**Avaliação de Efeitos Locais da Construção de Usinas Eólicas nos Municípios Brasileiros: Uma Abordagem por Controle Sintético**

**ANPEC- 46º Encontro Nacional de Economia**

**Área 11 – Economia Agrícola e do Meio Ambiente**

**Ricardo Agostini Martini[[1]](#footnote-1) (BNDES)**

**Marília de Figueiredo Jordão[[2]](#footnote-2) (BNDES)**

**Daniel da Silva Grimaldi[[3]](#footnote-3) (BNDES)**

**Resumo:** O presente trabalho tem o objetivo de verificar possíveis impactos da construção de usinas eólicas nas economias municipais. Mais especificamente, pretende-se aqui analisar a evolução do PIB per capita dos municípios que receberam investimentos na construção de usinas eólicas no período 2008-2014 em comparação a um grupo de outros municípios com características semelhantes, mas que não recebeu essa intervenção. O estudo utilizou uma base de dados municipais construída como uma consolidação de 250 variáveis quantitativas provenientes de 12 fontes diferentes, sendo utilizada uma metodologia apropriada 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 compilados são mais favoráveis em 2 anos após as construções e em parques eólicos com investimentos mais elevados.

**Palavras-chave:** Usinas Eólicas, Avaliação de Impacto, Controle Sintético

**Classificação JEL:** C23, Q42, R58

**Absract:** The present work aims to verify possible impacts of the construction of wind power plants in the municipal economies. More specifically, it intends to analyze the evolution of GDP per capita of the municipalities that received investments in the construction of wind power plants in the period 2008-2014 in comparison to a group of other municipalities with similar characteristics, but that did not receive this intervention. The study used a municipal database that consolidates 250 quantitative variables from 12 different sources, using an appropriate methodology to reduce the dimensionality of the parametric space, called adaLASSO, in order to select the most important covariates. The synthetic control method was applied for each case, adding 37 individual evaluations, which were then compiled and analyzed in an aggregate manner. The results obtained were heterogeneous, but with median positive effects, in agreement with the bibliography raised. The compiled effects are more favorable in 2 years after the constructions and in wind farms with higher investments.

**Keywords:** Wind Power Plants, Impact Avaliation, Synthetic Control Methods

**Classificação JEL:** C23, Q42, R58

**1. 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 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 o 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.

Nesse sentido, o presente 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 PIB per capita dos municípios que receberam investimentos na construção de usinas eólicas em comparação a um grupo de outros municípios com características semelhantes, mas que não recebeu essa intervenção. O período coberto pela avaliação vai de 2008 a 2014, devido à melhor disponibilidade de dados nesse período.

É necessário levar em consideração o fato de que, para que os resultados estimados no exercício empírico possam ser expandidos para uma análise da política efetuada, é necessário que as estimativas obtidas por essa técnica permitam inferências estatísticas confiáveis. Isto é, é preciso que se tenha confiança de que os resultados obtidos para a amostra reflitam o efeito da intervenção para a população como um todo. Para isso, é fundamental que a amostra de unidades tratadas tenha um volume apropriado. Esse ponto é sensível para muitas intervenções públicas, como, por exemplo, para o caso do impacto local, em nível geográfico, das grandes obras de infraestrutura. Nesse sentido, o presente texto apresenta uma metodologia adequada para os casos em que as intervenções atingem um número pequeno 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[[4]](#footnote-4).

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 compilados são mais favoráveis em 2 anos após as construções e em parques eólicos com investimentos mais elevados.

O estudo aqui realizado utilizou uma base de dados municipais construída como uma consolidação de 250 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), informações do Instituto Nacional de Pesquisas Espaciais, 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 o cruzamento dessas variáveis brutas. Por fim, o banco de dados da Agência Nacional de Energia Elétrica (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 no presente 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 *Adaptative 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. Essa 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, serão apresentados os resultados das estimações individuais compiladas. Por fim, serão apresentadas as considerações finais do trabalho e possíveis desdobramentos futuros dessa linha de pesquisa.

**2. Panorama do Setor de Energia Eólica no Brasil**

Conceitualmente, pode-se entender um parque eólico ou usina eólica como um conjunto de moinhos de vento ou turbinas que são usados para gerar energia elétrica através 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 que contribuem para o aquecimento global. 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, diferentemente do que ocorre com combustíveis fósseis, assim como há baixos riscos ambientais que podem haver em usinas hidrelétrica e nucleares, por exemplo (COSTA ET AL, 2009).

Atualmente, a maior fonte energética do Brasil advém de usinas hidrelétricas. Apesar de sua produção não poluente e barata, essa dependência pode acarretar danos. Isso porque, diante de períodos de seca, os reservatórios de água podem esvaziar e vê-se 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 possui um dos maiores volumes de ventos do mundo, assim como baixa probabilidade de ocorrência de fenômenos climáticos extremos, o país tem possibilidades concretas de ampliar seu uso de energia eólica. Segundo o Atlas do Potencial Eólico (Amarante et al, 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. Atualmente, 14 estados produzem energia elétrica através de usinas eólicas. Até 2014, a quantidade de energia eólica produzida era de 4 terawatt-hora por ano. O Brasil iniciou o mês de fevereiro de 2018 com capacidade instalada de 12,95 GW, 8,3% na matriz elétrica total do país (ABEEólica, 2018). Nesse mês, havia ainda 4,8 GW de capacidade em construção.

Os incentivos no setor de energia eólica no Brasil iniciaram-se no ano de 2002, através do Programa de Incentivo às Fontes Alternativas de Energia Elétrica (Proinfa). Os investimentos ocorrem por meio de leilões especificamente para a geração de eólicas, e também outras fontes de energias alternativas, por serem pouco competitivas em relação às fontes tradicionais (hidrelétrica, gás natural, óleo combustível, etc.) (Lage e Processi, 2013). O BNDES apoia o setor de energia eólica com linhas especiais de financiamento para empresas do segmento de geração, transmissão e distribuição de energia. Para os fabricantes de componentes de aerogeradores, o banco disponibiliza dois produtos: o FINEM, que visa apoiar investimentos em aumento da capacidade e construção de novas plantas; e o FINAME, que visa financiar a venda de máquinas e equipamentos já negociados com as respectivas compradoras (Costa et al, 2009).

O gráfico abaixo 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 os anos de 2014 e 2017. Dessa maneira, a maior parte dos parques eólicos não pôde ser avaliada neste trabalho por falta de dados nas variáveis de comparação referentes aos anos subsequentes a 2015 no banco de dados do IBGE.

**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 a seguir 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 do presente trabalho, definida pela disponibilidade de dados. A região Nordeste apresenta a maioria das usinas eólicas no país, com 78% do total. O estado que apresenta 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 (ANEEL). O mapa também evidencia uma concentração de usinas no período mais recente.

**Gráfico 2: 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 o desenvolvimento econômico de regiões agrícolas, com capacidade de melhorar, assim, esse setor, aumentando a qualidade de vida e reduzindo desigualdades sociais (AIE, 2002; Costa e 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. Isto se deve aos arrendamentos de suas terras para a instalação de torres eólicas (ABEEL, 2014).

Na literatura há uma crescente série de estudos empíricos que procuraram 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 no Brasil, tomando como base de dados primários uma série de entrevistas com dezoito parques eólicos e empresas de componentes para o setor e utilizando uma análise baseada em matriz insumo-produto. O trabalho chegou à conclusão que o setor de energia eólica tem o potencial de geração de 330 mil empregos até 2020, principalmente no setor de construção.

Estudando especificamente os efeitos em nível municipal, Resende (2015) procurou verificar o impacto da construção de usinas eólicas sobre os preços dos alugueis nos entornos. O trabalhou utilizou dados do Censo Demográfico (IBGE) de 1991 e 2010 para cobrir dezoito municípios que receberam alguma usina eólica entre esses anos. Por meio da metodologia do controle sintético, o estudo concluiu que os alugueis ficaram 8,3% mais caros nos municípios que receberam usinas eólicas do que foi previsto pelo grupo de controle.

Especificamente para os municípios da Região Nordeste, Rodrigues et al (2016) procurou analisar o impacto dos parques eólicos no nível de emprego, na massa salarial e no número de firmas em cada unidade. A análise partiu do cruzamento entre dados da ANEEL com dados de mercado de trabalho da Relação Anual de Informações Sociais (RAIS), do Ministério do Trabalho. Por meio de um pareamento por escore de propensão, o estudo verificou que os municípios com usinas eólicas apresentaram maior massa salarial e número de firmas que as demais, ao passo que o efeito sobre o emprego não foi significante.

Já para o caso da Região Sul, Rintzel et al (2017) cruzou dados de mercado de trabalho da RAIS com 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) em 1999, 2006 e 2013. O estudo procurou verificar efeitos da instalação de parques eólicos nos municípios por meio do método estrutural diferencial, o qual é baseado na comparação do desempenho dos municípios tratados com a média da região para cada variável. O trabalho verificou resultados positivos para todas as variáveis, especialmente para o nível de emprego.

Ou seja, os estudos empíricos levantados destacaram potenciais efeitos positivos da construção de usinas eólicas nas suas localidades. Todavia, esses estudos observaram casos específicos e não controlaram o problema da micronumerosidade de casos tratados, o que compromete inferências estatísticas sobre a avaliação. O único estudo que tratou esse problema pelo método do controle sintético trabalhou com apenas um ponto no período pré-tratamento, de modo que suas previsões podem não ter consistência. Dessa maneira, a contribuição do presente trabalho é exatamente utilizar uma metodologia apropriada para avaliação com poucos 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 et al (2016), o qual, ainda que tenha focado em usinas hidrelétricas, foi a principal inspiração metodológica do presente artigo. 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, tais como PIB per capita, taxa de crescimento do PIB, número de empregos formais, número de empresas formais e tamanho da população. Os autores utilizaram a metodologia do controle sintético, a qual, por basear-se na construção de 82 estudos de caso comparáveis, permite a estimação dos efeitos dinâmicos desde o início da construção das usinas, assim como o cálculo do efeito mediano da construção e a distribuição desses efeitos por ano, no curto e no médio prazo. O estudo observou que o impacto da construção das usinas hidrelétricas é estimulado apenas no curto prazo, tendendo a zero no 5º ano após o início das obras. O único impacto de médio prazo foi observado no nível de emprego formal no município. Por fim, os impactos nos municípios são muito heterogêneos, em todos os indicadores estimados.

**3. Base de Dados**

Com o objetivo de verificar possíveis efeitos locais da política de construção de usinas eólicas no Brasil, o presente 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 250 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:

- O Sistema de Informações sobre Mortalidade (SIM), do Departamento de Informática do Sistema Único de Saúde (DATASUS), com dados de causas de mortes;

- O Cadastro Nacional de Estabelecimentos de Saúde (CNES), também do DATASUS, com informações sobre número de estabelecimentos e equipamentos de saúde;

- Informações de economia e população do Instituto Brasileiro de Geografia e Estatística (IBGE), incluindo o PIB do município e sua desagregação nos setores de agropecuária, indústria, serviços e administração pública, tanto em proporção, quanto em valor agregado;

- Informações da Estatística Bancária por Município (ESTBAN), incluindo o número de agências bancárias e o montante de empréstimos, financiamentos e outras operações bancárias, além do ativo e passivo total do setor;

- Informações sobre importações e exportações da Secretaria de Comércio Exterior (SECEX), órgão do Ministério do Desenvolvimento, Indústria e Comércio Exterior (MDIC);

- Informações sobre desmatamento e cobertura florestal na Amazônia Legal do Instituto Nacional de Pesquisas Espaciais (INPE);

- A Relação Anual de Informações Sociais (RAIS), do Ministério do Trabalho e Emprego, com informações sobre número de empregados, vínculos empregatícios e de estabelecimentos, salário médio e idade média dos empregados, com informações consolidadas em nível municipal;

- Informações financeiras dos municípios (FinBra), banco de dados criado pela Secretaria do Tesouro Nacional (STN), em convênio com a Caixa Econômica Federal (CEF). Essa base inclui informações sobre receitas e despesas públicas desagregadas por função. Inclui também ativos, passivos, patrimônio líquido e resultados acumulados da administração pública;

- Dados do Censo Escolar, do Instituto Nacional de Estudos e Pesquisas (INEP), com informações sobre número de escolas, turmas e alunos matriculados, assim como sua média de idade;

- Dados do Censo do Ensino Superior, também do INEP, com informações sobre o número de estabelecimentos e de alunos, distribuídos por turno e por condição (entrantes ou concluintes), assim como sua média de idade;

- Dados do Sistema Nacional de Informações sobre Saneamento (SNIS) sobre os sistemas de água, esgoto e coleta de resíduos e o acesso aos mesmos.

- Além disso, a base municipal conta com variáveis derivadas, construídas a partir do cruzamento de dois ou mais indicadores de diferentes fontes. A maior parte dessas variáveis indicam valores integrais como proporção da população, das receitas do governo ou da área do município. Por exemplo, tem-se o PIB per capita, as despesas com habitação como proporção da receita corrente líquida da prefeitura, a área desmatada sobre a área total e a densidade demográfica do município.

As principais informações sobre essas fontes de dados estão registradas na Tabela 1. Nessa tabela 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.

**Tabela 1: 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 |
| INPE | 3 | 2000 | 2015 | 13,6 | 13,6 |
| 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.

Do total das 250 variáveis numéricas presentes na base de dados, o destaque são as contas públicas municipais, com um total de 76 indicadores, seguido do registro de unidades do SUS, com 37 indicadores. Em termos de cobertura temporal, a base de dados municipais consolidada abrange um período total de 1999 a 2016. O destaque são as bases do IBGE e as variáveis derivadas, que cobrem o período inteiro. Por outro lado, os censos das escolas e do ensino superior são relativamente mais restritos. De uma forma geral, a base é mais completa no período de 2002 a 2014.

Uma avaliação de impacto exige informações completas sobre os indicadores de interesse, assim como de suas covariadas. Essa necessidade é ainda mais importante em estimativas realizadas por controle sintético, as quais são mais consistentes quanto mais extenso for o período de ajuste. O Censo Escolar, a RAIS, o Sistema de Informações sobre Mortalidade e o IBGE apresentam as informações mais completas. Já o INPE parece apresentar as informações menos completas. Todavia, isso se deve ao fato de que suas variáveis são observadas apenas para os municípios da Amazônia Legal.

Em termos de cobertura por municípios, o percentual de preenchimento foi determinado como a proporção de municípios brasileiros com pelo menos uma informação ao longo do período de cobertura, para cada variável. Novamente, o Censo Escolar, a RAIS, o Sistema de Informações sobre Mortalidade e o IBGE apresentam as informações mais completas. De forma geral, observa-se que os indicadores médios de preenchimento e de cobertura por município são semelhantes em cada fonte. Isso sugere que os problemas de falta de dados em algumas séries devem-se mais à inexistência de dados para alguns municípios – sobretudo os de menor porte – do que à descontinuidade das séries. A exceção a essa tendência é a base do SNIS, que é presente em mais de 80% dos municípios brasileiros, mas com uma taxa de preenchimento total de menos de 50%.

O presente estudo de caso pretende calcular o impacto da construção de usinas eólicas sobre o PIB per capita dos municípios beneficiados. Devido à disponibilidade de dados, a avaliação considera como unidades tratadas os 37 municípios beneficiados de 2008 a 2014 por entrada em operação de sua primeira usina eólica, expostos na Tabela 2. Já os potenciais controles (isto é, o *donor pool*) inclui um total de 5.490 municípios que não possuem usinas eólicas. Outros 43 municípios que receberam sua primeira usina eólica antes de 2008 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. Para melhor capturar os efeitos da construção das usinas eólicas, o controle sintético foi aplicado para cada caso com um ano de defasagem em relação ao tratamento tal como definido.

**Tabela 2: Lista de Municípios Avaliados, Ano e Potência de sua Primeira Usina Eólica**

|  |  |  |
| --- | --- | --- |
| **Município** | **Tratamento** | **Potência Outorgada (MW)** |
| Acaraú - CE | 2010 | 70.800 |
| Amontada - CE | 2009 | 54.600 |
| Aracati - CE | 2008 | 10.500 |
| Areia Branca - RN | 2013 | 20.000 |
| Barra dos Coqueiros - SE | 2012 | 34.500 |
| Beberibe - CE | 2008 | 25.600 |
| Boituva - SP | 2012 | 2,24 |
| Brotas de Macaúbas - BA | 2012 | 95.190 |
| Cabo de Santo Agostinho - PE | 2012 | 2.000 |
| Caetité - BA | 2014 | 296.820 |
| Camocim - CE | 2009 | 105.000 |
| Cururupu - MA | 2008 | 22,5 |
| Galinhos - RN | 2014 | 118.570 |
| Gravatá - PE | 2010 | 14.850 |
| Guamaré - RN | 2010 | 51.000 |
| Guanambi - BA | 2014 | 167.840 |
| Igaporã - BA | 2014 | 143.840 |
| Itarema - CE | 2014 | 30.000 |
| Iturama - MG | 2012 | 156 |
| João Câmara - RN | 2012 | 39.600 |
| Macaparana - PE | 2010 | 4.950 |
| Palmares do Sul - RS | 2010 | 9.200 |
| Paracuru - CE | 2008 | 25.200 |
| Parazinho - RN | 2014 | 466.000 |
| Parnaíba - PI | 2008 | 18.000 |
| Pedra Grande - RN | 2014 | 118.400 |
| Pelotas - RS | 2014 | 1,98 |
| Pombos - PE | 2010 | 4.950 |
| Sant'Ana do Livramento - RS | 2011 | 60.000 |
| São Francisco de Itabapoana - RJ | 2010 | 28.050 |
| São Miguel do Gostoso - RN | 2014 | 51.200 |
| Sento Sé - BA | 2013 | 90.000 |
| Sobradinho - BA | 2013 | 48.000 |
| Trairi - CE | 2013 | 55.392 |
| Tramandaí - RS | 2011 | 70.000 |
| Tubarão - SC | 2014 | 2.099,5 |
| Xangri-lá - RS | 2014 | 27.675 |

Fonte: Elaboração própria com base em dados da ANEEL.

A tabela a seguir apresenta estatísticas descritivas sobre os municípios tratados em relação ao total de municípios elegíveis como unidades de comparação (*donor pool*), considerando-se o ano imediatamente anterior aos primeiros tratamentos (2007). Apresenta-se a média e o desvio-padrão de 16 indicadores com o objetivo de apresentar um perfil dos dois grupos de municípios. Conforme mostra a tabela, os tratados estavam, em 2007, com uma renda anual média por habitante ligeiramente acima da média nacional. Além disso, apresentam maior porte do que os não-tratados em população, número de estabelecimentos, receitas totais, despesas com saúde e educação e extensão da rede de água. Já os potenciais controles apresentam vantagem em termos de PIB, receitas tributárias, menores homicídios per capita e extensão da rede de esgotos. Pelos dados apresentados, não é possível inferir diferenciais de níveis de desenvolvimento entre os dois grupos.

**Tabela 3: Perfil das Unidades (2007)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variável** | ***Donor Pool*** | | **Tratados** | |
| **Média** | **Desvio-Padrão** | **Média** | **Desvio-Padrão** |
| PIB (R$) | 488.272.016,86 | 5.279.110.957,38 | 449.717.972,97 | 665.679.901,85 |
| PIB per Capita (R$) | 9.194,31 | 10.845,34 | 9.760,24 | 10.238,24 |
| Adm. Pública/PIB (%) | 34,04 | 17,69 | 32,17 | 16,71 |
| Agropecuária/PIB (%) | 22,10 | 15,29 | 11,30 | 8,72 |
| Indústria/PIB (%) | 13,52 | 15,15 | 23,02 | 23,74 |
| Serviços/PIB (%) | 30,34 | 12,55 | 33,52 | 15,89 |
| População | 32.522,25 | 196.462,44 | 49.570,16 | 60.348,03 |
| N. Estabelecimentos | 460,69 | 3.702,55 | 555,22 | 1.044,92 |
| N. Vínculos de Emprego | 6.687,53 | 71.966,34 | 6.465,81 | 11.050,27 |
| Receita Total (R$) | 38.876.013,24 | 331.122.457,73 | 51.145.459,31 | 61.111.215,44 |
| Receita Tributária (R$) | 7.419.508,94 | 138.466.444,76 | 5.037.272,07 | 7.530.259,74 |
| Desp. Educação (R$) | 9.571.826,20 | 70.131.023,63 | 14.224.404,77 | 13.447.551,95 |
| Desp. Saúde (R$) | 8.318.443,56 | 59.459.079,07 | 10.252.666,07 | 14.744.707,09 |
| Óbitos por Homicídios por 100 mil habitantes | 12,89 | 17,39 | 16,06 | 17,90 |
| Extensão da Rede de Água (km) | 99,45 | 429,73 | 137,74 | 179,28 |
| Extensão da Rede de Esgotos (km) | 89,90 | 142,44 | 54,11 | 90,08 |
| Total de Municípios | 5.490 |  | 37 |  |

Fonte: Elaboração própria com base em dados do IBGE, RAIS, Finbra, SIM-DATASUS e SNIS.

**4. Metodologia**

*4.1. Controle Sintético*

A análise de impacto de intervenções públicas em nível de localidade, como, por exemplo, da construção de usinas eólicas sobre o desempenho econômico dos municípios contemplados, traz uma série de desafios de ordem empírica (ASSUNÇÃO ET AL, 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 se construir obras de infraestrutura em localidades específicas é baseada em fatores políticos, econômicos e ambientais. Isto é, como a escolha 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.

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 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 et al (2003, 2010, 2015).

Supõe-se uma amostra de unidades (por exemplo, municípios) indexadas por j. A unidade é a unidade de interesse, isto é, a unidade tratada por uma intervenção. As demais unidades, de a constituem o chamado *donor pool*, isto é, 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 é um painel balanceado, isto é, que é composta por dados longitudinais e que todas as suas unidades são observadas no período de . A amostra inclui um número positivo de períodos pré-intervenção T0, assim como de períodos pós-intervenção T1, de modo que . A intervenção que será avaliada consiste na exposição da unidade ao tratamento durante os períodos , considerando que essa intervenção não tem efeitos durante o período pré-tratamento . 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 comparado com a unidade tratada. O controle sintético é representado por um vetor de pesos , tal que para e Dessa maneira, a escolha de qualquer valor particular de W é equivalente à escolha de um controle sintético.

Seja X1 um vetor contendo as características pré-tratamento da unidade tratada, as quais pretende-se aproximar o máximo possível. X0, por sua vez, é uma matriz contendo 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. 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 , 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. Para , seja X1m o valor da variável m para a unidade tratada e X0m um vetor que contém os valores da variável m para as unidades do *donor pool*, deve-se escolher o W\* que minimiza:

Nessa equação, vm é um peso que reflete a importância relativa atribuída à variável m quando se mede a discrepância entre X1 e X0W.

Seja Yjt a variável de interesse da unidade j no tempo t. Y1 é um vetor dos valores pós-intervenção da variável de interesse para a unidade tratada, de modo que . Y0 é uma matriz em que a coluna j contém os valores pós-intervenção da variável de interesse para a unidade . Dessa maneira, a variável de interesse do controle sintético é .

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:

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 ET AL, 2010). Segundo esse estudo, esse fator contribui por 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.

Em caso 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 ET AL, 2016). Ou seja, ao invés de se considerar uma única unidade tratada , considera-se um conjunto de G unidades tratadas por uma intervenção, as quais são indexadas por . Seja T0g o ano em que houve o tratamento em cada unidade tratada. Para melhor comparar o tratamento em diferentes anos, esses são normalizados em , em que é o ano do tratamento de cada unidade em G.

Após a normalização temporal, seja o efeito estimado da intervenção na unidade , pertencente ao conjunto J de unidades tratadas e não tratadas no período τ. Por exemplo, pode referir-se 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:

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:

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 estar descolada do seu controle sintético desde antes da intervenção. Em segundo lugar, pode haver heterocedasticidade de choques idiossincráticos. 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, poderá haver superestimação do efeito do tratamento. Se os choques atingirem o controle sintético com maior variância do que a unidade tratada, pode haver subestimação do tratamento.

Os dois problemas – viés e heterocedasticidade de choques – são detectados no exercício empírico por meio de um mau ajuste 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 se controlar a influência desses casos na estimação do efeito compilado das intervenções é a eliminação das unidades tratadas com um nível de *RMSPE* pré-tratamento acima de um patamar escolhido.

*4.2. AdaLASSO –Método de Seleção de Covariadas*

Os modelos de alta dimensionalidade estão cada vez mais presentes na literatura, já que a inclusão de um grande número de variáveis pode contribuir para ganhos de capacidade preditiva dos modelos (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 com a identificação das covariadas mais importantes, assim como com a obtenção de um modelo mais parcimonioso. Da mesma maneira, reduz a complexidade do modelo e o seu custo em termos computacionais.

O LASSO (*Least Absolute Shrinkage and Selection Operator*) 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. O 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 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:

Aqui, é uma função de um parâmetro de ajuste t, de modo que . Quanto maior o valor de λ, maior será a penalidade imposta ao somatório dos coeficientes. Por outro lado, se , 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, e o modelo é iterativamente estimado eliminando-se uma subamostra. Assim, fazem-se previsões com base em cada estimação e compara-se 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 encolher 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 E 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:

Ou seja, wj é um vetor de pesos individuais para auxiliar a seleção das variáveis relevantes para o modelo. Esse vetor é função de , isto é, dos coeficientes estimados por meio de uma regressão que controla possíveis efeitos de multicolinearidade (regressão *ridge*). Uma variável relevante xj terá um valor comparativamente elevado, o que reduz o peso wj atribuído ao seu coeficiente. Por outro lado, uma variável irrelevante xj terá um valor comparativamente baixo, o que aumenta o peso wj 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 wj apropriados, o adaLASSO possui 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.

**5. Resultados**

O método do controle sintético foi aplicado individualmente para comparar cada município tratado com o seu contrafactual. Foram realizadas um total de 37 avaliações individuais. O intuito foi fazer uma análise agregada de todos os municípios brasileiros que tiveram usinas eólicas entrando em operação de 2008 a 2014 a partir de cada avaliação individual. Para melhor capturar o efeito da construção das usinas, o controle sintético foi defasado em um ano para cada tratamento. 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, foram desconsiderados do trabalho três municípios: Guamaré (RN), Sobradinho (BA) e Parazinho (RN). Os dois primeiros apresentaram resultados discrepantes dos outros por razões alheias à produção de energia eólica. Guamaré sofreu sua a maior crise no ano de 2011 por conta da queda dos preços do petróleo na região, setor ao qual sua economia é muito dependente. Sobradinho teve uma de suas maiores secas em 2013, a qual esvaziou seu reservatório hídrico e comprometeu o fornecimento de energia advinda de sua usina hidrelétrica. Para o caso do município de Parazinho, houve problemas com o ajuste no período pré-tratamento no controle sintético, de modo não foi encontrado um contrafactual consistente para a verificação de impacto. Seu *RMSPE* pré-tratamento foi igual a 51,5, muito acima da média da amostra (5,9). Restaram 34 casos para avaliação. Os resultados das avaliações individuais encontram-se em apêndice ao presente 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 circular no entorno de 50% ao longo do tempo – tal qual ocorre com a proporção de caras após n lançamentos de uma moeda não viciada.

A tabela a seguir 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[[5]](#footnote-5). Para julgar se esses valores são estatisticamente significantes, um intervalo de confiança foi construído com base numa 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 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 a tabela, é possível observar o descolamento da proporção de efeitos positivos em relação ao intervalo de confiança a partir de 2 anos após o tratamento, nova evidência favorável ao impacto positivo das usinas sobre as economias locais.

**Tabela 4: 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 | 18 | 16 | 0,53 | 0,34 | 0,66 | 0,5 |
| t+1 | 22 | 12 | 0,65 | 0,34 | 0,66 | 0,5 |
| t+2 | 26 | 8 | 0,76 | 0,34 | 0,66 | 0,5 |
| t+3 | 17 | 7 | 0,71 | 0,31 | 0,69 | 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 a 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 et al (2016), os quais também depararam-se com resultados individuais heterogêneos sobre os efeitos municipais de obras de usinas hidrelétricas. Os resultados encontrados estão representados na tabela a seguir.

**Tabela 5: Resultados Compilados por Ano de Referência (%)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Ano** | **Média** | **DP** | **P. 0,25** | **Mediana** | **P. 0,75** | **Mínimo** | **Máximo** |
| t-3 | -0,13 | 4,22 | -2,34 | 0,00 | 1,71 | -9,71 | 13,54 |
| t-2 | -0,64 | 5,49 | -2,59 | 0,00 | 2,15 | -16,84 | 10,23 |
| t-1 | -1,87 | 6,96 | -3,61 | -0,81 | 0,00 | -28,12 | 16,33 |
| t | 1,80 | 8,72 | -3,86 | 0,86 | 5,55 | -16,16 | 26,08 |
| t+1 | 9,82 | 24,39 | -3,73 | 7,13 | 15,05 | -30,42 | 122,18 |
| t+2 | 23,49 | 61,01 | 1,68 | 9,37 | 16,22 | -25,19 | 338,07 |
| t+3 | 9,33 | 23,66 | -0,47 | 7,42 | 13,39 | -27,32 | 80,40 |

Fonte: elaboração própria.

Construiu-se um gráfico que 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 terceiro quartis, respectivamente. Observa-se que, em geral, há efeito mediano positivo no período pós-tratamento dos municípios sob análise, de 7,1% no primeiro ano, 9,4% no segundo ano e 7,4% no terceiro ano.

**Gráfico 3: Diferenças Percentuais de cada Unidade Tratada em Relação ao seu Controle Sintético**



Fonte: elaboração própria.

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 potência outorgada fixado no período em que as primeiras usinas estavam em construção e o indicador foi normalizado pela sua média de modo a permitir uma medida mais intuitiva de comparação.

Com base no valor mediano desse indicador de dose, equivalente a -0,35, ou aproximadamente 29.025 MW/R$ per capita, a amostra foi separada em dois grupos: dose baixa e dose alta. O gráfico abaixo relaciona a dose com o impacto mediano sobre o PIB per capita. Observa-se que o gráfico destaca efeito positivo maior em doses mais altas do que em doses baixas, com destaque para o segundo ano desde a construção da usina.

**Tabela 6: Resultados Compilados por Dose de Tratamento (%)**

|  |  |  |
| --- | --- | --- |
| **Ano** | **Dose Alta** | **Dose Baixa** |
| t-3 | 0,00 | 0,00 |
| t-2 | 1,83 | 0,00 |
| t-1 | -2,86 | 0,00 |
| t | 1,84 | 0,78 |
| t+1 | 11,87 | 0,48 |
| t+2 | 13,64 | 8,09 |
| t+3 | 8,81 | 5,56 |

Fonte: elaboração própria.

**6. Conclusão**

O presente 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. Após isso, os resultados individuais foram compilados. Com base nos 34 casos aqui avaliados, foi possível observar impacto positivo dos parques eólicos sobre o PIB per capita dos municípios afetados, ainda que haja expressiva variabilidade. 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 7,1% e 9,4%, foram mais claros para municípios relativamente mais pobres ou que receberam parques maiores e entre 2 a 3 anos após o início da construção.

A metodologia aqui adotada é inovadora em relação às avaliações dos impactos locais das obras de usinas eólicas no Brasil[[6]](#footnote-6). Para controlar a dimensionalidade do modelo empírico, foi adotada a metodologia de seleção de covariadas denominada adaLASSO. Esse procedimento visa identificar e selecionar as variáveis mais importantes para explicar cada indicador de interesse escolhido. A estimação dos efeitos locais foi realizada por meio da metodologia de controle sintético. Esse método não apenas é mais consistente para avaliações com micronumerosidade de unidades tratadas, como também, ao contrário da literatura levantada, 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 termos de sua heterogeneidade entre as unidades avaliadas, são semelhantes àqueles observados por Assunção et al (2016), em um estudo sobre efeitos locais da construção de usinas hidrelétricas. Esse trabalho também observou que os efeitos locais das obras tendem a sumir no 5º ano após a construção. Essa constatação não pôde ser observada no presente trabalho devido à indisponibilidade de dados, principalmente devido ao fato de que as usinas eólicas são de construção relativamente mais recente, a partir de 2014. Todavia, há motivos para que as usinas eólicas tenham efeitos locais mais duradouros que as usinas hidrelétricas, já que não implicam necessariamente na perda de terras que poderiam ser utilizadas na produção agropecuária.

É importante destacar que, de acordo com os dados da ANEEL, o período de maior crescimento de usinas eólicas entrando em operação no Brasil foi a partir de 2014. O presente estudo investiga parques eólicos instalados até 2014, ou seja, conta com uma base de dados pequena e um curto período para análise. Não obstante, há municípios que só foram analisados apenas um ano após o tratamento por falta de dados na base municipal.

Portanto, há uma potencial agenda de estudos futuros sobre a avaliação das eólicas sobre os municípios beneficiados. Espera-se que essa agenda seja cada vez mais consistente e mostre maiores evidências de impacto, por haver maior dosagem de investimento nos municípios e também por conter um banco de dados com maior quantidade de tratados.

**7. Referências Bibliográficas**

ABEEólica - Associação Brasileira de Energia Eólica. *Dados Mensais*. Fevereiro de 2018.

ANEEL - Agência Nacional de Energia Elétrica

ABADIE, A., & GARDEAZABAL, J. The economic costs of conflict: A case study of the Basque Country. *The American Economic Review*, 93(1), 113-132, 2003.

ABADIE, A., DIAMOND, A., & HAINMUELLER, J. Synthetic control methods for comparative case studies: Estimating the effect of California’s tobacco control program. *Journal of the American Statistical Association*, 105(490), 493-505, 2010.

ABADIE, A., DIAMOND, A., & HAINMUELLER, J. Synth: An R Package for Synthetic Control Methods in Comparative Case Studies. *Journal of Statistical Software*, v. 42, n. 13, jun. 2011.

ABADIE, A., DIAMOND, A., & HAINMUELLER, J. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495-510, 2015.

ALBUQUERQUE, B.; BUCHBINDER, F.; GRIMALDI, D.; MARTINI, R.; PEREIRA, J. P.; PINTO, A.; TORTORELLI, M. *Uma solução automatizada para avaliações quantitativas de impacto: primeiros resultados do MARVIm*. 2017.

AMARANTE, O.; ZACK, M.; SÁ, A. *Atlas do Potencial Eólico Brasileiro*. 2001.

ASSUNÇÃO, J.; COSTA, F.; SZERMAN, D. *Local Socioeconomic Impacts of Brazilian Hydroelectric Power Plants*. 2016.

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*, n. 21, p. 5-30, mar. 2005.

COSTA, R.; CASOTTI, B.; AZEVEDO, R. Um Panorama da Indústria de Bens de Capital Relacionados à Energia Eólica - *BNDES Setorial*, 2009.

GRIMALDI, D.; MARTINI, R.; PEREIRA, J. P.; JORDÃO, M.; TORTORELLI, M. *Uma solução automatizada para avaliações de impacto em estudos de caso: o MARVIm Módulo de Controle Sintético*. 2018, no prelo.

KONZEN, E. *Penalizações tipo Lasso na seleção de covariáveis em séries temporais*. Dissertação de Mestrado em Economia. Universidade Federal do Rio Grande do Sul, Faculdade de Ciências Econômicas. 2014.

LAGE, E. S.; PROCESSI, L. D. Panorama do Setor de Energia Eólica. *Revista do BNDES*, v. 39, junho, 2013.

RESENDE, B. C. M. *O Efeito da Implantação de Usinas Eólicas sobre o Preço dos Alugueis*. Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, Monografia de Final de Curso. 2015.

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*. Anais do XX Encontro de Economia da Região Sul. 2017.

RODRIGUES, T. P.; GONÇALVES, S. L.; CHAGAS, A. L. S. *Usinas Eólicas e o Mercado de Trabalho nos Municípios do Nordeste Brasileiro*. 2016.

SIMAS, M. S. *Energia Eólica e Desenvolvimento Sustentável no Brasil: Estimativa da Geração de Empregos por meio de uma matriz Insumo-Produto Ampliada*. Dissertação apresentada ao Programa de Pós-Graduação em Energia da Universidade de São Paulo para a obtenção do título de Mestre em Ciências. 2012.

TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society*. Series B (Methodological), v. 58, 267-288, 1996.

ZHAO, P.; YU, B. On model selection consistency of Lasso. *The Journal of Machine Learning Research*, v. 7, 2541-2563, 2016.

ZOU, H. *The adaptive lasso and its oracle properties*. Journal of the American Statistical Association, v. 101,1418-1429, 2006.

**Apêndice: Resultados das Estimações Individuais de Controle Sintético**

**Tabela 7: Resultados das Estimações Individuais de Controle Sintético (%)**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Município** | **t-3** | **t-2** | **t-1** | **t** | **t+1** | **t+2** | **t+3** |
| Acaraú - CE | 0,00 | 0,00 | 0,00 | -2,80 | 6,53 | 9,41 | 5,62 |
| Amontada - CE | 5,53 | -2,33 | -3,55 | 1,84 | -4,10 | 2,23 | 7,91 |
| Aracati - CE | 3,69 | -3,59 | -1,23 | -9,00 | -8,48 | -7,22 | -0,34 |
| Areia Branca - RN | -6,79 | -13,14 | -16,43 | -16,16 | -6,46 | -17,61 | -27,32 |
| Barra dos Coqueiros - SE | 0,00 | 0,00 | 0,00 | -8,72 | -11,44 | -9,80 | -19,97 |
| Beberibe - CE | 0,00 | 0,00 | 0,00 | -0,15 | 18,24 | 16,71 | 15,68 |
| Boituva - SP | 0,00 | 0,00 | 0,00 | 0,78 | 9,34 | 12,94 | 18,79 |
| Brotas de Macaúbas - BA | -4,11 | 9,12 | -4,50 | -1,39 | 6,52 | -6,88 | -8,31 |
| Cabo de Santo Agostinho - PE | 0,08 | 0,42 | 0,02 | -5,63 | -0,78 | 12,82 | 11,03 |
| Caetité - BA | -0,65 | 1,92 | -0,49 | -1,70 | 32,01 | 44,68 | #N/D |
| Camocim - CE | 4,99 | -5,25 | -8,54 | -4,21 | -4,76 | 5,56 | -1,11 |
| Cururupu - MA | -3,65 | -2,34 | -1,35 | 2,97 | -2,60 | 9,32 | 7,73 |
| Galinhos - RN | 13,54 | -16,84 | -28,12 | -9,57 | -30,42 | -10,95 | #N/D |
| Gravatá - PE | -2,28 | -2,17 | -1,45 | 3,05 | 8,97 | 13,98 | 5,63 |
| Guamaré - RN | 1,86 | 3,77 | 17,28 | 20,73 | 35,54 | -58,11 | -102,57 |
| Guanambi - BA | -2,45 | 2,28 | 3,54 | 4,55 | 8,39 | 13,64 | #N/D |
| Igaporã - BA | 1,57 | 3,97 | 8,55 | 6,85 | 122,18 | 84,10 | #N/D |
| Itarema - CE | 4,37 | 2,81 | -4,63 | 24,47 | 29,05 | 26,93 | #N/D |
| Iturama - MG | -7,8 | -11,1 | -6,62 | 11,46 | 25,97 | 22,34 | -0,87 |
| João Câmara - RN | -4,55 | 2,80 | 2,32 | 4,30 | 33,29 | 73,93 | 71,00 |
| Macaparana - PE | 0,00 | 0,00 | 0,00 | 0,93 | 0,48 | -4,62 | 3,30 |
| Palmares do Sul - RS | -0,31 | -2,92 | -0,25 | -4,51 | -7,49 | -25,19 | -20,35 |
| Paracuru - CE | -3,23 | 2,22 | -1,31 | 4,93 | 22,77 | 4,09 | 7,11 |
| Parazinho - RN | -75,69 | -32,54 | 114,68 | 30,32 | 70,51 | 45,84 | #N/D |
| Parnaíba - PI | 0,00 | 0,00 | 0,00 | -5,65 | -6,72 | -0,47 | 5,56 |
| Pedra Grande - RN | 2,49 | 10,23 | 16,33 | 26,08 | 13,13 | 338,07 | #N/D |
| Pelotas - RS | -0,66 | 4,77 | 2,68 | -5,68 | -2,57 | 1,50 | #N/D |
| Pombos - PE | 4,25 | 1,19 | -1,13 | 9,58 | 7,73 | 12,32 | 8,68 |
| Sant'Ana do Livramento - RS | -0,16 | 5,17 | -2,87 | 9,86 | 11,87 | 9,32 | 8,81 |
| São Francisco de Itabapoana - RJ | 2,10 | -3,59 | 2,54 | 12,05 | 10,85 | 8,09 | 80,40 |
| São Miguel do Gostoso - RN | -0,77 | 1,74 | -2,86 | -1,27 | 15,35 | 84,36 | #N/D |
| Sento Sé - BA | 0,57 | 1,83 | -7,56 | -0,08 | -12,10 | 14,73 | 12,63 |
| Sobradinho - BA | -5,68 | 2,51 | -1,37 | -0,79 | -3,77 | -27,25 | -24,22 |
| Trairi - CE | -9,71 | -7,33 | -3,63 | 5,75 | 26,84 | 34,02 | 16,10 |
| Tramandaí - RS | 1,76 | 1,15 | -4,06 | 7,07 | 3,82 | 7,10 | 16,29 |
| Tubarão - SC | 0,00 | 0,00 | 0,00 | 1,36 | 14,15 | 6,64 | #N/D |
| Xangri-lá - RS | -2,36 | -2,67 | 1,04 | -0,01 | 4,42 | 12,67 | #N/D |

Fonte: Elaboração própria.

1. Economista do BNDES. Mestre em Economia pelo CEDEPLAR/UFMG. [↑](#footnote-ref-1)
2. Estagiária do BNDES. Graduanda em Estatística na UFF. [↑](#footnote-ref-2)
3. Economista do BNDES. Mestre em Economia pela FEA/USP. [↑](#footnote-ref-3)
4. Ver, por exemplo, Abadie e Gardeazabal (2003) e Abadie et al (2010, 2011, 2015). [↑](#footnote-ref-4)
5. O Ano “t” equivale ao ano de construção do parque eólico, sendo “t+1” o ano em que ele entrou em operação. [↑](#footnote-ref-5)
6. Esse trabalho integra uma iniciativa mais ampla, desenvolvida pelo Departamento de Avaliação e Promoção da Efetividade (DEAPE), da Área de Planejamento Estratégico (AP) do Banco Nacional de Desenvolvimento Econômico e Social (BNDES), com o objetivo de dar escala a suas atividades de monitoramento e avaliação. O produto dessa iniciativa é a construção do Modelo Automatizado em R para Verificação de Impacto (MARVIm), uma metodologia de avaliação de efeitos de intervenções públicas de maneira automatizada. Para isso, combina uma série de scripts e funções em linguagem R que automatizam exercícios de avaliação quantitativa de impacto. O primeiro módulo do MARVIm foi voltado para métodos baseados em pareamento e diferenças-em-diferenças (ALBUQUERQUE ET AL, 2017), e o segundo, para avaliações baseadas em controle sintético (GRIMALDI ET AL, 2018 , no prelo). [↑](#footnote-ref-6)