BA) e Galinhos (RN). Os dois primeiros apresentaram resultados discrepantes dos outros por razões alheias à produção de energia eólica. Guamaré sofreu sua maior crise no ano de 2011 em razão da queda dos preços do petróleo na região, setor do qual sua economia é muito dependente. Sobradinho teve uma de suas maiores secas em 2013, o que esvaziou seu reservatório hídrico e comprometeu o fornecimento de energia advinda de sua usina hidrelétrica. No caso do município de Galinhos, houve problemas com o ajuste no período pré-tratamento no controle

sintético, de modo que não foi encontrado um contrafactual consistente para a verificação de impacto. Seu RMSPE pré-tratamento foi igual a 23,9, muito acima da média da amostra (5,1). Restaram trinta casos para a avaliação compilada. Os resultados das avaliações individuais estão no Apêndice B deste trabalho.

A primeira análise feita com base nos resultados consistiu numa comparação da proporção de casos em que o PIB *per capita* dos tratados ficou acima dos valores observados para seus respectivos controles. Intuitivamente, assumindo que o método não é viesado para a definição dos controles, deveríamos observar que, se a construção das usinas eólicas não tiver nenhum impacto sobre as localidades, essa proporção deveria ficar sempre próxima de 50% ao longo do tempo – tal como ocorre com a proporção de caras após determinada quantidade de lançamentos de uma moeda não viciada.

O Quadro 1 mostra as estatísticas descritivas dos municípios tratados e a proporção de casos positivos por ano de referência, em que t é o ano de tratamento para cada caso.³ Para julgar se esses valores são estatisticamente significantes, um intervalo de confiança foi construído com base em uma distribuição de Bernoulli, assumindo um parâmetro p de 50%. Dessa maneira, a hipótese nula assume que em cada momento do tempo há 50% de chance de um município tratado estar melhor do que seu controle. Caso a proporção observada fique acima do limite superior dos intervalos de confiança, há evidência de que a proporção de comparações positivas é estatisticamente superior à de comparações negativas. De acordo com o quadro, é possível observar o descolamento da proporção de efeitos positivos em relação ao intervalo de confiança a partir do ano do

3 O ano t equivale ao ano de construção do parque eólico, sendo $t+1$ o ano em que ele entrou em operação.

tratamento. Em todos os anos de referência, a proporção de impactos positivos supera os 70%, mantendo-se sempre acima do limite superior do intervalo de confiança (68%).

Quadro 1 • **Proporção de efeitos positivos e intervalo de confiança por ano de referência**

Ano	N. de casos positivos	N. de casos negativos	Proporção de casos positivos	Limite inferior	Limite superior	p
t	22	8	0,73	0,33	0,67	0,5
t+1	21	9	0,70	0,33	0,67	0,5
t+2	21	7	0,75	0,33	0,67	0,5
t+3	19	8	0,70	0,32	0,68	0,5

Fonte: Elaboração própria.

Outra análise levou em consideração a magnitude dos impactos estimados, mensurados como a diferença percentual entre a trajetória de cada tratado e seu respectivo controle sintético. Os resultados individuais foram normalizados, de modo que o ano de tratamento para todos os casos foi alinhado em *t*, e os demais anos de análise seguiram sua referência. Para cada ano de análise, os resultados individuais foram compilados, sendo calculadas algumas estatísticas. Devido à heterogeneidade dos resultados individuais, observados pelos elevados desvios-padrão, optou-se por observar primeiramente as medianas, assim como os percentis 0,25 e 0,75 da distribuição. Esse procedimento já havia sido adotado por Assunção, Costa e Szerman (2016), os quais também se depa-
raram com resultados individuais heterogêneos sobre os efeitos municipais de obras de usinas hidrelétricas. Os resultados encontrados estão representados no Quadro 2.

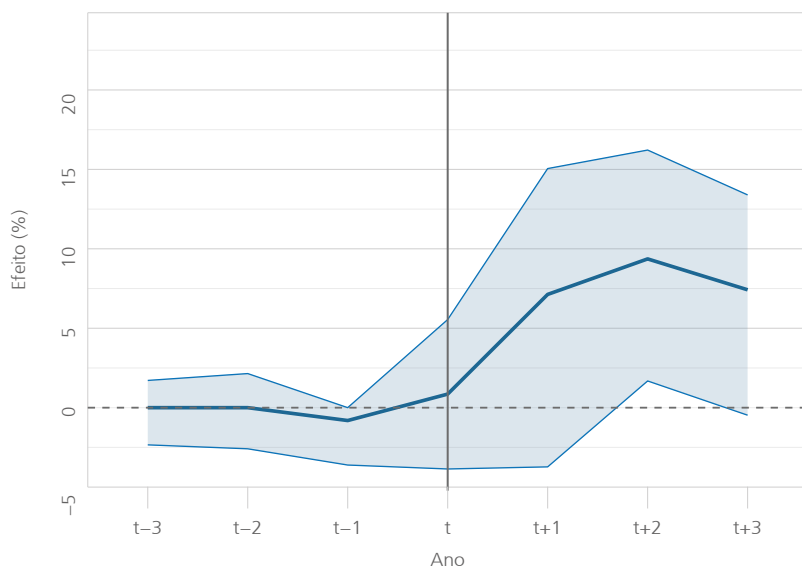
Quadro 2 • Resultados compilados por ano de referência (%)

Ano	Média	DP	P. 0,25	Mediana	P. 0,75	Mínimo	Máximo
t-3	-1,84	4,82	-3,86	0,00	0,76	-18,71	3,89
t-2	-0,66	4,97	-2,33	0,00	1,43	-11,08	12,60
t-1	-1,30	3,90	-2,62	-0,07	1,26	-15,61	4,83
t	11,78	50,54	-0,91	3,28	8,23	-24,47	273,48
t+1	41,08	190,66	-1,69	5,72	17,99	-37,05	1046,45
t+2	29,75	94,41	0,37	9,27	15,03	-48,25	484,52
t+3	45,30	114,15	-2,85	10,08	24,05	-46,65	475,68

Fonte: Elaboração própria.

O Gráfico 2 desenha o comportamento do efeito percentual das diferenças da variável PIB *per capita* entre tratados e controles ao longo do tempo. A linha azul central mostra o efeito mediano entre todos os municípios tratados. As linhas secundárias abaixo e acima da mediana indicam os efeitos no primeiro e no terceiro quartil, respectivamente. Observa-se que, em geral, há efeito mediano positivo no período pós-tratamento dos municípios sob análise, de +3,3% no ano do tratamento, +5,7% no primeiro ano após o tratamento, +9,2% no segundo ano e +10,1% no terceiro ano. Também se observa um aumento da dispersão das linhas no último ano. Novamente, isso pode estar associado à desaceleração dos efeitos das obras mais curtas sobre as economias locais.

Gráfico 2 • Diferenças percentuais de cada unidade tratada em relação ao seu controle sintético



Linhas em azul: efeito do tratamento nos quantis 0,25, 0,5 e 0,75.

Fonte: Elaboração própria.

Outra questão importante a ser estudada se refere aos componentes do impacto econômico das usinas eólicas nos municípios. Conforme levantado na bibliografia, os efeitos podem ocorrer tanto por causa da mobilização de insumos para as obras de construção dos parques eólicos como em razão das receitas oriundas do arrendamento de terras após sua implantação. Esse ponto foi abordado por este trabalho a partir de um exercício baseado em regressões. Nessas regressões, o efeito percentual calculado sobre cada município de t até $t+3$ foi modelado em função de um indicador de potência referente à obra em curso e da potência

já instalada no município. Além disso, a fim de obter a avaliação de acordo com o potencial econômico do município tratado em relação ao investimento recebido, foi criada uma variável de dose do tratamento. Essa variável equivale à razão entre o total de potência outorgada do parque eólico, em megawatts (MW), e o PIB *per capita* do município. A dose foi calculada com o valor de PIB *per capita* fixado no período em que as primeiras usinas estavam em construção. Todos os indicadores utilizados nas regressões foram normalizados pelas suas médias, de modo a permitir uma medida mais intuitiva de comparação.

Para cada par de variáveis independentes foi estimado um modelo baseado em mínimos quadrados ordinários empilhados (MQOE) e outro modelo baseado em efeitos fixos (EF). Intuitivamente, o modelo de mínimos quadrados trata cada par indivíduo-tempo como uma unidade independente na amostra. O modelo de efeitos fixos, por sua vez, controla o efeito de possíveis heterogeneidades individuais ao normalizar o valor de cada variável pela média do indivíduo ao qual está associado. A Tabela 3 mostra os resultados das regressões estimadas. Em nenhuma delas as medidas referentes à estrutura instalada no município foram estatisticamente significantes. A potência da obra foi significativa nos dois modelos, sendo que o coeficiente estimado pelo MQOE teve magnitude superior ao estimado por EF. O indicador de dose da obra só foi estatisticamente significativo no modelo de MQOE, o que provavelmente decorre do fato de que o modelo de EF controla o efeito do PIB *per capita* fixado no primeiro ano da obra.

Tabela 3 • Resultados das estimações dos determinantes dos efeitos das usinas eólicas sobre as economias municipais

	Variável dependente: impacto no PIB <i>per capita</i> (%)			
	EF	MQOE	EF	MQOE
Potência obra	20,04** (9,83)	53,08*** (8,53)		
Potência instalada	-6,67 (9,03)	-18,19 (11,99)		
Dose obra			10,30 (12,13)	24,66* (12,78)
Dose instalada			4,95 (8,80)	-7,27 (12,78)
Intercepto		9,26 (10,46)		31,67*** (11,39)
Observações	115	115	115	115
R2	0,05	0,27	0,02	0,03
R2 ajustado	-0,31	0,26	-0,35	0,02
Estatística F	2,10 (df = 2; 83)	21,06*** (df = 2; 112)	0,77 (df = 2; 83)	1,92 (df = 2; 126)

Fonte: Elaboração própria.

Nota: *p<0,1; **p<0,05; ***p<0,01.

Conclusão

Este artigo procurou fazer uma análise de impacto da política de construção de usinas eólicas sobre o PIB *per capita* dos municípios beneficiados. Para isso, foi realizada uma avaliação individual baseada em controle sintético para os municípios que receberam sua primeira usina eólica no período de 2008 a 2014. Depois disso, os resultados individuais foram compilados. Com base nos trinta casos

avaliados, foi possível observar o impacto positivo dos parques eólicos sobre o PIB *per capita* dos municípios beneficiados, ainda que haja expressiva variabilidade.⁴ Os exercícios de avaliação concordaram em relação ao crescimento do efeito e aos anos em que se apresentou maior impacto positivo. Os efeitos, com mediana estimada entre +3,3% e +10,1%, foram mais claros entre dois e três anos após o início da construção, que estão relacionados aos maiores parques instalados. Também se verificou maior dispersão de impactos no terceiro ano, sinalizando possível desaceleração dos impactos das obras de menor duração. A análise por regressões baseadas em modelos de efeitos fixos e de mínimos quadrados empilhados destacou que o principal determinante dos efeitos é a magnitude da obra realizada, e não a estrutura instalada no município. Portanto, as evidências empíricas apontam para a hipótese de que o impacto das usinas sobre as economias locais vem da mobilização de recursos para as obras de instalação, e não do arrendamento das terras.

A metodologia adotada nesta pesquisa é inovadora em relação às avaliações dos impactos locais das obras de usinas eólicas no Brasil. A estimação dos efeitos locais foi realizada por meio da metodologia de controle sintético. Esse método, além de ser mais consistente para avaliações com micronumerosidade de unidades tratadas, permitiu observar que os efeitos são diferentes ao longo do tempo após a construção e em função da dose de tratamento. Os resultados verificados, tanto em termos de efeitos como em relação à heterogeneidade entre as unidades avaliadas, são semelhantes aos observados por Assunção, Costa e Szerman (2016), em um estudo sobre efeitos

4 Uma possível explicação para essa variabilidade pode ser algum efeito de transbordamento dos impactos para municípios limítrofes a aqueles que receberam investimentos em parques eólicos. Esse é um fator que não foi identificado pela literatura correlata no Brasil, e é um ponto importante para futuras investigações.

locais da construção de usinas hidrelétricas. Esse trabalho também observou que os impactos locais das obras, mensurados em indicadores de produção econômica e de mercado de trabalho, tendem a se dissipar no quinto ano após a construção. Essa constatação não pôde ser observada neste trabalho em razão da indisponibilidade de dados, principalmente devido ao fato de que as usinas eólicas foram construídas, de certa forma, mais recentemente, a partir de 2014. Todavia, as evidências empíricas aqui obtidas indicam que as obras têm efeitos econômicos superiores aos gerados pelo arrendamento de terras proveniente da potência instalada no município.

É importante destacar que o período de maior crescimento de usinas eólicas entrando em operação no Brasil foi a partir de 2014. Este estudo investigou parques eólicos instalados até 2014, devido à disponibilidade de dados do acompanhamento das centrais geradoras eólicas da Aneel. Não obstante, há municípios que só foram analisados um ano após o tratamento por falta de informações na base municipal. Alterações nas características dos empreendimentos realizados após 2014, seja em termos de porte de equipamentos ou de políticas e práticas adotadas pelos empreendedores, podem ter acarretado mudanças nos impactos dos parques eólicos sobre as economias municipais. Portanto, qualquer extrapolação dos resultados encontrados para o período mais recente deve ser feita com a devida cautela.

Conclui-se que há potencial agenda de estudos futuros acerca dos efeitos das eólicas sobre os municípios beneficiados. Espera-se que essa agenda seja cada vez mais consistente e mostre mais evidências de impacto, por haver mais investimentos nos municípios e bancos de dados com maior quantidade de tratados.

Referências

- ABADIE, A.; DIAMOND, A.; HAINMUELLER, J. Comparative politics and the synthetic control method. *American Journal of Political Science*, [s. l.], v. 59, n. 2, p. 495-510, 2015. Disponível em: <https://economics.mit.edu/sites/default/files/publications/Comparative%20Politics%20and%20the%20Synthetic%20Control.pdf>. Acesso em: 8 mar. 2024.
- ABADIE, A.; DIAMOND, A.; HAINMUELLER, J. Synth: an r package for synthetic control methods in comparative case studies. *Journal of Statistical Software*, [s. l.], v. 42, n. 13, 2011. Disponível em: <https://www.jstatsoft.org/article/view/v042i13#:~:text=Abstract,Diamond%2C%20and%20Hainmueller%202010>). Acesso em: 8 mar. 2024.
- ABADIE, A.; DIAMOND, A.; HAINMUELLER, J. Synthetic control methods for comparative case studies: estimating the effect of California's tobacco control program. *Journal of the American Statistical Association*, London, v. 105, n. 490, p. 493-505, 2010. Disponível em: <https://www.tandfonline.com/doi/abs/10.1198/jasa.2009.ap08746>. Acesso em: 8 mar. 2024.
- ABADIE, A.; GARDEAZABAL, J. The economic costs of conflict: a case study of the Basque Country. *American Economic Review*, [s. l.], v. 93, n. 1, p. 113-132, 2003. Disponível em: <https://www.aeaweb.org/articles?id=10.1257/000282803321455188>. Acesso em: 8 mar. 2024.
- AIE – Agência Internacional de Energia. *Renewable energy working party: renewable energy into the mainstream*. Sittard, Oct., 2002.
- AMARANTE, O. A. C.; ZACK, M. B. J.; SÁ, A. L. *Atlas do potencial eólico brasileiro*. Brasília, DF: Ministério de Minas e Energia, 2001. Disponível em: http://www.cresesb.cepel.br/publicacoes/download/atlas_eolico/Atlas%20do%20Potencial%20Eolico%20Brasileiro.pdf. Acesso em: 8 mar. 2024.
- ANEEL – AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA. *Micro e minigeração distribuída: Sistema de Compensação de Energia Elétrica*. Brasília, DF: Centro de Documentação (Cedoc), 2014.
- ASSUNÇÃO, J.; COSTA, F.; SZERMAN, D. *Local socioeconomic impacts of Brazilian hydroelectric power plants*. Rio de Janeiro: Fundação Getulio Vargas, 2017. Disponível em: <https://repositorio.fgv.br/server/api/core/bitstreams/9bfc787c-1642-45d8-a099-e6bf22e72681/content>. Acesso em: 8 mar. 2024.

COSTA, R. A.; CASOTTI, B. P.; AZEVEDO, R. L. S. Um panorama da indústria de bens de capital relacionados à energia eólica. *BNDES Setorial*, Rio de Janeiro, n. 29, p. 229-277, 2009. Disponível em: https://web.bndes.gov.br/bib/jspui/bitstream/1408/2183/2/BS%2029%20Um%20panorama%20da%20ind%3bastria%20de%20bens%20de%20capital_P.pdf. Acesso em: 8 mar. 2024.

COSTA, R. C.; PRATES, C. P. T. O papel das fontes renováveis de energia no desenvolvimento do setor energético e barreiras à sua penetração no mercado. *BNDES Setorial*, Rio de Janeiro, n. 21, p. 5-30, 2005. Disponível em: https://web.bndes.gov.br/bib/jspui/bitstream/1408/2436/2/BS%2021%20O%20papel%20das%20fontes%20renov%3a%1veis_P.pdf. Acesso em: 8 mar. 2024.

EPE – EMPRESA DE PESQUISA ENERGÉTICA. *Anuário estatístico de energia elétrica 2023: ano base 2022*. Rio de Janeiro: EPE, 2023. Disponível em: <https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-160/topico-168/anuario-factsheet.pdf>. Acesso em: 8 mar. 2024.

GOUVÊA, R. L. P.; SILVA, P. A. Desenvolvimento do setor eólico no Brasil. *Revista do BNDES*, Rio de Janeiro, v. 25, n. 49, p. 81-118, 2018. Disponível em: https://web.bndes.gov.br/bib/jspui/bitstream/1408/16081/1/PRArt_Desenvolvimento%20do%20setor%20e%3b%3lico%20no%20Brasil_compl.pdf. Acesso em: 8 mar. 2024.

KONZEN, E. *Penalizações tipo Lasso na seleção de covariáveis em séries temporais*. 2014. Dissertação (Mestrado em Economia) – Faculdade de Ciências Econômicas, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2014. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/103896/000939155.pdf?sequence=1&isAllowed=y>. Acesso em: 8 mar. 2024.

LAGE, E. S.; PROCESSI, L. D. Panorama do setor de energia eólica. *Revista do BNDES*, Rio de Janeiro, n. 39, p. 183-205, 2013. Disponível em: https://web.bndes.gov.br/bib/jspui/bitstream/1408/2926/1/RB%2039%20Panorama%20do%20setor%20de%20energia%20e%3b%3lica_P.pdf. Acesso em: 8 mar. 2024.

MARTINI, R. A. *et al.* *Uma solução automatizada para avaliações de impacto em estudos de caso: o Modelo Automatizado em R para Verificação de Impacto (MARVIm): módulo de controle sintético*. Rio de Janeiro: BNDES, 2018. (Textos para discussão, n. 130). Disponível em: https://web.bndes.gov.br/bib/jspui/bitstream/1408/15781/1/TD_130__Web%20%281%29.pdf. Acesso em: 8 mar. 2024.

RESENDE, B. C. M. *O efeito da implantação de usinas eólicas sobre o preço dos aluguéis*. 2015. Monografia (Graduação em Economia) – Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2015. Disponível em: https://www.econ.puc-rio.br/uploads/adm/trabalhos/files/Bruno_Cesar_Mariano_Resende.pdf. Acesso em: 8 mar. 2024.

RINTZEL, L. T.; ALVES, T. W.; MASSUQUETTI, A. Análise dos impactos econômicos decorrentes da instalação dos parques eólicos nos municípios da região Sul do Brasil. In: ENCONTRO DE ECONOMIA DA REGIÃO SUL, 20., 2017, Porto Alegre. *Anais [...]*. Niterói: Associação Nacional dos Centros de Pós-Graduação em Economia, 2017. Disponível em: https://www.anpec.org.br/sul/2017/submissao/files_I/i4-b152e3c00741c640d281063ece8999c0.pdf. Acesso em: 8 mar. 2024.

RODRIGUES, T. P.; GONÇALVES, S. L.; CHAGAS, A. L. S. Usinas eólicas e o mercado de trabalho nos municípios do Nordeste brasileiro. In: ENCONTRO NACIONAL DE ECONOMIA, 44., 2016, Foz do Iguaçu. *Anais [...]*. Niterói: Associação Nacional dos Centros de Pós-Graduação em Economia, 2016. Disponível em: https://www.anpec.org.br/encontro/2016/submissao/files_I/i10-d8cc205ee197d4af658822148a2874e7.pdf. Acesso em: 8 mar. 2024.

SIMAS, M. S. *Energia eólica e desenvolvimento sustentável no Brasil: estimativa da geração de empregos por meio de uma matriz insumo-produto ampliada*. 2012. Dissertação (Mestrado em Ciências) – Universidade de São Paulo, São Paulo, 2012. Disponível em: <https://www.teses.usp.br/teses/disponiveis/86/86131/tde-10092012-095724/en.php>. Acesso em: 8 mar. 2024.

TIBSHIRANI, R. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Oxford*, v. 58, n. 1, p. 267-288, 1996. Series B (Methodological). Disponível em: <https://www.jstor.org/stable/2346178>. Acesso em: 8 mar. 2024.

ZHAO, P.; YU, B. On model selection consistency of Lasso. *Journal of Machine Learning Research*, [s. l.], v. 7, p. 2541-2563, 2006. Disponível em: <https://www.jmlr.org/papers/volume7/zhao06a/zhao06a.pdf>. Acesso em: 8 mar. 2024.

ZOU, H. The adaptive Lasso and its oracle properties. *Journal of the American Statistical Association*, [s. l.], v. 101, n. 476, p. 1418-1429, 2006. Disponível em: <http://users.stat.umn.edu/~zouxx019/Papers/adalasso.pdf>. Acesso em: 8 mar. 2024.

APÊNDICE A – Fontes dos dados municipais

As 12 fontes de informações municipais que integram este trabalho estão descritas no Quadro 3.

Quadro 3 • Fontes de informações municipais

Sigla	Nome	Informações disponíveis
SIM-Datasus	Sistema de Informações sobre Mortalidade, do Departamento de Informática do Sistema Único de Saúde	Causas de mortes.
CNES-Datasus	Cadastro Nacional de Estabelecimentos de Saúde, do Departamento de Informática do Sistema Único de Saúde	Número de estabelecimentos e equipamentos de saúde.
IBGE	Instituto Brasileiro de Geografia e Estatística	População e PIB do município e sua desagregação nos setores de agropecuária, indústria, serviços e administração pública, tanto em proporção quanto em valor agregado.
ESTBAN	Estatística Bancária por Município	Número de agências bancárias e montante de empréstimos, financiamentos e outras operações bancárias, além do ativo e passivo total do setor.
Secex	Secretaria de Comércio Exterior, órgão do Ministério do Desenvolvimento, Indústria e Comércio Exterior	Volume de importações e exportações.
Rais	Relação Anual de Informações Sociais, do Ministério do Trabalho e Emprego	Número de empregados, vínculos empregatícios e estabelecimentos, salário médio e idade média dos empregados, com informações consolidadas em nível municipal.

Continua

Continuação

Sigla	Nome	Informações disponíveis
Finbra	Informações financeiras dos municípios, da Secretaria do Tesouro Nacional	Receitas e despesas públicas desagregadas por função. Inclui também ativos, passivos, patrimônio líquido e resultados acumulados da administração pública.
CE-Inep	Censo Escolar, do Instituto Nacional de Estudos e Pesquisas	Número de escolas, turmas e alunos matriculados, assim como sua média de idade.
CES-Inep	Censo do Ensino Superior, do Instituto Nacional de Estudos e Pesquisas	Número de estabelecimentos e de alunos, distribuídos por turno e por condição (entrantes ou concluintes), assim como sua média de idade.
SNIS	Sistema Nacional de Informações sobre Saneamento	Sistemas de água, esgoto e coleta de resíduos e acesso a eles.
Derivadas	Variáveis derivadas	Variáveis construídas a partir da comparação entre dois ou mais indicadores de diferentes fontes. A maior parte dessas variáveis indica valores integrais como proporção da população, das receitas do governo ou da área do município.

As principais informações sobre essas fontes de dados estão registradas na Tabela 4. Nela, estão representados, por fonte de dados, o número de variáveis, o menor ano inicial de cobertura, o maior ano final e as médias de preenchimento, tanto das variáveis em si como a proporção dos municípios cobertos com pelo menos uma observação. De forma geral, a base é mais completa no período de 2002 a 2014.

Tabela 4 • Informações gerais sobre as fontes das variáveis municipais

Fonte	Variáveis	Ano inicial	Ano final	Preenchimento (%)	Cobertura nos municípios (%)
Censo escolar	6	2007	2015	99,9	100,0
Derivadas	22	1999	2016	71,0	77,8
Ensino superior	11	2009	2014	85,5	86,0
ESTBAN	21	1999	2015	61,7	69,5
Finbra	76	2002	2015	83,7	96,3
IBGE	10	1999	2016	99,5	99,6
Rais	15	2002	2014	99,8	100,0
Secex	8	2000	2016	56,3	67,5
SNIS	36	1999	2014	49,8	80,5
Óbitos	5	1999	2014	93,9	100,0
Unidades SUS	37	2005	2015	56,4	56,4

Fonte: Elaboração própria.

APÊNDICE B – Resultados das estimações individuais de controle sintético

Quadro 4 • Resultados das estimações individuais de controle sintético (%)

Município	t-3	t-2	t-1	t	t+1	t+2	t+3
Acaraú – CE	2,47	2,56	-4,56	-11,64	-26,41	-21,66	-16,17
Amontada – CE	3,31	-8,08	1,06	3,22	4,91	12,97	-1,00
Aracati – CE	3,13	0,85	-0,46	-14,29	8,47	1,28	7,27
Areia Branca – RN	-18,71	-6,49	-1,25	3,33	-8,94	-15,01	-46,65
Barra dos Coqueiros – SE	-1,00	-11,08	-9,52	-7,93	-18,65	-9,92	-23,45
Beberibe – CE	0,16	0,04	0,08	16,33	13,40	11,78	19,99

Continua

Continuação

Município	t-3	t-2	t-1	t	t+1	t+2	t+3
Brotas de Macaúbas – BA	-4,11	9,12	-4,50	-1,39	6,52	-6,88	-8,31
Cabo de Santo Agostinho – PE	0,27	-0,18	0,14	6,12	13,02	11,22	14,47
Caetité – BA	0,12	0,03	0,13	11,22	0,69	1,53	28,10
Camocim – CE	1,11	-10,03	-15,61	-5,54	1,60	4,72	1,39
Galinheiros – RN	20,62	14,57	-26,94	-41,14	-45,41	-40,33	-33,31
Guamaré – RN	5,50	2,64	-5,53	6,33	-61,82	-102,59	-99,49
Guanambi – BA	0,49	-5,76	-3,56	4,77	4,44	9,87	10,08
Igaporã – BA	-6,09	3,69	1,38	5,45	14,63	11,19	129,35
Itarema – CE	-4,68	-3,16	-0,68	-3,61	19,45	8,67	8,13
João Câmara – RN	-3,12	2,38	1,70	2,94	23,60	62,32	62,45
Macaparana – PE	0,19	-1,19	0,50	2,02	-4,14	-4,17	-5,64
Mataraca – PB	-5,91	-6,37	-3,63	-24,47	-37,05	-48,25	-22,90
Palmares do Sul – RS	-7,38	-0,14	-1,16	13,04	-1,32	8,31	15,06
Paracuru – CE	-0,57	0,32	-0,14	0,54	18,35	4,94	16,64
Parazinho – RN	2,71	0,81	-5,40	273,48	1046,45	484,52	475,68
Parnaíba – PI	0,72	12,60	-1,59	-11,11	-7,04	-2,36	-4,69
Pedra Grande – RN	-1,19	4,34	1,52	32,71	36,85	16,81	344,73
Pombos – PE	0,00	0,00	0,00	5,29	16,52	13,60	5,44
Sant'Ana do Livramento – RS	3,89	-2,07	1,58	3,19	1,22	14,44	13,06
São Francisco de Itabapoana – RJ	-8,95	-2,03	4,83	14,62	29,70	105,47	145,20
São Miguel do Gostoso – RN	-0,78	-2,42	2,03	2,48	16,90	93,33	-
Sento Sé – BA	0,77	1,46	-2,78	5,55	-6,55	19,26	19,67
Sobradinho – BA	-7,18	-3,48	-14,66	-2,03	-48,41	-55,49	-44,22
Trairi – CE	-10,33	-1,85	1,59	8,43	41,49	42,88	32,04
Tramandaí – RS	0,77	1,33	-2,15	0,68	1,21	2,05	3,23
Tubarão – SC	0,00	0,00	0,00	7,62	-1,81	-	-
Xangri-lá – RS	-2,38	1,52	1,32	10,29	25,02	-	-

Fonte: Elaboração própria.

