

Complexidade econômica e crescimento do PIB per capita: uma análise de diferenças-em-diferenças para os municípios brasileiros

Danielle Evelyn de Carvalho¹
Alexandre de Queiroz Stein²
Arthur Ribeiro Queiroz³
João Prates Romero⁴

RESUMO

Os trabalhos sobre complexidade econômica mostraram, nos últimos anos, que regiões e países com estruturas produtivas mais avançadas tendem a obter maior crescimento. Esta pesquisa possui o objetivo de avaliar se aumento de complexidade econômica implicou crescimento do PIB *per capita* nos municípios brasileiros, considerando o período de 2009 a 2019. Para responder a isso, unimos métodos de avaliação de impacto com o arcabouço de complexidade econômica, utilizando a metodologia do Propensity Score Matching (PSM) em conjunto com o modelo de diferenças-em-diferenças (DID). Foram definidos como municípios que receberam o tratamento aqueles cuja posição no ranking de complexidade aumentou mais que 0,3, 1,1, 1,5 ou 2 desvios-padrão em relação à média das variações no ranking. Os resultados indicam que, independentemente do ponto de corte estabelecido, o aumento de complexidade econômica gerou um crescimento no PIB *per capita* dos municípios no período.

Palavras-chave: complexidade econômica; avaliação de impacto; *Propensity Score Matching*; diferenças-em-diferenças.

ABSTRACT

Studies on economic complexity showed in the last years that regions and countries with more advanced productive structures tend to achieve greater growth. This research aims to assess whether an increase in economic complexity implied in greater GDP per capita growth in Brazilian municipalities, considering the years between 2009 and 2019. We combine impact assessment methods with the economic complexity perspective. We used the Propensity Score Matching (PSM) methodology in conjunction with the differences-in-differences (DID) model. Municipalities that received the treatment were defined as those whose position in the economic complexity ranking increased by more than 0.3, 1.1, 1.5 or 2 standard deviations from the mean of the ranking variations. The results show that an increase economic complexity implied in growth in the per capita GDP of the municipalities in the period, regardless of the cut-off points established.

Keywords: economic complexity; impact evaluation; *Propensity Score Matching*; difference-in-differences.

JEL classification: O14, O25

Área 6 - Crescimento, Desenvolvimento Econômico e Instituições

¹ Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR-UFMG).

² Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR-UFMG).

³ Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR-UFMG).

⁴ Centro de Desenvolvimento e Planejamento Regional (CEDEPLAR-UFMG).

1. Introdução

Desde a publicação dos trabalhos que nortearam a abordagem da complexidade econômica, diversos autores vêm avaliando a relação da complexidade com as demais variáveis que caracterizam o processo de desenvolvimento econômico. Conforme aponta Hidalgo (2021), vários trabalhos examinaram a relação da complexidade com o crescimento da renda (HIDALGO; HAUSMANN, 2009; HAUSMANN *et al.*, 2011; FELIPE *et al.*, 2012), com a desigualdade de renda (HARTMANN *et al.*, 2017) e com a emissão de gases de efeito estufa (ROMERO; GRAMKOW, 2021). Estes trabalhos indicam que a complexidade econômica seria uma variável importante para compreender o processo de desenvolvimento e uma boa preditora para o crescimento da renda.

Contudo, os estudos sobre a relação entre complexidade e crescimento da renda não foram capazes de estabelecer com clareza uma relação causal entre as variáveis. Nos estudos seminais sobre o tema, os métodos utilizados não lidam de forma adequada com possíveis problemas de endogeneidade, de forma que seus resultados precisam ser interpretados com cuidado.

O presente artigo busca contribuir para a literatura sobre complexidade e crescimento aplicando, pela primeira vez, metodologias de avaliação de impacto para investigar essa relação. Diversos trabalhos utilizam a metodologia do *Propensity Score Matching* (PSM) em conjunto com o modelo de diferenças-em-diferenças (DID) para avaliação de políticas públicas. Almeida (2014), por exemplo, analisou o impacto dos royalties do petróleo sobre os gastos com educação e saúde nos municípios brasileiros que os receberam. Jacques (2013) investigou o impacto das cooperativas de crédito sobre o PIB *per capita* nos municípios brasileiros. Filho e Sousa (2018), por sua vez, avaliaram a implantação de Guardas Municipais sobre indicadores de segurança das cidades.

O objetivo do presente artigo é avaliar o efeito do aumento da complexidade econômica sobre o crescimento do PIB *per capita* a partir de dados de municípios brasileiros para o período entre 2009 e 2019. Para isso, foram criados critérios para identificar os municípios que tiveram avanço em termos da complexidade em relação aos demais. Quatro níveis de variação no ranking de complexidade foram utilizados (0,3 desvio-padrão (D.P.), 1,1 D.P., 1,5 D.P. e 2 D.P.) para classificar os municípios em grupos de tratado e controle. A partir disso, encontramos municípios que possuem características observáveis parecidas, por meio da metodologia de *Propensity Score Matching* (PSM), com o intuito de captar o que teria acontecido com a cidade caso ela não tivesse aumentado seu nível de complexidade – grupo controle. Em seguida, foi então utilizado o modelo de diferenças-em-diferenças (DID) para verificar as diferenças de taxas de crescimento da renda entre os dois grupos.

2. Análise empírica

2.1 Indicadores de complexidade

A avaliação realizada neste artigo sustenta-se nos indicadores propostos por Hausmann *et al.* (2007) e Hidalgo e Hausmann (2009). A abordagem da complexidade econômica, inaugurada por tais autores, compreende que o crescimento da renda dos países é consequência da transformação estrutural em direção a setores que requerem maior quantidade de conhecimento para sua produção. Para isso, os trabalhos propõem diversos indicadores para a mensuração do nível de conhecimento requerido para a produção competitiva de cada produto e também do nível de conhecimento produtivo existente em cada país a partir dos conceitos de diversificação e ubiquidade. Os autores constroem tais medidas embasando-se no indicador de vantagens comparativas reveladas (VCR), estipulado por Balassa (1965). Tal indicador expressa o comparativo entre a participação de um setor na economia local e a participação desse mercado local na economia como um todo. Com isso, verifica-se a capacidade das regiões em atuar em um determinado setor de maneira competitiva. O índice assume a seguinte forma:

$$VCR_{pc} = \frac{X_{pc}/\sum_p X_{pc}}{\sum_p X_{pc}/\sum_c \sum_p X_{pc}} \quad (1)$$

onde o X_{pc} representa a quantidade exportada do produto p e do país c . Compreende-se, portanto, que, se o indicador resultar em um número maior ou igual a um, a produção local do setor é competitiva em relação às demais localidades. Por outro lado, para um valor inferior a 1, a atividade produtiva de referência não exerce importância relevante no mercado analisado.

De forma complementar, Hidalgo e Hausmann (2009) enfatizam o papel do conhecimento e das capacidades produtivas sobre a determinação dos setores em que as economias se especializam. Conforme argumentam os autores, o nível de produtividade é resultado do conhecimento, sobretudo o tácito, e das capacidades internas que as economias concentram. Logo, as diferenças na renda entre os países são explicadas por diferenças de conhecimento e de capacidades. Embora não seja possível medir o nível de conhecimento produtivo existente em cada economia, é possível inferir esse nível a partir das informações contidas na estrutura produtiva das economias. Para isso, os autores partem do cálculo de dois indicadores que fundamentam a análise da complexidade dos produtos e dos países:

$$D_c = k_{c,0} = \sum_p M_{pc} \quad (2)$$

$$U_p = k_{p,0} = \sum_c M_{pc} \quad (3)$$

onde, conforme definem Hidalgo e Hausmann (2009), a quantidade de bens exportados com VCR é definida como o grau de diversificação dos países (D_c) e a quantidade de países que exportam uma determinada mercadoria com VCR reflete o grau de ubiquidade (U_p) desse setor. M_{pc} refere-se a uma matriz binária para representar em quais setores, p , os países, c , possuem VCR. Logo, quando há VCR, a entrada assume o valor 1 e, caso contrário, 0. Nesses indicadores primários, os países com maior quantidade de conhecimento produtivo são aqueles caracterizados por elevados patamares de diversificação, uma vez que a produção competitiva diversificada requer maior nível de conhecimento. Analogamente, quanto mais baixo o grau de ubiquidade dos produtos, maior deve ser o requerimento de conhecimento produtivo para sua produção competitiva.

A partir desses dois indicadores, Hidalgo e Hausmann (2009) calculam, então, os índices que definem o nível de complexidade dos países – Índice de Complexidade dos Produtos (ECI) - e dos produtos – Índice de Complexidade dos Produtos (ICP). Estes índices são construídos pela combinação iterada entre os indicadores de diversificação e ubiquidade. Tal combinação é utilizada para que as localidades com alta (baixa) diversificação, mas concentradas na produção de bens muito (pouco) ubíquos, sejam consideradas menos (mais) complexas e para que os produtos pouco (muito) ubíquos, mas produzidos competitivamente por países menos (mais) diversificados, sejam também considerados menos (mais) complexos. Essa iteração é feita continuamente para que a mensuração da complexidade dos produtos e dos países, ao final, constitua uma informação robusta. Formalmente, essa combinação é representada por:

$$k_{c,N} = \left(\frac{1}{k_{c,0}}\right) \sum_p M_{cp} k_{p,N-1} \quad (4)$$

$$k_{p,N} = \left(\frac{1}{k_{p,0}}\right) \sum_c M_{cp} k_{c,N-1} \quad (5)$$

onde o N se refere ao número das iterações. Substituindo (4) em (5), formalmente, tem-se que:

$$k_{c,N} = \sum_{c'} \varpi_{cc'} k_{c',N-2} \quad (6)$$

Em que:

$$\varpi_{cc'} = \sum_p (M_{cp} M_{c'p}) / (k_{c,0} k_{p,0}) \quad (7)$$

A equação (6) é satisfeita no momento em que $k_{c,N} = k_{c,N-2} = 1$ - o que representa o autovetor associado ao maior autovalor de ϖ_{cc} . Entretanto, o autovetor de referência deve ser aquele associado ao segundo maior autovalor de ϖ_{cc} , uma vez que o primeiro é composto apenas por valores unitários, não sendo, portanto, informativo. A necessidade de destacar os autovetores associados aos maiores autovalores objetiva encontrar os indicadores que incorporam a maior parte da variância dos dados iniciais. Dessa maneira, o ICE é, formalmente, definido como:

$$ICE = \frac{(K^* - \langle K^* \rangle)}{stdev(K^*)} \quad (8)$$

onde K é o autovetor associado ao segundo maior autovalor de ϖ_{cc} , o operador $\langle \rangle$ denota a média e, $stdev$, o desvio padrão. Seguindo o mesmo raciocínio, mas substituindo (5) em (4), pode-se encontrar o autovetor (Q) associado ao segundo maior autovalor de ϖ_{pp} . Com isso, estipula-se, formalmente, o ICP como:

$$ICP = \frac{(Q^* - \langle Q^* \rangle)}{stdev(Q^*)} \quad (9)$$

2.2 Base de dados

A principal fonte de dados utilizada na investigação apresentada neste artigo foi a Relação Anual de Informações Sociais (RAIS) da Secretaria de Trabalho do Ministério da Economia. A RAIS é uma base de registros administrativos que traz informações sobre todos os estabelecimentos formais e vínculos celetistas e estatutários do Brasil. A declaração das informações é anual, de responsabilidade dos estabelecimentos e de caráter obrigatório. A partir da RAIS é possível obter informações sobre os estabelecimentos formais, empregados e sobre o tipo de relação entre estabelecimentos e empregados, os chamados vínculos empregatícios. Os dados permitem a identificação do município em que está localizada cada empresa, sendo possível, a partir disso, obter o número de vínculos empregatícios em cada município, em cada ano, de forma discriminada para cada setor econômico.

Os indicadores de complexidade econômica para municípios e setores, utilizados nos exercícios apresentados neste trabalho, foram calculados utilizando dados de emprego setoriais por município em nível de divisão CNAE (2 dígitos). Para a mensuração, foi utilizado o método dos reflexos, de acordo com a metodologia proposta por Hidalgo *et al.* (2009).

Tabela 1 – Variáveis utilizadas no PSM e nas regressões de DD

Variável	Definição	Fonte	PSM	DD
PIB <i>per capita</i>	Produto Interno Bruto <i>per capita</i> .	IBGE	Sim	Sim
ICE	Índice de Complexidade Econômica - calculado a partir da equação (8).	RAIS	Sim	Sim
População	População do município.	Datasus	Sim	Sim
Emprego Superior	Porcentagem do emprego formal com nível superior em relação ao total empregado.	RAIS	Sim	Sim
Densidade Demográfica	População dividida pela área do município.	Datasus e IBGE	Sim	Sim
E.F. de Unidades Federativas	Dummies dos 27 Estados brasileiros	IBGE	Sim	Não
E.F. de municípios	Dummies dos municípios brasileiros	IBGE	Não	Sim

Fonte: elaboração própria.

Ademais, outras quatro variáveis foram utilizadas para a configuração dos modelos estimados. Para o pareamento – via PSM – e para a estimação do DID, utilizamos a população, a densidade demográfica e os vínculos empregatícios com a escolaridade definida por ensino superior completo. Enquanto a última variável é também disponibilizada pela RAIS, as duas primeiras são acessadas por meio da plataforma do departamento de informática do Sistema Único de Saúde do Brasil (DATASUS) – e do Instituto Brasileiro de Geografia e Estatística (IBGE). Por fim, a variável dependente da pergunta de pesquisa, o PIB *per capita*, é também disponibilizada pelo IBGE.

As variáveis utilizadas nas estimações apresentadas no restante do artigo e suas respectivas fontes são apresentadas de forma resumida na Tabela 1.

2.3 Estratégia empírica

A maneira ideal de se avaliar o impacto de uma melhoria na estrutura produtiva sobre o nível de renda local seria realizar um experimento em que o aumento da complexidade fosse um processo exógeno. Isso implica que as decisões dos municípios de avançar para a produção e comercialização de produtos mais complexos fosse independente das características e condições da localidade. No entanto, as decisões e/ou escolhas dos municípios (isto é, de suas empresas e agentes) para estruturas produtivas mais complexas possuem natureza multidimensional e, por isso, sofrem com o problema de endogeneidade. Nesse sentido, uma metodologia quase experimental parece adequada quando não é possível controlar a atribuição dos grupos tratados.

O objetivo deste artigo, portanto, consiste em analisar o que teria acontecido com o PIB *per capita* dos municípios onde se verificou aumento da complexidade caso não tivesse ocorrido esse aumento. Para avaliar a evolução do PIB nesses municípios, é preciso considerar também as características dos municípios pré-tratamento para identificar a diferença entre os níveis de PIB *per capita* dos municípios com aumento ou não nos níveis de complexidade econômica.

A dificuldade das avaliações de impacto está relacionada à impossibilidade de observar as mesmas cidades em duas situações no mesmo período de tempo: na situação de tratado – quando aumentam seus níveis de complexidade – e de não tratamento – quando não aumentam seus níveis de complexidade. Na tentativa de solucionar este problema, é possível utilizar contrafactuais do grupo tratado, considerando aqueles municípios com características semelhantes que não melhoraram no ranking de complexidade durante o período de tempo analisado. No entanto, a simples comparação entre esses dois grupos de municípios geraria resultados viesados, pois existem outras variáveis que diferenciam ambos e influenciam tanto na condição de tratado, quanto na variável de resultado (PIB *per capita*). Para definir o grupo de tratamento relevante, portanto, torna-se necessário um pareamento com municípios que carregam algumas características semelhantes. A técnica de *Propensity Score Matching* (PSM), de Rosebaum e Robin (1983), é amplamente utilizada para definir um grupo estatístico de comparação. No presente caso, busca-se um grupo de comparação para os municípios que aumentaram seus níveis de complexidade. Esta técnica utiliza variáveis observáveis para fazer o pareamento entre os municípios, encontrando, dentro do grupo de municípios de controle, aqueles mais parecidos em termos de características observáveis. O grupo controle identificado foi então utilizado para estimação de um modelo de diferenças-em-diferenças (DID), possibilitando a comparação do crescimento do PIB *per capita* dos municípios tratados e controle.

2.3.1 Grupos tratado e controle

A escolha dos grupos tratado e controle é crucial para o delineamento do trabalho. Uma vez que o índice de complexidade econômica é uma variável sempre relativa e somente comparável dentro de um mesmo ano, optou-se por não analisar o valor da variável e sim as variações que os municípios sofreram no ranking de complexidade. Além disso, foram realizados testes de robustez utilizando diferentes níveis de aumento e queda no ranking, de maneira a validar os resultados.

Para definir um número de posições de aumento no ranking e obter um ponto de corte que defina quais municípios serão considerados como tendo apresentado um aumento de complexidade ao longo do tempo, optou-se por realizar as avaliações a partir dos desvios padrões (D.P.) em relação à média das variações no ranking. A primeira estimação foi realizada considerando os municípios que subiram 0,3 D.P.

no ranking de complexidade, em comparação com os que não tiveram esse aumento. Ou seja, se um município subiu um número de posições no ranking de complexidade equivalente à, no mínimo, 0,3 desvios padrão acima da média, ele será considerado como um município em que houve aumento de complexidade. As estimações de robustez compreendem os municípios que tiveram variações em 1,1, 1,5 e 2 D.P. Esses valores de desvios-padrões foram escolhidos a partir dos resultados do PSM. Primeiro foi realizado o pareamento para todos os valores de desvios-padrão entre 0 e 2. Em seguida, foram então escolhidos os modelos nos quais foram encontrados grupos tratado e controle semelhantes estatisticamente nas características observáveis, pressuposto para que os grupos sejam comparáveis.

2.3.2 Modelos PSM e DID

Para identificar as verdadeiras diferenças de médias entre tratado e controle é necessário encontrar a estratégia que corrige o viés de seleção. No presente artigo utilizou-se o método *Difference-in-Differences* (DID). Denotando $Y(D)$ como o resultado potencial para o município, sendo $Y(1)$ a situação de tratamento no resultado e $Y(0)$ a situação para esse mesmo município na condição contrafactual de não ter passado pela intervenção, a equação $E[Y(1) - Y(0)]$ representaria o Efeito Médio do Tratamento (ATE). Entretanto, como já mencionado, não é possível observar a mesma cidade simultaneamente nas situações de tratado e controle. Portanto, a diferença de médias $E[Y(1) | D = 1] - E[Y(0) | D = 0]$ não é um estimador adequado do ATE.

Ao realizar a soma e a subtração do contrafactual na equação de diferenças de médias observadas é possível determinar o viés⁵. Após realizar as operações e rearranjar os termos, temos o efeito do tratamento sobre os tratados mais um viés de seleção, que se refere ao impacto médio do tratamento, como observado na equação abaixo:

$$E[Y(1) - Y(0) | D=1] + E[Y(0) | D=1] - E[Y(0) | D(0)] \quad (10)$$

A técnica de emparelhamento resolve este problema de viés de seleção substituindo a randomização pelo condicionamento dos regressores. O viés de seleção é removido apenas se o tratamento for puramente aleatório entre municípios com o mesmo escore de propensão. A correspondência com o escore de propensão pode reduzir o viés de seleção e, portanto, melhorar a estimativa do efeito do tratamento com dados observáveis. Assim, condicionando em X , para dar conta dos efeitos observáveis que determinam a auto-seleção, chega-se ao Efeito Médio sobre os Tratados (ATT), que é o primeiro termo da equação (10). A alternativa proposta por Rosenbaum e Rubin (1983) contorna o problema da dimensionalidade por meio do uso do propensity score $p(x_i)$, que dá ao município i a probabilidade de receber o tratamento baseado nas suas características X_i , o que permite os tratados e controle com semelhantes *propensity score* serem comparados.

O PSM é um método que utiliza um modelo Probit, com a variável dependente referente à complexidade econômica - em que o valor "1" indica que o município subiu no ranking de complexidade econômica, e "0" caso contrário. A seleção das variáveis explicativas consiste no entendimento sobre as possíveis características que influenciam na probabilidade do município aumentar ou diminuir seu nível de complexidade. De outra forma, objetiva-se com a inclusão destas informações controlar as possíveis dissonâncias existentes entre tratado e controle, identificando especificidades de cada município e corrigindo-as de maneira a isolar o possível impacto de um aumento no nível de complexidade do tratado.

A constituição do grupo de controle se pauta pela técnica de Nearest Neighbor Matching (NNM), que combina os resultados de todos os componentes do grupo de controle que possuem escores de propensão estatisticamente mais similares ao escore do componente do grupo de tratamento (CAMERON; TRIVEDI, 2005). Além disso, utilizou-se um vizinho próximo e com reposição.

A hipótese do estudo é a de que um aumento da complexidade econômica nos municípios gera maior crescimento do PIB *per capita* nestes lugares. O modelo de DID pode ser especificado de forma genérica como:

⁵ $E[Y(0) | D=1]$ é a média dos tratados caso não tivessem sido tratados e $E[Y(1) | D=0]$ é a média para os não tratados caso eles tivessem sido tratados.

$$Y_{it} = \alpha + \delta D_i + \rho T_t + \gamma D_i * T_t + \beta_1 X_{it} + \beta_2 \omega + \mu_t + \tau_i + \epsilon_{it} \quad (11)$$

em que Y representa o PIB *per capita* dos municípios. O vetor X representa as características observadas (covariadas ou variáveis de controle) do município, denotado pelo subscrito *i*. O subscrito *t* é o período de tempo. A variável T assume valor 1 no período após o crescimento ou decréscimo da complexidade econômica, e 0 caso contrário. D é a variável binária que indica se o município *i* foi, de fato, tratado, assumindo o valor 1, e valor 0 se não foi tratado (não teve aumento no ranking de complexidade econômica). O termo D*T representa a interação entre D e T e γ é o efeito do tratamento - melhoria nos níveis de complexidade econômica - obtido por DID. O termo ω representa os pesos do PSM. Este é um método alternativo para obter o ATT ao utilizar o score de propensão como uma variável explicativa na regressão de DID, de modo a reduzir ainda mais a influência de variáveis que possam estar influenciando na variável de resultado. Por fim, μ_t , τ_i e ϵ_{it} representam os efeitos fixos de tempo, município e os resíduos da regressão, respectivamente. Ressalta-se que as variáveis utilizadas no PSM quando se considera variações no ranking de complexidade de 0,3, 1,1 D.P. e 2 D.P. foram iguais, de modo manter a comparabilidade entre os modelos.

Para a utilização do método, existem algumas hipóteses que precisam ser verificadas. A primeira hipótese, de suporte comum, assegura que todo grupo tratado possua pelo menos um grupo de controle para servir de contrafactual (SCORZAFAVE et al., 2015). Já a segunda hipótese requer que a tendência temporal da variável de resultado dos dois grupos seja a mesma antes do tratamento – que seria o ponto de início da análise da evolução da complexidade econômica nos municípios. A ideia é que uma trajetória temporal semelhante indica que ambos os grupos vinham reagindo de forma similar a todo e qualquer fator que afetasse a variável de resultado antes do período de início da análise. Realizamos, portanto, o teste para verificar se havia a mesma trajetória temporal para os dois grupos (FILHO; PINTO, 2017).

Outra hipótese é a de que a composição dos grupos de tratamento e controle não se altere de forma significativa entre os períodos anterior e posterior ao momento de análise. No caso do presente artigo isso não ocorre, pois as unidades de análise são os municípios e os dados foram obtidos para todo o período de análise (FILHO; PINTO, 2017).

Por fim, a última condição requerida pelo método DID é de que os grupos de tratamento e controle não sejam afetados de forma específica por mudanças de qualquer natureza que ocorram após o tratamento. De forma a diminuir as chances de que isso ocorra, utilizamos várias variáveis relevantes no PSM, de maneira a fazer com que, ao menos nas variáveis observáveis, o tratado e o controle sejam o mais semelhantes possível (FILHO; PINTO, 2017). Nesse sentido, dada a hipótese de seleção em observáveis, o *Propensity Score Matching* remove o viés de seleção. Além disso, o método de DID controla o viés proveniente de variáveis não observáveis fixas no tempo. Os dois métodos lidam com diferentes problemas, mas não simultaneamente. Portanto, a combinação do DID com o PSM trata o viés de seleção nas observáveis e não observáveis fixas no tempo.

Em resumo, as etapas realizadas neste trabalho foram: a) definição dos grupos de tratado e controle - a partir de aumento (em 0,3, 1,1, 1,5 e 2 D.P.) no ranking de complexidade econômica; b) estimação da decisão de tratamento - *propensity score* -, de maneira a delimitar o grupo de controle que seja o mais semelhante possível antes do ponto de análise da complexidade com o grupo de tratamento; c) estimação do DID ponderado pelo *propensity score*. O período de análise compreende a mudança dos municípios no ranking de complexidade do ano de 2009 para 2019.

3. Resultados

3.1 Análise descritiva

A Tabela 2 apresenta as médias das variáveis utilizadas para realizar o pareamento e encontrar os municípios que fazem parte do grupo controle, bem como os testes de médias entre os grupos antes e depois do pareamento. Assim, verifica-se que antes do pareamento os dois grupos possuíam características observáveis diferentes. Apenas em termos da população observa-se que o teste de médias para os dois grupos sugeriu que eles eram semelhantes com 95% de significância. No entanto, após o pareamento, o teste de médias indicou que não se pode rejeitar a hipótese nula de que a diferença de média entre os grupos

é igual a zero, confirmando que, considerando as variáveis utilizadas, o grupo tratado e controle são semelhantes com 95% de confiança.

Tabela 2 – Teste de médias entre os grupos tratado e controle sobre as variáveis do PSM antes e após o pareamento

	Grupo	Antes do pareamento			Após o pareamento (0,3 D.P.)		
		Média	Estat. t	P-valor	Média	Estat. t	P-valor
Log(População)	Tratado	9,04	0,60	0,54	9,13	0,55	0,58
	Controle	9,05			9,11		
Log(PIB <i>per capita</i>)	Tratado	2,12	5,27	0,00	1,814	-0,09	0,92
	Controle	2,20			1,816		
ECI	Tratado	-0,52	12,02	0,00	-0,34	-0,59	0,11
	Controle	-0,357			-0,29		
Emprego Superior	Tratado	0,183	-12,24	0,00	0,119	1,12	0,26
	Controle	0,154			0,116		
Densidade demográfica	Tratado	32,95	3,09	0,00	35,969	1,28	0,22
	Controle	35,71			32,524		

Fonte: elaboração própria.

A Tabela 3 apresenta as estatísticas descritivas dessas variações para os diferentes valores de desvios-padrão no ranking de complexidade considerados neste artigo. Observa-se que, em média, tem-se que uma variação no ranking para 0,3 D.P. representa um aumento de 690 posições para o grupo tratado e queda de 352 para o grupo controle. Para 1,1 D.P. verifica-se, em média, um ganho de 1198 posições para os municípios tratados e uma queda de 219 para o grupo controle. Para 1,5 D.P. observa-se um aumento médio de 1455 posições para o primeiro grupo e uma queda de 124 para o segundo grupo. E, por último, considerando 2 D.P., verifica-se uma média de aumento de 1796 posições para o grupo tratado e ganho de 80 posições para o grupo controle.

Tabela 3 – Estatísticas descritivas das variações no ranking por cada desvio padrão

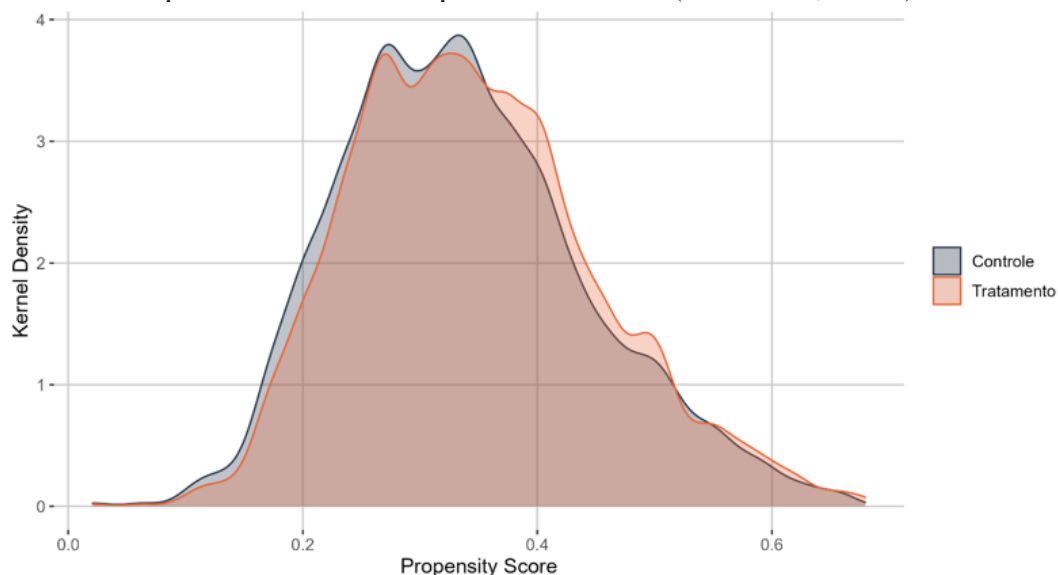
D.P.	Grupo	Mínima	Máxima	Média.	Desv. Pad.	N. de municípios
0,3	Controle	-2738	191	-351,7	469	1.668
0,3	Tratado	193	4916	690	496	1.161
1,1	Controle	-3265	710	-219,3	589	514
1,1	Tratado	717	4916	1197,6	499	597
1,5	Controle	-3265	955	-124	651	318
1,5	Tratado	981	4916	1454,6	508	352
2	Controle	-1283	1307	80,7	578	165
2	Tratado	1309	4916	1795,8	521	177

Fonte: elaboração própria.

É importante ressaltar que quanto maior o desvio-padrão considerado como ponto de corte, maior a média de ganho de posições no ranking de complexidade, dado que estão sendo estabelecidos limites mais à direita da distribuição das variações no ranking de complexidade. Em outras palavras, com maiores valores de D.P., estamos escolhendo como tratados apenas municípios que tiveram expressivo aumento no ranking de complexidade.

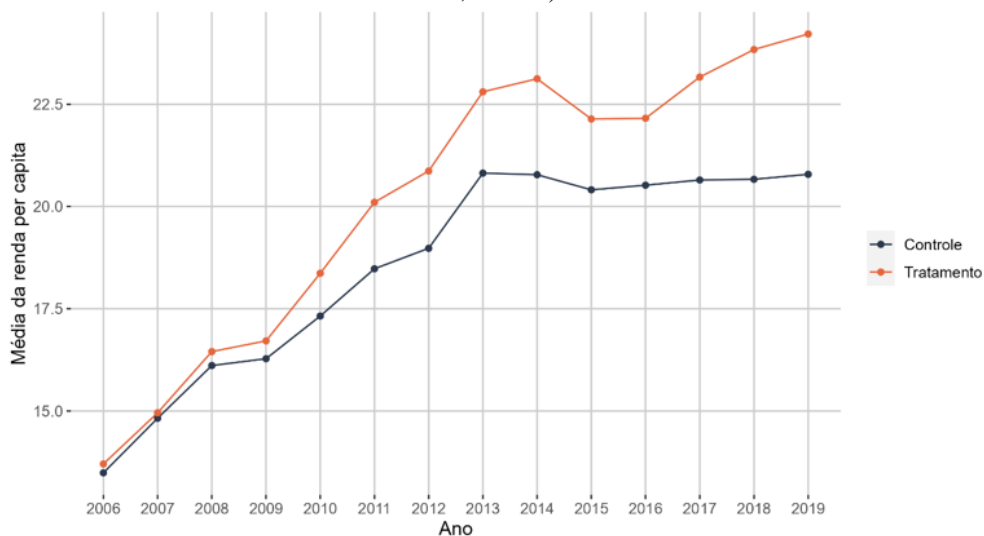
Para a efetividade do PSM, como destacado por Scorzafave et al. (2015), é preciso confirmar a hipótese de suporte comum, de modo a atestar que as observações de tratamento possuam unidades de comparação – grupo controle - nas proximidades da distribuição do escore de propensão. A Figura 2 apresenta a região de suporte comum para a amostra total (que independe da localização dos municípios) usada neste estudo assumindo variações no ranking de complexidade de 0,3 D.P. para os anos de 2006 a 2019. É possível identificar que existe uma sobreposição das duas curvas de distribuição (do grupo de controle e de tratamento), o que evidencia que os municípios que tiveram aumento de complexidade possuem compatibilidade com as unidades não tratadas em termos das características observáveis, facilitando, portanto, o processo de pareamento.

Figura 1 – Suporte comum do escore de propensão (Densidade Kernel) para os tratados e não tratados pelo aumento de complexidade – Total (Corte de 0,3 D.P.)



Fonte: elaboração própria.

Figura 2 – Média do PIB *per capita* para os grupos de tratamento e controle após o pareamento (Corte de 0,3 D.P.)



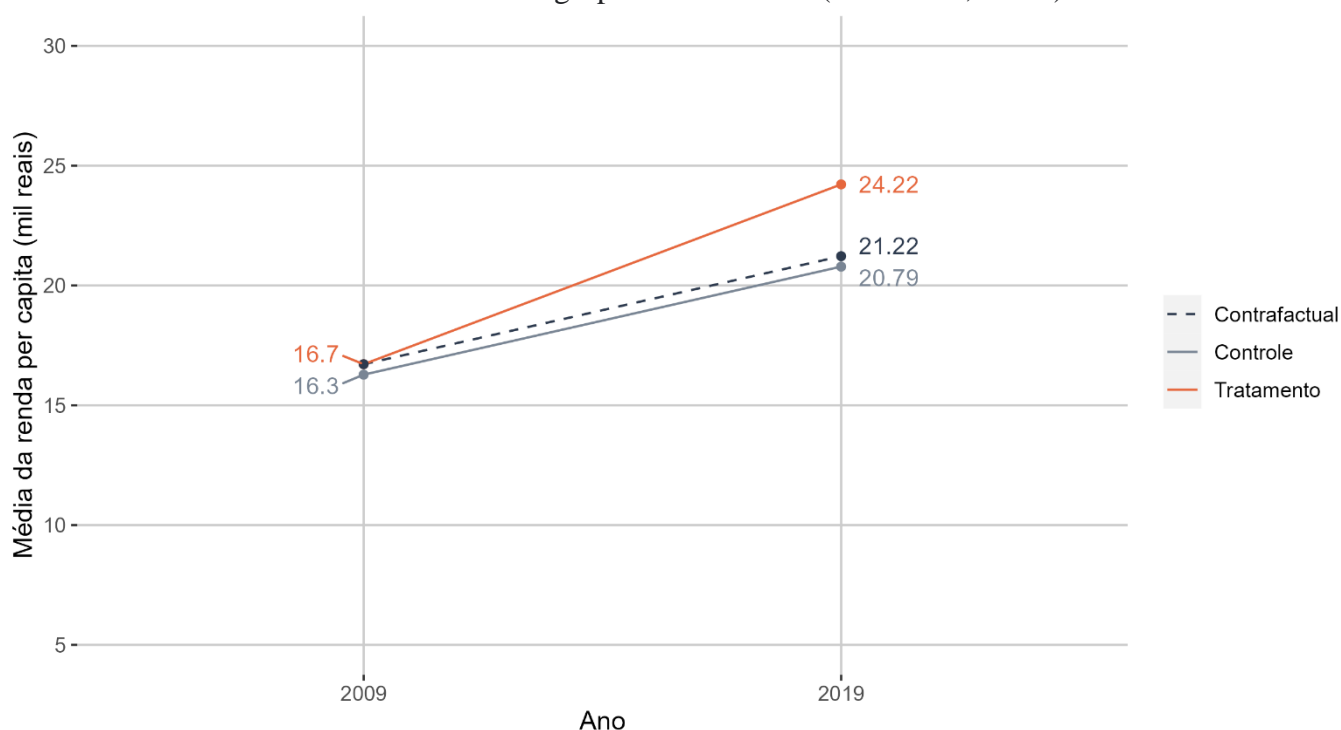
Fonte: elaboração própria.

A Figura 2 mostra a média da variável resultado (PIB *per capita*) para o grupo de tratamento e controle considerando 0,3 D.P. nas variações no ranking de complexidade para os anos de 2006 a 2019. É possível perceber que os dois grupos crescem da mesma forma entre 2006 e 2008. Apenas em 2008 e 2009 a taxa de crescimento do grupo de tratamento se mostra levemente inferior à do grupo controle, fazendo com que a média dos dois grupos se aproxime.

Esta análise é importante pois ajuda a confirmar a segunda hipótese, como destacado por Scorzafave et al. (2015), que postula que a tendência temporal da variável de resultado dos dois grupos seja a mesma antes do tratamento. A variação ocorrida entre 2008 e 2009 do grupo de tratados é bem menor do que as oscilações que ocorreram após o tratamento – após 2009. Dessa forma, acreditamos que, de maneira geral, a trajetória temporal parece semelhante para ambos os grupos antes do tratamento, confirmando a segunda hipótese.

A Figura 3 apresenta como ficaria o gráfico de crescimento do PIB *per capita* se os municípios do grupo tratado não tivessem tido um aumento no ranking de complexidade (contrafactual). Isso quer dizer que se os municípios que estão no grupo de tratamento não tivessem aumentado no ranking de complexidade eles seguiriam a mesma tendência de crescimento do PIB *per capita* dos municípios do grupo controle e teriam uma renda média menor no valor de 300 reais *per capita*. Ressalta-se que nessa figura ainda não são consideradas outras variáveis que podem influenciar nas diferenças de renda entre os dois grupos e que estão nas estimações.

Figura 3 – Média do PIB *per capita* para os grupos de tratamento e controle após o pareamento e projeção do contrafactual do grupo de tratamento (Corte de 0,3 D.P.)



Fonte: elaboração própria.

3.2 Resultados do modelo

A Tabela 4 mostra o resultado das estimações do modelo de *diferenças-em-diferenças* do tratamento aumento no ranking de complexidade (0,3 D.P.) nos municípios brasileiros para os anos 2006 e 2019. Este foi considerado o modelo principal por conter um maior número de municípios e também por ser o mais parcimonioso com relação à escolha do grupo controle, considerando também cidades que tiveram menores aumentos no ranking de complexidade.

Assim, mesmo para as duas estimações – a com menos parâmetros (I) e a com mais parâmetros (II) –, os resultados indicam que ter o tratamento de melhoria no ranking de complexidade gerou um aumento significativo no PIB *per capita* dos municípios entre 2009 e 2019. A diferença do valor do coeficiente dos dois modelos para a variável de interesse do trabalho – tratamento * tempo – pode ser em razão da influência de outras variáveis que afetam o PIB per capita e que foram adicionadas na estimação com mais parâmetros (II). Por isso, na estimação (II), o valor do coeficiente de interesse se reduziu e, portanto, também o efeito do tratamento.

A partir dos resultados, os municípios que tiveram aumento no ranking de complexidade tiveram um aumento médio anual de 3,25%⁶ no PIB *per capita* em comparação às cidades que caíram no ranking de complexidade no mesmo período. A utilização do pareamento faz com que, considerando as características observáveis do PSM, os grupos tratado e controle sejam semelhantes. Portanto, estamos comparando municípios que são pareados em termos das variáveis observáveis que consideramos e, assim, a variável resultado (PIB *per capita*) não seria influenciada por diferenças entre os grupos dessas características importantes. Além disso, ao recorrer à estimação de DID, conseguimos eliminar variáveis que não são observáveis, mas que são constantes no tempo, como alguns atributos institucionais dos municípios. Salienta-se que foram utilizados efeitos fixos de tempo e de município, no intuito de captar diferenças entre os anos, bem como características diferentes entre os municípios que são fixas ao longo do tempo.

Tabela 4 – Estimações do aumento no ranking de complexidade nos municípios brasileiros após o pareamento para 0,3 D.P. entre 2006 e 2019

	(I)	(II)
	0,3 D.P.	
Tratamento * Tempo	0,040*** (0,0054)	0,032*** (0,0038)
Tratamento	0,171** (0,0837)	0,637*** (0,0324)
Tempo	0,696*** (0,0039)	1,214*** (0,0060)
Pesos PSM	0,073 (0,6087)	-0,359*** (0,0283)
ICE		0,056*** (0,0043)
População		-0,482*** (0,0265)
Emprego superior		-0,041*** (0,0125)
Densidade demográfica		-0,000*** (0,0001)
Constante	1,897*** (0,0536)	6,450*** (0,2705)
E.F. Tempo	Sim	Sim
E.F. Município	Sim	Sim
Observações	39.606	39.606
R ² ajustado	0,87	0,95

Fonte: elaboração própria. Nota: Variável dependente: log do PIB *per capita*. Erro padrão entre parênteses. Significância dos coeficientes: *** (p < 0,01), ** (p < 0,05), * (p < 0,1).

O coeficiente da variável apenas de tratamento – considerando o período todo – sugere que os dois grupos possuem diferenças, em média, no PIB *per capita*. O coeficiente ter apresentado significância estatística pode ser explicada pelo período analisado, dado que se leva em conta os anos antes do tratamento e depois do tratamento. Como observado na Tabela 2, as médias do PIB *per capita* são estatisticamente iguais antes do tratamento. Contudo, após o tratamento a média entre os dois grupos se altera, como evidenciado no coeficiente da variável de interesse.

⁶ Para a interpretação de variáveis *dummies* em regressões semilogarítmicas precisamos tomar o antilogaritmo (para base e) do coeficiente binário estimado, subtrair 1 dele e multiplicar a diferença por 100 (GUJARATI; PORTER, 2011).

Além disso, o coeficiente da variável de tempo também foi significativo e positivo para os dois modelos, indicando que há uma tendência de aumento do PIB *per capita* de 2009 a 2019, comparativamente ao período todo (2006 a 2019) que independe do efeito do tratamento. Por isso, ao adicionar esta variável, retira-se a influência dessa tendência.

Observa-se também que a variável ICE é significativa e positiva, indicando que maiores níveis de complexidade nos municípios estão associados com o maior crescimento do PIB *per capita*, como já observado na literatura para diferentes anos e regiões (HIDALGO; HAUSMANN, 2009; FELIPE *et al.*, 2012; ROMERO; SILVEIRA, 2018).

3.3 Testes de robustez

As hipóteses de suporte comum e de mesma tendência temporal são validadas quando se assume variações no ranking de complexidade de 1,1 D.P., 1,5 D.P. e 2 D.P. Os gráficos verificando as hipóteses para a variável de resultado antes do tratamento se encontram nas Tabelas A1, A2 e A3 do Apêndice.

A Tabela 5 mostra as estimações de aumento no ranking de complexidade de 1,1, 1,5 e 2 D. P. para os municípios brasileiros após o pareamento, entre 2006 e 2019. Os resultados indicam que a escolha dos municípios que fazem parte do grupo de tratamento a partir dos desvios-padrão de aumento e redução no ranking de complexidade não influenciaram de forma relevante nos resultados encontrados. Em todas as regressões os municípios que tiveram aumento no ranking de complexidade desfrutaram de um aumento no PIB *per capita* comparativamente àqueles municípios semelhantes em termos de características observáveis, mas que não tiveram aumento no ranking.

Tabela 5 – Estimações do aumento no ranking de complexidade nos municípios brasileiros após o pareamento para 1 D.P. entre 2006 e 2019

	(I)	(II)	(III)	(IV)	(V)	(VI)
	1,1 D.P.		1,5 D.P.		2 D.P.	
Tratamento * Tempo	0,068*** (0,0090)	0,052*** (0,0062)	0,065** (0,0120)	0,047*** (0,0087)	0,084*** (0,0176)	0,072*** (0,0125)
Tratamento	0,350*** (0,0867)	0,643*** (0,1103)	-0,109 (0,0816)	0,582*** (0,0879)	0,241*** (0,0934)	-0,491*** (0,1633)
Tempo	0,695*** (0,0059)	1,238*** (0,0100)	0,707*** (0,0078)	1,245*** (0,0134)	0,716*** (0,0102)	1,275*** (0,0200)
Pesos PSM	-1,124*** (0,1126)	-1,510*** (0,0694)	0,094 (0,1599)	-0,138 (0,0900)	0,804*** (0,0994)	0,371*** (0,1343)
ICE		0,064*** (0,0064)		0,069*** (0,0081)		0,066*** (0,0121)
População		-0,522*** (0,0507)		-0,529*** (0,0637)		-0,580*** (0,1079)
Emprego superior		-0,056*** (0,0176)		-0,037* (0,0209)		-0,0046 (0,0369)
Densidade demográfica		-0,000*** (0,0005)		-0,0017 (0,0012)		-0,006*** (0,00194)
Constante	3,860*** (0,1657)	8,694*** (0,5004)	2,086*** (0,1688)	6,150*** (0,5226)	0,885*** (0,1584)	6,402*** (1,0928)
E.F. Tempo	Sim	Sim	Sim	Sim	Sim	Sim
E.F. Município	Sim	Sim	Sim	Sim	Sim	Sim
Observações	15.554	15.554	9.380	9.380	4.788	4.788
R ² ajustado	0,84	0,94	0,83	0,93	0,82	0,92

Fonte: elaboração própria. Nota: Variável dependente: log do PIB *per capita*. Erro padrão entre parênteses. Significância dos coeficientes: *** (p < 0,01), ** (p < 0,05), * (p < 0,1).

Para as estimações assumindo variações no ranking de complexidade de 1,1 D.P. verifica-se uma diferença no PIB *per capita* anual de 5,33%. Considerando 1,5 D.P., os municípios do grupo tratado, comparativamente aos do grupo controle, aumentaram seu PIB *per capita*, em média, em 4,81%. Tendo em

vista variações no ranking de 2 D.P. verifica-se que municípios que aumentaram a complexidade tiveram uma elevação média anual de 7,46% no PIB *per capita*.

Por fim, é importante ressaltar que quanto maior a restrição para se considerar os municípios que tiveram oscilações no ranking de complexidade, maior é o resultado no aumento do PIB *per capita*. Isto é, quanto mais se limita a escolha dos municípios que serão tratados e controle (antes do pareamento) - para que eles do grupo tratamento são considerados apenas aqueles municípios que aumentaram de maneira mais significativa no ranking - maiores são as diferenças no PIB *per capita* dos dois grupos. Esse resultado já era esperado, porque com a limitação dessa escolha, passamos a considerar apenas municípios que tiveram uma profunda mudança estrutural em termos da complexidade de suas estruturas produtivas. Assim, a consequência é que a diferença de renda entre os dois grupos seja intensificada.

4. Considerações finais

A abordagem da complexidade econômica tem crescido e se desenvolvido com vários trabalhos demonstrando que regiões ou países com maior nível de complexidade em suas estruturas produtivas tendem a obter maior crescimento econômico. Essas pesquisas utilizam vários modelos e dados para realização das estimações, bem como a inclusão de diferentes tipos de controles, efeitos fixos e análises de robustez. Entretanto, os resultados exigem cautela para afirmações de causalidade entre complexidade e crescimento econômico, devido a problemas de endogeneidade nesses modelos.

O presente artigo utiliza a metodologia do *Propensity Score Matching* (PSM) em conjunto com o modelo de diferenças-em-diferenças (DID) para avaliar o aumento da complexidade econômica em municípios brasileiros sobre o crescimento do PIB *per capita* nos anos de 2009 a 2019.

Os resultados indicam que ter o tratamento de melhoria no ranking de complexidade gerou um aumento no PIB *per capita* dos municípios entre 2009 e 2019. Para 0,3, 1,1, 1,5 e 2 D.P. foram encontradas variações em média de 3,25%, 5,33%, 4,81% e 7,46% no PIB *per capita* dos municípios, respectivamente.

Nossos resultados, portanto, corroboram a literatura que sugere uma relação positiva entre complexidade e renda *per capita* das regiões. No entanto, como existem variáveis que podem influenciar em conjunto tanto o aumento da complexidade, quanto no aumento da renda, as análises por meio de metodologias de avaliação de impacto trazem indícios mais sólidos sobre a conexão entre essas variáveis.

Por fim, é importante destacar que não se pode afirmar que foi eliminada completamente a endogeneidade. Contudo, metodologia aplicada neste artigo certamente contribui para a diminuição desse viés, já que compara grupos semelhantes em termos da média das variáveis consideradas, além de, por meio do DID, eliminar variáveis não observáveis e fixas ao longo do tempo que podem influenciar as estimativas finais.

Em trabalhos futuros, mais análises de robustez para validar os resultados podem ser feitas, como a utilização de diferentes anos, como, por exemplo, nos anos 70 e 80, em que o Brasil teve consideráveis melhorias em termos de estrutura produtiva industrial. Outra alternativa é considerar formas diferentes de classificar o grupo controle por meio do PSM. Além disso, utilizar alguma variável instrumental em conjunto com as estimações de DID possibilitaria ainda mais confiança na robustez dos resultados encontrados.

REFERÊNCIAS

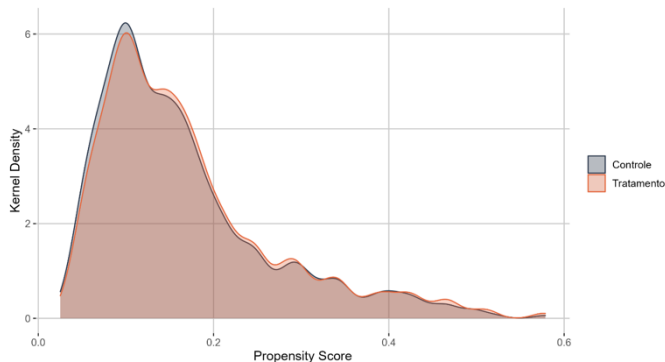
- ALMEIDA, A. N. Os impactos dos Royalties do Petróleo em gastos sociais no Brasil: Uma análise usando Propensity Score Matching. *Revista Economia & Tecnologia*, v. 10, n. 2, 2014.
- BALASSA, B. *Trade liberalization and revealed comparative advantage*, Manchester School of Economics and Social Studies, 1965.
- BRITTO, G.; ROMERO, J. P.; FREITAS, E.; COELHO, C. *The Great Divide: Economic Complexity and Development Paths in Brazil and South Korea*. *Cepal Review*, 2019.
- CAMERON, A. C.; TRIVEDI, P. K. *Microeconometrics: methods and applications*. Cambridge university press, 2005.

- FELIPE, J.; KUMAR, U.; ABDON, A.; BACATE, M. Product complexity and economic development. *Structural Change and Economic Dynamics*, v. 23, n. 1, p. 36-68, 2012.
- FILHO, N. A. M. F.; PINTO, C. C. X. *Avaliação econômica de projetos sociais*. 3. ed. São Paulo: Fundação Itaú Social, 2017.
- FILHO, O. A. P.; SOUSA, M. C. S. Avaliação De Impacto Das Guardas Municipais Com O Uso De Tratamentos Binários, Multivalorados E Contínuos. In: *Anais do XLIV Encontro Nacional de Economia*, 2018.
- GUJARATI, D. N.; PORTER, D. C. *Econometria básica*. Amgh Editora, 5. ed, 2011.
- HAUSMANN, R.; HIDALGO, C.A.; BUSTOS, S.; COSCIA, M.; CHUNG, S.; JIMENEZ, J.; SIMÕES, A.; YILDIRIM, M. A. *The Atlas of Economics Complexity – Mapping Paths to prosperity*. Puritan Press: New York, 2011.
- HAUSMANN, R.; HIDALGO, C.A.; BUSTOS, S.; COSCIA, M.; CHUNG, S.; JIMENEZ, J.; SIMÕES, A.; YILDIRIM, M. A. *The Atlas of Economics Complexity – Mapping Paths to prosperity*. Puritan Press: New York, 2011.
- HAUSMANN, R.; HWANG, J.; RODRIK, D. What you export matters. *Journal of economic growth*, v. 12, n. 1, p. 1-25, 2007.
- HARTMANN, D., GUEVARA, M. R., JARA-FIGUEROA, C., ARISTARÁN, M., & HIDALGO, C. A. Linking economic complexity, institutions, and income inequality. *World development*, v. 93, p. 75-93, 2017.
- HIDALGO, C.; KLINGER, B.; BARABASI, A. L.; HAUSMANN, R. The product space conditions the development of nations. *Science*, v. 317, n. 5837, p. 482-487, 2007.
- HIDALGO, C. A.; HAUSMANN, R. The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 2009 106(26), 2009. p. 10570-10575.
- JACQUES, E. R. Impacto das cooperativas de crédito sobre o PIB per capita dos municípios brasileiros. Dissertação - Programa de Pós-Graduação em Desenvolvimento Econômico, 2013.
- ROMERO, J. P.; GRAMKOW, C. Economic complexity and greenhouse gas emissions. *World Development*, v. 139, p. 105317, 2021.
- ROMERO, J. P.; SILVEIRA, F. Mudança estrutural e complexidade econômica: identificando setores promissores para o desenvolvimento dos estados brasileiros. In: Marcos Vinicius Chialliato Leite (Org.). *Alternativas para o desenvolvimento brasileiro: novos horizontes para a mudança estrutural com igualdade*. 201ed. Santiago: Nações Unidas, CEPAL, 2019, 252p., p. 137-160.
- ROSENBAUM, P. R.; RUBIN, D. B. The central role of the propensity score in observational studies for causal effects. *Biometrika*, v. 70, n. 1, p. 41-55, 1983.
- SCORZAFAVE, L. G.; SOARES, M. K.; DORIGAN, T. A. Vale a Pena Pagar para Desarmar? Uma Avaliação do Impacto da Campanha de Entrega Voluntária de Armas sobre as Mortes com Armas de Fogo. *Estudos Econômicos*, v. 45, n. 3, p. 475-497, 2015.

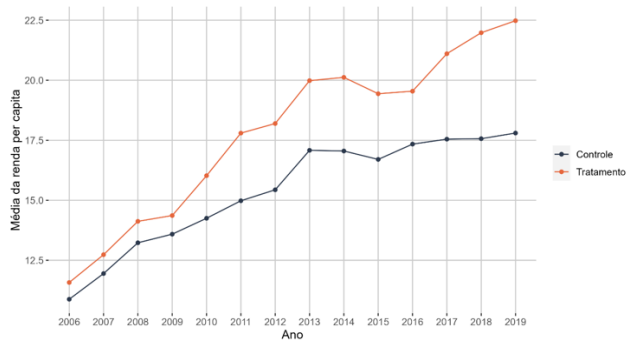
APÊNDICE

Tabela A1 – Corte de 1,1 D.P.

A. Suporte comum



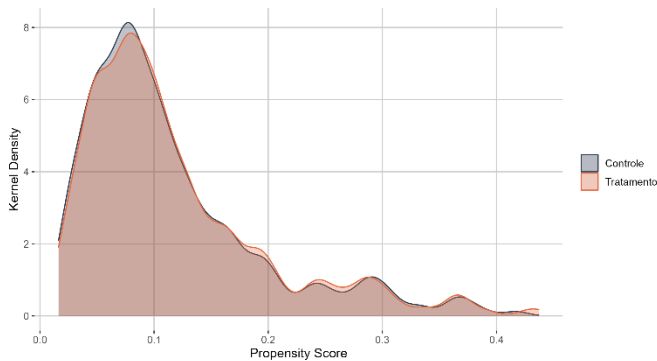
B. Diferenças de médias



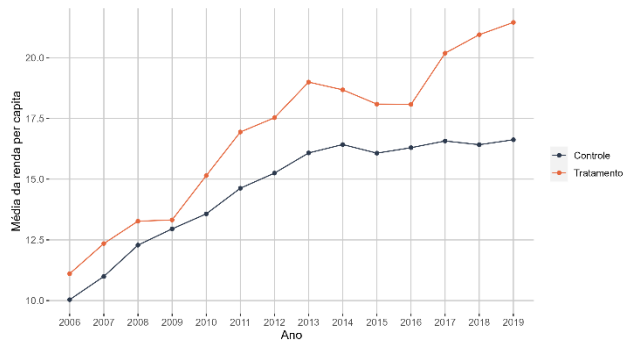
Fonte: elaboração própria.

Tabela A2 – Corte de 1,5 D.P.

A. Suporte comum



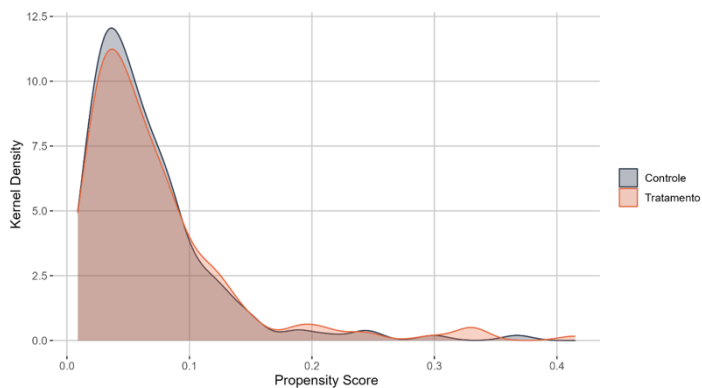
B. Diferenças de médias



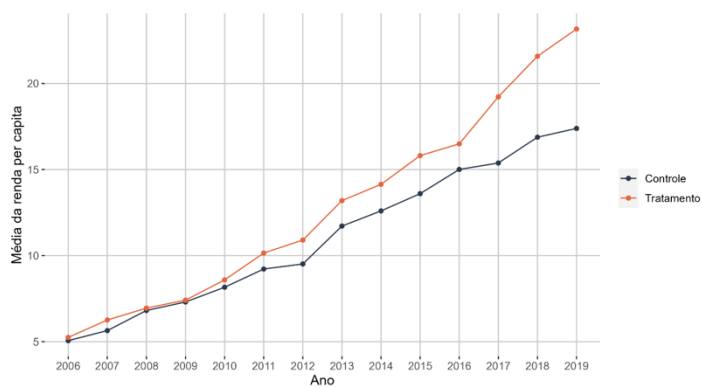
Fonte: elaboração própria.

Tabela A3 – Corte de 2 D.P.

A. Suporte comum



B. Diferenças de médias



Fonte: elaboração própria.