ISSN 0034-7140 00004 >
9 770034 714013

FGV

ISSN 0034-7140 (print)
ISSN 1806-9134 (online)
Rio de Janeiro • v. 72 • n. 4 • Out–Dez 2018

FGV

v. 72 • n. 4 • Out–Dez 2018

Revista Brasileira de Economia

# rbe

## Revista Brasileira de Economia

Regimes Inflacionários e Ciclos Econômicos: A Experiência Brasileira Pós–Plano Real
*Luckas Sabioni Lopes, Thiago Costa Soares*

Assessing Interdependence Among Countries' Fundamentals and its Implications for Exchange Rate Misalignment Estimates: An Empirical Exercise Based on GVAR
*Emerson Fernandes Marçal, Beatrice Zimmermann, Diogo de Prince, Giovanni Merlin*

Trajetória Longitudinal dos Alunos de Escolas Profissionalizantes para o Mercado de Trabalho
*Francisca Zilania Mariano, Ronaldo Albuquerque Arraes*

Grants and Marginal Cost of Public Funding: Empirical Evidence for Local Governments in Brazil
*Enlinson Mattos, Rafael Cardim, Ricardo Politi*

Impacto do Financiamento Habitacional Sobre o Mercado de Trabalho na Construção Civil
*Fábio Nishimura, Carlos Eduardo de Freitas, Roselaine Almeida*

Avaliação de Impacto das Guardas Municipais Sobre a Criminalidade com o uso de Tratamentos Binários, Multivalorados e Contínuos
*Oliveira Alves Pereira Filho, Maria da Conceição Sampaio de Sousa, Patrick Franco Alves*

Errata

**Patrocinadores (Financiadores)**

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)
Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes)
Ministério da Educação (MEC)
Programa de Apoio a Publicações Científicas do Ministério da Ciência e Tecnologia (MCT).

CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico
CAPES
Ministério da Educação
Ministério da Ciência e Tecnologia
BRASIL UM PAÍS DE TODOS GOVERNO FEDERAL

r b e

*Revista Brasileira de Economia*

**FUNDAÇÃO GETULIO VARGAS**

Instituição de caráter técnico-científico, educativo a filantrópico, criada em 20 de dezembro de 1944 como pessoa jurídica de direito privado, tem por finalidade atuar no âmbito das ciências sociais, particularmente econômica e administração, bem como contribuir para a proteção ambiental e o desenvolvimento sustentável.

**Presidente fundador:** Luiz Simões Lopes.
**Presidente:** Carlos Ivan Simonsen Leal.
**Vice-presidentes:** Francisco Oswaldo Neves Dornelles, Marcos Cintra Cavalcanti de Albuquerque e Sergio Franklin Quintella.

**CONSELHO DIRETOR**
**Presidente:** Carlos Ivan Simonsen Leal. **Vice-presidentes:** Francisco Oswaldo Neves Dornelles, Marcos Cintra Cavalcanti de Albuquerque e Sergio Franklin Quintella.
**Vogais:** Armando Klabin, Carlos Alberto Pires de Carvalho e Albuquerque, Cristiano Buarque Franco Neto, Ernane Galvêas, José Luiz Miranda, Lindolpho de Carvalho Dias, Marcílio Marques Moreira e Roberto Paulo Cezar de Andrade.
**Suplentes:** Aldo Floris, Antonio Monteiro de Castro Filho, Ary Oswaldo Mattos Filho, Eduardo Baptista Vianna, Gilberto Duarte Prado, Jacob Palis Júnior, José Ermírio de Moraes Neto, Marcelo José Basílio de Souza Marinho e Mauricio Matos Peixoto.

**CONSELHO CURADOR**
**Presidente:** Carlos Alberto Lenz César Protásio. **Vice-presidente:** João Alfredo Dias Lins (Klabin Irmãos & Cia).
**Vogais:** Alexandre Koch Torres de Assis, Andrea Martini (Souza Cruz S.A.), Antonio Alberto Gouveia Vieira, Eduardo M. Krieger, Estado da Bahia, Estado do Rio de Janeiro, Estado do Rio Grande do Sul, Luiz Chor, Luiz Ildefonso Simões Lopes, Marcelo Serfaty, Marcio João de Andrade Fortes, Miguel Pachá, Murilo Portugal Filho (Federação Brasileira de Bancos), Pedro Henrique Mariani Bittencourt, Ronaldo Vilela (Sindicato das Empresas de Seguros Privados, de Previdência Complementar e de Capitalização nos Estados do Rio de Janeiro e do Espírito Santo), Sandoval Carneiro Junior, Tarcísio Godoy (IRB-Brasil Resseguros S.A) e Willy Otto Jordan Neto.
**Suplentes:** Almirante Luiz Guilherme Sá de Gusmão, Clóvis Torres (Vale S.A.), General Joaquim Maia Brandão Júnior, José Carlos Schmidt Murta Ribeiro, Luiz Ildefonso Simões Lopes (Brookfield Brasil Ltda), Luiz Roberto Nascimento Silva, Manoel Fernando Thompson Motta Filho, Nilson Teixeira (Banco de Investimentos Crédit Suisse S.A.), Olavo Monteiro de Carvalho (Monteiro Aranha Participações S.A.), Patrick de Larragoiti Lucas (Sul América Companhia Nacional de Seguros), Rui Barreto, Sergio Lins Andrade e Victório Carlos De Marchi.

**Sede:** Praia de Botafogo, 190
Rio de Janeiro, RJ, Brasil · CEP 22250-900 · Tel. 21 3799-6000
Caixa Postal 62591 · CEP 22252-970

# *Normas para Publicação*

r b e

**Escopo / missão**

1) Artigos: trabalhos, de conteúdo analítico, que apresentem contribuições originais, tanto de cunho teórico como de avaliação empírica, no campo geral da economia;

2) Resenha bibliográfica: análise crítica de livros, editados no Brasil e no exterior, que digam respeito à economia.

**Processo de avaliação**

1) Todos os trabalhos submetidos à apreciação da RBE devem ser inéditos (nacional ou internacionalmente), não estando sob consideração para publicação em qualquer outro veículo de divulgação. Os trabalhos publicados em anais podem ser enviados depois de transformados em artigos;

2) Todos os trabalhos submetidos à publicação na RBE são examinados pelo editor e, ocasionalmente, por editores associados que auxiliam na escolha de alguns pareceristas e em decisões editoriais. O editor poderá rejeitar artigos que não estejam próximos da fronteira teórica e/ou empírica, sem a necessidade de apresentar pareceres, de modo a agilizar a submissão do artigo em outras revistas;

3) Em caso de aprovação, o trabalho é editorado e submetido para evisão do autor principal. No pedido de revisão será claramente informado o prazo para envio da revisão. Caso o autor não encaminhe a revisão no prazo estipulado, a publicação do trabalho poderá ser adiada.

**Forma de apresentação dos originais**

Os artigos submetidos à RBE devem obedecer às seguintes regras de apresentação dos originais:

1) O formato dos arquivos deverá ser, preferencialmente, LaTeX (classe *article*), mas artigos encaminhados em RTF ou Microsoft Word também serão aceitos, desde que observado o item 5;

2) Os artigos deverão ser acompanhados de resumos em português e em inglês, com no máximo 100 palavras com indicação de palavras-chave e códigos de classificação JEL;

3) Na primeira página devem constar as seguintes informações sobre cada autor: nome, instituições a que está vinculado e endereço para correspondência;

4) As referências bibliográficas dos artigos devem ser elaboradas de acordo com as normas da Associação Brasileira de Normas Técnicas (ABNT – NBR-6023) e apresentadas no final do texto. Quando na utilização do LaTeX, deverá ser utilizado o programa BibTeX;

5) As figuras e gráficos do artigo deverão constar em arquivos separados, nos formatos PDF, PS ou EPS. Os arquivos deverão estar claramente identificados e numerados, e sua localização no corpo do texto claramente identificada;

6) Os trabalhos deverão ser submetidos através do sistema de submissão *online* `http://bibliotecadigital.fgv.br/ojs/`.

FGV

ISSN 0034-7140
Rio de Janeiro • v. 72 • n. 4 • Out–Dez 2018

# r b e

## *Revista Brasileira de Economia*

**Sumário**


| | |
|---|---|
| 410 | Regimes Inflacionários e Ciclos Econômicos: A Experiência Brasileira Pós–Plano Real<br>*Luckas Sabioni Lopes, Thiago Costa Soares* |
| 429 | Assessing Interdependence Among Countries' Fundamentals and its Implications for Exchange Rate Misalignment Estimates: An Empirical Exercise Based on GVAR<br>*Emerson Fernandes Marçal, Beatrice Zimmermann, Diogo de Prince, Giovanni Merlin* |
| 451 | Trajetória Longitudinal dos Alunos de Escolas Profissionalizantes para o Mercado de Trabalho<br>*Francisca Zilania Mariano, Ronaldo Albuquerque Arraes* |
| 479 | Grants and Marginal Cost of Public Funding: Empirical Evidence for Local Governments in Brazil<br>*Enlinson Mattos, Rafael Cardim, Ricardo Politi* |
| 497 | Impacto do Financiamento Habitacional Sobre o Mercado de Trabalho na Construção Civil<br>*Fábio Nishimura, Carlos Eduardo de Freitas, Roselaine Almeida* |
| 515 | Avaliação de Impacto das Guardas Municipais Sobre a Criminalidade com o uso de Tratamentos Binários, Multivalorados e Contínuos<br>*Oliveira Alves Pereira Filho, Maria da Conceição Sampaio de Sousa, Patrick Franco Alves* |
| 545 | Errata |

A **Revista Brasileira de Economia** é indexada nas bases de dados JEL, EconLit e SciELO.

Os artigos assinados são de responsabilidade exclusiva dos autores. É permitida a reprodução total ou parcial dos artigos, desde que seja citada a fonte.

**Projeto gráfico:** ConsulLaTeX
**Editoração Eletrônica:** Sálvio Marcelo Soares

## CONSELHO EDITORIAL

**Fundadores:** Eugênio Gudin e Arizio de Viana
**Editor:** Ricardo de Oliveira Cavalcanti (FGV/EPGE)
**Editores associados:** Alexandre B. Cunha (UFRJ), Arilton Teixeira (FUCAPE), Bernardo Guimaraes (EESP/FGV), Bruno Sultanum (Federal Reserve Bank of Richmond), Carlos Eugênio da Costa (FGV/EPGE), Cezar Santos (University of Mannheim), Felipe Schwartzman (Federal Reserve Bank of Richmond), Gabriel de Abreu Madeira (USP), Jefferson Donizeti Pereira Bertolai (FEA-RP/USP), José Fajardo (EBAPE/FGV), Juliano J. Assunção (PUC-Rio), Marcelo Cunha Medeiros (PUC-Rio), Paulo Klinger Monteiro (FGV/EPGE) e Vladimir Ponczek (EESP/FGV).

## EDITORA FGV

**Editora Executiva:** Marieta de Moraes Ferreira
**Assinaturas:** Paulo César Peres Appolinário

## ESCOLA DE PÓS-GRADUAÇÃO EM ECONOMIA

**Diretor Geral:** Rubens Penha Cysne
**Vice-Diretor:** Aloisio Araujo
**Vice-Diretor de Ensino:** Carlos Eugênio da Costa
**Vice-Diretor de Pesquisa:** Humberto Moreira
**Vice-Diretores de Graduação e Regulação Institucional:** Luis Henrique Bertolino Braido & André Arruda Villela

## CORRESPONDÊNCIA REDATORIAL

**Revista Brasileira de Economia**
Praia de Botafogo, 190 sala 1100
Rio de Janeiro, RJ · CEP 22250-900
E-mail: rbe@fgv.br
Site: http://bibliotecadigital.fgv.br/ojs/index.php/rbe

## CORRESPONDÊNCIA COMERCIAL

**Fundação Getulio Vargas / Editora**
Caixa Postal 62591
Rio de Janeiro, RJ · CEP 22250-900
Tel.: 0800-21-7777
Fax: (21) 3799-5532
E-mail: assine@fgv.br

## LIVRARIAS

**Livraria J C Felipe**
Praia de Botafogo, 190
Rio de Janeiro, RJ
Tel.: (21) 3799-5535

**Livraria Faria Lima**
Av. Nove de Julho, 2029
São Paulo, SP
Tel.: (11) 3281-7875

Revista Brasileira de Economia, ano 1 n. 1 set. 1947.

Rio de Janeiro, Fundação Getulio Vargas

v. 23 cm trimestral

Revista da Escola de Pós-Graduação em Economia da Fundação Getulio Vargas

1. Economia Política - periódicos. I. Fundação Getulio Vargas, Rio de Janeiro. Escola de Pós-Graduação em Economia.

CDD 33(0.5)
CDU 330.5

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180020
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Regimes Inflacionários e Ciclos Econômicos: A Experiência Brasileira Pós–Plano Real

Luckas Sabioni Lopes*
Thiago Costa Soares*

**Sumário**

1. Introdução ..... 410
2. A formação de expectativas inflacionárias e sua relação com as mudanças nos regimes ..... 412
3. Resultados ..... 417
4. Conclusão ..... 425

**Palavras-chave**
Inflação, incerteza, ciclos econômicos, cadeias de Markov

**JEL Codes**
E3, E58

**Resumo • Abstract**

This paper analyzes the dynamics of Brazilian inflation between 1995 and 2016. By estimating Markovian models, we derive measures of uncertainties related to changes in monetary policy (regime uncertainty) and to random shocks (certainty equivalent). Our results show that the level and variability of inflation have reduced after the adoption of inflation targeting. In the same period, there were declines in regime uncertainties, with the exception of 2015 onwards, when inflation exceeded its upper limit and instabilities rose abruptly. It is also found that impulses in regime uncertainty impacts negatively production, consumption, and investment in the country.
e o investimento no país.

## 1. Introdução

O Brasil conviveu com elevadas taxas de inflação durante décadas. Por exemplo, em 1989 o índice de preços cresceu 1.973% e, em 1993, um máximo histórico de aproximadamente 2.500% foi alcançado.[1] Embora a reforma monetária do Plano Real em 1994 tenha conseguido reduzir a aceleração dos preços (média atual em torno de 7% a.a.), o país ainda exibe números para a inflação que são quatro, ou cinco vezes maiores que os observados em países mais desenvolvidos, como os Estados Unidos da América (EUA, 1,3% a.a.), Inglaterra (1,7% a.a.) e França (0,84% a.a.).[2]

Com o intuito de manter a inflação controlada, a política monetária no Brasil mudou de um sistema de câmbio administrado (CA), em operação entre 1995 e meados de 1999, para o de metas para a inflação (IT[3]), utilizado desde então. Os valores médios da inflação caíram no período, passando de 9% (CA) para 7% a.a. (IT), e o desvio padrão se reduziu de 11 (CA) para cinco pontos percentuais (IT). Assim, percebe-se que as variações nos preços se tornaram menores e menos voláteis desde a adoção do IT.

*Universidade Federal de Juiz de Fora, Departamento de Economia, Av. Dr. Raimundo Monteiro de Rezende, 330, Centro, Governador Valadares, MG, CEP 35010-177, Brasil.
luckas.lopes@ufjf.edu.br thiago.costa@ufjf.edu.br

[1] Índice de Preços ao Consumidor Amplo, IPCA.

[2] Médias do período 2012–2015. Dados obtidos no FRED® Economic Data (https://fred.stlouisfed.org).

[3] Do termo em inglês inflation targeting.



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Recebido em 15 de agosto de 2017
Aceitação final em 28 de abril de 2018

Dessa forma, pesquisas buscaram examinar a contribuição do novo sistema para a diminuição da inflação brasileira. Por exemplo, Barbosa-Filho (2008), ao estudar diversas séries de tempo entre 1994 e 2006, concluiu que o IT obteve êxito na redução do nível de preços após as crises de confiança de 1999 e 2002. Por outro lado, Mello e Moccero (2011), estimando um sistema de equações novo keynesiano em um período similar, argumentaram que o ambiente macroeconômico favorável foi mais importante do que a alteração da política monetária para a estabilização dos preços.

A divergência entre os resultados de Barbosa-Filho (2008) e Mello e Moccero (2011) pode ser atribuída, em outros fatores, ao curto período analisado desde a implementação do sistema IT no Brasil. Carvalho e Minella (2012) ponderaram que a credibilidade do sistema de metas para a inflação ainda está em processo de construção. Suas evidências econométricas mostraram que as metas perseguidas pelo Banco Central do Brasil (BCB) tenderam a guiar as expectativas inflacionárias mais significativamente após os choques cambiais de 2002–2003. Dessa forma, estudos anteriores que consideraram séries de tempo que se estendem pouco após esses eventos podem não capturar os desdobramentos da política monetária no IT.

Nesse contexto, o presente artigo procura novas evidências acerca da dinâmica inflacionária brasileira que podem estar interligadas com a mudança ocorrida na política monetária em 1999. A estratégia empírica foi desenvolvida em duas etapas. Na primeira, estimam-se probabilidades de ocorrência de três regimes inflacionários (inflação baixa, média e alta) com o método de cadeias de Markov (Hamilton, 1989). Em seguida, constroem-se indicadores de incerteza inflacionária que se associam às alterações da política monetária e aos choques exógenos que atingem a economia (com o método de Evans & Wachtel, 1993). Dessa forma, um tópico adicional de pesquisa é abordado aqui, qual seja: a relação recente entre incertezas inflacionárias e atividade econômica no país.

A literatura econômica com o foco sobre as adversidades da inflação no contexto macroeconômico tem crescido rapidamente.[4] Entre os estudos aplicados à economia brasileira, tem-se Teixeira (1991), que analisou dados de 1970 a 1985 com modelos de heterocedasticidade condicional univariados (ARCH) para a inflação e o produto real, encontrando que a volatilidade da inflação teve impacto negativo, porém não significativo, sobre as taxas de crescimento da economia. Vale (2005) avançou ao estimar sistemas GARCH para a inflação e o produto no período de 1975 a 2001. Seus resultados também não comprovaram efeitos estatisticamente significativos da volatilidade inflacionária sobre a atividade econômica real. Silva Filho (2007) procurou estimar os efeitos da incerteza inflacionária sobre o investimento em máquinas e equipamentos entre 1980 e 2003. O autor considerou como proxies da incerteza os erros de previsão da inflação derivados de modelos autorregressivos simples. Os resultados obtidos mostraram que a incerteza inflacionária afetou negativa e significativamente a acumulação de capital no país. Por sua vez, Costa Filho (2014) analisou os efeitos da incerteza em um sentido geral sobre a atividade econômica, empregando como proxy o número de vezes em que a palavra incerteza foi utilizada em artigos de veículos de mídia selecionados. Suas funções de impulso-reposta para o período de 2004 a 2013

[4]É válido destacar que muitos bancos centrais têm dado crescente ênfase à estabilidade de preços como objetivo central da política monetária, baseados em pesquisas que indicam que inflação elevada e volátil produz efeitos danosos à economia (por exemplo, Friedman, 1977; Gregorio, 1992; Fischer, 1993; e Evans & Wachtel, 1993, entre outros).

indicaram que a produção industrial e o emprego responderam negativamente à elevação da incerteza.

É importante ressaltar, contudo, que modelos autorregressivos clássicos e heterocedásticos, do tipo AR, VAR, ARCH e GARCH, entre outros mais simples, podem produzir estimativas não satisfatórias da incerteza inflacionária porque eles ignoram as instabilidades estruturais causadas por mudanças nos regimes da inflação. Como demonstrado por Evans e Wachtel (1993), os modelos markovianos proporcionam a divisão do erro de previsão da inflação em dois componentes, um associado com as incertezas das realizações futuras da inflação por si, e outro relacionado com as dificuldades de se realizar previsões quando a condução da política monetária pode se alterar. Assim, a presente pesquisa contribui fornecendo estimativas inéditas dessas medidas de incerteza, as quais estão diretamente associadas com as instabilidades geradas por um processo inflacionário sujeito a mudanças de regime. Tal discussão assume revelada importância, uma vez que, mesmo sob o IT, as expectativas de inflação se distanciaram consistentemente da meta de 4,5% definida pelo BCB nos últimos anos da amostra estudada.[5]

O restante do artigo está organizado como a seguir. A seção 2 analisa as fontes de incerteza inflacionária e as relaciona com o modelo markoviano estimado. A seção 3 apresenta os principais resultados e discussões. Por fim, a seção 4 as conclusões da pesquisa.

## 2. A formação de expectativas inflacionárias e sua relação com as mudanças nos regimes

Diversos estudos indicam que os erros de previsão da inflação, presentes nas pesquisas realizadas pelo BCB, são serialmente correlacionados. Este resultado tem sido interpretado de duas maneiras. Na primeira, encontra-se que a formação das expectativas não atende à hipótese de racionalidade, uma vez que os agentes estariam ignorando informações relevantes em seu processo decisório. Carvalho e Bugarin (2006) e Kohlscheen (2012) apresentam argumentos nessa direção. Lima e Alves (2011) sugerem, inclusive, cautela por parte do BCB ao considerar as previsões de mercado em sua formulação de política.

Na segunda linha interpretativa, defende-se que as incertezas quanto ao regime inflacionário corrente e futuro dificultam a realização de boas previsões, mesmo quando os agentes usam racionalmente toda a informação disponível. Segundo Evans e Wachtel (1993), se os participantes do mercado estão antecipando uma mudança no processo inflacionário, ou se eles ainda estão tomando conhecimento de alguma alteração já ocorrida, seus erros de previsão, embora racionais, serão autocorrelacionados e diferentes de zero na média. O restante desta seção irá desenvolver a lógica subjacente a esse raciocínio. Cabe destacar agora que um número crescente de pesquisas na literatura nacional vem especificando a inflação como um processo temporal com mudanças de regimes, modelagem assumida nesta pesquisa (Chauvet, 2001; Figueiredo & Marques, 2011; e Guillén & Garcia, 2014, entre outros).

Evans e Wachtel (1993) descrevem os efeitos de diferentes regimes inflacionários por meio de um exemplo simples: suponha-se que a inflação oscile entre dois processos temporais

[5] As expectativas ficam mais nitidamente desancoradas após meados de 2012, atingindo um pico de 7,3% a.a. em janeiro de 2016. Desde a segunda metade deste ano, entretanto, as previsões de mercado têm convergido para o centro da meta definida pelo Banco Central — em janeiro de 2018, a inflação esperada para doze meses à frente foi de 3,98%.

de acordo com uma variável (estado) não observável, $s_t$, que assume os valores um ($s_t = 1$) ou zero ($s_t = 0$). Denota-se por $P_t = \mathrm{P}\left(s_{t+1} = 1 \middle| \Omega_t\right)$ a probabilidade de que a inflação, $\pi$, seja determinada pelo processo temporal associado ao estado 1, dado o conjunto de informações disponíveis para os agentes no tempo $t$, $\Omega_t$. Nesse sentido, a taxa de inflação esperada é

$$\mathbb{E}_t\, \pi_{t+1} = P_t\, \mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 1, \Omega_t\right] + (1 - P_t)\, \mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 0, \Omega_t\right]. \tag{1}$$

A equação (1) indica que a inflação prevista para o próximo período é uma média ponderada de seu valor condicionada ao regime inflacionário em que a economia se encontra, isto é, estado um ou estado zero. Os pesos são as probabilidades imputadas a cada diferente estado. De acordo com Evans e Wachtel (1993), a equação (1) pode ser originada quando os agentes se deparam com incertezas associadas ao regime inflacionário. Por exemplo, caso ocorra uma mudança na conduta do banco central, os agentes podem levar algum tempo averiguando se a política realmente está em curso. Neste caso, as probabilidades associadas a cada estado serão atualizadas com a introdução de novas observações e a previsão da inflação será dada por (1). Quando os agentes estão, finalmente, certos de que o regime inflacionário se alterou, a sua probabilidade converge para a unidade.

A principal questão que surge com a equação (1) é que expectativas formadas racionalmente podem apresentar viés sistemático. Por exemplo, se a inflação segue o processo temporal do estado um, o erro de previsão *ex post* é definido por

$$\begin{aligned}\pi_{t+1} - \mathbb{E}_t\, \pi_{t+1} = {} & \left(\pi_{t+1} - \mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 1, \Omega_t\right]\right) \\ & + (1 - P_t)\left(\mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 1, \Omega_t\right] - \mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 0, \Omega_t\right]\right). \end{aligned} \tag{2}$$

O primeiro termo entre os parênteses do lado direito da equação (2) tem média zero e é não correlacionado, porque se assumiu que em $t + 1$ a inflação está no regime um.[6] Em outras palavras, espera-se que os erros de previsão sejam independentes com expectativas racionais e que os agentes tenham conhecimento do processo temporal que a inflação segue.

O segundo elemento do lado direito de (2) diz respeito ao erro de previsão relacionado com a incerteza quanto ao regime inflacionário. Percebe-se que ele é a diferença da inflação esperada em cada estado, ponderado pela probabilidade de a inflação estar, de fato, em $s_{t+1} = 0$. Este termo é diferente de zero, enquanto os agentes considerarem a fração $(1 - P_t)$ positiva e que a inflação se comporta de maneira diferente nos dois estados da economia. De maneira geral, a diferença $\mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 1, \Omega_t\right] - \mathbb{E}\left[\pi_{t+1} \mid s_{t+1} = 0, \Omega_t\right]$ pode ser correlacionada com as informações disponíveis em $t$ ($\Omega_t$), as quais incluem erros de previsão passados. Assim, os erros de previsão não serão independentes e irão apresentar viés *ex post*, mesmo quando os agentes utilizam toda a informação disponível ao estimar a inflação futura.

É possível decompor a incerteza inflacionária na presença de mudanças de regimes em dois componentes. Como mostra a equação (3), a variância da inflação para $k$ períodos à frente é representada por (Evans & Wachtel, 1993):

$$\mathrm{Var}\left(\pi_{t+k} \middle| \Omega_t\right) = \mathbb{E}\left[\mathrm{Var}\left(\pi_{t+k} \middle| \Omega_t,\ s_{t+k} \middle| \Omega_t\right)\right] + \mathrm{Var}\left(\mathbb{E}\left[\pi_{t+k} \middle| \Omega_t,\ s_{t+k} \middle| \Omega_t\right]\right). \tag{3}$$

Na equação anterior, o primeiro termo do lado direito é a expectativa da variância da inflação ao longo dos diferentes estados, dado que os agentes sabem em qual regime a inflação estará

[6] $\mathbb{E}\left[\pi_{t+1} - \mathbb{E}\left[\pi_{t+1} \middle| s_{t+1} = 1, \Omega_t\right] \middle| \Omega_t\right] = 0.$

em $t + k$. O segundo termo mede a variância condicional da inflação esperada quando o estado futuro é desconhecido. Se as expectativas são independentes dos estados futuros, tem-se que $\text{Var}\left(\mathbb{E}\left[\pi_{t+k}\middle|\Omega_t, s_{t+k}\middle|\Omega_t\right]\right) = 0$ e, com isso, o primeiro termo do lado direito de (3) se resumirá em $\text{Var}\left(\pi_{t+k}\middle|\Omega_t\right)$. Neste caso, a única fonte de incerteza seria aquela originada dos choques não antecipados na inflação entre $t$ e $t + k$. A literatura se refere a este componente da incerteza como certo equivalente (CE).

Contudo, quando a possibilidade de mudanças no regime inflacionário é considerada pelos agentes econômicos, o segundo termo do lado direito de (3) é positivo. Como ele identifica os efeitos das mudanças de regimes sobre a incerteza das previsões da inflação futura, Evans e Wachtel (1993) o chamam de incerteza do regime (RU, da expressão em inglês *regime uncertainty*). Logo, a variância da inflação pode ser constituída de dois componentes: i) aquele comumente associado aos choques exógenos da inflação (CE); e outro que surge em função da incerteza do regime em que a inflação se encontra hoje e em $k$ períodos à frente (RU).

Com o intuito de distinguir mais claramente a diferença existente entre as duas fontes de incerteza inflacionária (CE e RU), é válido pensar em dois casos extremos. Primeiro, imagina-se que não exista incerteza do regime inflacionário da economia. O público reconhece a aversão à inflação dos encarregados pela política monetária e confia plenamente que um valor estabelecido para a inflação será perseguido (no caso brasileiro, a meta para a inflação). Neste caso, a única fonte de variabilidade da inflação será devido aos choques estocásticos, isto é, os agentes reconhecem que a inflação é uma variável aleatória com certa distribuição de probabilidade. Ao longo do tempo, choques positivos e negativos se cancelariam e o erro de previsão $(\pi_{t+k} - \mathbb{E}_t\left[\pi_{t+k}\right])$ deveria se comportar como um ruído branco. Este raciocínio é representado pela fração CE da variabilidade inflacionária em (3).[7]

Por outro lado, assuma-se que a inflação é completamente determinada pelo Banco Central, que escolhe seu valor de acordo com suas preferências e funções de bem estar social. Porém, o público não sabe exatamente se o(a) presidente do banco é avesso(a) à inflação ou se ele ou ela aceita uma taxa de inflação mais alta. Nesta situação, a economia iria oscilar entre conjunturas inflacionárias. Em um ambiente de máxima incerteza, a inflação esperada seria igual à média das taxas de inflação vigente nos dois regimes, ou seja, $\mathbb{E}\left[\pi_{t+k}\right] = \pi^e = 0{,}5 \times \pi_{\text{baixa}} + 0{,}5 \times \pi_{\text{alta}}$, e a sua variabilidade seria igual à $\text{Var}\left(\pi_{t+k}\right) = 0{,}5 \times (\pi^e - \pi_{\text{baixa}})^2 + 0{,}5 \times (\pi^e - \pi_{\text{alta}})^2 > 0$.

Tal componente está associado com a incerteza inflacionária ges pelas mudanças na política. O erro de previsão oriundo de RU não é necessariamente igual a zero, uma vez que os agentes podem utilizar erros passados para atualizar as probabilidades de que o regime mude no futuro, um processo que pode ser demorado.

Algumas pesquisas recentes mostram que a incerteza do regime inflacionário parece ser um importante componente dos desvios da inflação prevista de seus valores observados. Por exemplo, Carvalho e Minella (2012) encontraram que as previsões dos profissionais de mercado têm um desempenho melhor do que as resultantes da aplicação de modelos estatísticos uni e multivariados (AR, ARMA, VAR, etc.). Segundo os autores, o conjunto de informações utilizado pelos profissionais pode ser mais amplo, combinando o julgamento a

[7]Este raciocínio só é estritamente válido caso o banco central consiga (e queira) ajustar o desvio inflacionário imediatamente e para o horizonte de planejamento $k = 1$. Nas situações em que o banco central pondera os desvios do produto ao longo do potencial e para horizontes mais longos, o erro de previsão pode assumir um comportamento do tipo média móvel.

resultados de modelos econométricos diversos. Adicionalmente Carvalho e Minella (2012) ressaltam que o erro de previsão nas pesquisas realizadas pelo BCB é dominado por um componente que afeta os respondentes de maneira similar, o que pode ser explicado pelas incertezas geradas por mudanças nos regimes inflacionários.

Na próxima seção apresenta-se um modelo de séries temporais para a inflação que permite mudanças nos regimes (um modelo markoviano), capaz de distinguir as frações da variabilidade inflacionária que se referem ao certo equivalente e à incerteza do regime. Em seguida, estas medidas são utilizadas na análise do impacto das incertezas inflacionárias sobre a atividade econômica real por meio do método de vetores autorregressivos (VAR).

## 2.1 Um modelo markoviano para a inflação brasileira e seus estimadores de CE e RU

As medidas de CE e RU dependem dos parâmetros do modelo markoviano selecionado. Neste artigo, após uma série de testes com especificações alternativas ter sido realizada, a versão estimada e analisada foi uma com três estados para a média e a variância, apresentando a seguinte formulação:[8]

$$\pi_{0,t} = \pi_0 + \upsilon_{0,t}, \quad \upsilon_{0,t} \sim \mathcal{N}\left(0, \sigma_0^2\right), \tag{4}$$

$$\pi_{1,t} = \pi_1 + \upsilon_{1,t}, \quad \upsilon_{1,t} \sim \mathcal{N}\left(0, \sigma_1^2\right), \tag{5}$$

$$\pi_{2,t} = \pi_2 + \upsilon_{2,t}, \quad \upsilon_{2,t} \sim \mathcal{N}\left(0, \sigma_2^2\right). \tag{6}$$

Veja-se que as equações (4), (5) e (6) consideram que a série de tempo da inflação brasileira pode oscilar entre três regimes estacionários, 0 (zero), 1 (um) e 2 (dois), com médias e variâncias diferentes.[9] Esta formulação permite, ainda, que os agentes respondam de maneira assimétrica aos choques de baixa e alta inflação, como encontrado por Silva Filho (2007). Por fim, os resultados obtidos com a estimação deste modelo pareceram reproduzir bem as características dos erros de previsão dos analistas de mercado no país, como se mostra na próxima seção, e a associação empírica positiva entre inflação e sua volatilidade (averiguada por Vale, 2005).

No modelo de regimes markovianos, a variável estado não observável ($s_t = 0, 1, 2$) segue uma cadeia em que a probabilidade de transição é definida por

$$p_{i,j} = \mathrm{P}\left(s_{t+1} = i \middle| s_t = j\right), \tag{7}$$

para $i, j = 0, 1, 2$. Em (7), a probabilidade de mudança do regime $j$, em $t$, para $i$, em $t + 1$, depende apenas do estado atual em que a variável se encontra. A matriz completa de

[8] Modelos mais parametrizados para a dinâmica inflacionária foram testados. No caso em que foi possibilitado um termo AR(1) diferente em cada regime, as regressões apresentaram informações não condizentes com fatos estilizados pela literatura, como a correlação positiva entre inflação e sua volatilidade, além de regimes com probabilidades que não correspondiam com a realidade brasileira. A inclusão de um termo AR(1) comum proporcionou um coeficiente estatisticamente significativo e resíduos mais bem comportados. Porém, as medidas de interesse, CE e RU, são muito pouco afetadas. Como pode ser verificado com as equações (9) e (10), as únicas alterações se dão nas probabilidades k períodos à frente. Ainda assim, as séries obtidas foram similares àquelas apresentadas na pesquisa, principalmente em horizontes mais longos, em função da suavização das probabilidades e do termo autorregressivo no tempo. Em suma, incluir um termo AR(1) comum na regressão markoviana não altera as conclusões centrais do artigo.

[9] Testes prévios do tipo ADF padrão e com quebras estruturais indicaram que a inflação no período de 1995 a 2016 é um processo temporal estacionário.

probabilidades de transição $P_3$ é escrita como

$$P_3 = \begin{bmatrix} p_{0,1} & p_{0,2} & p_{0,2} \\ p_{1,0} & p_{1,1} & p_{1,2} \\ p_{2,0} & p_{2,1} & p_{2,2} \end{bmatrix}. \tag{8}$$

O conjunto de parâmetros das equações (4)–(6) e as probabilidades definidas em (8) são obtidos com o método de máxima verossimilhança proposto por Hamilton (1989) e com o algoritmo de suavização de Kim (1994).

De posse das estimativas, o termo certo equivalente (CE) apresentado em (3) é escrito em termos dos parâmetros do modelo como

$$\begin{aligned} &\mathbb{E}\left[\mathrm{Var}\left(\pi_{t+k}\middle|\Omega_t, s_{t+k}\right)\middle|\Omega_t\right] \\ &\quad= \sum_{i=0,1,2} \mathrm{Var}\left(\pi_{i,t+k}\middle|\Omega_t\right) \mathrm{P}\left(s_{t+k}=i\middle|\Omega_t\right) \\ &\quad= \sigma_0^2\,\mathrm{P}\left(s_{t+k}=0\middle|\Omega_t\right) + \sigma_1^2\,\mathrm{P}\left(s_{t+k}=1\middle|\Omega_t\right) + \sigma_2^2\,\mathrm{P}\left(s_{t+k}=2\middle|\Omega_t\right). \end{aligned} \tag{9}$$

A parcela RU é composta do somatório do quadrado das diferenças da inflação esperada em cada par de regimes, ponderada por suas probabilidades de ocorrência. Matematicamente,

$$\begin{aligned} &\mathrm{Var}\left\{\mathbb{E}\left[\pi_{t+k}\middle|\Omega_t, s_{t+k}\right]\middle|\Omega_t\right\} \\ &\quad= \left\{\mathbb{E}\left[\pi_{0,t+k}\middle|\Omega_t\right] - \mathbb{E}\left[\pi_{1,t+k}\middle|\Omega_t\right]\right\}^2 \times \mathrm{P}\left(s_{t+k}=0\middle|\Omega_t\right)\mathrm{P}\left(s_{t+k}=1\middle|\Omega_t\right) \\ &\quad+ \left\{\mathbb{E}\left(\pi_{0,t+k}\middle|\Omega_t\right) - \mathbb{E}\left(\pi_{2,t+k}\middle|\Omega_t\right)\right\}^2 \times \mathrm{P}\left(s_{t+k}=0\middle|\Omega_t\right)\mathrm{P}\left(s_{t+k}=2\middle|\Omega_t\right) \\ &\quad+ \left\{\mathbb{E}\left(\pi_{1,t+k}\middle|\Omega_t\right) - \mathbb{E}\left(\pi_{2,t+k}\middle|\Omega_t\right)\right\}^2 \times \mathrm{P}\left(s_{t+k}=1\middle|\Omega_t\right)\mathrm{P}\left(s_{t+k}=2\middle|\Omega_t\right). \end{aligned} \tag{10}$$

Ao substituir em (10) a configuração do modelo markoviano assumida nesta pesquisa, encontra-se que

$$\begin{aligned} &\mathrm{Var}\left\{\mathbb{E}\left[\pi_{t+k}\middle|\Omega_t, s_{t+k}\right]\middle|\Omega_t\right\} \\ &\qquad= (\pi_0 - \pi_1)^2 \times \mathrm{P}\left(s_{t+k}=0\middle|\Omega_t\right)\mathrm{P}\left(s_{t+k}=1\middle|\Omega_t\right) \\ &\qquad+ (\pi_0 - \pi_2)^2 \times \mathrm{P}\left(s_{t+k}=0\middle|\Omega_t\right)\mathrm{P}\left(s_{t+k}=2\middle|\Omega_t\right) \\ &\qquad+ (\pi_1 - \pi_2)^2 \times \mathrm{P}\left(s_{t+k}=1\middle|\Omega_t\right)\mathrm{P}\left(s_{t+k}=2\middle|\Omega_t\right). \end{aligned} \tag{11}$$

Em (9) e (11), as probabilidades de ocorrência de cada estado $k$ períodos à frente são obtidas da matriz de probabilidades de transição ($P_3$) e das probabilidades suavizadas encontradas pelo algoritmo de estimação. A título de ilustração, a equação (12) mostra o procedimento para os cálculos referentes ao regime zero (Evans & Wachtel, 1993):

$$\mathrm{P}\left(s_{t+k}=0\middle|\Omega_t\right) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} P_3^k \begin{bmatrix} \mathrm{P}\left(s_t=0\middle|\Omega_t\right) \\ \mathrm{P}\left(s_t=1\middle|\Omega_t\right) \\ \mathrm{P}\left(s_t=2\middle|\Omega_t\right) \end{bmatrix}, \tag{12}$$

Em resumo, as equações (9) e (11) permitem estimar as frações do erro de previsão para $k$ períodos à frente relacionadas ao erro originado por choques exógenos nas realizações da inflação (CE) e pelas incertezas quanto aos regimes que a inflação pode se encontrar (RU). Ressalta-se que CE e RU são séries temporais, as quais podem ser utilizadas na avaliação dos efeitos das incertezas sobre a atividade econômica real no país, inclusive por pesquisas posteriores. As seções seguintes apresentam os resultados obtidos com tais análises.

## 3. Resultados

Uma primeira questão analisada diz respeito a possível ocorrência de mudanças estruturais na dinâmica da inflação brasileira. Especificamente, pretende-se mostrar, por meio da estimação de modelos markovianos, que a implantação do Plano Real introduziu uma forte quebra estrutural na série temporal da inflação medida pelo IPCA. A adoção do sistema de metas, por outro lado, produziu resultados menos expressivos nesse sentido. Dessa forma, mesmo tendo como foco o período entre 1995 e 2016, é interessante estender brevemente a amostra de dados para os anos de 1980. A Tabela 1 traz os resultados das estimações dos modelos com dois e três regimes para a média e a variância, considerando a série de valores mensais da taxa de inflação — MS(2) e MS(3), respectivamente.[10]

A regressão com dois estados divide a dinâmica inflacionária brasileira em períodos de inflação baixa e estável (regime zero) e de inflação alta e instável (regime um).[11] Como pode ser visto na Figura 1(a), há uma clara concentração do estado de inflação alta (áreas cinza) nos primeiros anos da amostra. As únicas exceções ocorrem nos meses de ajustamento entre 1995 e 1996 e no episódio do choque cambial de 2002/2003. Quanto às probabilidades de transição, o modelo estimado mostra que a duração esperada dos diferentes regimes é elevada, respectivamente, 125 e 65 meses nos regimes R0 e R1. Ademais, o baixo valor para $p_{1,0}$ (menor que 1%) indica que a probabilidade de a economia brasileira visitar um regime de inflação tal como observado nos anos de 1980 e início da década de 1990 é extremamente

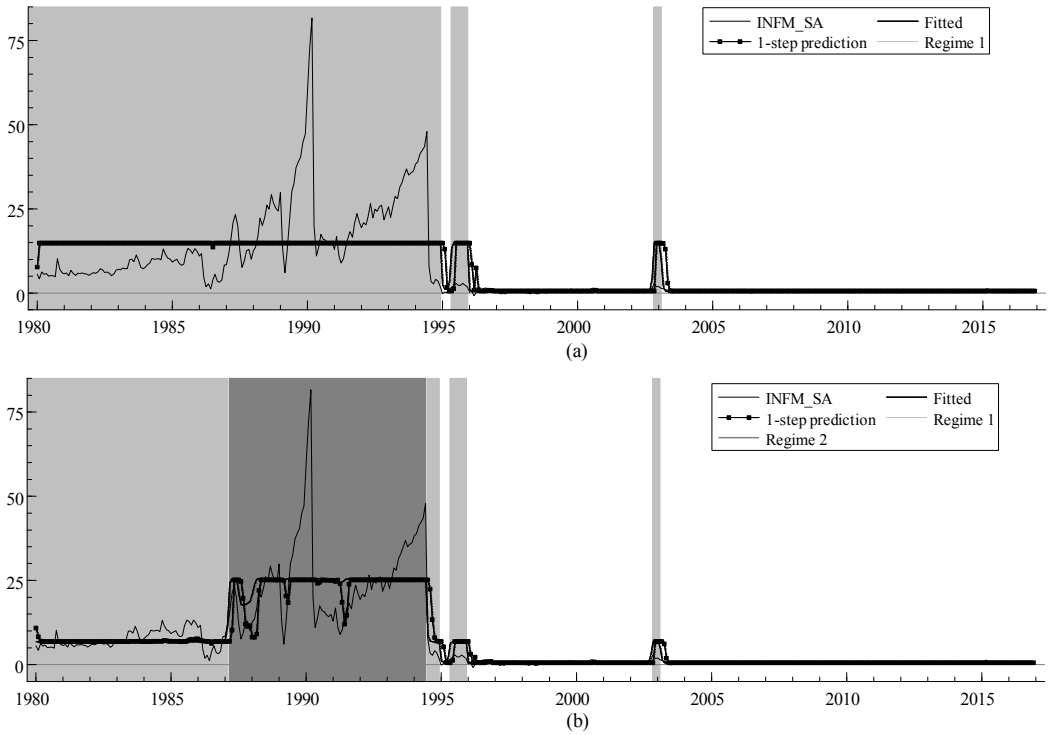
**Tabela 1.** Resultados dos modelos markovianos, amostra de dados 1980–2016.

| | MS(2) | | MS(3) | |
|---|---|---|---|---|
| Coeficiente | Valor estimado | Valor-p | Valor estimado | Valor-p |
| $\pi_0$ | 0,499 | 0,000 | 0,498 | 0,000 |
| $\pi_1$ | 15,113 | 0,000 | 6,758 | 0,000 |
| $\pi_2$ | – | – | 25,501 | 0,000 |
| $\sigma_0$ | 0,327 | 0,000 | 0,325 | 0,000 |
| $\sigma_1$ | 13,021 | 0,000 | 3,195 | 0,000 |
| $\sigma_2$ | – | – | 13,182 | 0,000 |
| $p_{0,0}$ | 0,992 | 0,000 | 0,992 | 0,000 |
| $p_{0,1}$ | 0,015 | 0,086 | 0,028 | 0,085 |
| $p_{1,1}$ | – | – | 0,958 | 0,000 |
| $p_{1,2}$ | – | – | 0,018 | 0,412 |
| Estatística AIC | 3,932 | – | 3,324 | – |
| Teste LR de linearidade | | | | |
| $\chi^2(4)$ | 1673,3 | 0,000 | | |
| $\chi^2(8)$ | | | 1950,9 | 0,000 |

*Notas:* Dados dessazonalizados com o método aditivo *Census* X-12. As probabilidades não apresentadas podem ser obtidas de (8) e pelo fato de que suas linhas e, ou, colunas somam um. Testes LR de linearidade indicam que os estados markovianos são significativos.

[10] Como os valores para a taxa de inflação no período de 1985 a 1994 são elevados, a anualização dos valores mensais observados não produziu bons resultados nos modelos econométricos.

[11] Doravante, refere-se aos regimes $i$, $i = 0, 1, 2$, como R0, R1 ou R2.




*Notas:* O eixo vertical representa os valores da inflação; as áreas em escala de cinza, os diferentes estados markovianos.

**Figura 1.** Regimes inflacionários nos modelos markovianos de dois (a) e três (b) estados, 1980–2016.

baixa. Pode-se dizer, portanto, que a implantação do Plano Real introduziu uma quebra estrutural na dinâmica da inflação no país.

Por sua vez, o modelo MS(3) caracteriza, nos regimes R0, R1 e R2, estados de inflação baixa, média e alta na amostra. É possível perceber, ainda, que a variância se eleva conjuntamente com a estimativa da inflação em cada estado, refletindo uma associação positiva entre a taxa de inflação e sua volatilidade, amplamente documentada na literatura como um fato estilizado (Okun, 1971; Friedman, 1977; Ball & Cecchetti, 1990; Evans & Wachtel, 1993, entre outros).

A Figura 1(b) representa graficamente as probabilidades suavizadas do modelo MS(3). As áreas cinza claras e escuras denotam os períodos de inflação média e alta, respectivamente. Nota-se que a datação obtida é muito similar à proposta por Minella (2003). O principal ponto que emerge dessa figura é a dificuldade de se visualizar claramente diferenças da dinâmica inflacionária no início do Plano Real e no período de implementação do sistema de metas. A fim de analisar esta questão mais a fundo, estimam-se os modelos de cadeias de Markov com dois e três estados na amostra restrita aos anos mais recentes.

Como descrito na Tabela 2, o modelo de dois regimes na amostra 1995–2016 traça períodos de inflação baixa e estável (R0) e de inflação alta e instável (R1). Em termos anuais, esses estados estão associados com taxas de inflação de 5,7% e 16,4% a.a. A duração esperada de R0 é de 60 meses e a de R1 é bastante inferior, de oito meses. Dessa forma, o modelo markoviano com dois estados indica que a inflação brasileira oscila na maior parte do tempo ao longo de uma taxa de inflação de 6% a.a., com períodos mais curtos em que a inflação se acelera em função de choques. A Figura 2(a) representa as probabilidades suavizadas

**Tabela 2.** Resultados dos modelos markovianos, amostra de dados 1995-2016.

| | MS(2) | | MS(3) | |
|---|---|---|---|---|
| Coeficiente | Valor estimado | Valor-p | Valor estimado | Valor-p |
| $\pi_0$ | 0,464 | 0,000 | 0,449 | 0,000 |
| $\pi_1$ | 1,276 | 0,000 | 0,507 | 0,000 |
| $\pi_2$ | – | – | 1,862 | 0,000 |
| $\sigma_0$ | 0,261 | 0,000 | 0,177 | 0,000 |
| $\sigma_1$ | 0,857 | 0,000 | 0,385 | 0,000 |
| $\sigma_2$ | – | – | 0,592 | 0,000 |
| $p_{0,0}$ | 0,983 | 0,000 | 0,989 | 0,000 |
| $p_{0,1}$ | 0,120 | 0,068 | 0,013 | 0,341 |
| $p_{1,1}$ | – | – | 0,959 | 0,000 |
| $p_{1,2}$ | – | – | 0,183 | 0,000 |
| Estatística AIC | 0,678 | – | 0,524 | – |
| Teste LR de linearidade | | | | |
| $\chi^2(4)$ | | | | |
| $\chi^2(8)$ | | | | |

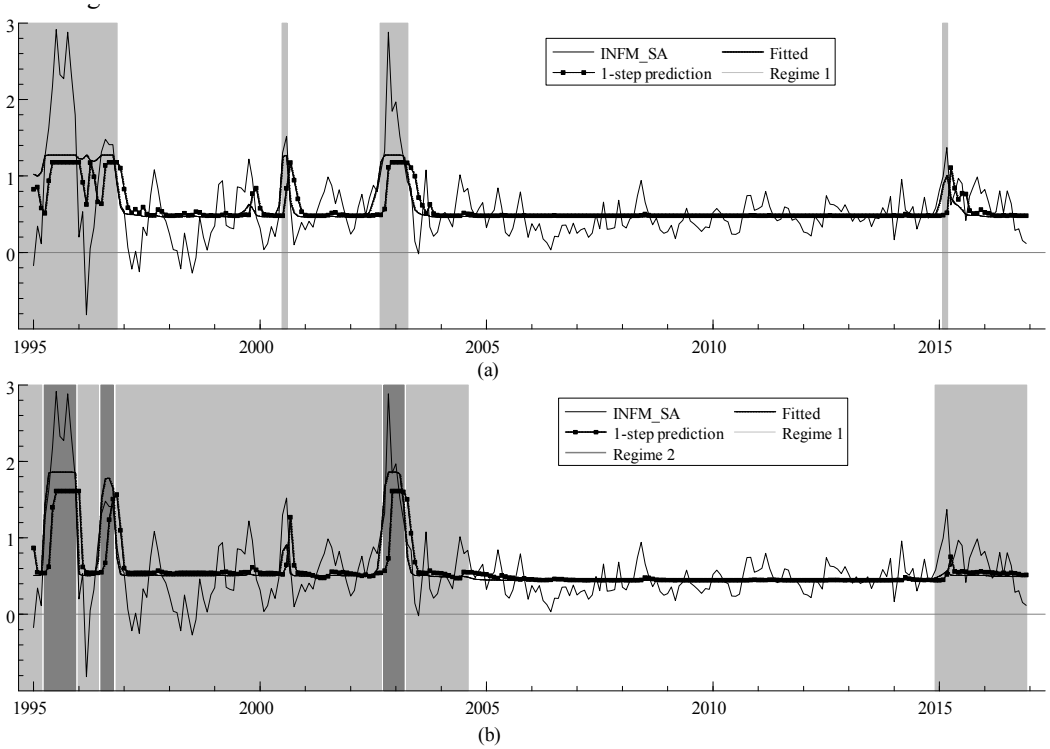
*Notas:* Dados dessazonalizados com o método aditivo *Census* X-12. As probabilidades não apresentadas podem ser obtidas de (8) e pelo fato de que suas linhas e, ou, colunas somam um. Testes LR de linearidade indicam que os estados markovianos são significativos.

de R0 (áreas claras) e R1 (áreas cinza). Ainda, veja-se que o modelo MS(2) da Tabela 2 e Figura 2(a) também não indica qualquer alteração substancial no comportamento da inflação após a implantação do IT no país. Os resultados do modelo MS(3) apresentados na Tabela 2 caracteriza a inflação brasileira variando em regimes de inflação baixa e estável (R0, 5,5% a.a.), média (R1, 6,3% a.a.) e alta e instável (R2, 24,8% a.a.). As durações de cada regime foram de, respectivamente, 91, 24 e seis meses, ou seja, R0 e R1 são mais frequentes na amostra.[12]

Diferente dos demais modelos estimados, o MS(3) na amostra 1995–2016 indica que o sistema de metas pode ter, de fato, produzido efeitos favoráveis sobre a inflação no país. Por exemplo, a Figura 2(b) descreve que o regime de inflação baixa e estável vigorou apenas após a implantação do IT. Especificamente, R0 foi verificado no período que variou de setembro de 2004 a dezembro de 2014. Além disso, o regime de inflação alta e instável ocorreu em 19 meses da amostra. Destes, 13 (ou 68%) se localizaram entre 1995 e 1996. A Tabela 3 resume a classificação dos regimes durante as duas abordagens de política monetária no Brasil pós-Real, a saber, de câmbio administrado (1.995 a junho de 1999) e de metas para a inflação (julho de 1999 em diante).

A Tabela 3 traz informações relevantes em termos da comparação do desempenho das duas abordagens de política monetária no país. Percebe-se que, nos 45 meses de adoção do sistema de câmbio administrado, a inflação oscilou entre os regimes de inflação média e alta. Após a adoção do sistema de metas, a inflação observada passa a oscilar, em 97% dos

[12]Destaca-se que todos os modelos estimados calcularam a inflação no regime de baixos preços em torno de 5,5 a 6% a.a. Tal fato pode estar associado com um viés inflacionário do BCB, uma vez que a meta para a inflação é de 4,5% desde 2005.




*Notas:* O eixo vertical representa os valores da inflação; as áreas em escala de cinza, os diferentes estados markovianos.

**Figura 2.** Regimes inflacionários nos modelos markovianos de dois (a) e três (b) estados, 1995–2016.

**Tabela 3.** Classificação dos regimes de acordo com MS(3) e as políticas monetárias no Brasil – 1995–2016.

| Sistema monetário | Inflação baixa (R0) | Inflação média (R1) | Inflação Alta (R2) | Total |
|---|---|---|---|---|
| *Painel (a): contagem das ocorrências* | | | | |
| Câmbio administrado | 0 | 32 | 13 | 45 |
| Metas para inflação | 123 | 90 | 6 | 219 |
| Total | 123 | 122 | 19 | 264 |
| *Painel (b): percentuais das ocorrências (%)* | | | | |
| Câmbio administrado | 0 | 71 | 29 | 100 |
| Metas para inflação | 56 | 41 | 3 | 100 |

casos, entre os regimes de inflação baixa e média.[13] Dessa forma, é possível concluir que a adoção do sistema de metas em meados de 1999 conseguiu reduzir a inflação do patamar médio/alto para o baixo/médio.

Das especificações econométricas anteriores, aquela que parece representar melhor as mudanças ocorridas na dinâmica recente da inflação brasileira é, portanto, a MS(3).[14]. Derivam-se deste modelo as séries de tempo da incerteza inflacionária (RU) e do certo equivalente (CE), a fim de associar tais medidas ao desempenho da atividade econômica no país. A Figura 3 ilustra as estimativas de RU e CE, obtidas com as equações (9) e (11), no eixo direito dos gráficos, contra os valores observados da inflação entre 1995 e 2016, no eixo esquerdo.

Como pode ser visto nos painéis (a)–(d), ambas as medidas de incerteza têm picos que coincidem com os da inflação, mas o componente RU exibe oscilações que acompanham mais de perto as movimentações observadas do IPCA. Ademais, é possível perceber que há uma queda de patamar e maior estabilização dos indicadores de incerteza quando se comparam os períodos de 1995 a 2003 e de 2004 a 2014. Ou seja, o amadurecimento do regime de metas no país trouxe ganhos no que diz respeito à redução da incerteza inflacionária. Contudo, com a aceleração da inflação ocorrida em 2015, as variâncias retornam a valores muito próximos aos observados no início do Plano Real. Estes resultados mostram que a confiança

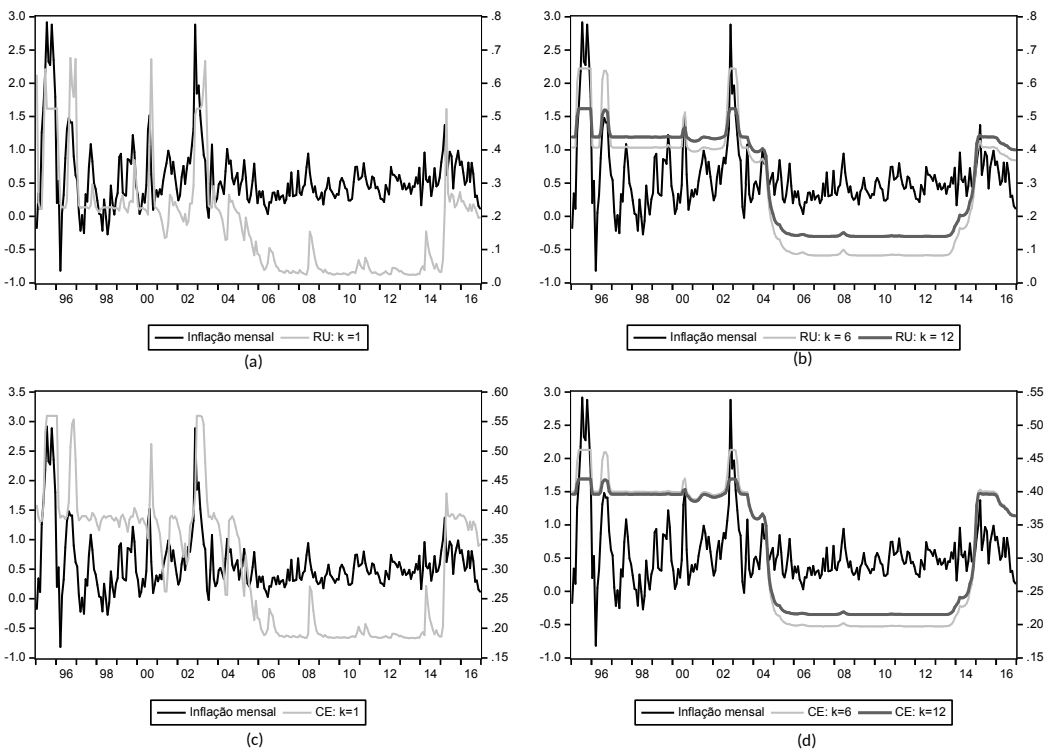



**Figura 3.** Medidas dos componentes da incerteza inflacionária. (a) RU um período à frente; (b) RU seis e doze períodos à frente; (c) CE um período à frente; (d) CE seis e doze períodos à frente.

[13] As únicas ocorrências de regimes de inflação alta e instável no sistema de metas se dão no período da crise de confiança que gerou alto repasse cambial em 2002/2003.

[14] Em função de sua boa aderência à história recente da inflação e dos critérios de informação de Akaike (AIC).

do público quanto à capacidade do BCB em manter a inflação estabilizada ao longo da meta pode ter sido abalada pelo descontrole ocorrido nos últimos anos da amostra (dado o crescimento de RU).

Uma análise da participação relativa de CE e RU na variância total da inflação um período à frente reforça tal constatação. Entre 1995 e 2003, esses componentes explicaram, respectivamente, 60% e 40% da variância. Por sua vez, no período de 2004 a 2014, a parcela devido a CE se elevou para 80% em média, mostrando que a maior parte das incertezas nesses anos era gerada por choques exógenos e não por dúvidas quanto ao regime inflacionário futuro. A partir de 2015 o percentual devido a RU dobra, alcançando novamente 40% (tal padrão se repete para os demais horizontes de planejamento).

O BCB pode ainda verificar o estado das incertezas da economia ao analisar o desvio padrão das expectativas coletadas junto aos especialistas de mercado, muito embora esta *proxy* não seja perfeita.[15] Verificou-se que as estimativas do certo equivalente e da incerteza do regime tem boa correlação com esse indicador. Especificamente, os coeficientes de correlação entre CE e o desvio padrão das expectativas[16] foi igual a 0,65, 0,59 e 0,55 nos horizontes de previsão de um, seis e 12 períodos à frente. No caso de RU, os valores das correlações foram, nesta mesma ordem, de 0,71, 0,67 e 0,58. Portanto, as medidas derivadas do modelo markoviano são bons indicadores da incerteza dos respondentes das pesquisas realizadas pelo BCB. Além disso, o procedimento empregado nesta pesquisa tem o benefício de caracterizar as incertezas quanto ao regime inflacionário, intimamente associadas com as ações dos formuladores de políticas.

Avalia-se, agora, o impacto das fontes de incertezas inflacionárias sobre a atividade econômica. Para tanto, estimam-se modelos VAR contendo cinco variáveis, a saber: os indicadores CE e RU,[17] o produto interno bruto (PIB), o consumo das famílias e o investimento (formação bruta de capital fixo).[18] Por questões de limitações da base de dados, a frequência das séries utilizadas para os agregados macroeconômicos é trimestral, variando de 1997 a 2016. Ademais, para obter modelos vetoriais estacionários, os dados foram logaritmizados e o componente cíclico (multiplicado por cem) foi extraído com o filtro-HP. Conforme procedimento padrão da literatura, assumiu-se o valor de 1.600 para o parâmetro de suavização (Hodrick & Prescott, 1997). Com respeito aos indicadores de incertezas CE e RU, a conversão da periodicidade mensal (original) para a trimestral se deu pela média aritmética.

---

[15] Por exemplo, em alguns momentos vários profissionais podem concordar em elevadas incertezas no futuro. Assim, a dispersão das expectativas será baixa, muito embora as incertezas sejam altas. Ver a esse respeito em Giordani e Söderlind (2003) e Lahiri e Sheng (2010).

[16] Inflação acumulada para os próximos 12 meses, retiradas do Sistema de Expectativas de Mercado, disponível em https://www3.bcb.gov.br/expectativas/publico/consulta/serieestatisticas

[17] Considera-se um par de indicadores de incerteza por turno, de acordo com os horizontes de previsão analisados. Esta abordagem se baseia em Evans e Wachtel (1993).

[18] Especificamente, as séries utilizadas foram:

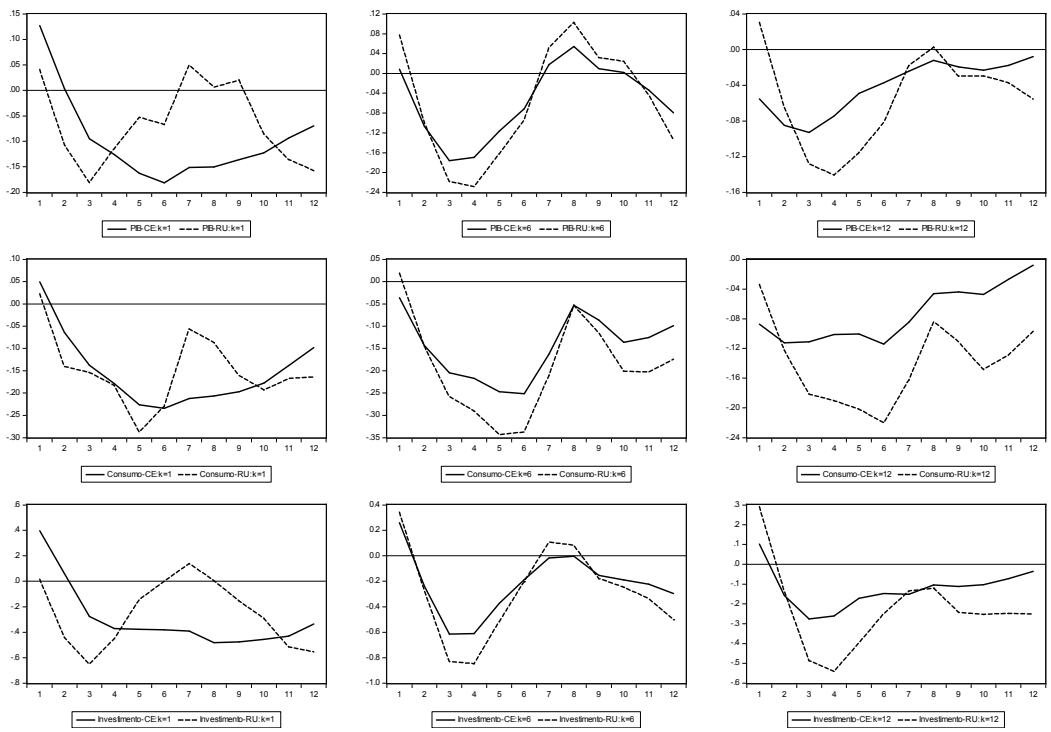
PIB – preços de mercado – índice encadeado – dessaz. (média 1995 = 100) – ref. 2010;
Consumo final – famílias – índice encadeado – dessaz. (média 1995 = 100) – ref. 2010; e,
Capital fixo – formação bruta – índice encadeado – dessaz. (média 1995 = 100) – ref. 2010.

As séries são medidas pelo Instituto Brasileiro de Geografia e Estatística (IBGE, 2016) e disponibilizadas pelo Instituto de Pesquisa Econômica Aplicada, IPEA (http://www.ipeadata.gov.br/Default.aspx).

As análises a seguir mostram o efeito de choques de um desvio-padrão generalizado das incertezas sobre as variações cíclicas da atividade econômica.[19] Como a questão da significância estatística destes impactos é de suma importância, as funções impulso-resposta (FIR) foram obtidas com o procedimento de projeções locais proposto por Jordà (2005, 2009).[20]

As FIR obtidas por projeções locais possuem algumas vantagens. Por exemplo, não é necessário conhecer *a priori* o sistema multivariado gerador dos dados. Assim, possíveis erros de má-especificação são minimizados (se o processo gerador dos dados for, de fato, um VAR, então as FIR do método de projeções locais são equivalentes às obtidas pelo método clássico). Além disso, dada a linearidade do método, as projeções locais podem ser obtidas com estimadores robustos à heterocedasticidade e autocorrelação (como discutido em Jordà, 2009).

A Figura 4 ilustra as respostas dos ciclos dos agregados PIB, consumo e investimento aos choques nas incertezas associadas às flutuações aleatórias da inflação (CE) e às políticas do




*Notas:* A primeira linha representa as respostas dos ciclos do PIB; a segunda, do consumo; e, a terceira, do investimento. Nas colunas as variáveis respondem à incerteza nos horizontes de um, seis e doze períodos à frente, respectivamente.

**Figura 4.** Respostas dos agregados macroeconômicos a choques em CE ou RU.

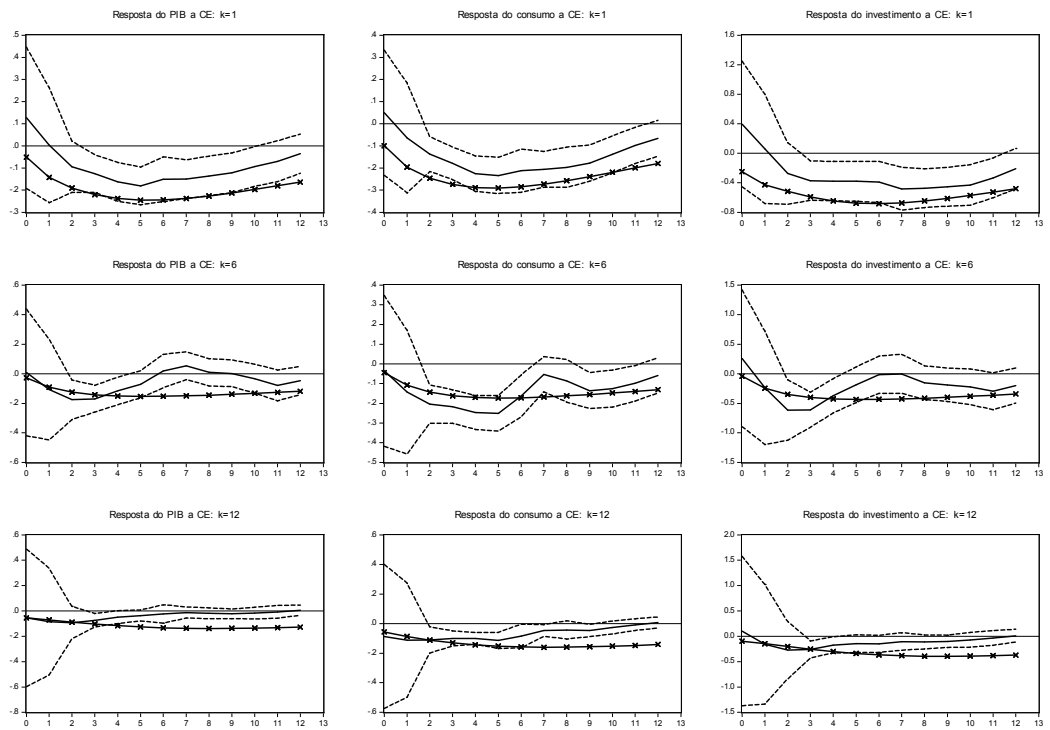
[19] Uma vez que as incertezas podem afetar e serem afetadas pela atividade econômica, optou-se pela construção de uma matriz de inovações ortogonais que independe da ordenação das variáveis para a obtenção das FIR.

[20] Este método utiliza a matriz triangular de coeficientes do VAR em sua forma estrutural e regressões das variáveis endógenas projetadas *h* períodos à frente contra seus valores correntes e defasados para construir suas FIR. Para cada horizonte definido na função impulso-resposta, realiza-se uma projeção com esse formato (Jordà, 2005). Como o modelo é linear por essência, o cômputo de intervalos de significância e a realização de testes estatísticos são facilitados. Neste artigo, seguindo o critério de seleção de Schwartz (BIC), utiliza-se modelos VAR(1) em todos os casos.

banco central (RU). Ela está organizada em formato matricial, com três linhas e três colunas. As linhas trazem as respostas aos choques do PIB, consumo e investimento, nesta ordem. Nas colunas, alteram-se os horizontes em que a incerteza é mensurada ($k = 1, 6, 12$). As linhas sólidas representam a resposta aos choques em CE e as pontilhadas, em RU.

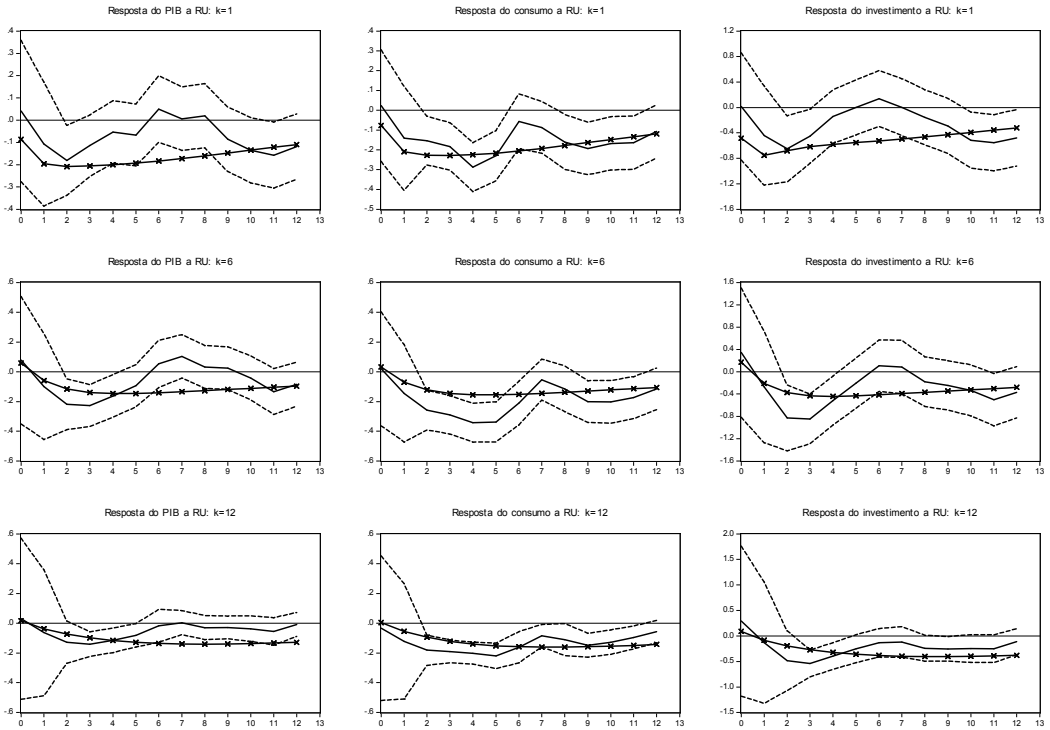
As funções impulso-resposta indicam efeitos negativos das incertezas sobre todos os agregados. Estes são, em grande parte, significativos ao nível de 5% — as significâncias individuais podem ser atestadas pelas Figuras 5 e 6. Uma informação interessante destacada pela Figura 4 é a de que os efeitos negativos das incertezas geradas pela política monetária tendem a ser superiores àqueles ocasionados por choques aleatórios na inflação, especialmente nos horizontes de previsão de seis e 12 meses.

A Tabela 4 traz o efeito de choques em CE e RU sobre os ciclos dos agregados analisados, acumulados em 12 períodos. Como as séries de tempo das variáveis macroeconômicas passaram por filtragens HP, os valores descritos na tabela mostram os desvios percentuais em torno da tendência advindos de elevações nas incertezas. Nota-se que o PIB, o consumo e o investimento respondem negativa e significativamente às perturbações em vários dos casos descritos. Assim, surpresas monetárias, mesmo surtindo algum efeito estimulador sobre a economia em prazos mais curtos, como pode ser visto nos primeiros trimestres dos gráficos da Figura 4, produzem impactos acumulados de longo prazo negativos e significativos sobre a atividade econômica. Esta argumentação levanta críticas ao uso de políticas monetárias excessivamente discricionárias e intervencionistas. Nessa mesma linha de raciocínio, Cunha e Ferreira (2004) e Couto e Gomes (2017) argumentam que instabilidades crescentes na trajetória do consumo geram perdas significativas de bem-estar no Brasil.




*Notas:* A linha contínua representa as respostas das variáveis ao impulso em CE; as linhas tracejadas, os intervalos de confiança de 90%; a linha marcada com cruzes, as FIRs obtidas com o método clássico.

**Figura 5.** Resposta dos agregados aos choques em CE.




*Notas:* A linha contínua representa as respostas das variáveis ao impulso em CE; as linhas tracejadas, os intervalos de confiança de 90%; a linha marcada com cruzes, as FIRs obtidas com o método clássico.

**Figura 6.** Resposta dos agregados aos choques em RU.

**Tabela 4.** Efeitos dos choques das incertezas sobre os agregados acumulados 12 períodos à frente.

| Resposta dos ciclos do | CE ($k = 1$) | RU ($k = 1$) | CE ($k = 6$) | RU ($k = 6$) | CE ($k = 12$) | RU ($k = 12$) |
|---|---|---|---|---|---|---|
| PIB | −1,16* | −0,78*** | −0,66*** | −0,69 | −0,50 | −0,67 |
| Consumo | −1,82* | −1,80** | −1,76* | −2,31* | −0,88 | −1,68** |
| Investimento | −3,52* | −3,04** | −2,65*** | −3,40*** | −1,49 | −2,77 |

*, ** e *** denotam significância dos efeitos acumulados a 1, 5 e 10%, respectivamente.

## 4. Conclusão

A presente pesquisa analisou a inflação brasileira no período entre 1995 e 2016, por meio de regressões com estados markovianos, objetivando fornecer informações adicionais acerca da dinâmica recente dos preços no país. No âmbito da análise das ações do BCB, a metodologia empregada apresentou o benefício de gerar medidas de incertezas que estão diretamente relacionadas com a condução da política monetária, as quais foram associadas com variáveis que medem o desempenho da atividade econômica agregada.

Os resultados obtidos mostraram que houve redução no nível e na variabilidade da inflação, a qual oscilou na faixa médio-alta no período de câmbio administrado (1.995 a junho de 1999) e na faixa baixo-média após a adoção do sistema de metas (de julho de 1999 em diante). De fato, no primeiro período, a inflação observada esteve em 71% dos casos no estado de inflação média; e, 29% no de alta. Por outro lado, na amostra do regime de metas,

os estados com variações baixas e médias nos preços foram detectados em 56% e 41% do tempo, respectivamente. Assim, argumentou-se que a nova forma de conduzir a política monetária no país produziu resultados positivos em termos de redução inflacionária, ainda que não tão expressivos como os obtidos com a implantação do Plano Real por si.

Talvez essa seja uma razão para a controvérsia existente nas análises empíricas consultadas que procuraram comparar o desempenho da política monetária no Brasil pós-1995, com alguns estudos que indicaram melhorias advindas da adoção do IT (por exemplo, Barbosa-Filho, 2008; BCB, 2011; e Carvalho & Minella, 2012), enquanto outros não evidenciaram diferenças entre os resultados da política monetária de câmbio administrado e de metas (Mello & Moccero, 2011; Arestis, Ferrari-Filho, & de Paula, 2011). Mostra-se neste artigo que ambas as conclusões podem ser obtidas dos dados, uma vez que essas análises não levaram em consideração os diferentes estados inflacionários e suas probabilidades de ocorrência.

Ademais, a adoção do sistema de metas para a inflação esteve associada com uma redução significativa das incertezas quanto ao regime inflacionário (RU) escolhido pelo banco central. Verificou-se que entre 1995 e 1999 o nível desta variável era alto. O período que seguiu à adoção do IT foi perturbado por uma série de choques exógenos (cambiais, energéticos e políticos) que afetaram as taxas de inflação e levantaram dúvidas sobre a manutenção da política, mantendo RU elevada. Contudo, desde o fim de 2003 as incertezas do regime inflacionário chegaram a mínimos históricos, não sendo afetadas, inclusive, pela última crise financeira internacional em 2007/08. Tal dinâmica só foi alterada em 2015, quando os controles artificiais sobre os preços, a situação fiscal problemática e a (provável) menor aversão à inflação do banco central colocaram em cheque a condução da política monetária brasileira.

Concomitantemente, no segundo semestre de 2014 o país adentrou em uma de suas piores recessões econômicas. As funções impulso-resposta estimadas permitiram afirmar que o recrudescimento e a desestabilização da inflação brasileira dos últimos anos contribuíram para tal processo, uma vez que as instabilidades na condução da política podem produzir efeitos negativos sobre a produção agregada, o consumo e o investimento. Além disso, os efeitos das incertezas quanto ao regime se mostraram superiores aos de movimentações originadas por choques exógenos nos índices de preços. Dessa forma, mostrou-se que a política monetária, ao prezar por seu objetivo primordial da estabilidade da inflação, tem o potencial de gerar um importante efeito colateral, qual seja, o de estimular a atividade produtiva no país.

## Referências bibliográficas

Arestis, P., Ferrari-Filho, F., & de Paula, L. F. (2011). Inflation targeting in Brazil. *International Review of Applied Economics*, *25*(2), 127–148. http://dx.doi.org/10.1080/02692171.2010.483465

Ball, L., & Cecchetti, S. G. (1990). Inflation and uncertainty at short and long horizons. *Brookings Papers on Economic Activity*, *1990*(1), 215–254. https://www.brookings.edu/bpea-articles/inflation-and-uncertainty-at-short-and-long-horizons

Barbosa-Filho, N. H. (2008). Inflation targeting in Brazil: 1999–2006. *International Review of Applied Economics*, *22*(2), 187–200. http://dx.doi.org/10.1080/02692170701880684

BCB – Banco Central do Brasil. (2011). *Dez anos de metas para a inflação — 1999–2009*. Brasília, DF: Banco Central do Brasil. https://www.bcb.gov.br/pt-br/#!/n/LIVRO10ANOSMI

Carvalho, F. A. d., & Bugarin, M. S. (2006). Inflation expectations in Latin America. *Economía*, *6*(2), 101–145. http://dx.doi.org/10.1353/eco.2006.0014

Carvalho, F. A. d., & Minella, A. (2012). Survey forecasts in Brazil: A prismatic assessment of epidemiology, performance, and determinants. *Journal of International Money and Finance*, *31*(6), 1371–1391. http://dx.doi.org/10.1016/j.jimonfin.2012.02.006

Chauvet, M. (2001). Indicadores antecedentes da inflação brasileira. *Pesquisa e Planejamento Econômico*, *31*(1). http://ppe.ipea.gov.br/index.php/ppe/article/view/160

Costa Filho, A. E. d. (2014). Incerteza e atividade econômica no Brasil. *Economia Aplicada*, *18*(3), 421–453. http://dx.doi.org/10.1590/1413-8050/ea607

Couto, G. T., & Gomes, F. A. R. (2017). Custo de bem-estar da incerteza macroeconômica na América Latina. *Revista Brasileira de Economia*, *71*(2), 137–152. http://dx.doi.org/10.5935/0034-7140.20170007

Cunha, B. V. B. V. d., & Ferreira, P. C. (2004). Custo de ciclo econômico no Brasil em um modelo com restrição a crédito. *Estudos Econômicos*, *34*(2), 243–268. http://dx.doi.org/10.1590/S0101-41612004000200001

Evans, M., & Wachtel, P. (1993). Inflation regimes and the sources of inflation uncertainty. *Journal of Money, Credit and Banking*, *25*(3), 475–511. http://dx.doi.org/10.2307/2077719

Figueiredo, E. A. d., & Marques, A. M. (2011). Inflação inercial sob mudançaas de regime: Análise a partir de um modelo MS-ARFIMA, 1944–2009. *Economia Aplicada*, *15*(3), 443–457. http://dx.doi.org/10.1590/S1413-80502011000300005

Fischer, S. (1993). The role of macroeconomic factors in growth. *Journal of monetary economics*, *32*(3), 485–512. http://dx.doi.org/10.1016/0304-3932(93)90027-D

Friedman, M. (1977). Nobel lecture: Inflation and unemployment. *Journal of Political Economy*, *85*(3), 451–472. http://dx.doi.org/10.1086/260579

Giordani, P., & Söderlind, P. (2003). Inflation forecast uncertainty. *European Economic Review*, *47*(6), 1037–1059. http://dx.doi.org/10.1016/S0014-2921(02)00236-2

Gregorio, J. d. (1992). The effects of inflation on economic growth: Lessons from Latin America. *European Economic Review*, *36*(2-3), 417–425. http://dx.doi.org/10.1016/0014-2921(92)90098-H

Guillén, D., & Garcia, M. (2014). Expectativas desagregadas, credibilidade do Banco Central e cadeias de Markov. *Revista Brasileira de Economia*, *68*(2), 197–223. http://dx.doi.org/10.1590/S0034-71402014000200003

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, *57*(2), 357–384. http://dx.doi.org/0012-9682(198903)57:2<357:ANATTE>2.0.CO;2-2

Hodrick, R. J., & Prescott, E. C. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking*, *29*(1), 1–16. http://dx.doi.org/10.2307/2953682

IBGE – Instituto Brasileiro de Geografia e Estatística. (2016). *Contas nacionais trimestrais: Ano de referência 2010* (3ª ed.) [Série Relatórios Metodológicos]. Rio de Janeiro: IBGE.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review*, *95*(1), 161–182. http://dx.doi.org/10.1257/0002828053828518

Jordà, Ò. (2009). Simultaneous confidence regions for impulse responses. *Review of Economics and Statistics*, *91*(3), 629–647. https://www.jstor.org/stable/25651363

Kim, C.-J. (1994). Dynamic linear models with Markov-switching. *Journal of Econometrics*, *60*(1-2), 1–22. http://dx.doi.org/10.1016/0304-4076(94)90036-1

Kohlscheen, E. (2012). Uma nota sobre erros de previsão da inflação de curto-prazo. *Revista Brasileira de Economia*, *66*(3), 289–297. http://dx.doi.org/10.1590/S0034-71402012000300002

Lahiri, K., & Sheng, X. (2010). Measuring forecast uncertainty by disagreement: The missing link. *Journal of Applied Econometrics*, *25*(4), 514–538. http://dx.doi.org/10.1002/jae.1167

Lima, E. C. R., & Alves, P. P. (2011, junho). *O desempenho do mercado (Focus) na previsão da inflação: Novos resultados?* (Texto para Discussão Nº 1621). Brasília, DF: Instituto de Pesquisa Econômica Aplicada (IPEA). http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=9801

Mello, L. d., & Moccero, D. (2011). Monetary policy and macroeconomic stability in Latin America: The cases of Brazil, Chile, Colombia and Mexico. *Journal of International Money and Finance*, *30*(1), 229–245. http://dx.doi.org/10.1016/j.jimonfin.2010.08.002

Minella, A. (2003). Monetary policy and inflation in Brazil (1975–2000): a VAR estimation. *Revista Brasileira de Economia*, *57*(3), 605–635. http://dx.doi.org/10.1590/S0034-71402003000300005

Okun, A. M. (1971). The mirage of steady inflation. *Brookings Papers on Economic Activity*, *1971*(2), 485–498. https://www.brookings.edu/bpea-articles/the-mirage-of-steady-inflation

Silva Filho, T. N. T. d. (2007, dezembro). *Is the investment-uncertainty link really elusive? The harmful effects of inflation uncertainty in Brazil* (Working Paper Nº 157). Brasília, DF: Banco Central do Brasil. https://www.bcb.gov.br/pec/wps/ingl/wps157.pdf

Teixeira, E. (1991). Incerteza inflacionária e crescimento do produto e incerteza do produto e crescimento inflacionário. *Nova Economia*, *2*(2), 123–133. https://revistas.face.ufmg.br/index.php/novaeconomia/article/view/2320

Vale, R. V. (2005). Inflation, growth and real and nominal uncertainty: Some bivariate Garch-in-mean evidence for Brazil. *Revista Brasileira de Economia*, *59*(1), 127–145. http://dx.doi.org/10.1590/S0034-71402005000100006

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180021
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Assessing Interdependence Among Countries' Fundamentals and its Implications for Exchange Rate Misalignment Estimates: An Empirical Exercise Based on GVAR*

Emerson Fernandes Marçal[†], Beatrice Zimmermann[‡], Diogo de Prince[§], Giovanni Merlin[¶]

**Contents**

1. Introduction ..... 429
2. A short review of exchange rate misalignment literature ..... 430
3. Methodologies to calculate exchange rate misalignment ..... 432
4. Results ..... 436
5. Final remarks ..... 446

**Keywords**
Real effective exchange rate, cointegration, global VAR

**JEL Codes**
C52, F31, F37

**Abstract · Resumo**

Taxas de câmbio é um preço macroeconômico importante e as mudanças nessa taxa afetam a atividade econômica, o nível de preços, as taxas de juros e os fluxos comerciais. Metodologias foram desenvolvidas em estudos empíricos de desalinhamento cambial para avaliar se a taxa de câmbio real efetiva real está sobre ou sub valorizada. Existe uma literatura sobre os determinantes das taxas de câmbio reais de longo prazo e sobre estratégias empíricas para implementar as normas de equilíbrio obtidas a partir de modelos teóricos. A contribuição deste é mostrar que o modelo de global vetorial autoregressivo (GVAR) proposto por Pesaran, Schuermann, and Weiner (2004) pode agregar informações relevantes à mensuração do desalinhamento da taxa de câmbio. O nosso exercício empírico sugere que o desalinhamento da taxa de câmbio estimada obtido a partir do GVAR pode ser bastante diferente daquele que utiliza as técnicas tradicionais de séries temporais, que tratam os países como unidades independentes. As diferenças entre as duas abordagens são mais pronunciadas para os países pequenos e em desenvolvimento. Nossos resultados também sugerem uma forte interdependência entre os países da Zona do Euro.

## 1. Introduction

The exchange rate is an important macroeconomic price and changes in these rates affect economic activity, prices, interest rates, and trade flows. Large changes in an exchange rate always generate debate on whether the movements are "excessive", reflect "fundamentals", or are "rational". Empirical studies have developed models to assess the long-term determinants of real exchange rates. Empirical strategies are then formulated based on these models, using the doctrine of purchasing power parity (PPP), or based on a fundamentals analysis.

*The authors acknowledge the comments from the participants of the first International Annual Applied Econometric Conference, fourth Emerging Market Group Conference, fourteenth OxMetrics User Conference. The authors also acknowledge Joseph Gagnon, Luis Catão, Irineu Evangelista, Pedro Valls, Megumi Kubota, Cesar Calderon, Aluisio de Lima Campos and Vera Thorstensen for helpful comments on a very preliminary version of this work. All remaining errors are the responsibility of the authors.

[†]Fundação Getulio Vargas, Escola de Economia de São Paulo, Centro de Macroeconomia Aplicada (FGV/EESP/CEMAP). Rua Itapeva 286, 10º andar, São Paulo, SP, CEP 01332-000, Brasil. This author acknowledges partial financial support received from CNPQ (Grant number 311044/2013-1) and from WTO Chair Project.

[‡]World Bank, DIME. Setor Comercial Norte, Quadra 2, Lote A, Edifício Corporate Financial Center, 7° andar, Brasília, DF, CEP 70712-900, Brasil.

[§]Universidade Federal de São Paulo, Escola Paulista de Política, Economia e Negócios (Unifesp/EPPEN). Address: Rua Angélica, 100, Jd. das flores, Osasco, SP, 06110-295, Brazil.

[¶]PhD in Economics at Escola de Economia de São Paulo. Dorfstrasse 9, 8800, Thalwil, Switzerland.

emerson.marcal@fgv.br beatrice_aline@hotmail.com dioted@gmail.com giovanni.merlin@fgv.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Received on Jul 4th, 2017
Final Acceptance on Mar 6th, 2018

Many studies have attempted to construct more accurate estimates of the magnitude and sign of exchange rate misalignment. Exchange rate misalignment is defined as the difference between a measure of the real exchange rate and some equilibrium norm. Discussions on exchange rate misalignment can be divided into two levels. The first focuses on which is the best norm to use to evaluate exchange rate equilibrium. Economic models give a better understanding of the determinants of the real exchange rate. These models attempt to determine the best set of fundamentals that explain real effective exchange rates in the long run. The second level of debate revolves around the best empirical strategy to measure exchange rate equilibrium norms. This is an econometric debate.

Empirical studies also need to choose between a time series or panel approach. The time series approach has the advantage of allowing a particular structure to be estimated for each country. However, the approach does not allow a broader set of variables to be analyzed at the same time because the available macroeconomic samples are not long enough. Panel techniques allow analysts to enlarge the spectrum of variables, but at the cost of imposing untested similarities between the parameters of different countries' models. Hossfeld (2010) reviews exchange rate misalignment literature, and evaluates the benefits and limits of the time series and panel approaches.

This study seeks to contribute to the current body of literature by showing that the global vector autoregressions model (GVAR) proposed by Pesaran, Schuermann, and Weiner (2004) can be used to model the interdependence between countries. In addition, the model can add relevant information to the literature on measuring exchange rate misalignment. As far as the authors are aware, this approach has not been applied to exchange rate misalignment estimation.

This paper is divided into five sections. The first is this introduction. Section 2 provides a brief review of current literature on exchange rate misalignment determinants and describes the challenges faced by empirical studies in trying to determine whether a country's exchange rate is overvalued or undervalued. Section 3 presents the global vector autoregressive model and explains how to adapt Gonzalo and Granger's methodology to this framework. Section 4 describes the results of an empirical exercise that models real effective exchange rates for a selected group of countries. Here, we also present a comparative analysis of the traditional time series approach and the GVAR approach to exchange rate misalignment. Our results suggest that the GVAR approach is worth considering. section 5 applies the limits and the merits of the GVAR approach to the exchange rate misalignment problem and suggests possible extensions to our work. This section also concludes the paper.

## 2. A short review of exchange rate misalignment literature

The literature on real exchange rates is extensive (Froot & Rogoff, 1995). The classical doctrine, and perhaps the oldest one on real exchange rate determinants is that of purchasing power parity (PPP). Reference to this theory can be found in classic studies. Recent studies confirm the validity of PPP for tradable goods, although the adjustment towards equilibrium is quite slow. Ahmad and Craighead (2011) obtained strong evidence of a mean reversion with a high half-life using a secular consumer price index dataset for the United States and United Kingdom. Their work investigates the point made by Taylor (2001) on the effects of temporal aggregation on PPP tests.

There is also much theoretical discussion on which variables drive the real exchange rate in the long term. Older studies include those of Edwards (1989, 1991), who analyzes

the causes and consequences of exchange rate misalignment, and Dornbusch (1976), who developed the classic flexible exchange rate model approach under which monetary policy shocks cause deviations from PPP fundamentals.

The studies of Bilson (1979) and Mussa (1976) are also classics. These are key references for the monetary approach to exchange rates. Under this approach, the exchange rate would be primarily driven by two fundamentals: the difference between domestic and foreign income, and the money supply. The approach assumes that PPP and uncovered interest parity (UIP) hold continuously, and that the demand for money is stable in all countries. However, the research by Meese and Rogoff (1983) casts doubt on the explanatory power of this theory by showing that the predictions of this approach are not superior to a "naive" forecast model for exchange rates, such as a pure random walk. Rossi (2013) shows that the random walk can be outperformed by an econometric model that uses information based on net foreign investment position. "Predictability is most apparent when one or more of the following hold: the predictors are Taylor rule and net foreign assets fundamentals; the model is linear; and a small number of parameters are estimated" Rossi (2013).

Stein (1995) formulated the natural exchange rate approach (NATREX). According to the author, the equilibrium exchange rate is one that is equal to the level of investment and savings generated by economic fundamentals.

Williamson (1994) had a significant impact on exchange rate misalignment theory. Here, the equilibrium exchange rate is the one that allows a country to sustain a desirable result in its external accounts. This is referred to as the fundamental real exchange rate approach (FRER). A more recent reference to this approach is that of Cline and Williamson (2008). A limitation of this approach is that choosing the target of foreign accounts is highly arbitrary and subjective. As a result, the results may not be robust to different targets. In addition, this approach focuses on flows, not stocks.

Faruqee (1995) incorporates the evolution of stocks and constructed a model that allows flows and stocks to interact. Thus, there must be a stable relationship between the real exchange rate and the net foreign asset position between residents and non-residents. This is referred to as the behavioral real exchange rate (BRER) approach. The model was subsequently extended by Alberola, Cervero, Lopez, and Ubide (1999).

Kubota (2009) includes a representative agent who maximizes intertemporal consumption and accumulates capital. This study indicates that the real exchange rate is a function of terms of trade, net external position, and the relative productivity of the tradable and non-tradable sectors. This approach seeks to reduce the degree of subjectivity when estimating exchange rate misalignment. To this end, she establishes a link between the real exchange rate and a set of fundamentals derived from a theoretical model. She then decomposes the series of real exchange rates into transitory and permanent components using the time series econometric technique.

Recently, the International Monetary Fund (IMF) began to systematically disseminate its research efforts into measuring the exchange rate misalignment in several of its member countries (IMF, 2013). Two documents were recently released. These works are an important advance towards transparency. The codes and dataset used to calculate the exchange rate misalignment are available on the IMF website, and the results are easy to replicate. The methodology is also a step forward in incorporating the role of policy gaps in exchange rate misalignment estimates.

The External Balance Assessment (EBA) methodology, developed by the IMF's research department, is based on two panel estimations: one for the current account and one for the

real effective exchange rate (REER) indices.[1] The basic idea is that the REER can be written as a function of the output gap, real interest rate differential, and factors that may affect saving, investment, current account, capital flows, and changes in foreign currency reserves.

The explanatory variables included in the EBA model are the commodity terms of trade, trade openness, share of administered prices, VIX,[2] the share of own currency in world reserves, financial home bias, population growth, expected GDP growth over the next five years, productivity, and changes in foreign reserves. The following policy-related regressors are also included: health expenditure to GDP, foreign exchange interventions, real short-term interest rate differential, private credit to GDP, and capital controls. Most of the variables described are relative to the country's trade partners. They use the same weights as the REER calculation and/or interact with capital account openness. In addition, some variables are lagged to control for endogeneity. The sample data covers 40 countries over the period 1990–2010. The model includes countries fixed effects. To guarantee multilateral consistency in the results, the exchange rate misalignment must be adjusted.

Given the results of the estimation, the "Total REER Gap" can be calculated as the sum of the regression residual and the "Total Policy Gap". The policy gap is a measure of a cyclical gap (over a benchmark) in six policy areas: fiscal balance, capital controls, social spending, foreign exchange market intervention, financial policies, and monetary policy. The gap is calculated as the difference between the actual level of the variable and its "desirable" level, multiplied by the value of the estimated coefficient. The "desirable" levels are supplied by the desk of each IMF country.

## 3. Methodologies to calculate exchange rate misalignment

### 3.1 Traditional time series approach

The analysis starts with an estimation of a vector error correction model (VECM), as suggested by Johansen (1988, 1995). The model is given by equation (1):

$$\Delta x_{it} = \alpha_i \beta_i' x_{i,t-1} + \Gamma_{i1} \Delta x_{i,t-1} + \cdots + \Gamma_{i,k-1} \Delta x_{i,t-k+1} + \Phi_i D_t + \varepsilon_{it}, \tag{1}$$

where $\varepsilon_{it}$ are not correlated random errors. The vector $x_{it}$ contains the variables for the real exchange rate and the fundamentals (e.g., net foreign investment position, etc.) and has dimension $p$; $D_t$ contains deterministic terms; and $\theta = \left\{\alpha_i, \beta_i, \Gamma_{i1}, \ldots, \Gamma_{i,k-1}, \Phi_i\right\}$ is the set of parameters to be estimated.

#### The Gonzalo and Granger decomposition

Several decompositions have been proposed to decompose the series into transitory and permanent components. The decomposition of Gonzalo and Granger (1995) is widely used in exchange rate misalignment empirical literature.[3] In their decomposition, the transitory components do not cause changes in the permanent component in the long term. In other words, exchange rate misalignment — defined as the transitory component of the real exchange rate in a multivariate equation system — does not contain relevant information for predicting the changes of the permanent components in the long term.

[1] A full description of the methodology, data, and routines are available at http://www.imf.org/external/np/spr/2013/esr/

[2] Chicago Board Options Exchange Market Volatility Index.

[3] For example, Alberola et al. (1999) and Kubota (2009).

Using the parameters from (1), it is possible to calculate the transitory component ($T_{it}$) and the permanent component ($P_{it}$) from the following equations:

$$P_{it} = \beta_i \left(\alpha'_{i\perp}\beta_{i\perp}\right)^{-1} \alpha'_{i\perp} x_{it}, \tag{2}$$

$$T_{it} = \alpha_i \left(\beta'_i \alpha_i\right)^{-1} \beta'_i x_{it}. \tag{3}$$

The estimate of exchange rate misalignment is the component associated with the position of the real exchange rate in vector $x_{it}$. Assuming that the real exchange rate is in the first position of the vector, and using the value of the error correction mechanism centered on their own means, one can calculate the misalignment using the following equation:

$$mis_{it} \equiv \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix} \alpha_i \left(\beta'_i \alpha_i\right)^{-1} \left(\beta'_i x_{it} - \mathbb{E}\left[\beta'_i x_{it}\right]\right). \tag{4}$$

## 3.2 Motivation for a global model

The severity of the U.S. economic crisis in 2008 brought the fear of a strong negative contagion to the rest of the world. The U.S. authorities have subsequently adopted an aggressive monetary policy with a strong reduction in nominal interest rates and monetary expansion, among other measures. Some analysts may argue that this policy could have generated strong pressure to depreciate the U.S. dollar against currencies whose domestic interest rates did not follow the same movement. Countries that did not follow such a reduction and opted to accumulate reserves to prevent the appreciation of their currency could have faced inflationary pressures. Some authors argue that the United States was using its monetary policy to depreciate its currency, thereby fostering aggregate demand to reduce the intensity and duration of the economic slowdown. This policy may have generated repercussions around the world. There is much discussion about the extension of these effects and whether they are deleterious.

We construct a global model to assess the magnitude of effects, similar to those discussed in the previous paragraph. In this context, the GVAR appears to be an interesting option, as the relevance and magnitude of global factors, vis-a-vis domestic components, can be explicitly and properly evaluated and tested.

### 3.2.1 GVAR model

In this study, we apply the GVAR methodology to ascertain whether there is any external factor affecting the real exchange rate in the long or short run for a group of selected countries. In this sense, the measure of exchange rate misalignment may have two components. The first is related to domestic fundamentals and the second to global factors. The GVAR explains the source of external influences on the domestic economy by including external variables in VARX.[4] External variables are usually assumed to be weakly exogenous for each country, as defined in Engle, Hendry, and Richard (1983) and Hendry (1995).

In general, the GVAR can be described as a two-step approach. In the first step, a specific model for each country is estimated using the variables of the country and the trading partners weighted average of external variables. Then, all individual models are stacked and grouped into a system of equations, which are solved. Once this is done, the model provides options for different types of analyses, such as the forecast evaluation and impulse response

[4]VARX is the vector autoregression (VAR) model that contains exogenous variables.

analysis. Overall, there is a set of individual models represented as VARX that are combined to obtain the GVAR.

Following the notation of Pesaran et al. (2004), we restrict our discussion to the specification with first-order dynamics, as represented by VAR(1, 1). We consider a set of $N$ countries and the VAR(1, 1) below:

$$x_{it} = a_{i,0} + a_{i,1}t + \Phi_i x_{i,t-1} + \Lambda_{i,0}x^*_{i,t} + \Lambda_{i,1}x^*_{i,t-1} + \varepsilon_{it}, \tag{5}$$

where $x_{it}$ is a vector of $k_i \times 1$ specific variables for each country; $x^*_{i,t}$ is a vector of $k^*_i \times 1$ foreign variables, $i = 1, 2, \ldots, N$, $t = 1, 2, \ldots, T$; $\Lambda_{i,0}$ and $\Lambda_{i,1}$ are respectively matrices with parameters of the contemporaneous and lagged terms; $a_{i,0}$ is a vector containing the constant; and $a_{i,1}$ is the coefficient associated with the time trend. The term $\varepsilon_{it}$ is a vector of idiosyncratic shocks for each country.

We construct the external variables using the trade weights, $w$,

$$x^*_{it} = \sum_{j=1}^{N} w_{i,j}x_{jt}, \tag{6}$$

where $\sum_{j=1}^{N} w_{ij} = 1$ with $j = 1, \ldots, N$ and $w_{ii} = 0$ with $i = 1, \ldots, N$. From (5), we can see that the domestic variable, $x_{it}$, depends on the external variable $x^*_{it}$. The system from equation (5) needs to be solved for all domestic variables.

We can obtain the global VAR defining $z_{it} = \begin{pmatrix} x_{it} \\ x^*_{it} \end{pmatrix}$ and we rewrite (5) as

$$A_i z_{it} = a_{i0} + a_{i1}t + B_i z_{i,t-1} + \varepsilon_{it}, \tag{7}$$

where $A_i = \left(I_{ki}, -\Lambda_{i0}\right)$, and $B_i = \left(\Phi_i, \Lambda_{i1}\right)$. The terms $A_i$ and $B_i$ have dimension $k_i \times \left(k_i + k^*_i\right)$, and $A_i$ has full column rank, $k_i$.

We can define the external variables as

$$z_{it} = W_i z_t, \tag{8}$$

where $W_i$ is a weight matrix with dimension $\left(k_i + k^*_i\right) \times k$, where $k = \sum_{i=1}^{N} k_i$. The matrix $W_i$ reflects the relationship between countries and allows us to unify the model into a complete global model. Using (8) and (7), we obtain

$$A_i W_i z_t = a_{i0} + a_{i1}t + B_i W_i z_{t-1} + \varepsilon_{it}, \tag{9}$$

where $A_i W_i$ and $B_i W_i$ have dimension $k_i \times k$. Finally, the stacked equations can be written as a GVAR(1):

$$G z_t = a_0 + a_1 t + H z_{t-1} + \varepsilon_t, \tag{10}$$

where $a_o = \begin{bmatrix} a_{10} \\ a_{20} \\ \vdots \\ a_{N0} \end{bmatrix}$, $a_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{N1} \end{bmatrix}$, $\varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}$, $G = \begin{bmatrix} A_1 W_1 \\ A_2 W_2 \\ \vdots \\ A_N W_N \end{bmatrix}$, $H = \begin{bmatrix} B_1 W_1 \\ B_2 W_2 \\ \vdots \\ B_N W_N \end{bmatrix}$.

Assuming that $G$ is not singular and has dimension $k \times k$, the reduced form of (10) can be rewritten as:

$$z_t = b_0 + b_1 t + F z_{t-1} + v_t, \tag{11}$$

where $F = G^{-1}H$, $b_0 = G^{-1}a_0$, $b_1 = G^{-1}a_1$ and $v_t = G^{-1}\varepsilon_t$.

After estimating the models for each country separately from (5), we can solve the global model in (11) to obtain recursively the future values of all endogenous variables ($z_t$).

### 3.2.2 The GVAR and Gonzalo and Granger decomposition

For the remainder of the paper, we will rewrite the GVAR as the global vector error correction model (GVECM). Both are equivalent, but the GVECM allows us to deal with permanent and transitory decompositions.

Assuming the following GVECM and stacking the models, we obtain that

$$\Delta X_t = AB' Z_{t-1} + \tilde{\Gamma}_1 \Delta Z_{t-1} + \cdots + \tilde{\Gamma}_{k-1} \Delta Z_{t-k+1} + \tilde{\Phi} D_t + \tilde{\Gamma}_{0,1} \Delta \bar{X}_t + \tilde{\varepsilon}_t, \tag{12}$$

where

$$A = \begin{bmatrix} \tilde{\alpha}_1 & 0 & \cdots & 0 & 0 \\ 0 & \tilde{\alpha}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \tilde{\alpha}_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & \tilde{\alpha}_N \end{bmatrix}; \qquad B = \begin{bmatrix} \tilde{\beta}_1 & 0 & \cdots & 0 & 0 \\ 0 & \tilde{\beta}_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \tilde{\beta}_{N-1} & 0 \\ 0 & 0 & \cdots & 0 & \tilde{\beta}_N \end{bmatrix};$$

$\tilde{\varepsilon}_t$ are random errors, not time correlated;

$X_t' = \begin{bmatrix} X_{1,t}' & X_{2,t}' & \cdots & X_{N-1,t}' & X_{N,t}' \end{bmatrix}'$;

$Z_{it-1}' = \begin{bmatrix} x_{i,t-1}^1 & \bar{x}_{i,t-1}^1 & \cdots & \cdots & x_{i,t-1}^p & \bar{x}_{i,t-1}^p \end{bmatrix}'$;

and $\bar{x}_{i,t-1}^j$ is the average of variable $x^j$ for country $i$.

Using $Z_t = WX_t$, we can now write the Global VECM as

$$\Delta X_t = AB' WX_{t-1} + \tilde{\Gamma}_1 \Delta Z_{t-1} + \cdots + \tilde{\Gamma}_{k-1} \Delta Z_{t-k+1} + \tilde{\Phi} D_t + \tilde{\Gamma}_{0,1} \Delta \bar{X}_t + \tilde{\varepsilon}_t. \tag{13}$$

Defining $\Gamma_{0,1}^* W \Delta X_t \equiv \tilde{\Gamma}_{0,1} \Delta \bar{X}_t$, and after some algebra, we obtain equation (14):

$$\left[ I - \Gamma_{0,1}^* W \right] \Delta X_t = AB' WX_{t-1} + \tilde{\Gamma}_1 \Delta Z_{t-1} + \cdots + \tilde{\Gamma}_{k-1} \Delta Z_{t-k+1} + \tilde{\Phi} D_t + \tilde{\varepsilon}_t. \tag{14}$$

Now, assume that we can calculate the inverse of matrix $\left[ I - \Gamma_{0,1}^* W \right]$, we define

$$A^* = \left[ I - \Gamma_{0,1}^* W \right]^{-1} A,$$

$$\tilde{\Gamma}_l^* = \left[ I - \Gamma_{0,1}^* W \right]^{-1} \tilde{\Gamma}_l, \qquad l = 1, \ldots, k-1,$$

and

$$\tilde{\varepsilon}_t^* = \left[ I - \Gamma_{0,1}^* W \right]^{-1} \tilde{\varepsilon}_t.$$

Then, we can solve the global model, yielding the solution to the global VECM, as shown in (15):

$$\Delta X_t = A^* B' WX_{t-1} + \tilde{\Gamma}_1^* \Delta Z_{t-1} + \cdots + \tilde{\Gamma}_{k-1}^* \Delta Z_{t-k+1} + \tilde{\Phi}^* D_t + \tilde{\varepsilon}_t^*. \tag{15}$$

So we obtain the transitory component as

$$T_t^{\text{GVAR}} = A^* \left( B' WA^* \right)^{-1} B' WX_t - \mathbb{E}\left[ A^* \left( B' WA^* \right)^{-1} B' WX_t \right]. \tag{16}$$

The permanent component is defined as the difference between the actual values of the series and the transitory component given in (16). The matrix given by (17) contains the weights that each cointegrated relationship will contribute to the transitory component:

$$LF^{\text{GVAR}} = A^* \left( B' WA^* \right)^{-1}. \tag{17}$$

The exchange rate misalignment can be calculated for country $i$ by picking the country's real exchange rate in vector $X_t$,

$$mis_{i,t}^{\text{GVAR}} \equiv \begin{bmatrix} 0 & \cdots & 0_{p(i-1)} & 1 & 0 & \ldots & 0 \end{bmatrix} T_t^{\text{GVAR}}. \tag{18}$$

In the following section, we estimate the exchange rate misalignment of equations (4) and (18), and we compare the misalignment measures.

# 4. Results

## 4.1 Dataset

The database of this study is annual and covers the period from 1970 to 2012. The foreign trade figures were collected from the IMF Direction of Trade Statistics (DOTS-IMF). We use these weights to calculate the external variables in the GVAR model. We collect the real effective exchange rate from the IMF's International Financial Statistics (IFS-IMF). We obtain the values of net foreign assets from Lane and Milesi-Ferretti (2007) and the IFS-IMF. The full sample consists of 33 countries, but we only analyze countries with data for all series and years during 1970 to 2012. This restricts the sample to 27 countries. Finally, we opt to work with end-of-period figures to avoid problems caused by data aggregation. In certain contexts, the temporal aggregation can cause significant distortions (Ghysels & Miller, 2015; Taylor, 2001).

## 4.2 Is there evidence of cointegration between real exchange rates and fundamentals?

To assess the existence of cointegration between variables, we run the following cointegration tests: Engle and Granger (1987), Shin (1994), and Johansen (1988). The first two tests are univariate whereas the third is multivariate. The null hypothesis of the first test is absence of cointegration and alternative hypothesis is existence of cointegration. Second test is a confirmatory procedure whre the null hypothesis is existence of cointegration and alternative hypothesis is no cointegration. Trace and maximum eigenvalue tests proposed by Johansen if and how many cointegration relationships may exist.

Table 1 shows the results of the Engle and Granger (1987) and Shin (1994) cointegration tests, while Table 2 shows the results of the Johansen test for the sample countries. The null hypothesis of no cointegration is not rejected for many countries when looking at the cointegration of Engle and Granger or Johansen's tests. However, this conclusion is not corroborated by Shin's test. We do not reject the null hypothesis of cointegration at the 1% level of significance for most countries when both domestic and external variables are used. The rejection of the null hypothesis of cointegration is more frequent when we investigate only the relationship between real exchange rate and domestic fundamentals. These results suggests that external factors may explain real exchange rates in the long term for some countries. Since no formal rejection of the null hypothesis of cointegration occurs at the 1% level when all variables are used, we choose to work with the hypothesis that there exists one cointegrated relationship between the variables in specific countries' models. Another point to justify this choice is the cost of keeping irrelevant variable in econometric model compared to exclude a relevant variable. Omitting a relevant variable can induce bias in estimates of remaining parameters.

**Table 1.** Results of univariate cointegration test.

| | Engle and Granger test | | | | Shin cointegration test | | |
|---|---|---|---|---|---|---|---|
| Country\Lag | 0 | 1 | 2 | N. of regressors | 1 | 2 | 3 |
| Australia | −1.33 | −1.38 | −3.27 | | 0.276 | 0.149 | 0.131 |
| Austria | −2.29 | −2.99 | −3.40 | | 0.287 | 0.167 | **0.161** |
| Belgium | −2.97 | **−3.86** | **−3.84** | | **0.925** | **0.510** | **0.380** |
| Brazil | −2.41 | −2.51 | −3.06 | | 0.300 | 0.069 | 0.070 |
| Canada | −2.23 | −2.16 | −2.63 | | 0.171 | 0.178 | 0.055 |
| China | −1.57 | −1.61 | −2.45 | | **0.367** | **0.372** | **0.199** |
| Colombia | **−3.35** | −2.51 | −2.50 | | **0.821** | **0.814** | 0.092 |
| Denmark | **−3.49** | **−3.58** | −3.55 | | 0.163 | **0.306** | **0.289** |
| Finland | −1.53 | −1.92 | −2.01 | | 0.054 | 0.066 | 0.066 |
| France | −2.05 | −3.16 | −3.10 | | **0.576** | **0.290** | **0.272** |
| Germany | −2.11 | −2.66 | −2.68 | | **0.788** | **0.312** | **0.355** |
| Greece | −2.29 | −2.23 | −2.33 | | **0.350** | 0.087 | 0.088 |
| India | −1.26 | −1.86 | −1.84 | | 0.202 | **0.237** | 0.146 |
| Indonesia | −1.28 | −1.64 | −1.93 | | **0.344** | 0.059 | 0.053 |
| Ireland | −1.91 | −2.90 | −3.09 | | **1.441** | **0.703** | **0.619** |
| Italy | **−3.37** | **−3.48** | **−3.94** | | **0.929** | **0.472** | 0.101 |
| Japan | −2.33 | **−3.63** | −3.55 | | 0.082 | 0.108 | 0.074 |
| Mexico | −2.74 | −2.74 | −2.97 | | **0.417** | 0.179 | 0.068 |
| Netherlands | −2.78 | −3.28 | −3.80 | | 0.307 | **0.349** | 0.089 |
| Singapore | −1.98 | −2.96 | −2.98 | | 0.128 | **0.392** | 0.126 |
| South Africa | −2.96 | −2.35 | −2.59 | | **0.365** | **0.389** | **0.384** |
| Spain | **−3.08** | −3.29 | −3.27 | | 0.244 | 0.170 | 0.070 |
| Sweden | −1.31 | −2.28 | −2.28 | | 0.167 | 0.093 | 0.085 |
| United Kingdom | −2.87 | −2.85 | −2.58 | | **1.192** | **0.234** | **0.245** |
| United States | −3.04 | −3.04 | −3.31 | | 0.200 | **0.223** | **0.164** |
| Uruguay | −2.48 | −2.50 | −2.46 | | 0.116 | 0.172 | 0.148 |
| Turkey | −1.43 | −1.97 | −1.96 | | **0.390** | 0.159 | 0.138 |
| *Critical Values* | | | | *Critical Values* | | | |
| 10% | −3.07 | −3.45 | −3.83 | | | | |
| 5% | −3.37 | −3.77 | −4.11 | 5% | 0.314 | 0.221 | 0.159 |
| 1% | −3.96 | −4.31 | −4.73 | 1% | 0.533 | 0.380 | 0.271 |

**Table 2.** Results of multivariate cointegration test.

| | Countries Trace Statistics: Null Hypothesis | | |
|---|---|---|---|
| | N. of vectors | Absence | One vector |
| Australia | 1 | 40.21 | 10.14 |
| Austria | 1 | 30.01 | 12.93 |
| Belgium | 0 | 15.88 | 3.21 |
| Brazil | 1 | 35.69 | 12.60 |
| Canada | 0 | 11.20 | 4.58 |
| China | 1 | 39.88 | 8.38 |
| Colombia | 0 | 26.91 | 12.70 |
| Denmark | 0 | 24.09 | 5.30 |
| Finland | 0 | 14.84 | 5.42 |
| France | 0 | 21.10 | 2.48 |
| Germany | 0 | 14.94 | 2.64 |
| Greece | 0 | 17.70 | 5.46 |
| India | 1 | 28.66 | 5.90 |
| Indonesia | 0 | 25.60 | 3.30 |
| Ireland | 0 | 15.78 | 2.99 |
| Italy | 0 | 18.21 | 5.73 |
| Japan | 0 | 26.89 | 6.55 |
| Mexico | 0 | 27.25 | 5.93 |
| Netherlands | 0 | 24.22 | 7.22 |
| Singapore | 0 | 17.37 | 6.40 |
| South Africa | 1 | 32.31 | 10.03 |
| Spain | 0 | 23.35 | 9.63 |
| Sweden | 0 | 22.55 | 6.32 |
| United Kingdom | 0 | 21.30 | 6.17 |
| United States | 0 | 23.25 | 8.83 |
| Uruguay | 0 | 16.05 | 5.16 |
| Turkey | 1 | 28.11 | 6.66 |

The specification may be improved by adding other variables. For example, a variable that controls the possible Balassa–Samuelson effect may alter the results towards finding stronger evidence of cointegration.[5]

## 4.3 Is there evidence of global effects?

This section attempts to answer the question of whether the model with external factors is better than the model without these factors. We compare eight different specifications. We consider models with complete interdependence, in other words, that have external factors, in the short and long term dynamics, similar to the equation (12). There are models in which interdependence is allowed only in the long term. In another specification, we allow interdependence only on short-term dynamics. Finally, there are models in which no interdependence is allowed. We estimate the models also allowing a structure with or without

[5]In this study, we could have analyzed a broader set of information using variables to control for the Balassa–Samuelson effect, similar to Kubota (2009) and Alberola et al. (1999). However, the number of countries in the sample would have been further reduced. We opt to explore a longer sample with a wider number of countries rather than a restricted sample with more variables. The inclusion of a variable to control for the Balassa–Samuelson effect reduces the sample in both the temporal and cross-sectional dimensions.

common cycles.[6] We compare these eight models using the Schwarz (SC), Hannan-Quinn (HQ), and Akaike information criteria (AIC).

Table 3 presents a detailed description of the models and all information criteria. Options 3 and 4 consist of models with no interdependence. The evidence in favor of interdependence varies between countries. According to all information criteria, there is no evidence of interdependence for the following countries: Australia, Ireland, India, Netherlands, and Turkey. For other countries, there is evidence of interdependence in the short and/or long term. Thus, we can conclude that statistical evidence corroborates the hypothesis of interdependence for a large group of countries. The next step in the analysis is to assess the relevance of global and domestic factors to exchange rate misalignment estimate.

## 4.4 Calculating exchange rate misalignment using the GVECM

This section describes the results of the GVECM estimation. Table 6 shows the estimated cointegrated vectors for each country. Table 4 shows the results of the estimate loading factor given by (17) for real exchange rate equation. The value in line $i$ and column $j$ represents the weight of the error correction mechanism of country $j$ that will be used to calculate the misalignment for country $i$. For example, we can check that the United States and Germany columns contain many non-zero terms. This suggests strong linkages between these economies and others economies analyzed in the sample. In the case of Germany, there seems to be a strong effect in eurozone countries. Brazil is an example of the opposite case. Here, the Brazilian exchange rate misalignment causes minor effects on all countries other than Uruguay. Although Brazil is a large economy, its global share is relatively small. Intuitively, the Brazilian economy is affected by others countries' disequilibrium, but its own disequilibrium does not affect others countries. The United States exchange rate misalignment may generate quite small effects on eurozone countries.

Table 5 compares the results of the exchange rate misalignment using the traditional and GVAR methodologies. In general, the estimate misalignment tends to have the same sign for almost 67% of the sample. There are 1,161 ($27 \times 43$) estimate of exchange rate misalignments, across all countries and periods. For the United States, the results are virtually the same in terms of sign and magnitude. The overall picture does not change when the comparison is made using the magnitude rather than the sign of the exchange rate misalignment. We compute the proportion of each case out of the total, where the estimate misalignments for both models have the same sign and absolute value above 10%, or different signs but absolute an below 10%. In the 57% of cases, these criteria were satisfied. However, in about 43% of cases, the estimates are not the same. The results in Table 5 suggest that quite different results can be obtained from the GVAR. The dynamics of real exchange rates in these countries cannot be seen as detached from the rest of world, or at least from their main trading partners. Table 6 shows the estimates of all parameters necessary to solve the GVAR.

Table 7 provides information about the source of the differences between the two methodologies. First, the coefficient of the NFA variable changes quite significantly for many countries. This may explain part of the change in the results and highlights the importance of investigating the main drivers of real exchange rates in the long run. However, there are countries where the coefficient hardly changes at all, but the external variables cause

[6] See Hecq, Palm, and Urbain (2000) and Hecq, Palm, and Urbain (2002) for a common cycle definition, a discussion, and its relationship to permanent and transitory decomposition.

**Table 3.** Results of the tests for interdependence.

| | | External Variables in the Model | | |
|---|---|---|---|---|
| **Models** | Common Features | *Short Run* | *Long Run* | Interdependence |
| 1 | No | Yes | Yes | Yes, Both types |
| 2 | Yes | Yes | Yes | Yes, Both types |
| 3 | No | No | No | No |
| 4 | Yes | No | No | No |
| 5 | No | Yes | No | Yes, Short Run |
| 6 | Yes | Yes | No | Yes, Short Run |
| 7 | No | No | Yes | Yes, Long Run |
| 8 | Yes | No | Yes | Yes, Long Run |

| | Best Model – Information Criteria | | |
|---|---|---|---|
| **Country** | SC | HQ | AIC |
| Australia | 8 | 8 | 2 |
| Austria | 6 | 6 | 6 |
| Belgium | 6 | 6 | 1 |
| Brazil | 8 | 8 | 8 |
| Canada | 4 | 4 | 5 |
| China | 8 | 8 | 8 |
| Colombia | 6 | 6 | 2 |
| Denmark | 4 | 4 | 2 |
| Finland | 3 | 3 | 2 |
| France | 8 | 1 | 1 |
| Germany | 4 | 4 | 4 |
| Greece | 4 | 4 | 4 |
| India | 8 | 8 | 8 |
| Ireland | 6 | 6 | 6 |
| Italy | 8 | 8 | 8 |
| Japan | 6 | 1 | 1 |
| Mexico | 4 | 8 | 8 |
| Netherlands | 2 | 2 | 2 |
| Singapore | 4 | 4 | 5 |
| South Africa | 8 | 2 | 2 |
| Spain | 4 | 4 | 2 |
| Sweden | 6 | 2 | 2 |
| United Kingdom | 6 | 5 | 5 |
| United States | 6 | 6 | 5 |
| Uruguay | 4 | 6 | 6 |
| Turkey | 6 | 2 | 2 |

**Table 4.** Loading factor for calculating exchange rate misalignment from the GVAR model.

| | Error Correction Mechanism | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | AUS[a] | AUT | BEL | BRA | CAN | CHN | COL | DEN | FIN | FRA | DEU | GRC | IND | IRL |
| Australia | 0.55 | 0.00 | 0.04 | −0.01 | −0.04 | 0.01 | 0.00 | 0.02 | 0.00 | 0.24 | 0.08 | 0.01 | −0.01 | 0.03 |
| Austria | 0.00 | 0.13 | −0.01 | 0.00 | 0.01 | 0.00 | 0.00 | −0.04 | 0.00 | −0.17 | 0.34 | 0.01 | 0.00 | 0.01 |
| Belgium | 0.01 | 0.00 | 0.97 | 0.00 | 0.01 | 0.00 | 0.00 | −0.09 | 0.00 | −0.32 | 0.31 | 0.01 | 0.02 | 0.04 |
| Brazil | 0.02 | −0.01 | −0.16 | 1.26 | 0.24 | −0.04 | 0.02 | −0.18 | 0.00 | −1.94 | −0.02 | 0.00 | 0.03 | 0.09 |
| Canada | 0.00 | 0.00 | 0.00 | 0.00 | 1.02 | 0.00 | 0.00 | 0.00 | 0.00 | −0.03 | −0.01 | 0.00 | 0.00 | 0.00 |
| China | −0.05 | 0.00 | 0.00 | −0.06 | −0.07 | −0.22 | 0.00 | −0.02 | 0.00 | 0.34 | −0.09 | 0.00 | −0.05 | 0.04 |
| Colombia | 0.01 | 0.00 | 0.06 | 0.06 | 0.19 | −0.01 | 0.53 | −0.04 | 0.00 | −0.36 | 0.36 | 0.01 | 0.02 | 0.01 |
| Denmark | 0.00 | 0.03 | 0.25 | 0.02 | 0.05 | −0.01 | 0.00 | 0.92 | −0.03 | 1.46 | 1.02 | 0.08 | 0.01 | 0.00 |
| Finland | 0.01 | 0.04 | 0.47 | 0.03 | 0.08 | −0.01 | 0.00 | −0.37 | 0.00 | 2.83 | 1.11 | 0.12 | 0.02 | 0.07 |
| France | 0.02 | 0.08 | 1.11 | 0.06 | 0.16 | −0.02 | 0.00 | −0.41 | −0.01 | 7.01 | 2.55 | 0.28 | 0.05 | 0.13 |
| Germany | 0.00 | 0.01 | 0.10 | 0.01 | 0.02 | 0.00 | 0.00 | −0.03 | 0.00 | 0.29 | 1.20 | 0.02 | 0.00 | 0.01 |
| Greece | 0.00 | 0.00 | −0.04 | 0.00 | 0.00 | 0.00 | 0.00 | 0.01 | 0.00 | −0.33 | 0.00 | 0.05 | 0.00 | −0.01 |
| India | 0.05 | 0.00 | −0.01 | 0.04 | 0.18 | −0.03 | 0.01 | 0.02 | 0.00 | −1.57 | −0.13 | −0.01 | 1.09 | −0.18 |
| Ireland | −0.04 | −0.02 | −0.40 | −0.03 | −0.09 | 0.01 | 0.00 | −0.10 | 0.01 | −1.72 | −0.89 | −0.07 | −0.04 | 0.83 |
| Italy | −0.02 | −0.02 | −0.87 | −0.07 | −0.11 | 0.02 | 0.00 | 0.23 | 0.01 | −5.04 | −2.04 | −0.05 | −0.04 | 0.01 |
| Japan | −0.02 | 0.00 | −0.03 | −0.02 | −0.14 | 0.04 | 0.00 | 0.04 | 0.00 | −0.08 | −0.12 | −0.02 | −0.01 | −0.01 |
| Mexico | −0.01 | 0.00 | 0.00 | −0.01 | −0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.02 | 0.00 | 0.00 | 0.02 |
| Netherlands | 0.01 | 0.00 | 0.06 | 0.00 | 0.00 | 0.00 | 0.00 | −0.10 | 0.00 | −0.39 | 0.32 | 0.01 | 0.00 | 0.04 |
| Singapore | 0.02 | 0.00 | 0.01 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.03 | 0.02 | 0.00 | 0.01 | 0.00 |
| Spain | 0.00 | −0.07 | −0.94 | −0.06 | −0.11 | 0.02 | −0.01 | 0.34 | 0.01 | −6.16 | −2.10 | −0.21 | −0.03 | 0.03 |
| Sweden | 0.02 | 0.03 | 0.51 | 0.03 | 0.09 | −0.01 | 0.00 | −1.68 | −0.06 | 3.00 | 0.60 | 0.11 | 0.03 | 0.15 |
| UK | −0.05 | −0.03 | −0.43 | −0.03 | −0.15 | 0.01 | 0.00 | 0.18 | 0.01 | −1.30 | −0.99 | −0.09 | −0.05 | −0.27 |
| USA | 0.01 | 0.00 | 0.01 | 0.01 | 0.13 | 0.00 | 0.00 | −0.01 | 0.00 | 0.02 | 0.01 | 0.00 | 0.00 | −0.01 |
| Uruguay | 0.02 | 0.01 | 0.29 | −0.72 | 0.06 | −0.02 | 0.00 | 0.08 | −0.01 | 1.24 | 0.77 | 0.15 | 0.04 | 0.11 |
| South Africa | −0.02 | 0.01 | 0.07 | −0.01 | 0.00 | 0.05 | 0.00 | −0.03 | 0.00 | 1.11 | 0.07 | 0.02 | −0.10 | 0.05 |
| Indonesia | −0.05 | 0.00 | 0.00 | 0.00 | 0.02 | 0.01 | 0.00 | −0.03 | 0.00 | 0.25 | 0.02 | 0.01 | −0.06 | 0.02 |
| Turkey | 0.01 | 0.02 | −0.04 | 0.00 | 0.04 | −0.01 | 0.00 | −0.05 | 0.00 | −0.75 | 0.62 | 0.08 | 0.01 | 0.05 |
| | ITA | JAP | MEX | NLD | SGP | ESP | SWE | GBR | USA | URY | ZAF | IDN | TUR | |
| Australia | 0.04 | 0.02 | −0.02 | 0.02 | 0.05 | 0.02 | 0.00 | −0.03 | −0.08 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Austria | 0.05 | 0.00 | 0.00 | 0.02 | 0.00 | 0.00 | 0.00 | −0.01 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Belgium | −0.05 | 0.01 | 0.00 | 0.30 | 0.00 | 0.00 | 0.00 | −0.05 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Brazil | −0.35 | 0.16 | 0.23 | 0.54 | −0.02 | −0.02 | 0.00 | −0.09 | 0.39 | 0.05 | 0.01 | 0.00 | 0.00 | |
| Canada | 0.00 | 0.02 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.11 | 0.00 | 0.00 | 0.00 | 0.00 | |
| China | −0.06 | −0.96 | −0.05 | −0.03 | 0.06 | 0.01 | 0.00 | −0.03 | −0.29 | 0.00 | −0.01 | 0.01 | 0.00 | |
| Colombia | 0.02 | 0.15 | 0.14 | 0.14 | −0.01 | 0.07 | 0.00 | 0.00 | 0.66 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Denmark | 0.36 | 0.06 | 0.03 | 0.32 | 0.01 | 0.17 | −0.03 | −0.03 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Finland | 0.52 | 0.02 | 0.04 | 0.42 | 0.02 | 0.29 | 0.03 | −0.09 | 0.11 | 0.00 | 0.01 | 0.00 | 0.00 | |
| France | 1.27 | 0.06 | 0.09 | 0.90 | 0.04 | 0.68 | 0.00 | −0.18 | 0.23 | 0.00 | 0.02 | −0.01 | 0.01 | |
| Germany | 0.12 | 0.00 | 0.01 | 0.09 | 0.00 | 0.05 | 0.00 | −0.01 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Greece | 0.09 | 0.00 | 0.00 | −0.02 | 0.00 | −0.03 | 0.00 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | |
| India | −0.04 | 0.87 | 0.06 | 0.03 | 0.18 | −0.13 | 0.01 | 0.12 | 0.57 | 0.00 | 0.02 | −0.02 | 0.00 | |
| Ireland | −0.48 | −0.07 | −0.04 | −0.20 | 0.01 | −0.22 | 0.02 | 0.25 | −0.14 | 0.00 | −0.03 | 0.00 | 0.00 | |
| Italy | 0.16 | −0.05 | −0.08 | −0.68 | −0.03 | −0.55 | 0.00 | 0.04 | −0.20 | 0.00 | −0.02 | 0.00 | 0.00 | |
| Japan | −0.01 | 1.16 | −0.07 | −0.06 | −0.15 | −0.02 | 0.00 | 0.01 | −0.22 | 0.00 | 0.00 | 0.01 | 0.00 | |
| Mexico | 0.00 | −0.01 | 1.42 | 0.03 | 0.04 | 0.00 | 0.00 | −0.02 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Netherlands | −0.07 | 0.00 | 0.00 | 0.87 | 0.00 | −0.01 | 0.00 | −0.04 | 0.01 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Singapore | 0.01 | −0.01 | 0.01 | 0.01 | 1.13 | 0.00 | 0.00 | 0.00 | 0.03 | 0.00 | 0.00 | −0.01 | 0.00 | |
| Spain | −1.49 | −0.04 | −0.10 | −0.70 | −0.05 | 0.39 | −0.01 | 0.02 | −0.17 | −0.01 | −0.01 | 0.00 | 0.00 | |
| Sweden | 0.42 | −0.04 | 0.05 | 0.38 | 0.03 | 0.28 | 0.02 | −0.16 | 0.13 | 0.00 | 0.02 | 0.00 | 0.00 | |
| UK | −0.39 | −0.09 | −0.06 | −0.43 | −0.03 | −0.23 | 0.00 | 0.16 | −0.24 | 0.00 | −0.03 | 0.00 | 0.00 | |
| USA | 0.01 | 0.03 | 0.06 | 0.00 | −0.01 | 0.00 | 0.00 | 0.01 | 0.43 | 0.00 | 0.00 | 0.00 | 0.00 | |
| Uruguay | 0.20 | 0.10 | 0.04 | 0.24 | 0.01 | 0.15 | 0.00 | −0.11 | 0.20 | 0.94 | 0.01 | 0.00 | 0.00 | |
| South Africa | 0.15 | −0.22 | 0.00 | −0.04 | −0.01 | 0.04 | 0.00 | −0.02 | −0.02 | 0.00 | 0.41 | 0.00 | 0.00 | |
| Indonesia | −0.01 | −1.10 | 0.01 | 0.00 | 0.05 | 0.00 | 0.00 | −0.02 | −0.11 | 0.00 | 0.00 | −0.16 | 0.00 | |
| Turkey | −0.01 | 0.01 | 0.01 | 0.06 | 0.01 | 0.00 | 0.00 | −0.06 | 0.07 | 0.00 | 0.00 | 0.00 | 0.02 | |

[a] ISO alpha-3 country codes: AUS (Australia), AUT (Austria), BEL (Belgium), BRA (Brazil), CAN (Canada), CHN (China), COL (Colombia), DEN (Denmark), FIN (Finland), FRA (France), DEU (Germany), GRC (Greece), IND (India), IRL (IRELAND), ITA (Italy), JPN (Japan), MEX (Mexico), NLD (Netherlands), SGP (Singapore), ESP (Spain), SWE (Sweden), GBR (United Kingdom), USA (United States of America), URY (Uruguay), ZAF (South Africa), IDN (Indonesia), TUR (Turkey).

**Table 5.** Comparing exchanges rate misalignment estimates.

| | Percentage of case where: | |
|---|---|---|
| | Sign of misalignment is the same (%) | Magnitude of misalignment is the same[a] (%) |
| Australia | 58.14 | 76.74 |
| Austria | 48.84 | 86.05 |
| Belgium | 65.12 | 81.40 |
| Brazil | 74.42 | 60.47 |
| Canada | 79.07 | 51.16 |
| China | 62.79 | 25.58 |
| Colombia | 67.44 | 76.74 |
| Denmark | 69.77 | 48.84 |
| Finland | 55.81 | 39.53 |
| France | 67.44 | 48.84 |
| Germany | 93.02 | 79.07 |
| Greece | 76.74 | 55.81 |
| India | 60.47 | 44.19 |
| Ireland | 60.47 | 51.16 |
| Italy | 51.16 | 34.88 |
| Japan | 60.47 | 32.56 |
| Mexico | 81.40 | 74.42 |
| Netherlands | 65.12 | 83.72 |
| Singapore | 88.37 | 90.70 |
| Spain | 62.79 | 27.91 |
| Sweden | 74.42 | 46.51 |
| United Kingdom | 62.79 | 41.86 |
| United States | 79.07 | 86.05 |
| Uruguay | 81.40 | 60.47 |
| South Africa | 55.81 | 41.86 |
| Indonesia | 41.86 | 58.14 |
| Turkey | 55.81 | 44.19 |
| Total | 66.67 | 57.36 |

[a] If both estimates give a value of misalignment above 10% in modulus or below 10% in modulus simultaneously.

changes in the magnitude of the exchange misalignment estimate. Brazil is a good example. Although there is a minor change in the NFA coefficient when external variables are included in the model, they add relevant information to the long run level of the exchange rate. The misalignment of the exchange rate for both models is different after introducing foreign variables.

We show exchange rate misalignment estimates from traditional and GVAR methodologies in Figures 1 and 2. The case of Brazil and USA are shown as examples. For the USA, estimates of real exchange rate misalignment by traditional and GVAR methodologies differ in magnitude for most of the sample according to Figure 1. The amplitude of misalignment is lower by GVAR model. The Brazilian estimates of misalignment by traditional and GVAR methodologies are quite similar until 2008, according to Figure 2. In the end of the sample, particularly after the financial crisis of 2008, the results are different for both of the countries.

**Table 6.** Coefficient estimates for each country model used to solve the GVAR.

| | Loading Matrix | | Cointegrated Vectors | | | | | Contemporaneous Effect | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Country | $\alpha_{i1}$ | $\alpha_{i2}$ | $\beta_{i,\text{RER}}$ | $\beta_{i,\text{NFA}}$ | $\beta_{i,\text{RERw}}$ | $\beta_{i,\text{NFAw}}$ | $\beta_{i,\text{Constant}}$ | $C_{i,11}$ | $C_{i,12}$ | $C_{i,21}$ | $C_{i,22}$ |
| Australia | −0.45 | 0.27 | 1 | −1.34 | 0.71 | −2.07 | −8.21 | −0.55 | 0.55 | 0.54 | −0.66 |
| Austria | −0.07 | −0.37 | 1 | 1.31 | 0.23 | −2.25 | −5.43 | 0.90 | −0.10 | −0.04 | 0.06 |
| Belgium | −0.39 | −0.03 | 1 | −0.13 | −0.86 | −0.77 | −0.60 | 1.11 | −0.23 | −0.29 | 0.47 |
| Brazil | −0.89 | −0.18 | 1 | −1.19 | 0.85 | −3.21 | −9.18 | −0.25 | 2.74 | 0.82 | −0.44 |
| Canada | −0.16 | −0.03 | 1 | 0.00 | −0.37 | 0.11 | −2.79 | −0.02 | 1.22 | 0.10 | −0.49 |
| China | 0.06 | 0.04 | 1 | −8.88 | −18.54 | 14.25 | 80.15 | 0.42 | 0.10 | −0.39 | 0.41 |
| Colombia | −0.08 | 0.01 | 1 | −6.45 | 4.47 | 0.48 | −26.80 | 0.39 | 1.58 | 0.10 | −0.14 |
| Denmark | −0.05 | 0.29 | 1 | 0.10 | −2.26 | 1.03 | 6.12 | −0.01 | 0.46 | −1.09 | −0.78 |
| Finland | 0.00 | 0.41 | 1 | −1.27 | −2.48 | 7.74 | 6.92 | 0.62 | 0.47 | 0.41 | 5.56 |
| France | −0.82 | 0.59 | 1 | 0.18 | −1.49 | 0.18 | 2.24 | 1.15 | −0.62 | 0.42 | −0.52 |
| Germany | −0.45 | −0.18 | 1 | −0.15 | −0.80 | 0.14 | −0.89 | 0.52 | −0.96 | 0.33 | 0.15 |
| Greece | 0.00 | −0.31 | 1 | 0.36 | −1.84 | −1.81 | 4.28 | 0.61 | −0.25 | −0.49 | −0.69 |
| India | −0.32 | −0.03 | 1 | −1.42 | −7.74 | 0.49 | 31.01 | 1.87 | 0.63 | 0.09 | −0.27 |
| Ireland | −0.52 | 0.17 | 1 | 0.30 | −0.68 | 2.44 | −1.19 | −0.44 | −0.90 | 0.75 | −4.21 |
| Italy | −0.46 | 0.14 | 1 | 0.09 | 2.25 | −2.56 | −15.09 | −0.17 | 0.62 | −0.39 | −0.01 |
| Japan | −0.48 | −0.11 | 1 | 0.14 | 1.41 | 0.70 | −11.19 | −1.36 | −1.91 | −0.19 | −0.53 |
| Mexico | −1.05 | −0.40 | 1 | −0.83 | −0.41 | 0.84 | −2.95 | 0.98 | 0.42 | 0.06 | −0.13 |
| Netherlands | −0.21 | −0.55 | 1 | 0.00 | −0.56 | −1.85 | −1.94 | 1.33 | −0.05 | −1.37 | 1.72 |
| Singapore | −0.20 | −0.45 | 1 | −0.05 | −0.51 | 0.08 | −2.19 | 0.64 | 0.24 | −0.80 | −2.65 |
| Spain | −0.36 | −0.17 | 1 | 0.16 | 2.97 | −4.67 | −18.32 | 0.29 | 0.82 | −0.88 | 1.40 |
| Sweden | −0.04 | −0.27 | 1 | 2.07 | −8.08 | −1.09 | 32.92 | 0.59 | −0.23 | 2.26 | 0.13 |
| UK | −0.04 | 0.09 | 1 | −3.56 | 9.64 | −12.90 | −50.20 | −2.27 | 1.22 | −0.10 | −2.19 |
| USA | −0.12 | −0.06 | 1 | 2.35 | −2.59 | 3.40 | 7.75 | −0.25 | 0.46 | 0.59 | −0.68 |
| Uruguay | −0.42 | 0.02 | 1 | −1.85 | 2.16 | −3.16 | −15.47 | −0.60 | −0.53 | 0.09 | 1.19 |
| South Africa | −0.41 | 0.14 | 1 | −4.45 | −7.39 | −1.89 | 28.86 | 1.98 | −3.85 | −0.72 | −0.32 |
| Indonesia | 0.05 | 0.07 | 1 | −5.11 | −10.19 | 4.05 | 39.97 | −1.60 | −0.59 | −1.94 | −3.51 |
| Turkey | −0.02 | −0.05 | 1 | 17.13 | −53.64 | 9.75 | 249.77 | 1.10 | 0.88 | 1.42 | −1.75 |

**Table 7.** Comparing the cointegrated coefficient of the GVAR and traditional methodologies.

| | Cointegrated Relationship | | | | | | | | | Difference between NFA long run coefficients in both models |
|---|---|---|---|---|---|---|---|---|---|---|
| | GVAR | | | | | Traditional Methodologhy | | | | |
| | RER | NFA | RERw | NFAw | Constant | RER | NFA | RERw | NFAw | Constant | |
| Australia | 1 | −1.3 | 0.7 | −2.1 | −8.2 | 1 | −0.9 | 0 | 0 | −4.9 | −0.4 |
| Austria | 1 | 1.3 | 0.2 | −2.2 | −5.4 | 1 | −0.1 | 0 | 0 | −4.6 | 1.4 |
| Belgium | 1 | −0.1 | −0.9 | −0.8 | −0.6 | 1 | 0.0 | 0 | 0 | −4.6 | −0.2 |
| Brazil | 1 | −1.2 | 0.9 | −3.2 | −9.2 | 1 | −1.3 | 0 | 0 | −5.1 | 0.1 |
| Canada | 1 | 0.0 | −0.4 | 0.1 | −2.8 | 1 | 1.0 | 0 | 0 | −4.3 | −1.0 |
| China | 1 | −8.9 | −18.5 | 14.2 | 80.2 | 1 | −2.1 | 0 | 0 | −4.5 | −6.8 |
| Colombia | 1 | −6.5 | 4.5 | 0.5 | −26.8 | 1 | −5.0 | 0 | 0 | −5.5 | −1.4 |
| Denmark | 1 | 0.1 | −2.3 | 1.0 | 6.1 | 1 | −0.2 | 0 | 0 | −4.6 | 0.3 |
| Finland | 1 | −1.3 | −2.5 | 7.7 | 6.9 | 1 | −0.3 | 0 | 0 | −4.8 | −0.9 |
| France | 1 | 0.2 | −1.5 | 0.2 | 2.2 | 1 | −0.3 | 0 | 0 | −4.6 | 0.5 |
| Germany | 1 | −0.2 | −0.8 | 0.1 | −0.9 | 1 | −0.1 | 0 | 0 | −4.6 | −0.1 |
| Greece | 1 | 0.4 | −1.8 | −1.8 | 4.3 | 1 | 0.0 | 0 | 0 | −4.3 | 0.3 |
| India | 1 | −1.4 | −7.7 | 0.5 | 31.0 | 1 | −8.1 | 0 | 0 | −5.6 | 6.7 |
| Ireland | 1 | 0.3 | −0.7 | 2.4 | −1.2 | 1 | 0.2 | 0 | 0 | −4.4 | 0.1 |
| Italy | 1 | 0.1 | 2.3 | −2.6 | −15.1 | 1 | 0.2 | 0 | 0 | −4.6 | −0.1 |
| Japan | 1 | 0.1 | 1.4 | 0.7 | −11.2 | 1 | 0.3 | 0 | 0 | −4.5 | −0.1 |
| Mexico | 1 | −0.8 | −0.4 | 0.8 | −3.0 | 1 | −0.4 | 0 | 0 | −4.7 | −0.4 |
| Netherlands | 1 | 0.0 | −0.6 | −1.8 | −1.9 | 1 | −0.2 | 0 | 0 | −4.6 | 0.2 |
| Singapore | 1 | −0.1 | −0.5 | 0.1 | −2.2 | 1 | 0.0 | 0 | 0 | −4.6 | 0.0 |
| Spain | 1 | 0.2 | 3.0 | −4.7 | −18.3 | 1 | 0.1 | 0 | 0 | −4.4 | 0.0 |
| Sweden | 1 | 2.1 | −8.1 | −1.1 | 32.9 | 1 | −1.0 | 0 | 0 | −4.9 | 3.1 |
| United Kingdom | 1 | −3.6 | 9.6 | −12.9 | −50.2 | 1 | −0.2 | 0 | 0 | −4.8 | −3.3 |
| United States | 1 | 2.4 | −2.6 | 3.4 | 7.8 | 1 | −0.5 | 0 | 0 | −4.7 | 2.8 |
| Uruguay | 1 | −1.9 | 2.2 | −3.2 | −15.5 | 1 | −1.3 | 0 | 0 | −4.9 | −0.5 |
| South Africa | 1 | −4.5 | −7.4 | −1.9 | 28.9 | 1 | 1.5 | 0 | 0 | −4.5 | −5.9 |
| Indonesia | 1 | −5.1 | −10.2 | 4.1 | 40.0 | 1 | −4.0 | 0 | 0 | −6.6 | −1.1 |
| Turkey | 1 | 17.1 | −53.6 | 9.7 | 249.8 | 1 | 0.2 | 0 | 0 | −4.2 | 16.9 |

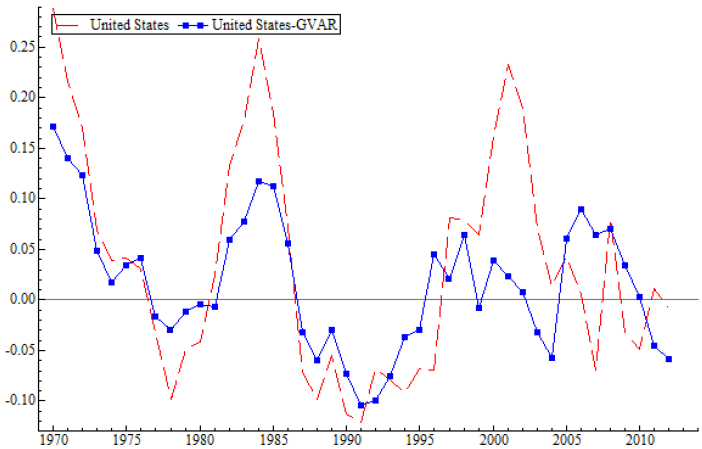



**Figure 1.** Real exchange rate misalignment estimate – GVAR versus traditional methodology – United States.

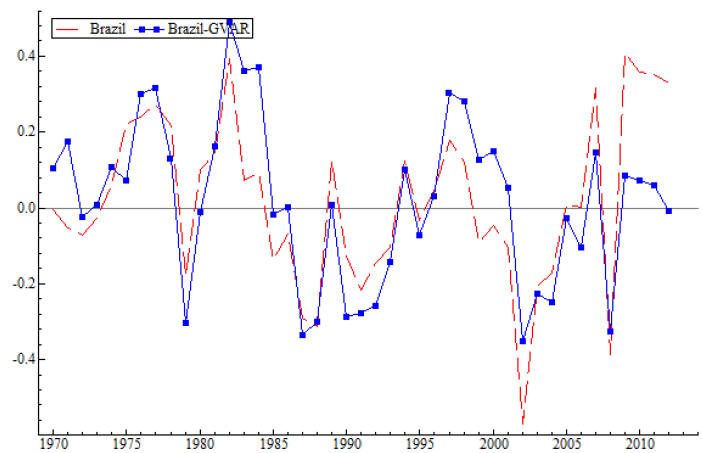



**Figure 2.** Real exchange rate misalignment estimate – GVAR versus traditional methodology – Brazil.

In a recent paper, Ericsson and Reisman (2012) proposes a refinement of the GVAR approach using a model selection procedure called *Autometrics*.[7] This refinement allows one to search for points of instability and structural changes in the modeling process. It is also possible to incorporate a wider range of information sets in each country model. However, this must occur within a rigorous general-to-specific econometric modeling approach in the spirit of the London School of Economics (LSE) tradition developed by David Hendry. This reduces the possibility of criticism that a time series analysis can only manage a few real exchange rate determinants at a time, as compared to panel models. By using the GVAR refinement proposed by Ericsson and Reisman (2012), it is possible, for example, to match the merits of the IMF approach, with its wide range of variables, to a panel with the merits of a global model. One can opt to model different set fundamentals for each country.

[7]See Doornik (2009).

The interdependence among countries is introduced in GVAR by adding trade weighted foreign variables in the model of each specific country. We opt to use *Autometrics* search algorithm to check if these variables explain the dynamics of each country models. The traditional methodology can be extended in many directions. The GVAR is one possible way but not the only one. We decide to construct indicators functions (impulse, step and time trend dummies) to allow for structural change in mean of the process. We also run principal components analysis on all first differences of the logarithm of real effective exchange rate and net foreign assets and use the factors as possible regressors. The rationality for this relies on possible existence of another global factor not well captured by GVAR averages variables.

In order to evaluate the GVAR we must compare it to a more general model. We opt to include others sources of interdependence and structural change in the mean parameters of the model. This general model nests the GVAR structure. We run the *Autometrics* in two steps.

In the first step the general unrestricted model (GUM) is given by the following equation:

$$\Delta x_{i,t} = \sum_{h=1}^{q} \theta_{1h}^{1} V_{ht}^{i} + \sum_{h=1}^{T_i} \theta_{2h}^{1} IS(T_h) + \sum_{h=1}^{T_i} \theta_{2h}^{1} DS(T_h) + \sum_{i=1}^{T_i} \theta_{3h}^{1} DT(T_h) + u_i^1, \tag{19}$$

where

$$IS(T_i) = \begin{cases} 1, & \text{if } t = T_i, \\ 0, & \text{if } t \neq T_i; \end{cases} \quad DS(T_i) = \begin{cases} 1, & \text{if } t \geq T_i, \\ 0, & \text{if } t < T_i; \end{cases} \quad DT(T_i) = \begin{cases} t, & \text{if } t \geq T_i, \\ 0, & \text{if } t < T_i; \end{cases}$$

and $V^i = V_1^i, V_2^i, \ldots, V_q^i$ contains all GVAR's country $i$ variables. The superscript denotes the step.

In the second step the general unrestricted model contains the variables of the final model in the first step and broader set of variables:

$$\Delta x_{i,t} = \sum_{h=1}^{q} \theta_{1i}^{2} FM_{it} + \sum_{h=1}^{N} \sum_{j=1}^{k_i} \theta_{2ij}^{2} \Delta x_{h,t-j} + \sum_{h=1}^{N} \sum_{j=1}^{k_i} \theta_{3ij}^{2} \Delta x_{h,t-j}^{*} + \sum_{h=1,h\neq j}^{N} \theta_{4ij}^{2} ecm_{h,t-1} + \sum_{h=1}^{2*N} \sum_{j=1}^{k_i} \theta_{3ij}^{2} PC_{h,t-j} + u_i^2, \tag{20}$$

where *FM* denotes the variables in the final model of step 1 and *PC* denotes the principal components of all RER and NFA variables, *ecm* denotes error correction mechanism and the remainder variables were defined previously in the paper.

The GUM of the first step in our exercise contains a total of $\text{TOT}_1 = T*3+2*6+2+1$ variables. In our application $\text{TOT}_1$ equals $75 = 30*3+12+2+1$. The GUM in the second step will contain at least $\text{TOT}_2$ variables $(2*[N+(N-1)]+2*[N+(N-1)]+(N-1)+2*N)$. In our example the number of countries is 27 and $\text{TOT}_2$ equals 292. The GUM upper limit will be $292+75=367$. The *Autometrics* will retain under the null that all coefficients are zero an average of irrelevant regressors given by $\text{TOT} = (\text{TOT}_1 + \text{TOT}_2)$ times the chosen nominal size. If we set the nominal size to 0.0001 (minute), *Autometrics* will retain 0.0367 irrelevant regressors on average and about 0.367 irrelevant regressors if the nominal size is set to 0.001 (tiny).

Autometrics allows us to check whether the GVAR approach is a congruent statistical model and a rival model that contains broader information set can encompass the GVAR. If the final model contains only the variables of the traditional approach it is reasonable to conclude that neither GVAR nor the broader GUM are worth taking.

Table 8 provides information on each country final model chosen by *Autometrics* search algorithm. In some case there is evidence of structural change (e.g. Australia, Greece, Uruguay, Turkey). In these cases, a closer look to other domestic variables can be an alternative. For some countries our evidence suggests that the GVAR can not capture all possible sources of interdependence. For example, in the case of Ireland, a closer link to the dynamics of United Kingdom's real effective exchange was revealed. However the foreign NFA weigthed variable seems to explain the dynamics in United Kingdom and United States cases quite well. These countries have the leading financial centers of the world, and GVAR is probably capturing this interdependence channel. In many countries the final model selected by *Autometrics* contains the foreign trade weighted variable. This results suggest that GVAR improves of our understanding on the dynamics real exchange rate and net foreign asset. The principal components variables were selected only in some few models. For some countries, the results of *Autometrics* using minute instead of tiny significance level seems more intuitive (e.g. Brazil or Germany). Spain is the only country that *Autometrics* results are not easy to rationalize.

### 4.5 Discussion, limitations and possible extensions

The previous discussion on the merits and limitations of the time series and panel approaches is addressed in this section, as well as whether the GVAR model can be a bridge linking both approaches. The time series approaches allow little room to introduce fundamentals because of the sample size available in macroeconomic datasets. The panel approach allows a more flexible structure and the inclusion of a larger group of fundamentals in the analysis. However, this approach must limit heterogeneity to a manageable level. It is not clear to what extent this can lead to distortions, since the main goal is to make assertions on specific units, not to assess the relevance of a group of variables in explaining real exchange rate movements and their average effect. A GVAR model can reconcile the merits of the two approaches, allowing us to map directly the effect of trading partner shocks on a country.

In the same way, it is possible to adapt the decomposition of Gonzalo and Granger (1995) to the GVAR environment. The same can be done for a Beveridge and Nelson decomposition, as shown in the work of Proietti (1997), under a VECM framework. The development of a model for the eurozone is a natural extension of our work. A regional factor can be easily added to the GVAR to map directly the interdependence between countries in the region. To the best of our knowledge, this has not been done before. Although we did not do so, our GVAR model was able to capture a strong interdependence effect among eurozone countries.

Even the IMF approach does not directly tackle the question of interdependence between countries. However, the IMF approach does have the benefit of considering a wide range of fundamentals, and incorporates the role of policy gaps in determining the misalignment.

## 5. Final remarks

In this study, we estimated a global VAR to investigate the interdependence hypothesis among countries in terms of their real effective exchange rates and fundamentals. We were

**Table 8.** Summary of Autometrics results.

| Countries | Type of search | Variables in the final model | | | | | | Evidence of structural change | Are specification tests good? | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | ECM | RER | Weighted RER | NFA | Weighted NFA | Factors | | At 5% | At 1% |
| Australia | Tiny | 0 | No | No | No | No | No | $DT, DI, DS$ | No | Yes |
| | Minute | 0 | No | No | No | No | No | $DI$ | Yes | Yes |
| Austria | Tiny | 0 | No | 0 | No | No | No | $DI$ | No | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Belgium | Tiny | 0 | No | No | No | No | No | No | No | Yes |
| | Minute | 0 | No | No | No | No | No | No | No | Yes |
| Brazil | Tiny | 0 | No | India, USA | No | No | NFA | $DI$ | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | $DI$ | No | Yes |
| Canada | Tiny | 0 | No | No | No | 0 | No | No | No | Yes |
| | Minute | No | No | No | No | No | No | No | Yes | Yes |
| China | Tiny | 0, Sweden | No | No | Germany | No | NFA | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Colombia | Tiny | 0 | USA | 0, Mexico | No | No | No | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Denmark | Tiny | 0 | No | No | No | No | No | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Finland | Tiny | 0 | No | No | 0, Ireland | No | No | DT | No | Yes |
| | Minute | 0 | No | No | No | No | No | DT | No | No |
| France | Tiny | 0 | No | 0 | No | No | NFA | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | No | Yes |
| Germany | Tiny | 0, South Africa | No | No | No | No | No | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Greece | Tiny | No | No | No | No | No | No | $DT, DI, DS$ | No | Yes |
| | Minute | 0 | No | No | No | No | No | No | No | No |
| India | Tiny | 0 | No | No | No | No | No | No | No | Yes |
| | Minute | 0 | No | No | No | No | No | No | No | Yes |
| Ireland | Tiny | 0 | UK | No | No | 0, Singapore | NFA | No | Yes | Yes |
| | Minute | 0 | UK | 0 | No | 0 | No | No | Yes | Yes |
| Italy | Tiny | 0, South Africa | No | No | No | No | NFA | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Japan | Tiny | 0 | Spain | 0 | No | 0 | No | No | Yes | Yes |
| | Minute | 0 | No | 0 | No | 0 | No | No | No | Yes |
| Mexico | Tiny | 0 | No | 0 | No | 0 | No | No | No | Yes |
| | Minute | 0 | 0 | No | No | No | No | No | No | No |
| Netherlands | Tiny | 0 | No | 0 | No | No | No | No | No | No |
| | Minute | 0 | No | 0 | No | No | No | $DI$ | No | Yes |
| Singapore | Tiny | No | No | No | No | No | No | No | No | Yes |
| | Minute | No | No | No | No | No | No | No | No | Yes |
| Spain | Tiny | 0, Australia | USA | No | No | No | No | No | No | Yes |
| | Minute | 0, Australia, Turkey, Uruguay | USA | Colombia | No | No | No | No | No | Yes |
| Sweden | Tiny | 0 | No | 0 | No | No | No | No | No | Yes |
| | Minute | 0 | No | No | No | No | No | No | No | No |
| UK | Tiny | 0 | No | No | No | 0 | No | No | No | Yes |
| | Minute | 0 | No | No | No | 0 | No | No | No | Yes |
| USA | Tiny | 0 | No | No | No | 0 | No | No | No | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Uruguay | Tiny | 0 | No | No | No | No | No | $DT, DI$ | No | No |
| | Minute | 0 | No | No | No | No | No | DT | No | No |
| South Africa | Tiny | 0 | No | No | No | No | No | No | Yes | Yes |
| | Minute | 0 | No | No | No | No | No | No | Yes | Yes |
| Turkey | Tiny | No | No | No | No | No | No | $DT, DI, DS$ | No | Yes |
| | Minute | No | No | No | No | No | No | $DT, DI, DS$ | No | Yes |

*Notes:* "0" refers to own country, $DT$ is break time trend dummy, $DI$ is an impulse indicator, and $DS$ is a step dummy.

able to find evidence in favor of interdependence in both the short and long run for some countries. In only a few cases in the sample could the null hypothesis of no interdependence not be rejected.

We also discussed the impact that the GVAR may have on exchange rate misalignment estimates. Here, we adapted the Gonzalo and Granger decomposition to a GVAR framework and conducted an empirical exercise to try to explain the relevance of global effects to exchange rate misalignment estimates. Our findings show that the effects are greater for small or developing countries, because they tend to be more affected by global economy conditions.

Our global model was also capable of detecting important linkages between eurozone countries, as expected. The United States and Germany, two leading economies in the world, seem to have an effect on the real exchange rate of other countries. However, their exchange rate misalignment estimates are only marginally affected in terms of magnitude and sign when both models' estimates are compared. The reason for this has to do with the dynamics of their real effective exchange rates, which are almost not affected by others countries' variables.

Finally, possible extensions to our approach include improving on country-specific models by using recent advances in time series model selection to investigate the role of a broader set of domestic variables. We can access not only the statistical significance of external factors but their relative importance in relation to domestic factors.

## References

Ahmad, Y., & Craighead, W. D. (2011). Temporal aggregation and purchasing power parity persistence. *Journal of International Money and Finance*, *30*(5), 817–830. http://dx.doi.org/10.1016/j.jimonfin.2011.05.008

Alberola, E., Cervero, S. G., Lopez, H., & Ubide, A. (1999). *Global equilibrium exchange rate: Euro, dolar, "ins", "outs," and other major currencies in a panel cointegration framework* (IMF Working paper No. 175). International Monetary Fund. https://www.imf.org/en/Publications/WP/Issues/2016/12/30/Global-Equilibrium-Exchange-Rates-Euro-Dollar-Ins-Outs-and-Other-Major-Currencies-in-a-Panel-3369

Bilson, J. F. O. (1979). Recent developments in monetary models of exchange rate determination. *Staff Papers (International Monetary Fund)*, *26*(2), 201–223. http://dx.doi.org/10.2307/3866509

Cline, W. R., & Williamson, J. (2008). Estimates of the equilibrium exchange rate of the renminbi: Is there a consensus and if not, why not? In M. Goldstein & N. R. Lardy (Eds.), *Debating China's exchange rate policy* (pp. 131–168). Peterson Institute for International Economics. https://piie.com/publications/chapters_preview/4150/04iie4150.pdf

Doornik, J. A. (2009). Autometrics. In J. L. Castle & N. Shephard (Eds.), *The methodology and practice of econometrics: A Festschrift in honour of David F. Hendry.* Oxford University Press.

Dornbusch, R. (1976). Expectations and exchange rate dynamics. *Journal of Political Economy*, *84*(6), 1161–1176. https://www.jstor.org/stable/1831272

Edwards, S. (1989). Exchange rate misalignment in developing countries. *The World Bank Research Observer*, *4*(1), 3–21. http://documents.worldbank.org/curated/en/737331468739271547/Exchange-rate-misalignment-in-developing-countries

Edwards, S. (1991). *Real exchange rates, devaluation, and adjustment: Exchange rate in developing countries*. MIT Press.

Engle, R. F., & Granger, C. W. J. (1987). Co-integration and error correction: Representation, estimation, and testing. *Econometrica*, *55*(2), 251–276. http://dx.doi.org/10.2307/1913236

Engle, R. F., Hendry, D. F., & Richard, J.-F. (1983). Exogeneity. *Econometrica*, *51*(2), 277–304. http://dx.doi.org/10.2307/1911990

Ericsson, N. R., & Reisman, E. L. (2012). Evaluating a global vector autoregression for forecasting. *International Advances in Economic Research*, *18*(3), 247–258. http://dx.doi.org/10.1007/s11294-012-9357-0

Faruqee, H. (1995). Long-run determinants of the real exchange rate: A stock-flow perspective. *IMF Staff Papers*, *42*(1), 80–107. http://dx.doi.org/10.2307/3867341

Froot, K. A., & Rogoff, K. (1995). Perspectives on PPP and long-run real exchange rates. In G. M. Grossman & K. Rogoff (Eds.), *Handbook of international economics* (Vol. 3, pp. 1647–1688). North Holland.

Ghysels, E., & Miller, I. J. (2015). Testing for cointegration with temporally aggregated and mixed-frequency time series. *Journal of Time Series Analysis*, *36*(6), 797–816. http://dx.doi.org/10.1111/jtsa.12129

Gonzalo, J., & Granger, C. (1995). Estimation of common long-memory components in cointegrated systems. *Journal of Business & Economic Statistics*, *13*(1), 27–35. http://dx.doi.org/10.2307/1392518

Hecq, A., Palm, F. C., & Urbain, J.-P. (2000). Permanent-transitory decomposition in VAR models with cointegration and common cycles. *Oxford Bulletin of Economics and Statistics*, *62*(4), 511–532. http://dx.doi.org/10.1111/1468-0084.00185

Hecq, A., Palm, F. C., & Urbain, J.-P. (2002). Separation, weak exogeneity, and P-T decomposition in cointegrated VAR systems with common features. *Econometric Reviews*, *21*(3), 273–307. http://dx.doi.org/10.1081/ETC-120015785

Hendry, D. F. (1995). *Dynamic econometrics*. Oxford University Press.

Hossfeld, O. (2010). *Equilibrium real exchange rates and real exchange rate misaligments: Time series vs. panel estimates* (Tech. Rep. No. 65). FIW – Forschungsschwerpunkt Internationale Wirtschaft. http://hdl.handle.net/10419/121070 (Working Paper)

IMF – International Monetary Fund. (2013). *2013 pilot external sector report*. Washington, DC: International Monetary Fund. https://www.imf.org/en/Publications/Policy-Papers/Issues/2016/12/31/2013-Pilot-External-Sector-Report-PP4789

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, *12*(2-3), 231–254. http://dx.doi.org/10.1016/0165-1889(88)90041-3

Johansen, S. (1995). Identifying restrictions of linear equations with applications to simultaneous equations and cointegration. *Journal of Econometrics*, *69*(1), 111–132. http://dx.doi.org/10.1016/0304-4076(94)01664-L

Kubota, M. (2009). *Real exchange rate misaligments* (PhD Thesis). University of York, York, North Yorkshire.

Lane, P. R., & Milesi-Ferretti, G. M. (2007). The external wealth of nations mark II: Revised and extended estimates of foreign assets and liabilities, 1970–2004. *Journal of International Economics*, *73*(2), 223–250. http://dx.doi.org/10.1016/j.jinteco.2007.02.003

Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics*, *14*(1-2), 3–24. http://dx.doi.org/10.1016/0022-1996(83)90017-X

Mussa, M. (1976). The exchange rate, the balance of payments and monetary and fiscal policy under a regime of controlled floating. *The Scandinavian Journal of Economics*, *78*(2), 229–248. http://dx.doi.org/10.2307/3439926

Pesaran, H. M., Schuermann, T., & Weiner, S. M. (2004). Modeling regional interdependencies using a global error-correcting macroeconometric model. *Journal of Business & Economic Statistics*, *22*(2), 129–162. http://dx.doi.org/10.1198/073500104000000019

Proietti, T. (1997). Short-run dynamics in cointegrated systems. *Oxford Bulletin of Economics and Statistics*, *59*(3), 405–422. http://dx.doi.org/10.1111/1468-0084.00073

Rossi, B. (2013). Exchange rate predictability. *Journal of Economic Literature*, *51*(4), 1063–1119. http://dx.doi.org/10.1257/jel.51.4.1063

Shin, Y. (1994). A residual-based test of the null of cointegration against the alternative of no cointegration. *Econometric Theory*, *10*(1), 91–115. http://dx.doi.org/10.1017/S0266466600008240

Stein, J. L. (1995). *The fundamental determinants of the real exchange rate of the U.S. dollar relative to other G-7 currencies* (IMF Working Paper No. 95/81). International Monetary Fund. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=883229

Taylor, A. M. (2001). Potential pitfalls for the purchasing-power-parity puzzle? Sampling and specification biases in mean-reversion tests of the Law of One Price. *Econometrica*, *69*(2), 473–498. https://www.jstor.org/stable/2692239

Williamson, J. (Ed.). (1994). *Estimating equilibrium exchange rates*. Peterson Institute for International Economics.

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180022
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Trajetória Longitudinal dos Alunos de Escolas Profissionalizantes para o Mercado de Trabalho*

Francisca Zilania Mariano†
Ronaldo Albuquerque Arraes‡

**Sumário**

1. Introdução....451
2. Metodologia....455
3. Resultados....458
4. Conclusões....464
Apêndice A. Procedimentos Metodológicos ...470
Apêndice B. Descrição das Variáveis ....473
Apêndice C. Resultados Complementares....475

**Palavras-chave**
Ensino profissionalizante, mercado de trabalho, dados longitudinais

**JEL Codes**
C21, I21, I28

**Resumo · Abstract**

The article evaluates the effect of the Vocational Schools (VS) on the occupations and earnings of their graduates in the labor market. Combining longitudinal data from the School Census (2011), SPAECE (2011) and RAIS (2014), estimates of the earning equations, weighted by entropy and corrected for sample selectivity, indicated insignificant effects of VSs on the earnings. Additional estimates of a discrete multinomial model, weighted by entropy for five levels of occupancy, also pointed to insignificant effects of VSs, except for the health area in intermediate occupations.

## 1. Introdução

Avaliação de políticas públicas direcionadas para a inserção de jovens no mercado de trabalho tem se mantido na agenda de pesquisadores devido à persistência de altas taxas daqueles que nem estudam e nem trabalham. Bassi, Busso, Urzúa, e Vargas (2012) verificam que, no Brasil e em outros países da América Latina, aproximadamente 15% dos jovens entre 16 e 24 anos inserem-se nessa categoria. Segundo relatório do Banco Mundial (de Hoyos, Rogers, & Székely, 2016), essa proporção aumentou para 19% nesta região. Além disso, nessa fase existem elevadas taxas de empregos informais, indicando que estes jovens não possuem habilidades exigidas em empregos formais de qualidade (Araújo, Chein, & Pinto, 2014). No Brasil, cerca de 4 milhões de jovens trabalham em atividades informais, das quais 90% correspondem a menos de um salário mínimo (Simões, 2010).

Segundo Guimarães e Almeida (2013), esse grupo de jovens torna-se um público vulnerável que enfrentam maiores dificuldades de encontrar emprego, e tendem a trabalhar

*Os autores agradecem a Secretaria de Educação do Estado do Ceará e ao Ministério do Trabalho pela disponibilização dos dados.

†Universidade Federal do Ceará, Campus Sobral – Mucambinho, Bloco I. Rua Coronel Estanislau Frota, s/n, Centro, Sobral, CE, CEP 62010-560, Brasil.

‡Universidade Federal do Ceará (UFC), Programa de Pós-Graduação em Economia (CAEN). Av. da Universidade, 2762, Fortaleza, CE, CEP 60020-181, Brasil.

zilania@ufc.br ronald@ufc.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Recebido em 10 de janeiro de 2018
Aceitação final em 28 de agosto de 2018

em ocupações mais precárias devida à incipiente formação educacional adquirida. Araújo et al. (2014) afirmam, juntamente com outros estudos realizados para países da América Latina, que grande parte dos jovens egressos do ensino médio apresenta condições piores de trabalho, como, por exemplo, menores salários e menos estabilidade, em comparação aos trabalhadores mais experientes. Segundo os autores, isso ocorre porque a maioria dos jovens inicia a trajetória profissional no mercado de trabalho informal e, posteriormente, aprimoram as habilidades até conseguirem um emprego formal. Na visão desses autores, essas habilidades deveriam ser adquiridas na etapa de ensino de conclusão da educação básica, o nível médio.

Dentre os tipos de políticas que visam amenizar essa situação, encontra-se a educação profissional, a qual proporciona ao aluno um ensino específico direcionado ao mercado de trabalho. Para Silva, Pelissari, e Steimbach (2013), os jovens procuram os cursos técnicos visando uma formação de qualidade superior, profissionalização em áreas com melhores *status* e garantia de empregabilidade. A educação profissional pode ser uma opção para os países em desenvolvimento que buscam melhorar as oportunidades e os ganhos dos jovens no mercado de trabalho, ao prepara-los com técnicas e habilidades de trabalhos específicos. Porém, o debate sobre os benefícios deste tipo de ensino é inconclusivo na literatura internacional (Mahirda & Wahyuni, 2016). Em alguns países, estudantes do ensino profissional obtém maiores rendimentos quando comparados aos alunos de escolas gerais. Neuman e Ziderman (1991) encontraram que, em Israel, os graduados de escolas profissionais ganham 10% a mais que os estudantes das escolas tradicionais. No Egito, esse tipo de ensino resulta em prêmio de 29,3% para homens e 27,2% para as mulheres (El-Hamidi, 2006). Attanasio, Kugler, e Meghir (2011) avaliam o impacto de um programa de treinamento sobre jovens desfavorecidos na Colômbia em 2005, escolhidos de forma aleatória, e verificaram que a participação deste programa eleva as chances de adquirir emprego formal e eleva os ganhos em 19,6% para as mulheres. Newhouse e Suryadarma (2011) utilizaram uma rica base de dados contendo informações longitudinais de 1993 a 2007 para os jovens na Indonésia e verificaram que o prêmio estimado sobre os salários é maior para as mulheres graduadas do ensino profissional do que para os homens.

Por outro lado, na Tanzânia (Kahyarara & Teal, 2008) e no Suriname (Horowitz & Schenzler, 1999) foram constatados que estudantes das escolas gerais apresentam maiores rendimentos do que as profissionais, enquanto que na Alemanha (Lechner, 2000), Coréia do Sul (Chae & Chung, 2009), Romênia (Malamud & Pop-Eleches, 2010) e na Indonésia (Chen, 2009; Mahirda & Wahyuni, 2016), os resultados entre as escolas não diferiram significativamente.

Trabalhos feitos para o Brasil têm focado nos efeitos desse tipo de ensino sobre a inserção no mercado de trabalho, porém, alguns autores analisam apenas a probabilidade de estar trabalhando, não considerando os rendimentos ou ocupações (Araújo et al., 2014), enquanto outros negligenciam as devidas correções ao considerar somente os indivíduos que estão no mercado de trabalho (Severnini & Orellano, 2010; Assunção & Gonzaga, 2010; Amoroso Neto, Menezes-Filho, & Komatsu, 2017). Além disso, não se constata em nenhum deles análises com informações longitudinais a fim de identificar o jovem pós ensino médio no mercado de trabalho. Este fato, aliado à carência de tratamento adequado aos dados na maioria destes trabalhos, comprometem seus resultados.

Severnini e Orellano (2010) questionam se os programas de qualificação profissional realmente contribuem para os treinados obterem uma melhoria de bem-estar através de uma eventual elevação na probabilidade de inserção no mercado de trabalho. Para tanto,

os autores conduziram uma investigação comparativa entre os egressos e aqueles que não cursaram esse tipo de ensino. Através de microdados da Pesquisa sobre Padrões de Vida (PPV) do IBGE de 1996 encontraram que os egressos de cursos profissionalizantes de nível básico tinham renda esperada 37% maior do que a dos demais. Por outro lado, para os egressos do ensino profissional de nível tecnológico, constataram uma redução de 27% da renda esperada, comparativamente aos que não participaram desse tipo de curso no ensino superior.

Assunção e Gonzaga (2010), com base nos microdados do suplemento especial sobre educação profissional da PNAD/2007, observaram que a ocorrência de indivíduos que optam por educação profissional, e cujas famílias detém renda per capita inferior a dois salários mínimos, é bem menor do que nas demais faixas de renda, além do que, essa opção proporciona uma maior produtividade aos trabalhadores.

Araújo et al. (2014) avaliaram os efeitos da educação profissional e tecnológica (EPT) de nível médio sobre a inserção produtiva do aluno, e verificaram um aumento entre 1,2 e 1,3 pontos percentuais na probabilidade de estar empregado. Já a probabilidade de trabalhar em atividade para a qual se preparou é de, aproximadamente, 17 pontos percentuais maior do que os alunos que não realizaram EPT.

Apresentando similaridades com o presente estudo, Amoroso Amoroso Neto et al. (2017) compararam indivíduos com níveis de educação formal semelhante para cada modalidade de estudo (médio técnico, graduação tecnológica e qualificação profissional), além daqueles diferenciados que concluíram os cursos do PRONATEC.[1] Utilizando dados em *cross section* do suplemento da PNAD/2014, estimaram a equação de rendimentos e o *Propensity Score Macthing* – PSM, embora sem o devido controle pelo *backgound* educacional dos alunos, nem correções para seletividade amostral. Concluíram que os formandos no ensino médio técnico e dos cursos de qualificação profissional ganham salários, em média, 15% e 10% maiores, respectivamente, do que os indivíduos com características semelhantes que se graduaram na educação formal; ao considerar somente os que concluíram o ensino médio técnico, a superioridade é 19%. Além disso, encontraram que os egressos de cursos de qualificação profissional do PRONATEC têm retornos comparativamente menores.

Embora não sendo uma avaliação do ensino profissional, Lin, Lutter, e Ruhm (2016) usaram informações de uma pesquisa longitudinal de jovens e dados suplementares para avaliar como o desempenho cognitivo, medido no fim do ensino secundário, está relacionado com resultados no mercado de trabalho. Observam-se cinco resultados principais. Primeiro, o desempenho cognitivo está positivamente associada a resultados futuros do mercado de trabalho em todas as idades. Segundo, os retornos da habilidade cognitiva aumentam com a idade. Terceiro, o efeito sobre os rendimentos totais reflete uma combinação de impactos positivos do desempenho cognitivo tanto para salários/hora e horas de trabalho anuais. Em quarto lugar, os retornos da habilidade cognitiva são maiores para as mulheres que para os homens e para os negros e hispânicos do que para os brancos não-hispânicos. Por fim, os ganhos médios dos rendimentos ao longo da vida previsto para maiores níveis de desempenho cognitivo são, apenas, ligeiramente superiores aos reportados em estudos anteriores.

As Escolas Estaduais de Educação Profissional (EEEP) no Ceará fornecem, além do ensino com as disciplinas básicas do currículo do ensino médio tradicional, os cursos técnicos e profissionalizantes voltados para diversas áreas (saúde, agricultura, computação, finanças,

[1] Programa Nacional de Acesso ao Ensino e Emprego – PRONATEC

etc). Essa ação possibilita, não apenas o aluno concluir o ensino médio e se profissionalizar para o mercado de trabalho, como também capacitá-lo a ingressar em universidades. É, pois, centrado nessas políticas de ensino que este trabalho se desenvolve, com ênfase no ensino profissionalizante. Cabe destacar, todavia, que a opção pela escolha do estado do Ceará como objeto de estudo decorre de sua representatividade[2] nacional e pertinência na geração das amostras para verificações empíricas em nível longitudinal. Além disso, esta escolha foi fundamentada, principalmente, no acesso a permissibilidade da obtenção de uma rica base de dados identificada com disposições longitudinais sobre características pessoais e educacionais dos alunos, background familiar, infraestrutura das escolas e informações sobre o mercado de trabalho, as quais permitiram acompanhar o aluno no término do ciclo do ensino médio e a entrada no mercado de trabalho, através do cruzamento das seguintes bases: Sistema Permanente de Avaliação da Educação Básica do Ceará – SPAECE, Censo Escolar, e Relação Anual de Informações Sociais – RAIS.

Frente aos demais trabalhos consultados na literatura brasileira, acredita-se que este forneça inovações ao trazer evidências acerca do papel das escolas profissionais públicas sobre os rendimentos e os tipos de ocupações dos jovens, baseadas em uma composição inédita de dados. A contribuição do empirismo reside na construção de identificadores que tornaram possível a geração de uma amostra com informações longitudinais através do cruzamento de diferentes bases de dados — Censo Escolar 2011, SPAECE 2011 e Relação Anual de Informações Sociais, RAIS 2014 cujo procedimento permitiu identificar cada aluno concludente das escolas estaduais do Ceará em 2011, e sua trajetória para a inserção no mercado de trabalho em 2014. Porém, o fato de se observar apenas os jovens que estão inseridos no mercado de trabalho requereu que se conduzisse a correção da seletividade amostral através da aplicação dos procedimentos de Heckman (1979). Além disso, visando uma maior homogeneização da amostra, aplicou-se o método do balanceamento por entropia desenvolvido por Hainmueller (2012), o qual pondera as distribuições das variáveis a fim de satisfazer um conjunto de condições especiais predeterminadas.

Para verificar o efeito do ensino profissionalizante e dos cursos sobre as remunerações no mercado de trabalho, estimou-se a equação de rendimentos ponderada pelos pesos da entropia e corrigida pela razão inversa de Mills, e comparou-se com a mesma equação sem tais correções. Procedimento semelhante foi utilizado para as ocupações no mercado de trabalho, porém, dado o elevado número destas, aplicou-se os procedimentos de (Jannuzzi, 2001, 2003) para reagrupa-las em apenas cinco estratos sócio-ocupacionais. Em seguida, buscou-se estimar o impacto desse tipo de ensino e do curso escolhido sobre os cinco níveis de ocupações através de um modelo logit multinomial ponderado pela entropia e corrigido pela razão inversa de Mills.

Em sequência, o artigo está organizado com as seguintes seções: abordagem metodológica dividida em Seletividade Amostral e Balanceamento por Entropia; resultados e conclusões.

[2]Esse estado ganhou destaque em âmbito nacional, ao combinar o ensino integral com o profissionalizante e aumentar a participação de jovens no Exame Nacional de Ensino Médio – ENEM. Segundo informações disponibilizadas pelo INEP 2009 a 2014, este estado apresentou a 3ª maior participação de estudantes de escolas públicas estaduais no ENEM em 2014. A taxa de participação desses estudantes passou de 4,13% em 2009 para 7,91% em 2014, perfazendo a maior variação (92%) dentre os estados brasileiros.

## 2. Metodologia

Quando se pretende identificar o efeito do ensino profissionalizante adquirido durante o ensino médio sobre o mercado de trabalho utilizando pesquisas censitárias como bases de dados, deve-se levar em consideração diferentes fontes de vieses que lhes são inerentes, portanto, alguma ação deve ser tomada para corrigi-las, como a correção do problema de seletividade amostral. Ao fazer o cruzamento da lista de concludentes do ensino médio do Ceará com a Relação Anual de Informações Sociais (RAIS) como base de dados, deve-se considerar que as informações coletadas são fornecidas apenas para aqueles que estavam no mercado de trabalho formal no período da pesquisa, ou seja, os salários e as ocupações observados estão relacionados com a decisão de um indivíduo trabalhar ou não e esta pode estar correlacionada a fatores não observáveis que afetam seu rendimento.

Assim, quando se observa apenas os jovens que estão no mercado de trabalho, é possível que estes tenham sido pré-selecionados através de características que influenciaram esta decisão, e a não consideração dessa seleção pode levar a resultados duvidosos quando se pretende avaliar o efeito do ensino profissionalizante sobre os rendimentos, já que, o salário é observado apenas para quem ingressou no mercado de trabalho.

Em vista da composição da amostra aqui utilizada, há que se considerar incialmente a correção para solucionar o possível viés de seletividade amostral. Para tanto, aplica-se o procedimento de Heckman (1979) em dois estágios, o qual é formado por duas equações: a primeira descreve a participação no mercado de trabalho em 2014 inserindo todos os concludentes em 2011, ou seja, estejam ou não identificados no mercado de trabalho auferindo renda positiva; a outra equação segue a especificação minceriana, que considera somente os que possuem algum rendimento. No segundo estágio, acrescenta-se a variável Mills obtida através da estimação da primeira equação, a qual indicará se os indivíduos observados no mercado de trabalho foram ou não previamente selecionados.

Além de se verificar o efeito do ensino profissionalizante sobre os rendimentos, buscar-se-á também encontrar o seu impacto sobre o tipo de ocupação que o jovem, após o termino do ensino médio, ocupa no mercado de trabalho. Para tanto, será estimado um modelo logit multinomial com base nos cinco níveis de ocupações definidas em Jannuzzi (2001, 2003),[3] cujos procedimentos metodológicos encontram-se no Apêndice A.

Além do problema de seletividade amostral, existe uma fonte de viés inerente aos dados que está associado à variável que identifica os concludentes das escolas profissionalizantes, pois, dado o número limitado de escolas, não há vagas suficientes para atender todos os alunos que pretendem cursar o ensino médio no Ceará. Como decorrência, a Secretaria de Educação (SEDUC) estabelece critério para o processo seletivo de entrada em cada escola, o qual é baseado nas maiores médias das notas obtidas pelos pretendentes no ensino fundamental até o preenchimento das vagas. Os critérios de desempate, caso necessário, são a idade e a proximidade da residência com a escola. Dessa forma, não existe um valor pré-estabelecido da nota média mínima para garantia de uma vaga por parte do aluno, pois esta pode variar de acordo com as notas médias dos alunos inscritos em determinada escola, cuja oferta de vagas é limitada.

Cabe mencionar que existe a possibilidade do surgimento de vieses decorrentes de características não observáveis e observáveis associadas ao processo seleção. No primeiro caso poderia se inserir a não observação da característica relacionada à motivação do aluno

[3]Este procedimento será definido na seção 2.1.

em efetuar sua inscrição. Neste caso, uma forma de solucioná-lo seria considerar na análise todos os alunos inscritos em cada escola, distinguidos entre os que ingressaram e não ingressaram; porém, essa lista não é disponibilizada. A característica observável se relaciona à indisponibilidade de informações relativas ao desempenho dos alunos inscritos no ensino fundamental, pois, caso contrário, seria possível verificar a nota limite à entrada por cada escola e, dependendo do tamanho da amostra, poderia-se aplicar regressão descontínua com vários pontos de descontinuidade.

Tendo em vista a indisponibilidade dessas informações, optou-se por minimizar os vieses daí decorrentes através do paramento dos alunos por suas notas no desempenho do SPAECE e idades dos concludentes, cuja adequação é justificada pela suposição de haver uma correlação direta com a nota média no ensino fundamental. Ademais, como os candidatos têm conhecimento prévio que o processo seletivo se baseia no desempenho, acredita-se que essa variável seja um dos fatores que afetam a motivação para o aluno concorrer a uma vaga. Vale ressaltar que esse procedimento será utilizado, não para solucionar esses problemas, mas minimizá-los ao tornar a amostra mais homogênea. O procedimento escolhido para essa homogeneização amostral se baseia na aplicação do balanceamento por entropia[4] às variáveis proficiência em matemática e português e à variável idade.

Hainmueller e Xu (2013) descrevem o balanceamento por entropia como uma generalização da abordagem do pareamento por escore de propensão (PEP), embora com procedimentos distintos. Enquanto o PEP calcula os escores de propensão através de uma regressão logit ou probit para fazer o pareamento, e verificar se os pesos estimados equilibram as distribuições das covariadas, a entropia calcula os pesos diretamente para ajustar as distribuições amostrais conhecidas, integrando o balanceamento das covariáveis aos pesos. Embora ambos os procedimentos avaliem efeitos de tratamentos, o balanceamento por entropia também pode ser aplicado para ajustar a amostra, conforme demonstrado por Watson e Elliot (2016).

Segundo Hainmueller (2012), esse método permite ponderar um conjunto de dados, tais que, as distribuições das variáveis nas observações reponderadas satisfaçam um conjunto de condições especiais de momentos, de forma que exista equilíbrio exato sobre o primeiro (média), segundo (variância) e terceiro (assimetria) momentos das distribuições de variáveis independentes nos grupos de tratamento e controle. Assim, é possível se especificar um nível de equilíbrio desejável para as covariadas, usando um conjunto de condições associados aos momentos da distribuição. A vantagem deste método sobre os algoritmos logit/probit reside na capacidade de implementar diretamente o equilíbrio exato. Para tanto, será utilizada uma variável binária para o tratamento, a qual irá permitir encontrar o efeito das escolas profissionalizantes sobre a remuneração e as ocupações dos concludentes do ensino médio em 2011 no mercado de trabalho em 2014.

Para fazer o balanceamento por entropia, será utilizado como covariáveis as proficiências de matemática e português no SPAECE e a idade dos alunos, partindo do pressuposto que existe uma relação direta destas variáveis com o processo seletivo das escolas. Esse procedimento permitirá comparar indivíduos semelhantes em termos de desempenho escolar e idade. Assim, após equilibrar os grupos nessas características, espera-se que esses alunos egressos do ensino médio em 2011 que se encontram no mercado de trabalho em 2014 não possuam diferenças significativas em termos de habilidades escolares observadas e que estejam na mesma faixa etária.

[4]Especificação do método encontra-se no Apêndice A.

Tendo em vista que durante o ensino profissionalizante o aluno tem a opção de escolher um curso, dentre vários, para se especializar, optou-se por considerar desagregar a variável que indica se o jovem concluiu o ensino médio nas escolas profissionalizantes em variáveis binárias, as quais serão definidas de acordo com o curso escolhido durante o ensino profissionalizante presente na amostra, cuja descrição encontra-se na Tabela B-1 do Apêndice B.

## 2.1 Variáveis e Bases de Dados

Para avaliar o efeito da política de profissionalização no ensino médio das EEEP sobre o mercado de trabalho, construiu-se uma amostra com informações longitudinais a partir da junção das seguintes bases de dados: lista dos concludentes no ensino médio em 2011, SPAECE 2011, Censo Escolar 2011, e Relação Anual de Informações Sociais (RAIS) 2014. Dessa forma, foram identificados os alunos concludentes das escolas estaduais do Ceará em 2011,[5] e sua inserção no mercado de trabalho em 2014. Vale ressaltar que, como as informações da RAIS são de 2014, optou-se por identificar os egressos do ensino médio em 2011, para permitir uma defasagem temporal a fim de evitar o eventual impedimento para a entrada no mercado de trabalho devido à idade do jovem ou tempo gasto por procura de emprego.

Para compor a amostra, observou-se inicialmente a lista dos concludentes disponibilizada pela SEDUC, dos quais 7.606 concluíram em 2011, fez-se o cruzamento com o censo escolar e com o SPAECE para obter informações sobre características pessoais e escolares desses alunos, tais como, sexo, raça, idade, escola de conclusão, turma, curso de profissionalização, localização da escola, desempenho dos alunos nas proficiências de matemática e português, dos quais resultaram 6.796 concludentes com essas informações disponíveis. Destes, 3.830 estudaram nas escolas de ensino profissionalizante e 2.966 das escolas regulares.

Em seguida, a partir da RAIS 2014, identificou-se que, 1.956 (51,07%) e 1.032 (34,7%) alunos advindos das escolas profissionalizantes e regulares, respectivamente, foram para o mercado de trabalho, totalizando 2.988 alunos,[6] os quais constituem a amostra resultante para análise. A Tabela C-3 e os gráficos da Figura C-1 e Figura C-2 apresentam uma caracterização dos concludentes referentes aos desempenhos em português e matemática, para os que foram ou não para o mercado de trabalho. Destaca-se que os formados pelas escolas profissionalizantes em 2011, e que foram para o mercado de trabalho, possuem desempenho levemente inferior aos que concluíram nessas escolas, mas não estavam presentes na RAIS 2014, ao contrário do que se observa nos graduados das escolas regulares.

Para compor a análise do modelo (pareamento por entropia), cuja variável de tratamento é binária, foram criados os seguintes grupos: (a) Tratados — alunos que se profissionalizaram nesse período, estavam presentes no mercado de trabalho em 2014, em que foram observados 1.956 do ensino profissional; (b) Controle — alunos advindos do ensino regular, perfazendo um total de 1.032 alunos. Cabe destacar que os cursos profissionalizantes observados na amostra estão assim distribuídos: saúde (40%), informação (33,2%), gestão e negócios (10%), Hospitalidade e lazer (11,1%) e outros (5,7%).[7]

[5]Identificadores foram criados pela SEDUC.

[6]Aproximadamente 44% dos concludentes.

[7]Para mais detalhes ver Figura C-3 no Apêndice C.

As variáveis que representam as ocupações, descritas na Tabela B-1 do Apêndice B, serão utilizadas para verificar o impacto das escolas e dos cursos profissionalizante sobre o mercado de trabalho. Tendo em vista o elevado número de 596 ocupações criadas a partir da Classificação Brasileira de Ocupações (CBO 2002), e sua intratabilidade estatística, utilizou-se o procedimento de Jannuzzi (2001, 2003) para reagrupa-las em apenas cinco estratos sócio-ocupacionais, a partir de indicadores de renda, escolaridade e nível de precarização ocupacional, onde estes foram organizados com base no índice socioeconômico posicional (Isep),[8] na ordenação das ocupações proporcionada pelos índices socioeconômicos, na tipologia de grupos ocupacionais de Valle Silva (1992) e na proposta metodológica de estratificação social segundo grupos ocupacionais de Goldthorpe (1992).

O primeiro estrato compreende as ocupações em posto de comando, de direção ou com especialização técnica superior. O segundo estrato (Isep de 85%) é composto pelos pequenos proprietários, chefes e supervisores e empregados qualificados de escritório e técnicos de média especialização na indústria e serviços. O terceiro estrato (Isep de 69%) reúne a maioria de ocupações no Comércio, Serviços e postos qualificados da Indústria. O quarto grupo socioocupacional (Isep de 45%) compreende os empregados em prestação de serviços de baixa qualificação nos serviços, construção civil e indústria tradicional. Na última categoria (Isep de 18%), estão agrupados os trabalhadores rurais, além das ocupações urbanas de baixo status, ou seja, as que possuem remunerações e escolaridade mais baixas e elevada precariedade no mercado de trabalho.[9]

Para estimações das equações de participação no mercado de trabalho (Estimação I), dos rendimentos e as ocupações (Estimação II), foram utilizadas diferentes covariadas, tais como descritas na Tabela B-2 do Apêndice B.

# 3. Resultados

## 3.1 Análise Descritiva

Estatísticas descritivas das variáveis de resultados e de controles relacionadas aos alunos egressos de escolas profissionalizantes (EEEP) e das escolas regulares em 2011 estão explicitadas na Tabela 1. Observa-se que a remuneração média em 2014 dos alunos advindos das EEEP é superior aos alunos que concluíram nas escolas regulares em apenas 4,4%.

Referentes aos tipos de ocupações socioeconômicas, verifica-se que os concludentes das EEEP se concentram, principalmente, na categoria 2 com 48,1%, seguida das categorias 3 e 4 com 36,2% e 10,54%, respectivamente. Para os egressos das escolas regulares, a categoria 3 detém a maior proporção (39,84%), seguida da categoria 4 (29,25%). Por outro lado, a categoria 1 concentra as mais baixas proporções de trabalhadores de ambas as escolas,4,6% para as EEEP e 2,52% para as regulares.

Apesar dessas estatísticas apresentarem valores que favorecem aos alunos que concluíram nas escolas de ensino profissional, não se pode afirmar que o efeito destas sobre a remuneração e os tipos de ocupações no mercado de trabalho sejam positivas, pois o grupo de alunos advindos das escolas regulares não representam necessariamente o contrafactual de não tratamento. Dessa forma, se faz necessário técnicas mais apropriadas que busquem isolar o efeito desta política sobre as variáveis de resultado.

[8] Corresponde à porcentagem de indivíduos ocupados cujos níveis médios combinados de rendimento e escolaridade são menores ou iguais aos da ocupação considerada.

[9] Para mais detalhes ver Jannuzzi (2001, 2003).

**Tabela 1.** Estatística descritiva das variáveis da amostra por tipos de escolas.

| | Alunos egressos das EEEP | | | Alunos egressos das Regulares | | |
|---|---|---|---|---|---|---|
| Variáveis | Média | Proporção (%) | Desvio Padrão | Média | Proporção (%) | Desvio Padrão |
| *Mercado de trabalho em 2014* | | | | | | |
| Renda_media | 885,65 | – | 324,63 | 848,12 | – | 314,31 |
| Ocupação1 | – | 4,6 | 0,2097 | – | 2,52 | 0,1570 |
| Ocupação2 | – | 48,18 | 0,4997 | – | 26,33 | 0,4406 |
| Ocupação 3 | – | 36,20 | 0,4807 | – | 39,84 | 0,4898 |
| Ocupação 4 | – | 10,54 | 0,3072 | – | 29,25 | 0,4551 |
| Ocupação 5 | – | 0,48 | 0,0677 | – | 2,06 | 0,1414 |
| *Características dos Alunos em 2011* | | | | | | |
| Sexo – Masculino | – | 47,54 | 0,4995 | – | 56,87 | 0,4954 |
| Raça – Branca | – | 16,71 | 0,3732 | – | 13,46 | 0,3415 |
| Zona Urbana | – | 87,83 | 0,4995 | – | 50,67 | 0,5001 |
| Proficiência_MT | 299,2016 | – | 50,6841 | 268,6783 | – | 51,2099 |
| Proficiência_PT | 289,8749 | – | 41,4428 | 261,7721 | – | 46,4631 |
| Idade1 | – | 99,13 | 0,0928 | – | 89,53 | 0,3062 |
| Idade2 | – | 0,066 | 0,0812 | – | 9,20 | 0,2892 |
| Bolsa_familia | – | 55,36 | 0,4972 | – | 66,95 | 0,4705 |
| Escolaridade pai | – | 33,69 | 0,4727 | – | 48,64 | 0,5000 |
| Escolaridade mãe | – | 29,11 | 0,4544 | – | 42,58 | 0,4947 |

*Notas:* A) Ocupações: 1 – médicos, engenheiros, professores universitários, empresários, gerentes e postos superiores na administração pública (juízes, promotores, delegados, oficiais das forças armadas, etc.); 2 – técnicos de contabilidade e administração, mestre e contramestres na indústria, professores de ensino fundamental e médio, corretores de imóveis, inspetores de polícia, carteiros, comerciantes (proprietários) e agricultores; 3 – torneiro mecânico, montadores de equipamentos elétricos, vendedores, operadores de caixa, comerciantes conta-própria, professores de ensino pré-escolar, motoristas, inspetores de alunos, auxiliares de enfermaria, auxiliares administrativos e de escritório, policiais e praças das forças armadas; 4 – ocupações da indústria de alimentos, da indústria têxtil, pedreiros, pintores, garçons, vigias, porteiros, estivadores; 5 – trabalhadores rurais, além das ocupações urbanas de baixo status, como a de serventes de pedreiro, lavadeiras, empregadas domesticas e lixeiros. B) Covariadas: MT = matemática; LP = língua portuguesa; Idade1 = faixa etária 17–25; Idade2 = faixa etária 26–40; Escolaridade pai e Escolaridade mãe: se o pai(mãe) não estudou ou se possui até a 4° série do ensino fundamental.

*Fonte:* Elaboração dos autores com base nos dados da RAIS (2014); SPAECE (2011) e Censo escolar (2011).

Comparando estes grupos referentes às demais variáveis, observa-se, ainda na Tabela 1, que a proporção de homens advindos das escolares regulares é superior aos das profissionalizantes, em aproximadamente 20%. Quanto à raça e à zona residencial, a proporção de alunos considerados brancos e que moram em áreas urbanas são maiores nas EEEP, 24% e 73%, respectivamente.

Com relação às variáveis de proficiências, matemática e português, os alunos das escolas de ensino profissionalizante apresentaram desempenho superior em, aproximadamente, 11% comparados aos alunos das escolas regulares. Ademais, 99% desses alunos apresentavam idade entre 17 e 25 anos, enquanto que no outro grupo, 89% estava contido nessa faixa etária. Analisando as variáveis sobre background familiar, observa-se que 55,3% dos alunos das EEEP estão inseridos em famílias que recebem o bolsa família, enquanto que, para os alunos das regulares, essa proporção aumenta para, aproximadamente, 67%. Quanto à herança educacional, pais de alunos das EEEP possuem, proporcionalmente, grau de educação superior do que os de escolas regulares.

A Tabela 2 apresenta as estatísticas descritivas das variáveis de resultado no mercado de trabalho em 2014, remuneração e tipos de ocupação, relativas aos cursos escolhidos

**Tabela 2.** Estatística descritiva das variáveis de resultados no mercado de trabalho em 2014 por Curso Concluído em 2011.

| Cursos dos egressos das EEEP em 2011 | Renda Média (R$) em 2014 | Ocupações no Mercado de Trabalho – 2014 (%) | | | | |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 |
| Saúde | 924,06 | 2,30 | 55,11 | 32,48 | 9,46 | 0,65 |
| Gestão | 840,39 | 8,76 | 38,65 | 42,26 | 10,33 | – |
| Informação | 869,90 | 6,81 | 46,13 | 35,75 | 10,83 | 0,48 |
| Hospitalidade e Lazer | 831,62 | 4,12 | 38,99 | 44,05 | 12,84 | – |
| Outros | 891,27 | 1,76 | 46,03 | 38,95 | 12,38 | 0,88 |

*Nota:* Ocupações: 1 – médicos, engenheiros, professores universitários, empresários, gerentes e postos superiores na administração pública (juízes, promotores, delegados, oficiais das forças armadas, etc.); 2 – técnicos de contabilidade e administração, mestre e contramestres na indústria, professores de ensino fundamental e médio, corretores de imóveis, inspetores de polícia, carteiros, comerciantes (proprietários) e agricultores; 3 – torneiro mecânico, montadores de equipamentos elétricos, vendedores, operadores de caixa, comerciantes conta-própria, professores de ensino pré-escolar, motoristas, inspetores de alunos, auxiliares de enfermaria, auxiliares administrativos e de escritório, policiais e praças das forças armadas; 4 – ocupações da indústria de alimentos, da indústria têxtil, pedreiros, pintores, garçons, vigias, porteiros, estivadores; 5 – trabalhadores rurais, além das ocupações urbanas de baixo status, como a de serventes de pedreiro, lavadeiras, empregadas domesticas e lixeiros.

*Fonte:* Elaboração dos autores com base nos dados da RAIS (2014); SPAECE (2011) e Censo escolar (2011).

pelos egressos das EEEP em 2011. Verifica-se que as áreas de saúde provem a maior remuneração média (R$924,06) seguida de outros cursos (R$891,27), e cursos na área de gestão (R$869,90).

Dentre os cursos profissionalizantes, a área de gestão apresenta a maior proporção de trabalhadores situados na ocupação socioeconômica do tipo 1, 8,76%, seguido das áreas de informação e hospitalidade e lazer, com 6,81% e 4,12%, respectivamente. Referente a ocupação 2, destaca-se os cursos da área de saúde, com 55,11% da proporção dos concludentes das profissionalizantes, enquanto que os demais cursos apresentam proporções que variam entre 38,65% a 46,13%. Na ocupação 3, a área com maior proporção são os cursos da área de Hospitalidade e Lazer, com 44,05%, seguido dos outros tipos de cursos, com 38,95%, assim como para a categoria de ocupação 4, 12,84% e 12,38%, respectivamente, porém, na categoria 5, as áreas de saúde, informação e outros apresentam proporção semelhante, menos de 1%, enquanto as áreas de gestão e hospitalidade não apresentaram proporções de concludentes.

Cabe destacar que dentre os alunos formados nas áreas de saúde, 55,11% encontram-se nas ocupações de categoria 2, seguido da ocupação 3, com 32%, enquanto as ocupações 1, 4 e 5 perfazem os restantes 22,41%. Resultado semelhante, porém, com menor proporção na categoria 2, situa-se a área de informação e outros tipos de cursos, enquanto que as áreas de hospitalidade e lazer e gestão apresentam maiores proporções de trabalhadores nas ocupações socioeconômicas 3, seguidas da categoria 2.

## 3.2 Resultados dos Modelos

Ao propor verificar o efeito do ensino profissionalizante sobre variáveis relacionadas ao mercado de trabalho, faz-se necessário a correção do problema de seletividade amostral, uma vez que são observados apenas os jovens egressos das escolas estaduais do Ceará em 2011 que estão presentes na RAIS 2014. Ou seja, existe uma escolha feita pelos jovens após a conclusão do ensino médio que deve ser levada em consideração, em que, a tomada de decisão em ir para o mercado de trabalho formal pode ser ter sido afetada por diversas características observadas ou não durante o período anterior a esta escolha, por exemplo,

nem todo jovem egresso se sente preparado em ingressar no mercado de trabalho, alguns tiveram acesso ao ensino médio técnico que os propiciou a conseguir um emprego com mais facilidade; outros podem optar por ingressar no ensino superior e outros, simplesmente, não querem trabalhar. Além disso, as condições socioeconômicas da família podem influenciar essa decisão também.

Como primeira etapa deste modelo, especificou-se uma variável binária de participação no mercado de trabalho e calculou-se a razão inversa de Mills, utilizando como variáveis explicativas sexo, idade no período de conclusão do ensino médio, tipo de ensino utilizado na escola, zona de residência, variável indicativa referente ao bolsa família, escolaridade do pai e da mãe, cujos resultados podem ser observados na Tabela C-4 no Apêndice C. Destes resultados, cabe ressaltar que, o efeito do ensino profissionalizante sobre a participação no mercado de trabalho foi positivo e significativo, indicando que alunos advindos dessas escolas apresentam maiores chances de estarem inseridos em trabalhos formais em 2014.

Antes de estimar o efeito deste tipo de ensino sobre a renda e a categoria de ocupação no mercado de trabalho, optou-se por fazer um balanceamento por entropia, utilizando como covariáveis as proficiências de matemática e português no SPAECE e a idade dos alunos, partindo do pressuposto que existe uma relação direta destas variáveis com o processo seletivo das escolas. Esse procedimento permitirá comparar indivíduos semelhantes em termos de desempenho escolar e idade. Assim, após equilibrar os grupos, alunos que concluíram nas escolas profissionalizante e alunos advindos das escolares regulares, nessas características, espera-se que esses alunos egressos do ensino médio em 2011 que se encontram no mercado de trabalho em 2014 não possuam diferenças significativas em termos de habilidades escolares observadas e que estejam na mesma faixa etária. Para tanto, considerou-se a variável binária de tratamento para os concludentes de escolas profissionalizantes (tratados) e regulares (controle), referentes aos alunos que terminaram o ensino médio em 2011. Algumas estatísticas das covariadas nos dois grupos estão expostas na Tabela 3, a fim de verificar, inicialmente, se os grupos estão equilibrados e se divergem em características, para em seguida encontrar-se o efeito.

Uma vez que existem diferenças significativas na média, variância e assimetria dessas variáveis, conclui-se que os grupos não se apresentam adequadamente balanceados, requerendo-se, portanto, a aplicação do balanceamento por entropia, a fim de equilibrar as características entre os grupos.

O algoritmo busca equilibrar os valores dos momentos especificados na amostra de referência, neste caso, o grupo dos concludentes das profissionalizantes, e procura, a partir de um conjunto de pesos de entropia, ajustar o grupo de controle combinando-o com a amostra de referência. Esta aplicação tem por objetivo encontrar pesos que minimizem

**Tabela 3.** Condições de momentos das covariadas antes do balanceamento por entropia.

| | Tratados – Escolas EEEP | | | Controles – Regulares | | | Diferenças | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Covariadas | Média | Var | Ass. | Média | Var | Ass. | Média | Var | Ass. | p-valor |
| Proficiência_MT | 299,201 | 2,569 | −0,0914 | 268,678 | 2,622 | 0,2827 | 30,523 | −53 | −0,3741 | 0,0000 |
| Proficiência LP | 289,874 | 1,718 | −0,4711 | 261,772 | 2,159 | −0,1086 | 28,102 | −441 | −0,3625 | 0,0000 |
| Idade1 | 0,9913 | 0,0086 | −10,59 | 0,8953 | 0,0937 | −2,583 | 0,0959 | −0,0851 | −8,007 | 0,0000 |
| Idade2 | 0,0066 | 0,0066 | 12,14 | 0,0920 | 0,0836 | 2,822 | −0,0854 | −0,0770 | 9,318 | 0,0000 |

*Notas:* A) Var: Variância; Ass: Assimetria. B) Covariadas: MT = matemática; LP = língua portuguesa; Idade1 = faixa etária 17–25; Idade2 = faixa etária 26–40.

as diferenças entre tratados e controles, considerando os três momentos da distribuição. A convergência ocorre quando todos os momentos são pareados entre os grupos dentro de um determinado número de interações e nível de tolerância.[10]

Com base nos resultados após o balanceamento por entropia apresentados na Tabela 4, se pode inferir que, em todas as covariadas, os três momentos da distribuição não diferem entre os grupos, ou seja, não há diferenças estatísticas significativas entre eles. Dessa forma, conclui-se que os grupos estão perfeitamente equilibrados quanto às variáveis utilizadas na amostra, após a entropia, ou seja, o grupo formado pelos alunos das escolas regulares tornaram-se semelhantes ao grupo dos alunos das escolas profissionalizante em desempenho escolar medido pelas proficiências em matemática e português e em faixa etária. Assim, quando se estimar a regressão de rendimentos ou de ocupações considerando o balanceamento, as diferenças entre esses grupos são se darão por essas características.

Uma vez encontrado os pesos que tornam a amostra mais homogênea, estimou-se o efeito do ensino profissionalizante sobre os rendimentos, considerando a razão inversa de Mills para a verificação do problema de seletividade amostral e as variáveis sexo, raça e idade. Ademais, como existe a possibilidade de escolher um tipo de curso no qual se pretenda se especializar, estimou-se a equação de rendimentos utilizando variáveis binárias para cada curso presente na amostra, cujos resultados estão expostos na Tabela 5.

Ao considerar as correções no modelo referentes aos problemas de seleção tanto no mercado de trabalho quanto nas escolas, o efeito do ensino profissionalizante, independente

**Tabela 4.** Condições de Momentos das covariadas após o balanceamento por entropia.

| | Tratados – Escolas EEEP | | | Controles – Regulares | | | Diferenças | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Covariadas | Média | Var | Ass. | Média | Var | Ass. | Média | Var | Ass. | p-valor |
| Proficiência_MT | 299,201 | 2,569 | −0,0914 | 299,201 | 2,569 | −0,09144 | 0 | 0 | 4E-05 | 0,999 |
| Proficiância LP | 289,874 | 1,718 | −0,4711 | 289,874 | 1,718 | −0,4711 | 0 | 0 | 0 | 0,998 |
| Idade1 | 0,9913 | 0,00862 | −10,59 | 0,9913 | 0,008626 | −10,59 | 0 | −6E-06 | 0 | 0,999 |
| Idade2 | 0,00664 | 0,006605 | 12,14 | 0,00665 | 0,006613 | 12,14 | −1E-05 | −8E-06 | 0 | 0,978 |

*Notas:* A) Var: Variância; Ass: Assimetria. B) Covariadas: MT = matemática; LP = língua portuguesa; Idade1 = faixa etária 17–25; Idade2 = faixa etária 26–40.

**Tabela 5.** Estimação das Equações de Rendimentos em 2014 com Correções – EEEP e por Curso Concluído em 2011.

| Variáveis | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão |
|---|---|---|---|---|
| EEEP | −14,4464 | 38,4818 | – | – |
| Saúde | – | – | 35,75 | 40,46 |
| Gestão | – | – | −65,75 | 41,89 |
| Informação | – | – | −37,94 | 38,82 |
| Hospitalidade e Lazer | – | – | −66,62 | 43,75 |
| Outros | – | – | −19,86 | 49,84 |
| Mills | −256,41 | 110,64 | −255,16** | 111,29 |
| N° observações | | 2.980 | | |

**Coeficiente significante a 5%.

[10] O número de iteração e o nível de tolerância padrão são 20 e 0,015, respectivamente, e podem ser aumentados se a convergência falhar.

da escolha do curso, não se mostrou significativo sobre os rendimentos no mercado de trabalho em 2014. Mesmo após desagregar esse ensino nos tipos de cursos escolhido pelos alunos, saúde, gestão, informação, hospitalidade e lazer, e outros, o efeito permanece insignificante. Porém, observa-se que a razão inversa de Mills foi significativa nas duas análises, indicando que a não inclusão desta variável em equações de rendimentos fornecerá estimativas viesadas. Dessa forma, é possível afirmar que os indivíduos que estão no mercado de trabalho em 2014 foram pré-selecionados se comparados aos egressos que não estão trabalhando no emprego formal, ou seja, apresentam características comuns que levaram a tomar essa decisão de trabalhar.

Para efeitos de comparação, considerou-se também a estimação essas equações sem as correções, cujos resultados podem ser observados na Tabela 6. Caso as correções amostrais — razão inversa de Mills e o balanceamento por entropia — não fossem consideradas, o efeito do ensino profissionalizante sobre os rendimentos seria positivo e significante, com um retorno de R$61,15, equivalente a um aumento de 7,22% se considerar a média dos rendimentos no mercado de trabalho em 2014 dos alunos que concluíram nas escolas regulares em 2011, R$848,13. Os resultados para a equação de rendimentos sem correção corroboram com o estudo de Amoroso Amoroso Neto et al. (2017), que avaliaram o efeito de cursos de educação profissional no Brasil, em 2014, sobre os ganhos salariais dos egressos que concluíram o ensino médio através da estimação do método de pareamento por escore de propensão e encontraram uma estimativa positiva e significativa de 19% para o retorno do curso de ensino médio técnico. Todavia, os autores não consideraram variáveis de background educacional dos alunos, e não levaram em consideração o problema de seleção ao identificar somente os alunos egressos do ensino superior que estavam no mercado de trabalho, procedimento similar feito por este trabalho para o modelo sem correções.

Os resultados aqui apresentados fornecem fortes indícios de que a não consideração destes procedimentos corretivos — balanceamento dos grupos pelas proficiências em matemática e português e a não correção da seletividade amostral — tendem aviesar os resultados e, consequentemente, podem induzir os gestores de política educacional a tomadas de decisão inadequadas.

Além do efeito sobre os rendimentos, buscou-se avaliar o impacto desse tipo de ensino e do curso escolhido sobre os tipos de ocupações presentes no mercado de trabalho em 2014 ocupadas pelos concludentes do ensino médio em 2011. O Painel A da Tabela 7 apresenta

**Tabela 6.** Estimação das Equações de Rendimentos em 2014 sem Correções – EEEP e por Curso Concluído em 2011.

| Variáveis | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão |
|---|---|---|---|---|
| EEEP | 61,1481*** | 12,2535 | – | – |
| Saúde | – | – | 112,2651*** | 15,7828 |
| Gestão | – | – | 12,8299 | 18,4121 |
| Informação | – | – | 35,4416** | 17,0833 |
| Hospitalidade e Lazer | – | – | 9,4406 | 21,9208 |
| Outros | – | – | 56,9815* | 31,6416 |
| N° observações | | 2.988 | | |

*Nota:* Níveis de significância dos coeficientes: *10%; **5%; ***1%.

os resultados dos efeitos marginais após as estimações do modelo logit multinomial[11] considerando o ensino profissionalizante independente do curso escolhido na primeira análise e os tipos de cursos escolhidos pelos concludentes na segunda análise, cujos resultados estão expostos na Tabela C-5 e no Painel A da Tabela C-6, no Apêndice C.

Vale ressaltar que, para obtenção destas estimativas, aplicou-se os procedimentos da correção de Heckman e o balanceamento pela entropia. Dos resultados, observa-se que, para todas as categorias de ocupações no mercado de trabalho, o efeito do ensino profissionalizante, independente do curso, não foi significativo. Porém, ao considerar os tipos de cursos, apenas a área da saúde apresentou efeito significativo sobre a categoria de ocupação de nível 3, indicando que, fazer esse curso durante o ensino médio reduz em 8,44% a probabilidade de o jovem estar inserido nessa ocupação.

Assim como para a equação de rendimentos, foram também estimados os efeitos marginais das EEEP e dos cursos sobre os tipos de ocupações sem aplicar os procedimentos das correções. Para tanto, estimou-se o modelo multinomial sem a razão inversa de mills e sem ponderar pelos pesos da entropia, cujos resultados estão expostos nos painéis B e C da Tabela C-6, no Apêndice C. O Painel B da Tabela 7 mostra que, desconsiderando as correções, o efeito do ensino profissionalizante passa a ser significativo em todas as categorias ocupacionais, aumentando a probabilidade de o jovem pertencer às de níveis 1 e 2, porém, reduzindo nas de níveis 3 e 4. Resultado semelhante pode ser observado quando se desagrega esse tipo de ensino pelo curso escolhido nas áreas de saúde e informação, exceto para os níveis 1 e 3, respectivamente. Para o curso na área de gestão, o efeito seria significativo, positivo no nível 1 e negativo no nível 3. No curso de hospitalidade e lazer, o efeito teria sido significante apenas para o nível 4 das categorias ocupacionais. Semelhante a equação de rendimentos, esses resultados reforçam que a não consideração das correções tendem a gerar resultados viesados.

## 4. Conclusões

Espera-se que este trabalho contribua com o debate na literatura que relaciona educação e mercado de trabalho, ao trazer novas evidências, em nível longitudinal, que associam os alunos de escolas de ensino profissionalizante e suas inserções no mercado de trabalho comparativamente com os de escolas regulares. Isso é feito com procedimentos amostrais inéditos.

Uma das contribuições deste estudo reside na base de dados utilizada para mensurar este efeito, pois através da construção de identificadores providos pela SEDUC foi possível construir uma amostra com informações longitudinais a partir da junção das seguintes bases de dados: Censo Escolar (2.011), SPAECE (2011) e RAIS (2014). Dessa forma, foram identificados os alunos concludentes das escolas estaduais do Ceará em 2011, e sua inserção no mercado de trabalho em 2014.

Ao considerar apenas os jovens egressos das escolas estaduais do Ceará em 2011 que estão presentes na RAIS 2014 para composição da amostra, surge o inevitável problema de seletividade amostral, o qual se fez necessário corrigi-lo através dos procedimentos de Heckman (1979) a fim de eliminar vieses na estimação. Como primeira etapa deste modelo, estimou-se a equação de participação no mercado de trabalho, e verificou-se que o efeito do ensino profissionalizante foi positivo e significativo. A partir desta, calculou-se a razão

[11] Utilizou-se a categoria de ocupação 5 como base para a estimação.

**Tabela 7.** Efeito Marginal das EEEP e dos Cursos em 2011 sobre Tipos de Ocupações no Mercado de Trabalho em 2014.

| | Egressos das EEEP em 2011 | | Egressos do Curso na área Saúde em 2011 | | Egressos do Curso na área Gestão em 2011 | | Egressos do Curso na área Informação em 2011 | | Egressos do Curso na área Hospitalidade e lazer em 2011 | |
|---|---|---|---|---|---|---|---|---|---|---|
| Ocupações no Mercado de Trabalho em 2014 | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão |
| ***Painel A: Com correções*** | | | | | | | | | | |
| 1 | 0,0142 | 0,0242 | −0,0046 | 0,0174 | 0,0670 | 0,0466 | 0,0451 | 0,0288 | 0,0170 | 0,0277 |
| 2 | 0,0201 | 0,0638 | 0,0518 | 0,0467 | −0,0824 | 0,0501 | 0,0062 | 0,0461 | −0,0805 | 0,0501 |
| 3 | −0,0757 | 0,0603 | −0,0844** | 0,0421 | 0,0006 | 0,0541 | −0,0611 | 0,0424 | 0,0120 | 0,0526 |
| 4 | 0,0278 | 0,0375 | 0,0299 | 0,0363 | 0,0168 | 0,0457 | 0,0062 | 0,0332 | 0,0537 | 0,0493 |
| N° de Observações | | | | | 2.974 | | | | | |
| ***Painel B: Sem correções*** | | | | | | | | | | |
| 1 | 0,0211*** | 0,0069 | −0,0037 | 0,0105 | 0,0688** | 0,0284 | 0,0416*** | 0,0141 | 0,0202 | 0,0216 |
| 2 | 0,2070*** | 0,0187 | 0,2307*** | 0,0300 | 0,0665 | 0,1081 | 0,1736*** | 0,0520 | 0,0781 | 0,1145 |
| 3 | −0,0446** | 0,0196 | −0,1044** | 0,0498 | −0,0262 | 0,1046 | −0,0881 | 0,0637 | −0,0066 | 0,1113 |
| 4 | −0,1717*** | 0,0161 | −0,1218*** | 0,0281 | −0,1064** | 0,0429 | −0,1259*** | 0,0325 | −0,0885* | 0,0459 |
| N° de Observações | | | | | 2.982 | | | | | |

*Notas:* Ocupações: 1 – médicos, engenheiros, professores universitários, empresários, gerentes e postos superiores na administração pública (juízes, promotores, delegados, oficiais das forças armadas, etc.); 2 – técnicos de contabilidade e administração, mestre e contramestres na indústria, professores de ensino fundamental e médio, corretores de imóveis, inspetores de polícia, carteiros, comerciantes (proprietários) e agricultores; 3 – torneiro mecânico, montadores de equipamentos elétricos, vendedores, operadores de caixa, comerciantes conta-própria, professores de ensino pré-escolar, motoristas, inspetores de alunos, auxiliares de enfermaria, auxiliares administrativos e de escritório, policiais e praças das forças armadas; 4 – ocupações da indústria de alimentos, da indústria têxtil, pedreiros, pintores, garçons, vigias, porteiros, estivadores. Níveis de significância dos coeficientes: *10%; **5%; ***1%.

inversa de Mills, utilizando como variáveis explicativas sexo, idade no período de conclusão do ensino médio, tipo de ensino utilizado na escola, zona de residência, variável indicativa referente ao bolsa família, escolaridade do pai e da mãe, das quais, apenas idade não foi significativa para determinação da decisão do jovem de ingressar no mercado de trabalho.

Ademais, visando obter uma maior homogeneidade da amostra, antes de estimar o efeito do ensino profissionalizante sobre a renda e sobre o tipo de ocupação no mercado de trabalho, este trabalho optou por aplicar o método do balanceamento por entropia desenvolvido por Hainmueller (2012), o qual permite ponderar um conjunto de dados, tais que, as distribuições das variáveis nas observações reponderadas satisfaçam um conjunto de condições especiais de momentos, de forma que exista equilíbrio exato sobre o primeiro (média), segundo (variância) e terceiro (assimetria). Para tanto, utilizou-se como covariáveis, as proficiências de matemática e português no SPAECE e a idade dos alunos, partindo do pressuposto que existe uma relação direta destas variáveis com o processo seletivo das escolas.

Observou-se que, antes de aplicar o pareamento por entropia, existiam diferenças significativas na média, variância e assimetria entre os grupos de tratados e controle. Dessa forma, os grupos não se apresentam adequadamente balanceados, porém, após a aplicação desse método, todas as covariadas, nos três momentos da distribuição, não diferiram entre os grupos. Ou seja, o grupo formado pelos alunos das escolas regulares tornaram-se semelhantes ao grupo dos alunos das escolas profissionalizante em faixa etária e em desempenho escolar medido pelas proficiências em matemática e português.

Dessa forma, para verificar o efeito do ensino profissionalizante, independente do curso, sobre as remunerações no mercado de trabalho, estimou-se a equação de rendimentos ponderada pelos pesos da entropia e corrigida pela razão inversa de Mills e para fins de comparação, estimou-se a mesma equação, porém sem as devidas correções. Dos resultados, constatou-se que, ao considerar as correções no modelo, o efeito do ensino profissionalizante, independente da escolha do curso, não se mostrou significativo sobre os rendimentos no mercado de trabalho em 2014. Mesmo após desagregação desse ensino nos tipos de cursos escolhido pelos alunos, saúde, gestão, informação, hospitalidade e lazer, e outros, o efeito permaneceu insignificante. Porém, observa-se que a razão inversa de Mills foi significativa nas duas análises, indicando que a não inclusão desta variável em equações de rendimentos forneceria estimativas viesadas. Além disso, se não considerar as correções amostrais, razão inversa de mills e o balanceamento por entropia, o efeito do ensino profissionalizante sobre os rendimentos seria positivo e significante, assim como, os cursos nas áreas de saúde, informação e outros.

Considerando os diferentes tipos de ocupações exercidas pelos concludentes no mercado de trabalho em 2014, utilizou-se o procedimento de Jannuzzi (2001, 2003) para reagrupa-los em apenas cinco estratos sócio-ocupacionais, a partir de indicadores de renda, escolaridade e nível de precarização ocupacional. Em seguida, para avaliar o impacto desse tipo de ensino e do curso escolhido sobre os cinco níveis de ocupações, estimou-se um modelo logit multinomial para obtenção dos efeitos marginais, levando em consideração a correção de Heckman e o balanceamento pela entropia. Dos resultados, observa-se que, para todas as categorias de ocupações no mercado de trabalho, o efeito do ensino profissionalizante, independente do curso, não foi significativo. Porém, ao considerar os tipos de cursos, apenas a área da saúde apresentou efeito significativo sobre a categoria de ocupação de nível 3, indicando que, fazer esse curso durante o ensino médio reduz a probabilidade do jovem está inserido nessa ocupação em 8,44%.

Assim como para a equação de rendimentos, optou-se por estimar os efeitos marginais das EEEP e dos cursos sobre os tipos de ocupações sem aplicar os procedimentos das correções. Os resultados mostraram que o efeito do ensino profissionalizante passou a ser significativo em todas as categorias ocupacionais, aumentando a probabilidade do jovem pertencer as categorias de níveis mais elevados, e reduz a probabilidade de inserção nas categorias intermediárias e menos qualificadas. Quando se desagrega por área cursada no ensino profissionalizante, alguns cursos, tais como gestão, saúde e informação, se mostrariam significantes em elevar a probabilidade de inserção em categorias com melhores remunerações.

Cursar disciplinas técnicas durante o ensino médio nas escolas estaduais de ensino profissionalizante no Ceará eleva a probabilidade de inserção dos seus egressos no mercado de trabalho quando comparado aos alunos que concluíram nas escolas regulares. Porém, após aplicar os procedimentos de correção necessários, ou seja, considerar que os indivíduos que estão presentes no mercado de trabalho são pré-selecionados e que os egressos das escolas profissionalizantes foram selecionados via proficiência e idade, os resultados forneceram fortes indícios que o ensino público profissionalizante não se mostra significativo em elevar os rendimentos ou melhorar os níveis de ocupações dos jovens egressos do ensino médio que já se encontram no mercado de trabalho. Nesse sentido, cabe realçar que negligenciar tais procedimentos incorreria-se na provisão de resultados viesados e contraditórios, portanto, inadequados para fins de política educacional vis a vis o mercado de trabalho.

## Referências bibliográficas


Amoroso Neto, V., Menezes-Filho, N., & Komatsu, B. K. (2017). *Os efeitos da educação profissional e do PRONATEC sobre os salários* (Policy Paper Nº 25). São Paulo: Insper. https://www.insper.edu.br/wp-content/uploads/2018/09/Efeitos-Educacao-Profissional-Pronatec-Salarios.pdf

Araújo, A. J. N., Chein, F., & Pinto, C. (2014, 19–12 de dezembro). Ensino profissionalizante, desempenho escolar e inserção produtiva: Uma análise com dados do ENEM. *In 42º Encontro Nacional de Economia da ANPEC*, Natal, RN. Acessado em 08 agosto 2016: http://www.anpec.org.br/encontro/2014/submissao/files_I/i12-ac3a8f487db438fc6278e117ab468a01.pdf

Assunção, J., & Gonzaga, G. (2010). *Educação profissional no Brasil: Inserção e retorno* (Série Cenários Nº 3). Brasília: SENAI.DN.

Attanasio, O., Kugler, A., & Meghir, C. (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *3*(3), 188–220. http://dx.doi.org/10.1257/app.3.3.188

Bassi, M., Busso, M., Urzúa, S., & Vargas, J. (2012). *Desconectados: Habilidades, educación y empleo en América Latina*. Washington, DC. https://publications.iadb.org/handle/11319/427?scope=123456789

Chae, C., & Chung, J. (2009). *Pre-employment vocational education and training in Korea* (SP Discussion Paper Nº 0921). Washington, DC: The World Bank – Social Protection & Labor. http://documents.worldbank.org/curated/en/644501468045004776/Pre-employment-vocational-education-and-training-in-Korea

Chen, D. (2009). *Vocational schooling, labor market outcomes, and college entry* (Policy Research Working Paper Nº 4814). Washington: World Bank. http://documents.worldbank.org/curated/en/771881468049456978/Vocational-schooling-labor-market-outcomes-and-college-entry

de Hoyos, R., Rogers, H., & Székely, M. (2016). *Out of school and out of work: Risk and opportunities for Latin America's* Ninis. Washington, DC: World Bank. http://hdl.handle.net/10986/22349

El-Hamidi, F. (2006). General or vocational schooling? Evidence on school choice, returns, and 'sheepskin' effects from Egypt 1998. *The Journal of Policy Reform*, *9*(2), 157–176. http://dx.doi.org/10.1080/13841280600772861

Goldthorpe, J. H. (1992). *Social mobility and class structure in modern Britain*. New York: Oxford University Press.

Guimarães, A. Q., & Almeida, M. E. (2013). Os jovens e o mercado de trabalho: Evolução e desafios da política de emprego no Brasil. *Temas de Administração Pública*, *8*(2). http://seer.fclar.unesp.br/temasadm/article/view/6845

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political Analysis*, *20*(1), 25–46. http://dx.doi.org/10.1093/pan/mpr025

Hainmueller, J., & Xu, Y. (2013). `ebalance`: A Stata package for entropy balancing. *Journal of Statistical Software*, *54*(7). http://dx.doi.org/10.18637/jss.v054.i07

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1), 153–161. http://dx.doi.org/10.2307/1912352

Hirano, K., Imbens, G., & G., R. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, *71*(4), 1161–1189. http://dx.doi.org/10.1111/1468-0262.00442

Ho, D. E., Imai, K., King, G., & A., S. E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis*, *15*(3), 199–236. http://dx.doi.org/10.1093/pan/mpl013

Horowitz, A. W., & Schenzler, C. (1999). Returns to general, technical and vocational education in developing countries: Recent evidence from Suriname. *Education Economics*, *7*(1), 5–19. http://dx.doi.org/10.1080/09645299900000001

Imbens, G. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *The Review of Economics and Statistics*, *86*(1), 4–29. http://dx.doi.org/10.1162/003465304323023651

Jannuzzi, P. M. (2001). *Status* socioeconômico das ocupações brasileiras: Medidas aproximativas para 1980, 1991 e anos de 1990. *Revista Brasileira de Estatística*, *61*(2), 47–74. http://biblioteca.ibge.gov.br/visualizacao/periodicos/111/rbe_2000_v61_n216.pdf

Jannuzzi, P. M. (2003). *Indicadores sociais no Brasil* (2ª ed.). Campinas: Alínea.

Kahyarara, G., & Teal, F. (2008). The returns to vocational training and academic education: Evidence from Tanzania. *World Development*, *36*(11), 2223–2242. http://dx.doi.org/10.1016/j.worlddev.2007.11.011

Lechner, M. (2000). An evaluation of public-sector-sponsored continous vocational training program in East Germany. *Journal of Human Resources*, *35*(2), 347–375. http://dx.doi.org/10.2307/146329

Lin, D., Lutter, R., & Ruhm, C. J. (2016). *Cognitive performance and labor market outcomes* (Discussion Paper Nº 10075). IZA – Forschungsinstitut zur Zukunft der Arbeit. https://www.iza.org/de/publications/dp/10075/cognitive-performance-and-labor-market-outcomes

Mahirda, K., & Wahyuni. (2016). Returning to general and vocational high-schools in Indonesia. *Review of Economic and Business Studies*, *9*(2), 9–28. http://dx.doi.org/10.1515/rebs-2016-0031

Malamud, O., & Pop-Eleches, C. (2010). General education versus vocational training: Evidence from an economy in transition. *The Review of Economics and Statistics*, *92*(1), 43–60. http://dx.doi.org/10.1162/rest.2009.11339

Neuman, S., & Ziderman, A. (1991). Vocational schooling, occupational matching, and labour market earnings in Israel. *The Journal of Human Resources*, *26*(2), 256–281. http://dx.doi.org/10.2307/145923

Newhouse, D., & Suryadarma, D. (2011). The value of vocational education: High School type and labor market outcomes in Indonesia. *The World Bank Economic Review*, *25*(2), 296–322. http://dx.doi.org/10.1093/wber/lhr010

Rosenbaum, P. R., & Rubin, D. B. (1983). Assessing sensitivity to an unobserved binary covariate in an observational study with binary outcome. *Journal of the Royal Statistical Society. Series B (Methodological)*, *45*(2), 212–218. https://www.jstor.org/stable/2345524

Rubin, D. B. (2006). *Matched sampling for causal effects*. Cambridge University Press.

Sekhon, J. S. (2009). Opiates for the matches: Matching methods for causal inference. *Annual Review of Political Science*, *12*, 487–508. http://dx.doi.org/10.1146/annurev.polisci.11.060606.135444

Severnini, E. R., & Orellano, V. I. F. (2010). O efeito do ensino profissionalizante sobre a probabilidade de inserção no mercado de trabalho e sobre a renda no período pré-PLANFOR. *Revista EconomiA*, *11*, 155–174. http://www.anpec.org.br/revista/vol11/vol11n1p155_174.pdf

Silva, M. R., Pelissari, L. B., & Steimbach, A. A. (2013). Juventude, escola e trabalho: Permanência e abandono na educação profissional técnica de nível médio. *Educação e Pesquisa*, *39*(2), 403–417. http://dx.doi.org/10.1590/S1517-97022012005000022

Simões, C. A. (2010). Políticas públicas do ensino médio: Realidade e desafios. *In* C. A. Ferreira, S. O. Peres, C. N. Braga, & M. L. de Macedo Cardoso (Orgs.), *Juventude e iniciação científica: Políticas públicas para o ensino médio* (pp. 135–142). Rio de Janeiro: EPSJV, UFRJ. http://www.epsjv.fiocruz.br/publicacao/livro/juventude-e-iniciacao-cientifica-politicas-publicas-para-o-ensino-medio

Valle Silva, N. (1992). *Uma proposta de classificação das ocupações brasileiras* [Mimeografado]. LNCC.

Watson, S. K., & Elliot, M. (2016). Entropy balancing: A maximum-entropy reweighting scheme to adjust for coverage error. *Quality & Quantity*, *50*(4), 1781–1797. http://dx.doi.org/10.1007/s11135-015-0235-8

# Apêndice A. Procedimentos Metodológicos

## Seletividade Amostral

O primeiro estágio consiste em estimar a seguinte equação de participação:

$$L_i = \beta_i X_i + \mu_i, \tag{A-1}$$

em que $L_i$ é a variável que designa a participação no mercado de trabalho e $X_i$ representa um conjunto de variáveis que explicam a participação no mercado de trabalho observadas durante o ensino médio. Como $L_i$ não é observado, utiliza-se uma variável *dummy* definida como segue:

$$L_i = \begin{cases} 1, & \text{se } L_i^* > 0, \\ 0, & \text{se } L_i^* \leq 0. \end{cases} \tag{A-2}$$

Neste caso, $L_i$ assume o valor 1 para representar o indivíduo que está trabalhando e 0 caso contrário.

Após a estimação do modelo binário com a hipótese *probit*, o segundo procedimento para aplicar o método de Heckman (1979) consiste em estimar a equação de rendimentos, levando em consideração o viés de seleção amostral:

$$W_i = \alpha X_i + \gamma_i' Z_i + \varepsilon_i, \tag{A-3}$$

em que $W$ é a renda do indivíduo; $X$ representa uma variável binária, identificando se o aluno concluiu o ensino médio nas EEEP; $Z$ é o vetor de características pessoais; e $\varepsilon$ é o vetor de erros. Assim, o viés de seleção amostral pode ser observado da seguinte forma:

$$\begin{aligned} \mathbb{E}\left(W_i \middle| L_i^* > 0\right) &= \mathbb{E}\left(W_i \middle| \mu_i > -\beta_i X_i\right) \\ &= \gamma_i' Z_i + \mathbb{E}\left(\varepsilon_i \middle| \mu_i > -\beta_i X_i\right) \\ &= \gamma_i' Z_i + \frac{\mathrm{Cov}\left(\mu_i, \varepsilon_i\right)}{\sigma_\mu} \frac{\varnothing\left(\beta_i X_i\right)}{\Phi\left(\beta_i X_i\right)}, \end{aligned} \tag{A-4}$$

em que $\varnothing$ é a função de densidade de probabilidade e $\Phi$ é a função de distribuição normal padronizada acumulada. A função $\varphi_i = \varnothing\left(\beta_i X_i\right)/\Phi\left(\beta_i X_i\right)$ é denominada razão inversa de Mills e representa a variável que segundo Heckman (1979) é utilizada como explicativa na regressão principal para eliminar o viés de seletividade amostral, em que a significância desta função indicará a presença de viés de seletividade amostral. Logo, a equação[12] será representada da seguinte forma:

$$Y_i = \alpha X_i + \gamma_i' Z_i + \gamma_\varphi \varphi_i + \varepsilon_i, \tag{A-5}$$

$$Y_i = \gamma_i' K_i + u_i. \tag{A-6}$$

[12] Em vista da especificação da equação (A-5), cabe registrar que a estimação da equação (A-3) pelo método de mínimos quadrados ordinários forneceria estimativas viesadas e inconsistentes, devido à omissão da variável $\varphi_i$, resultando no erro de especificação descrito por Heckman (1979).

## Modelo Multinomial

Especificamente, a probabilidade é determinada por $p[(Y = j)|x]$, com $j = 1, 2, 3, 4, 5$:

$$p_{ij} = p\left[\left(Y_i - j\right)\middle| x_i\right] = \frac{e^{x_i'\beta_j}}{\sum_{j=1}^{5} e^{x_i'\beta_j}}, \tag{A-7}$$

em que $Y_i$ é a variável aleatória que indica a escolha da ocupação; $p_{ij}$ é a probabilidade de um indivíduo $i$ optar pela escolha $j$; $x_i$ é o vetor de características; $\beta$ é o vetor de parâmetros a serem estimados. A equação estimada fornece um conjunto de probabilidades para $J + 1$ escolhas, porém, aplicando uma normalização em (A-7), e fazendo $\beta_0 = 0$, $J$ parâmetros serão necessários para determinar as $J + 1$ probabilidades. Assim, a equação (A-7) pode ser substituída por

$$p_{ij} = p\left[\left(Y_i - j\right)\middle| x_i\right] = \frac{e^{x_i'\beta_j}}{1 + \sum_{j=1}^{5} e^{x_i'\beta_j}}, \qquad \beta_0 = 0. \tag{A-8}$$

As razões de chances podem ser calculadas da seguinte forma:

$$\ln \frac{p_{ij}}{p_{ik}} = x_i'\left(\beta_j - \beta_k\right) = x_i'\beta_j, \qquad k = 0. \tag{A-9}$$

Uma vez que os coeficientes estimados não representam diretamente as respostas marginais das variáveis explicativas, faz-se necessário calcular os efeitos marginais para analisar corretamente os resultados, os quais são obtidos por:

$$\delta_{ij} = \frac{\partial p_{ij}}{\partial x_i} = p_{ij}\left[\beta_j - \sum_{j=1}^{5} p_{ij}\beta_k\right] = p_{ij}\left[\beta_j - \bar{\beta}_i\right], \tag{A-10}$$

em que $\bar{\beta}_i = \sum_{j=1}^{5} p_{ij}\beta_k$ é a probabilidade média de $\beta_j$.

## Balanceamento por Entropia

Considere amostras aleatórias extraídas de unidades de tratados e de controle, nas quais inserem-se as seguintes definições:

a) $D_i$ é uma variável binária que assume o valor 1 se a unidade pertence ao tratamento, e 0 se pertencer ao controle;
b) $X$ é uma matriz composta pelos elementos $x_{ij}$, referentes aos valores da variável exógena pré-determinada $j$ na unidade $i$;
c) A densidade das covariadas nas amostras de tratamento e controle são dadas por $f_{X|D_i=1}(x)$ e $f_{X|D_i=0}(x)$, respectivamente;
d) A variável resultado observado é expressa por $Y_i = Y_i(1)D_i + \left(1 - D_i\right)Y_i(0)$.

O Efeito Médio Tratamento sobre os Tratados (EMTT) é dado por $\tau = \mathbb{E}\left[Y_i(1)\middle| D_i = 1\right] - \mathbb{E}\left[Y_i(0)\middle| D_i = 1\right]$, onde a primeira média pode ser diretamente identificada do grupo de tratados, mas a segunda corresponde ao contrafactual, o qual não é observado. Rosenbaum e

Rubin (1983) mostra que, assumindo seleção nos observáveis, $Y(0) \perp D|X$, e sobreposição, $\Pr\left(D_i = 1 \middle| X_i = x\right) < 1$ para todo $x$ no suporte de $f_{X|D_i=1}$, o EMTT é identificado como

$$\tau = \mathbb{E}\left(Y_i \middle| D_i = 1\right) - \int \mathbb{E}\left(Y_i \middle| X_i = x, D_i = 0\right) f_{X|D_i=1}(x)\,\mathrm{d}x. \tag{A-11}$$

Para estimar o contrafactual, a distribuição da covariável no grupo de controle necessita ser ajustada para torná-la semelhante à distribuição no grupo de tratamento, tal que o indicador de tratamento D se torne mais perto de ser ortogonal em relação às covariáveis. Uma variedade de métodos de pré-processamento de dados, tais como pareamento pelo vizinho mais próximo, cem, escore de propensão têm sido propostos para reduzir o desequilíbrio na distribuição de variáveis de controle. Uma vez ajustadas as distribuições de variáveis independentes, métodos de análise padrão, tais como a regressão, podem ser posteriormente utilizados para estimar o tratamento com menor erro e modelo de dependência (Ho, Imai, King, & A., 2007; Imbens, 2004; Rubin, 2006; Sekhon, 2009).

Considere o caso mais simples em que o efeito tratamento nos dados pré-processados é estimado usando a diferença nos resultados médios entre os grupos de tratados e controle ajustado, cujo método muito utilizado na literatura é o escore de propensão ponderado (Hirano, Imbens, & G., 2003), cuja a média contrafactual é estimada como segue:

$$\mathbb{E}\left[Y_i(0) \middle| \widehat{D}_i = 1\right] = \frac{\sum_{\{i|D=0\}} Y_i d_i}{\sum_{\{i|D=0\}} d_i}. \tag{A-12}$$

As unidades de controle recebem um peso dado por $d_i = \frac{\hat{p}(x_i)}{1-\hat{p}(x_i)}$; $\hat{p}(x_i)$ na equação (A-12) é o escore de propensão, o qual é comumente estimado através de uma regressão probit ou logit. Se este modelo estiver corretamente especificado, então o peso estimado $d_i$ assegura que a distribuição da covariável das unidades de controle reponderadas corresponda à distribuição no grupo de tratamento. No entanto, na prática, essa abordagem muitas vezes não consegue equilibrar conjuntamente todas as covariáveis.

O balanceamento por entropia generaliza a abordagem de ponderação do escore de propensão ao estimar os pesos diretamente de um conjunto de restrições de equilíbrio que exploram o conhecimento do pesquisador sobre os momentos amostrais. Considere $w_i$ o peso do balanceamento por entropia escolhido para cada unidade de controle, os quais foram encontrados pelo seguinte esquema de reponderação que minimiza a distancia métrica de entropia:

$$\min_{w_i} H(w) = \sum_{\{i|D=0\}} w_i \log(w_i/q_i), \tag{A-13}$$

sujeito as restrições de equilíbrio e normalização

$$\sum_{\{i|D=0\}} w_i c_{ri}\left(X_i\right) = m_r, \qquad \text{com } r \in 1, \ldots, R; \tag{A-14}$$

$$\sum_{\{i|D=0\}} w_i = 1; \tag{A-15}$$

$$w_i \geq 0, \qquad \text{para todo } i, \text{ tal que } D = 0, \tag{A-16}$$

em que $q_i = 1/n$ é um peso base, sendo n o tamanho da amostra das unidades de controle; $c_{ri}\left(X_i\right) = m_r$ descreve um conjunto de $R$ restrições referentes aos momentos das covariadas

no grupo de controle reponderados. Inicialmente, escolhe-se a covariada que será incluída na reponderação. Para cada covariada, especifica-se um conjunto de restrições de balanceamento (A-14) para equiparar os momentos das distribuições das covariadas entre os grupos de tratamento e controles reponderados. As restrições de momentos podem ser a média (primeiro momento), a variância (segundo momento), e a assimetria (terceiro momento).

Dessa forma, o balanceamento por entropia procura, para um conjunto de unidades, pesos $W = [w_i, \dots, w_{n_0}]'$ no qual minimiza a equação (A-13), distancia de entropia entre $W$ e o vetor base de pesos $Q = [q_i, \dots, q_{n_0}]'$, sujeita as restrições de balanceamento na equação (A-14), restrição de normalização (A-15), e restrição de não-negatividade (A-16). Devido à sua propriedade de minimizar os desequilíbrios entre os grupos de tratados e controle, este método será aqui aplicado.

## Apêndice B. Descrição das Variáveis

**Tabela B-1.** Agrupamentos de Ocupações.

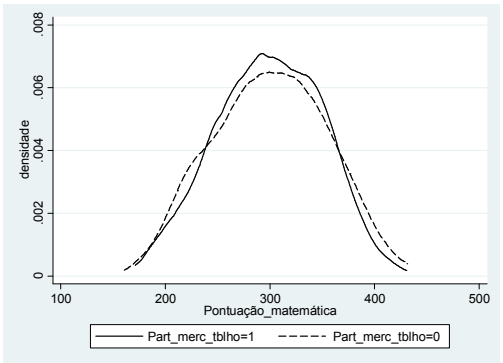
| Ocupações | Categorias Profissionais |
|---|---|
| 1 | Médicos, engenheiros, professores universitários, empresários, gerentes e postos superiores na administração pública (juízes, promotores, delegados, oficiais das forças armadas, etc.). |
| 2 | Técnicos de contabilidade e administração, Mestre e Contramestres na indústria, Professores de ensino fundamental e médio, Corretores de Imóveis, Inspetores de Polícia, Carteiros, Comerciantes (proprietários) e Agricultores. |
| 3 | Torneiro Mecânico, Montadores de Equipamentos Elétricos, Vendedores, Operadores de caixa, Comerciantes conta-própria, Professores de ensino pré-escolar, Motoristas, Inspetores de alunos, Auxiliares de enfermaria, Auxiliares administrativos e de escritório, Policiais e Praças das Forças Armadas. |
| 4 | Ocupações da Indústria de Alimentos, da Indústria Têxtil, Pedreiros, Pintores, Garçons, Vigias, Porteiros, Estivadores. |
| 5 | Trabalhadores rurais, além das ocupações urbanas de baixo status, como a de serventes de pedreiro, lavadeiras, empregadas domesticas e lixeiros. |

**Tabela B-2.** Variáveis na amostra e suas fontes.

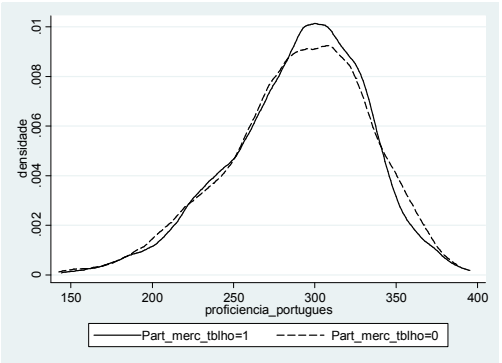
| Variável | Descrição | Fonte |
|---|---|---|
| **Estimação I** | | |
| *Dependente* | | |
| Participação no Mercado de Trabalho | 1 se o concludente em 2011 estava auferindo renda positiva no mercado de trabalho formal em 2014; 0 caso contrário | RAIS |
| *Explicativas* | | |
| EEEP[a] | 1 se o aluno concluiu ensino médio na escola profissionalizante; 0 se o aluno concluiu na escola regular | SEDUC/Censo escolar |
| Idade_2011 | Idade do concludente em 2011 | SPAECE |
| Sexo | 1 se masculino; 0 se feminino | SPAECE |
| Zona Residencial | 1 se urbana; 0 caso contrário | SPAECE |
| Bolsa família | 1 se alguém da família recebe o benefício do bolsa família; 0 caso contrario | SPAECE |
| Escolaridade Pai | 1 se o pai não estudou ou se possui até a 4° série do ensino fundamental; 0 caso contrário | SPAECE |
| Escolaridade Mãe | 1 se o pai não estudou ou se possui até a 4° série do ensino fundamental; 0 caso contrário | SPAECE |
| *Entropia* | | |
| Proficiência MT | Proficiência do aluno na prova de matemática em 2011 | SPAECE |
| Proficiência LP | Proficiência do aluno na prova de português em 2011 | SPAECE |
| Faixa_idade1 | 1 se o índividuo possui idade entre 17 e 25 anos | Censo escolar |
| Faixa_idade2 | 1 se o índividuo possui idade entre 26 e 40 anos | Censo escolar |
| **Estimação II** | | |
| *Dependente* | | |
| Renda média | Remuneração anual média | RAIS |
| Ocupações | Indicador categórico que reflete nível de status socioeconômico e grau de precarização das ocupações, variando de 1 (alto status) a 5 (baixo status) | RAIS |
| *Explicativas* | | |
| EEEP[a] | 1 se o aluno concluiu ensino médio na escola profissionalizante; 0 se o aluno concluiu na escola regular | SEDUC/Censo escolar |
| Saúde | 1 se participou do curso de profissionalizante na área da saúde; 0 Caso contrário | SEDUC/Censo escolar |
| Gestão | 1 se participou do curso de profissionalizante na área de gestão; 0 Caso contrário | SEDUC/Censo escolar |
| Informação | 1 se participou do curso de profissionalizante na área de informação; 0 Caso contrário | SEDUC/Censo escolar |
| Hospitalidade e Lazer | 1 se participou do curso de profissionalizante na área de Hospitalidade e Lazer; 0 Caso contrário | SEDUC/Censo escolar |
| Outros | 1 se participou do curso de profissionalizante em outras áreas; 0 Caso contrário | SEDUC/Censo escolar |
| Sexo | 1 se masculino; 0 caso contrário | RAIS |
| Raça | 1 se branca; 0 caso contrário | RAIS |
| Idade 2014 | Idade do concludente em 2014 | RAIS |
| Mills | Razão inversa de Mills | |

[a] Escolas Estaduais de Ensino profissionalizante.

# Apêndice C. Resultados Complementares

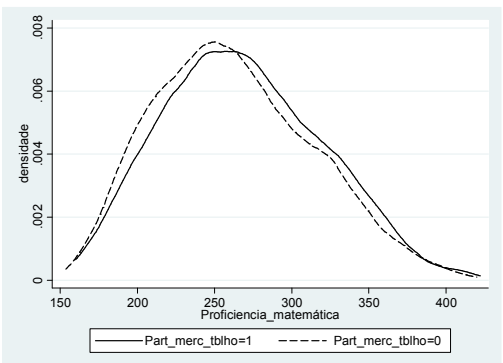



**(a)** Em Matemática

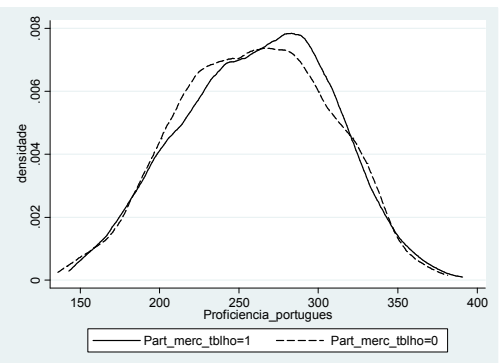



**(b)** Em Português

**Figura C-1.** Desempenho dos concludentes das EEEP

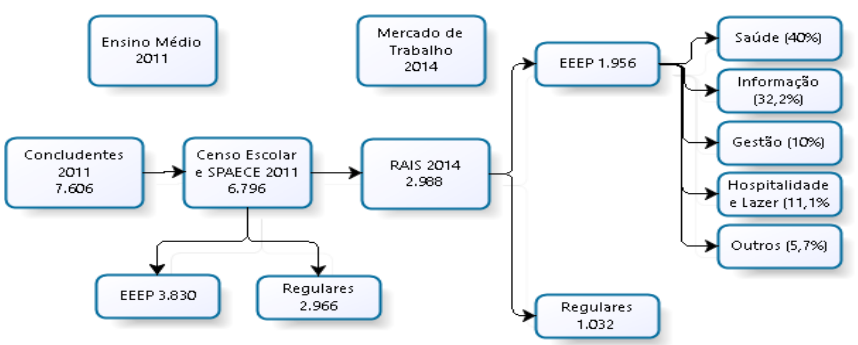



**(a)** Em Matemática




**(b)** Em Português

**Figura C-2.** Desempenho dos Concludentes das Escolas Regulares




*Fonte:* Elaboração dos Autores com base nos Microdados do SPAECE 2011, Censo Escolar 2011 e RAIS 2014.

**Figura C-3.** Construção da amostra.

**Tabela C-3.** Desempenho dos egressos de 2011 por Tipo de Escola e Participação no Mercado de Trabalho.

| | Participam no Mercado de Trabalho | | | | Não Participam do Mercado de Trabalho | | | |
|---|---|---|---|---|---|---|---|---|
| Proficiências | Média | Desvio-padrão | Valor Mín | Valor Máx | Média | Desvio-padrão | Valor Mín | Valor Máx |
| *Escolas Profissionalizantes* | | | | | | | | |
| Matemática | 299,20 | 50,68 | 171,3 | 430,7 | 299,94 | 54,87 | 160,7 | 431,7 |
| Português | 289,87 | 41,44 | 144,2 | 395,2 | 290,04 | 43,28 | 142,5 | 387,3 |
| *Escolas Regulares* | | | | | | | | |
| Matemática | 268,67 | 51,20 | 160,0 | 422,2 | 262,86 | 51,05 | 153,6 | 420,0 |
| Português | 261,77 | 46,46 | 142,8 | 390,8 | 259,29 | 46,97 | 136,0 | 381,3 |

**Tabela C-4.** Estimação do Probit – Participação no Mercado de Trabalho 2014.

| Variáveis | Coeficientes | Erro-padrão |
|---|---|---|
| EEEP | 0,3683 *** | 0,0355 |
| Idade | 0,0067 | 0,0055 |
| Gênero-Masculino | 0,3058 *** | 0,0313 |
| Residência – Urbano | 0,1456 *** | 0,0375 |
| Bolsa_familia | 0,0645 ** | 0,0324 |
| Escolaridade pai * | −0,0878 ** | 0,0348 |
| Escolaridade mãe * | −0,0507 | 0,0359 |
| Constante | −0,7063 *** | 0,1143 |
| N° de observações | 6.773 | |

*Notas:* Escolaridade pai e Escolaridade mãe: se o pai(mãe) não estudou ou se possui até a 4° série do ensino fundamental. Significância estatística dos coeficientes: *10%; **5%; ***1%.

**Tabela C-5.** Estimação do modelo multinomial das ocupações no mercado de trabalho – com correção da seletividade amostral e balanceamento por entropia.

| | Ocupação 1 | | Ocupação 2 | | Ocupação 3 | | Ocupação 4 | |
|---|---|---|---|---|---|---|---|---|
| | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão |
| EEEP | −2,1332 * | 1,1589 | −2,4679 *** | 0,9753 | −2,7061 *** | 0,9671 | −2,3298 ** | 0,9841 |
| Gênero–Masculino | −4,0100 *** | 1,1601 | −4,6980 *** | 1,0715 | −4,2341 *** | 1,0690 | −2,7722 *** | 1,0809 |
| Raça–Branca | 0,0349 | 0,5508 | −0,2121 | 0,5038 | −0,4720 | 0,5005 | −0,5815 | 0,5151 |
| Idade | −0,0428 | 0,0329 | −0,2112 *** | 0,0655 | −0,1063 *** | 0,0375 | −0,0424 | 0,0297 |
| Mills | −12,2297 *** | 3,9572 | −12,9144 *** | 3,5289 | −12,0601 *** | 3,5074 | −8,5183 ** | 3,5619 |
| Constante | 17,2877 *** | 5,1245 | 24,3611 *** | 4,8253 | 21,3264 *** | 4,6341 | 14,8591 *** | 4,6602 |
| N° de Observações | 2.974 | | | | | | | |

*Notas:* Ocupações: 1 – médicos, engenheiros, professores universitários, empresários, gerentes e postos superiores na administração pública (juízes, promotores, delegados, oficiais das forças armadas, etc.); 2 – técnicos de contabilidade e administração, mestre e contramestres na indústria, professores de ensino fundamental e médio, corretores de imóveis, inspetores de polícia, carteiros, comerciantes (proprietários) e agricultores; 3 – torneiro mecânico, montadores de equipamentos elétricos, vendedores, operadores de caixa, comerciantes conta-própria, professores de ensino pré-escolar, motoristas, inspetores de alunos, auxiliares de enfermaria, auxiliares administrativos e de escritório, policiais e praças das forças armadas; 4 – ocupações da indústria de alimentos, da indústria têxtil, pedreiros, pintores, garçons, vigias, porteiros, estivadores. Níveis de significância dos coeficientes: *10%; **5%; ***1%.

**Tabela C-6.** Estimação do modelo multinomial das ocupações no mercado de trabalho.

| | Ocupação 1 | | Ocupação 2 | | Ocupação 3 | | Ocupação 4 | |
|---|---|---|---|---|---|---|---|---|
| | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão | Coeficientes | Erro-padrão |
| ***Painel A:*** *Sem correção da seletividade amostral e sem balanceamento por entropia* | | | | | | | | |
| EEEP | 2,0141*** | 0,4675 | 1,9670*** | 0,4143 | 1,2859*** | 0,4128 | 0,4274 | 0,4177 |
| Gênero | −0,9651** | 0,4962 | −1,6426*** | 0,4626 | −1,3001*** | 0,4621 | −0,7130 | 0,4671 |
| Raça | −0,2508 | 0,4925 | −0,5568 | 0,4450 | −0,7655* | 0,4448 | −0,8063* | 0,4537 |
| Idade | −0,0255 | 0,0512 | −0,0433 | 0,0405 | −0,0450 | 0,0405 | −0,0208 | 0,0403 |
| Constante | 1,5415 | 1,2780 | 4,6983*** | 1,0516 | 5,0087*** | 1,0422 | 3,8026*** | 1,0426 |
| N° de Observações | | | | 2.982 | | | | |
| ***Painel B:*** *Cursos — com correção da seletividade amostral e balanceamento por entropia* | | | | | | | | |
| Saúde | −2,4910* | 1,3578 | −2,2279* | 1,2570 | −2,5729** | 1,2532 | −2,1620* | 1,2611 |
| Gestão | 12,2687*** | 1,3252 | 10,9355*** | 1,2216 | 11,1616*** | 1,2201 | 11,2674*** | 1,2385 |
| Informação | −0,6177 | 1,3190 | −1,5229 | 1,2439 | −1,6968 | 1,2397 | −1,4969 | 1,2471 |
| Hospitalidade e Lazer | 11,3455*** | 1,3423 | 10,7226*** | 1,2270 | 10,9702*** | 1,2220 | 11,2528*** | 1,2365 |
| Gênero | −3,6035*** | 1,3132 | −4,2515*** | 1,2641 | −3,7534*** | 1,2623 | −2,3388* | 1,2677 |
| Raça | 0,0614 | 0,5507 | −0,1862 | 0,5011 | −0,4378 | 0,4981 | −0,5519 | 0,5125 |
| Idade | −0,0254 | 0,0328 | −0,2017*** | 0,0667 | −0,0965*** | 0,0377 | −0,0342 | 0,0292 |
| Mills | −9,3644** | 4,3325 | −10,6385*** | 4,1488 | −9,5013** | 4,1372 | −6,2880 | 4,1587 |
| Constante | 13,5614** | 5,5978 | 21,4086*** | 5,5465 | 18,0445*** | 5,3951 | 11,9895** | 5,3864 |
| ***Painel C:*** *Cursos — sem correção da seletividade amostral e sem balanceamento por entropia* | | | | | | | | |
| Saúde | 0,7884 | 0,5946 | 1,4035*** | 0,5142 | 0,6108 | 0,5146 | −0,0886 | 0,5245 |
| Gestão | 16,8844 | 12,8702 | 15,9066 | 12,8702 | 15,6858 | 12,8702 | 14,6999 | 12,8702 |
| Informação | 2,5042*** | 0,6706 | 1,9911*** | 0,6284 | 1,3641** | 0,6281 | 0,5210 | 0,6364 |
| Hospitalidade e Lazer | 15,9715 | 11,8743 | 15,6838 | 11,8743 | 15,4905 | 11,8743 | 14,7166 | 11,8743 |
| Gênero | −1,1811** | 0,5045 | −1,6650*** | 0,4688 | −1,3578*** | 0,4685 | −0,7644 | 0,4736 |
| Raça | −0,2774 | 0,4952 | −0,5901 | 0,4458 | −0,7790* | 0,4459 | −0,8056* | 0,4549 |
| Idade | −0,0255 | 0,0507 | −0,0506 | 0,0402 | −0,0493 | 0,0398 | −0,0208 | 0,0396 |
| Constante | 1,7161 | 1,2635 | 5,0256*** | 1,0370 | 5,2065*** | 1,0287 | 3,8427*** | 1,0280 |

*Notas:* Ocupações: 1 – médicos, engenheiros, professores universitários, empresários, gerentes e postos superiores na administração pública (juízes, promotores, delegados, oficiais das forças armadas, etc.); 2 – técnicos de contabilidade e administração, mestre e contramestres na indústria, professores de ensino fundamental e médio, corretores de imóveis, inspetores de polícia, carteiros, comerciantes (proprietários) e agricultores; 3 – torneiro mecânico, montadores de equipamentos elétricos, vendedores, operadores de caixa, comerciantes conta-própria, professores de ensino pré-escolar, motoristas, inspetores de alunos, auxiliares de enfermaria, auxiliares administrativos e de escritório, policiais e praças das forças armadas; 4 – ocupações da indústria de alimentos, da indústria têxtil, pedreiros, pintores, garçons, vigias, porteiros, estivadores. Níveis de significância dos coeficientes: *10%; **5%; ***1%.

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180023
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Grants and Marginal Cost of Public Funding: Empirical Evidence for Local Governments in Brazil

Enlinson Mattos[*]
Rafael Cardim[†]
Ricardo Politi[‡]

**Contents**

1. Introduction .....479
2. Theoretical background ....481
3. Methodology .....484
4. Result .....488
5. Conclusion .....492
Appendix. .....495

**Keywords**
Price effect, income effect, transfers, taxes, fiscal federalism

**JEL Codes**
H2, H7, H25

**Abstract · Resumo**

Este documento documenta evidências empíricas sobre o efeito-preço causado por subsídios de montante fixo para governos locais no Brasil entre 2006 e 2010. Dahlby (2011) demonstra teoricamente que subsídios de quantia fixa podem reduzir o custo da provisão de bens públicos (efeito-preço), além do tradicional efeito renda. Nossas contribuições são triplas. Primeiro, estimamos a semi-elasticidade dos efeitos das alterações nas alíquotas de impostos na base de cálculo (0,016). Segundo, calculamos o MCF do imposto local imposto sobre a prestação de serviços (ISS) para os municípios brasileiros (média de 0,04). Finalmente, estimamos a estimativa de preço-efeito para o imposto ISS. Nossos resultados sugerem que, para toda a amostra, um aumento de R$ 1,00 nas transferências incondicionais per capita reduz o efeito de preço local (MCF) em torno de 0,07%, mas esse resultado não é consistentemente estimado em todas as subamostras.

## 1. Introduction

An important topic in the Public Finance literature is the discussion about the cost of raising additional resources from private sector to fund public provision. As discussed in the tax burden literature, in a true non-distortionary tax system, tax revenue is distributed directly from the private to the public sector, in such a manner that the income loss for the private sector is equivalent to the income earning for the public sector (Hakonsen, 1998). In this scenario, the marginal dead weight loss from taxation is null.

In a different formulation, Dahlby (2011), following Hamilton (1986), approaches tax burden in a model in which local governments are financed by distortive taxes. Under this scenario, the marginal loss caused to society by a tax rate increase is captured by the idea of marginal cost of public fund (MCF). The MCF can be stated as the ratio between the social marginal value of a monetary unit (i.e. $1 Real) raised by the government and its counterpart

[*]Fundação Getúlio Vargas, Escola de Economia de São Paulo (FGV/EESP). Rua Itapeva, 474, Bela Vista, São Paulo, SP, CEP 01332-000, Brasil.

[†]Mestre em economia. Fundação Getúlio Vargas, Escola de Economia de São Paulo (FGV/EESP).

[‡]Universidade Federal do ABC, Centro de Engenharia, Modelagem e Ciencias Sociais Aplicadas (UFABC/SECS). Avenida dos Estados, 5001, Santa Terezinha, Santo André, SP, CEP 09210-580, Brasil.

enlinson.mattos@fgv.br rafaelcardim1@hotmail.com ricardo.politi@ufabc.edu.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Received on December 17th, 2015
Final Acceptance on July 5th, 2018

in the private sector (the social marginal value of a monetary unit in the private sector). According to Jacobs (2018), the MCF is a measure of the scarcity of the public resouces in society.

In fact, the MCF aims to measure the distortion in the allocation of resources in an economy resulting from a response to a non lump-sum tax, such as the labor supply, consumption, and investment decisions. Beyond tax distortion, additional resources from intergovernamental grants can also affect the cost of public funds. In the Dahlby's (2011) model, lump-sum intergovernmental grants have both an income effect, which is widely reported in the "flypaper" literature (Gramlich, 1977; Fisher, 1982; Wyckoff, 1991; Inman, 2008), and a price effect. The mechanism that such price-effect works is via a decrease in the marginal cost of public fund.

This price effect refers to the situation in which additional resources from intergovernamental grants can impact negatively the local cost of public provision, as it allows the local government to decrease the tax rate and still to maintain the same level of public goods. Following this debate, the main objective of this paper is to estimate the price-effect in lump-sum grants using data from 5,173 Brazilian municipalities between 2006 and 2012 (seven years sample). In this article we (i) estimate the effects of tax rate changes on tax base (number of fims); (ii) calculate the MCF of the local business tax imposed on the supply of services for Brazilian municipalities; and (iii) estimate the effect of grants on the estimated MCF.

The fiscal federalism literature has provided considerable attention to the intergovernmental grants schemes, which are important tools for correcting fiscal imbalances caused by disequilibrium of taxation sources between government levels and by differences in local fiscal need.

Lump-sum grants seem to stimulate more the local public spending than increases in local private income, reinforcing the expression "*money sticks where it hits*". The idea is that the revenues raised by the public sector remains in that sector. Hines and Thaler (1995) charactherize as an anomaly the larger propensity of grants to increase public spending compared to private income resources and there seems to have many theoretical explanations for that.[1] More related to our work, Hamilton (1986) claims that the flypaper effect might be due to tax distortion, which is captured by the deadweight loss of taxation (Hakonsen, 1998), as also by the difference between income and grants effects.

This issue is further explored in Dahlby (2011) that argues that grants can reduce the cost of taxation for local governments through a change in the MCF. In particular, the author shows that lump-sum grants allow the recipient government to reduce its tax rate, which in turn, decreases the MCF in order to keep the same level of public service. Whether local MCFs are larger than national average counterpart, additional resources from grants could increase consumer welfare. According to Browning (1987) tax-response can produce very disperse estimates (1% increase of a particular tax can trigger a marginal loss in well-being between 9.9 and 300%).[2]

[1] For a detailed survey, see Inman (2008) and Gamkhar and Shah (2007). For Brazil, Siqueira, Nogueira, Souza, and Carvalho (2010) found 0.002 to 2.2 using 27 groups of goods. Lanzer and Porto (2011) estimated the MCF for the whole Brazilian tax system in between 1.167 and 1.173. See also Cossio (2002) and Cossio and Carvalho (2001).

[2] Recently Auriol and Warlters (2012) estimated a mean MCF of 1.21 for 38 African countries using a general equilibrium model, finding that taxes on production factors have higher MCFs than do taxes on imports and domestic goods.

As discussed in Dahlby (2011, p.309), Brazil offers a interesting case to discuss the flypaper effect as also the cost of public fund as the Brazilian system of intergovernmental grants is broad; it accounts for 12% of gross domestic product (GDP) and it corresponds about 70% of local revenue streams in 2008. Thus, it is relevant to use Brazilian data in order to provide empirical evidence of the price-effect of transfers.

Overall, our main results indicate that the tax base (number of firms) reacts negatively to a local tax rate increase. Moreover, we find limited evidence that grants pushes down the marginal cost of public fund. Interestingly, our results suggest that this price effect seems greater for larger cities. We argue that this finding is associated to greater distortionary taxation found in larger municipalities. Considering the entire sample, our estimates suggests that an increase in R$ 1.00 in per capita unconditional transfers reduces the local price effect (MCF) around 0.00003 (0.07% of the average MCF).

In addition to this introduction, this paper is organized in four sections. Section 2 presents a modified version of Dahlby's (2011) model that allows to compare the price effect of grants to the price effect of income. Additionally, section 3 describes the empirical approach and the data used. Section 4 discusses our main findings. Finally, section 5 concludes the paper.

## 2. Theoretical background

### 2.1 Marginal Cost od Public Finance (MCF) Model

In this section, in order to develop an empirical approach to obtain the price effect of unconditional grants, we explore Dahlby's (2011) model of Marginal Cost of Public Funds (MCF). First, we assume that all local governments have a homogeneous, fixed population that may be represented by a single agent. The local government is in charge of choosing a tax rate $\tau$ that affects the tax base, $B$, and the amount of public services provided, $g$, at a constant production cost per capita $c$. The local government receives a lump-sum transfer $T$. Therefore, the government's budget constraint follows $\tau B + T = cg$.

The utility of the representative resident is $U = u(x, B) + w(g)$, where $x$ is the private consumption of goods (price equal to 1), $u(\cdot)$ is a quasi-concave function, $w' > 0$, and $w'' < 0$. In our case, the tax base ($B$) depends only on the local tax rate, $\tau$. Thus, it follows $B = B(\tau)$. Dahlby (2011, equation (1), p.307) derives MCF as

$$MCF_{\tau i} = \frac{B_i}{\frac{\mathrm{d}R}{\mathrm{d}\tau_i}} = \frac{B_i}{B_i + \tau_i \frac{\mathrm{d}B_i}{\mathrm{d}\tau_i}} = \frac{1}{1 + \tau_i \eta}, \tag{1}$$

where $\eta = \frac{\partial \ln B}{\partial \tau} < 0$ is the elasticity of the tax base with respect to the tax rate and $R$ stands for local tax revenue. Because of tax avoidance by firms, it is expected a negative association (elasticity) between the tax base and the tax rate, captured by the term $\eta$. Equation (1) indicates that the MCF is a measurement that reflects the cost of a tax rate increase in terms of tax revenue, or equivalently, how much it costs to obtain an additional R$ 1 in terms of tax revenue. For example, if the tax base is not affected by the increase in tax rate (i.e., $\eta = 0$), then an increase in the tax rate reflects the same measurement in tax revenue. However, it is expected that the tax base shrinks with a tax rate increase (as we expect that $\eta < 0$), which leads to MCF greater than one ($MCF > 1$).[3]

[3]Furthermore, the municipalities are assumed to be in the increasing segment of the Laffer Curve, i.e., $1 + \tau_i \eta > 0$. It can be shown that MCF is increasing with respect to tax rate under some conditions.

Assuming that local government intends to maximize the well-being of the residents, the provision of public goods goes up to the point at which the marginal benefit of public goods equals the effective price, $P$; e.g., $P \equiv MCF \cdot c$, where $c$ is the constant cost of provision of the public good per capita. If this government receives a lump-sum grants, $T$ from the central government, the provision of public goods could increase and the tax rate might decrease. Then, the MCF decreases because tax rate has decreased, which in turn, allows for a reduction in the effective price of providing the public good. Deriving the effective price in relation to $T$ and keeping $g$ constant ($\frac{\partial \tau_i}{\partial T} = -B^{-1}MCF$), the effect of an increase in grants to the relative price effect to may be expressed as

$$\frac{\partial P}{\partial T} = \frac{\partial P}{\partial \tau}\frac{\partial \tau}{\partial T} = \frac{c\eta}{(1+\tau\eta)^3}\frac{1+E}{B}, \tag{2}$$

where $E = \left(\frac{\mathrm{d}\eta_i}{\mathrm{d}\tau_i}\right)\left(\frac{\tau_i}{\eta_i}\right)$, which corresponds to the change of the elasticity of the tax base in relation to the tax rate. If the local government is financed with lump-sum taxes we have $MCF = 1$. If $E = -1$, then there would be no price effect for lump-sum grants. Supposing that $MCF > 1$ and $E > -1$, a lump-sum grant presents a greater price effect compared to the ratio between the lump-sum grant over the local tax revenue, as well as that of the local MCF.

To find the expression that reflects the relationship between those two forms of variation in the price-effect (grants and income), we first have to completely differentiate the budget constraint, while keeping $g$ constant, then one has to derive $P$ with respect to $Y$ and to $\tau$. After some algebra, one can find the price effect of income as indicated in function

$$\frac{\partial P}{\partial Y} = \frac{c\eta}{(1+\tau\eta)^3}\frac{1+E}{B}B_y\tau. \tag{3}$$

Function (3) represents the change in the price of the public good resulting from the variation in income.[4] Comparing the price-effects of income (3) to its counterpart of the grants' function (2), we have:

$$\frac{\partial P}{\partial T} = \frac{c\eta}{(1+\tau\eta)^3}\frac{1+E}{B} = \frac{\partial P}{\partial Y}\frac{1}{\tau B_y}. \tag{4}$$

Equation (4) shows that if $1 > \tau B_y > 0$, one has $\frac{\partial P}{\partial \tau} > \frac{\partial P}{\partial Y}$. This means that intergovernmental grants have a greater ability compared income to decrease the effective price of the public good. Thus, we should expect a negative relantionship between grants and MCF. To summarize, in our empirical approach, we investigate two main hyphothesis derived from our analytical framework: i) if the the association between tax rate and tax base is negative or not; ii) if the effect of grants in MCF is also negative and its magnitude. Next, we related the MCF model with Brazilian institutional framework.

[4]For example, if we consider $c = 1$, $E = 0$ (i.e., there is no variation in the elasticity of the tax base), $B = 100$, and $t = 1.01$ (tax rate of 1%), one has a price effect resulting from the grants that is equal to −0.00017 and a price effect of income equal to −0.0000002, less (in magnitude) because the price effect of income is multiplied by the variation in the tax base in relation to income (0.001).

## 2.2 Local tax revenues and unconditional transfers in Brazil

Regarding local tax revenues, the tax on business services (ISS) is the main source of public funding. This tax is collect over services acivities (excluding any type of manufactured activity) as small retail stores, hairdressers, repair services, restaurants, financial and educational services. The statutory business tax rate ranges from 2% to 5%.[5] Some sectors are tax exempt. Additionally, municipalities also obtain resources from local property tax (IPTU, in Portuguese). However, statutory tax rate in property tax in Brazil are historically low. In fact, average tax rate is less than 2% and most households live in areas or houses that are tax free.[6] According to Carvalho (2006), only 30% of the households collect property tax and property tax revenues represents only approximatelly 0.5% from Gross Domestic Product (GDP). In our sample, the local business tax revenues represent on average twice the revenue collected in local property tax. Thus, it is appropriate to consider the business tax as the main source of local tax revenue and source of potential distortionary taxation.

Regarding intergovernamental grants, municipalities obtain resources from States and from the Federation. Each State return approximatelly from 20% to 25% of the Valeu Added Tax (VAT) and of the Vehicle Tax revenue to the locality where the tax was collected.[7] However, it is noteworthy to clarify that the main source of revenues for local governments in Brazil are unconditional grants. Table 1 shows that meanwhile intergovernmental grants from the Municipalities Participation Fund (FPM, in Portuguese) accounts for on average 60% of total revenue resources, local tax revenue corresponds on average to only 20% of fiscal resources for municipal jurisdictions. Because of this large share in municipalities' budgets, it is important to comprehend the effect of unconditional grants on the MCF.

Finally, it is important to remind that very small firms in Brazil collect tax in a differentiated regime. In the "Simples Nacional" tax regime, firms collect local, State and National tax

**Table 1.** Fiscal variables by populational ranges – per capita in 2010 R$.

| | | Tax Revenue | | | | |
|---|---|---|---|---|---|---|
| Population Range | Unconditional Grant (FPM) | ISS | IPTU | State Grant | GDP | Nº Munic. |
| up to 5 k (thous.) | 1,367.5 | 44.1 | 10.4 | 567.4 | 11,918 | 1,178 |
| > 5 to 10 k | 630.8 | 42.8 | 14.1 | 404.6 | 10,873 | 1,172 |
| > 10 to 25 k | 473.0 | 41.0 | 16.9 | 319.4 | 10,299 | 1,612 |
| > 25 to 50 k | 343.8 | 59.0 | 30.1 | 322.1 | 12,305 | 650 |
| > 50 to 250 k | 245.2 | 89.7 | 48.8 | 359.8 | 15,483 | 466 |
| > 250 k | 132.8 | 156.7 | 97.6 | 410.2 | 20,873 | 95 |

*Note:* Number of observations is 36,211.

[5] It follows Lei Complementar 116/2003 and Emenda Constitucional 37/2002.

[6] See Carvalho (2006) for a comprehensive discussion in Portuguese.

[7] There are two other additional types of intergovernmental transfers: Federal Support for Education Fund (FUNDEB, in Portuguese), and Federal Support for Health System (SUS, in Portuguese). Together, those grants account for on average approximatelly 10% of municipalities revenue resources. However, as opposed to FPM grants, they are denominated matching grant transfers as they produce different allocative effects because they impose a counterpart spending for local governments. Because of that, we do not include those grants in our analysis.

in a unified way.[8] The implication of this tax regime for this paper is that is that a share of the business tax revenue, mostly in small cities, is collect in this special tax regime. On average, around 33% of the tax collect in the "Simples" regime was redistributed to municipalities as local business tax. In our methodological approach, we try to circumvent this issue including a variable to control for fiscal capacity (see section 3.3). It is important to note that the "Simples" has reached more economic activities and firms after 2014 (as a consequence of Lei Complementar 147/2014), and because of that our sample should be less affected by this tax regime.

An important characteristic of FPM transfers is that it follows exogenous criteria based on populations cutoffs. Excluding the 27 State capitals and some large cities, municipalities which follows different distributive rules (based on both population and income inverse rules), municipalities with less than 142 thousands inhabitants receive FPM different amounts based only on populational ranges. Those ranges are on Table A-1 in Appendix. What makes this interesting for our methodological approach is that it allow us to explore this legal framework in a regression discontinuity design[9] as discussed in further detail in the next section.

# 3. Methodology

## 3.1 Empirical strategy

As discussed by Dahlby (2008), there are several methodologies to calculate the MCF. The method we follow consists of estimating the sensitivity of the tax base in relation to the tax rate as in equation (1).[10] Thus, the parameter of interest is the sensivity of the tax base to tax rates as depicted in function (5):

$$\log B_{i,t} = \alpha_i + \eta\tau_{it} + \pi X_{it} + dtime_t + municipalities + \xi_i, \tag{5}$$

where $B_{i,t}$ is the tax base of municipality $i$ at time $t$; $\tau$ corresponds to the tax rate; $\xi$ is a vector of control variables; and *dtime* are time dummies. Therefore, the coefficient of the tax rate in this regression is the semi-elasticity of the tax base in relation to the tax rate ($\eta$). From our analytical discussion in section 2, we expect a negative coefficient for $\eta$ as a tax rate increase should present a negative effect on tax base.

The next step is to identify whether the increase in grants and income acts to reduce the municipal MCF according to the equation below:

$$MCF_{it} = \alpha_i + \beta grants_{it} + \gamma Y_{it} + \rho X_{it} + dtime_t + municipalities + u_{it}, \tag{6}$$

where $MCF_{it}$ corresponds to MCF of municipality $i$ at time $T$; $grants_{it}$ is the transfers received by the municipalities; $Y$ captures exogenous income; and $X$ includes the same vector of control variables used in equation (5). Regarding our empirical approach, since

[8] It follows Lei Complementar number 123/06 and Lei Complementar 155 (2016). For example, in year 2010 firms were shared in 20 different groups according their total revenue. Statutory tax rate varied from 4.5% to 16.85%.

[9] We should note that this fact has being previously explored by and Brollo and Nannicini (2012) in a regression discontinuity design (RDD).

[10] There is a large literature on the determination of the response of the tax base in relation to a tax variation. Gruber and Saez (2002) perform this estimation for the USA, and Mintz and Smart (2004) do so for Canada. Dahlby (2008) provides a good review of the topic and also shows results in the context of fiscal federalism.

we have obtained the MCF from equation (5) we should calculate standard errors using bootstrapping technique.

According to public finance literature, we should expect that municipalities decisions over local tax rate (tax revenue) and revenues on intergovernamentals grants are jointly determined which might bias the estiation of equation (6). For this reason, Ordinary Least Squares (OLS) regression would produce inconsistent and biased coefficient estimates. In order to obtain valid estimates of MCF, we explore the fact that the FPM grants follow populationonal ranges and build binary variables based on FPM populational cutoffs. Thus, to estimate the effect of the grants on the municipalities's MCF, we use a regression discontinuity design approach.

## 3.2 FPM thresholds

The most important transfer to Brazilian Municipalities is called the Municipalities' Participation Fund (Fundo de Participação dos Municípios, FPM). It is a mandatory, unconditional, nonmatching, revenue-sharing grant. According to Law nº 5.172/1966, out of the total amount of FPM transfers, 10% belongs to the capitals, 86.4% belongs to the municipalities of the countryside, and the remaining 3.6% constitutes the Reserve Fund, which is distributed among countryside municipalities with more than 142,633 inhabitants (Decree Law nº 1.881/1981 and Complementary Law nº 91/1997, art. 3).

Because the amount that must be allocated to each municipality is clearly established, FPM transfers are transparent and can be separated from political pressure. Furthermore, local governments have autonomy in spending those transfers.[11] On one hand, such autonomy is desirable because the municipality can allocate the transferred resources according to local needs and preferences, which they are in a better position to understand. On the other hand, this autonomy can lead to diminished returns to scale in spending and less tax effort. The Brazilian Census measures the official number of inhabitants in any particular municipality every ten years, and Instituto Brasileiro de Geografia e Estatistica (IBGE) provides annual estimates of local populations between census years. The central government then uses these estimates to distribute FPM funds to municipalities. We explore the yearly fluctuation of local populations across thresholds to quantify the causal effects of government spending on economic activity.

One important aspect of government expenditures and budget resources is that they can be simultaneously determined (Dahlberg, Mörk, Rattsø, & Ågren, 2008). To circumvent this issue, the empirical approach reconciles two exogenous aspects of categorical and block grants using a two-stage least squares (TSLS) method. Regarding block grants, the distribution rules based on the population threshold can used as IV in the first-stage regression.

More specifically, the method uses a regression discontinuity strategy in which each population cut-off is used to build a binary variable that denotes the unconditional transfer range. This allows us to build 17 variables to use as instruments in the first-stage regression, as follows: $D_i = 1$ if $Population_i > Cutoff_r$ at each cutoff $r$ and municipality $i$.

[11] Although Litschig and Morrison (2013) shows that the population measure could be manipulated by municipalities in the 1990s, such manipulation has not been observed since 2000, when the federal governments started to use census data and IBGE estimates for the population. See also Brollo, Nannicini, Perotti, and Tabellini (2013) and Ferraz and Finan (2011) for further evidence and discussion.

As discussed by Litschig and Morrison (2013), there is no reason to believe that sorting around cut-offs (manipulations of the true size of the population, creation of new municipalities above the cut-offs or incentive effects to migrate) affects the identification strategy. Thus, TSLS estimations use selected cut-offs as excluded instruments to obtain consistent estimates of the causal effect of grants on per pupil spending (Van der Klaauw, 2002).

Equation (7) provides the first-stage regression for a IV approach in which FPM grants are treated as an endogenous regressor. It uses the population dummies (17 population dummies $Z$) as valid instruments for the unconditional grant. All other variables are used as valid instruments for themselves:

$$FPM_{it} = \alpha_0 + \beta_1 Z_1 + \beta_2 Z_2 + \cdots + \beta_{17} Z_{17} + \rho X_{it} + dtime_t + municipalities_i + u_{it}, \quad (7)$$

where $FPM$ is instrumented by the population dummies ($Z$); all other controls ($X$), time effects; *municipalities*, fixed effects; and $u_{it}$ is an error term clustered at the municipality level.

The causal effect that we are identifying is local for two reasons: first, the RDD identifies the effect only to observations around the thresholds; and second, the to compliers (that is, municipalities that received larger transfers because of the exogenous FPM revenue-sharing mechanism). However, the identification of compliers neglects a subpopulation that may also be of interest: the always-takers, or municipalities that receive larger transfers independently of their position above or below each population threshold.

Next we run this instrumented FPM on the computed MCF to identify the price effect and equation (6) would modify to:

$$MCF_{it} = \alpha_0 + \delta_1 \widehat{FPM}_1 + \cdots + \delta_{17} \widehat{FPM}_{17} + \rho X_{it} + dtime_t + municipalities_i + v_{it}$$

We should note that this formulation of the MCF relies on the following hypotheses: i) changes in ISS rates produce very small effect over IPTU tax revenues; ii) each municipality's ISS tax burden is supported only by the its own; iii) there is no interaction between the tax bases—which makes sense for services in small municipalities; iv) the elasticity of the tax base may be the same in all of the local governments.

## 3.3 Data

Regarding local tax and revenue information, we note that annual data is available in a systematic way until year 2012 from Brazilian Treasury (Secretaria do Tesouro Nacional, STN). Data on number of service firms (which compound ISS tax base), and their number of employes and firms's payroll total costs is available in the Brazilian Bureau of Geography and Statistics' reports on Services Firms Census (Cadastro Central de Empresas) from 2006 to 2015.[12] Considering the time period availability of both datasets, our sample includes data for 5,173 (from a total of 5,565) municipalities in Brazil from year 2006 to 2012 (seven years period). We also collect data on demographic variables as described Table 2.[13] We note that all monetary variables are deflated to year 2010 in Brazilian currency (R$ Reais) values.

[12] Availabe at https://sidra.ibge.gov.br/pesquisa/cempre/tabelas/brasil/2015 as in January 2018.

[13] Datasus data is available at http://www2.datasus.gov.br/DATASUS/index.php?area=0206 and HDI at http://www.firjan.com.br/ifdm/

**Table 2.** Descriptive statistics of observable variables from municipalities – from 2006 to 2012, in R$, year 2010.

| Variable | Definition | Source | Mean | Std. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| FPM | Total FPM transfers revenues | Finbra dataset | 9,182,481 | 17,600,000 | 3,293 | 802,000,000 |
| ISS | Total service tax revenue | Finbra dataset | 4,966,098 | 106,000,000 | 11 | 11,200,000,000 |
| Payroll | Total payroll costs for firms in service sector | Cadastro Central de Empresas | 139,940 | 2,199,693 | 2 | 206,000,000 |
| HDI | Human Development Index | FIRJAN | 0.616 | 0.119 | 0.237 | 0.916 |
| Firms | Number of firms in service sector | Cadastro Central de Empresas | 934 | 8,555 | 4 | 599,434 |
| Ln Employees | Natural logarithm of employees in the service sector | Cadastro Central de Empresas | 7.21 | 1.46 | 1.61 | 15.62 |
| Tax capacity | State grant from VAT (ICMS) normalized | Authors | 0.00 | 1.00 | −1.02 | 43.86 |
| AETR | Average effective tax rate (ISS revenue/ payroll costs) | Authors | 0.040 | 0.225 | 0.000 | 30.076 |
| MCF | Marginal costs of public fund | Authors | 0.001 | 0.006 | 1.000 | 1.946 |
| Metropolitan region | Binary variable (=1) if municipality is in metropolitan area | IBGE | 0.138 | 0.345 | 0 | 1 |
| Ln Population | Natural logarithm of local population | IBGE | 9.42 | 1.14 | 6.69 | 16.25 |
| Density | Populational density | IBGE | 103.8 | 546.0 | 0.1 | 13,347 |
| Share of young | Individuals from 5 to 15 years old | DATASUS | 0.266 | 0.052 | 0.074 | 0.513 |
| Share of elderly | Individuals above 60 years old | DATASUS | 0.116 | 0.031 | 0.021 | 0.294 |

*Note:* Number of observations is 36,211.

As previously debated by several authors (see Hakonsen (1998) for a discussion), one of the main limitations of the applied literature is to identify a feasible measure of tax base for local tax to properly approach the marginal cost of public funds. One reasonable measure of firms tax burden would be the local business tax (ISS). However, there is not systematic data available for ISS tax rates across municipalities. Differently, we have information available about the payroll expenses (firms expenses with employee earnings) by municipality. Based on this information, we build a measure of effective average tax rate (AETR) as suggested by Hayashi and Boadway (2001). Thus, instead of using statutory tax rates, our measure of AETR follows the ratio between ISS tax revenues and annual payroll of firms services under the municipalities in the analysis. The idea is that in the service sector, firms' revenue is strongly correlated to firms' labor spending or to payroll spending. The main advantage of using a AETR approach is that it is less rigid than statutory tax rate and provides more variation in time and across units. Additionally, it should capture better the effect of tax rate changes on number of firms and firms revenues.

Finally, following previous work on tax competition literature, we include a covariate to control for fiscal capacity. We build a measure of relative fiscal capacity based on the amount each municipality obtains from Value Added Tax (ICMS,[14] in Portuguese). We explore the conception that fiscal capacity is positively associated to VAT collection at the municipal level (see also section 2). We normalize the revenues originated from this devolutive grant to obtain a national average of zero. The idea is that if a municipality present a VAT refund revenue above (below) the national average it should present a higher (lower) fiscal capacity compared to their counterparts. This covariate should account for local fiscal capacity as also should minimize a potential measurement error in AETR in smaller cities, originated from the "Simples Nacional" tax regime. It should also account for informality in the municiplaities. Next Section brings our main findings.

# 4. Result

## 4.1 Tax base

We first estimate the tax base elasticity as described in equation (5). We consider as measure of tax base the natural logarithm of the number of firms in the service sector. As discussed in section 2, it is expected a negative correlation between the average effective tax rate (AETR) and the tax base ($\eta < 0$). Table 3 brings the results for the RDD approach for 5 cutoffs with three different bandwhidth ranges and for the total of observations in our sample. We note that after the $5^{th}$ cutoff the number of cases around the cutoff is small. RDD approach would be possible only with very large bandwidths. All regressions include demographic variables (population, squared population, population density, share of younger and share of elderly, and a binary variable indicating if the municipality is in a metropolitan region or not) as covariates.

Overall, as expected, results on Table 3 shows that a tax rate increase has a negative effect on the number of firms in the municipalties. In column I, results in RDD on the first cutoff indicate that an one standard deviation increase in average tax rate reduces in 4.5% ($0.2 \times 0.225 \times 100$) the number of service firms. Interestingly, results in Table 3 suggests that the negative effect is neither homogenous nor linear around the different cutoffs. Estimated coeffcients suggests that the results are negative and significant for the smaller (first

[14] Imposto Sobre Circulação de Mercadorias e Serviços.

**Table 3.** Regression on tax base elasticity. Dependent variable: Ln of number of firms.

| Bandwidth size | 500 | 1,000 | 1,500 | 2,000 | 3,000 | 4,000 | Total |
|---|---|---|---|---|---|---|---|
| Cut-off 1 | −0.200*** | −0.107** | −0.052 | | | | |
| | (−5.310) | (−2.530) | (−1.321) | | | | |
| Number of cases | 1,355 | 2,687 | 3,904 | | | | |
| Cut-off 2 | | −0.010 | 0.002 | −0.002 | | | |
| | | (−1.448) | (0.374) | (−0.175) | | | |
| Number of cases | | 2,003 | 2,947 | 3,914 | | | |
| Cut-off 3 | | | 0.175*** | 0.110 | −0.007 | | |
| | | | (3.222) | (0.980) | (−0.150) | | |
| Number of cases | | | 2,214 | 2,476 | 4,298 | | |
| Cut-off 4 | | | | −0.250 | 0.002 | 0.194 | |
| | | | | (−1.191) | (0.011) | (0.971) | |
| Number of cases | | | | 1,769 | 2,596 | 3,520 | |
| Cut-off 5 | | | | −0.630** | −0.570** | −0.476** | |
| | | | | (−2.359) | (−2.465) | (−2.437) | |
| Number of cases | | | | 997 | 1,519 | 2,023 | |
| Total | | | | | | | −0.016** |
| | | | | | | | (−2.463) |
| Number of cases | | | | | | | 36,211 |

*Notes:* All regressions include fixed effect and time effects. In parentheses are robust $t$ statistics. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

cutoff) and larger (fifth cutoff) municipalities in our sample. Those cutoffs corresponds to municipalities with approximatelly 10 and 30 thousand inhabitants, respectively. Differently, coefficients in cutoff 3 and 4 suggest (municipalities with approximately 16 and 23 thousand inhabitants, respectively) that a tax rate increase present a positive on the number of firms in those jurisdictions (although this positive effect is significant only in one out of fifteen regression).

This trend is depicted in Figure 1 which brings the scatter plot between tax rate and the logarithm of the number of firms for three different bandwidth ranges for the five cutoffs as in Table 3. If someone depicts a linear trend line, she will note that the effect on tax base is negative as expected. Contrastingly, if someone depicts a polynomial trend line, she will obtain an inverted U shape relationship. Importantly, results suggest that, after the 2$^{nd}$ cutoff, the negative tax rate effect on tax base increases as municiplaities gets larger. One possible explanation for those findings is that the most populated municipalities presents the largest firms which are more affected by a tax rate increase. On the other hand, as developed by Carbonnier (2008), the effect of tax competition can be greater for smaller jurisdictions with less fixed factor, as land for example. The idea is that very small jurisdictions can be more affected by tax rate spill-over if the externality is driven by a fixed factor.

It is important to note that very small firms in Brazil collect taxes in a diferentied mechanism, called "Simples Nacional" (see section 3 for a brief discussion). To verify if

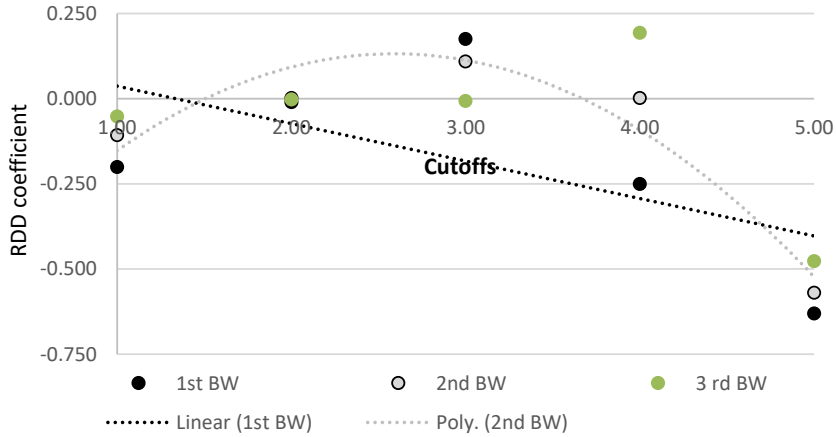



**Figure 1.** Tax Rate Effect on Tax Base.

the results in smaller municipalities could be biased because of this tax regime, we include in all regresions a covariate to control for tax capacity. Interestingly, if someone compare the results of the RDD regression with and without this variable (Table 3 and Table A-2 in Appendix, respectivelly), she will note that the inclusion of this variable leads to similar coefficient in most regressions but to a greater coefficient in the first cutoff. This result suggests that our measure of tax rate could not capture all the effects of tax rate on tax base because of this differentiated tax regime. However, this limitation seems circumnscrit to the smallest municiplaities (1$^{st}$ cutoff) in our sample, since the estimated coefficient over the other cutoffs are similar with or without this covariate.

Finally, we also report results for a standard panel fixed effect model. We note that regression on our total sample (column VI) seems to reflect the relantionship between tax rate and number of firms in the smaller and greater municipalities, and the estimated coefficient on the tax rate indicates that one standard deviation increase in tax rate reduces 0.36% the number of firms in the municipality. It is important to remind that, as discussed in section 2, we use the estimated coefficient of tax rate on tax base to build our measure of Marginal Cost of Fund (MCF) as described in function (1). In the next section, we discuss the results over this dependent variable.

## 4.2 MCF's Price-effect

In this section we estimate the unconditional grants price effect on the MCF using our model described in section 2. We expect that FPM grants present a negative effect on the MCF since this should reduce the price of public provision as discussed in Dahlby (2011). Table 4 brings the results for our RDD approach around the same five cutoffs as in Table 3 and for two stage least squares instrumental variable (2SLS-IV) with bootstrapping standard errors (last column Table 4). We note that, as in Table 3, all regressions control for demographic variables as also for Human Development Index (HDI) indicator, and fixed and time effects.

Overall, we find limited evidence of a negative effect of unconditional grants on MCF. Estimated coefficents on grants ranges from −0.007 to 0.002. In 9 out 15 RDD results, we find a negative coeffcients for grants, but only one coeffcient is significant at the 10% significance level (larger bandwidth in the 5$^{th}$ cut-off). As the MCF reflects the marginal loss in the economy due to the marginal increase in tax rate, for example, a negative coefficient in the

**Table 4.** Regression on tax price effect. Dependent variable: Ln of number of firms. Dependent variable: Marginal Cost of Public Fund (MCF).

| Bandwidth size | 500 | 1,000 | 1,500 | 2,000 | 3,000 | 4,000 | Total |
|---|---|---|---|---|---|---|---|
| Cut-off 1 | 0.020 | 0.002 | 0.002 | | | | |
| | (0.963) | (0.864) | (1.166) | | | | |
| F-test | 1.01 | 31.73 | 65.16 | | | | |
| Cut-off 2 | | 0.003 | −0.002 | −0.002 | | | |
| | | (0.941) | (−0.434) | (−0.585) | | | |
| F-test | | 8.87 | 25.80 | 56.80 | | | |
| Cut-off 3 | | | −0.000 | 0.000 | 0.001 | | |
| | | | (−0.069) | (0.006) | (0.877) | | |
| F-test | | | 4.60 | 12.27 | 30.56 | | |
| Cut-off 4 | | | | −0.001 | −0.001 | | −0.001 |
| | | | | (−0.641) | (−1.074) | | (−0.955) |
| F-test | | | | 7.24 | 23.77 | | 38.82 |
| Cut-off 5 | | | | −0.007 | −0.005 | | −0.004* |
| | | | | (−0.946) | (−1.347) | | (−1.779) |
| F-test | | | | 1.18 | 4.25 | | 10.16 |
| Total | | | | | | | −0.003** |
| | | | | | | | (−2.276) |
| F-test | | | | | | | 48.37 |

*Notes:* All regressions include fixed effect and time effects. In parentheses are robust bootstrapped $t$ statistics. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

last column (total)[15] in Table 4 indicates that each R$ 1 increase in unconditional grants is associated with a reduction in the cost of public fund about 0.003 (1% of its standard deviation).

To examine if the effect of grants on MCF presents some trend, we plot a graph with the five cutoff and the three different bandwidths. Interestingly, Figure 2 suggests that, although not very large, the negative effects of grants on MCF increases the larger is the municipality. This seems a counter-intuitive result, since small communities tend to present higher marginal cost of public fund and they should present a stronger reaction to additional grants resources (Buettner & Fabritz, 2011).

However, there two possible explanations for our results in Table 4. First, the negative effect of tax rate increase on tax base could be greater in larger cities (Buettner & Holm-Hadulla, 2008), as we have previously observed in the results in Table 3. Second, grants effects on MCF tend to be larger in recipients communities with a large distortionary taxation. For example, using data for Argentinian Provinces. Vegh and Vuletin (2015) show that the

[15]Table A-3 in Appendix brings the coefficents for the first stage regression and it shows that all binary variables are significant. They present negative signal which reflects the fact that the smaller the population the higher the FPM per capita received.

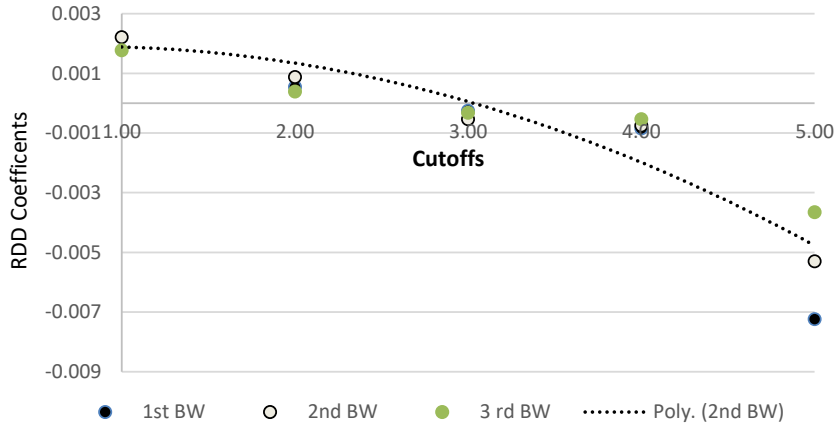



**Figure 2.** Grant Effect on MCF.

flypaper effect (and the stimulative effect of grants on public spending) is larger in localities with tax rate above the national average. In our sample although the average tax rate (AETR) is similar accross municipalites, larger municipalities tend to host larger firms, in terms of number of employees and total revenues (see Figure A-1 in Appendix).

Thus, overall, our results reflects most of the public funding scheme in Brazil, where public provision in smaller cities are financed mostly by grants (this resource accounts for 80% of total public resources) and in larger cities mostly by local taxes. Results in total sample could be capturing this higher price effect of grants on municipalities that present higher distortionary taxation over business services.

## 5. Conclusion

This paper investigates the effect of unconditional grants on the Marginal Cost of Public Funding (MCF) using Brazilian local governments data. In particular, we explore the model proposed in Dahlby (2011), to design a theoretical extension that shows the difference between the price effect resulting from grants and from income attempt to show how these different methods of increasing income change the effective price of public goods.

Using data from Brazilian national Treasury for the years 2006 to 2012 we construct average effective ISS local taxes and unconditional (FPM – the recipient government has complete freedom to decide on the allocation of the resources) transfers to estimate such difference.

We first estimate the effect of tax rates on tax basis in order to construct MCF estimations. That estimated semi-elasticity (−0.016) is used to calculate the MCF, representing the price of the public good for each municipality.

Next, we estimate the effect of transfers and income on the constructed MCF, correcting the standard error of these estimations. We estimate that for R$ 1 increase in unconditional transfers seems to reduce (in average) the local price effect (MCF) but not in a consistent fashion. Considering the entire sample, our estimates suggests that an increase in R$ 1.00 in per capita unconditional transfers reduces the local price effect (MCF) around 0.00003 (0.07% of the average MCF), but this result is not consistently estimated across all subsamples.

## References


Auriol, E., & Warlters, M. (2012). The marginal cost of public funds and tax reform in Africa. *Journal of Development Economics*, *97*(1), 58–72. http://dx.doi.org/10.1016/j.jdeveco.2011.01.003

Brollo, F., & Nannicini, T. (2012). Tying your enemy's hands in close races: The politics of federal transfers in Brazil. *American Political Science Review*, *106*(4), 742–761. http://dx.doi.org/10.1017/S0003055412000433

Brollo, F., Nannicini, T., Perotti, R., & Tabellini, G. (2013). The political resource curse. *American Economic Review*, *103*(5), 1759–1796. http://dx.doi.org/10.1257/aer.103.5.1759

Browning, E. K. (1987). On the marginal welfare cost of taxation. *The American Economic Review*, *77*(1), 11–23. https://www.jstor.org/stable/1806725

Buettner, T., & Fabritz, N. (2011, June 30). Responses to grants and the marginal cost of public funds. In *IV Workshop on Fiscal Federalism*, Barcelona.

Buettner, T., & Holm-Hadulla, F. (2008, November). *Cities in fiscal equalization* (CESifo Working Paper No. 2447). Munich, Germany: CESifo Group. http://www.cesifo-group.de/ifoHome/publications/docbase/DocBase_Content/WP/WP-CESifo_Working_Papers/wp-cesifo-2008/wp-cesifo-2008-11/12012008002447.html

Carbonnier, C. (2008, February). *Fiscal competition between decentralized jurisdictions, theoretical and empirical evidence* (Documents de travail No. 2008-17). Cergy-Pontoise: THEMA – Université de Cergy-Pontoise. http://thema.u-cergy.fr/IMG/documents/2008-17.pdf

Carvalho, P. H. B. d., Jr. (2006, December). *IPTU no Brasil: Progressividade, arrecadação e aspectos extra-fiscais* (Texto para Discussão No. 1251). Brasília, DF: Instituto de Pesquisa Econômica Aplicada (IPEA). http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=4814

Cossio, F. A. B. (2002). *Ensaios sobre federalismo fiscal no Brasil* (PhD Thesis, PUC-Rio, Rio de Janeiro). http://dx.doi.org/10.17771/PUCRio.acad.3712

Cossio, F. A. B., & Carvalho, L. M. d. (2001). Os efeitos expansivos das transferências intergovernamentais e transbordamentos espaciais de despesas públicas: Evidências para os municípios brasileiros – 1996. *Pesquisa e Planejamento Econômico*, *31*(1), 75–124. http://ppe.ipea.gov.br/index.php/ppe/article/view/161

Dahlberg, M., Mörk, E., Rattsø, J., & Ågren, H. (2008). Using a discontinuous grant rule to identify the effect of grants on local taxes and spending. *Journal of Public Economics*, *92*(12), 2320–2335. http://dx.doi.org/10.1016/j.jpubeco.2007.05.004

Dahlby, B. (2008). *The marginal cost of public funds: Theory and applications*. Cambridge, MA: MIT Press.

Dahlby, B. (2011). The marginal cost of public funds and flypaper effect. *International Tax and Public Finance*, *18*(3), 304–321. http://dx.doi.org/10.1007/s10797-010-9160-x

Ferraz, C., & Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, *101*(4), 1274–1311. http://dx.doi.org/10.1257/aer.101.4.1274

Fisher, R. C. (1982). Income and grants effects on local expenditure: The flypaper effect and other difficulties. *Journal of Public Economics*, *17*(3), 324–345. http://dx.doi.org/10.1016/0094-1190(82)90021-3

Gamkhar, S., & Shah, A. (2007). The impact of intergovernmental transfers: A synthesis of the conceptual and empirical literature. In R. Boadway & A. Shah (Eds.), *Intergovernmental fiscal transfers: Principles and practice* (pp. 225–258). Washington, DC: World Bank. http://hdl.handle.net/10986/7171

Gramlich, E. (1977). Intergovernmental grants: A review of the empirical literature. In W. E. Oates (Ed.), *The political economy of fiscal federalism.* Lexington Press.

Gruber, J., & Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics*, *84*(1), 1–32. http://dx.doi.org/10.1016/S0047-2727(01)00085-8

Hakonsen, L. (1998). An investigation into alternative representations of the marginal cost of public funds. *International Tax and Public Finance*, *5*(3), 329–343. http://dx.doi.org/10.1023/A:1008686227749

Hamilton, J. H. (1986). The flypaper effect and the deadweight loss from taxation. *Journal of Urban Economics*, *19*(2), 148–155. http://dx.doi.org/10.1016/0094-1190(86)90036-7

Hayashi, M., & Boadway, R. (2001). An empirical analysis of intergovernmental tax interaction: The case of business income taxes in Canada. *Canadian Journal of Economics*, *34*(2), 481–503. http://dx.doi.org/10.1111/0008-4085.00085

Hines, J. R., & Thaler, R. H. (1995). The flypaper effect. *Journal of Economic Perspectives*, *9*(4), 217–226. http://dx.doi.org/10.1257/jep.9.4.217

Inman, R. P. (2008, December). *The flypaper effect* (Working Paper No. 14579). National Bureau of Economic Research (NBER). http://dx.doi.org/10.3386/w14579

Jacobs, B. (2018). The marginal cost of public funds is one at the optimal tax system. *International Tax and Public Finance*, *25*(4), 883–912. http://dx.doi.org/10.1007/s10797-017-9481-0

Lanzer, B. N., & Porto, S. d. S., Jr. (2011, December). A economia informal e o custo marginal da tributação no Brasil. In *39º Encontro Nacional de Economia da ANPEC*, Foz do Iguaçu. http://anpec.org.br/encontro/2011/inscricao/arquivos/000-614536d09ffbc91573943c164f21ad09.pdf

Litschig, S., & Morrison, K. M. (2013). The impact of intergovernmental transfers on education outcomes and poverty reduction. *American Economic Journal: Applied Economics*, *5*(4), 206–240. http://dx.doi.org/10.1257/app.5.4.206

Mintz, J., & Smart, M. (2004). Income shifting, investment, and tax competition: Theory and evidence from provincial taxation in Canada. *Journal of Public Economics*, *88*(6), 1149–1168. http://dx.doi.org/10.1016/S0047-2727(03)00060-4

Siqueira, R. B. d., Nogueira, J. R. B., Souza, E. S. d., & Carvalho, D. B. (2010, December). O custo marginal social da tributação indireta no Brasil: Identificando direções de reforma. In *38º Encontro Nacional de Economia da ANPEC*, Salvador, BA. http://www.anpec.org.br/encontro2010/inscricao/arquivos/271-8dbd10a9e0435a7e286905e9dda8141a.doc

Van der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review*, *43*(4), 1249–1287. http://dx.doi.org/10.1111/1468-2354.t01-1-00055

Vegh, C. A., & Vuletin, G. (2015). Unsticking the flypaper effect in an uncertain world. *Journal of Public Economics*, *131*, 142–155. http://dx.doi.org/10.1016/j.jpubeco.2015.09.001

Wyckoff, P. G. (1991). The elusive flypaper effect. *Journal of Urban Economics*, *30*(3), 293–305. http://dx.doi.org/10.1016/0094-1190(91)90052-9

# Appendix.

**Table A-1.** Population cutoffs for FPM (transfers) distribution.

| Population Range | Factor |
|---|---|
| Up to 10,188 | 0.6 |
| From 10,189 to 13,584 | 0.8 |
| From 13,585 to 16,980 | 1.0 |
| From 16,981 to 23,772 | 1.2 |
| From 23,773 to 30,564 | 1.4 |
| From 30,565 to 37,356 | 1.6 |
| From 37,357 to 44,148 | 1.8 |
| From 44,149 to 50,940 | 2.0 |
| From 50,941 to 61,128 | 2.2 |
| From 61,129 to 71,316 | 2.4 |
| From 71,317 to 81,504 | 2.6 |
| From 81,505 to 91,692 | 2.8 |
| From 91,693 to 101,880 | 3.0 |
| From 101,881 to 115,464 | 3.2 |
| From 115,465 to 129,048 | 3.4 |
| From 129,049 to 142,632 | 3.6 |
| From 142,633 to 156,216 | 3.8 |
| > 156,216 | 4.0 |

*Note:* Law by Decree nº 1.881/1981.

**Table A-2.** Regression on tax base elasticity (without tax capacity covariate). Dependent variable: Ln of number of firms.

| **Bandwidth size** | **500** | **1,000** | **1,500** | **2,000** | **3,000** | **4,000** | **Total** |
|---|---|---|---|---|---|---|---|
| Cut-off 1 | −0.140*** | −0.094** | −0.050 | | | | |
| | (−3.129) | (−2.453) | (−1.430) | | | | |
| Number of cases | 1,355 | 2,687 | 3,904 | | | | |
| Cut-off 2 | | −0.023 | −0.009 | −0.013 | | | |
| | | (−1.114) | (−0.507) | (−0.640) | | | |
| Number of cases | | 2,003 | 2,947 | 3,914 | | | |
| Cut-off 3 | | | 0.177*** | 0.111 | −0.041 | | |
| | | | (3.253) | (0.993) | (−0.934) | | |
| Number of cases | | | 2,214 | 2,476 | 4,298 | | |
| Cut-off 4 | | | | −0.249 | −0.004 | | 0.191 |
| | | | | (−1.182) | (−0.021) | | (0.957) |
| Number of cases | | | | 1,769 | 2,596 | | 3,520 |
| Cut-off 5 | | | | −0.629** | −0.570** | | −0.476** |
| | | | | (−2.357) | (−2.467) | | (−2.437) |
| Number of cases | | | | 997 | 1,519 | 2,023 | |
| Total | | | | | | | −0.016** |
| | | | | | | | (−2.463) |
| Number of cases | | | | | | | 37,708 |

*Notes:* All regressions include fixed effect and time effects. In parentheses are robust $t$ statistics. ***$p < 0.01$, **$p < 0.05$, *$p < 0.1$.

**Table A-3.** First Stage (IV) regression results.

| Population cutoffs (binary variables) | | |
|---|---|---|
| dfxpop1 | −0.446*** | (−19.360) |
| dfxpop2 | −0.310*** | (−15.240) |
| dfxpop3 | −0.211*** | (−11.530) |
| dfxpop4 | −0.147*** | (−9.220) |
| dfxpop5 | −0.102*** | (−7.590) |
| dfxpop6 | −0.064*** | (−6.300) |
| dfxpop7 | −0.029*** | (−4.100) |
| dfxpop9 | 0.020** | (2.040) |
| dfxpop10 | 0.007 | (0.820) |
| dfxpop12 | 0.008 | (0.710) |
| dfxpop14 | −0.010 | (−0.440) |
| dfxpop16 | −0.006 | (−0.360) |

*Notes:* All regressions include fixed effect and time effects. In parentheses are $t$ statistics. Number of obs is 36,211. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

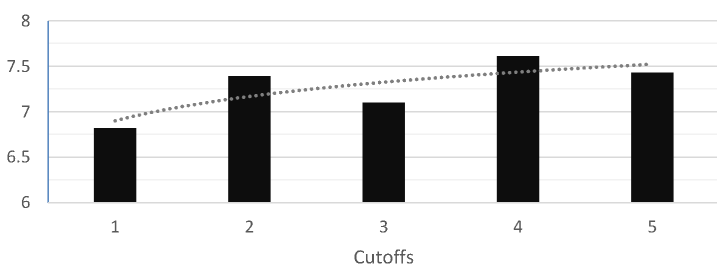



**Figure A-1.** Firms' Size (number of Employees).

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180024
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Impacto do Financiamento Habitacional Sobre o Mercado de Trabalho na Construção Civil

Fábio Nishimura*
Carlos Eduardo de Freitas*
Roselaine Almeida†

**Sumário**

1. Introdução ..... 497
2. O Programa de Arrendamento Residencial ..... 498
3. Dados ..... 500
4. Estratégia Empírica ..... 503
5. Análise dos Resultados ..... 507
6. Considerações Finais ..... 511

**Palavras-chave**
Financiamento habitacional, trabalho e salário, regressão descontínua

**JEL Codes**
J20, J30, J48

**Resumo · Abstract**

In 2001, the Brazilian government, concerned about housing deficit problems, created the Housing Leasing Program (PAR). As a main prerogative, we believe that the program, in addition to reducing the housing deficit, also influences the labor market in the construction sector. In order to check this hypothesis, we use as an empirical strategy a Regression Discontinuity Design, besides applying several tests of robustness in our model, which guarantees us more accurate results. We have verified that the PAR can increase the jobs and the wages of the actors inserted in the labor market of the civil construction.
inseridos no mercado de trabalho da construção civil.

## 1. Introdução

O déficit habitacional sempre foi um problema social para diversos países no mundo. A falta de moradia leva as pessoas a apresentarem baixos níveis de qualidade de vida, provocando a deterioração da condição humana (Onibokun, 1983 e Salau, 1990). Segundo Bouillon (2012), uma em cada três famílias na América Latina, bem como no Caribe, vive em moradias que são inadequadas para a habitação. Países como Nicarágua, Honduras, Guatemala, Peru e Bolívia possuem mais de 50% de suas famílias morando em habitações subnormais devido ao déficit habitacional (Bouillon, 2012).

O Brasil, preocupado com seu déficit, elaborou e aplicou ações voltadas à política habitacional ao longo dos anos. Por volta de 1946, foi criado a instituição "Fundação Casa Popular" que visava, de forma específica, diminuir as populações moradoras das favelas e de mocambos no Distrito Federal. Atualmente existe no país o programa "Minha Casa, Minha Vida" que objetiva, essencialmente, a redução do déficit habitacional dos moradores de baixa renda no Brasil, com reflexos no crescimento econômico nacional. Essas ações foram criadas com o intuito de amortizar o déficit habitacional, porém, mesmo com tais medidas ainda persiste, em grau elevado, a questão da falta de moradia e suas consequências como falta de segurança, aumento de doenças e exclusão social.

*Universidade Federal de Mato Grosso, Instituto de Ciências Humanas e Sociais (UFMT/ICHS), Campus Rondonópolis. Avenida dos Estudantes 5055, Bairro Sagrada Família, Rondonópolis, MT, CEP 78735-901, Brasil.

†Universidade Federal da Grande Dourados, Faculdade de Administração, Ciências Contábeis e Economia (UFGT/FACE). Rua João Rosa Góes 1761, Vila Progresso, Dourados, MS, CEP 79825-070, Brasil.

f_nobuo@yahoo.com.br cefreitas@ufmt.br rose2almeida@yahoo.com.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Recebido em 16 de março de 2017
Aceitação final em 30 de abril de 2018

Esses planos e programas habitacionais, além das metas de redução dos déficits de moradias, causam efeitos inesperados (ou até certo ponto esperado) no ambiente, e que por muitas vezes, não foram planejados. Pode-se afirmar que um programa dessa natureza, criado com intuito de reduzir o déficit habitacional, acaba também promovendo uma série de acontecimentos, tais como: alteração na arrecadação dos impostos (Teixeira & Carvalho, 2005), no direito de uso da terra (Pollakowski & Wachter, 1990; Peng & Wheaton, 1994) e, principalmente, a criação de novos empregos formais e informais e o aumento de salários dos empregados em setores chaves (Henley, 2005; Saks, 2008; Zhao, Lü, & De Roo, 2011).

O aumento do número de empregos e dos salários das pessoas empregadas na construção civil é muito benéfico para economia como um todo, pois gera renda, consumo e aumento do produto interno bruto dos países, o que, resumindo, leva a um crescimento econômico devido a seu efeito multiplicador (Teixeira & Carvalho, 2005).

Nesse sentido, o objetivo deste trabalho é identificar se, de fato, os programas voltados para construção habitacional, conseguem melhorar o nível de empregos e salários na economia através de seu efeito multiplicador, como fora levantado por diversos trabalhos, entretanto, a estratégia empírica adotada foi o nosso diferencial. Desta forma, analisamos o Programa de Arrendamento Residencial (PAR), criado pelo Governo Federal através da Lei 10.188 de 12 de fevereiro de 2001, onde a dinâmica do programa é financiar, por meio da modalidade de leasing financeiro, a construção e reforma de imóveis, cuja meta principal é destinar moradias de qualidade para pessoas e a meta secundária é a criação de postos de empregos diretos e indiretos.

Para realizar a análise do efeito do Programa de Arrendamento Residencial sobre o ramo da construção civil, utilizamos como instrumento estatístico um modelo de Regressão Descontínua (RD), onde através de um ponto de corte (*cutoff*) existente no programa (população mínima de 100.000 habitantes) podemos construir um contrafactual que responda nossa hipótese de que o PAR consegue aumentar os empregos e o valor dos salários. Dessa forma os resultados mensurados por este método nos garantem uma defesa contra problemas de endogeneidade.

Buscando a relação causal do trabalho, nos preocupamos com possíveis problemas de viés de seleção e alguma existência de manipulação em nosso *cutoff*, fator principal para a estratégia empírica, assim, aplicam-se testes para verificar a possibilidade de rejeição de tais hipóteses. Por último, para garantir os resultados estimados, são realizados testes de robustez.

Para um maior aprofundamento da análise, este trabalho está dividido em 6 partes além desta introdução. A seção 2 discute o Programa de Arrendamento Residencial através da sua forma de legalização, seus objetivos, suas características além de seus critérios de adoção por parte do município e da população. Na seção 3, apresenta-se os dados utilizados no modelo de regressão: onde foram extraídos, qual o período, e quais as especificidades das variáveis de dependentes e independentes. Na seção 4 encontra-se a estratégia empírica utilizada, sendo o desenho de regressão descontinua e seus testes de robustez. Ao final segue a análise dos resultados e considerações finais.

## 2. O Programa de Arrendamento Residencial

A preocupação em relação a questão habitacional brasileira sempre motivou profundas discussões entre os especialistas da área em relação aos efeitos ocorridos e as ações executadas pelos Governos. Algumas propostas relacionadas às políticas habitacionais, como

Fundação da Casa Popular em 1946, o Banco Nacional de Habitação (BNH) e o Sistema Nacional de Habitação (SNH) em 1964, o Programa de Financiamento de Lotes Urbanizados (PROFILURB), o Financiamento de Construção, Conclusão ou Melhoria de Habitação de Interesse Popular (FICAM) e o Programa de Erradicação de Subhabitação (PROMORAR), todos em vigor na década de 1970, foram instituídas com o propósito similar de suprir a necessidade habitacional da camada mais pobre da sociedade. Os anos da década de 1980 e de 1990 não foram expressivos se tratando de políticas habitacionais, gerando um déficit ainda maior no problema há anos existente.

Seguindo a vertente da preocupação habitacional, o Governo Federal, cria, através da Lei 10.188 de 12 de fevereiro de 2001, o Programa de Arrendamento Residencial (doravante PAR), de origem do executivo, por meio dos órgãos da Secretaria de Desenvolvimento Urbano e do Ministério da Previdência e Assistência Social. Segundo CAIXA (2003), o objetivo do programa é o fomento à oferta de unidades habitacionais e à melhoria das condições do estoque de imóveis existentes, a promoção da melhoria da qualidade de vida das famílias beneficiadas, a intervenção em áreas objeto de Planos Diretores, a criação de novos postos de trabalho diretos e indiretos, o aproveitamento de imóveis públicos ociosos em áreas de interesse habitacional e o atendimento aos idosos e portadores de deficiência física.

Com o passar do tempo o PAR sofreu várias alterações e reedições para a sua adequação em relação ao cenário econômico. Quanto a responsabilidade do projeto, passando a ser responsável pela gestão o Ministério das Cidades, mas mantendo sua execução sob domínio da Caixa Econômica Federal (CAIXA). Em relação ao órgão responsável projeto também houve mudanças, pois, o programa passou a ser controlado pelo Ministério das Cidades, porém mantendo sua execução sob domínio da Caixa Econômica Federal (CAIXA).

Além do Ministério das Cidades e da CAIXA, há o envolvimento dos Estado, DF e Municípios, identificando os locais para implantação dos projetos; indicação das famílias a serem beneficiadas; promovendo ações facilitadoras e redutoras dos custos de implantação dos projetos, tais como, redução de tributos, contribuições e taxas; trazendo recursos financeiros, bens ou serviços economicamente mensuráveis, necessários à realização das obras e serviços do empreendimento. Também participam do programa as empresas da construção civil e as empresas do ramo de administração imobiliária, executando atividades técnicas e operacionais de construção e do gerenciamento de contratos de arrendamento dos imóveis e dos condomínios (CAIXA, 2003).

O PAR financia recursos para pessoas que têm interesse em construir novos empreendimentos, mesmo estando em etapas diferentes de construção. Assim, imóveis ainda na planta, em fases intermediária de construção ou em fase de recuperação/restauração do imóvel, desde que o arrendatário contemple os critérios que o torne elegível ao programa, será beneficiado pela ação. O PAR tem como prerrogativa, atender as famílias com renda mensal de até R$1.800,00 (um mil e oitocentos reais). No caso de profissionais da área de segurança pública, especialmente os policiais civis e militares, admite-se renda mensal de até R$2.400,00 (dois mil e quatrocentos reais).

Como condições para acesso ao recurso, o Programa de Arrendamento Residencial (Brasil, 2001) apresenta o seguinte processo:

- A proposta de aquisição e produção do empreendimento é apresentada à CAIXA pela empresa construtora proponente;

- A proponente construtora e o projeto do empreendimento são submetidos às análises técnica e de risco;
- É efetuada análise jurídica do vendedor do imóvel, da construtora proponente, bem como da regularidade e legalidade da documentação do empreendimento;
- A habilitação definitiva da proposta deve respeitar o limite do orçamento do Fundo de Arrendamento Residencial (FAR) para o Programa, por Unidade da Federação;
- A liberação dos recursos pelas obras executadas na construção ou recuperação é feita em parcelas mensais, creditadas na conta corrente da empresa construtora, condicionadas ao cumprimento do cronograma físico-financeiro da obra;
- Após a conclusão do empreendimento, as unidades são arrendadas às famílias que atendem aos requisitos de enquadramento no Programa;
- O Poder Público local identifica as famílias a serem beneficiadas;
- A CAIXA realiza a seleção dos arrendatários por meio da análise cadastral, da apuração da renda familiar bruta e da margem de renda disponível para comprometimento com as despesas de arrendamento.

O Programa é operado com recursos do Fundo de Arrendamento Residencial (FAR), criado exclusivamente para aplicação no PAR, composto com recursos onerosos provenientes de empréstimo junto ao FGTS e recursos não onerosos provenientes dos fundos FAS, FINSOCIAL, FDS e PROTECH e da rentabilidade das disponibilidades do FAR (Brasil, 2001).

O programa adota a modalidade de "arrendamento mercantil" ou leasing. Segundo Brito (2009), o imóvel é parte do patrimônio do Fundo de Arrendamento Residencial (FAR) e permanece como propriedade fiduciária da Caixa Econômica Federal, principal agente financiador da política pública, gestora do fundo e representante do arrendador até que as unidades habitacionais estejam quitadas. Nesse sentido, o mecanismo de leasing favorece quanto a possíveis processos judiciais em caso de inadimplência, ainda o arrendatário tem um período de 15 anos para decidir pela aquisição do imóvel. É importante frisar que o PAR é uma política que visa reduzir os déficits habitacionais, tirar as famílias em condições de submoradias e inserir em condições mínimas de bem-estar, por esse fator o programa é aplicado em municípios com mais de 100 mil habitantes, onde as questões de problemas de moradia ocorrem com mais intensidade.

Na próxima seção apresentaremos os dados utilizados e suas estatísticas descritivas para avaliar os impactos do programa de arrendamento residencial sobre o emprego e os salários.

## 3. Dados

Nossas variáveis dependentes — número de empregados e salários na construção civil — foram coletadas do banco de dados do Instituto Brasileiro de Geografia e Estatística (IBGE), mais precisamente do Cadastro Central de Empresas (CEMPRE). Segundo IBGE, o banco de dados do CEMPRE é formado por empresas e outras organizações e suas respectivas unidades locais formalmente constituídas, registradas no CNPJ (Cadastro Nacional de Pessoa Jurídica). Sua atualização ocorre anualmente, a partir das pesquisas econômicas anuais do IBGE, nas áreas de Indústria, Comércio, Construção e Serviços, e de registros administrativos, como a Relação Anual de Informações Sociais (RAIS). Ambas as variáveis dependentes são dos anos de 2000 a 2003, para todos os municípios brasileiros e, no caso dessa pesquisa, foi utilizado o log natural dos dados coletados.

Analisando os dados da Tabela 1, relativos ao número de empregados no setor da construção civil, existe um aumento de 66% no número de pessoas em trabalhos pertencentes ao setor da construção em municípios que participam do PAR, frente a um aumento de 46% nos empregos em municípios sem o PAR. A variação mostra que municípios contemplados pelo programa possuem um aumento 20% maior no número de servidores, o que demonstra um indicativo de que o programa está correlacionado com um aumento nos empregos do setor.

Em relação aos salários, a dinâmica é a mesma, conforme dados da Tabela 1. Os municípios com a presença do programa possuem um aumento de 15% a mais nos salários do que nos municípios que não recebem o PAR.

Para analisar o efeito do PAR, foram coletados dados do programa junto a Caixa Econômica Federal (CAIXA), órgão responsável pelo gerenciamento operacional do programa, e do Ministério das Cidades. Como forma de identificar os municípios com o PAR utiliza-se uma dummy, onde denominaremos os municípios que receberam o programa de tratados (e recebem o valor 1) e para os locais que não participam do PAR daremos o nome de controles (e recebem o valor 0).

Através do conjunto de dados levantados na Caixa Econômica Federal e apresentados na Tabela 2, observamos que em 2001 o programa contemplou 340 dos 5.570 municípios

**Tabela 1.** Estatisticas Descritivas.

| | Controle | | | Tratado | | | Diferença | |
|---|---|---|---|---|---|---|---|---|
| Variável | Nº Obs. | Média | D.P. | Nº Obs. | Média | D.P. | Diff | D.P. |
| Ln Empregos na Construção Civil | 15.468 | 0,454 | 0,168 | 1.628 | 0,662 | 0,099 | -0,208* | 0,004 |
| Ln Salários na Construção Civil | 12.587 | 0,278 | 0,120 | 1.616 | 0,431 | 0,075 | -0,152* | 0,003 |
| Índice de Theil | 37.759 | 0,504 | 0,129 | 1.667 | 0,515 | 0,133 | -0,011* | 0,003 |
| Anos de Escolaridade | 36.557 | 0,024 | 0,146 | 1.642 | 0,0006 | 0,001 | 0,024* | 0,003 |
| Ln Consultas Médicas | 36.950 | 11,065 | 1,332 | 1.641 | 14,141 | 1,337 | -3,075* | 0,033 |

**Tabela 2.** Quantidade de Municípios Contemplados pelo PAR.

| UF | Quantidade de Municípios com PAR | % em relação ao total de municípios da UF | UF | Quantidade de Municípios com PAR | % em relação ao total de municípios da UF |
|---|---|---|---|---|---|
| AC | 1 | 4,54% | PB | 5 | 2,17% |
| AL | 2 | 1,87% | PE | 15 | 7,69% |
| AM | 1 | 1,53% | PI | 2 | 0,86% |
| AP | 1 | 5,88% | PR | 31 | 7,56% |
| BA | 17 | 3,95% | RJ | 30 | 30,00% |
| CE | 6 | 3,09% | RN | 4 | 2,31% |
| DF | 1 | 100,00% | RO | 3 | 5,55% |
| ES | 9 | 10,84% | RR | 1 | 6,66% |
| GO | 12 | 4,83% | RS | 35 | 6,81% |
| MA | 8 | 3,59% | SC | 11 | 3,66% |
| MG | 37 | 4,25% | SE | 2 | 2,63% |
| MS | 2 | 2,47% | SP | 91 | 13,80% |
| MT | 4 | 2,70% | TO | 2 | 1,40% |
| PA | 6 | 4,16% | | | |

brasileiros existentes, sendo o estado do Rio de Janeiro o de maior atendimento (30%), seguido de São Paulo e Espírito Santo com 13,8% e 10,8% respectivamente. Os estados com menor inserção são Piauí (0,86%), Tocantins (1,4%) e Amazonas (1,53%).

Na Tabela 3 estão dispostos os valores em reais e o número de unidades contempladas pelo PAR, no período de 2001 a 2005. O valor financiado foi da ordem de 3,5 bilhões de reais, com 139.109 unidades contempladas pelo programa. O seu financiamento máximo foi atingido no ano de 2003, com 43.827 unidades, alcançando o valor de 1,1 bilhões de reais.

Os dados da Tabela 4, com estatísticas a nível de regiões, revelam que o Sudeste apresenta 45,4% das unidades totais financiadas do PAR, em segundo temos o Nordeste com 30,3% e em terceiro o Sul, com 13,5% do total de unidades contempladas pelo programa entre 2001 a 2005.

As covariáveis ou variáveis de controle utilizadas foram o índice de Theil, anos de estudo e número de consultas médicas. O Índice de Theil foi utilizado para controlar as possíveis desigualdades de renda existentes entre os municípios. A coleta se deu junto ao banco de dados (IPEADATA) do Instituto de Pesquisa Econômica Aplicada (IPEA), no período de 2000 a 2003. Para este indicador a média brasileira entre os municípios de controle ficou em 0,50 e para os municípios de tratamento o índice ficou em 0,51, o que demonstra pouca diferença entre os dois grupos.

A covariável anos de estudo foi coletada do banco de dados do Instituto Brasileiro de Geografia e Estatística (IBGE), através dos censos de 1999 e 2000. Esta variável foi utilizada para controlar possíveis efeitos advindos da educação e do avanço em capital humano, assim observamos na Tabela 1 que os municípios de controle possuem 0,024 anos a mais que

**Tabela 3.** Quantidade de Recursos Alocados para o PAR e Unidades Contratadas no Brasil.

| Ano | Valor do Investimento | Conjuntos | Unidades |
|---|---|---|---|
| 2001 | 426.349.064,66 | 165 | 20.315 |
| 2002 | 729.538.679,03 | 200 | 30.261 |
| 2003 | 1.113.903.023,06 | 293 | 43.827 |
| 2004 | 869.800.000,00 | 188 | 33.556 |
| 2005[a] | 307.400.000,00 | 61 | 11.150 |
| Total | 3.446.990.766,75 | 907 | 139.109 |

*Fonte:* Dados da Caixa Econômica Federal e Bonates (2008).
[a] Até julho de 2005.

**Tabela 4.** Quantidade de Unidades Financiadas pelo PAR no Brasil, por região.

| | Regiões | | | | | |
|---|---|---|---|---|---|---|
| Ano | Sudeste | Nordestes | Sul | Centro-Oeste | Norte | Total |
| 2001 | 11.280 | 5.151 | 2.398 | 864 | 622 | 20.315 |
| 2002 | 16.739 | 6.297 | 3.892 | 1.677 | 1.656 | 30.261 |
| 2003 | 20.562 | 12.049 | 7.093 | 3.547 | 576 | 43.827 |
| 2004 | 13.177 | 10.274 | 5.068 | 4.247 | 790 | 33.556 |
| 2005[a] | 4.284 | 3.786 | 2.146 | 436 | 576 | 11.228 |

*Fonte:* Dados da Caixa Econômica Federal e Bonates (2008).
[a] Até julho.

municípios tratados, o que evidencia uma certa possibilidade de que escolaridade não esteja correlacionada com um aumento dos empregos e dos salários na construção civil.

Por último, utilizou-se a variável consultas médicas para controlar condições de saúde das pessoas, pois o bem-estar clínico faz com que o contingente de oferta de trabalho aumente e provoque mudanças no setor em análise. Este dado foi retirado do Departamento de Informática do Sistema Único da Saúde (DATASUS), através das informações geradas pela Produção Hospitalar (SIH/SUS). Desta forma, conforme a análise da Tabela 1, infere-se que as pessoas moradoras de municípios que possuem o programa fazem mais visitas ao médico se comparado a população dos municípios que não possuem o PAR.

Salientamos que estas variáveis de controle são utilizadas para dar maior precisão aos valores estimados, conforme Imbens e Lemieux (2008), porém não devem interferir ou provocar qualquer viés nos resultados encontrados do modelo.

Conhecidas as variáveis do trabalho, na próxima seção demonstraremos a estratégia empírica adotada neste trabalho.

## 4. Estratégia Empírica

### 4.1 Teste de Manipulação do Cutoff pelos Municípios

Antes de aplicarmos a estratégia empírica deste estudo e chegarmos aos resultados, analisemos o critério imposto pelo programa, onde os municípios contemplados devem ter mais do que 100.000 habitantes. Esta imposição numérica gera margem para discussão sobre a questão da contagem da população dos municípios, o que também foi observado no trabalho de Monasterio (2014). A hipótese que elaboramos é de que gestores municipais interessados em receber o PAR, podem, de certa forma, tentar "manipular" as informações populacionais e serem contemplados com o programa. Para testar essa hipótese, utilizamos a estratégia apresentada em Cattaneo, Jansson, e Ma (2018) (daqui em diante CJM) denominado "Manipulation Test", baseado na densidade da descontinuidade. Assim, segundo CJM, para implementar um teste de manipulação, o pesquisador precisa estimar a densidade de unidades perto do ponto de corte e realizar um teste de hipótese sobre a densidade da descontinuidade.

Segundo demonstra CJM,[1] assume-se que $X_1, X_2, \ldots, X_n$ é uma amostra aleatória de tamanho $n$ da variável aleatória $X$ com função de distribuição cumulativa (f.d.c.) e função de densidade de probabilidade (f.d.p.) dada por $F(x)$ e $f(x)$, respectivamente.

A variável aleatória $X_i$ significa a pontuação, índice ou variável de análise da unidade $i$ na amostra. Cada unidade é atribuída ao controle ou tratamento, dependendo se o índice observado exceder um corte conhecido denotado por $\overline{x}$. Nesse caso a atribuição de grupo ou "tratamento" é dada por

$$\begin{aligned} &\text{Unidade } i \text{ atribuída ao grupo de controle se } X_i < \overline{x}, \\ &\text{Unidade } i \text{ atribuída ao grupo de tratamento se } X_i \geq \overline{x}, \end{aligned}$$

onde o ponto de corte $\overline{x}$ é conhecido e, é claro, elencamos observações suficientes para cada grupo que estão disponíveis. Um teste de manipulação neste contexto é um teste de hipóteses sobre a continuidade da densidade $f(\cdot)$ no ponto de corte $x$. Formalmente, estamos interessados no seguinte problema: Um teste de manipulação neste contexto é um

[1] Teste formal extraído de Cattaneo et al. (2018).

teste de hipóteses sobre a continuidade da densidade $f(\cdot)$ no ponto de corte $\overline{x}$. Formalmente, estamos interessados no seguinte problema:

$$H_0 : \lim_{x\uparrow\overline{x}} f(x) = \lim_{x\downarrow\overline{x}} f(x) \quad vs \quad H_1 : \lim_{x\uparrow\overline{x}} f(x) \neq \lim_{x\downarrow\overline{x}} f(x)$$

Para construir uma estatística para este teste de hipóteses, seguimos CJM e estimamos a densidade $f(x)$ usando um estimador de densidade polinomial local com base no f.d.c. da amostra observada. Este estimador tem várias propriedades interessantes, incluindo o fato de que não requer pré-binning dos dados, além de também permitir incorporar restrições no f.d.c., derivadas de ordem superior da densidade, induzindo a novos testes de manipulação com propriedades mais poderosas nas aplicações.

A classe de estatísticas do Teste de Manipulação implementadas assume a seguinte forma:

$$T_p(h) = \frac{\hat{f}_{+,p}(h) - \hat{f}_{-,p}(h)}{\hat{V}_p(h)}, \qquad \hat{V}_p^2 = \hat{K}\left[\hat{f}_{+,p}(h) - \hat{f}_{-,p}(h)\right],$$

onde $T_p(h) \sim \mathcal{N}(0,1)$ significa sob suposições apropriadas e a notação $\hat{V}[\cdot]$ é designado por algum estimador consistente da quantidade de população $V[\cdot]$. O parâmetro $h$ é a largura(s) de banda usada(s) para localizar os procedimentos de estimativa e inferência perto do ponto de corte $\overline{x}$.

As estatísticas podem ser construídas de várias maneiras diferentes, em particular, dada uma escolha de largura de banda. Dois ingredientes principais são usados para construir a estatística de teste $Tp(h)$ considerando: i) os estimadores de densidade polinomial local $f(h)$; e ii) o erro padrão correspondente ao estimador $+$, $p-$, $pV_p(h)$.

Esses estimadores também dependem da escolha da ordem polinomial $p$, da escolha da função kernel $K(\cdot)$ e das restrições impostas no modelo, entre outras possibilidades. As fórmulas de erro padrão $V_p(h)$ podem ser baseadas em um plug-in assintótico ou uma abordagem jackknife, e sua forma específica dependerá de restrições adicionais ao modelo.

Um elemento crucial é, naturalmente, a escolha da largura de banda h, que determina quais as observações próximas ao ponto de corte $\overline{x}$ que são usadas para estimação e inferência. Essa escolha pode ser especificada pelo usuário ou estimada usando os dados disponíveis. A estimação permite, quando possível, opções de largura de banda diferentes de cada lado do ponto de corte $\overline{x}$. Uma largura de banda comum em ambos os lados do ponto de corte é sempre possível.

## 4.2 Desenho de Regressão Descontínua

O trabalho objetiva analisar o efeito do financiamento social, tomando como objeto de pesquisa o Programa de Arrendamento Residencial (PAR), sobre o setor da construção civil. Assim, como estratégia empírica, utilizamos o desenho de Regressão Descontínua (RD), conforme o trabalho desenvolvido por Thistlethwaite e Campbell (1960). Adotamos essa estratégia empírica por possuirmos um ponto de salto de probabilidade, que chamamos de ponto de corte, e que faz parte do critério de adesão do município ao programa. Esse ponto de corte é exatamente em municípios com mais de 100.000 habitantes.

Segundo Rocha e Belluzzo (2010) o pressuposto de descontinuidade formaliza a ideia de que indivíduos um pouco acima e abaixo do corte precisam ser "comparáveis", exigindo

que eles tenham uma média similar dos resultados possíveis, ao receber ou não tratamento. Assim, estimamos a seguinte equação:

$$Y_{ip} = \beta_0 + \beta_1 PAR_{ip} + \beta_2 T_{ip} + \varepsilon_{ip}, \tag{1}$$

onde $Y_i$ é variável de interesse do modelo, no município $i$ para o ano $p$; $PAR$ é o programa de arrendamento residencial que leva valor igual a 1 caso o município recebe o programa e 0 caso contrário, no município $i$ para o ano $p$. $T_{ip}$ é o valor que indica se o município está acima ou abaixo do valor de corte citado anteriormente, no município $i$ para o ano $p$ e por fim $\varepsilon_{ip}$ é um termo de erro.

Porém, observando os dados, existem municípios que com o passar do tempo aumentam o números de seus habitantes devido a dinâmica socioeconômica natural, promovendo variações no número de seus moradores, ou seja, em alguns anos estão acima do corte e em outros anos estão abaixo do corte, o que gera correlação entre o termo de erro e a variável de interesse. Desta forma, foi escolhido o modelo de regressão descontínua fuzzy (FRD), onde, segundo Trochim (1984), tem a sensibilidade de considerar um aumento de probabilidade, mas não de zero para um, pois a atribuição ao tratamento pode depender de fatores adicionais. Assim, para estimar os efeitos do PAR em um modelo FRD, usamos a abordagem de variáveis instrumentais (IV) proposta por Angrist e Pischke (2008) através do modelo de mínimos quadrados em dois estágios (2SLS), desta forma temos:

$$Y_{ip} = \beta_0 + \beta_1 PAR_{ip} + f\left(Pop_{ip}, Cut_{ip}\right) + X_{ip}\Theta + \eta_{1ip}, \tag{2}$$

$$PAR_{ip} = \delta_0 + \delta_1 Cut_{ip} + f\left(Pop_{ip}, Cut_{ip}\right) + X_{ip}\Omega + \eta_{2ip}, \tag{3}$$

onde $Cut_{ip}$ é uma variável dummy que possui valor igual 1 se o município tem população acima do ponto de corte no município $i$ para o ano $p$; $f\left(Pop_{ip}, Cut_{ip}\right)$ é um polinômio de segunda ordem que interage com $Cut_{ip}$; $X_{ip}$ é um de vetor covariáveis[2] com características municipais, dos serviços na saúde e das condições socioeconômicas.

Continuando nossa estratégia, a estimação está em sua forma não paramêtrica. Para isso determinamos que a nossa função Kernel será a triangular, seguindo a mesma estratégia adotada por Smith (2016); Toro, Tigre, e Sampaio (2015); e Fujiwara (2015). Dois bandwidths foram utilizados, um de 60.000 e outro de 100.000, para testar se alterações significativas no tamanho da janela interferem na robustes dos resultados. Ainda, o modelo foi calculado em sua forma linear e quadrática.[3]

## 4.3 Viés de Seleção Amostral: Modelo de Correção de Heckman

Ao analisar o efeito que o programa provoca nos empregos e nos salários no setor da construção civil, deve-se estar atento à outras possíveis causas que possam surgir, uma vez que a dinâmica do PAR poderia também gerar outro efeito no rendimento financeiro e nas quantidades de serviços.

Esses fatores eliminam a característica de aleatoriedade de nossa amostra, o que provocaria viés de seleção em nosso modelo. Os fatores podem estar ligados a pessoas com mais habilidade, redes de contatos para a vaga de emprego, uma entidade de classe mais

[2]Estamos incluindo covariáveis no modelo pois segundo Imbens e Lemieux (2008) inserir covariáveis aumenta a sua precisão, como citado anteriormente.

[3]A especificação do polinômio do modelo segue a orientação dada no artigo de Gelman e Imbens (2017).

atuante, entre outros, que facilitam a entrada de novos empregados no mercado de trabalho, bem como alterações nos salários da categoria. Nesse sentido, utilizou-se neste trabalho a estratégia denominada Heckit, elaborado por Heckman (1974) para corrigir o problema de seletividade amostral. É importante frisar, que este modelo foi usado apenas quando a hipótese de seletividade amostral fora identificada, visto que causaria viés de seleção em nossos resultados, caso contrário, mantém-se a estratégia inicial do modelo de regressão descontinua sem a utilização do Heckit.

Assim, abaixo demonstra-se a forma precisa do método correção de seleção de Heckman em dois passos (Heckit) para identificar possíveis causa de viés de seleção, onde[4]

$$Y_2 = \alpha Z + \delta, \tag{4}$$

$$Y_1 = \beta_0 + \beta_1 X + \sigma \rho_{\varepsilon\delta} \lambda(T - \alpha Z) + \sigma' \varepsilon'. \tag{5}$$

Na equação de seleção, que é estimada com um probit, $Y_2$ é a variável dependente dicotômica, $Z$ é a variável independente, $\alpha$ é o coeficiente de $Z$, e $\delta$ é o termo de erro normalmente distribuído. Na equação de regressão, o valor de $Y_1$ é observado quando $Y_2$ é maior do que um limite $T$ e é censurado (ou seja, falta) se $Y_2 \geq T$. A estimativa de (5) simplesmente regredindo $Y$ em $X$ será tendenciosa por causa do termo sigma, que representa a variável omitida.

Este problema pode ser resolvido em duas etapas. Primeiro, a equação de seleção (6) é estimada usando probit e os valores previstos são retidos como estimativas de $(T - \alpha Z)$. A proporção de Mills inversa é então estimada para cada caso dividindo a função de densidade normal, avaliada em $-\phi(T - \alpha Z)$, por 1 menos a função de distribuição cumulativa normal, estimada em $-\phi(T - \alpha Z)$ (5). O segundo passo é uma regressão de mínimos quadrados ordinária com $X$ e a proporção de Mills inversa incluída como regressores. O estimador é consistente quando os pressupostos são atendidos:

$$\lambda(T - \alpha Z) = \frac{\phi(T - \alpha Z)}{1 - \Phi(T - \alpha Z)}. \tag{6}$$

## 4.4 Testes de Robustez

Para testar a especificação de nosso modelo, foram aplicados quatro testes de robustez. O primeiro foi um teste placebo em que testamos para tendências anteriores. Assim, os resultados dos parâmetros estimados, do ano imediatamente anterior a entrada do PAR, não podem apresentar significância estatística.

O segundo teste verificou a largura da janela, sendo de 60.000 habitantes e de 100.000 habitantes. Como resultado para este teste em ambas janelas deve haver significância estatística.

Para o terceiro teste foi feito uma alteração no *cutoff*, onde arbitrariamente alteramos o ponto de corte para 95.000 e 105.000 habitantes, trabalhando somente com quem é considerado não tratado e da outra forma, tratado. Como resultado, os parâmetros estimados também não podem ser estatisticamente significantes, o que provocaria um enfraquecimento da estratégia argumentativa de nosso trabalho.

No quarto e último teste, foram utilizadas as covariáveis relacionada com o modelo especificado (Calonico, Cattaneo, Farrell, & Titiunik, 2016). Neste caso, o Índice de Theil,

[4]Demonstração apresentada no trabalho de Bushway, Johnson, e Slocum (2007).

Escolaridade e Consultas Médicas, através dos seus estimadores não podem apresentar significância estatística, pois acredita-se que os municípios bem perto do corte possuem as mesmas características, o que anulariam seus efeitos.

Na próxima seção apresentaremos a análise do efeito do PAR sobre os empregos e salários na construção civil.

## 5. Análise dos Resultados

Antes de iniciarmos a análise das regressões, apontamos que os teste realizados para identificar viés de seleção amostral (Heckit) sinalizaram que a variável dependente salários determinou que podemos rejeitar a hipótese de viés de seleção amostral, conforme resultado apresentado na Tabela 5. Com isso, utilizamos a estratégia de Regressão Descontínua para estimar nossos parâmetros, porém, na variável dependente empregos não podemos rejeitar a hipótese de viés de seleção, o que nos obriga a realizar nossas estimações utilizando o modelo de correção de Heckmam, junto a nossa estratégia de RD.

Outro teste que também foi aplicado, como pode ser observado o seu resultado na Tabela 5, foi o "Manipulation Test" de Cattaneo et al. (2018). Este teste apontou que a hipótese de manipulação do ponto de corte do programa pelos municípios pode ser rejeitadae com isso podemos dizer que empiricamente não há evidencias de manipulação de nosso *cutoff*.

Dando início às análises de nossas estimações, com o intuito de verificar o efeito do PAR sobre o emprego e o salário na construção civil, observemos a Figura 1. Esta figura avalia a relação entre a densidade populacional dos municípios e sua relação com o PAR.

Observamos que existe uma descontinuidade em relação a municípios com mais de 100.000 habitantes comparados a municípios com menos de 100.000 habitantes e a diferença de densidade populacional entre estes dois grupos separados pela descontinuidade chega a

**Tabela 5.** Financiamento Habitacional sobre Mercado de Trabalho na Construção Civil – Sem Variáveis de Controle.

| | Pessoas Empregadas | | | | | Salários dos Empregados | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variável | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| **Painel A:** ***Ano de Início*** | | | | | | | | | | |
| $PAR_t$ | 0,148*** | 0,139*** | 0,145*** | 0,134*** | −0,024 | 0,001 | 0,010 | 0,005 | 0,006 | −0,032 |
| | (0,018) | (0,015) | (0,017) | (0,015) | (0,056) | (0,046) | (0,033) | (0,050) | (0,041) | (0,038) |
| **Painel B:** ***Anos Posteriores*** | | | | | | | | | | |
| $PAR_{t+1}$ | 0,168*** | 0,150*** | 0,166*** | 0,146*** | – | 0,064* | 0,061** | 0,071* | 0,064** | – |
| | (0,018) | (0,015) | (0,018) | (0,015) | – | (0,035) | (0,025) | (0,038) | (0,031) | – |
| $PAR_{t+2}$ | 0,167*** | 0,149*** | 0,165*** | 0,145*** | – | 0,054* | 0,052* | 0,082** | 0,070** | – |
| | (0,018) | (0,015) | (0,018) | (0,015) | – | (0,029) | (0,028) | (0,038) | (0,029) | – |
| Especificação | Linear | Linear | Quad. | Quad. | Quad. | Linear | Linear | Quad. | Quad. | Quad. |
| Bandwidth | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 |
| Heckit | Sim | Sim | Sim | Sim | Sim | Não | Não | Não | Não | Não |
| Manipulation Test for $PAR$ ($P > \|T\|$) | 0,2254 | | | | | | | | | |
| Nº de Observações | 5.107 | 5.320 | 5.107 | 5.320 | 5.085 | 476 | 1.983 | 476 | 1.984 | 450 |

*Notas:* Vars. Dependente log Quantidade de Empregados e Log Salários. Todas as especificações usam Kernel Triangular. As colunas 1 a 4 são relativas as pessoas empregadas. As colunas 6 a 9 são relativas ao salário das pessoas empregadas. O PAR (Programa de Arrendamento Residencial) estima a descontinuidade de municípios logo acima de 100.000 habitantes. As colunas 5 e 10 refere-se a um teste placebo de um ano antes do início do PAR (2000). Modelo de Correção de Heckman (Heckit) aplicado ao modelo de regressão quando há significância estatística. Erro-Padrão robusto em parênteses.
*** $p < 0{,}01$, ** $p < 0{,}05$ e * $p < 0{,}1$.

0,437. Isso demonstra que o critério de seleção para adoção do PAR nos municípios existe e é exatamente como descrito nas condições pré-estabelecidas no programa.

Na Figura 2 nota-se uma descontinuidade tanto no gráfico a esquerda, relativo as vagas de empregos na construção civil, quanto no gráfico a direita, relativa aos salários na construção civil. Isso demonstra que o programa provoca uma variação próximo ao ponto de corte e essa variação é relativa a um aumento tanto dos empregos quanto dos salários no setor.

Outra forma de comprovarmos e medir a magnitude do aumento dos empregos e dos salários através de ações do PAR é observando os resultados da Tabela 5. Estimou-se a relação entre o PAR e os empregos e os salários da construção civil, e a princípio, não utilizou-se variáveis de controle no modelo, pois independentemente da existência dessas covariáveis, deve haver significância estatística em nossos parâmetros, o que de fato confirmou os resultados.

Conforme Tabela 5, nas colunas de 1 a 4, temos os resultados relativos a pessoas empregadas na construção civil e nas colunas 6 a 9 temos os resultados do efeito sobre os salários do setor em questão. O Painel A refere-se ao ano de entrada do PAR, ou seja, 2001. O Painel B demonstra anos posteriores, sendo 2002 ($t+1$) e 2003 ($t+2$). Os resultados foram testados com *Bandwidth* distintos para garantir que nossos estimadores sejam robustos.

Nas colunas 1 e 2, do Painel A, Tabela 5, observamos um aumento de 14,8% e 13,9% de pessoas empregadas em municípios cobertos pelo PAR. Esses valores são relativos a modelos lineares e com um *Bandwidth* de 60.000 e 100.000 habitantes. A tendências desses resultados permanecem, mesmo quando são especificados em sua forma quadrática e também com *Bandwidth* distintos, ocorrendo um aumento de 14,5% a 13,4% em relação as pessoas empregadas na construção civil, o que é demonstrado nas colunas 3 e 4.

Já os resultados encontrados nas colunas 6 a 9, do Painel A, Tabela 5, referentes aos salários dos empregados na construção civil, apontaram uma relação positiva, porém não apresentaram significância estatística. Estes resultados são compreensíveis, pois poderia haver excesso de oferta de mão de obra, o que não provocaria aumento dos salários de forma instantânea. Ressaltamos que não temos a pretensão de buscar os mecanismos, mas apenas os possíveis efeitos do PAR sobre os empregos e os salários.

Quando analisamos os resultados para anos subsequentes, conforme dados da Tabela 5, o efeito positivo, ou seja, o aumento ocasionado pelo PAR, se mantém evidente, tanto para os empregos, quanto para os salários na construção civil. Verifica-se que esses aumentos cheguem a 16,8% para os empregos um ano após a criação do programa, e a 16,7% de aumento após dois anos.

Em relação aos salários, no primeiro ano após o início do programa ocorre um aumento salarial de até 6,4% e quando se trata do segundo ano após o início temos um aumento de até 5,4%. Todos os resultados possuem significância estatística.

Contudo, incluímos covariáveis em nosso modelo, de acordo com a estratégia de Imbens e Lemieux (2008), para verificar se os resultados se mantém e se nossos estimadores são realmente robustos e precisos. Sendo assim, na coluna 1 e 2 (forma linear) da Tabela 6 temos como resultados de nossas estimações um percentual positivo de 13% e 12% em relação aos empregos, ambos significantes estatisticamente.

Em sua forma quadrática (colunas 3 e 4) os resultados se mantém positivos e significantes, com um percentual de 13,2% e 12,1% respectivamente. Os anos subsequentes também apontaram uma variação positiva e relevante de 12,9% a 15,2% nos dois primeiros anos, nos modelos lineares. Nos modelos quadráticos apresentaram variações também positivas com significância de 12,9% a 15,3% nos dois primeiros anos. Quanto aos resultados sobre

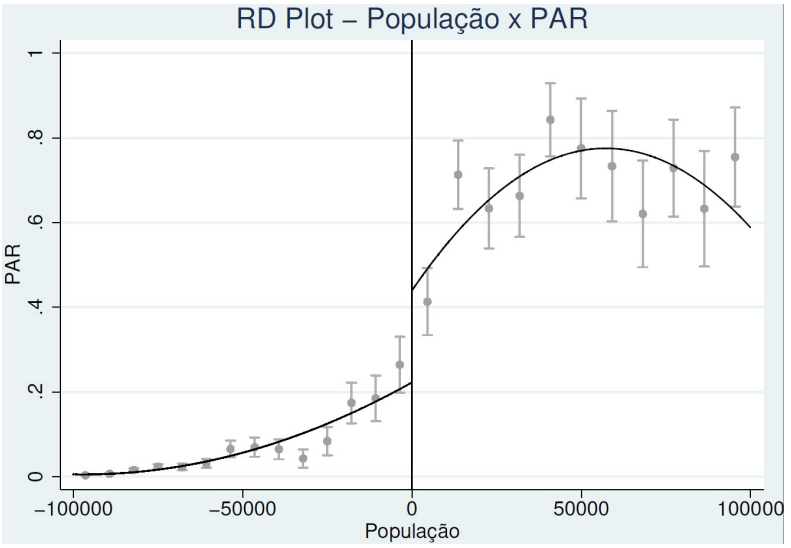



**Figura 1.** População sobre o Par.

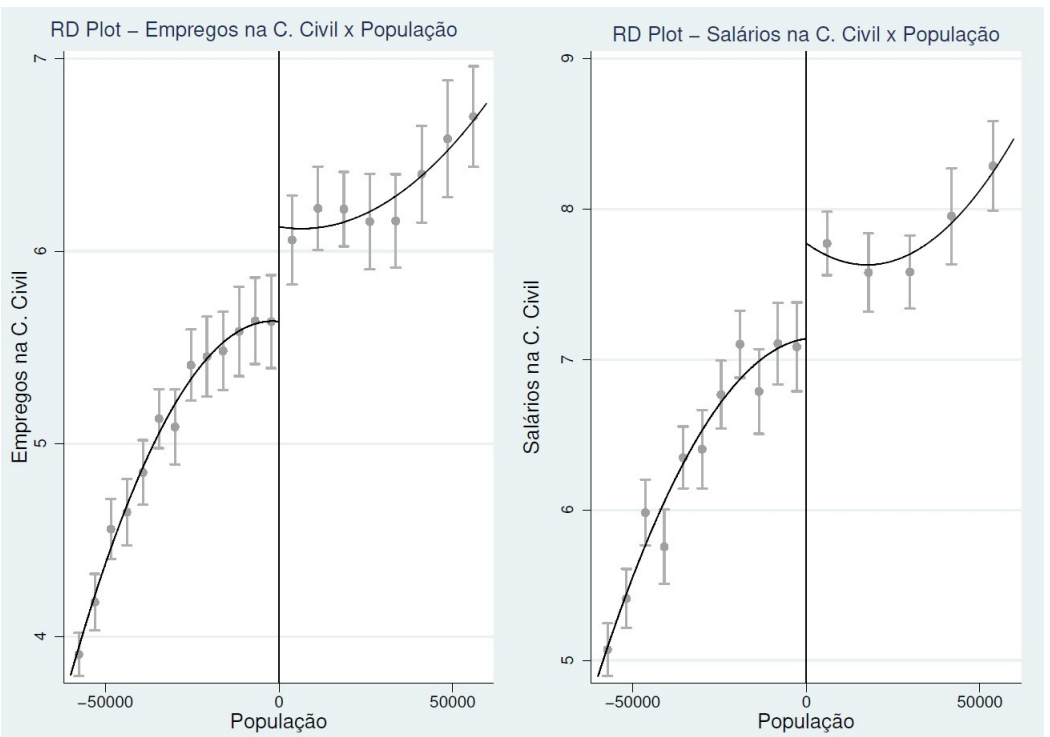



**Figura 2.** Financiamento Habitacional sobre Empregos e Salários.

**Tabela 6.** Financiamento Habitacional sobre Mercado de Trabalho na Construção Civil – Com Variáveis de Controle.

| | Pessoas Empregadas | | | | | Salários dos Empregados | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Variável | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) |
| **Painel A:** ***Ano de Início*** | | | | | | | | | | |
| $PAR_t$ | 0,130*** | 0,120*** | 0,132*** | 0,121*** | −0,227 | 0,013 | 0,012 | 0,009 | 0,010 | −0,022 |
| | (0,018) | (0,015) | (0,018) | (0,015) | (0,193) | (0,037) | (0,033) | (0,049) | (0,041) | (0,038) |
| **Painel B:** ***Anos Posteriores*** | | | | | | | | | | |
| $PAR_{t+1}$ | 0,146*** | 0,129*** | 0,147*** | 0,129*** | – | 0,060** | 0,052** | 0,069* | 0,063** | – |
| | (0,018) | (0,015) | (0,018) | (0,015) | – | (0,030) | (0,025) | (0,038) | (0,031) | – |
| $PAR_{t+2}$ | 0,152*** | 0,132*** | 0,153*** | 0,132*** | – | 0,051* | 0,050* | 0,076** | 0,066** | – |
| | (0,018) | (0,018) | (0,015) | (0,015) | – | (0,029) | (0,028) | (0,038) | (0,029) | – |
| Especificação | Linear | Linear | Quad. | Quad. | Quad. | Linear | Linear | Quad. | Quad. | Quad. |
| Bandwidth | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 |
| Heckit | Sim | Sim | Sim | Sim | Sim | Não | Não | Não | Não | Não |
| Manipulation Test for $PAR$ $(P > \|T\|)$ | 0,2254 | | | | | | | | | |
| Nº de Observações | 5.084 | 5.295 | 5.084 | 5.295 | 5.085 | 476 | 1.983 | 476 | 1.984 | 450 |

*Notas:* Vars. Dependente log Quantidade de Empregados e Log Salários. Todas as especificações usam Kernel Triangular. As colunas 1 a 4 são relativas as pessoas empregadas. As colunas 6 a 9 são relativas ao salário das pessoas empregadas. O PAR (Programa de Arrendamento Residencial) estima a descontinuidade de municípios logo acima de 100.000 habitantes. As colunas 5 e 10 refere-se a um teste placebo de um ano antes do início do PAR (2000). Modelo de Correção de Heckman (Heckit) aplicado ao modelo de regressão quando há significância estatística. Erro-Padrão robusto em parênteses.
*** $p < 0{,}01$, ** $p < 0{,}05$ e * $p < 0{,}1$.

salários, os estimadores apontaram variação positiva e significante somente após o primeiro ano de implantação do PAR, atingindo um aumento entre 5% a 6% em modelos lineares e entre 6,3 a 7,6% em modelos quadráticos.

## 5.1 Testes de Robustez

Para validar os parâmetros estimados de nosso modelo, aplicamos quatro testes de robustez. O primeiro teste é referente a resultados obtidos a um ano anterior ao início do PAR, porém espera-se que não apresente significância estatística. O segundo teste, foi uma alteração no *Bandwidth* em relação ao *cutoff*, assim, se houver aumento da janela o modelo deve permanecer estatisticamente significante. No terceiro teste alteramos o ponto de corte (*cutoff*) estabelecido pelo programa para um outro ponto em 95.000 habitantes e 105.000 habitantes. Como resultado desta troca espera-se que os resultados não tenham significância estatística, demonstrando que o ponto de corte do programa seja único. O quarto e último teste de robustez foi verificar se as covariáveis do modelo não possuem significância estatística.

Com relação ao primeiro teste, observamos nas Tabelas 5 e 6, nas colunas 5 e 10, que os resultados obtidos não apresentaram significância estatística. Com isso verificamos que nenhum outro programa ou atividade antes do PAR possa ter causado alguma influência que alterasse a quantidade de pessoas empregadas e os salários no setor da construção civil.

O segundo teste foi empregado em todas estimações das Tabelas 5 e 6. Verificou-se que, mesmo com a alteração do *Bandwidth*, o modelo permaneceu empiricamente com seus parâmetros significantes estatisticamente.

O terceiro teste, representado através dos resultados da Tabela 7, permite observar que quando alteramos o ponto de corte, tanto para 95.000 habitantes quanto para 105.000 habitantes, os resultados não apresentam significância estatística, demonstrando que o ponto de corte do programa pode ser confirmado como um fator preponderante quanto ao salto de descontinuidade necessário à criação de um contrafactual de qualidade do nosso modelo. Os

**Tabela 7.** Teste de Robustez do Financiamento Habitacional sobre Mercado de Trabalho na Construção Civil – Alteração do Cutoff.

| | Pessoas Empregadas | | | | Salários dos Empregados | | | |
|---|---|---|---|---|---|---|---|---|
| Variável | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| **Painel A:** ***Ano de Início*** | | | | | | | | |
| $PAR_t$ | −4,035 | −0,508 | 0,502 | −8,708 | −0,260 | 0,045 | −18,533 | 0,032 |
| | (6,130) | (0,983) | (0,922) | (7,209) | (0,350) | (0,158) | (180,05) | (0,824) |
| **Painel B:** ***Anos Posteriores*** | | | | | | | | |
| $PAR_{t+1}$ | 0,053 | −0,694 | 2,829 | −0,387 | −0,522 | −0,184 | 1,996 | −0,403 |
| | (0,033) | (0,517) | (3,856) | (0,273) | (0,286) | (0,263) | (1,408) | (0,427) |
| $PAR_{t+2}$ | −0,011 | −0,026 | −0,015 | 0,136 | −4,609 | 0,066 | 0,177 | 0,143 |
| | (0,040) | (0,251) | (0,342) | (0,345) | (15,953) | (0,149) | (0,308) | (0,905) |
| Especificação | Linear | Linear | Quad. | Quad. | Linear | Linear | Quad. | Quad. |
| Cutoff | 95.000 | 105.000 | 95.000 | 105.000 | 95.000 | 105.000 | 95.000 | 105.000 |
| Bandwidth | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 100.000 |
| Controles | Sim | Sim | Sim | Sim | Sim | Sim | Sim | Sim |
| Nº de Observações | 503 | 2.418 | 503 | 2.418 | 474 | 1.968 | 474 | 1.968 |

*Notas:* Vars. Dependente log Quantidade de Empregados e Salários. Todas as especificações usam Kernel Triangular. As colunas 1 a 4 são relativas as pessoas empregadas. As colunas 5 a 8 são relativas ao salário das pessoas empregadas. O PAR (Programa de Arrendamento Residencial) estima a descontinuidade de municípios logo acima de 100.000 habitantes. Erro-Padrão robusto em parênteses. *** $p < 0{,}01$, ** $p < 0{,}05$ e * $p < 0{,}1$.

resultados do quarto teste está apresentado na Tabela 8. Contata-se que quando alteramos a variável dependente para Índice de Theil não observamos qualquer sinal de significâncias nos resultados, o que evidencia que o aumento dos financiamentos está diretamente ligado aos empregos e salários na construção civil e não tem efeito com a covariável utilizada no modelo.

Outra variável dependente que substituímos foi nível de estudo. Novamente não observamos para essa variável qualquer nível de significância estatística que faça aceitar seus resultados, comprovando, assim, que o PAR não exerce influência qualquer sobre o aumento ou redução do nível de escolaridade. Por último, a variável consultas médicas também foi testada, e não sendo significante estatisticamente. Desse modo, com base nos resultados, pode-se afirmar que o PAR realmente produz um efeito que garante avanços positivos nos empregos e salários na construção civil.

Portando, observa-se que o PAR consegue ser o fator causal de alteração ou de crescimento dos empregos e dos salários no setor da construção civil nos municípios que são tratados pelo programa, proporcionando um aumento considerável e demonstrando que as políticas implementadas vão além de seu escopo principal, tomado dimensões, como neste caso, favoráveis ao desenvolvimento do país.

## 6. Considerações Finais

O presente trabalho analisou o efeito do financiamento habitacional, tomando como objeto de estudo o Programa de Arrendamento Residencial (PAR), sobre o mercado de trabalho da construção civil. O objetivo principal foi verificar se a existência do programa de financiamento habitacional no município causou alterações na quantidade de pessoas empregadas e no salário das pessoas inseridas no mercado de trabalho da construção civil.

**Tabela 8.** Teste de Robustez do Financiamento Habitacional sobre Mercado de Trabalho na Construção Civil – Teste sobre Covariáveis.

| | Índice de Theil | | | | Nível de Estudos | | | | Consultas Médicas | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variável | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| **Painel A:** ***Ano de Início*** | | | | | | | | | | | | |
| $PAR_t$ | 1,507 | −0,197 | 2,592 | 5,799 | −0,003 | −0,002 | − 0,004 | −0,003 | 0,189 | 0,360 | 0,193 | 0,267 |
| | (1,409) | (0,354) | (2,239) | (5,814) | (0,002) | (0,001) | (0,003) | (0,002) | (0,592) | (0,403) | (0,719) | (0,527) |
| **Painel B:** ***Anos Posteriores*** | | | | | | | | | | | | |
| $PAR_{t+1}$ | −2,612 | −7,881 | 2,632 | −0,909 | 0,003 | 0,002 | 0,003 | 0,003 | −0,515 | −0,331 | −0,610 | −0,569 |
| | (2,712) | (10,881) | (2,889) | (1,151) | (0,002) | (0,002) | (0,003) | (0,002) | (0,713) | (0,609) | (0,712) | (0,699) |
| $PAR_{t+2}$ | 0,176 | −3,317 | 10,759 | 4,082 | 0,001 | 0,001 | 0,001 | 0,001 | 0,064 | 0,017 | 0,107 | −0,008 |
| | (0,193) | (8,143) | (10,337) | (4,027) | (0,001) | (0,001) | (0,001) | (0,0008) | (0,277) | (0,300) | (0,271) | (0,305) |
| Especificação | Linear | Linear | Quad. | Quad | Linear | Linear | Quad. | Quad. | Linear | Linear | Quad. | Quad. |
| Bandwidth | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 100.000 | 60.000 | 100.000 |
| Nº de Obs. | 541 | 5.443 | 541 | 5.443 | 538 | 5.348 | 538 | 5.348 | 541 | 5.417 | 541 | 5.417 |

*Notas:* Todas as especificações usam Kernel Triangular. Erro-Padrão robusto em parênteses.
$^{***}p < 0{,}01$, $^{**}p < 0{,}05$ e $^{*}p < 0{,}1$.

Para tal análise, adotamos uma estratégia empírica utilizando um desenho de regressão descontínua. Essa abordagem foi possível, pois o programa possui um critério de que apenas municípios que possuem 100.000 habitantes podem receber os benefícios do financiamento e repassá-los a seus moradores. Esse ponto de corte, instituído pelo programa, permite estabelecer um contrafactual para a análise sobre o efeito proporcionado pelo PAR. Porém, alguns cuidados foram tomados, como por exemplo, problemas de viés de seleção amostral e a manipulação de informações a respeito desse critério de implementação do programa, pois poderiam proporcionar graves erros à nossa estimação. Para o primeiro caso, utilizamos o teste de correção de seleção de Heckman (Heckit), para o segundo caso utilizamos o "*Manipulation Test*" de Cattaneo et al. (2018).

Além, do cuidado metodológico quanto ao instrumento de regressão utilizado, também nos preocupamos com possíveis resultados inconsistentes de nosso modelo devido a problemas de ordem endógena. Para isso, aplicamos testes de robustez como forma de garantir os parâmetros estimados e assegurar uma análise causal ao choque ocorrido nos tratados.

Como resultados de nossas regressões pode-se inferir que o PAR consegue aumentar de 13,4% a 16,8% o número de pessoas empregadas na construção civil, isso já no ano de implementação do programa no município, conseguindo obter um efeito positivo e duradouro nos anos subsequentes. Em relação aos salários das pessoas que trabalham na construção civil, o efeito só aparece após um ano de existência do PAR no município, promovendo um aumento entre 5,2% a 8,2%, se mantendo crescente nos anos subsequentes. Ressaltamos que os resultados apresentados foram todos estatisticamente significativos.

Como complemento da pesquisa, os resultados de todos os testes de robustez aplicados foram condizentes com nossas expectativas, o que garantiu que os estimadores calculados refletissem que os efeitos dos financiamentos habitacionais ocorridos nos municípios tratados realmente são consistentes.

Por fim, podemos concluir que o programa de financiamento habitacional, PAR, consegue promover no mercado de trabalho da construção civil uma melhora considerável nos postos de trabalhos. E nos anos subsequentes também proporciona o desenvolvimento econômico deste setor com o aumento de renda dos trabalhadores.

## Referências bibliográficas


Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Bonates, M. F. (2008). O Programa de Arrendamento Residencial – PAR: Acesso diferenciado à moradia e à cidade. *Risco – Revista de Pesquisa em Arquitetura e Urbanismo*, *7*, 147–164. http://dx.doi.org/10.11606/issn.1984-4506.v0i7p147-164

Bouillon, C. P. (Org.). (2012). *Un espacio para el desarrollo: Los mercados de vivienda en América Latina y el Caribe*. Banco Interamericano de Desarrollo. https://publications.iadb.org/handle/11319/3472

Brasil – Presidência da República. (2001). *Lei nº 10.188, de 12 de fevereiro de 2001: Cria o Programa de Arrendamento Residencial, institui o arrendamento residencial com opção de compra e dá outras providências.* http://www.planalto.gov.br/ccivil_03/LEIS/LEIS_2001/L10188.htm

Brito, J. N. d. S. (2009). *Retroalimentação do processo de desenvolvimento de empreendimentos de habitação de interesse social a partir de reclamações de usuários: Estudo no Programa de Arrendamento Residencial* (Dissertação de Mestrado, Universidade Federal do Rio Grande do Sul, Porto Alegre). http://hdl.handle.net/10183/21267

Bushway, S., Johnson, B. D., & Slocum, L. A. (2007). Is the magic still there? The use of the Heckman two-step correction for selection bias in criminology. *Journal of Quantitative Criminology*, *23*(2), 151–178. http://dx.doi.org/10.1007/s10940-007-9024-4

CAIXA – Caixa Economica Federal. (2003). *Programa de arrendamento residencial.* www.caixa.gov.br

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2016). *Regression discontinuity designs using covariates*.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *Stata Journal*, *18*(1), 234–261. https://www.stata-journal.com/article.html?article=st0522

Fujiwara, T. (2015). Voting technology, political responsiveness, and infant health: Evidence from Brazil. *Econometrica*, *83*(2), 423–464. http://dx.doi.org/10.3982/ECTA11520

Gelman, A., & Imbens, G. (2017). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*. (Published online) http://dx.doi.org/10.1080/07350015.2017.1366909

Heckman, J. J. (1974). Sample selection bias as a specification error. *Econometrica*, *42*(1), 679–94. http://dx.doi.org/10.2307/1912352

Henley, A. (2005). Job creation by the self-employed: The roles of entrepreneurial and financial capital. *Small Business Economics*, *25*(2), 175–196. http://dx.doi.org/10.1007/s11187-004-6480-1

Imbens, G. W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635. http://dx.doi.org/10.1016/j.jeconom.2007.05.001

Monasterio, L. M. (2014). A estranha distribuição da população dos pequenos municípios brasileiros. *Revista Econômica do Nordeste*, *45*(4), 109–117. http://dx.doi.org/https://ren.emnuvens.com.br/ren/article/view/112

Onibokun, A. G. (1983). *Issues in Nigerian housing: A bibliographic review*. Nigerian Institute of Social and Economic Research (NISER), University of Ibadan.

Peng, R., & Wheaton, W. C. (1994). Effects of restrictive land supply on housing in Hong Kong: An econometric analysis. *Journal of Housing Research*, *5*(2), 263–291. https://www.jstor.org/stable/24832796

Pollakowski, H. O., & Wachter, S. M. (1990). The effects of land-use constraints on housing prices. *Land Economics*, *66*(3), 315–324. http://dx.doi.org/10.2307/3146732

Rocha, V. A. d., & Belluzzo, W. (2010). Avaliação do programa de descentralização de gastos públicos no sistema municipal de ensino fundamental de São Paulo. *In 38º encontro nacional de economia da anpec*, Salvador, BA. http://www.anpec.org.br/encontro2010/inscricao/arquivos/514-02770290564494fd9766bfc02488caa0.pdf

Saks, R. E. (2008). Job creation and housing construction: Constraints on metropolitan area employment growth. *Journal of Urban Economics*, *64*(1), 178–195. http://dx.doi.org/10.1016/j.jue.2007.12.003

Salau, A. T. (1990). The environmental context of urban housing-public services and infrastructural facilities in Nigerian urban centres. *Urban Housing in Nigeria Ibadan: NISER*, 58–88.

Smith, A. C. (2016). Spring forward at your own risk: Daylight saving time and fatal vehicle crashes. *American Economic Journal: Applied Economics*, *8*(2), 65–91. http://dx.doi.org/10.1257/app.20140100

Teixeira, L. P., & Carvalho, F. M. A. d. (2005). A construção civil como instrumento do desenvolvimento da economia brasileira. *Revista Paranaense de Desenvolvimento*(109), 9–26. http://www.ipardes.pr.gov.br/ojs/index.php/revistaparanaense/article/view/138

Thistlethwaite, D. L., & Campbell, D. T. (1960). Regression-discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*(6), 309–317. http://dx.doi.org/10.1037/h0044319

Toro, W., Tigre, R., & Sampaio, B. (2015). Daylight saving time and incidence of myocardial infarction: Evidence from a regression discontinuity design. *Economics Letters*, *136*, 1–4. http://dx.doi.org/10.1016/j.econlet.2015.08.005

Trochim, W. M. K. (1984). *Research design for program evaluation: The regression-discontinuity approach* (Vol. 6). Sage.

Zhao, P., Lü, B., & De Roo, G. (2011). Impact of the jobs-housing balance on urban commuting in Beijing in the transformation era. *Journal of Transport Geography*, *19*(1), 59–69. http://dx.doi.org/10.1016/j.jtrangeo.2009.09.008

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180025
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Avaliação de Impacto das Guardas Municipais Sobre a Criminalidade com o uso de Tratamentos Binários, Multivalorados e Contínuos

Oliveira Alves Pereira Filho[*]
Maria da Conceição Sampaio de Sousa[†]
Patrick Franco Alves[‡]

**Sumário**

1. Introdução ..... 515
2. As Guardas Municipais no Brasil . 517
3. Metodologia ..... 521
4. Bases de Dados ..... 523
5. Resultados ..... 528
6. Conclusões ..... 540
Apêndice. ..... 543

**Palavras-chave**
Guardas municipais, segurança pública, PSM, diferenças-em-diferenças, tratamentos multivalorados, GPS, dose-resposta

**JEL Codes**
K42, K32, H41, H76, C23

**Resumo · Abstract**

We evaluated whether the creation of Municipal Guards (GMs) improved the public safety indicators of the local governments that adopted them. Using municipal data, we tested several identification strategies based on the potential results model, focusing on binary treatments, discrete multivalued treatments and continuous treatments. The main results indicate that, in small and medium municipalities, having a GM with some years of activity may represent up to 30% fewer homicides than in their control group, which is equivalent to up to -4.8 deaths per 100 thousand inhabitants. On the contrary, the existence of GMs does not seem to impact property crimes.
não impactar as taxas de roubos e furtos de veículos, nossa *proxy* para os crimes contra o patrimônio.

## 1. Introdução

Pouco ainda há para se debater sobre o quão importante a segurança pública é para o cidadão brasileiro. Em conjunto com o recado não favorável passado pelas estatísticas do setor, pesquisa de opinião recente do Datafolha a elegeu como a 2º maior preocupação da população, atrás somente da saúde (JN, 2014). Se por um lado está claro qual é o incômodo, as estratégias sobre como viabilizar menores índices de criminalidade ainda continuam objeto de debate.

Fator agravante para esse quadro parece ser a divisão de competências imposta pela Constituição Federal de 1988 (CF, 1988). Por esse diploma, coube aos Estados e ao DF a parcela mais expressiva das atividades de segurança pública, relegando-se aos Municípios o papel subsidiário e facultativo de implantação de Guardas Municipais civis (GMs) em seus limites territoriais e com atribuições circunscritas de "... proteção de seus bens, serviços e

[*]Ministério do Planejamento, Desenvolvimento e Gestão – MPDG. SEPN 516 D 8, Asa Norte, Brasília, DF, CEP 70770-524, Brasil.

[†]Universidade Federal da Paraíba, Programa de Pós-Graduação em Economia (UFPB/PPGE). Jardim Cidade Universitária, João Pessoa, PB, CEP 58051-830, Brasil.

[‡]Instituto de Pesquisa Econômica Aplicada – IPEA. Setor Bancário Sul Q. 1, Ed. BNDES, sala 1223, Brasília, DF, CEP 70076-900, Brasil.

oliveira.ap.filho@gmail.com csampaiodesousa@gmail.com patrickfrancoalves@yahoo.com.br



This is an open-access article distributed under the terms of the Creative Commons Attribution License.

Recebido em 14 de maio de 2017
Aceitação final em 10 de julho de 2018

instalações, conforme dispuser a lei." Ocorre, no entanto, que aos olhos dos eleitores/cidadãos os problemas acontecem em seus quintais, não sendo plenamente possível desassociar criminalidade e atuação política local.

Nesse sentido, um número crescente de Municípios tem optado por criar e gerir seus próprios aparatos de GMs, esperançosos de que possam responder aos anseios da sociedade em termos de redução na criminalidade. Todavia, esse movimento tem sido posto em prática de maneira predominantemente empírica, sem avaliações de eficiência ou de impacto que o referendem. Se tais iniciativas forem, por algum motivo, inócuas, ter-se-á o pior dos cenários, a saber: continuidade da sensação de insegurança com a prevalência de maiores gastos orçamentários, antes inexistentes.[1]

Objetivando contribuir com o preenchimento dessa lacuna acadêmica, este artigo tem por finalidade avaliar e mensurar os impactos advindos da criação e da manutenção dessas estruturas municipais de segurança pública (GMs) sobre os principais indicadores de criminalidade disponíveis — taxa de homicídios por 100 mil habitantes e taxa de roubos e furtos de veículos por 100 mil emplacamentos. A hipótese básica é de que essas corporações possuem pouco potencial de afetar esses indicadores e isso se daria por conta do seu pequeno tamanho relativo (em relação ao papel dos Estados e do DF e também em relação à União) e, principalmente, pela indefinição em torno do seu modelo de atuação, ora identificado como de segurança patrimonial ou de policiamento comunitário ou ainda possivelmente como força repressora auxiliar.

A base de dados utilizada é proveniente de diversos anos da Pesquisa de Informações Básicas Municipais – Munic (Perfil dos Municípios Brasileiros) do Instituto Brasileiro de Geografia e Estatística – IBGE, e consiste em dois painéis fortemente balanceados, sendo o maior deles (1991–2012) utilizado sempre que a variável dependente for a taxa de homicídios por 100 mil habitantes e o menor (2009–2011) para quando a variável de resultado for a taxa de roubos e furtos de veículos por 100 mil veículos emplacados.

As estratégias de identificação utilizadas para se determinar a relação causal entre GMs e criminalidade devem, inicialmente, superar a possível existência de um viés de autosseleção nos entes que decidiram criar tais instituições, bem como se atentar para a possibilidade de endogeneidade por simultaneidade, algo recorrente em mensurações relacionadas à economia do crime.[2] Em suma, com o intuito de superar tais patologias, foca-se na criação de grupos de controle compostos por Municípios estatisticamente similares aos do grupo de tratamento, exceto pela decisão de criarem ou não a sua Guarda Municipal, o que permitiria estimativas confiáveis para o impacto causal avaliado.

Além de se propor mensurações focadas nos impactos do programa como uma variável binária, isto é, se o ente recebeu ou não o tratamento — Diferenças-em-Diferenças (DID), *Propensity Score Matching* (PSM) simples e em conjunto com Diferenças-em-Diferenças (PSM-DID)—, ainda foram investigados os casos de tratamentos multivalorados, sejam esses por categorias ou aplicados de maneira contínua — *Generalized Propensity Score* (GPS) e função Dose-Resposta (DR). A justificativa reside no fato de que mesmo entre os

[1] Conforme será debatido nas seções posteriores, as despesas orçamentárias dos Municípios com segurança pública — predominantemente criação e manutenção de Guardas Municipais — elevaram-se substancialmente no período 2000–2011 (crescimento de mais de 250% acima da inflação do período).

[2] As abordagens propostas neste artigo estão fundamentadas na denominada Economia do Crime, isto é, um conjunto teórico, sujeito a testes empíricos, que relaciona criminalidade com racionalidade econômica. Nesse caso, o trabalho seminal é de Becker (1968) e, em língua portuguesa, há a instigante resenha de Cerqueira e Lobão (2003) sobre o papel do crime segundo teorias de diversas áreas de conhecimento.

participantes de um dado programa (grupo de tratamento) existem diferentes níveis de exposição e/ou intensidade que poderiam interferir nos resultados obtidos. Os indicadores de múltiplos tratamentos considerados neste estudo foram os anos de exposição à GM.

Isso posto, este artigo está subdividido em mais 4 seções, além desta Introdução. Imediatamente a seguir são debatidas as principais características legais e funcionais dessas corporações. Na seção 3 há uma sucinta apresentação dos métodos econométricos anteriormente citados, ao passo que nas seções 4 e 5 detalham-se, respectivamente, a base de dados utilizada e os diversos resultados estimados. O estudo se encerra na seção 6 com uma síntese dos principais achados.

## 2. As Guardas Municipais no Brasil

O instituto das Guardas Municipais – GMs, visto de maneira ampla como a possibilidade de participação dos Municípios nos esforços de segurança pública, não é um fenômeno recente no Brasil, contudo toma relevo apenas após a consolidação legal trazida pela Constituição Federal de 1988 (CF, 1988) que, em seu artigo 144, §8, especifica e delimita a abrangência de tais iniciativas,[3] conforme segue:

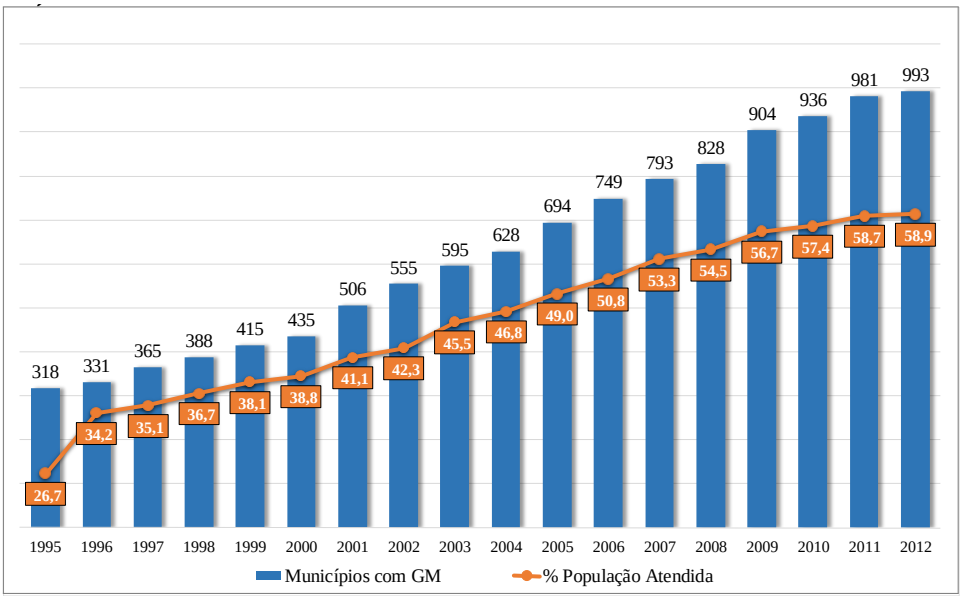
> Art. 144. A segurança pública, dever do Estado, direito e responsabilidade de todos, é exercida para a preservação da ordem pública e da incolumidade das pessoas e do patrimônio, através dos seguintes órgãos:
> ...
> §8 Os Municípios poderão constituir guardas municipais destinadas à proteção de seus bens, serviços e instalações, conforme dispuser a lei.

Restou inconteste o caráter subsidiário relegado pelos Constituintes aos Municípios ("poderão"), assim como, pela letra fria da lei, deverão atuar numa dimensão notadamente restrita ("proteção de seus bens, serviços e instalações"). Afora esse caráter normativo, impende ponderar as mudanças no tecido social ocorridas após 1988, em especial o cenário desfavorável de crescimento da criminalidade, sobretudo nos grandes centros urbanos nos anos 1990. Essas questões de cunho prático passam a pautar as criações e a atuação de tal aparato, como uma possível resposta em termos de política pública aos anseios da sociedade que invariavelmente identifica também o Prefeito como um possível responsável pelo quadro deteriorado da segurança e não apenas Governadores ou Presidentes.

No gráfico da Figura 1 é possível verificar que houve incremento no quantitativo de Municípios que optaram pela instalação desses aparatos de segurança pública, haja vista serem apenas 318 em 1995 e em 2012 passaram para 993 (17,8% do total de Municípios existentes), isto é, foram criadas 675 novas GMs no período. Como essas corporações são proporcionalmente mais presentes em cidades maiores, tem-se que o percentual da população atendida por tal serviço atingiu em 2012 o maior valor da série, 58,9% (em 1995 eram apenas 26,7%).

Em termos de recursos financeiros, a trajetória ascendente é praticamente a mesma, conforme esclarecido pelas informações da Tabela 1. Ainda que a participação relativa dessas alocações orçamentárias no total do setor de segurança pública brasileiro seja discreta, houve

[3] De acordo com dados da Munic 2012 (IBGE, 2013), dos 993 Municípios que declararam possuir GM, cerca de 9,8% informaram que elas existem desde antes de 1988.




*Source:* Elaboração dos autores a partir das informações da Munic 2012 (IBGE, 2013).

**Figura 1.** Evolução do Número de Municípios com GM e % da População Atendida (1995–2012).

**Tabela 1.** Despesas na Função Segurança Pública por Entes Federados – R$ Bilhões Constantes de 2011 (IPCA Médio)

| Unidade da Federação | 2000 | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 | 2011 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TOTAL | 29,2 | 35,5 | 38,3 | 36,5 | 36,2 | 39,7 | 44,2 | 48,4 | 52,2 | 57,3 | 60,1 | 59,6 |
| Governo Federal | 4,4 | 4,9 | 3,9 | 3,7 | 4,0 | 4,0 | 5,0 | 6,6 | 7,7 | 9,0 | 10,4 | 7,7 |
| Polícia Federal | 1,7 | 1,9 | 2,0 | 1,8 | 2,1 | 2,1 | 2,5 | 3,0 | 3,3 | 3,3 | 3,4 | 3,3 |
| Polícia Rod. Federal | 0,8 | 1,0 | 0,1 | 1,0 | 1,1 | 1,2 | 1,5 | 1,7 | 1,6 | 1,7 | 1,8 | 1,8 |
| FNSP | 0,5 | 0,8 | 0,6 | 0,5 | 0,4 | 0,3 | 0,4 | 1,0 | 0,4 | 0,2 | 0,2 | 0,3 |
| Outros | 1,5 | 1,3 | 1,1 | 0,4 | 0,4 | 0,4 | 0,5 | 0,9 | 2,5 | 3,7 | 5,0 | 2,4 |
| Estados | 22,2 | 27,8 | 31,1 | 29,4 | 28,8 | 32,1 | 35,1 | 37,3 | 39,2 | 42,7 | 43,9 | 45,5 |
| Distrito Federal | 1,8 | 1,8 | 2,2 | 2,3 | 2,1 | 2,3 | 2,5 | 2,8 | 3,1 | 3,4 | 3,4 | 3,6 |
| Próprios | 1,8 | 1,8 | 2,2 | 0,4 | 0,2 | 0,2 | 0,1 | 0,1 | 0,2 | 0,2 | 0,3 | 0,3 |
| FCDF | – | – | – | 1,9 | 2,0 | 2,1 | 2,4 | 2,7 | 3,0 | 3,2 | 3,1 | 3,3 |

*Fonte:* Elaboração dos autores a partir dos dados da STN, SIAFI e IBGE.

um crescimento real de 251,9% nessas despesas, as quais saíram de R$0,8 bilhão em 2000 para R$2,8 bilhões em 2011.

A ocorrência de GMs se intensificou, em termos relativos, nos Municípios pequenos, ao passo que caiu nas grandes cidades, conforme pode ser vislumbrado no gráfico da Figura 2. Em 1991, cerca de 27,5% das GMs existentes encontravam-se em localidades com menos de 20.000 habitantes e em 2012 essa participação já era de 32,9%. Em caminho contrário, nos Municípios maiores (população acima de 250.000) o percentual relativo retraiu-se de 12,6% (em 1991) para 8,6% em 2012.

No contexto evolutivo dessas GMs, Vargas e Oliveira (2010) conseguem visualizar indicativos de que essas instituições de fato têm extravasado seu restrito mandato constitucional e, sob um enfoque positivo, propõem um modelo para a análise do funcionamento e da atuação desses entes pautado por três critérios, não necessariamente excludentes:

1. Polícia Municipal: Aplicador da lei e repressor;
2. Polícia Comunitária: Apaziguador e mediador;
3. Guarda Patrimonial: Protetor de bens e de pessoas.

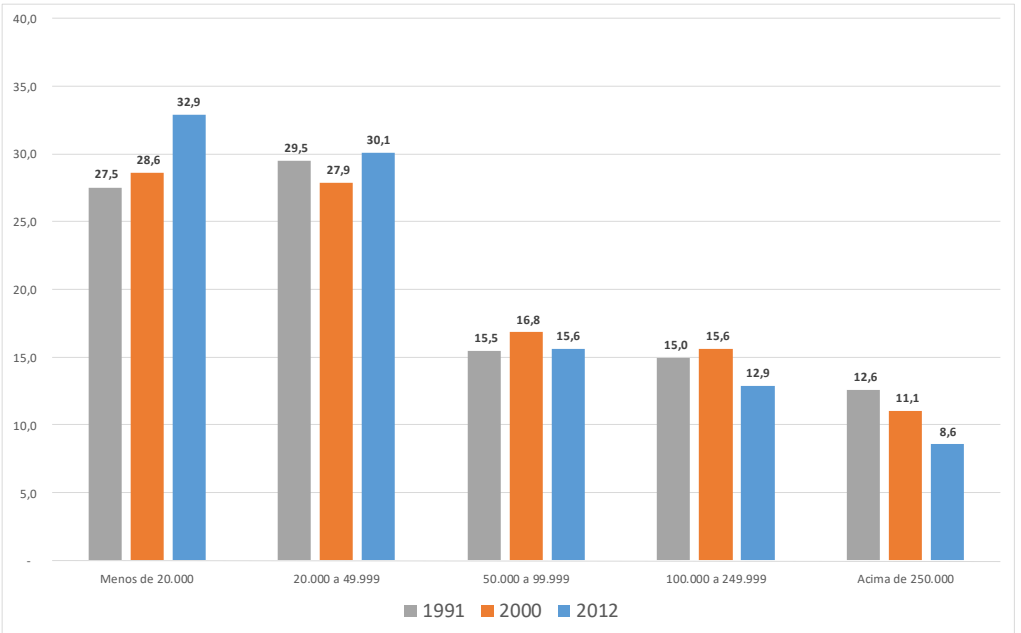
De forma complementar, o questionário da Munic 2012 (IBGE, 2013) dá números a esse tipo de debate, na medida em que registra quais são as principais atividades exercidas por essas GMs. Conforme se depreende do gráfico da Figura 3, a atuação dessas corporações é diversificada e com características convergentes aos três critérios estabelecidos por Vargas e Oliveira (2010). Nota-se que há prevalência de atividades de proteção a bens (93,4%), mas também se observam componentes de policiamento, propriamente dito (por exemplo, Auxílio à Polícia Militar, 72,9%; Patrulhamento ostensivo, 67,7%; e Auxílio à Polícia Civil, 59,3%).

Portanto, parecem existir evidências suficientes em favor do entendimento de que as GMs não se comportam apenas como vigilantes ou porteiros do patrimônio público e que a cada ano o seu potencial de influenciar os indicadores do sistema de segurança pública brasileiro vem aumentando. Se isso é verdade, convém sejam envidados esforços para se estimar em qual magnitude se dá esse impacto, seara em que este artigo pretende contribuir por meio das estratégias detalhadas em seções posteriores.

Apesar do crescente apelo envolvendo tais atividades locais de segurança pública, ainda são poucas as contribuições empíricas que tomam como base o modelo de racionalidade econômica aplicado a atividades criminais para mensurar o impacto das GMs brasileiras nos diversos indicadores de criminalidade.

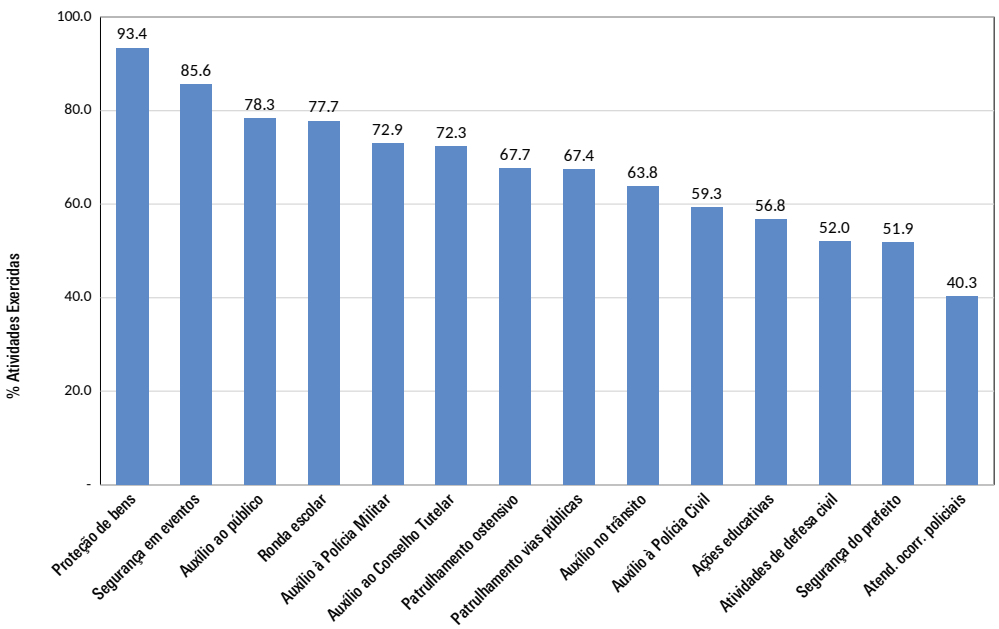
O trabalho de Resende e Andrade (2011), ainda que não tenha como foco o papel das GMs, propõe, sob uma estrutura de dados *cross-section* para o ano de 2004 ($n = 225$), duas medidas para as atividades das Guardas, sendo a primeira uma variável *dummy* $= 1$ se o ente possui essa instituição implantada, zero em caso contrário, e a segunda refere-se a um índice de qualidade para essas atividades, composto de três dimensões: i) escolaridade do efetivo; ii) razão entre o piso salarial e a renda *per capita* do Município; e iii) razão entre o número de viaturas e o número de guardas. Utilizando tais variáveis apenas como controles em regressões OLS explicativas de diversos tipos de delitos criminais, os autores rejeitam a significância estatística em ambas as variáveis.[4]

[4]Exceção foi a estimação que levou em consideração como variável dependente os homicídios declarados pelas polícias estaduais à SENASP. Contudo, as mesmas variáveis não foram significativas quando a variável




*Source:* Elaboração dos autores a partir das informações da Munic 2012 (IBGE, 2013).

**Figura 2.** Localização das GMs em %, segundo Grupos Populacionais e Anos Selecionados.




*Source:* Elaboração dos autores a partir das informações da Munic 2012 (IBGE, 2013).

**Figura 3.** Principais atividades exercidas pelas Guardas Municipais (em %) – 2012.

A principal contribuição dada pela dissertação de mestrado de Ferreira (2012) foi no sentido de pioneiramente propor avaliação centrada no papel das GMs na redução de um crime em específico: os homicídios. A base de dados utilizada foi um painel contendo informações dos anos de 2004, 2006 e 2009 (anos em que a Munic analisou a questão da segurança pública), sendo 3 (três) as variáveis explicativas diretamente relacionadas às GMs: i) *dummy* de presença; ii) contagem para os anos de existência de cada GM; e iii) efetivo por 100 mil habitantes. As estimações realizadas apenas por POLS e por Efeitos Fixos (*within*), comparando Municípios com e sem GMs, não foram capazes de superar a patologia de simultaneidade recorrente em modelos econômicos de criminalidade, o que inviabilizou a análise dos resultados apresentados, mas reforçou a necessidade de estratégias de identificação mais robustas para avaliações desse contexto.[5]

A busca na literatura internacional retornou evidências de que o modelo de GMs utilizado no Brasil é *sui generis*, em especial pelo seu componente de participação facultativo. Até onde foi possível pesquisar, não foram encontradas experiências empíricas similares, inclusive em modelos regionalizados de policiamento. Mesmo para esses, notou-se a obrigatoriedade de sua existência como uma característica marcante, relacionada ao fato de que os serviços de segurança pública são monopólios estatais resguardando de maneira inalienável o uso da força ao Poder constituído. Para situações como essa, métodos baseados no arcabouço de resultados potenciais (que em geral necessitam de grupos de controle) não se mostram tão eficazes como, por exemplo, estimações comparativas por meio de fronteiras de eficiência.[6]

## 3. Metodologia

Conforme debatido nas seções anteriores, a decisão de implantar uma GM é uma faculdade inerente aos próprios Municípios, diferindo, portanto, das demais ações da área de segurança pública que são constitucionalmente mandatórias (por exemplo, existência de polícias civis e militares nos Estados e no DF). Essa nuance indica, entre outras consequências, que a designação ao tratamento (criação de uma GM) não é um processo aleatório. Mais ainda, pode existir um componente de autosseleção associado com características não observáveis desses participantes e isso os tornaria uma espécie de subamostra especializada, com possíveis impactos sobre suas variáveis de resultado. Como esse parece ser o caso, a tarefa de se isolar o efeito causal desse tipo de programa se torna ainda mais dificultosa.

Nesse contexto comparativo, busca-se definir como os Municípios tratados (com GM) se sairiam sem possuí-las e também como seria o desempenho daqueles entes que não foram tratados, caso o fossem. Como não é possível observar os mesmos indivíduos simultaneamente em duas situações, há a necessidade de se estimar resultados potenciais com o auxílio de contrafactuais. O trabalho de Imbens e Wooldridge (2009) resenha de maneira crítica os desenvolvimentos recentes desse campo de pesquisa.

---

dependente foram os homicídios registrados pela base do DATASUS, mais ampla e menos sujeita ao subregistro, o que demanda cautela nas interpretações.

[5] Posteriormente, essa dissertação foi publicada com melhorias, dentre as quais a opção por uma análise restrita aos Municípios do Estado de São Paulo, mas com a inclusão dos crimes de roubos e furtos na análise (Ferreira, Mattos, & Terra, 2016). Nesse novo contexto, os autores apresentam evidências de que as GMs acarretaram reduções tanto na taxa de homicídios quanto na taxa de roubos e furtos dos Municípios paulistas.

[6] Um pouco mais de informação sobre os modelos europeus e norte-americano de policiamento pode ser obtido em Bayley (1990) e em Monet (2001).

Neste artigo optou-se por avaliar os possíveis impactos advindos da criação de GMs, uma lacuna na pesquisa econômica do crime, por três enfoques distintos, cada um com suas respectivas estratégias de identificação. No primeiro caso investigamos apenas o aspecto binário (se possui ou não possui GM) e logo após analisamos de duas maneiras distintas também a intensidade do tratamento (anos com a GM): i) como se composto por múltiplos níveis;[7] e ii) como um processo contínuo.

Para a análise binária, a estratégia de identificação é composta de estimações comparativas envolvendo o método DID (Meyer, 1995, Bertrand, Duflo, & Mullainathan, 2004), seguido pelo PSM (Rosenbaum & Rubin, 1983, Caliendo & Kopeinig, 2005) e por uma combinação de ambos (DID-PSM) (Heckman, Ichimura, & Todd, 1998, 1997). Os Municípios são divididos em dois grupos, o de tratamento composto por aqueles que possuem GM e o de controle, formado por aqueles entes que não possuem GM. Em adição, buscou-se contornar o possível problema de endogeneidade por simultaneidade, comum em atividades de segurança pública, por meio da criação de grupos de controle e de tratamento defasados, isto é, foram criadas diversas variáveis binárias definindo há quantos anos o Município conta com a GM e para cada uma dessas situações foram rodadas regressões individualizadas, com e sem covariadas de controle.

Por outro lado, nem sempre é crível imaginar que participantes de um dado tratamento recebam doses homogêneas dessa intervenção, superiores apenas àquelas dos que não participaram do programa, como é o suposto implícito das avaliações de impacto com variáveis binárias. No caso prático deste artigo, é de se conjecturar que, mesmo entre os tratados, a longevidade das GMs em cada Município (anos em funcionamento) possa afetar o seu desempenho.[8]

Para situações multivaloradas como essa (categóricas ou discretas), a contribuição pioneira de Imbens (2000) registra que para a estimação de efeitos causais não é de fato necessário dividir a amostra de indivíduos em subpopulações em que as comparações causais seriam válidas, como faz o PSM, sendo suficiente repartir essa base de dados em subpopulações em que os produtos potenciais médios possam ser estimados.

Respeitadas as hipóteses debatidas em Wooldridge (2010, p.962), as rotinas de estimação para o efeito causal de tratamentos multivalorados discretos, como os utilizados neste artigo, consistem no uso, em separado, de modelos preditores para: i) a variável de resultado (ajustamento por regressão – RA, do tipo log-linear e de Poisson, dada a natureza não negativa das taxas criminais); e ii) a variável de designação ao tratamento (estimadores baseados em ponderação pelo inverso da probabilidade – IPW, obtidos com o uso de um *logit multinomial*). Isso feito, são adicionalmente testados estimadores do tipo duplamente robustos (*doubly robust*), associando os benefícios de cada uma das abordagens anteriores (IPWRA),[9] conforme definição e apresentação detalhada de Hirano e Imbens (2001), Wooldridge (2010, p.930–934), e de StataCorp (2013).

[7] Nesse caso, as variáveis de tratamento são agrupadas em 5 categorias distintas, conforme sua dosagem, sendo o primeiro nível referente aos Municípios sem GM e o último, por gradação, remete àqueles entes que possuem maior exposição ao programa.

[8] Essa variável está dividida em cinco categorias, denotando sempre dosagens maiores a partir do primeiro estrato (de não tratados).

[9] De acordo com Robins e Rotnizky (1995) *apud* Hirano e Imbens (2001, p.264), após combiná-los (IPWRA), basta que apenas um dos dois modelos (RA ou IPW) estejam corretamente especificados para que os estimadores resultantes sejam consistentes.

Uma maneira alternativa de verificar se o aumento na intensidade do tratamento (longevidade das GMs) acarreta impactos diferenciados sobre as variáveis de resultado (taxas de criminalidade) é encarando essas intervenções como contínuas. Como em outras ocasiões, tal estratégia necessita de unidades com características observáveis suficientemente similares, mas que agora contem com níveis diferenciados nas doses dos respectivos tratamentos.[10]

Os trabalhos pioneiros de Imbens (2000) e de Hirano e Imbens (2004) fornecem o arcabouço teórico para aplicações desse tipo, baseando-as inicialmente na noção de Escore de Propensão Generalizado – GPS, isto é, na probabilidade condicional de se receber um nível particular de tratamento ($t$) dadas as variáveis pré-tratamento ($x \in X$).

A rotina de operacionalização econométrica seguiu as contribuições de Guardabascio e Ventura (2013), cujo principal enfoque foi flexibilizar a estimação do GPS antes proposta por Bia e Mattei (2008), em especial para situações em que a variável de tratamento não é normalmente distribuída.

## 4. Bases de Dados

Neste artigo as rotinas econométricas encontram-se pautadas na utilização dos Municípios (exceto o Distrito Federal) como unidades de análise, portanto o universo potencial são todos os entes locais que existiam em cada ano do painel analisado (em 2012, por exemplo, eram 5.564). Dado o número satisfatório de observações, foi possível investigar os fenômenos de interesse por meio de classes de Municípios, isto é, agrega-los por similaridades populacionais a fim de minimizar problemas de regressão à média. Assim, utilizamos os seguintes agrupamentos de cidades:

**Grupo 1:** Municípios Pequenos (população de 20.000 a 49.999);
**Grupo 2:** Municípios Médios (de 50.000 a 99.999);
**Grupo 3:** Municípios Intermediários (de 100.000 a 249.999); e
**Grupo 4:** Municípios Grandes (acima de 250.000).[11]

Utilizam-se nas estimações dos modelos econométricos dois conjuntos distintos de dados, a saber: i) um painel fortemente balanceado compreendendo os anos de 1991 a 2012, nas situações em que a variável de resultado sejam as taxas de homicídios por 100 mil habitantes (*proxy* para os crimes contra a pessoa); e ii) um painel não equilibrado com informações dos anos de 2009 a 2011 para quando a variável dependente tratar das taxas de roubos e furtos de veículos para cada 100 mil veículos emplacados (estimativa para os crimes contra o patrimônio).

Os registros de homicídios são provenientes do Sistema de Informações sobre Mortalidade do Ministério da Saúde – SIM. Optou-se pela sua utilização basicamente por possuírem adequada abrangência espacial e temporal e por ser a variável criminal menos sujeita ao problema de subregistro presente em ocorrências policiais, garantindo com isso a

[10] É oportuno ressaltar que nesse arcabouço, ao contrário daquele focado em tratamentos multivalorados discretos, não se está mais analisando as diferenças em diversos níveis de tratamento em relação ao grupo de controle (dosagem = 0), mas sim averiguando se a intensidade do tratamento, entre uma subamostra de tratados, repercute sobre as respectivas variáveis de resultado.

[11] Os Municípios muito pequenos (abaixo de 20.000 habitantes) foram propositalmente excluídos da amostra porque nessas localidades a dinâmica da criminalidade ainda é muito distante do que existe nos seus congêneres. Por exemplo, em 1991 existiam 3.093 desses entes na nossa base, sendo que em 2.992 não haviam sido registrados nenhum homicídio (96,7% do total). Em 2012, tínhamos 3.896 entes com população menor do que 20.000, sendo que em 2.866 observações não ocorreram homicídios (73,6% do total).

comparabilidade entre os entes avaliados.[12] O mesmo se aplica à série de roubos e furtos de veículos, com o benefício adicional de possibilitar a utilização de uma *proxy* também para as atividades de segurança pública relacionadas aos crimes contra o patrimônio. Por fim, até onde foi possível investigar, a compilação dessa série em nível municipal para os anos de 2009 a 2011 configura-se como uma contribuição inédita para a pesquisa de segurança pública nacional.[13]

Consoante a cada estratégia de identificação adotada, foram 2 (duas) as variáveis de interesse utilizadas para se estimar o possível impacto das guardas municipais na criminalidade: i) $dummy = 1$ caso o Município possua guarda municipal em dado ano, zero caso contrário; e ii) tempo de existência (em anos) da guarda municipal em um dado município em cada ano. Todas foram elaboradas a partir de diversas edições da Pesquisa de Informações Básicas Municipais – Munic do Instituto Brasileiro de Geografia e Estatística – IBGE (Munic 2012 e a Munic 2014).

Os controles utilizados, adiante pormenorizados, seguem as recomendações da teoria da racionalidade econômica aplicada a atividades criminosas e também de componentes da teoria da desorganização social. Nas Tabelas 2 (grupos de tratamento e de controle) e 3 (apenas grupo de tratamento) são apresentadas, com uma breve definição e com as respectivas estatísticas descritivas, as variáveis utilizadas nos modelos que explicam os resultados (taxas de criminalidade) e nas especificações que investigam os condicionantes ao tratamento (existência ou não de guarda municipal e a sua longevidade). Todas estão separadas pelos grupos populacionais estipulados para o agrupamento dos Municípios da amostra.

Inicialmente, buscou-se controlar para a possível existência de um dado "custo moral" em se associar com eventos criminais, conforme previsto pela teoria. Sendo a criminalidade um ilícito, é de se ponderar que a refração a tal atitude pode também variar conforme considerações éticas e psicológicas, agindo na prática como uma barreira ou um estímulo à entrada em tal mercado, tudo o mais constante. Segundo alguns autores, a difusão dos televisores pelos lares tem o potencial de propagar padrões culturais mais frouxos e condutas ético-morais contestáveis, o que tenderia a tornar, na margem, atividades ilícitas menos reprováveis ou até mesmo desejáveis.[14] Nossa variável para tal fenômeno (*cmoral*) foi construída com dados do IBGE referentes ao percentual de pessoas que vivem em domicílios com aparelho de televisão (preto e branco ou colorido).

No tocante às condições demográficas, que quando saturadas podem estimular ambientes favoráveis a ilícitos, foram utilizadas medidas para a densidade demográfica (pessoas por km²) de um dado Município (*dens*), para o percentual de pessoas que residem nas áreas urbanas dessas localidades (*txurbana*) e para a presença de jovens do sexo masculino na

[12] Até 1995 esses registros são os constantes do Grande Grupo da CID9 "E55 Homicídios e lesões prov intenc p/outr pessoas". A partir de 1996 extraímos os homicídios dos Grandes Grupos da CID-10 "X85-Y09 Agressões" e "Y35-Y36 Intervenções legais e operações de guerra".

[13] Essa série foi obtida diretamente da base de registros "BIN Roubos e Furtos" do Departamento Nacional de Trânsito – DENATRAN em extração pontual realizada pela Confederação Nacional de Seguros – CNSeg, uma das poucas entidades a ter acesso a tais dados. As informações estão disponíveis apenas para os Municípios que em 2009 possuíam frota igual ou superior a 5.000 veículos (1.292 observações anuais). Não foram utilizados os dados provenientes dos Estados da Bahia e do Amazonas, pois neles todos os registros locais são armazenados nos Municípios da Capital (Salvador e Manaus), o que prejudicaria a comparabilidade.

[14] Sobre o assunto, algumas evidências podem ser encontradas em Hennigan et al. (1982); Centerwall (1992); e Robertson, McAnally, e Hancox (2013).

**Tabela 2.** Estatísticas Descritivas (1991–2012) –- Média Geral (Grupos de Tratamento e de Controle).

| | Grupo 1 (20.000 a 49.999) | | | | | Grupo 2 (50.000 a 99.999) | | | | | Grupo 3 (100.000 a 249.999) | | | | | Grupo 4 (Acima de 250.000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variável | Obs. | Média | Desv. Padrão | Mín. | Máx. | Obs. | Média | Desv. Padrão | Mín. | Máx. | Obs. | Média | Desv. Padrão | Mín. | Máx. | Obs. | Média | Desv. Padrão | Mín. | Máx. |
| *txhom* [1] | 21.621 | 14,4 | 15,9 | 0,0 | 160,6 | 6.651 | 19,6 | 18,1 | 0,0 | 212,4 | 3.245 | 27,7 | 21,9 | 0,0 | 183,8 | 1.856 | 36,1 | 24,1 | 0,0 | 157,6 |
| *txroubfurt* [2] | 1.612 | 211,5 | 205,9 | 4,3 | 2.380,7 | 787 | 284,4 | 278,7 | 5,5 | 2.268,2 | 499 | 489,0 | 411,1 | 1,9 | 2.419,2 | 283 | 734,2 | 462,4 | 1,0 | 2.713,2 |
| *gm* [3] | 21.621 | 0,2 | 0,4 | 0 | 1 | 6.651 | 0,3 | 0,5 | 0 | 1 | 3.245 | 0,5 | 0,5 | 0 | 1 | 1.856 | 0,6 | 0,5 | 0 | 1 |
| *tgm* [4] | 21.621 | 1,5 | 4,8 | 0 | 48 | 6.651 | 3,1 | 7,2 | 0 | 52 | 3.245 | 7,3 | 11,9 | 0 | 65 | 1.856 | 10,9 | 16,6 | 0 | 88 |
| *efetivo* [5] | 1.073 | 141,6 | 150,8 | 0,0 | 2.460,6 | 515 | 99,0 | 93,7 | 1,0 | 794,6 | 431 | 98,4 | 82,8 | 3,3 | 645,0 | 295 | 63,4 | 35,8 | 6,4 | 214,6 |
| *cmoral* [6] | 21.610 | 75,4 | 22,1 | 2,8 | 100,0 | 6.648 | 84,4 | 16,1 | 7,4 | 99,8 | 3.234 | 91,2 | 9,7 | 16,6 | 99,8 | 1.856 | 94,5 | 5,1 | 52,2 | 99,3 |
| *dens* [7] | 21.580 | 63,0 | 96,1 | 0,2 | 1.594,3 | 6.636 | 185,8 | 434,5 | 0,4 | 5.808,4 | 3.230 | 768,9 | 1.522,4 | 0,6 | 12.223,0 | 1.853 | 2.240,5 | 2.881,2 | 5,2 | 13.500,8 |
| *txurbana* [8] | 21.586 | 64,6 | 22,2 | 3,4 | 100,0 | 6.648 | 78,8 | 18,9 | 1,6 | 100,0 | 3.245 | 90,6 | 11,1 | 14,5 | 100,0 | 1.856 | 95,1 | 6,6 | 31,1 | 100,0 |
| *txhomensj* [9] | 21.597 | 20,7 | 1,8 | 10,4 | 63,2 | 6.645 | 21,0 | 1,5 | 13,8 | 33,6 | 3.245 | 21,6 | 1,3 | 16,2 | 32,1 | 1.856 | 21,5 | 1,0 | 17,5 | 24,6 |
| *capital* [10] | 21.621 | 0,0 | 0,0 | 0 | 1 | 6.651 | 0,0 | 0,0 | 0 | 1 | 3.245 | 0,0 | 0,1 | 0 | 1 | 1.856 | 0,3 | 0,4 | 0 | 1 |
| *rm* [11] | 21.621 | 0,2 | 0,4 | 0 | 1 | 6.651 | 0,3 | 0,5 | 0 | 1 | 3.245 | 0,5 | 0,5 | 0 | 1 | 1.856 | 0,8 | 0,4 | 0 | 1 |
| *fronteira* [12] | 21.621 | 0,0 | 0,1 | 0 | 1 | 6.651 | 0,0 | 0,1 | 0 | 1 | 3.245 | 0,0 | 0,1 | 0 | 1 | 1.856 | 0,0 | 0,1 | 0 | 1 |
| *atraso* [13] | 19.707 | 13,3 | 9,6 | 0,0 | 91,5 | 6.616 | 11,4 | 6,8 | 0,0 | 53,3 | 3.245 | 10,0 | 5,7 | 0,0 | 48,3 | 1.856 | 8,9 | 4,0 | 1,1 | 26,4 |
| *nopai* [14] | 21.621 | 17,8 | 7,8 | 1,2 | 88,3 | 6.651 | 16,9 | 6,9 | 0,0 | 57,5 | 3.245 | 15,7 | 5,5 | 4,0 | 51,2 | 1.856 | 14,1 | 4,3 | 3,9 | 30,0 |
| *densdom* [15] | 21.621 | 39,9 | 16,5 | 3,7 | 94,7 | 6.651 | 37,4 | 15,1 | 6,7 | 91,0 | 3.245 | 36,4 | 13,5 | 8,1 | 83,4 | 1.856 | 34,9 | 12,5 | 8,4 | 72,3 |
| *gini* [16] | 21.621 | 54,5 | 5,9 | 24,2 | 84,0 | 6.651 | 54,0 | 5,5 | 32,5 | 74,0 | 3.245 | 52,5 | 5,7 | 33,7 | 68,0 | 1.856 | 53,9 | 6,7 | 33,2 | 72,5 |
| *abandono* [17] | 3.085 | 13,3 | 7,3 | 0,1 | 53,1 | 954 | 13,3 | 7,2 | 0,2 | 44,5 | 542 | 11,7 | 6,6 | 0,2 | 37,9 | 295 | 12,0 | 6,4 | 0,7 | 28,0 |
| *recprop* [18] | 14.672 | 203,0 | 341,0 | 0,0 | 18.989,0 | 4.727 | 276,4 | 301,2 | 0,4 | 4.116,7 | 2.398 | 432,7 | 448,4 | 5,6 | 4.403,8 | 1.455 | 544,6 | 391,9 | 26,4 | 2.502,7 |
| *armas* [19] | 21.417 | 35,2 | 32,0 | 0,0 | 100,0 | 6.651 | 43,6 | 25,6 | 0,0 | 100,0 | 3.245 | 51,2 | 22,0 | 0,0 | 100,0 | 1.856 | 58,4 | 19,9 | 0,0 | 100,0 |
| *prefgov* [20] | 19.322 | 0,2 | 0,4 | 0 | 1 | 6.074 | 0,2 | 0,4 | 0 | 1 | 3.064 | 0,2 | 0,4 | 0 | 1 | 1.818 | 0,2 | 0,4 | 0 | 1 |
| *prefcol* [21] | 19.322 | 0,3 | 0,4 | 0 | 1 | 6.074 | 0,3 | 0,5 | 0 | 1 | 3.064 | 0,3 | 0,5 | 0 | 1 | 1.818 | 0,3 | 0,5 | 0 | 1 |

*Fonte:* DATASUS (MSaúde), CNSEG, DENATRAN, IBGE, INEP, TSE e STN. Elaboração dos autores.

[1] Taxa de homicídios por 100 mil habitantes
[2] Taxa de roubos e furtos de veículos por 100 mil veículos emplacados
[3] Dummy = 1 se o Município possui guarda municipal, zero caso contrário
[4] Tempo de existência da guarda municipal (em anos)
[5] Efetivo das guardas municipais por 100 mil habitantes
[6] Proxy para o custo moral (% de pessoas que vivem em domicílios com aparelho de televisão)
[7] Densidade demográfica (pessoas por km$^2$)
[8] % da população vivendo em áreas urbanas (em %)
[9] % homens na faixa etária dos 15 aos 39 anos
[10] Dummy = 1 se o Município é capital de Estado
[11] Dummy = 1 se o Município faz parte de uma região metropolitana
[12] Dummy = 1 se o Município faz fronteira internacional
[13] % de 15 a 17 anos no ensino médio com 2 anos de atraso
[14] % mães chefes de famílias sem fundamental completo e com filhos menores de 15 anos
[15] % da população vivendo em domicílios com densidade maior que 2 pessoas por dormitório
[16] Índice de desigualdade de Gini da renda domiciliar per capita
[17] Taxa de abandono escolar no ensino médio (em %)
[18] Receita própria por habitante (exceto transferências)
[19] Proxy para a difusão de armas de fogo (% de suicídios e de homicídios com armas de fogo, defasado)
[20] Dummy = 1 se o Prefeito é do mesmo partido que o Governador do Estado, zero caso contrário
[21] Dummy = 1 se o Prefeito é da mesma coligação que elegeu o Presidente, zero caso contrário

**Tabela 3.** Estatísticas Descritivas (1991–2012) – Grupo de Tratamento.

| | Grupo 1 (20.000 a 49.999) | | | | | Grupo 2 (50.000 a 99.999) | | | | | Grupo 3 (100.000 a 249.999) | | | | | Grupo 4 (Acima de 250.000) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Variável | Obs. | Média | Desv. Padrão | Mín. | Máx. | Obs. | Média | Desv. Padrão | Mín. | Máx. | Obs. | Média | Desv. Padrão | Mín. | Máx. | Obs. | Média | Desv. Padrão | Mín. | Máx. |
| *txhom* [1] | 3.627 | 17,0 | 18,4 | 0,2 | 149,3 | 1.926 | 21,6 | 18,3 | 0,1 | 124,6 | 1.653 | 30,3 | 22,8 | 0,0 | 183,8 | 1.188 | 36,6 | 22,5 | 0,0 | 157,6 |
| *txroubfurt* [2] | 360 | 300,2 | 276,8 | 8,0 | 2.380,7 | 338 | 332,9 | 298,4 | 6,1 | 2.172,2 | 330 | 541,3 | 407,5 | 1,9 | 2.061,9 | 232 | 780,9 | 470,7 | 45,1 | 2.713,2 |
| *gm* [3] | 3.627 | 1,0 | 0,0 | 1 | 1 | 1.926 | 1,0 | 0,0 | 1 | 1 | 1.653 | 1,0 | 0,0 | 1 | 1 | 1.188 | 1,0 | 0,0 | 1 | 1 |
| *tgm* [4] | 3.627 | 9,2 | 8,1 | 0 | 48 | 1.926 | 10,8 | 9,7 | 0 | 52 | 1.653 | 14,4 | 13,4 | 0 | 65 | 1.188 | 17,0 | 18,0 | 0 | 88 |
| *efetivo* [5] | 873 | 144,8 | 158,2 | 0,0 | 2.460,6 | 445 | 101,9 | 95,0 | 1,0 | 794,6 | 404 | 96,9 | 81,2 | 3,3 | 645,0 | 282 | 64,3 | 35,3 | 12,3 | 214,6 |
| *cmoral* [6] | 3.621 | 83,9 | 16,3 | 2,8 | 100,0 | 1.923 | 89,9 | 10,8 | 17,3 | 99,5 | 1.646 | 93,5 | 6,4 | 51,9 | 99,8 | 1.188 | 95,6 | 3,8 | 70,1 | 99,3 |
| *dens* [7] | 3.621 | 99,3 | 143,7 | 0,2 | 1.594,3 | 1.923 | 229,8 | 457,7 | 0,5 | 5.596,1 | 1.648 | 969,0 | 1.863,0 | 0,6 | 12.223,0 | 1.188 | 2.643,9 | 2.971,3 | 45,9 | 13.500,8 |
| *txurbana* [8] | 3.626 | 70,6 | 22,0 | 4,0 | 100,0 | 1.926 | 81,4 | 17,3 | 14,9 | 100,0 | 1.653 | 92,0 | 8,8 | 22,2 | 100,0 | 1.188 | 95,8 | 5,9 | 35,4 | 100,0 |
| *txhomensj* [9] | 3.624 | 21,2 | 1,7 | 12,5 | 42,9 | 1.926 | 21,3 | 1,3 | 13,8 | 29,7 | 1.653 | 21,6 | 1,2 | 16,2 | 31,0 | 1.188 | 21,4 | 1,1 | 17,5 | 24,6 |
| *capital* [10] | 3.627 | 0,0 | 0,0 | 0 | 1 | 1.926 | 0,0 | 0,0 | 0 | 1 | 1.653 | 0,0 | 0,1 | 0 | 1 | 1.188 | 0,3 | 0,5 | 0 | 1 |
| *rm* [11] | 3.627 | 0,3 | 0,5 | 0 | 1 | 1.926 | 0,4 | 0,5 | 0 | 1 | 1.653 | 0,6 | 0,5 | 0 | 1 | 1.188 | 0,8 | 0,4 | 0 | 1 |
| *fronteira* [12] | 3.627 | 0,0 | 0,2 | 0 | 1 | 1.926 | 0,0 | 0,1 | 0 | 1 | 1.653 | 0,0 | 0,1 | 0 | 1 | 1.188 | 0,0 | 0,1 | 0 | 1 |
| *atraso* [13] | 3.406 | 12,1 | 8,2 | 0,0 | 57,6 | 1.922 | 9,8 | 5,9 | 0,0 | 53,3 | 1.653 | 8,4 | 4,4 | 0,0 | 25,4 | 1.188 | 8,2 | 3,5 | 1,2 | 20,8 |
| *nopai* [14] | 3.627 | 19,4 | 9,3 | 1,2 | 64,4 | 1.926 | 18,1 | 7,9 | 0,0 | 57,5 | 1.653 | 15,3 | 5,2 | 4,0 | 42,4 | 1.188 | 14,0 | 4,2 | 3,9 | 30,0 |
| *densdom* [15] | 3.627 | 38,8 | 13,8 | 9,6 | 91,7 | 1.926 | 35,8 | 12,2 | 7,6 | 80,6 | 1.653 | 34,7 | 11,9 | 8,8 | 75,4 | 1.188 | 34,2 | 11,7 | 8,4 | 67,0 |
| *gini* [16] | 3.627 | 53,5 | 6,5 | 32,7 | 81,6 | 1.926 | 53,1 | 5,5 | 32,5 | 71,0 | 1.653 | 52,2 | 5,7 | 33,7 | 68,0 | 1.188 | 53,6 | 7,3 | 33,2 | 72,5 |
| *abandono* [17] | 828 | 13,5 | 7,5 | 0,1 | 41,4 | 419 | 12,4 | 7,0 | 0,3 | 38,7 | 365 | 11,1 | 6,4 | 0,2 | 37,2 | 241 | 11,6 | 6,6 | 0,7 | 27,6 |
| *recprop* [18] | 2.807 | 318,2 | 606,5 | 0,9 | 18.989,0 | 1.576 | 339,9 | 383,9 | 3,4 | 4.116,7 | 1.324 | 514,0 | 516,0 | 9,1 | 4.403,8 | 993 | 610,2 | 421,7 | 26,4 | 2.502,7 |
| *armas* [19] | 3.619 | 38,2 | 31,8 | 0,0 | 100,0 | 1.926 | 45,5 | 25,7 | 0,0 | 100,0 | 1.653 | 53,1 | 22,1 | 0,0 | 100,0 | 1.188 | 61,6 | 18,5 | 0,0 | 100,0 |
| *prefgov* [20] | 3.508 | 0,2 | 0,4 | 0 | 1 | 1.869 | 0,2 | 0,4 | 0 | 1 | 1.617 | 0,2 | 0,4 | 0 | 1 | 1.181 | 0,2 | 0,4 | 0 | 1 |
| *prefcol* [21] | 3.508 | 0,3 | 0,4 | 0 | 1 | 1.869 | 0,3 | 0,4 | 0 | 1 | 1.617 | 0,3 | 0,5 | 0 | 1 | 1.181 | 0,3 | 0,5 | 0 | 1 |

*Fonte:* DATASUS (MSaúde), CNSEG, DENATRAN, IBGE, INEP, TSE e STN. Elaboração dos autores.

[1] Taxa de homicídios por 100 mil habitantes
[2] Taxa de roubos e furtos de veículos por 100 mil veículos emplacados
[3] Dummy = 1 se o Município possui guarda municipal, zero caso contrário
[4] Tempo de existência da guarda municipal (em anos)
[5] Efetivo das guardas municipais por 100 mil habitantes
[6] Proxy para o custo moral (% de pessoas que vivem em domicílios com aparelho de televisão)
[7] Densidade demográfica (pessoas por km$^2$)
[8] % da população vivendo em áreas urbanas (em %)
[9] % homens na faixa etária dos 15 aos 39 anos
[10] Dummy = 1 se o Município é capital de Estado
[11] Dummy = 1 se o Município faz parte de uma região metropolitana
[12] Dummy = 1 se o Município faz fronteira internacional
[13] % de 15 a 17 anos no ensino médio com 2 anos de atraso
[14] % mães chefes de famílias sem fundamental completo e com filhos menores de 15 anos
[15] % da população vivendo em domicílios com densidade maior que 2 pessoas por dormitório
[16] Índice de desigualdade de Gini da renda domiciliar per capita
[17] Taxa de abandono escolar no ensino médio (em %)
[18] Receita própria por habitante (exceto transferências)
[19] Proxy para a difusão de armas de fogo (% de suicídios e de homicídios com armas de fogo, defasado)
[20] Dummy = 1 se o Prefeito é do mesmo partido que o Governador do Estado, zero caso contrário
[21] Dummy = 1 se o Prefeito é da mesma coligação que elegeu o Presidente, zero caso contrário

população (15 a 39 anos), haja vista serem os que mais matam e os que mais morrem pela criminalidade. Em situações em que a técnica de efeitos fixos não foi indicada, utilizaram-se *dummies* para segregar os Municípios que: i) são capital de um dado Estado (*capital*), ii) fazem parte de uma região metropolitana (*rm*); e iii) contam com área de fronteira internacional (*fronteira*).

A literatura também indica possível relação entre condições sociais desfavoráveis e criminalidade. Quatro variáveis oriundas do IBGE foram ofertadas para captar essa dinâmica, sendo: o percentual de jovens de 15 a 17 anos que estão no ensino médio, mas que apresentam pelo menos 2 (dois) anos de atraso nas séries (*atraso*); uma medida paras lares uniparentais, focada no percentual de mães chefe de família sem ensino fundamental completo e com filhos menores de 15 anos (*nopai*); métrica para a tensão familiar oriunda da densidade domiciliar (*densdom*); e o consagrado índice de desigualdade de Gini da renda domiciliar *per capita*.[15]

Nas estimações com crimes contra o patrimônio mais duas variáveis foram incluídas, ambas relacionadas com as condições econômicas dos entes. A taxa de abandono escolar no ensino médio (*abandono*), compilada pelo INEP, foi aventada como uma medida das oportunidades no mercado legal que estão sendo perdidas por esses jovens, aproximando-os, pelo custo de oportunidade, das atividades ilícitas. Como uma quantificação do desenvolvimento econômico de cada Município, e das oportunidades ali disponíveis, foi utilizada a receita pública própria *per capita*, exceto transferências (*recprop*), proveniente das bases regionais do Tesouro Nacional (atualizada a preços constantes de 2012 pelo IPCA médio do IBGE).

Ainda sobre a escolha das demais variáveis explicativas, uma das grandes limitações que se impõe em modelagens para o comportamento da criminalidade em nível municipal está em como se controlar para os diversos tipos de efeitos dissuasórios reconhecidos pela teoria (por exemplo, tamanho do efetivo policial estadual, taxa de encarceramento, ações do Poder Judiciário, entre outros). Em especial porque no caso brasileiro a maior parcela dessa provisão está a critério dos Estados e do DF, os quais não disponibilizam esses dados regionalizados por Municípios. Nesse tipo de encruzilhada três possíveis vias podem ser tomadas: i) negligenciar parte desses possíveis efeitos e entender que as demais variáveis elencadas são suficientes para captar essa heterogeneidade; ii) interpretar a intensidade das atividades de cada Estado como diferentes entre si, mas fixas entre todos os seus Municípios, utilizando para isso apenas um conjunto de 26 (vinte e seis) *dummies*; ou ii) utilizar *proxies* ou instrumentos minimamente aceitáveis para captar tal dinâmica. Neste artigo optou-se por uma combinação entre os três enfoques.

Em primeiro lugar, cumpre destacar que de acordo com a Munic 2014 (IBGE, 2015), os principais critérios utilizados naquele ano pelas Secretarias Estaduais de Segurança Pública para a alocação municipal de seus efetivos da Polícia Militar foram a densidade demográfica (23 respondentes), a incidência de criminalidade (21 respostas), a existência de presídios (7 respostas); se em localidade de fronteira (7 respostas); se é polo industrial/comercial (6 respostas); e outros (11 respostas). Nesse sentido, é possível verificar que as variáveis de

[15]Esses indicadores, assim como o percentual de lares com televisores, existem em nível municipal apenas para os anos de Censo (1991, 2000 e 2010). Objetivando utilizar toda a série disponível nas variáveis dependentes e nas variáveis de interesse (1991 a 2012), optou-se por construir uma série anual completa para esses controles por meio de interpolação linear. Como são variáveis com comportamentos estruturalmente bem definidos (de queda ou de ascensão), entendeu-se que os ganhos em termos de graus de liberdade (e de eficiência nos estimadores) superam os riscos inerentes a tal imputação.

controle sugeridas já englobam boa parte desses critérios (densidade demográfica, cidades de fronteira e condições econômicas), o que fortalece todas as justificativas para o seu uso.

Logo após, foram utilizadas, sempre que possível, as *dummies* estaduais com vistas a capturar as heterogeneidades estaduais fixas no tempo. Assim, ainda que entre os Estados exista valoração diferente no que concerne aos seus efeitos repressivos, se sua política interna de atendimento municipal segue critérios fixos no tempo (digamos, tamanho do ente), a estratégia será efetiva para reconhecer tais idiossincrasias.

Como precaução adicional para mitigar a omissão de variáveis dissuasórias possivelmente relevantes, foi feito uso de uma medida defasada para a difusão de armas de fogo nas localidades, seguindo a ampla resenha literária de Cerqueira e Mello (2012) sobre o tema. Trata-se da proporção entre os óbitos perpetrados por armas de fogo (suicídios e homicídios) em relação ao total desses tipos de mortes (*armas*). A racionalidade por trás da opção sugere que Municípios que observaram num passado recente uma expansão em seu estoque de armas de fogo são também aqueles em que a atuação policial, responsabilidade dos Estados, foi menos presente, uma vez que tais indicadores estão associados a mercados mais ativos de criminalidade, bem como com maiores níveis de criminalidade.

Por fim, espera-se que as variáveis *dummy* referentes ao ciclo político apoiem subsidiariamente tal estratégia, minimizando possível viés de omissão de variáveis na medida em que captam diferenças entre as orientações políticas dos respectivos Prefeitos, bem como suas proximidades relativas com Governadores (*prefgov*) e Presidentes da República (*prefcol*). Em termos apenas de conjectura, se um daqueles critérios para alocação dos efetivos policiais dos Estados nos seus respectivos Municípios, classificados na Munic 2014 como "Outros", for a realização ou o desfazimento de uma aliança política, este fenômeno possivelmente estaria contemplado pelas variáveis dicotômicas sugeridas.

## 5. Resultados

Nesta seção são apresentadas as estimativas para o impacto nas taxas de criminalidade (homicídios e roubos e furtos de veículos) provenientes da criação de GMs nos Municípios brasileiros. Optou-se pela apresentação desses resultados em subseções distintas, cada uma direcionada a uma das estratégias de estimação escolhidas (tratamentos binários, multivalorados por categorias e contínuos).

### 5.1 Tratamentos Binários (DID, PSM e PSM-DID)

As estimativas para o método de diferenças-em-diferenças (DID) com efeitos fixos (*within*), tendo como variável dependente a taxa de homicídios por 100 mil habitantes, foram obtidas a partir do seguinte modelo, rodado em separado para cada um dos 4 (quatro) grupos populacionais de Municípios:

$$\begin{aligned}\ln txhom_{it} = {} & \beta_0 + \beta_1 gm_{it} + \beta_2 \ln armas_{it} + \beta_3 \ln cmoral_{it} + \beta_4 prefgov_{it} + \beta_5 prefcol_{it} + \beta_6 \ln dens_{it} \\ & + \beta_7 \ln txurbana_{it} + \beta_8 \ln txhomensj_{it} + \beta_9 \ln atraso_{it} + \beta_{10} \ln densdom_{it} + \beta_{11} \ln nopai_{it} \\ & + \beta_{12} \ln gini_{it} + \theta_2 d_{i2} + \cdots + \theta_\tau d_{i\tau} + \alpha_i + \varepsilon_{it},\end{aligned}$$

em que o subscrito $i$ ($i = 1, 2, \ldots, n$) refere-se a cada um dos entes das subamostras municipais e $t$ ($t = 1, 2, \ldots, 22$) denota a dimensão temporal do painel. Todas as variáveis foram previamente definidas na seção 4 (Tabela 2), com exceção de $\theta_2 d_{i2}, \cdots, \theta_\tau d_{i\tau}$, que

representam o conjunto de *dummies* anuais (sem o primeiro ano), e de $\alpha_i$ e de $\varepsilon_{it}$ que apontam, respectivamente, para os efeitos fixos e para o elemento tradicional do termo de erro. Por fim, a variável de interesse (estimador DID) é $\beta_1$, *dummy* que indica a existência de GM nas localidades em determinado ano.[16]

Isso posto, a Tabela 4 sintetiza os diversos coeficientes obtidos em três contextos distintos: OLS1 (sem controles, sem efeitos fixos e sem *dummies* de tempo), OLS2 (com controles e com *dummies* de tempo e de Estado, mas sem efeitos fixos)[17] e o DID completo, por efeitos fixos.[18]

Por meio do modelo OLS1 é possível notar que os grupos são de fato distintos no que diz respeito à variável de resultados, haja vista que em quase todas as especificações há significância estatística para o teste de médias, isto é, os Municípios do grupo de tratamento convivem com taxas de homicídio sistematicamente maiores do que os seus equivalentes do grupo de controle, reforçando a suspeita de que possa existir simultaneidade. Nos Municípios pequenos essa diferença é de cerca de 28%, ao passo que é bem menor ou inexistente nos entes mais populosos (acima de 250.000).

O modelo OLS2 (sem efeitos fixos, mas com controles e com *dummies* de tempo e de Estado) sugere cenário qualitativamente diferente na medida em que as novas estimativas para o impacto da existência de GMs na criminalidade foram em sua maioria estatisticamente iguais a zero ou mesmo benéficas (no Grupo 2, entes com GM há pelo menos 3 anos e acima). Pode-se interpretar que o conjunto de controles sugeridos possui o potencial de captar razoavelmente bem a heterogeneidade dos entes da amostra, promovendo adequado tratamento ao problema de simultaneidade. Ainda, a inclusão dos efeitos fixos com as variáveis de controle e com as *dummies* de tempo em um cenário de diferenças-em-diferenças (modelo DID) sugere ainda mais fortemente que a existência de GMs (mesmo defasadas) pouco interferiu na dinâmica da variável dependente, isto é, não parece existir impacto oriundo da decisão de se implantar tal programa, isolados os demais fatores.

Contudo, há sempre a possibilidade de que as hipóteses subjacentes ao método DID (em especial o paralelismo nas tendências) não estejam presentes na amostra analisada, o que acarretaria estimadores ainda viesados para a relação debatida. Como medida de prudência, operacionalizaram-se mais duas estratégias de identificação, sendo a primeira baseada no uso do pareamento por PSM e a segunda na combinação do DID com esse mesmo PSM (DID-PSM).

Para o PSM estimam-se modelos que buscam explicar a designação ao tratamento, isto é, não se tenciona mais explicar a variável de resultado diretamente (como no DID) e sim qual é o escore de propensão associado a cada ente no que diz respeito à probabilidade de possuir ou não uma GM, condicional às suas características observáveis (pré-intervenção). Entes

[16] Como anteriormente esclarecido, foram criadas diversas defasagens dessas *dummies* com o fito de se mitigar o problema de endogeneidade por simultaneidade, presente em atividades de segurança pública. Dessa forma, existem *dummies* indicando Municípios que possuem GM há pelo menos 1 ano (situação contemporânea), há pelo menos 2 anos (defasagem 1), há pelo menos 3 anos (defasagem 2) e assim sucessivamente.

[17] Além das variáveis citadas, essa especificação inclui mais quatro *dummies* invariantes no tempo, suprimidas automaticamente em efeitos fixos. São elas: i) indicação para Municípios que possuem GM; ii) entes com fronteira internacional; iii) que fazem parte de regiões metropolitanas; e iv) que são capital de Estado.

[18] Todos os modelos foram, em conjunto, estatisticamente significantes ao nível de 1% ou menos. Os diversos testes de Hausman indicaram as especificações com efeitos fixos em detrimento daquelas com efeitos aleatórios. As magnitudes individuais dos controles, coerentes em termos de sinais e significâncias, foram propositalmente omitidas por questões de fluidez no texto, mas podem ser requisitadas a qualquer tempo junto ao autor.

**Tabela 4.** Estimativas DID do Impacto da Existência de GMs (1991–2012). Variável dependente = log da Taxa de Homicídios.

| Possui GM há pelo menos $X$ anos | OLS1 | | | | OLS2 | | | | DID | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta$ | *P*-value | *C* | *T* | $\beta$ | *P*-value | *C* | *T* | $\beta$ | *P*-value | *C* | *T* |
| *Grupo 1 = População (20.000 a 49.999)* | | | | | | | | | | | | |
| gm1 | 0,271 | 0,000 | 17.994 | 3.627 | −0,006 | 0,906 | 14.344 | 3.316 | 0,089 | 0,157 | 14.344 | 3.316 |
| gm2 | 0,276 | 0,000 | 18.500 | 3.121 | −0,018 | 0,721 | 14.794 | 2.866 | 0,053 | 0,388 | 14.794 | 2.866 |
| gm3 | 0,287 | 0,000 | 18.766 | 2.855 | −0,005 | 0,925 | 15.030 | 2.630 | 0,074 | 0,229 | 15.030 | 2.630 |
| gm4 | 0,284 | 0,000 | 19.018 | 2.603 | −0,017 | 0,732 | 15.253 | 2.407 | 0,052 | 0,396 | 15.253 | 2.407 |
| gm5 | 0,287 | 0,000 | 19.255 | 2.366 | −0,012 | 0,808 | 15.463 | 2.197 | 0,060 | 0,351 | 15.463 | 2.197 |
| gm10 | 0,351 | 0,000 | 20.231 | 1.390 | 0,024 | 0,687 | 16.352 | 1.308 | 0,143 | 0,074 | 16.352 | 1.308 |
| gm15 | 0,283 | 0,000 | 20.829 | 792 | −0,069 | 0,360 | 16.944 | 716 | −0,071 | 0,414 | 16.944 | 716 |
| gm20 | 0,229 | 0,000 | 21.213 | 408 | −0,122 | 0,242 | 17.320 | 340 | −0,176 | 0,108 | 17.320 | 340 |
| *Grupo 2 = População (50.000 a 99.999)* | | | | | | | | | | | | |
| gm1 | 0,241 | 0,000 | 4.725 | 1.926 | −0,053 | 0,404 | 4.174 | 1.861 | 0,030 | 0,700 | 4.174 | 1.861 |
| gm2 | 0,231 | 0,000 | 4.954 | 1.697 | −0,081 | 0,183 | 4.388 | 1.647 | 0,046 | 0,518 | 4.388 | 1.647 |
| gm3 | 0,218 | 0,000 | 5.073 | 1.578 | −0,101 | 0,094 | 4.501 | 1.534 | 0,034 | 0,616 | 4.501 | 1.534 |
| gm4 | 0,211 | 0,000 | 5.193 | 1.458 | −0,110 | 0,065 | 4.614 | 1.421 | 0,040 | 0,540 | 4.614 | 1.421 |
| gm5 | 0,193 | 0,000 | 5.306 | 1.345 | −0,131 | 0,028 | 4.722 | 1.313 | 0,021 | 0,751 | 4.722 | 1.313 |
| gm10 | 0,155 | 0,001 | 5.806 | 845 | −0,199 | 0,002 | 5.200 | 835 | −0,108 | 0,182 | 5.200 | 835 |
| gm15 | 0,070 | 0,223 | 6.129 | 522 | −0,227 | 0,005 | 5.523 | 512 | −0,124 | 0,195 | 5.523 | 512 |
| gm20 | −0,124 | 0,092 | 6.346 | 305 | −0,264 | 0,015 | 5.740 | 295 | −0,298 | 0,045 | 5.740 | 295 |
| *Grupo 3 = População (100.000 a 249.999)* | | | | | | | | | | | | |
| gm1 | 0,238 | 0,000 | 1.592 | 1.653 | −0,029 | 0,716 | 1.442 | 1.614 | 0,130 | 0,113 | 1.442 | 1.614 |
| gm2 | 0,224 | 0,000 | 1.746 | 1.499 | −0,008 | 0,916 | 1.584 | 1.472 | 0,099 | 0,202 | 1.584 | 1.472 |
| gm3 | 0,222 | 0,000 | 1.826 | 1.419 | −0,021 | 0,790 | 1.657 | 1.399 | 0,090 | 0,271 | 1.657 | 1.399 |
| gm4 | 0,213 | 0,000 | 1.908 | 1.337 | −0,035 | 0,661 | 1.735 | 1.321 | 0,087 | 0,291 | 1.735 | 1.321 |
| gm5 | 0,204 | 0,000 | 1.984 | 1.261 | −0,066 | 0,399 | 1.806 | 1.250 | −0,007 | 0,931 | 1.806 | 1.250 |
| gm10 | 0,166 | 0,000 | 2.359 | 886 | −0,116 | 0,120 | 2.174 | 882 | −0,204 | 0,019 | 2.174 | 882 |
| gm15 | 0,213 | 0,000 | 2.643 | 602 | −0,048 | 0,548 | 2.458 | 598 | −0,187 | 0,083 | 2.458 | 598 |
| gm20 | 0,296 | 0,000 | 2.844 | 401 | 0,046 | 0,568 | 2.659 | 397 | −0,085 | 0,314 | 2.659 | 397 |
| *Grupo 4 = População (Acima de 250.000)* | | | | | | | | | | | | |
| gm1 | 0,202 | 0,000 | 668 | 1.188 | 0,096 | 0,320 | 635 | 1.181 | 0,097 | 0,373 | 635 | 1.181 |
| gm2 | 0,142 | 0,001 | 765 | 1.091 | −0,002 | 0,981 | 730 | 1.086 | 0,002 | 0,987 | 730 | 1.086 |
| gm3 | 0,117 | 0,006 | 819 | 1.037 | −0,055 | 0,523 | 783 | 1.033 | −0,054 | 0,532 | 783 | 1.033 |
| gm4 | 0,094 | 0,028 | 873 | 983 | −0,083 | 0,319 | 837 | 979 | −0,079 | 0,335 | 837 | 979 |
| gm5 | 0,080 | 0,062 | 926 | 930 | −0,106 | 0,199 | 890 | 926 | −0,090 | 0,250 | 890 | 926 |
| gm10 | 0,078 | 0,079 | 1.181 | 675 | −0,120 | 0,130 | 1.145 | 671 | −0,061 | 0,412 | 1.145 | 671 |
| gm15 | 0,066 | 0,180 | 1.396 | 460 | −0,085 | 0,308 | 1.360 | 456 | 0,034 | 0,680 | 1.360 | 456 |
| gm20 | 0,002 | 0,975 | 1.556 | 300 | −0,085 | 0,343 | 1.520 | 296 | 0,076 | 0,452 | 1.520 | 296 |
| Dummies de Tempo | | Não | | | | Sim | | | | Sim | | |
| Efeitos Fixos | | Não | | | | Não | | | | Sim | | |
| Controles | | Não | | | | Sim | | | | Sim | | |

*Notas:* Erros-Padrão robustos à heteroscedasticidade e à correlação serial. $C$ = *untreated* (grupo controle) e $T$ = *treated* (grupo de tratamento).
*Fonte:* Elaboração dos autores com base nas estimativas fornecidas pelo programa Stata 13.

não tratados, mas com escores de propensão suficientemente parecidos com os dos tratados, são ditos bons contrafactuais e têm suas médias da variável de resultado comparadas para o cálculo do ATT. Portanto, para cada um dos quatro grupos populacionais estima-se, por meio de um modelo Probit, a seguinte relação:

$$gm_{it} = \beta_0 + \beta_1 txhom_{it} + \beta_2 armas_{it} + \beta_3 cmoral_{it} + \beta_4 prefgov_{it} + \beta_5 prefcol_{it} + \beta_6 dens_{it} + \beta_7 txurbana_{it} + \beta_8 txhomensj_{it} + \beta_9 atraso_{it} + \beta_{10} densdom_{it} + \beta_{11} nopai_{it} + \beta_{12} gini_{it} + \beta_{13} rm_i + \beta_{14} fronteira_i + \beta_{15} capital_i + \varepsilon_{it},$$

em que as variáveis utilizadas são, em essência, as mesmas já apresentadas, mas sem a transformação logarítmica e defasadas em pelo menos um período para captar heterogeneidades pré-intervenção.[19] Os resultados pertinentes estão sintetizados na Tabela 5.

Novamente, por questões de fluidez na apresentação foram omitidos os resultados da regressão por *probit*, bem como os testes de balanceamento realizados em cada uma das especificações propostas. Contudo, vale ressaltar que em todos os modelos abordados houve significância conjunta para as variáveis utilizadas (sempre a menos de 1%) e a condição de balanceamento foi amplamente satisfeita, seja nos critérios individuais (variável a variável) e, especialmente, no quesito da significância conjunta. Dessa forma, os testes de diferenças de médias sugerem, em termos estatísticos, que os grupos de controle e de tratamento, antes sensivelmente diferentes, se tornaram suficientemente similares após o pareamento, o que legitima os resultados potenciais estimados.[20]

Ainda sobre as estimações realizadas, tendo em vista o tamanho satisfatório da base de dados, foi possível fazer uso de opções mais restritivas em relação à formação do grupo de controle, com destaque para a utilização do critério de suporte comum e a imposição adicional de descarte (*trimming*) em 10% das observações do grupo de tratamento que coincidiram com as menores densidades do escore de propensão dos controles. Para o algoritmo de NNM utilizaram-se 3 (três) vizinhos, com reposição, e *caliper* de 0,01. No caso do PSM por *kernel*, a opção repousou sobre o tipo Epanechnikov com banda de 0,06.

Os resultados apresentados sumarizam, em termos qualitativos, comportamentos similares em todos os modelos apresentados, remetendo à inexistência de impactos provenientes da implantação de GMs nas taxas de homicídios dos diversos grupos populacionais analisados. A presença de simultaneidade também parece contornada, haja vista que os sinais e as significâncias permanecem basicamente os mesmos quando defasamos a variável de tratamento. Dito em outras palavras, controlando-se por estratégias de resultados potenciais que mitigam possíveis endogeneidades, as diferenças nas médias de homicídios entre aqueles entes que possuem GM e aqueles que não as implantaram não é, em termos estatísticos, diferente de zero.[21]

É possível conjecturar sobre o porquê desse comportamento, em particular uma das razões pode ser a já citada amplitude de atribuições com as quais essas corporações se

[19] A taxa de homicídios e a *proxy* para a difusão de armas de fogo (e para a atuação da polícia estadual) possuem duas defasagens por interpretação de que as políticas públicas de segurança acontecem sempre reativamente e sujeitas a uma institucionalidade que demanda tempo adicional para ser atualizada (por exemplo, debates legislativos, prazos para a realização de concursos e para treinamento do efetivo, atraso na publicação de estatísticas de interesse etc.).

[20] Os resultados detalhados podem ser requisitados a qualquer tempo junto aos autores.

[21] Existem algumas estimativas PSM (NNM e *kernel*) que foram significativas, entretanto não conseguiram demonstrar um padrão de estabilidade. Ademais, quando testadas em conjunto com o DID, todas perderam a significância.

**Tabela 5.** Estimativas do Impacto da Existência de GMs (1991–2012). Variável dependente = Taxa de Homicídios (PSM) e a sua Diferença (PSM-DID).

| Possui GM há pelo menos $X$ anos | PSM 1 (NNM) | | PSM 2 (kernel) | | PSM 1 - DID | | PSM 2 - DID | | Observações | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | att | *P*-value | att | *P*-value | att | *P*-value | att | *P*-value | *n* | *C* | *T* |
| *Grupo 1 = População (20.000 a 49.999)* | | | | | | | | | | | |
| gm1 | 0,492 | 0,239 | 0,572 | 0,059 | 0,301 | 0,279 | 0,242 | 0,253 | 16.123 | 13.427 | 2.696 |
| gm2 | 0,667 | 0,157 | 0,700 | 0,037 | 0,160 | 0,370 | 0,243 | 0,273 | 16.165 | 13.849 | 2.316 |
| gm5 | 1,055 | 0,075 | 1,106 | 0,003 | 0,391 | 0,267 | 0,386 | 0,197 | 16.225 | 14.448 | 1.777 |
| gm10 | 0,631 | 0,260 | 0,853 | 0,073 | −0,122 | 0,390 | 0,131 | 0,375 | 16.304 | 15.235 | 1.069 |
| gm15 | 0,013 | 0,399 | 0,123 | 0,389 | −0,110 | 0,393 | −0,106 | 0,389 | 16.357 | 15.765 | 592 |
| gm20 | −0,365 | 0,375 | −0,158 | 0,389 | −0,711 | 0,286 | 0,082 | 0,395 | 16.390 | 16.093 | 297 |
| *Grupo 2 = População (50.000 a 99.999)* | | | | | | | | | | | |
| gm1 | −0,809 | 0,215 | −0,381 | 0,275 | −0,259 | 0,374 | −0,095 | 0,384 | 5.531 | 3.950 | 1.581 |
| gm2 | −0,848 | 0,213 | −0,622 | 0,179 | −0,431 | 0,259 | −0,285 | 0,273 | 5.551 | 4.153 | 1.398 |
| gm5 | −1,013 | 0,177 | −1,032 | 0,035 | −0,122 | 0,389 | −0,283 | 0,297 | 5.583 | 4.467 | 1.116 |
| gm10 | −1,832 | 0,065 | −1,246 | 0,032 | −0,466 | 0,311 | −0,414 | 0,252 | 5.627 | 4.911 | 716 |
| gm15 | −0,859 | 0,299 | −0,940 | 0,133 | −0,453 | 0,329 | −0,263 | 0,353 | 5.538 | 5.099 | 439 |
| gm20 | −1,214 | 0,247 | −2,080 | 0,015 | 0,065 | 0,398 | 0,039 | 0,398 | 5.558 | 5.306 | 252 |
| *Grupo 3 = População (100.000 a 249.999)* | | | | | | | | | | | |
| gm1 | −0,313 | 0,384 | −0,325 | 0,353 | 0,282 | 0,364 | 0,198 | 0,369 | 2.737 | 1.358 | 1.379 |
| gm2 | −0,758 | 0,322 | −0,775 | 0,213 | −0,232 | 0,371 | −0,076 | 0,395 | 2.751 | 1.495 | 1.256 |
| gm5 | 0,039 | 0,399 | −0,445 | 0,311 | 0,471 | 0,299 | 0,181 | 0,375 | 2.772 | 1.706 | 1.066 |
| gm10 | 0,021 | 0,399 | −0,533 | 0,303 | 0,486 | 0,307 | 0,341 | 0,322 | 2.807 | 2.059 | 748 |
| gm15 | 1,217 | 0,229 | 0,593 | 0,303 | 1,206 | 0,279 | 0,746 | 0,156 | 2.781 | 2.273 | 508 |
| gm20 | −0,753 | 0,361 | 0,396 | 0,363 | −0,857 | 0,272 | 0,103 | 0,394 | 2.800 | 2.460 | 340 |
| *Grupo 4 = População (Acima de 250.000)* | | | | | | | | | | | |
| gm1 | −1,655 | 0,274 | −2,326 | 0,069 | −0,589 | 0,311 | −1,009 | 0,147 | 1.619 | 602 | 1.017 |
| gm2 | −1,997 | 0,190 | −2,295 | 0,061 | −0,507 | 0,317 | −0,898 | 0,181 | 1.628 | 691 | 937 |
| gm5 | −3,406 | 0,035 | −2,950 | 0,003 | −0,837 | 0,295 | −0,409 | 0,316 | 1.732 | 843 | 889 |
| gm10 | −1,280 | 0,278 | −1,035 | 0,194 | 0,364 | 0,357 | 0,079 | 0,395 | 1.668 | 1.089 | 579 |
| gm15 | −0,178 | 0,396 | −0,901 | 0,230 | −0,473 | 0,354 | −0,059 | 0,397 | 1.689 | 1.296 | 393 |
| gm20 | −3,172 | 0,101 | −1,818 | 0,089 | 0,454 | 0,357 | −0,193 | 0,385 | 1.704 | 1.449 | 255 |

*Notas:* Intervalos de confiança obtidos por *bootstrap* (200 replicações). $C$ = *untreated* (grupo controle) e $T$ = *treated* (grupo de tratamento).
*Fonte:* Elaboração dos autores com base nas estimativas fornecidas pelo programa Stata 13.

defrontam. Nesse sentido, pode acontecer que na prática estejam mesmo mais próximas de serviços de portaria e/ou vigilância (interpretação literal da CF, 1988) do que de potenciais repressoras de ilícitos (compreensão mais focada no espírito da lei). Outra sinalização pode vir do fato de que, ainda que tentem de fato executar serviços de segurança pública, suas estratégias e os seus recursos não sejam os mais adequados para tal tarefa e quando se controlam para o ambiente em que operam e para o papel de outras instâncias (por exemplo, os Estados), sua contribuição se torna negligenciável. Há ainda a possibilidade de que a natureza de atuação das GMs, ainda que realizadora de serviços de segurança pública, tenha poucas condições de produzir resultados devido à própria natureza dos crimes de homicídios, motivados por fatores múltiplos e que não se restringem apenas aos materiais.

Sobre essa última suposição, a base de dados explorada permite que exercício similar aos realizados até agora seja feito tendo a taxa de roubos e furtos de veículos por 100 mil emplacamentos como variável de resultado ao invés das taxas de homicídios. Nesse caso,

ocorre a utilização de uma *proxy* aceitável para os crimes contra o patrimônio, que em geral tendem a ser determinados por condicionantes não exatamente iguais àquelas relacionadas ao fenômeno do homicídio. Os modelos especificados seguem a mesma estrutura daqueles utilizados para a taxa de homicídios, mas contam com duas variáveis explicativas adicionais: a taxa de abandono escolar no ensino médio (*abandono*) e a receita corrente própria *per capita* dos Municípios (*recprop*), destinadas a captar as possibilidades de renda no mercado formal e a situação financeira da localidade, respectivamente.

As estimativas OLS e DID, constantes da Tabela A-1 do Apêndice, relatam cenário análogo àquele relacionado aos crimes contra a pessoa. Em primeiro lugar, sem os controles as taxas de roubos e furtos de veículos chegam a ser até 60% maiores nos Municípios com GM. Quando se incluem as covariadas, esse indício de simultaneidade é dissipado e, sob DID, pode-se depreender que não há evidência robusta em favor da hipótese de redução nos crimes contra o patrimônio advinda da existência de GMs em nenhum dos grupos populacionais.[22]

Contudo, quando se observam os resultados dos modelos PSM e PSM-DID aplicados às taxas de roubos e furtos de veículos (Tabela A-2 do Apêndice), é possível notar um quadro ligeiramente diferente do comportamento exposto para a taxa de homicídios.[23] Isso ocorre porque o PSM aparentemente não foi capaz de, mesmo selecionando em observáveis coerentemente balanceados, reduzir completamente as diferenças entre os grupos, haja vista a prevalência de estimativas sugestivas de endogeneidade (impactos positivos e significantes). Todavia, a sua utilização em conjunto com o arcabouço de diferenças-em-diferenças (PSM-DID) revelou impactos causais mais confiáveis, revertendo sensivelmente o padrão de simultaneidade e apresentando, em sua maioria, estimativas não significativas em termos estatísticos, inclusive com o sinal negativo esperado pela teoria em alguns casos (grupos 1 e 3, principalmente).

Em síntese, além das conjecturas já expostas para a possível falta de impacto das GMs na criminalidade, os resultados obtidos sob condições de tratamento binárias, ainda que meritórios e instigadores de novas considerações, nada informam sobre possíveis diferenças na dosagem dessas iniciativas, isto é, ainda podem existir impactos diferenciados nos resultados advindos de níveis desiguais de exposição ao programa. Seguindo essa indicação, as próximas duas subseções tratam dessas possibilidades por meio de estratégias de identificação distintas.

## 5.2 Tratamentos Multivalorados (Categóricos)

Uma alternativa às limitações impostas pela abordagem binária está na interpretação do tratamento como se administrado em níveis ou em múltiplos valores (ou ainda, em categorias). Sob essa ótica, o grupo de controle continuaria sendo composto por aqueles que não receberam dosagem alguma, contudo os membros do grupo de tratamento estariam agora sujeitos a múltiplas medidas de intensidade para referenciar a sua participação no

[22] Uma vez que a base para crimes contra o patrimônio conta com apenas três anos (2009–2011), optou-se por fundir os dois últimos grupos populacionais. Assim, para a situação em que a variável dependente é a taxa de roubos e furtos de veículos, temos agrupados, no último estrato, os Municípios acima de 100.000 habitantes.

[23] Todos os testes de balanceamento foram realizados e mostraram-se coerentes, indicando que os grupos de controle e de tratamento se tornaram, após o pareamento, estatisticamente similares. Por conta da amostra reduzida, utilizou-se, juntamente com a opção de suporte comum, a opção *trimming* a somente 2,5%. As demais configurações permaneceram as mesmas do exercício com a taxa de homicídios.

programa. Neste artigo propomos como medida para a dosagem do tratamento oriundas da existência de GMs o número de anos em que o Município contou com o funcionamento dessas corporações. Se o programa for de fato efetivo, registrar-se-á uma relação benéfica (redução da criminalidade) em função do aumento da intensidade nessas dosagens.

Os múltiplos valores para o tratamento foram estabelecidos como segue:

| | |
|---|---|
| Dose = 0 | Municípios sem GM; |
| Dose = A | Municípios em que a GM possui entre 1 e 3 anos de funcionamento; |
| Dose = B | Idem, mas entre 4 e 6 anos; |
| Dose = C | Idem, mas entre 7 e 10 anos; |
| Dose = D | Idem, mas acima de 10 anos. |

Conforme antecipado na seção 3, as estratégias de estimação envolvem os modelos para determinação das variáveis de resultado (taxas de criminalidade) por meio de ajustamento por regressão (RA) e, no tocante à determinação do escore de propensão ao tratamento, pelo uso das técnicas de IPW (pesos pelo inverso da probabilidade). As ferramentas também são combinadas (IPWRA) com o intuito de fornecer estimadores duplamente robustos, conforme aspectos detalhados em Hirano e Imbens (2001) e em Wooldridge (2010, p.930–934). A equação estimada em RA (especificações *log-log* e Poisson) é idêntica ao procedimento estipulado para DID, ao passo que IPW é obtido com as mesmas variáveis utilizadas no PSM, mas aplicando-se um modelo *multinomial logit* por conta dos diferentes níveis de dosagem que agora compõem a variável dependente (tempo com GM).[24]

Isso posto, a Tabela 6 sumariza as estimativas e as inferências levando em consideração os múltiplos níveis de tratamento para os anos de existência das GMs e os seus possíveis impactos sobre as taxas de homicídios por 100 mil habitantes.

Em primeiro lugar, cumpre ressaltar que as condições de sobreposição foram devidamente checadas por meio de inspeções gráficas nas densidades de cada um dos níveis de tratamento e em cada uma das especificações aventadas e demonstraram-se fortemente satisfatórias, sugerindo boa qualidade de ajustamento nos potenciais contrafactuais. Os efeitos individuais e as significâncias conjuntas de cada uma das equações estimadas também se mostraram adequados, mas foram omitidos para não tornar o texto enfadonho.[25]

Logo após, tem-se que as estimativas apresentadas podem ser consideradas coerentes em termos de manutenção de sinais, magnitudes e significâncias, sugerindo que os problemas de endogeneidade por simultaneidade e a autosseleção foram adequadamente contornados pelas técnicas abordadas. Tomando como base os modelos mais completos (IPWRA), é possível verificar que as GMs acarretaram impactos positivos sobre a criminalidade, isto é, redutores da taxa de homicídios, em todas as faixas populacionais (em pelo menos um nível da dosagem).

No tocante aos grupos populacionais, a existência de GMs parece produzir proporcionalmente mais impacto nos Municípios menores (grupos 1 e 2). Nesses entes, os tratados apresentaram reduções nas taxas de homicídios que variaram de 11,1% a 30,5% em comparação com o grupo de controle de não tratados (IPWRA *log-log*). O modelo Poisson correspondente estima a mesma relação, mas em termos de pontos na taxa de homicídios,

[24] Todas as rotinas foram realizadas no âmbito do programa STATA 13 e por meio do comando `teffects`, opções `ra`, `ipw`, `ipwra` e `overlap`.

[25] Cada especificação gera 5 modelos (um para cada nível de tratamento) em cada grupo populacional, sendo que esses valores dobram em IPWRA. Todos esses resultados podem ser requisitados aos autores.

**Tabela 6.** Estimativas para o Impacto Multivalorado das GMs (1991–2012). Variável dependente = Taxa de Homicídios.

| Níveis de Tratamento | RA | | | | | IPW | | | IPWRA | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | *log-log* | | *Poisson* | | | | | | *log-log* | | *Poisson* | | |
| | att | *P*-value | att | *P*-value | *N* | att | *P*-value | *N* | att | *P*-value | att | *P*-value | *N* |
| *Grupo 1 = População (20.000 a 49.999)* | | | | | | | | | | | | | |
| (A vs 0) | −0,051 | 0,316 | −0,269 | 0,674 | 17.695 | 0,176 | 0,755 | 16.422 | −0,111 | 0,032 | −1,234 | 0,041 | 15.740 |
| (B vs 0) | −0,218 | 0,001 | −1,217 | 0,094 | | 0,855 | 0,227 | | −0,236 | 0,000 | −1,753 | 0,011 | |
| (C vs 0) | −0,167 | 0,015 | −1,703 | 0,013 | | 1,481 | 0,046 | | −0,246 | 0,001 | −2,498 | 0,000 | |
| (D vs 0) | −0,217 | 0,001 | −1,792 | 0,003 | | −0,034 | 0,963 | | −0,305 | 0,000 | −2,895 | 0,000 | |
| *Grupo 2 = População (50.000 a 99.999)* | | | | | | | | | | | | | |
| (A vs 0) | −0,025 | 0,692 | −0,221 | 0,819 | 6.018 | 0,716 | 0,379 | 5.706 | −0,029 | 0,610 | −0,696 | 0,420 | 5.673 |
| (B vs 0) | −0,112 | 0,159 | −1,466 | 0,162 | | 1,134 | 0,227 | | −0,136 | 0,090 | −1,935 | 0,059 | |
| (C vs 0) | −0,214 | 0,006 | −3,715 | 0,000 | | −0,585 | 0,490 | | −0,279 | 0,000 | −4,847 | 0,000 | |
| (D vs 0) | −0,266 | 0,001 | −4,036 | 0,000 | | −0,718 | 0,542 | | −0,240 | 0,005 | −3,325 | 0,003 | |
| *Grupo 3 = População (100.000 a 249.999)* | | | | | | | | | | | | | |
| (A vs 0) | 0,100 | 0,157 | 3,473 | 0,015 | 3.054 | −0,414 | 0,713 | 2.890 | 0,019 | 0,759 | 0,272 | 0,820 | 2.885 |
| (B vs 0) | 0,071 | 0,299 | 2,265 | 0,133 | | −0,647 | 0,600 | | −0,028 | 0,686 | −0,150 | 0,914 | |
| (C vs 0) | 0,040 | 0,555 | 0,217 | 0,873 | | −1,048 | 0,396 | | −0,106 | 0,154 | −3,332 | 0,013 | |
| (D vs 0) | 0,194 | 0,001 | 6,293 | 0,000 | | 0,350 | 0,791 | | 0,015 | 0,815 | 1,509 | 0,264 | |
| *Grupo 4 = População (Acima de 250.000)* | | | | | | | | | | | | | |
| (A vs 0) | 0,028 | 0,644 | −1,742 | 0,396 | 1.816 | −3,138 | 0,052 | 1.732 | −0,004 | 0,933 | −1,435 | 0,388 | 1.731 |
| (B vs 0) | 0,000 | 0,994 | −2,847 | 0,201 | | −2,589 | 0,313 | | −0,068 | 0,281 | −2,469 | 0,241 | |
| (C vs 0) | −0,202 | 0,003 | −9,687 | 0,000 | | 3,212 | 0,723 | | −0,181 | 0,012 | −7,967 | 0,000 | |
| (D vs 0) | 0,156 | 0,023 | −5,861 | 0,007 | | −3,890 | 0,079 | | −0,163 | 0,021 | −3,066 | 0,189 | |

*Notas:* Erros padrão robustos. Nível 0 = sem GM; A = GM de 1 a 3 anos; B = GM de 4 a 6 anos; C = GM de 7 a 10 anos; D = GM há mais de 10 anos.
*Fonte:* Elaboração dos autores com base nas estimativas fornecidas pelo programa Stata 13.

e nos diz que possuir uma GM, nos grupos comentados e com determinada faixa etária, pode reduzir esse indicador em até 4,8 mortes por 100 mil habitantes, se comparado com a situação de inexistência desses aparatos. Para as cidades entre 20.000 e 49.999 habitantes, o impacto em relação ao grupo de controle se dá em todos os múltiplos valores do tratamento, inclusive no primeiro nível (GM de 1 a 3 anos), segmento em que se imagina que tais corporações ainda estejam se estruturando como resposta às demandas da sociedade local e que, portanto, podem apresentar resultado nulo ou até mesmo alguma simultaneidade. De fato, é o que acontece no grupo de Municípios imediatamente superior (população de 50.000 a 99.999), em que os impactos das GMs marginalmente se fazem sentir apenas a partir do segundo nível do tratamento (a partir de 4 anos de existência).

Uma explicação possível para esse fenômeno é que em localidades menores a disseminação de uma dada política pública pode ser mais intensa em função das facilidades em se mobilizar as comunidades envolvidas e também dos controles sociais mais rígidos. Em termos de um modelo Agente-Principal, pode ser dito que em tais cidades problemas de ação coletiva podem ser mais facilmente contornados. Outra conjectura, baseada no enquadramento proposto por Vargas e Oliveira (2010) e nos argumentos de Beato Filho (2002), é que nessas cidades pode de fato existir, pelas condições expostas, um processo mais

acentuado de policiamento comunitário via GMs, prática amplamente reconhecida como indutora de maiores níveis de segurança e mudanças de comportamentos.[26]

Por outro lado, o comportamento nos demais grupos de Municípios não é, aparentemente, tão homogêneo. Para os componentes do Grupo 3 (população entre 100.000 e 249.999) não parece haver ganhos advindos de tal estratégia, com exceção da dosagem "C" (GM de 7 a 10 anos) do modelo IPWRA Poisson. Contudo, esse comportamento não é reprisado em mais nenhum outro modelo dessa mesma dosagem de tratamento (RA, IPW e IPWRA log-log), o que pode sugerir se tratar de um resultado espúrio associado unicamente à modelagem citada.

Nas grandes cidades (mais de 250.000 habitantes), as estimativas IPWRA log-log são de uma taxa de homicídios cerca de 17% menor (categorias "C" e "D") do que a dos Municípios com dosagem do programa nula. Uma suposição inicial seria de que nesses entes a dinâmica social e econômica seria mais complexa e, portanto, a curva de aprendizado é mais acentuada, o que levaria essas corporações a demorarem um pouco mais de tempo para apresentarem resultados. Além disso, o argumento de capacidade de coordenação exposto para os Municípios pequenos pode atuar aqui, mas de maneira contrária. Em outra abordagem, há que se ponderar que foram nesses entes, em geral com melhor estrutura, que o processo de instalação de GMs foi, em termos relativos, iniciado e, por isso, contam atualmente com a maior parte de suas corporações nas categorias "C" ou "D".

Portanto, diferentemente do que fora observado junto aos resultados dos tratamentos binários, a existência de GMs parece de fato exibir o potencial de contribuir para a redução da taxa de homicídios nas localidades que decidiram por implantá-las, possuindo uma resposta à dosagem proporcionalmente maior nos Municípios menores (de 20.000 a 49.999 habitantes e de 50.000 a 99.999).

Por fim, há que se comentar os resultados da metodologia proposta em relação à base de dados de crimes contra o patrimônio. Novamente, a dimensão reduzida desse painel (2009–2011) demandou adaptações nos grupos populacionais, sendo a separação neste caso refeita em termos de apenas dois segmentos: i) localidades com população maior do que 20.000 e menor do que 99.999; e ii) cidades com mais de 100.000 habitantes. Além disso, como já informado, só foi possível combinar as taxas de roubos e furtos de veículos como variável de resultado para a dosagem baseada nos anos de existência das GMs. Sob o mesmo conjunto de variáveis utilizado no DID e no PSM, tem-se os resultados da Tabela A-3 do Apêndice.

Conforme disposto, os três modelos são qualitativamente similares informando que não há efeitos significativos da existência de GMs, e das suas diversas dosagens, na redução da criminalidade relacionada com crimes contra o patrimônio, ao contrário do que foi relatado para as taxas de homicídios (em algumas faixas populacionais). Aventa-se a possibilidade de que dada a natureza operacional das GMs possa existir um efeito diferenciado em relação aos crimes contra a pessoa e aos crimes contra o patrimônio. Uma explicação factível pode estar no fato de que os roubos e furtos de veículos estão relacionados primordialmente com atividades de crime organizado e que as atuações mais circunscritas das GMs não os atinjam. Se isso for verdade, tal variável provavelmente não seria uma boa *proxy* para os crimes contra

[26] Poder-se-ia argumentar também que nessas localidades há maior ausência de policiamento estadual, criminalidade de menor potencial ofensivo e estruturas urbanas menos complexas, contudo tais fatores já foram, de uma maneira ou outra, controlados nas regressões propostas pela inclusão de *proxies* específicas, o que, desejavelmente, relegaria a causa dos efeitos verificados somente ao tratamento.

o patrimônio sob os quais essas corporações poderiam exercer alguma melhoria, cogitando a sua reanálise para trabalhos futuros que sigam essa mesma abordagem.

Além disso, por ser um crime altamente especializado, é provável que os fatores para a sua redução passem por atuações investigativas mais incisivas a cargo das respectivas Polícias Civis e que em nossos modelos já estariam devidamente controlados. Outro aspecto relacionado com esse indicador criminal é o seu controle por meio da inclusão tecnológica (rastreadores, bloqueadores etc.). Se esses fatores alheios às atuações das GMs forem de fato de alguma relevância para o comportamento desses indicadores, ter-se-á que já estariam provavelmente incluídos no modelo pelos controles propostos (em especial, as *dummies* temporais). Por fim, deve-se destacar que a base de dados utilizada para esse fim, em que pese se configurar numa contribuição ao debate, pode ainda ser suficientemente limitada para inferências focadas em subpopulações e subdivisões da amostra, como foi o caso aqui apresentado.

Na subseção seguinte tenciona-se explorar um pouco mais desse aspecto multivalorado dos tratamentos, mas agora sob um enfoque contínuo, possibilitando a aplicação de testes adicionais às inferências aqui debatidas.

## 5.3 Tratamentos Contínuos (GPS e Dose-Resposta)

Os resultados obtidos por meio das estimações do GPS e da função Dose-Resposta, doravante denominada apenas DR, seguem os procedimentos operacionais estabelecidos por Bia e Mattei (2008), com os ajustes sugeridos por Guardabascio e Ventura (2013).[27] Consistem, como referenciado, em três etapas:

1. As estimativas do GPS, $r(t, X)$, que fizeram uso do mesmo conjunto de variáveis explicativas do PSM e dos modelos IPW;[28]

2. Logo após, computa-se o resultado esperado de um indivíduo sob o nível de tratamento $t$, dado o GPS, $\beta(t, r)$, com formas funcionais flexíveis (quadrática para ambos os termos, com interação); e

3. Tomando-se a média dessas expectativas condicionais estimadas, obtém-se a função DR.

Entre o primeiro e o segundo passo é feita a checagem da condição de balanceamento. Contudo, a rotina computacional implementada não o faz por meio de testes com viés padronizado usuais (tanto para variáveis individuais quanto para a significância conjunta). De acordo com o exposto por Guardabascio e Ventura (2013, p.9)), os critérios de inferência desse comando (e do seu antecessor `doseresponse` e `gpscore`) baseiam-se apenas no valor mais extremo das diversas estatísticas $t$ obtidas, o que pode levar a rejeições precipitadas. Nas estimações propostas, por exemplo, foram analisados os testes de média de até 56 parâmetros estimados em cada especificação, sendo que, em média, em apenas 4 situações as estatísticas $t$ superaram o valor consagrado de 1,96 (significância a 5%, bicaudal), limitando-se a um

[27] Comando `doseresponse2` do STATA 13.

[28] Guardabascio e Ventura (2013) flexibilizam o pressuposto de que o tratamento deva ser normalmente distribuído, condicional às covariadas (Bia & Mattei, 2008), incluindo nas possibilidades de estimação distribuições da família exponencial. Nos resultados apresentados optou-se pela distribuição *gamma* para o tratamento, após a realização de testes que rejeitaram o uso da distribuição normal (Kolmogorov-Smirnov).

máximo de 2,45. Na ausência de uma medida de significância conjunta para o balanceamento e em face das situações relatadas, assumiu-se que essa condição foi satisfatoriamente obtida para a base de dados em questão.

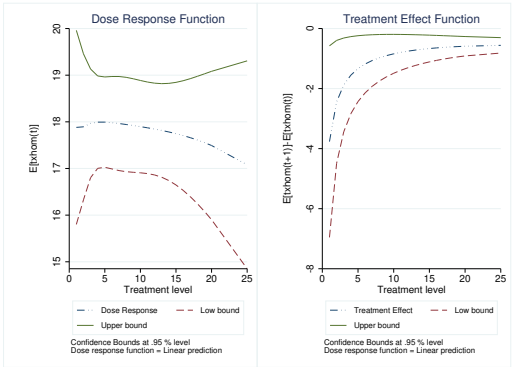
Os grupos de Municípios, as medidas de intensidade e as variáveis de resultado seguem as mesmas considerações feitas na subseção anterior referente a múltiplos tratamentos categóricos. Cabe ainda acentuar que as análises adiante não têm mais como base o grupo de Municípios que não implantou GMs, passando a se restringir somente àqueles que de fato receberam alguma dosagem positiva do programa.

Com relação à intensidade do tratamento pelos anos de funcionamento das GMs sobre a taxa de homicídios, é possível verificar, por meio dos painéis do gráfico da Figura 4, que, em geral, a criminalidade está negativamente relacionada com essa modalidade de participação dos Municípios na segurança pública. Em alguns casos mais intensamente e em outros de maneira mais moderada ou até mesmo estável (Municípios Grandes). Nas localidades com população entre 20.000 e 49.999 (pequenos) e entre 50.000 e 99.999 (médios), os intervalos de confiança a 95% para os efeitos marginais não tocam o zero e, portanto, tem-se uma relação estatisticamente significante por toda a curva. Pode-se inferir, nesses casos, que há um impacto negativo (benéfico) da existência prolongada de GMs sobre a taxa de homicídios, resultados similares aos obtidos e debatidos na subseção anterior (que leva em consideração também o grupo de não tratados). As magnitudes dessas reduções nas taxas de homicídios variaram na continuidade proposta, contudo estima-se que foram, em média, de −1,1 ponto nos Municípios pequenos e de −2,1 nos Municípios médios.

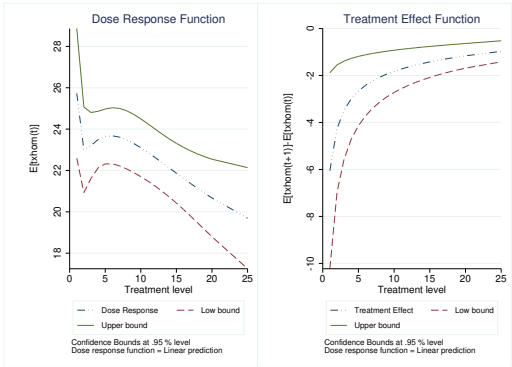
Para os demais grupos populacionais, no entanto, a dinâmica é um pouco diferente. Nos Municípios de porte intermediário (população entre 100.000 e 249.999) os resultados também equivalem aos obtidos com tratamentos multivalorados, isto é, não se vislumbram ganhos da longevidade das corporações locais de segurança pública sobre a taxa de homicídios. No que concerne aos entes mais populosos, o limite superior do intervalo de confiança encontra-se relativamente próximo do zero e o restante todo abaixo, em especial nos níveis mais altos de tratamento, o que poderia ensejar interpretações menos rígidas de significância na margem (próximas a 10%) e, se assim fosse possível conjecturar, ter-se-iam resultados que novamente se aproximariam daqueles obtidos pelo método constante da subseção anterior (benéficos).

Por fim, a base de dados disponível também permite um olhar sobre o comportamento da função DR no contexto dos crimes contra o patrimônio, em especial a *proxy* de roubos e furtos de veículos (gráficos da Figura 5). Em ambos os grupos de Municípios estipulados — pequenos e médios na Figura 5(a) e intermediários e grandes na Figura 5(b) —, esse tipo de criminalidade tende a ser maior nas localidades que já contam há mais tempo com GMs. É provável que, em cada uma das subdivisões populacionais, esses entes sejam aqueles que contam proporcionalmente com maiores oportunidades econômicas para que esses ilícitos aconteçam e que em face dessa mesma dotação benéfica sejam também aqueles que iniciaram o ciclo de criação/implantação de GMs ainda no início dos anos 1990.

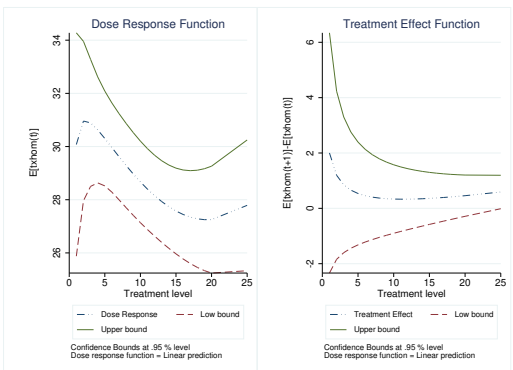
Os efeitos marginais da iniciativa, no entanto, parecem ser nulos em ambos os casos, com leve possibilidade de simultaneidade no primeiro painel (limite inferior do intervalo de confiança exatamente sobre o zero). Essas inferências são em grande parte compatíveis com o que foi debatido na subseção anterior, evidenciando novamente uma dinâmica própria para esse tipo de crime, possivelmente distante da esfera de influência das atividades das GMs, inclusive daquelas implantadas há vários anos.



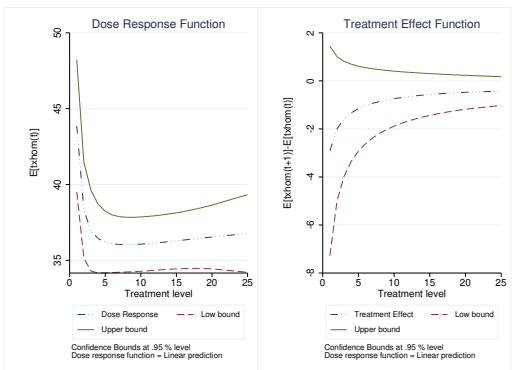

**(a)** Municípios Pequenos




**(b)** Municípios Médios



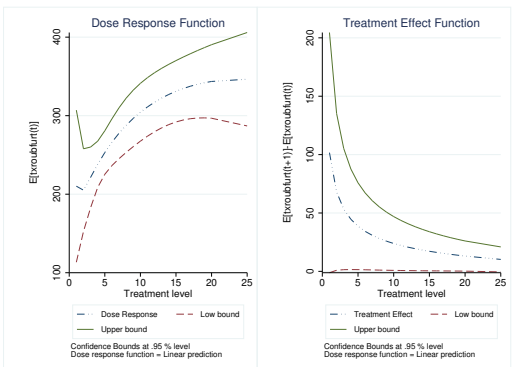

**(c)** Municípios Intermediários



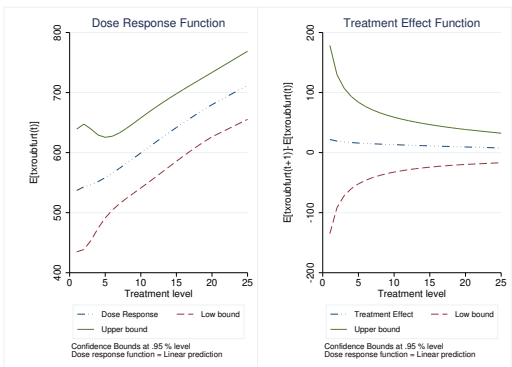

**(d)** Municípios Grandes

*Nota:* Intervalos de confiança obtidos com *bootstrap* de 500 replicações.
*Fonte:* STATA 13, comando `doseresponse2`.

**Figura 4.** Função DR e Efeitos do Tratamento (Anos com GM). Variável dependente = Taxa de Homicídios (1991–2012).




**(a)** Municípios Pequenos e Médios




**(b)** Municípios Intermediários e Grandes

*Nota:* Intervalos de confiança obtidos com *bootstrap* de 500 replicações.
*Fonte:* STATA 13, comando `doseresponse2`.

**Figura 5.** Função DR e Efeitos do Tratamento (Anos com GM). Variável dependente = Taxa de Roubos e Furtos de Veículos (2009–2011).

## 6. Conclusões

Neste estudo foram propostas estratégias baseadas no arcabouço de resultados potenciais para se estimar o impacto causal da existência de Guardas Municipais nos indicadores de segurança pública selecionados. Foi visto que, apesar dessa iniciativa estar arregimentado cada vez mais usuários, ainda não existem, em quantidades desejáveis, contribuições empíricas destinadas a mensurar suas possíveis repercussões, algo de especial interesse na medida que os modelos de atuação escolhidos por tais iniciativas são díspares, podendo ter como foco tanto a vigilância patrimonial quanto o policiamento comunitário e até mesmo atividades policias tipicamente repressoras.

Por ocasião desse vácuo acadêmico, pleiteou-se a utilização de abordagens para o tratamento que não o descrevessem apenas como um fenômeno binário, mas que também estivessem sujeitos a diferentes níveis de intensidade. Especial atenção foi dedicada ao tratamento dos problemas de autosseleção e de simultaneidade, fortemente associados ao contexto explorado. Outra contribuição julgada relevante foi a compilação da base de dados em formato de painel, congregando com ineditismo variáveis de resultado tanto para os crimes contra a pessoa (homicídios) quanto para aqueles relacionados ao patrimônio (roubos e furtos de veículo).

Com o auxílio dos métodos de DID, PSM e a combinação de ambos (PSM-DID), todos baseados na abordagem binária do fenômeno, não foi possível rejeitar a hipótese de que essas corporações possuem pouco potencial para interferir beneficamente nos indicadores de segurança pública. Ainda que as estratégias propostas tenham se mostrado confiáveis para reverter os problemas clássicos de simultaneidade, os parâmetros estimados por todos esses métodos apontaram, em essência, para efeitos nulos decorrentes da implantação da iniciativa.

No entanto, ao se expandir essa abordagem para tratamentos com critérios multivalorados ou, equivalentemente, com dosagens diferenciadas entre os participantes do programa (categóricas e contínuas), surgem evidências de que, para algumas localidades, a implantação das GMs pode sim gerar impactos redutores sobre a criminalidade. Sob a égide dos tratamentos multivalorados por categorias, destacam-se os ganhos relativos sobre a taxa de homicídios em favor de Municípios pequenos e médios tratados em comparação com seus congêneres que não o fizeram (até menos 30% nos homicídios em relação ao grupo de controle ou, de maneira equivalente, até −4,8 mortes por 100 mil habitantes). A principal conjectura para isso remete à possibilidade de que nessas localidades, por serem menores, haja uma disseminação mais efetiva das iniciativas de segurança pública em função da mobilização das comunidades envolvidas e de controles sociais mais rígidos.

Ao se considerarem modelos com tratamentos contínuos, o foco comparativo passou a se dar apenas em relação àqueles Municípios com algum nível de tratamento e a sua resposta em termos das variáveis de resultados. Contudo, de maneira análoga ao modelo anterior, persistem evidências de que nas localidades menores há impactos benéficos da existência de GMs sobre a taxa de homicídios por 100 mil habitantes, contando com reduções médias que variam entre −1,1 e −2,1 pontos em função da exposição contínua a esse programa.

No que concerne à avaliação do tratamento pelos anos de existência das GMs em função da variável de resultados proxy para os crimes contra o patrimônio (taxas de roubos e furtos de veículos), não parecem existir impactos consistentes sobre a variável de resultado em nenhum dos modelos multivalorados abordados (categóricos e contínuos).

Em face das considerações expostas, é possível inferir que as iniciativas municipais relacionadas à segurança pública parecem dispor de algum potencial redutor sobre a criminalidade, em especial nas localidades menores e que, portanto, convém sejam mais bem estruturadas do ponto de vista operacional com o intuito de maximizarem essa efetividade. Se mesmo com o arcabouço jurídico e administrativo vigente, complexo e com organizações cujas funções práticas são pouco nítidas, aventa-se a possibilidade de que as GMs reduzam a criminalidade contra a pessoa, resta projetar os ganhos ainda maiores que existiriam caso operassem sob regras mais racionais.

Em síntese, as prescrições em termos de política pública parecem se concentrar no estímulo para a criação desses aparatos nos Municípios pequenos e médios, onde a redução da criminalidade contra a pessoa (homicídios) de fato parece ter ocorrido, e um desestímulo às iniciativas dos entes de tamanho intermediário e grande, cujas especificidades parecem ser mais bem atendidas pelo policiamento provido pelos Estados e/ou pela União.

## Referências bibliográficas


Bayley, D. H. (1990). *Patterns of policing: A comparative international analysis*. New Jersey: Rutgers University Press.

Beato Filho, C. C. (2002). Reinventando a polícia: A implementação de um programa de policiamento comunitário. *In Policiamento comunitário: Experiências no Brasil 2000–2002* (pp. 113–166). São Paulo: Página Viva.

Becker, G. S. (1968). Crime and punishment: An economic approach. *The Journal of Political Economy*, *76*(2), 169–217. http://dx.doi.org/10.1086/259394

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, *119*(1), 249–275. http://dx.doi.org/10.1162/003355304772839588

Bia, M., & Mattei, A. (2008). A Stata package for the estimation of the dose–response function through adjustment for the generalized propensity score. *The Stata Journal*, *8*(3), 354–373. http://www.stata-journal.com/article.html?article=st0150

Caliendo, M., & Kopeinig, S. (2005, maio). *Some practical guidance for the implementation of propensity score matching* (Discussion Paper Nº 1588). Forschungsinstitut zur Zukunft der Arbeit (IZA). https://www.iza.org/publications/dp/1588/some-practical-guidance-for-the-implementation-of-propensity-score-matching

Centerwall, B. S. (1992). Television and violence : The scale of the problem and where to go from here. *JAMA*, *267*(22), 3059–3063. http://dx.doi.org/10.1001/jama.1992.03480220077031

Cerqueira, D. R. d. C., & Lobão, W. (2003, junho). *Determinantes da criminalidade: Uma resenha dos modelos teóricos e resultados empíricos* (Texto para Discussão Nº 956). Brasília, DF: Instituto de Pesquisa Econômica Aplicada (IPEA). http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=4186

Cerqueira, D. R. d. C., & Mello, J. M. P. d. (2012, março). *Menos armas, menos crimes* (Texto para Discussão Nº 1721). Brasília, DF: Instituto de Pesquisa Econômica Aplicada (IPEA). http://www.ipea.gov.br/portal/index.php?option=com_content&view=article&id=15101

CF – Brasil. (1988). *Constituição da República Federativa do Brasil de 1988*. Brasília, DF. http://www.planalto.gov.br/ccivil_03/Constituicao/Constituicao.htm

Ferreira, L. R. C. (2012). *O papel das guardas municipais na redução de homicídios: Evidências empíricas para o Brasil* (Dissertação de Mestrado, Escola de Economia de São Paulo da Fundação Getúlio Vargas (EESP/FGV), São Paulo, SP). http://hdl.handle.net/10438/9948

Ferreira, L. R. C., Mattos, E., & Terra, R. (2016). O papel das guardas municipais na redução da criminalidade: Evidências empíricas para um painel de municípios paulistas. *Pesquisa e Planejamento Econômico*, *46*(2), 151–180. http://dx.doi.org/http://ppe.ipea.gov.br/index.php/ppe/article/view/1484

Guardabascio, B., & Ventura, M. (2013, março). *Estimating the dose-response function through the GLM approach* (MPRA Paper N$^{o}$ 45013). Munich: Munich Personal RePEc Archive (MPRA). http://mpra.ub.uni-muenchen.de/45013/

Heckman, J. J., Ichimura, H., & Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, *65*(2), 261–294. http://dx.doi.org/10.1111/1467-937X.00044

Heckman, J. J., Ichimura, H., & Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The Review of Economic Studies*, *64*(4), 605–654. http://dx.doi.org/10.2307/2971733

Hennigan, K. M., Del Rosario, M. L., Heath, L., Cook, T. D., Wharton, J. D., & Calder, B. J. (1982). Impact of the introduction of television on crime in the United States: Empirical findings and theoretical implications. *Journal of Personality and Social Psychology*, *42*(3), 461–477. http://dx.doi.org/10.1037/0022-3514.42.3.461

Hirano, K., & Imbens, G. W. (2001). Estimation of causal effects using propensity score weighting: A application to data on right heart catheterization. *Health Services & Outcomes Research Methodology*, *2*(3-4), 259–278. http://dx.doi.org/10.1023/A:1020371

Hirano, K., & Imbens, G. W. (2004). The propensity score with continuous treatments. *In* A. Gelman & X.-L. Meng (Orgs.), *Applied bayesian modeling and causal inference from incomplete-data perspective.* Nova York: Wiley.

IBGE – Instituto Brasileiro de Geografia e Estatística. (2013). *Pesquisa de informações básicas municipais: Perfil dos municípios brasileiros – 2012*. Rio de Janeiro: IBGE. https://ww2.ibge.gov.br/home/estatistica/economia/perfilmunic/2012/default.shtm

IBGE – Instituto Brasileiro de Geografia e Estatística. (2015). *Perfil dos estados e dos municípios brasileiros*. Rio de Janeiro: IBGE.

Imbens, G. W. (2000). The role of the propensity score in estimating dose-response functions. *Biometrika*, *87*(3), 706–710. http://dx.doi.org/10.1093/biomet/87.3.706

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 05–86. http://dx.doi.org/?

JN – Jornal Nacional. (2014, 05 de maio). *Segurança é a 2ª maior preocupação dos brasileiros, segundo pesquisa.* Globo Comunicação e Participações. Acessado em 10 de novembro de 2015: http://g1.globo.com/jornal-nacional/noticia/2014/08/seguranca-e-2-maior-preocupacao-dos-brasileiros-segundo-pesquisa.html

Meyer, B. D. (1995). Natural and quasi-experiments in economics. *Journal of Business & Economic Statistics*, *13*(2), 151–161. http://dx.doi.org/10.1080/07350015.1995.10524589

Monet, J. C. (2001). *Polícias e sociedades na Europa*. São Paulo: Edusp.

Resende, J. P. d., & Andrade, M. V. (2011). Crime social, castigo social: Desigualdade de renda e taxas de criminalidade nos grandes municípios brasileiros. *Estudos Econômicos*, *41*(1), 173–195. http://dx.doi.org/10.1590/S0101-41612011000100007

Robertson, L. A., McAnally, H. M., & Hancox, R. J. (2013). Childhood and adolescent television viewing and antisocial behavior in early adulthood. *Pediatrics*, *131*(3), 439–446. http://dx.doi.org/10.1542/peds.2012-1582

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*(1), 41–55. http://dx.doi.org/10.1093/biomet/70.1.41

StataCorp. (2013). *Stata treatment–effects reference manual: Potential outcomes/counterfactual outcomes – Release 13.* College Station, Texas: StataCorp LP.

Vargas, J. D., & Oliveira, A., Jr. (2010). As guardas municipais no Brasil: Um modelo de análise. *DILEMAS – Revista de Estudos de Conflito e Controle Social*, *3*(7), 85–108. https://revistas.ufrj.br/index.php/dilemas/article/view/7201/0

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2ª ed.). Cambridge, MA: MIT Press.

## Apêndice.

**Tabela A-1.** Estimativas DID do Impacto da Existência de GMs (2009–2011). Variável dependente = log da Taxa de Roubos e Furtos de Veículos.

| Possui GM há pelo menos $X$ anos | OLS1 | | | | OLS2 | | | | DID | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $\beta$ | *P*-value | *C* | *T* | $\beta$ | *P*-value | *C* | *T* | $\beta$ | *P*-value | *C* | *T* |
| *Grupo 1 = População (20.000 a 49.999)* | | | | | | | | | | | | |
| gm1 | 0,464 | 0,000 | 1.252 | 360 | −0,351 | 0,090 | 1.040 | 280 | 0,034 | 0,891 | 1.040 | 280 |
| gm2 | 0,470 | 0,000 | 1.266 | 346 | −0,389 | 0,006 | 1.053 | 267 | −0,599 | 0,040 | 1.053 | 267 |
| gm3 | 0,473 | 0,000 | 1.276 | 336 | −0,340 | 0,014 | 1.059 | 261 | −0,402 | 0,001 | 1.059 | 261 |
| gm4 | 0,498 | 0,000 | 1.291 | 321 | −0,100 | 0,522 | 1.070 | 250 | −0,141 | 0,487 | 1.070 | 250 |
| gm5 | 0,512 | 0,000 | 1.312 | 300 | −0,043 | 0,769 | 1.088 | 232 | −0,099 | 0,655 | 1.088 | 232 |
| gm10 | 0,611 | 0,000 | 1.407 | 205 | 0,076 | 0,551 | 1.159 | 161 | −0,045 | 0,812 | 1.159 | 161 |
| gm15 | 0,653 | 0,000 | 1.454 | 158 | 0,083 | 0,504 | 1.200 | 120 | −0,059 | 0,861 | 1.200 | 120 |
| gm20 | 0,590 | 0,000 | 1.501 | 111 | 0,023 | 0,841 | 1.238 | 82 | 0,306 | 0,374 | 1.238 | 82 |
| *Grupo 2 = População (50,000 a 99,999)* | | | | | | | | | | | | |
| gm1 | 0,287 | 0,000 | 449 | 338 | 0,120 | 0,593 | 374 | 300 | 0,221 | 0,022 | 374 | 300 |
| gm2 | 0,282 | 0,000 | 472 | 315 | 0,279 | 0,134 | 397 | 277 | 0,280 | 0,150 | 397 | 277 |
| gm3 | 0,277 | 0,000 | 486 | 301 | 0,249 | 0,144 | 408 | 266 | 0,060 | 0,727 | 408 | 266 |
| gm4 | 0,308 | 0,000 | 505 | 282 | 0,242 | 0,096 | 424 | 250 | −0,046 | 0,747 | 424 | 250 |
| gm5 | 0,340 | 0,000 | 525 | 262 | 0,203 | 0,149 | 442 | 232 | 0,007 | 0,963 | 442 | 232 |
| gm10 | 0,459 | 0,000 | 611 | 176 | −0,076 | 0,589 | 523 | 151 | −0,045 | 0,738 | 523 | 151 |
| gm15 | 0,460 | 0,000 | 657 | 130 | −0,055 | 0,694 | 563 | 111 | 0,063 | 0,696 | 563 | 111 |
| gm20 | 0,297 | 0,016 | 708 | 79 | −0,181 | 0,243 | 611 | 63 | −0,072 | 0,340 | 611 | 63 |
| *Grupo 3 = População (Acima de 100,000)* | | | | | | | | | | | | |
| gm1 | 0,565 | 0,000 | 220 | 562 | 0,192 | 0,201 | 192 | 491 | 0,058 | 0,312 | 192 | 491 |
| gm2 | 0,503 | 0,000 | 257 | 525 | −0,068 | 0,547 | 224 | 459 | 0,031 | 0,726 | 224 | 459 |
| gm3 | 0,463 | 0,000 | 273 | 509 | −0,126 | 0,273 | 240 | 443 | −0,118 | 0,203 | 240 | 443 |
| gm4 | 0,452 | 0,000 | 285 | 497 | −0,088 | 0,418 | 251 | 432 | −0,179 | 0,044 | 251 | 432 |
| gm5 | 0,455 | 0,000 | 303 | 479 | −0,052 | 0,622 | 267 | 416 | −0,061 | 0,474 | 267 | 416 |
| gm10 | 0,447 | 0,000 | 409 | 373 | −0,075 | 0,380 | 357 | 326 | −0,062 | 0,466 | 357 | 326 |
| gm15 | 0,509 | 0,000 | 507 | 275 | 0,080 | 0,328 | 453 | 233 | 0,378 | 0,288 | 453 | 233 |
| gm20 | 0,447 | 0,000 | 606 | 176 | 0,049 | 0,481 | 533 | 150 | −0,078 | 0,162 | 533 | 150 |
| Dummies de Tempo | | Não | | | | Sim | | | | Sim | | |
| Efeitos Fixos | | Não | | | | Não | | | | Sim | | |
| Controles | | Não | | | | Sim | | | | Sim | | |

*Notas:* Erros-Padrão robustos à heteroscedasticidade e à correlação serial. *C* = *untreated* (grupo controle) e *T* = *treated* (grupo de tratamento).
*Fonte:* Elaboração dos autores com base nas estimativas fornecidas pelo programa Stata 13.

**Tabela A-2.** Estimativas do Impacto da Existência de GMs (2009–2011). Variável Dependente = Taxa de Roubos e Furtos de Veículos (PSM) e a sua diferença (PSM-DID).

| Possui GM há pelo menos *X* anos | PSM 1 (NNM) | | PSM 2 (kernel) | | Observações | | | PSM 1 - DID | | PSM 2 - DID | | Observações | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | att | *P*-value | att | *P*-value | *n* | *C* | *T* | att | *P*-value | att | *P*-value | *n* | *C* | *T* |
| *Grupo 1 = População (20.000 a 49.999)* | | | | | | | | | | | | | | |
| gm1 | 64,41 | 0,008 | 65,57 | 0,001 | 1.370 | 1.085 | 285 | −19,73 | 0,174 | −33,17 | 0,016 | 917 | 726 | 191 |
| gm2 | 57,34 | 0,017 | 65,83 | 0,001 | 1.372 | 1.098 | 274 | −32,58 | 0,049 | −34,01 | 0,018 | 918 | 737 | 181 |
| gm5 | 69,44 | 0,006 | 64,31 | 0,004 | 1.378 | 1.137 | 241 | −35,50 | 0,058 | −30,10 | 0,056 | 922 | 758 | 164 |
| gm10 | 95,91 | 0,002 | 98,27 | 0,000 | 1.378 | 1.215 | 163 | −23,21 | 0,225 | −26,17 | 0,109 | 927 | 811 | 116 |
| gm15 | 107,09 | 0,002 | 107,09 | 0,000 | 1.383 | 1.255 | 128 | −34,04 | 0,131 | −40,63 | 0,063 | 927 | 843 | 84 |
| gm20 | 107,25 | 0,019 | 88,02 | 0,012 | 1.352 | 1.265 | 87 | −29,84 | 0,229 | −20,55 | 0,269 | 905 | 848 | 57 |
| *Grupo 2 = População (50,000 a 99,999)* | | | | | | | | | | | | | | |
| gm1 | 44,34 | 0,122 | 70,52 | 0,006 | 717 | 419 | 298 | 8,35 | 0,357 | 2,91 | 0,391 | 469 | 266 | 203 |
| gm2 | 44,69 | 0,127 | 78,50 | 0,002 | 718 | 441 | 277 | −5,46 | 0,383 | −2,86 | 0,392 | 467 | 280 | 187 |
| gm5 | 81,19 | 0,005 | 90,28 | 0,001 | 711 | 478 | 233 | 13,53 | 0,298 | 1,87 | 0,396 | 466 | 311 | 155 |
| gm10 | 34,25 | 0,270 | 87,44 | 0,011 | 705 | 557 | 148 | 13,85 | 0,326 | 8,28 | 0,358 | 475 | 368 | 107 |
| gm15 | 34,94 | 0,299 | 53,16 | 0,166 | 716 | 601 | 115 | 16,07 | 0,318 | 7,07 | 0,376 | 475 | 399 | 76 |
| gm20 | −2,22 | 0,399 | 8,16 | 0,392 | 721 | 650 | 71 | −13,45 | 0,358 | −0,78 | 0,399 | 481 | 431 | 50 |
| *Grupo 3 = População (Acima de 100,000)* | | | | | | | | | | | | | | |
| gm1 | 80,63 | 0,103 | 115,90 | 0,070 | 722 | 202 | 520 | 1,08 | 0,398 | 0,632 | 0,399 | 485 | 135 | 350 |
| gm2 | 131,45 | 0,002 | 143,72 | 0,000 | 723 | 239 | 484 | −5,85 | 0,381 | 9,604 | 0,353 | 485 | 158 | 327 |
| gm5 | 78,19 | 0,060 | 111,32 | 0,005 | 724 | 281 | 443 | −14,02 | 0,300 | −9,891 | 0,351 | 486 | 185 | 301 |
| gm10 | 106,96 | 0,012 | 114,80 | 0,000 | 727 | 381 | 346 | 1,98 | 0,397 | −2,967 | 0,395 | 487 | 253 | 234 |
| gm15 | 87,40 | 0,037 | 84,25 | 0,008 | 729 | 474 | 255 | −1,97 | 0,397 | −9,378 | 0,337 | 489 | 314 | 175 |
| gm20 | 43,26 | 0,270 | 55,78 | 0,100 | 731 | 567 | 164 | −5,30 | 0,388 | −8,411 | 0,360 | 490 | 371 | 119 |

*Notas:* Intervalos de confiança obtidos por *bootstrap* (200 replicações). *C = untreated* (grupo controle) e *T = treated* (grupo de tratamento).
*Fonte:* Elaboração dos autores com base nas estimativas fornecidas pelo programa Stata 13.

**Tabela A-3.** Estimativas para o Impacto Multivalorado das GMs (2009–2011). Variável dependente = log da Taxa de Roubos e Furtos de Veículos.

| Níveis de Tratamento | RA | | | IPW | | | IPWRA | | |
|---|---|---|---|---|---|---|---|---|---|
| | att | *P*-value | *N* | att | *P*-value | *N* | att | *P*-value | *N* |
| *Grupo 1 = População (20.000 a 99.999)* | | | | | | | | | |
| (A vs 0) | -0,13 | 0,286 | 1.994 | -0,10 | 0,481 | 2.126 | -0,20 | 0,145 | 1.815 |
| (B vs 0) | 0,15 | 0,195 | | 0,00 | 0,978 | | 0,13 | 0,328 | |
| (C vs 0) | 0,06 | 0,641 | | -0,15 | 0,366 | | 0,00 | 0,991 | |
| (D vs 0) | 0,09 | 0,368 | | 0,10 | 0,587 | | 0,18 | 0,164 | |
| *Grupo 4 = População (Acima de 100.000)* | | | | | | | | | |
| (A vs 0) | 0,24 | 0,034 | 683 | -0,29 | 0,527 | 735 | 0,23 | 0,161 | 640 |
| (B vs 0) | 0,21 | 0,261 | | -0,56 | 0,235 | | 0,10 | 0,654 | |
| (C vs 0) | 0,18 | 0,173 | | -0,34 | 0,463 | | 0,27 | 0,134 | |
| (D vs 0) | -0,01 | 0,973 | | -0,40 | 0,382 | | 0,02 | 0,928 | |

*Notas:* Erros padrão robustos. Nível 0 = sem GM; A = GM de 1 a 3 anos; B = GM de 4 a 6 anos; C = GM de 7 a 10 anos; D = GM há mais de 10 anos.
*Fonte:* Elaboração dos autores com base nas estimativas fornecidas pelo programa Stata 13.

r b e
REVISTA BRASILEIRA DE ECONOMIA
DOI 10.5935/0034-7140.20180026
ISSN 1806-9134 (online)
FGV
FUNDAÇÃO GETULIO VARGAS

# Errata

No artigo "Impactos Econômicos de Políticas Climáticas Europeias e Americanas Sobre a Economia Brasileira", com número de DOI 10.5935/0034-7140.20180004, publicado no periódico *Revista Brasileira de Economia*, v. 72 n. 1, pp. 61–79, no rodapé da página 61:

*Onde se lia:*

> FEARP/USP. Rua Itapeva, 474, 6º andar, GVAgro Bela Vista, Sao Paulo, SP, Brasil. CEP 01332-000.

*Leia-se:*

> EESP/FGV. Rua Itapeva 474, 6º andar, GVAgro, Bela Vista, São Paulo, SP, Brasil. CEP 01332-000.



This is an open-access article distributed under the terms of the Creative Commons Attribution License.