UNIVERSIDADE FEDERAL DE PERNAMBUCO
DEPARTAMENTO DE ECONOMIA
Programa de Pós-graduação em Economia - PIMES

Júlio César Silva Becher

Ensaios em Econometria Espacial: aplicações em economia urbana e economia política

Recife/PE

2022

Júlio César Silva Becher

Ensaios em Econometria Espacial: aplicações em economia urbana e economia política

Tese de Doutorado submetida ao Programa de Pós-graduação em Economia, da Universidade Federal de Pernambuco como parte dos requisitos para a obtenção do título de Doutor em Economia.
Linha de Pesquisa: Teoria Econômica

Orientador: Marcelo Eduardo Alves da Silva
Co-orientador: André Matos Magalhães

Recife/PE

2022

Catalogação na Fonte
Bibliotecária Ângela de Fátima Correia Simões, CRB4-773

B394e Becher, Júlio César Silva
Ensaios em Econometria espacial: aplicações em economia urbana e economia política / Júlio César Silva Becher. - 2022.
91 folhas: il. 30 cm.

Orientador: Prof. Dr. Marcelo Eduardo Alves da Silva e Coorientador Prof. Dr. André Matos Magalhães.
Tese (Doutorado em Economia) – Universidade Federal de Pernambuco, CCSA, 2022.
Inclui referências.

1. Econometria. 2. Programa Bolsa Família (Brasil). 3. Tática política. I. Silva, Marcelo Eduardo Alves da (Orientador). II. Magalhães, André Matos (Coorientador). III. Título.

336 CDD (22. ed.) UFPE (CSA 2022 – 068)

Júlio César Silva Becher

Ensaios em Econometria Espacial: aplicações em economia urbana e economia política

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Economia - PIMES, do Departamento de Economia da Universidade Federal de Pernambuco como requisito parcial para obtenção do título de Doutor em Ciências Econômicas.

Aprovado em: 24/08/2022

BANCA EXAMINADORA

Prof. Dr. André Matos Magalhães (Co-orientador)
Universidade Federal de Pernambuco

Prof. Dr. Raul da Mota Silveira Neto (Examinador 1)
Universidade Federal de Pernambuco

Prof. Dr. Luis Eduardo Barbosa Carazza (Examinador 2)
Universidade Federal Rural de Pernambuco

Prof. Dra. Roberta Rocha (Examinador 3)
Universidade Federal de Pernambuco

Prof. Dra. Sónia Maria F. P. Oliveira Gomes (Examinador 4)
Universidade Federal Rural de Pernambuco

À Andressa, Petrus e Matteo.

Resumo

Estudar questões espaciais é uma abordagem largamente utilizada quando se objetiva enteder o efeito de transbordamento de políticas no âmbito regional. Nessa pesquisa busca-se contribuir com esse debate na literatura empírica que analisa a relação entre políticas públicas e dinâmica regional. Para isso, serão realizadas três ensaios empíricos. O primeiro ensaio objetiva tem como propósito avaliar a relação entre o tempo de deslocamento (commuting) e o acesso ao emprego feminino. No segundo artigo buscou-se mensurar a influência da distribuição de beneficiários do bolsa família nos municípios nas eleições presidenciais de 2018 e, por fim, o terceiro ensaio se analisa o efeito de uma reforma metodológica na definição das cotas de beneficiários ocorrida em 2009, a nível de município.

Palavras-chave: Econometria Espacial; Commuting time; Bolsa Família.

Abstract

Spatial econometric models is a widely used approach when the objective is to understand the spillover effect of policies at the regional level. This research aim to contribute to this debate in the empirical literature that analyzes the relationship between public policies and regional dynamics. For this, three empirical essays are carried out. First one objective essay aims to evaluate the relationship between commuting time and access to female employment. The second article sought to measure the influence of the distribution of beneficiaries of Bolsa Família in the municipalities in the 2018 presidential elections and, finally, the third essay analyzes the effect of a methodological reform on the definition of beneficiary quotas that took place in 2009, at the municipal level.

Keywords: Spatial Econometrics; Commuting time; Bolsa Família.

Lista de figuras

Figura 1 – Primeiro Mapa de São Paulo 20
Figura 2 – Mapa dos Municípios da RMSP 24
Figura 3 – Densidade da Força de Trabalho Feminina 24
Figura 4 – Taxa de desemprego Feminina 25
Figura 5 – Tempo de deslocamento para os homens 27
Figura 6 – Mapa LISA da Força de Trabalho Feminina - Solteiras 28
Figura 7 – Mapa LISA da Força de Trabalho Feminina - Casadas 28
Figura 8 – Mapa LISA do Desemprego Feminino - Solteiras 29
Figura 9 – Mapa LISA do Desemprego Feminino - Casadas 30
Figura 10 – Proporção de votos presidenciais - 2018 (1 turno) 44
Figura 11 – Moran Scatterplot para (%) Votos no Haddad 2018_1 53
Figura 12 – Local Moran (LISA) para (%) Votos no Haddad 2018_1 53
Figura 13 – Moran Scatterplot para (%) Votos no Haddad 2018_1 58
Figura 14 – Local Moran (LISA) para (%) Votos no Bolsonaro 2018_1 59
Figura 15 – Dispersão do choque da Reforma 2009 68
Figura 16 – Filtros aplicados na base de Municípios 69
Figura 17 – Proporção de beneficiários do BF - 2006/2010 73
Figura 18 – Variação geográfica da Reforma de 2009 74
Figura 19 – Efeito da Reforma 2009 na proporção de beneficiários 77
Figura 20 – Filtros aplicados na base de Municípios - Robustez 1 77
Figura 21 – Efeito da Reforma 2009 na proporção de beneficiários - Robustez 1 . . 78
Figura 22 – Filtros aplicados na base de Municípios - Robustez 1 79
Figura 23 – Efeito da Reforma 2009 na proporção de beneficiários - Robustez 2 . . 80
Figura 24 – Dispersão entre o percentual de beneficiários do BF e o coeficiente de abertura comercial 80
Figura 25 – Efeito da Reforma 2009 na proporção de beneficiários - Falsificação 1 . 81
Figura 26 – Dispersão entre o percentual de beneficiários do BF e o coeficiente de abertura comercial 81
Figura 27 – Efeito da Reforma 2009 na proporção de beneficiários - Falsificação 2 . 82

Lista de tabelas

Tabela 1 – Descrição das variáveis 22
Tabela 2 – Sumário Descritivo das Variáveis Dependentes 25
Tabela 3 – Regressão Participação na Força de Trabalho Feminina 31
Tabela 4 – Estatística I de Moran - Participação na Força de Trabalho 32
Tabela 5 – Efeito marginal do Commuting na Participação da Força de Trabalho . 33
Tabela 6 – Regressão para Taxa de Desemprego Feminino no Mercado de Trabalho 34
Tabela 7 – Estatística I de Moran 35
Tabela 8 – Efeito marginal do Commuting na Taxa de Desemprego Feminina . . . 35
Tabela 9 – Descrição das variáveis de trabalho 48
Tabela 10 – Estatísticas descritivas das variáveis 49
Tabela 11 – Proporção de votos no Haddad 2018 via OLS 51
Tabela 12 – Resultados dos testes de dependência espacial – Haddad 2018_1 . . . 52
Tabela 13 – Modelo com correção espacial 55
Tabela 14 – Proporção de votos Bolsonaro 2018.1 via OLS 57
Tabela 15 – Resultados dos testes de dependência espacial – Bolsonaro 2018_1 . . 58
Tabela 16 – Modelo com correção espacial 60
Tabela 17 – Estatísticas Descritivas por Grupo de Municípios 73
Tabela 18 – Estimação via OLS 75
Tabela 19 – Estimação do efeito da expansão do BF no voto presidencial 76
Tabela 20 – Estimação do efeito da expansão do BF no voto presidencial 78
Tabela 21 – Estimação do efeito da expansão do BF no voto presidencial 79

Sumário

1 APRESENTAÇÃO 10

2 *COMMUTING* E ACESSO AO EMPREGO FEMININO: UM ESTUDO DA RM DE SÃO PAULO 11
2.1 Introdução 11
2.2 Revisão da Literatura 14
2.3 Estratégia Empírica 17
2.3.1 Análise da Dependência Espacial 18
2.3.2 Correção da Potencial Endogeneidade 19
2.4 Base de dados e Região de Estudo 21
2.5 Resultados 24
2.5.1 Análise do Padrão Espacial da Empregabilidade Feminina 24
2.5.2 Efeitos do Tempo de Commuting para a Participação na Força de Trabalho 30
2.6 Conclusão 36

3 VOTOS PRESIDENCIAIS E PBF: HÁ CORRELAÇÃO ESPACIAL? 38
3.1 Introdução 38
3.2 Revisão da Literatura 40
3.3 Estratégia Empírica 44
3.4 Base de Dados e Variáveis 48
3.5 Resultados 49
3.6 Conclusões 60

4 BOLSA FAMÍLIA AFETA VOTO PRESIDENCIAL? EVIDÊNCIAS PARA O CASO BRASIL 63
4.1 Introdução 63
4.2 O Programa Bolsa Família e a Reforma Metodológica de 2009 . . . 66
4.3 Estratégia Empírica 69
4.4 Dados 72
4.5 Resultados 74
4.6 Robustez 77
4.7 Conclusões 82

REFERÊNCIAS 84

1 Apresentação

Essa Tese de Doutorado encontra-se estruturada em três Ensaios aplicados, o primeiro tem como propósito avaliar a relação entre o tempo de deslocamento (*commuting*) e o acesso ao emprego feminino. A questão fundamental desse artigo é analisar como o afastamento físico das mulheres dos centros de emprego afeta sua dinâmica no mercado de trabalho. Utilizou-se uma base de dados única para a RMSP, com dados desagregados por área de ponderação e construída a partir do Censo de 2010. Foram estabelecidas quatro variáveis dependentes para a participação feminina no mercado de trabalho, considerando o estado civil e a presença de filhos. Com o intuito de corrigir o viés da dependência espacial, utilizou-se o *spatial Durbin model* como estratégia empírica, estimado via *Two Stage Least Square*. Os resultados evidenciam uma associação negativa e significante apenas para o grupo das mulheres casadas e com filhos, reforçando a hipótese da penalização mais intensa às mulheres na presença de filhos.

No segundo artigo buscou-se mensurar a influência da distribuição de beneficiários do bolsa família nos municípios nas eleições presidenciais de 2018. A distribuição de votos presidenciais apresentou um correlação espacial marcante e para corrigir essa dependência regional, fez-se uso do *spatial Error model* como estratégia empírica. A base de dados foi construída sob informações oriundas do TSE e do IBGE e segue uma linha de estudos voltados a compreender os efeitos de programas sociais na decisão de voto. Os resultados mostraram que os municípios com maior proporção de beneficiários do programa votaram mais no candidato do PT nas eleições. Por outro lado, tais municípios foram menos propícios ao então candidato, Jair Bolsonaro.

No último ensaio, a proposta foi analisar o efeito de uma reforma metodológica na definição das cotas de beneficiários ocorrida em 2009, a nível de município. Essa reforma, modificou exogenamente o percentual de beneficiários do programa Bolsa Família. Contudo, a redistribuição das cotas de beneficiários foi desproporcional intra municípios, devido ao ajuste metodológico na definição de pobreza. A fonte de variação exógena oriunda da reforma institucional permitiu, a partir de um modelo de diferenças em diferenças, que além de mitigar a potencial endogeneidade, identifica o efeito causal do Programa Bolsa Família na determinação do voto, em especial, no Partido dos Trabalhadores. Os resultados que a expansão no número de beneficiários impactou positivamente os votos no PT.

2 *Commuting* e acesso ao emprego feminino: um estudo da RM de São Paulo

2.1 Introdução

Associar o tempo de deslocamento à empregabilidade feminina vem tomando relevante importância na análise urbana. Isso por que o constrangimento promovido pelo tempo perdido em trajetos longos (ou demorados) traz consigo uma série de custos de oportunidade no mercado de trabalho, desde a redução na produtividade dos fatores (DURANTON; TURNER, 2012) ao aumento da probabilidade de crimes (SILVEIRA NETO; MOURA, 2019). Por exemplo, Kawabata e Abe (2018) documentaram uma associação negativa e significativa entre o *commuting* e a participação feminina no mercado de trabalho, sendo as casadas e com filhos mais sensíveis a áreas com maior tempo de deslocamento casa-trabalho, em consonância com um conjunto de trabalhos que também apontam evidências no mesmo sentido (CRANE, 2007; BLACK; KOLESNIKOVA; TAYLOR, 2014; FOGLI; VELDKAMP, 2011; SILVEIRA NETO; DUARTE; PÁEZ, 2015; FARRÉ; JOFRE-MONSENY; TORRECILLAS, 2020; BARBANCHON; RATHELOT; ROULET, 2021; CHIDAMBARAM; SCHEINER, 2020). Contudo, ainda existe uma lacuna de análise dessa relação, em especial para o caso de países em desenvolvimento, como expõem Black, Kolesnikova e Taylor (2014), Chauvin et al. (2017) e Kawabata e Abe (2018).

Impende destacar que o custo de moradia aumentou nas principais cidades brasileiras, contribuindo para a ampliação na dificuldade de acesso ao mercado de trabalho (HADDAD; VIEIRA, 2015; SEABRA; SILVEIRA NETO; MENEZES, 2016; DUARTE, 2020). No entanto, tais efeitos têm configurações diferentes em função do grau de desenvolvimento do país. Black, Kolesnikova e Taylor (2014), Kawabata e Abe (2018), Farré, Jofre-Monseny e Torrecillas (2020) e Chidambaram e Scheiner (2020) estudaram o caso de cidades em países considerados desenvolvidos[1] e encontraram que mulheres casadas e com filhos são mais afetadas em termos de acesso ao mercado de trabalho, dada decisão de moradia mais distante do centro. Contudo, ao analisar países cujo padrão de renda é mais baixo, evidencia-se o surgimento de fatores adicionais, tais como o tipo da ocupação predominante entre as mulheres, a diferença salarial entre gênero, a segurança urbana e a qualidade dos modais ofertados, potencializando a dificuldade no acesso feminino ao mercado de trabalho (PEREIRA; SCHWANEN, 2015; SILVEIRA NETO; DUARTE; PÁEZ, 2015; SILVEIRA NETO; MOURA, 2019; TEDESCO; SOUZA, 2020).

Pereira e Schwanen (2015) mostraram que existe uma evidente variabilidade na distribuição geográfica feminina no mercado de trabalho e, mesmo a despeito do investimento

[1] EUA, Japão, Espanha e Alemanha, respectivamente.

na ampliação de vias e da diversificação de modais de transporte, o longo tempo dispendido no trajeto casa-trabalho ainda se configura como uma relevante externalidade negativa nos principais centros urbanos. No caso específico do Brasil, a literatura aponta para algumas razões que buscam justificar essa diferença no tempo de deslocamento. Por exemplo, o tipo de postos de trabalho predominantemente ocupados pelas mulheres nos centros urbanos, a exemplo de escolas, unidades de saúde e o comércio local (TEDESCO; SOUZA, 2020). Mulheres ganham menos e, com isso, preferem não arcar com custos maiores de *commuting* (BARUFI; HADDAD, 2017; DUARTE, 2020). Outro fator é a Hipótese da responsabilidade familiar, que ainda exerce influência negativa, pois imputa atribuições adicionais e restringe a capacidade de deslocamento das mulheres (SILVEIRA NETO; DUARTE; PÁEZ, 2015). Concomitantemente, resta evidenciar o fator segurança no transporte coletivo do meio urbano, como mostram Silveira Neto e Moura (2019), tornando a mulher mais suscetível a ocorrência de assaltos e furtos.

Adicionalmente, fatores como a qualidade dos modais (SÁNCHEZ; ISABEL; GONZÁLEZ, 2014; HADDAD; VIEIRA, 2015), o *gap* salarial, inclusive associado a Hipótese de Responsabilidade Doméstica (ALESINA; GIULIANO; NUNN, 2013) e o longo tempo de *commuting*, que expande a segregação espacial e dificulta o acesso ao centro de emprego devido ao intenso movimento pendular dos trabalhadores (BARUFI; HADDAD, 2017; DUARTE, 2020), são exemplos anteriormente estudados pela literatura. Em São Paulo, para se ter uma dimensão, segundo os dados do Censo (2010), mais da metade (55,3%) dos trabalhadores despendiam tempo maior do que 30 minutos se deslocando de casa para o trabalho, e 23% mais de uma hora. Esse padrão assemelha-se ao observado em outras cidades, tais como: Rio de Janeiro, Belo Horizonte, Recife e Porto Alegre. Pode-se ressaltar que a cidade do Recife, por exemplo, obteve o maior crescimento no tempo de deslocamento casa-trabalho entre 2003 e 2018 (SILVEIRA NETO; DUARTE; PÁEZ, 2015; BARBOSA; SILVEIRA NETO, 2017; IPEA, 2018).

Considerando tais evidências documentadas pela literatura, o presente ensaio analisa a associação entre o tempo de deslocamento masculino e a empregabilidade feminina, porém sob uma prisma intra-municipal, utilizando como unidade espacial as áreas de ponderação da RMSP. A escolha desse caso se justifica por três razões principais. Primeiro, por se tratar de um país em desenvolvimento, a infraestrutura de transportes mostra-se mais precária, fato que afeta a decisão de *commuting* (RAPINO; COOKE, 2011). Além disso, a estrutura urbana é diferente em comparação com outras metrópoles desenvolvidas, com maior densidade de ricos na região central e pobres na periferia, tornando o efeito do longo *commuting* mais perverso (HADDAD; VIEIRA, 2015). Adicionalmente, segundo Chauvin et al. (2017), existem poucas evidências sobre as características urbanas nas cidades pertencentes a países em desenvolvimento. Por fim, no cenário de estudo, os mais pobres ainda enfrentam a ausência de serviços de suporte às famílias, a exemplo da oferta de creches (LAMEIRA; GOLGHER, 2021).

No presente ensaio foram analisadas duas medidas para mulheres no mercado de

trabalho: participação na força de trabalho e a taxa de desemprego, ambas por área de ponderação. Mais especificamente, considera-se uma categorização das mulheres pelo estado civil e pela presença de filhos. A análise foi realizada para uma amostra de mulheres entre 18 e 54 anos, devido ao fato dessa faixa de idade representar o grupo mais afetado pela presença de filhos pequenos (CRANE, 2007). Segundo o Censo (2010), a idade média das gestantes para a RMSP é de 29,1 anos, e há uma tendência de aumento na idade média, explicada pela mudança de comportamento feminino no mercado de trabalho (TEDESCO; SOUZA, 2020). A principal variável explicativa será o tempo casa-trabalho masculino obtido pelo Censo (2010), assim como um conjunto de variáveis de controle, descritas na Seção 2.4. Contudo, importante destacar que, ao abordar o tema proposto, torna-se necessário tratar dois problemas potenciais, quais sejam: a possibilidade de dependência espacial e a endogeneidade entre as principais variáveis analisadas.

Com relação ao primeiro ponto, a estimação via OLS se mostrou inadequada, devido a presença da dependência espacial detectada pela estatística I de Moran. Para lidar com essa questão, utiliza-se o *spatial Durbin model - SDM*, que inclui *lags* espaciais na variável dependente, assim como na variável explicativa (ANSELIN et al., 2001). A despeito disso, Elhorst (2010) relata que isso não corrige a potencial endogeneidade na relação estimada. Kelejian e Prucha (1997) sugerem que a estimação do SDM ocorra via *Spatial Two Stages Least Square - STSLS*, instrumentalizando a variável dependente. Em artigo recente, Haddad e Barufi (2017) estimaram o salário de mulheres na RMSP, considerando o *commuting* e um índice aglomeração urbana, e para lidar com a potencial endogeneidade, instrumentalizaram o índice construindo pela distância do local de moradia até o centro de São Paulo, via rios. Assim, consideramos esse instrumento para lidar com a potencial endogeneidade entre as variáveis dependentes observadas nesse estudo e o tempo de deslocamento masculino.

Esse artigo está conectado com a literatura que estuda a relação entre a mobilidade urbana e a dinâmica no mercado de trabalho por gênero. As contribuições podem ser observadas sob duas perspectivas principais. Em primeiro lugar, constrói-se evidências importantes sobre a relação entre mercado de trabalho feminino e tempo de *commuting* para o cenário de um país em desenvolvimento, todavia com alto grau de urbanização. Elabora uma análise única para duas variáveis de mercado de trabalho: taxa de participação na força de trabalho e taxa de desemprego, considerando quatro grupos em função do estado civil e presença de filhos, isto é, solteiras e casadas, com filhos e sem filhos. Por fim, reporta evidências a partir de um modelo empírico que visa corrigir a potencial relação endógena contida na estimação espacial, instrumentalizando o tempo de *commuting*.

Os resultados mostraram que, além de variabilidade no padrão espacial da empregabilidade feminina na RMSP, observou-se que existe uma associação negativa e significante entre a taxa de participação da força de trabalho feminina e o tempo de deslocamento casa-trabalho para o grupo de mulheres casadas e com filhos, revelando que o estado civil e a presença de filhos penalizam a mulher no tocante a empregabilidade. Ao analisar a taxa

de desemprego, o *communting* potencializa o desemprego entre as mulheres, e se mostrou significante somente para o caso das casadas com filhos. Esses resultados constituem mais uma peça de evidências na direção de que problemas de mobilidade urbana penalizam mais fortemente mulheres, em especial àquelas com filhos pequenos, em consonância com a literatura dedicada a estudar essa relação, porém adicionando valor por, além de lidar com o problema espacial, corrige a possível endogeneidade entre o tempo de *commuting* e as medidas de empregabilidade.

A seção seguinte se propõe a uma revisão da literatura. A seção 2.3 descreve os métodos utilizados na pesquisa. A região de estudo e os dados são expostos na Seção 2.4. Os resultados e as conclusões estão nas Seções 2.5 e 2.6, respectivamente.

2.2 Revisão da Literatura

Em cidades monocêntricas adensadas, baseadas no modelo estruturado por (ALONSO et al., 1964; MUTH, 1969; MILLS, 1967), a renda tem influência direta na decisão locacional das famílias. Ao se observar esse desenho baseado no exemplo de países desenvolvidos, as famílias de maior renda optam por se estabelecer mais distantes do centro (CBD - Central Business District) em busca de mais espaço e, com isso, arcam com o custo de deslocamento (ABE, 2011; OMMEREN; PUIGARNAU, 2011). No entanto, de acordo com Åslund, Östh e Zenou (2010), Duranton e Turner (2012) e Andersson et al. (2018), ao considerar o custo de oportunidade do tempo perdido nesse deslocamento, esse resultado é ambíguo, pois, famílias de maior renda desejam estar mais próximas do CBD, promovendo uma maior complexidade ao acesso a essas moradias.

Uma das consequências de cidades que experimentam um longo tempo de deslocamento casa-trabalho é o aumento da distorção de acesso às oportunidades de emprego para grupos minoritários (ONG; MILLER, 2005; CRANE, 2007) e o gap de gênero (ABE, 2011; FARRÉ; JOFRE-MONSENY; TORRECILLAS, 2020), penalizando mais intensamente as mulheres, em especial as casadas e com filhos no mercado de trabalho (BLACK; KOLESNIKOVA; TAYLOR, 2014; KAWABATA; ABE, 2018; BARBANCHON; RATHELOT; ROULET, 2021).

Crane (2007) documentou que o tempo de *commuting* das mulheres é menor do que o masculino. E expõe como potenciais mecanismos, tais como a hipótese de responsabilidade doméstica, a estrutura urbana das cidades e o gap salarial. Fogli e Veldkamp (2011), por exemplo, reforçam esse último ponto, expondo a diferenciação salarial como um desincentivo, haja vista a segregação estrutural de oportunidades por gênero. E quando associada a uma composição familiar com filhos, leva a escolha por moradias mais distantes do centro, e consequente alto custo oportunidade de deslocamento.

Abe (2011) buscou estudar o caso das áreas metropolitanas de Tóquio e implementou um modelo de oferta de trabalho para a escolha das mulheres casadas. Os resultados

mostram que há uma significativa variabilidade na oferta de trabalho nessas áreas, sendo o *commuting* a principal variável de explicação. Nessa mesma linha, Black, Kolesnikova e Taylor (2014) expandem essa evidência encontrando uma relação negativa entre a proporção de mulheres casadas na força de trabalho e o tempo de *commuting* para o mercado de trabalho Norte-americano. Os autores mostram que um aumento no tempo de deslocamento está associado a uma queda na participação feminina. Tal evidência também é observada considerando uma especificação de diferenças em diferenças, isto é, áreas metropolitanas nas quais pode se observar um crescimento no tempo de deslocamento, experimentaram uma queda mais acentuada na participação feminina no mercado de trabalho.

No contexto de existência de uma variabilidade significativa no tempo de deslocamento por gênero, Sánchez e González (2016) estudam essa problemática para o caso espanhol, cujo padrão de mobilidade é fortemente diferente intra-regiões. As autoras evidenciam que mulheres, apesar do tempo de deslocamento menor, acabam por percorrer distâncias similares a dos homens, compactuando com o canal de que as mulheres assumem mais papeis e responsabilidades familiares, exigindo mais deslocamentos, porém de menor duração.

Kawabata e Abe (2018) estudaram a região metropolitana de Tóquio sob um aspecto intra-metropolitano. As autoras avaliaram o padrão espacial da empregabilidade feminina e sua associação com o tempo de deslocamento. Nesse caso, foi considerado o tempo de deslocamento do homem, uma vez que dada a escolha residencial mais periférica, os homens mantém o deslocamento ao CBD e as mulheres orbitam em um raio menor. O estudo considerou três medidas para analisar a associação com o mercado de trabalho: i) taxa de participação na força de trabalho, ii) taxa de participação no emprego regular e iii) no emprego *part-time*. Os resultados mostraram uma associação significativa e negativa entre o tempo de deslocamento do homem e a participação feminina no mercado de trabalho. Sendo tal efeito mais intenso nas mulheres casadas e com filhos pequenos, e principalmente nas menos escolarizadas, reduzindo a participação feminina no mercado de trabalho.

Farré, Jofre-Monseny e Torrecillas (2020) também analisam o efeito do *commuting* na diferença de acesso ao mercado de trabalho por gênero. Os autores estudam o caso de cidades nos EUA, e medem a elasticidade na oferta de trabalho para homens e mulheres. Adicionalmente, para lidar com a endogeneidade dessa relação, instrumentalizam o tempo de deslocamento a partir do formato geográfico das cidades como variável instrumental. Os resultados mostram que o tempo de *commuting* vem crescendo e intensifica a diferença por gênero, sendo as mulheres casadas mais sensíveis ao longo commuting.

Por outro lado, a realidade do acesso ao emprego das mulheres nos países em desenvolvimento é marcadamente distinta, e tais evidências possuem implicações e desdobramentos diferentes em termos de desenho de mecanismos e de políticas públicas,

fato que torna essa comparação relevante e necessária, especialmente devido ao papel social e familiar imputado à mulher (COSTA; KAHN, 2000; CRANE, 2007; ALESINA; GIULIANO; NUNN, 2013).

A Hipótese da Responsabilidade Doméstica (HRH) foi estudada largamente na tentativa de explicar a conexão entre o acesso ao emprego e a decisão de moradia (RAPINO; COOKE, 2011; TEDESCO; SOUZA, 2020). Essa questão foi diretamente abordada por Silveira Neto, Duarte e Páez (2015), que estudaram se a variação no tempo de deslocamento é devido à HRH, analisando o caso da Região Metropolitana de São Paulo (RMSP). Os autores encontram que o estado civil exerce um papel significativo, constituindo em um bom preditor para o tempo de *commuting* feminino. Contudo, a hipótese de responsabilidade doméstica teve seu papel atenuado, à medida em que variáveis como a presença de suporte familiar (irmãos mais velhos, tios ou avós) e o aumento da disponibilidade de creches, influenciando não somente no tempo de deslocamento, mas, consequentemente, na possibilidade de locomoção aos centros de emprego mais distantes, elevando o raio de vizinhança de acesso da mulher.

Pereira e Schwanen (2015) mostram que, apesar da diferença no tempo de *commuting* por gênero entre as regiões metropolitanas do Brasil estar diminuindo, ainda é persistente por nível de renda, no qual famílias mais pobres se deslocam mais; além da alta variabilidade por região. Para os autores, essa queda na diferença por gênero é consistente com as mudanças sociodemográficas que vem ocorrendo no país, a exemplo da queda na taxa de natalidade, da alteração na composição familiar e do aumento no nível de educação das mulheres, fatores que melhoraram o acesso ao trabalho e, por sua vez, o aumento no tempo de deslocamento feminino. Esse efeito, potencializou o estabelecimento de moradias em áreas periféricas, cujo preço é mais acessível, e, além de afastar as famílias do centro de emprego, também constitui em um movimento pendular maior dos trabalhadores e consequente aumento do tempo médio de deslocamento, devido ao congestionamento de vias. Esse último ponto também foi abordado por Haddad e Vieira (2015) e Barbosa e Silveira Neto (2017), que destacam a relação entre o deslocamento de trabalhadores como forma de se captar o efeito na produtividade de cidades e sua conexão com o desenho urbano.

Apesar da queda no gap por gênero na empregabilidade documentada por Pereira e Schwanen (2015), a diferença do tempo de deslocamento ainda persiste como um problema urbano crônico em muitos dos países em desenvolvimento, penalizando mais intensamente mulheres casadas e com baixa escolaridade. Tais evidências foram expostas por Barufi e Haddad (2017), que conectaram a teoria do mismatch espacial com a dinâmica do mercado de trabalho no Brasil. Os autores mostram que o longo trajeto casa-trabalho afeta negativamente o nível de salário, mas não é significante para estimar a probabilidade de estar empregado. Outra evidência importante é que como a taxa de urbanização das cidades não foi acompanhada pelo crescimento do Produto Interno per capta (PIB_pc), os trabalhadores menos escolarizados se mantém distantes dos centros de emprego, dificultando

o acesso ao emprego qualificado desse grupo.

Associado ao aumento no adensamento dos centros urbanos está o aumento no nível de violência urbana. Segundo dados do IPEA (2018), o Brasil é um dos países mais violentos, em especial nas Regiões Metropolitanas. Silveira Neto e Moura (2019) documentaram que o crescimento na insegurança está associado ao problema da mobilidade urbana. Os autores ainda expõe que o aumento no tempo de exposição nos espaços públicos (maior tempo de deslocamento), eleva a probabilidade em ser vítima de violência urbana. Mais especificamente, ao aumentar o tempo de deslocamento, há um elevação na probabilidade (cerca de 27%) de ser vítima de assaltos. E essa probabilidade é maior quando condicionada ao indivíduo ser do gênero feminino.

Adicionalmente, no Brasil, o fato de ser mulher configura uma característica significativa na oferta de trabalho. Segundo Tedesco e Souza (2020), as mulheres têm menor probabilidade de ofertar trabalho, especialmente quando há iteração com o fato de possuir filhos pequenos e ser menos escolarizada. Outro achado importante das autores foi que, mesmo tendo aumentado a participação feminina no mercado de trabalho, a diferença por gênero permanecem significativas, assim como a desproporcionalidade das divisões de trabalho intra-familiares.

Percebe-se assim que há uma corpo de pesquisas que evidenciam um efeito mais intenso do longo *commuting* nas mulheres, em especial nas casadas. Esse artigo buscou observar esse padrão para o caso da RMSP, principal metropolitana brasileira e analisar o padrão da empregabilidade feminina, assim como o papel do tempo de deslocamento casa-trabalho na participação feminina no mercado de trabalho.

2.3 Estratégia Empírica

O problema central desse ensaio é investigar se a distância ao centro de empregos afeta a empregabilidade feminina. Para isto estima-se o seguinte modelo:

$$y_g = \beta CT_g + X_g'\alpha + \epsilon_g \tag{2.1}$$

As variáveis que representam a empregabilidade feminina y_g definidas no escopo deste estudo são a participação na força de trabalho e a taxa de desemprego. Adicionalmente, a mulheres foram divididas em quatro agrupamentos distintos: solteiras, com filho e sem filhos, e casadas, com e sem filhos. A unidade de observação dentro da RMSP foi a área de ponderação, que se caracteriza por um agrupamento de setores censitários, para a aplicação dos procedimentos de calibração das estimativas com as informações conhecidas para a população como um todo (CENSO, 2010; BARBOSA; SILVEIRA NETO, 2017).

X_g representa um conjunto de variáveis da área de ponderação: renda local, preço médio da moradia, percentual de famílias que possuem crianças em creche, percentual

de famílias que possuem mais de duas crianças, percentual de famílias que residem em domicílios subnormais, distância até o centróide da RMSP[2].

O primeiro exercício empírico foi estimar a dependência da empregabilidade feiminina via OLS. Contudo, dois potenciais problemas precisam ser analisados antes inferir conclusões a partir desse tipo de estimação. Faz-se necessário checar se existe autocorrelação espacial, isto é, se a empregabilidade das mulheres está relacionada com a oferta de trabalho das unidades vizinhas. Além disso, em função da escolha simultânea entre local de moradia e a decisão de participar do mercado de trabalho, é preciso analisar a potencial endogenia entre as variáveis principais. Segundo LeSage e Pace (2010), desconsiderar esses passos pode levar a estimadores ineficientes e consequente viés na estimação.

2.3.1 Análise da Dependência Espacial

Para obter o padrão espacial e a detecção da presença de *clusters* espaciais, foi utilizado a estatística I de Moran (MORAN, 1950). Caso a hipótese de dependência espacial seja significante, o passo seguinte é modelar a relação entre a empregabilidade feminina e o tempo de *commuting* considerando a presença de autocorrelação espacial (ANSELIN et al., 2001). A hipótese nula do Índice de Moran (I) é que a relação é randomicamente distribuída no espaço, i.e., inexiste autocorrelação espacial. Caso contrário, e $I > 0$, significa que há um *cluster* espacial; se $I < 0$, dispersão espacial.

LeSage e Pace (2009) sugerem iniciar a escolha do modelo mais apropriado para a correção da dependência espacial pelo *spatial Durbin model - SDM*. O modelo de correção espacial de Durbin admite a existência de dependência espacial esteja presente tanto na variável dependente como no erro, isto é, $y = \rho W_y + X\beta + \epsilon$, $\epsilon = \lambda W\epsilon + u$, onde ρ e λ são os parâmetros que captam, a presença de autocorrelação espacial; W é a matriz de pesos espaciais e X é a matriz de variáveis explicativas à variável de interesse y. Os autores recomendam ainda que, a partir dessa estimação, deve-se avaliar o ganho em restringir o modelo para *spatial lag model - SLM*, que considera a dependência espacial apenas na variável dependente. A forma funcional do SLM é $y = \rho W_y + X\beta + \epsilon$. Adicionalmente, Anselin et al. (2001) recomenda a escolha do melhor modelo via Critério de Informação de Akaike (AIC).

Considerando a região de estudo desse ensaio, a escolha do modelo SDM para a correção espacial é também justificada pela correlação entre a participação feminina no mercado de trabalho com variáveis locais, tais como a própria participação da mulher no mercado de trabalho e a taxa de desemprego em áreas vizinhas e que exercem um papel importante na explicação da dinâmica espacial da oferta de trabalho feminina (SILVEIRA NETO; DUARTE; PÁEZ, 2015). Por fim, Vega e Elhorst (2017) e Kawabata e Abe (2018) expõem que o uso da correção da dependência espacial via SDM é eficiente

[2] Ver (BARUFI; HADDAD, 2017)

quando existe iteração com variáveis endógenas ou impactos espaciais de vizinhança (*spillovers effects*).

Conforme Elhorst (2010) a estimação via SDM pode ser redefinida de forma a reescrever ρ como um índice para as variáveis independentes, tal que:

$$y = (1 - \rho W)^{-1} X\beta + (1 - \rho W)^{-1}\epsilon \tag{2.2}$$
$$y = \sum_{r=1}^{k} (1 - \rho W)^{-1}\beta_r X_r + (1 - \rho W)^{-1}\epsilon$$
$$y = \sum_{r=1}^{k} S_r W X_r + (1 - \rho W)^{-1}\epsilon$$

.

Kawabata e Abe (2018) explicita que a presença dos *spillovers* na variável defasada, ou seja, o efeito da variável X_r para um indivíduo i não é dado por β_r, mas sim por $S_r(W)_{ii}$. Com isso, se faz necessário estimar os efeitos marginais, divididos em direto, indireto e total. O impacto direto tem interpretação equivalente aos estimadores do modelo via OLS, e mede o efeito médio da observação i sobre ela mesma. Já o impacto total, avalia o efeito médio de todas as demais variáveis j contidas na matriz de contiguidade sobre a observação i. O indireto é a diferença entre o total e o direto, estimando o efeito nas áreas circunvizinhas à obervação i.

2.3.2 Correção da Potencial Endogeneidade

A endogeneidade promovida pelo termo de defasa espacial e que está contida nos modelos espaciais, inviabiliza a estimação via OLS. Em tais condições, a estimação via Máxima Verosimilhança (ML) é normalmente utilizada, contudo requer que os resíduos sejam normalmente distribuídos (SEABRA; SILVEIRA NETO; MENEZES, 2016). Afim de estimar a relação do tempo de *commuting* na empregabilidade feminina e minimizar o potencial viés contido na determinação simultânea da localização de moradia e das firmas, estimou-se em comparação ao modelo via OLS, o SDM via dois estágios, como sugerem Kelejian e Prucha (1997), na qual a estimação ocorre via *spatial two-stages least squares* (STSLS), e cuja hipótese da normalidade dos resíduos não é requerida e fornece estimadores consistentes, não viesados e assintoticamente normais (BIVAND; PIRAS, 2015; BETZ; COOK; HOLLENBACH, 2020; BIVAND; MILLO; PIRAS, 2021). A escolha da melhor estimação ocorreu via critério de informação de Akaike, conforme recomendam Anselin et al. (2001) e Bivand e Piras (2015).

O viés da escolha simultânea cresce haja vista que a empregabilidade feminina depende tanto da localização do emprego como da moradia, afetando o custo da habitação. O custo de moradia, por sua vez, afeta a escolha do local de moradia e também o acesso ao emprego, pois o longo tempo de deslocamento restringe a disponibilidade de participar do mercado de trabalho, em especial para o caso feminino. Logo, o tempo de *commuting*, nossa

variável preditora, depende de fatores que afetam o acesso ao mercado de trabalho, mas que também podem influenciar o próprio tempo de deslocamento casa-trabalho. Assim, observar o efeito de um choque no tempo de *commuting* como canal de explicação causal no acesso ao emprego requer que tal variação seja exógena. Para lidar com esse ponto, e obter o estimador proposto por Kelejian e Prucha (1997), instrumentalizou-se o tempo de *commuting* pela distância ao centro histórico da RMSP, demarcado pela localização do Colégio Pateo, primeira escola construída pelos Jesuítas em São Paulo, porém percorrida via margem dos rios, que atualmente está submersa pelas vias urbanas construídas ao longo da RMSP. Essa variável foi exposta por Haddad e Barufi (2017), a qual lidam com a potencial endogeneidade entre o índice de acessibilidade de emprego e nível de salários. A Figura 2 caracteriza a localização no mapa da RMSP.

Figura 1 – Primeiro Mapa de São Paulo

Fonte: São Paulo Antigo; Plantas da Cidade. Elaboração por Haddad e Barufi, 2017. Mapa da Região Metropolitana de São Paulo, com destaque para o Colégio Pateo, caracterizado como primeiro mapa da região e que estabelece o centróide da área metropolitana.

Haddad e Barufi (2017) mostram que a cidade de São Paulo sofreu um processo de expansão urbana em torno desse ponto, hoje conhecido como Praça da Sé, e nomeado como "marco zero" da cidade. A escolha desse instrumento foi motivada pela expansão rodoviária possuir relação com a malha fluvial e que, atualmente, estão submersos pelo asfalto. Com isso, é factível supor a exogeneidade do instrumento.

Para o estabelecimento dos pesos espaciais, Elhorst (2010) retrata que a escolha da matriz W é a maior fragilidade dos modelos espaciais, uma vez que W não pode ser estimada, mas sim especificada a priori. Por outro lado, apesar da escolha equivocada da matriz W possa distorcer os coeficientes estimados, essa probabilidade converge a zero se a dependência espacial for realmente significante. Para esse ensaio, a matriz W utilizada foi baseada no critério *Rook*, seguindo a estratégia adotada por Kawabata e Abe (2018). Vale

destacar que a critério de robustez, foram testadas duas outras especificações de pesos, critério *Queen* e a distância euclidiana. Os resultados não sofreram alterações relevantes e, por conveniência, foram omitidos nesse ensaio.

2.4 Base de dados e Região de Estudo

Segundo o Censo (2010), foram entrevistados 67,6 milhões de domicílios em todos os 5.656 municípios brasileiros. Na Região Metropolitana de São Paulo (RMSP), foco analítico desse ensaio, foram entrevistados 19,7 milhões de pessoas, das quais 52,1% eram do sexo feminino, distribuídas em 39 municípios e 5 mesorregiões. A RMSP possui 633 áreas de ponderação, que são distribuídas por todos os 39 municípios pertencentes à região. A amostra de trabalho foi realizada para as mulheres entre 18 e 54 anos, intervalo um pouco distinto da estratégia empírica utilizada por Black, Kolesnikova e Taylor (2014) e Kawabata e Abe (2018). Tedesco e Souza (2020) reforça o ponto do aumento da amplitude da idade em que a presença de filhos afeta a oferta de trabalho feminino. Se por um lado está havendo um aumento da idade média para a maternidade, em 2000 a idade média era 26,8 anos, atualmente é de 29,1 anos, por outro mais mulheres ingressam cada vez mais cedo no mercado de trabalho, especialmente no coorte de atividades econômicas mais desenvolvido.

Além do fato da RMSP possuir um longo *commuting* (DUARTE, 2020), o que já torna a escolha dessa região como alvo de estudo válida, justifica-se a decisão em avaliar a relação do tempo de deslocamento casa-trabalho no mercado de trabalho feminino na RMSP também por duas razões principais. Primeiro, a metropolitana de São Paulo representa 20% do PIB brasileiro e é o principal centro produtivo do país, colocando-a como extremamente relevante na perspectiva econômica, além de possuir similaridades com outras metrópoles no mundo (HADDAD; BARUFI, 2017). Em segundo lugar, sua estrutura urbana e populacional é marcada pelo crescimento urbano elevado e significativa variabilidade no efeito da mobilidade na dinâmica de acesso a oportunidades de emprego (PEREIRA; SCHWANEN, 2015; HADDAD; VIEIRA, 2015; BARUFI; HADDAD, 2017; DUARTE, 2020). Tal cenário permite comparabilidade dos resultados encontrados tanto com outros estudos que abordaram essa problemática na própria região, a exemplo de Silveira Neto, Duarte e Páez (2015), Barufi e Haddad (2017) e Duarte (2020), bem como a diferentes metrópoles nas quais já foram alvo de investigação e que consideraram o papel do *commuting* na análise do mercado de trabalho feminino, a exemplo de Tóquio em Kawabata e Abe (2018) e de Nova Iorque em Black, Kolesnikova e Taylor (2014) e Farré, Jofre-Monseny e Torrecillas (2020).

Como consequência da dinâmica econômica, a RMSP possui intenso fluxo pendular intra-municípios e concentra boa parte da atividade econômica no município de São Paulo. Haddad e Vieira (2015) mostraram que, em São Paulo, 95,6% das ocupações ocorrem no próprio município, sendo as regiões circunvizinhas classificadas como cidades dormitórios.

À medida em que a distância aumenta, a cidade de São Paulo perde relevância na ocupação. Por fim, os autores ainda enfatizam que 51% dos ocupados se deslocam de outros municípios para trabalhar em São Paulo e outros 10,5% se deslocam de São Paulo para trabalhar em outros municípios, demonstrando relevante fluxo pendular no deslocamento casa-trabalho.

Para prover uma nova perspectiva sob a influência espacial da mobilidade na dinâmica do mercado de trabalho, aborda-se essa relação utilizando áreas de ponderação como unidade espacial. O uso dessa agregação permite obter significância estatística, uma vez que há possibilidade de calibração por pesos amostrais. Em outras palavras, as áreas de ponderação são a menor unidade geográfica com possibilidade de obter representatividade estatística a partir das amostras do Censo Demográfico (BARBOSA; SILVEIRA NETO, 2017). As áreas de ponderação são compostas pela agregação de domicílios dentro de setores censitários contíguos contidos em um único município. Sendo o mínimo de 400, e no máximo 800, domicílios particulares ocupados. Tal abordagem permite visualizar associações interessantes e ainda pouco exploradas.

A Tabela 1 descreve as variáveis utilizadas nesse estudo. Utiliza-se na estimação empírica como variável dependente para explicar a empregabilidade feminina, a participação na Força de Trabalho das mulheres e a taxa de desemprego, ambas por área de ponderação e obtidas a partir dos microdados do Censo 2010. Analisa-se, ainda, os subconjuntos de mulheres considerando seu estado civil e a presença de filhos pequenos (menores de 4 anos). A principal variável explicativa é o tempo de deslocamento casa-trabalho para os homens, devido a hipótese de que, na presença de um longo *commuting*, os homens são menos afetados em termos de dinâmica de trabalho, continuando a acessar e, consequentemente, se deslocar aos CBD; enquanto que as mulheres restringem a mobilidade. Portanto, o tempo de deslocamento masculino se caracteriza como um boa *proxy* do *commuting* local (CRANE, 2007; ALESINA; GIULIANO; NUNN, 2013; BLACK; KOLESNIKOVA; TAYLOR, 2014; HADDAD; BARUFI, 2017; KAWABATA; ABE, 2018).

Tabela 1 – Descrição das variáveis

Variáveis Dependentes	Descrição
Participação na Força de Trabalho	Proporção de mulheres[1] que participação da força de trabalho local
Taxa de desemprego	Taxa média de desemprego[2] feminino local
Variáveis Independentes	
Tempo de deslocamento (homem)	Tempo médio de deslocamento local para os homens
Renda local (ln)	Renda média domiciliar (em log)
Preço moradia local (ln)	Preço médio aluguéis residenciais (em log)
Creches	% de famílias que possuem crianças matriculadas em creches
HH2Ch	% de famílias que possuem 2 ou mais crianças
Subnormais	% de famílias que residem em domicílios subnormais
$DistCBD$	Log da distância da AP até o centróide da RMSP[3]
Notas:	[1] FT equivale a relação entre PEA e o total de mulheres entre 18 e 54 anos. [2] Taxa de desemprego equivale a relação entre desocupados e PEA.

A variável de tempo de deslocamento casa-trabalho é disponibilizada no Censo

(2010) como categórica, sendo as frequências marcadas tais como: $F1$- Até 5 minutos; $F2$- De 6 a 30 minutos; $F3$- De 31 a 60 minutos; $F4$- Entre 1 e 2 horas; e $F5$- Mais de 2 horas. Para o propósito desse estudo a variável foi reconstruída como numérica e adotou-se a regra proposta por (BARBOSA; SILVEIRA NETO, 2017).

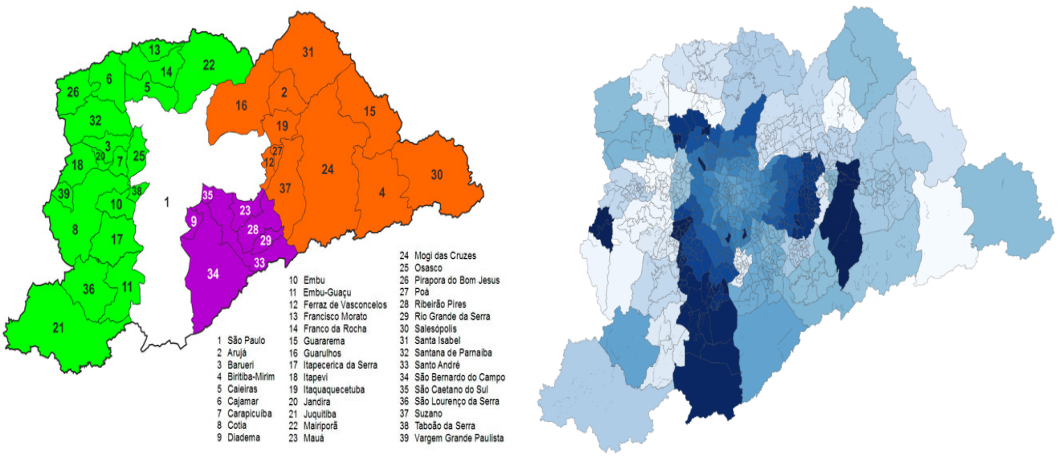
As variáveis de controle selecionadas foram a Renda local (ln), que captura a renda média domiciliar per capita, pois conforme Barufi e Haddad (2017), a condição econômica possui um padrão espacial relevante na determinação da participação feminina no mercado de trabalho, evidenciando o efeito *spillover* (FOGLI; VELDKAMP, 2011). Segundo Ong e Miller (2005) também é importante controlar para alguma variável que representa o efeito de custo de vida ou moradia, para isso utiliza-se a proporção de imóveis alugados na área de ponderação, que mapeia a distribuição da forma de moradia, dado que quanto maior for a proporção de habitações alugadas, mais fácil a mudança de local. Conforme Seabra, Silveira Neto e Menezes (2016), há uma relação positiva entre o espaço e a decisão de moradia, sendo intensificada em regiões cuja mobilidade é precária (HADDAD; VIEIRA, 2015) e (DURANTON; TURNER, 2012).

Seguindo a estrutura empírica documentada em Kawabata e Abe (2018), também são consideradas como variáveis explicativas, o percentual de famílias com duas ou mais crianças no domicílio e o percentual de crianças com até 4 anos que estão matriculadas em creches. Tais variáveis são importantes devido ao suporte familiar fornecido, principalmente às mulheres que possuem filhos pequenos, exercendo um papel positivo na decisão de oferta de trabalho (SILVEIRA NETO; DUARTE; PÁEZ, 2015).

Adicionalmente, com o intuito de observar características mais representativas da própria RMSP, inclui-se o percentual de domicílios subnormais, para captar a condição da moradia e das condições de habitabilidade, haja vista importância da acessibilidade para a decisão de emprego (DUARTE, 2020). E, por fim, para lidar com o viés estabelecido pela escolha simultânea entre a determinação do local de moradia e a localização de firmas, adotou-se, como variável exógena, o instrumento utilizado por Haddad e Barufi (2017), representado pela distância entre a área de ponderação e o centróide do emprego da RMSP (antigo Colégio Pateo). A medida foi calculada seguindo a metodologia dos autores, obtido pela distância (em log) seguindo a rota pelos rios atualmente submersos nas rodovias.

A Figura 2 ilustra a distribuição espacial dos municípios da Região Metropolitana de São Paulo, composta por 39 municípios e 633 áreas de ponderação. A capital São Paulo concentra o centro de atividades e emprego, assim como a maior taxa de concentração de empresas. Além disso, Haddad e Vieira (2015) mostram que existe um intenso fluxo de mobilidade de pessoas que residem fora do centro, mas realizam deslocamento diariamente, dado a caracterização de cidades dormitórios das cidades que orbitam o centro de São Paulo.

Figura 2 – Mapa dos Municípios da RMSP



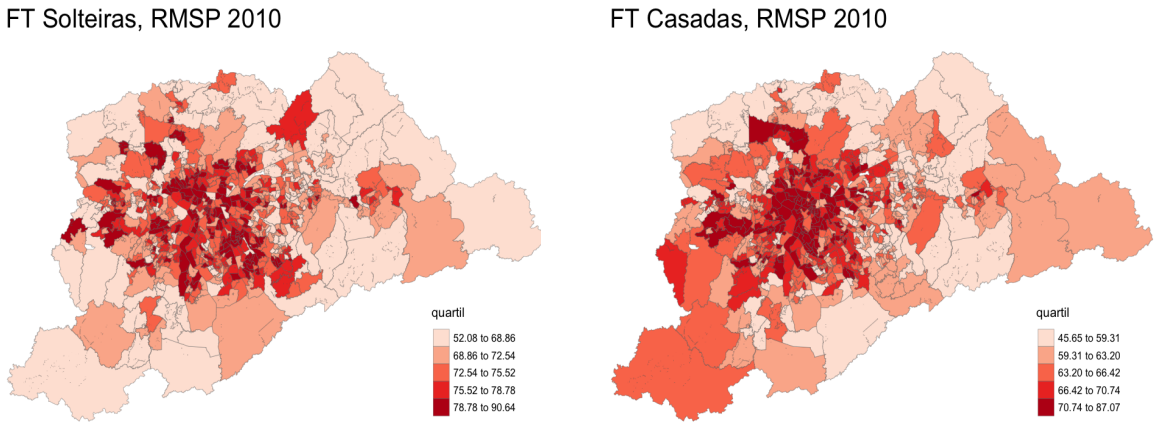
Fonte: Fundação Seade. Mapa da Região Metropolitana de São Paulo e respectivas áreas de ponderação.

2.5 Resultados

2.5.1 Análise do Padrão Espacial da Empregabilidade Feminina

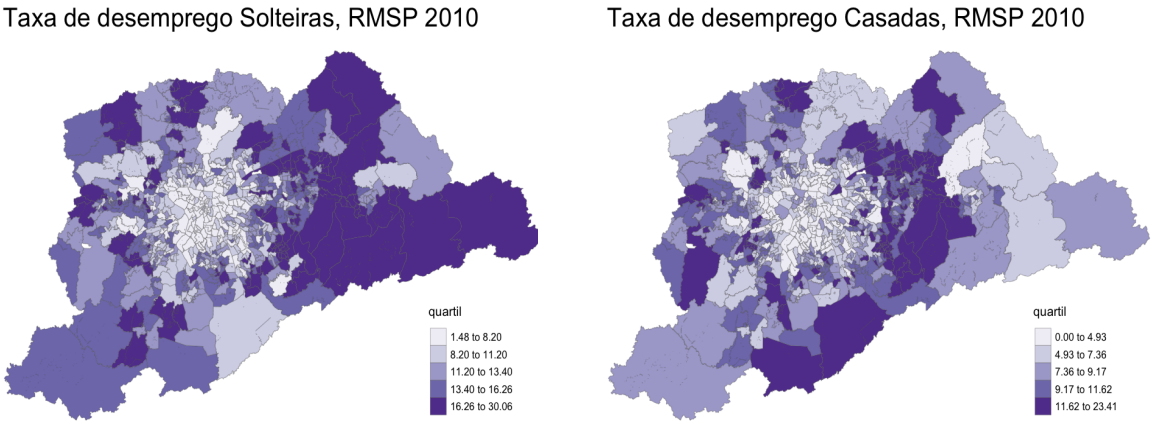
A figura 3 ilustra a distribuição da participação na força de trabalho feminina na região de estudo, verifica-se um padrão espacial distinto na oferta de trabalho, tanto para as mulheres solteiras quanto para as mulheres casadas. A concentração das mulheres solteiras orbita a região central da metropolitana de São Paulo, denotando maior inclinação desse grupo em estar mais próximo da área mais dinâmica de emprego. Ou seja, longos trajetos aparentam não serem desejáveis a este grupo. Já para o grupo das mulheres casadas, é possível observar um maior nível de espraiamento, demonstrando que o efeito do espaço pode ser uma variável relevante na decisão de moradia e, com isso, a criação de *clusters* espaciais de emprego fora da área central.

Figura 3 – Densidade da Força de Trabalho Feminina



Fonte: IBGE. Elaboração própria. Figura à esquerda ilustra a distribuição da força de trabalho das mulheres solteiras. A figura à direita, a distribuição da força de trabalho das mulheres casadas.

Figura 4 – Taxa de desemprego Feminina



Fonte: IBGE. Elaboração própria. Figura à esquerda ilustra a distribuição do desemprego das mulheres solteiras. A figura à direita, a distribuição do desemprego das mulheres casadas.

Por outro lado, para o caso da distribuição espacial do desemprego das mulheres, observado na figura 4, é possível ver que a taxa de desemprego é maior nas regiões periféricas para o caso das mulheres solteiras, em comparação com as casadas.

Assim, a primeira evidência a ser observada nesse estudo foi a diferença[4] entre a média de deslocamento casa-trabalho entre homens e mulheres na RMSP, exposta na Tabela 2. É possível ver que, em conformidade com a literatura que discute as diferenças de mobilidade entre gênero, na RMSP os homens se deslocam mais do que as mulheres. Em média, os homens gastam 49,4 minutos no percurso casa-trabalho, já as mulheres 45,9 minutos. Como comparação, os resultados são similares ao documentado em Kawabata e Abe (2018), no qual os homens gastam em média 49,9 minutos e as mulheres 38,8 minutos.

Tabela 2 – Sumário Descritivo das Variáveis Dependentes

	Média	Mediana	Desvio Padrão	Min.	Max.
Tempo de deslocamento					
Homem	49,4	48	10,9	25	85
Mulher	45,9	46	11,3	25	87
Participação Força de Trabalho					
Solteira sem filhos	75,8	76,2	6,5	48,8	90,8
Solteira com filhos	65,8	65,3	10,6	23,3	100
Casada sem filhos	62,8	63,1	9,6	32,2	89,8
Casada com filhos	74,8	76,3	11,1	28,4	100
Taxa de desemprego					
Solteira sem filhos	12,8	12,3	8,3	1,8	37,8
Solteira com filhos	17,3	16,9	5,6	0,2	58,2
Casada sem filhos	11,4	10,5	14,1	0,1	39,5
Casada com filhos	8,4	7,0	10,4	0,08	40,7

Fonte: IBGE/Elaboração própria. O tempo de deslocamento equivale ao tempo gasto (em min) no deslocamento casa-trabalho. Número de unidades espaciais foi de 633 para a RMSP, e a amostra foi constituída para homens e mulheres entre 18 e 54 anos.

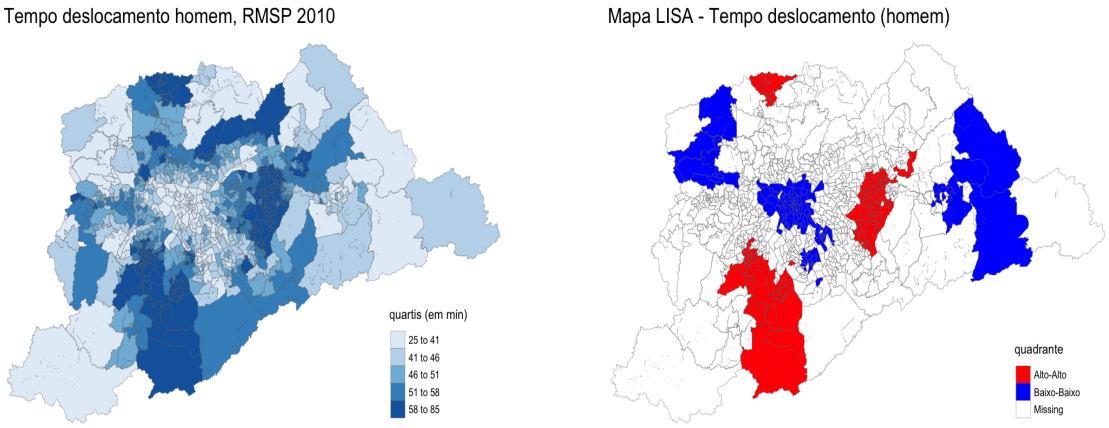
[4] Realizou-se um teste de médias e existe diferença estatisticamente significante, t = 22.38, df = 1129, p-value = 2.2e-16

Quando se observa a taxa de participação feminina no mercado de trabalho, há uma diferença clara quando se considera o estado civil e a presença de filhos. É possível ver que mulheres solteiras têm uma taxa de participação média de 75,8%, e tal percentual se reduz ao observar, dentre as solteiras, a presença de filhos (65,8%). À medida que é considerado o estado civil (se é casada) e, mais ainda, com a presença de filhos (casada e com filho). Respectivamente, a participação na força de trabalho são de 62,8% e 74,8%, para as mulheres casadas sem filhos e casadas com filhos. Tal como foi documentado por Silveira Neto, Duarte e Páez (2015), representa evidencia do peso da responsabilidade doméstica mais intenso sobre as mulheres na presença de filhos.

A taxa de desemprego também revela padrões distintos, considerando o estado civil e a presença de filhos. É possível notar, dentre as solteiras, que a presença de filhos exerce um papel limitador, reduzindo a possibilidade de empregabilidade desse grupo, o que implica em uma taxa de desemprego maior, 17,3%, quando comparado às solteiras sem filhos (12,8%). Já entre as mulheres casadas, o desemprego é mais intenso às mulheres sem filhos. A presença de um suporte familiar ou pelo acesso a creches, podem ser uma explicação para esse padrão distinto.

A Figura 5 revela a densidade do tempo de deslocamento para os homens na RMSP, assim como os *clusters* espaciais de alto valor (*hot spots*) e de baixo valor (*cold spots*), baseado na estatística de Getis-Ord. O tempo de deslocamento dos homens varia acentuadamente, revelando a existência de disparidade local, e se apresentando mais intenso nas regiões periféricas ao CBD. Fato compatível com o que documentam Haddad e Vieira (2015) e Barufi e Haddad (2017), que documentam a presença de um contingente de cidades dormitórios que orbitam o centro da RMSP, implicando em um movimento pendular intenso entre os municípios. Adicionalmente, também é possível notar que a concentração dos *clusters* de alto valor é localizada nos municípios mais periféricos, alguns em zonas mais distantes do CBD. É possível ver, inclusive, a sobreposição com as áreas de maior tempo de deslocamento, induzindo a criação de um *cluster* de empregos nessas regiões, ou seja, os homens trabalham e residem nesses mesmos municípios, acentuando o *commuting* nessas localidades.

Figura 5 – Tempo de deslocamento para os homens

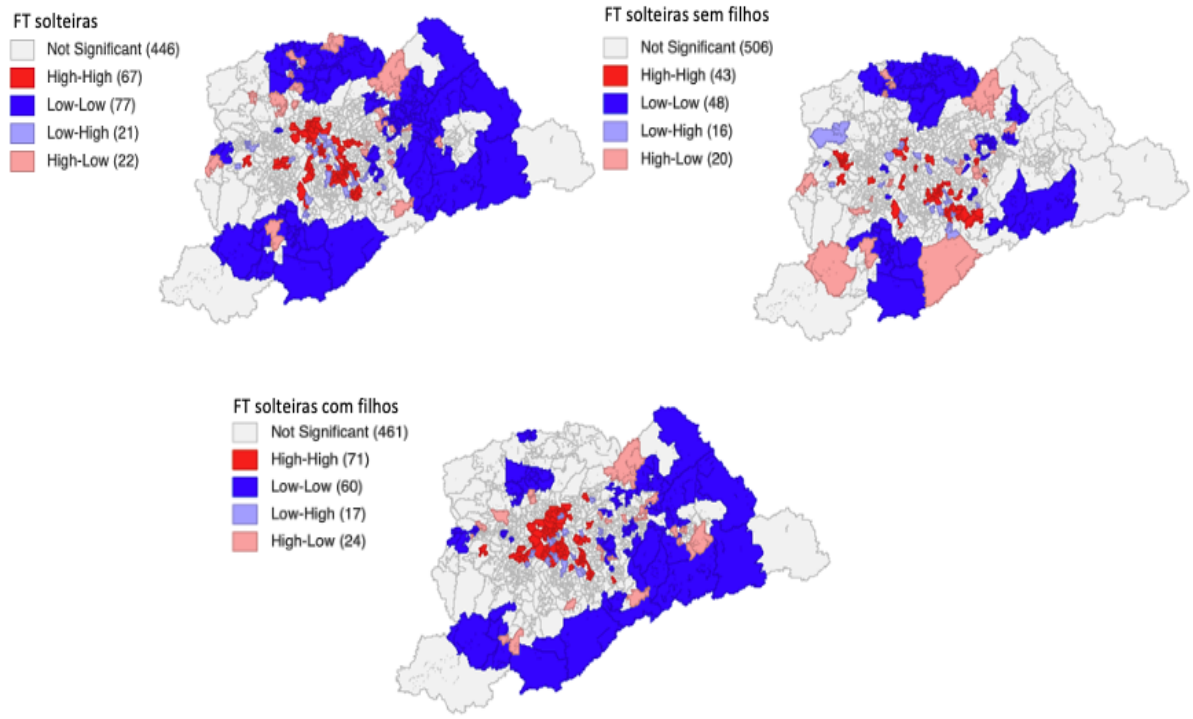


Fonte: IBGE. Elaboração própria. Figura 1 (à esquerda), expõe o Tempo médio de deslocamento casa-trabalho por município da RMSP. A Figura 2 (à direita) ilustra os *clusters* de alto (hot) e baixo (cold) valor na RMSP.

O padrão visualizado nas Figuras 6 e 7 expõem os mapas com os valores do Índice Local de Associação Espacial (LISA) e mostram que a distribuição da força de trabalho feminina é marcadamente diferente quando consideramos o estado civil e a presença de filhos. Nota-se que as mulheres solteiras, apesar de intensamente mais concentradas na região central (hot spots), estão mais espalhadas em torno do CBD, em especial na presença de filhos. Percebe-se que a presença de filhos induz a se aproximar do CBD, fato distinto do observado por Kawabata e Abe (2018), que revelou mais distanciamento àquelas com filhos pequenos.

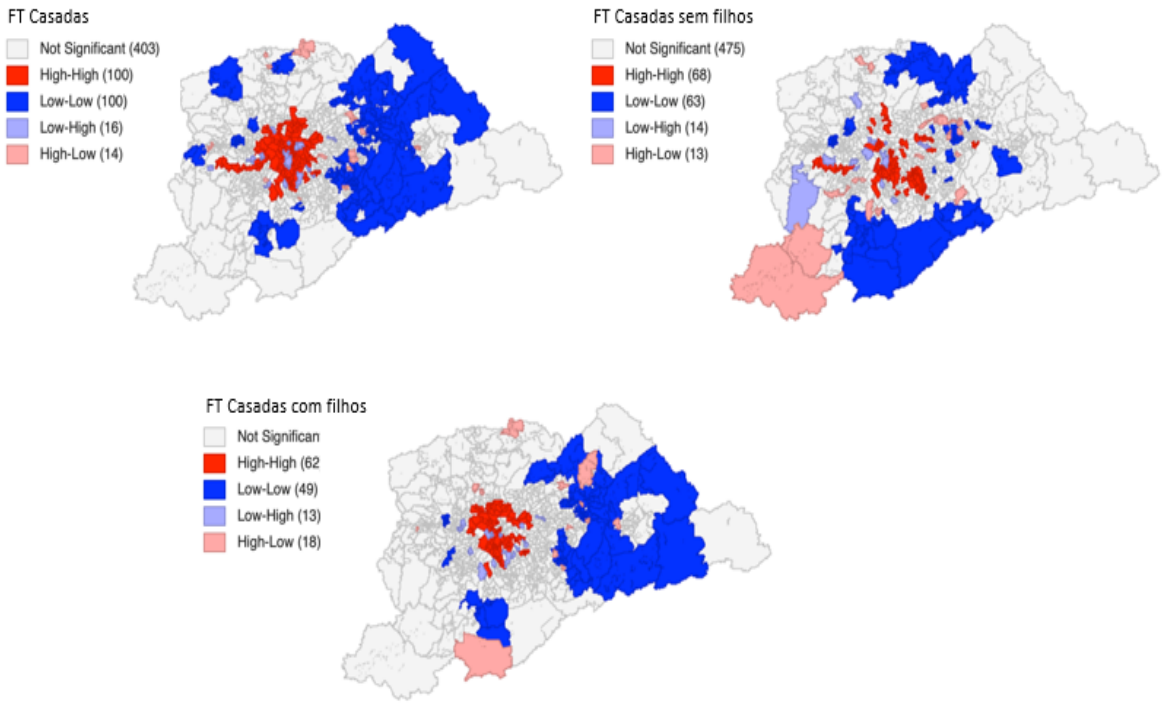
A análise da Figura 7 mostra padrão diferente tanto na comparação com as solteiras, assim como em relação ao estado civil e a presença de filhos. É possível ver que, para esse grupo, a presença de filhos exerce um papel mais relevante na definição dos *clusters* espaciais, com o surgimento de áreas intensas mais distantes do CBD. Já no caso das casadas sem filhos, percebe-se que os *clusters* espaciais de alto valor são menos frequentes e mais próximos ao CBD, e que os *clusters* de baixo valor estão situados nas regiões mais periféricas. Esse resultado é similar ao visualizado por Kawabata e Abe (2018). Tal distribuição indica um possível impacto do *commuting* mais suave nesse grupo, talvez explicado pela hipótese da responsabilidade familiar.

Figura 6 – Mapa LISA da Força de Trabalho Feminina - Solteiras



Fonte: IBGE. Elaboração própria. Figura à esquerda ilustra o Índice Local de Associação Espacial (LISA) da força de trabalho das mulheres solteiras. Figura à direita, o LISA da força de trabalho das mulheres solteiras sem filhos. E, por fim, a figura abaixo, o LISA da força de trabalho das mulheres solteiras com filhos menores de 4 anos.

Figura 7 – Mapa LISA da Força de Trabalho Feminina - Casadas

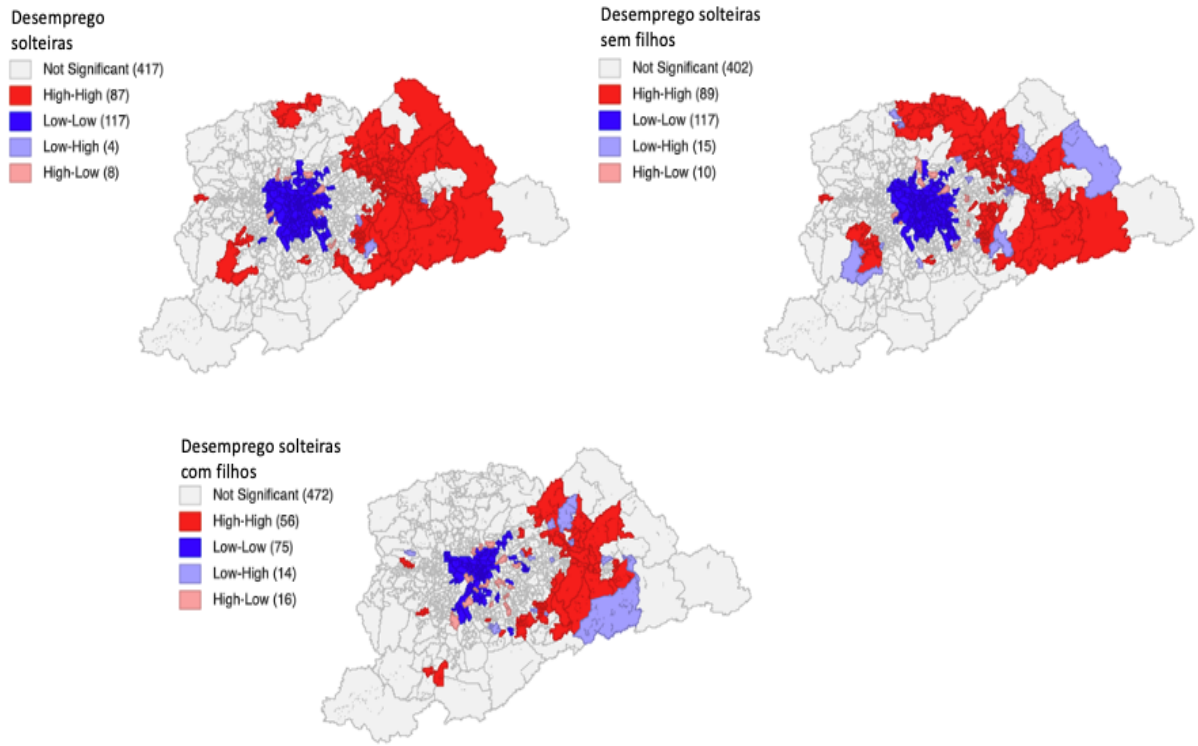


Fonte: IBGE. Elaboração própria. Figura à esquerda ilustra o Índice Local de Associação Espacial (LISA) da força de trabalho das mulheres casadas. Figura à direita, o LISA da força de trabalho das mulheres casadas sem filhos. E, por fim, a figura abaixo, o LISA da força de trabalho das mulheres casadas com filhos menores de 4 anos.

A análise da significância dos *clusters*, de alto e baixo valor, para a participação feminina na força de trabalho evidenciam que testar para existência de autocorrelação é um fator importante, haja vista iteração com os atributos dos vizinhos expostos nas figuras 6 e 7. Portanto, para analisar o padrão de empregabilidade feminina na RM de São Paulo se faz necessário verificar a necessidade de corrigir o viés espacial. O argumento intuitivo é que a dinâmica econômica local tem efeito na participação feminina no mercado de trabalho, sendo a estrutura de deslocamento nas cidades espacial-dependente, conforme documentado por Silveira Neto, Duarte e Páez (2015).

O padrão espacial da taxa de desemprego também se revelou não uniforme, seja na comparação entre o estado civil seja pela presença de filhos. Para o caso das mulheres solteiras, observa-se *clusters* espaciais mais intensos nas regiões periféricas, tanto para as solteiras sem filhos como para as com filhos, sendo esse último grupo com maior número de *clusters* de alto valor. Esse padrão expõe que a presença de filhos, para o caso de solteiras, penaliza mais aquelas mais distantes da região central.

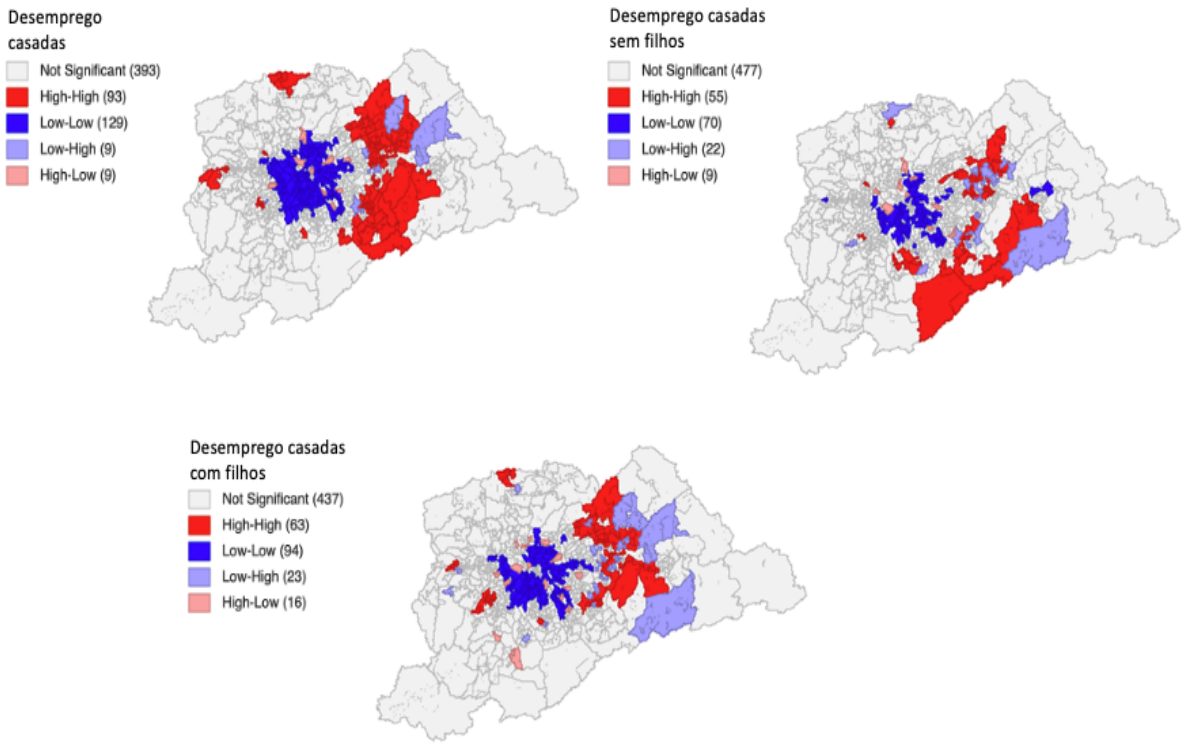
Figura 8 – Mapa LISA do Desemprego Feminino - Solteiras



Fonte: IBGE. Elaboração própria. Figura à esquerda ilustra o Índice Local de Associação Espacial (LISA) da taxa de desemprego das mulheres solteiras. Figura à direita, o LISA das mulheres solteiras sem filhos. E, por fim, a figura abaixo, o LISA das mulheres solteiras com filhos menores de 4 anos.

Para o grupo das casadas, a presença de filhos revela maior intensidade dos *clusters* espaciais fora do CBD, em comparação com as casadas sem filhos. Além disso, maior grupo dos *clusters* de alto valor estão mais distantes do centro de São Paulo, em municípios fronteira da RMSP.

Figura 9 – Mapa LISA do Desemprego Feminino - Casadas



Fonte: IBGE. Elaboração própria. Figura à esquerda ilustra o Índice Local de Associação Espacial (LISA) da taxa de desemprego das mulheres casadas. Figura à direita, o LISA das mulheres casadas sem filhos. E, por fim, a figura abaixo, o LISA das mulheres casadas com filhos menores de 4 anos.

2.5.2 Efeitos do Tempo de Commuting para a Participação na Força de Trabalho

A Tabela 3 reporta o resultado das estimações via OLS e SDM, todos robustos à heterocedasticidade, sendo a variável dependente a participação na força de trabalho feminina, considerando os quatro agrupamentos trabalhados nesse ensaio: solteiras e casadas, com filho e sem filho. A estatística I de Moran (Tabela 4) se mostrou positiva e significante para as todas as variáveis dependentes, o que torna os coeficientes estimados via OLS viesados e, portanto, o modelo com correção espacial deve ser considerado (ANSELIN et al., 2001). O SDM foi estimado, como sugerem LeSage e Pace (2010), via *spatial two stages least square*, seguindo os procedimentos descritos por Kelejian e Prucha (1997) e Bivand e Piras (2015). Para instrumentalizar o tempo de *commmuting* foi utilizada a variável instrumental descrita por Haddad e Barufi (2017).

Tabela 3 – Regressão Participação na Força de Trabalho Feminina

	Solteira sem filho		Solteira com filho		Casada sem filho		Casada com filho	
	OLS	*SDM*	*OLS*	*SDM*	*OLS*	*SDM*	*OLS*	*SDM*
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Tempo de deslocamento (homem)	-0.081^{***}	-0.050	-0.076^{*}	-0.080	-0.081^{**}	-0.093^{***}	-0.236^{***}	-0.201^{**}
	(0.026)	(0.048)	(0.039)	(0.072)	(0.036)	(0.067)	(0.044)	(0.082)
Desemprego do homem	-0.009	-0.028	-0.333^{***}	-0.213^{*}	-0.466^{***}	-0.305^{***}	-0.349^{***}	-0.274^{**}
	(0.074)	(0.081)	(0.112)	(0.121)	(0.104)	(0.113)	(0.127)	(0.137)
Renda pc (log)	0.007	0.003	-0.034	-0.012	-0.007	0.003	0.012	0.012
	(0.023)	(0.024)	(0.034)	(0.036)	(0.032)	(0.034)	(0.039)	(0.041)
Perc. domicílios alugados	-0.121^{***}	-0.136^{***}	-0.014	0.024	0.015	0.016	0.105^{*}	0.048
	(0.036)	(0.043)	(0.054)	(0.065)	(0.050)	(0.060)	(0.061)	(0.073)
Perc. crianças na creche	3.171	3.302	15.559^{***}	15.880^{***}	17.882^{***}	16.379^{***}	1.381	0.267
	(2.437)	(2.866)	(3.679)	(4.312)	(3.417)	(3.995)	(4.173)	(4.867)
Perc. famílias com mais duas crianças	0.009	-0.042	-0.434^{***}	-0.243^{*}	-0.080	0.065	-0.242^{*}	-0.193
	(0.079)	(0.091)	(0.119)	(0.136)	(0.111)	(0.126)	(0.136)	(0.154)
Perc. domicílios subnormais	-0.086^{***}	-0.072^{***}	-0.064^{***}	-0.089^{***}	-0.047^{**}	-0.077^{**}	-0.072^{***}	-0.045
	(0.015)	(0.022)	(0.022)	(0.033)	(0.021)	(0.030)	(0.025)	(0.037)
lag.ct_h		-0.044		0.045		0.049		-0.001
		(0.060)		(0.090)		(0.083)		(0.102)
lag.d_h		0.009		-0.352		-0.601^{***}		-0.197
		(0.147)		(0.223)		(0.210)		(0.252)
lag.ln_rend		-0.042		-0.046		0.001		0.051
		(0.045)		(0.068)		(0.063)		(0.077)
lag.per_alug		0.060		-0.012		0.084		0.160
		(0.070)		(0.104)		(0.096)		(0.118)
lag.creche		-0.138		-13.136^{*}		-10.291		-2.705
		(4.645)		(7.011)		(6.523)		(7.880)
lag.hh2ch		0.122		-0.507^{**}		-0.361		0.003
		(0.158)		(0.240)		(0.220)		(0.269)
lag.subn		-0.010		0.060		0.057		-0.034
		(0.031)		(0.046)		(0.043)		(0.052)
Constant	$87{,}87^{***}$	$73{,}06^{***}$	$80{,}93^{***}$	$89{,}51^{***}$	64.090^{***}	69.280^{***}	90.353^{***}	67.167^{***}
	(3,844)	(8,080)	(5,802)	(10,779)	(5,389)	(9,519)	(6,580)	(11,502)
ρ		0.11		0.46^{***}		0.44^{***}		0.41^{***}
N	633	633	633	633	633	633	633	633
R^2	0.179		0.292		0.260		0.179	
R^2 Adj.	0.170		0.284		0.251		0.170	
Akaike Inf. Crit.	4.088	4.067	4.588	4.581	4.506	4.484	4.746	4.739
LR Test (OLS ou SDM)		4.952^{**}		2.259		0.723		10.537^{***}

Note: *p<0.1; **p<0.05; ***p<0.01

O parâmetro ρ se mostrou significante para para todas as variáveis dependentes que representam a participação na força de trabalho feminina, com exceção do caso das solteiras sem filhos, corroborando com a estatística I de Moran. Adicionalmente, o modelo espacial se mostrou mais adequado, tanto via *Loglikelihood Test (LR) test*, assim como via critério de informação de Akaike (AIC). O modelo via SDM produz estimativas semelhantes à estimação OLS, porém com uma mudança relevante no tocante ao efeito do tempo de deslocamento. Para o caso da estimação via OLS, o tempo de deslocamento se mostrou significante em todos os casos. Por outro lado, ao corrigir para a dependência espacial e instrumentalizar a variável explicativa pela distância da área de ponderação ao centro de São Paulo pelas rotas dos rios, a significância observada no caso das mulheres solteiras deixa de existir, convergindo com os resultados obtidos por Black, Kolesnikova e Taylor (2014) e Kawabata e Abe (2018).

Tabela 4 – Estatística I de Moran - Participação na Força de Trabalho

Participação na Força de Trabalho	Moran I	p-valor
Solteira sem filhos	17.69	0.00
Solteira com filhos	27.53	0.00
Casada sem filhos	24.06	0.00
Casada com filhos	22.48	0.00

Fonte: Elaboração própria. Tabela expõe a Estatística I de Moran. Número de unidades espaciais foi de 633 para a RMSP, e a amostra foi constituída para homens e mulheres entre 18 e 54 anos.

A Tabela 5 reporta os efeitos marginais do tempo de deslocamento masculino sobre a participação feminina na força de trabalho considerando o modelo SDM, estimado via *spatial two stages least square*. Os efeitos totais diferem em relação ao estado civil, se mostrando não significante para o caso das mulheres solteiras. Entre as casadas, se configura muito mais intenso para os caso da presença de filhos. Por exemplo, para as casadas sem filhos, cujo impacto marginal total foi de -0.13, a redução em 10 minutos no tempo de deslocamento equivale a um aumento de $1,43\%$ na participação feminina do mercado de trabalho. Para o caso das casadas com filhos esse impacto é ainda mais intenso, com efeito total estimado em -0.31 pontos percentuais. Logo, 10 minutos a menos no tempo de deslocamento equivale a um incremento de $4,14\%$ na participação feminina. Isso representaria aproximadamente 22.174 mulheres a mais no mercado de trabalho. Essa evidência demonstra a importância do tempo de *commuting* como canal para impactar a empregabilidade feminina. Já o efeito indireto, que mensura o *spillover effect* não se mostrou significativo para nenhum dos grupos trabalhados nesse ensaio.

Tabela 5 – Efeito marginal do Commuting na Participação da Força de Trabalho

	Direto	Indireto	Total	N
Solteira sem filhos	-0.05	-0.01	-0.06	633
	(0.027)	(0.006)	(0.031)	
Solteira com filhos	-0.08	-0.04	-0.12	633
	(0.043)	(0.008)	(0.049)	
Casada sem filhos	-0.09***	-0.04	-0.13***	633
	(0.038)	(0.006)	(0.042)	
Casada com filhos	-0.20***	-0.11	-0.31***	633
	(0.048)	(0.018)	(0.061)	

Fonte: Elaboração própria. *p<0.1; **p<0.05; ***p<0.01.

A tabela 6 reporta os resultados da estimações via OLS e SDM, robustos à heterocedasticidade, para o caso da taxa de desemprego das medidas de empregabilidade observadas nesse ensaio. A estatística I de Moran (Tabela 7) assim como para a participação na força de trabalho, também se mostrou significativa, induzindo a corrigir a ineficiência dos estimadores obtidos via OLS, por um modelo que considere a dependência espacial (ANSELIN et al., 2001). É digno de nota que o SDM foi estimado sob os mesmos procedimentos descritos quando se analisou a participação na força de trabalho.

É possível ver que o parâmetro ρ se mostrou significante em todas as variáveis dependentes observadas. E similarmente ao caso anterior, após a correção espacial, o tempo de deslocamento se manteve significante apenas no caso das mulheres casadas e com filho. Esse resultado confirma a hipótese de que morar mais distante do centro de empregos piora a dinâmica de empregabilidade feminina, evidenciado pelo aumento na taxa de desemprego para o grupo das casadas e com filho, resultado compatível com o documentado por Silveira Neto, Duarte e Páez (2015) e mais intenso do que o observado por Kawabata e Abe (2018). Essa inferência reforça que o papel social da mulher em países em desenvolvimento penaliza mais sua participação no mercado de trabalho. A hipótese da responsabilidade doméstica (BARBANCHON; RATHELOT; ROULET, 2021), a criminalidade observada no trajeto casa-trabalho (SILVEIRA NETO; MOURA, 2019) e o tipo de ocupação mais predominante da mulher (TEDESCO; SOUZA, 2020) são canais que explicam esse efeito econômico.

Tabela 6 – Regressão para Taxa de Desemprego Feminino no Mercado de Trabalho

	Variáveis dependentes							
	Solteira sem filho		Solteira com filho		Casada sem filho		Casada com filho	
	OLS	SDM	OLS	SDM	OLS	SDM	OLS	SDM
Tempo deslocamento (homem)	0.052***	0.043	0.002	0.054	0.050*	0.051	0.127***	0.139***
	(0.019)	(0.035)	(0.037)	(0.068)	(0.027)	(0.050)	(0.026)	(0.048)
Renda pc (log)	0.050***	0.026	0.085***	0.017	0.043*	−0.005	0.027	0.014
	(0.017)	(0.018)	(0.033)	(0.034)	(0.024)	(0.025)	(0.023)	(0.024)
Perc. domicílios alugados	0.045*	0.013	0.023	−0.097	0.036	−0.057	−0.048	−0.059
	(0.027)	(0.031)	(0.052)	(0.061)	(0.038)	(0.045)	(0.036)	(0.043)
Perc. crianças na creche	−10.140***	−5.029**	−10.695***	1.213	−20.644***	−15.553***	−4.749**	−0.423
	(1.739)	(2.073)	(3.403)	(4.055)	(2.495)	(2.990)	(2.381)	(2.857)
Perc. famílias com mais duas crianças	0.454***	0.289***	0.831***	0.415***	0.465***	0.222**	0.335***	0.199**
	(0.056)	(0.065)	(0.110)	(0.128)	(0.081)	(0.094)	(0.077)	(0.090)
Perc. domicílios subnormais	0.013	0.033**	−0.043**	−0.025	−0.045***	−0.046**	−0.004	0.020
	(0.011)	(0.016)	(0.021)	(0.031)	(0.016)	(0.023)	(0.015)	(0.022)
lag.ct_h		−0.021		−0.132		−0.043		−0.062
		(0.043)		(0.084)		(0.062)		(0.059)
lag.ln_rend		−0.001		0.079		0.091*		0.030
		(0.033)		(0.064)		(0.047)		(0.045)
lag.per_alug		0.042		0.193**		0.155**		0.004
		(0.050)		(0.098)		(0.073)		(0.069)
lag.creche		−6.525**		−16.826***		−3.294		−4.964
		(3.289)		(6.339)		(4.815)		(4.453)
lag.hh2ch		0.156		0.639***		0.303*		0.193
		(0.113)		(0.221)		(0.160)		(0.152)
lag.subn		−0.046**		−0.067		−0.013		−0.049
		(0.022)		(0.044)		(0.032)		(0.031)
Constante	−3.203	−3.458	−6.206	−13.691	2.215	−5.361	−3.093	−3.452
	(2.866)	(4.389)	(5.608)	(8.608)	(4.112)	(6.325)	(3.924)	(6.047)
ρ		0.21**		0.12**		0.19**		0.25**
N	633	633	633	633	633	633	633	633
R^2	0.388		0.194		0.300		0.160	
R^2 Ajust.	0.382		0.187		0.293		0.152	
Akaike Inf. Crit.	3.687	3.663	4.537	4.509	4.144	4.127	4.085	4.073
LR Test		12.519***		3.618*		10.078***		16.676***

Note: *p<0.1; **p<0.05; ***p<0.01

Tabela 7 – Estatística I de Moran

Taxa de Desemprego	Moran I	p-valor
Solteira sem filhos	42.21	0.00
Solteira com filhos	24.77	0.00
Casada sem filhos	32.09	0.00
Casada com filhos	23.52	0.00

Fonte: Elaboração própria. Tabela expõe a Estatística *I* de Moran. Número de unidades espaciais foi de 633 para a RMSP, e a amostra foi constituída para homens e mulheres entre 18 e 54 anos.

Os efeitos direto, indireto e total do *commuting* na taxa de desemprego estão expostos na Tabela 8. Considerando o efeito total, a relação entre *commuting* e a taxa de desemprego feminina é positiva e significante em quase todas as variáveis, a única exceção o caso das solteiras sem filhos. Esse resultado é similar ao observado por Farré, Jofre-Monseny e Torrecillas (2020), Barbanchon, Rathelot e Roulet (2021) e Chidambaram e Scheiner (2020), que encontram resultados próximos para o caso dos EUA, Espanha e Alemanha, respectivamente, reforçando a ideia que um maior deslocamento aumento o gap de mercado trabalho por gênero. Mais uma vez, a hipótese é pelo efeito mais perverso da responsabilidade intra-familiar, que penaliza mais as mulheres na presença de filhos, e se mostra mais intensa nesse tipo de vínculo, cuja necessidade regularidade é mais evidente.

Tabela 8 – Efeito marginal do Commuting na Taxa de Desemprego Feminina

	Direto	Indireto	Total	N
Solteira sem filhos	0.04	-0.01	0.03**	633
	(0.027)	(0.006)	(0.031)	
Solteira com filhos	0.05	-0.14	-0.08	633
	(0.043)	(0.008)	(0.049)	
Casada sem filhos	0.05	-0.03	0.01***	633
	(0.038)	(0.006)	(0.042)	
Casada com filhos	0.13***	-0.03	0.10***	633
	(0.048)	(0.018)	(0.061)	

Fonte: Elaboração própria. *p<0.1; **p<0.05; ***p<0.01.

A comparação com o efeito marginal total obtido via SDM, com os coeficientes obtidos no modelo linear via OLS são consistentes, em termos de direção e magnitude. É possível estimar que um incremento de 10 minutos a mais no tempo médio de deslocamento provoca um aumento na taxa de desemprego em 2,3%, 0,9% e 11,9%, respectivamente para o caso das mulheres solteiras sem filhos e casadas sem filhos e com filhos. Nesse último caso, isso representaria cerca de 4.364 mulheres a mais desempregadas na RMSP.

Tais achados são mais uma peça de evidência que, além de convergir com o documentado por Kawabata e Abe (2018), contribui com o debate por duas razões diretas. Primeiro observa o efeito do *commuting* em um país em desenvolvimento, com características sociais e de mobilidade distantes ao caso japonês e americano. Além disso, utiliza a variável que mensura o tempo de deslocamento instrumentalizada, mitigando o efeito endógeno da escolha simultânea entre o local de moradia e a localização das firmas.

2.6 Conclusão

O ensaio buscou aprofundar a discussão sobre o padrão de empregabilidade feminino, com um foco especial sobre o efeito do tempo de *commuting*, comsiderando quatro medidas de empregabilidade: solteiras, com e sem filhos e casadas, com e sem filhos. Para isso foram constituídas duas medidas de mensuração: (i) taxa de participação na força de trabalho e (ii) taxa de desemprego. Foi considerado a Região Metropolitana de São Paulo (RMSP) como área de estudo, devido à sua relevância econômica, bem como pela sua capacidade de comparação com outras metrópoles que já tiveram essa temática abordada, a exemplo de Tóquio e de Nova Iorque.

Como contribuição à literatura, destaca-se o fato de se analisar a empregabilidade feminina sob a perspectiva tanto da participação da força de trabalho, similarmente ao observado em Kawabata e Abe (2018), mas também pela perspectiva da taxa de desemprego, considerando uma avaliação única para o estado civil e a presença de filhos. Outra adição ao corpo de pesquisa está na modelagem empírica, pois o efeito do tempo de deslocamento é estimado via *Spatial Two Stages Least Squate - STSLS*, que, baseado no estimador proposto por Kelejian e Prucha (1997), corrige a possível endogeneidade na relação estimada via variável instrumental, e cuja estimação se dá via *Generelized Method of Moments - GMM*. Por fim, essa pesquisa também contribui com a literatura por observar o caso de um país em desenvolvimento, pois para o caso de países desenvolvidos já existe um robusto corpo evidências e que convergem para o efeito negativo do tempo de *commuting* na participação feminina no mercado de trabalho, tais como Crane (2007), Abe (2011), Black, Kolesnikova e Taylor (2014), Kawabata e Abe (2018), Sánchez, Isabel e González (2014), Farré, Jofre-Monseny e Torrecillas (2020), Barbanchon, Rathelot e Roulet (2021) e Chidambaram e Scheiner (2020). No caso Brasil, já existem achados documentados e que mostram evidências de que o *commuting* afeta a empregabilidade feminina, tais como Silveira Neto, Duarte e Páez (2015), Pereira e Schwanen (2015), Silveira Neto e Moura (2019), Duarte (2020). Contudo, considerar a questão espacial em conjunto com uma estratégia empírica que corrige a potencial endogeneidade da escolha simultânea entre o local de moradia e a localização de moradia ainda não tinha sido testada, sendo uma peça valiosa no debate sobre o tema.

Os resultados da análise empírica expõe significante disparidade no padrão de empregabilidade feminina, com diferenciação pelo estado civil e a presença de filhos. O

Índice Global de Moran mostrou que o emprego feminino distribuído sob um *cluster* espacial, resultado similar ao verificado em Kawabata e Abe (2018). No caso das mulheres solteiras, os *clusters* de alto valor se concentram fortemente na região central do município de São Paulo (CBD), e os de baixo valor nas regiões mais periféricas. Para as mulheres casadas, a concentração diminui, sendo visualmente mais espalhada, criando zonas de emprego em áreas mais descentralizadas. Contudo, no caso das mulheres casadas com filhos pequenos, esse padrão é mais intenso, e se concentra longe do *core* do CBD, denotando a relevância da responsabilidade intra-doméstica e da restrição de mobilidade, esse resultado foi distinto do observado em kawabata2018intra, que demonstra clusterização mais intensa nas casadas e com filhos.

A análise de regressão espacial via SDM, estimado por STSLS, revelou uma associação positiva e significante entre o tempo de deslocamento masculino e a participação na força de trabalho e na taxa desemprego, mas apenas para o caso das mulheres casadas e com filhos. Isso significa dizer que quanto mais tempo dispendido no deslocamento casa-trabalho, menor será a participação na força de trabalho (queda 4,1%) e maior o impacto na taxa de desemprego (aumento de 11,9%), a cada 10 minutos a mais gasto no trajeto. Evidência em consonância com a literatura correlata, mostrando que a presença de filhos pequenos constitui-se como mais uma barreira na empregabilidade feminina. O desemprego desse grupo também é mais intenso fora do centro de empregos principal, o que corrobora a necessidade de políticas públicas mais direcionadas, a exemplo de expansão da rede de creches ou melhora na qualidade de modais. Fatores que criam uma seletividade positiva quanto ao comportamento da mulher. Ou seja, a presença de filhos e os longos trajetos seleciona aquelas mais dispostas a conseguir o emprego e assumir o custo de mobilidade dado o custo oportunidade da presença de filhos. Por fim, mostra-se que o padrão de empregabilidade feminino varia marcadamente considerando a geografia intra-municipal e as disparidades espaciais são variáveis importantes na condução de políticas públicas que almejam melhorar a empregabilidade desse grupo.

3 Votos Presidenciais e PBF: Há Correlação Espacial?

3.1 Introdução

O pleito eleitoral de 2018 restou marcado pelo surgimento de movimentos que interpretam a expansão de programas de transferência de renda como prejudiciais ao bem-estar coletivo (RENNÓ, 2020). Tais movimentos potencializaram a frustração associada aos governos predecessores, especialmente influenciada pela gestão no segundo mandato da ex-presidente Dilma Rousseff (AMARAL; RIBEIRO, 2015) e ensejaram a ascensão e a vitória do candidato Jair Bolsonaro que, originariamente, estabeleceu como uma de suas bases de campanha a defesa do fim de tais programas, em especial do bolsa família (SIMONI, 2021), estratégia em consonância com a implementação de uma política econômica distinta da observada nos mandatos do Lula I e II e da Dilma I (COELHO, 2017; AMARAL, 2020; BRAGA; ZOLNERKEVIC, 2020; RENNÓ, 2020).

A promulgação da Lei 14.282, de 29 de dezembro de 2021, que instituiu o programa Auxílio Brasil e decretou o fim do programa bolsa família, trouxe à tona um debate que parecia ter sido encerrado, em relação ao efeito das políticas de transferência de renda no comportamento de voto (AMARAL, 2020; SIMONI, 2021). Principalmente observando-se o contingente de trabalhos que já documentaram a associação positiva entre a escolha do candidato e a implementação de políticas sociais, em especial as de transferências de renda (ABENSUR et al., 2007; SHIKIDA et al., 2009; MARZAGÃO, 2013; MAGALHÃES; SILVA; DIAS, 2015; BUCHELI, 2015) e que tais políticas se constituem como uma variável preponderante para o resultado das eleições (LINS et al., 2016; PINHO NETO, 2019; FREY, 2019). Portanto, o fato de o atual governo implementar um programa nas bases do Auxílio Brasil, diante do atual cenário pandêmico e em ano eleitoral, vai diverge com a proposta política estabelecida na corrida eleitoral de 2018, a qual preconizava mais austeridade do que assistencialismo (RENNÓ, 2020).

Outro aspecto importante é que programa bolsa família não é uma unanimidade entre a população brasileira, apesar dos resultados positivos em diversas áreas, tais como melhoria na saúde infantil (PINHO NETO; BERRIEL, 2017), na redução da criminalidade (SHEI et al., 2014; CHIODA; MELLO; SOARES, 2016), no abandono escolar (FILMER; SCHADY, 2011), além de outros efeitos positivos. Mundim et al. (2020) mostram que o programa bolsa família não é uma unanimidade entre a população brasileira. Conforme expõem os autores, mesmo com o entendimento sobre o objetivo do programa - focado em reduzir a pobreza e combater desigualdades - há uma percepção de que o desenho do programa encontra-se equivocado, por não beneficiar o grupo de pessoas mais necessitado

e ter um caráter permanente, incentivando os beneficiários a se acomodarem, devido ao benefício recebido.

A caracterização dos efeitos do bolsa família na influência do voto continua em evidência, como relata (SIMONI, 2021). O autor documenta que o canal de transmissão dos benefícios do programa ocorre não somente por uma via direta, ou seja, situação na qual os efeitos afetam apenas aos participantes do programa e seus dependentes, mas também via efeitos indiretos, nos quais é possível observar um efeito de transbordamento na economia local, afetando variáveis de consumo e renda local, bem como dinâmica do mercado de trabalho, assim como demonstram Gerard, Naritomi e Silva (2021).

As evidências observadas mostram que o uso do programa bolsa família restou fortemente difundido na avaliação da dinâmica eleitoral. Contudo, o modelo de condução da política econômica brasileira parece não se distanciar da criação de instrumentos baseados em programas de transferência de renda, fato que ambienta o interesse em continuar a analisar essa temática. Ao recuperar o cenário do pleito eleitoral de 2018, não havia clareza sobre como o Bolsa Família seria utilizado na estratégia de campanha. Afinal, como demonstraram (ZUCCO, 2013; MAGALHÃES; SILVA; DIAS, 2015; AMARAL; RIBEIRO, 2015), o programa fora identificado dentre os principais canais de captação de voto presidencial. Adicionalmente, seria o primeiro pleito após o *impeachment* da ex-presidente Dilma Rousseff, com o ex-presidente Lula da Silva impedido de participar das campanhas, devido às condenações oriundas da Operação Lava Jato e sob um cenário de crescimento no apoio a um candidato que se mostrava contrário a expansão de políticas sociais de transferência de renda (RENNÓ, 2020).

Com essa perspectiva em foco, o ensaio se propõe a fazer uma análise da distribuição espacial dos votos do primeiro turno da eleição para presidente do Brasil no ano de 2018 com a distribuição municipal dos beneficiários do PBF. A estratégia empírica utilizada foi construída a partir do trabalho de Magalhães, Silva e Dias (2015), analisando os determinantes de voto na candidata Dilma Rousseff em 2010, considerando o modelo de correção espacial nos resíduos (Spatial Error Model - SEM). Analisou-se a associação das variáveis via modelo estimado por mínimos quadrados ordinários (OLS) e, em seguida, testou-se a presença da dependência espacial, através da estatística I de Moran. Na avaliação das duas variáveis de trabalho, a saber: a proporção de votos em Fernando Haddad (PT) e a proporção de votos em Jair Bolsonaro (PSL), a dependência espacial se mostrou significante, tanto para a variável defasada, como para o erro. Nessa situação, seguiu-se a regra exposta em Elhorst (2010), na qual é recomendado escolher o modelo com maior valor obtido no teste robusto para autocorrelação espacial. Para ambas as variáveis o teste robusto para o erro foi mais intenso, os resultados dos testes estão explicitados na seção de metodologia.

Os resultados revelam que tanto o Bolsa Família, quanto o "efeito Lula" ainda são significantes e exercem papel importante na decisão de voto. Contudo, diferentemente do

encontrado em Magalhães, Silva e Dias (2015), os resultados consolidam que o peso do Bolsa Família é maior em comparação com o "efeito Lula", provavelmente explicado pelos eventos judiciais à época associados ao ex-presidente, bem como pelo pífio desempenho da atividade econômica no segundo mandato de Dilma Rousseff (AMARAL; RIBEIRO, 2015; PINHO NETO, 2019). Na análise dos votos em Jair Bolsonaro, encontrou-se o esperado efeito negativo tanto do Bolsa Família, quanto do Lula na decisão do eleitor. Em complemento, utilizou-se uma variável para observar a influência da segunda gestão Dilma Rousseff nas eleições de 2018. Os resultados também evidenciam que essa variável se mostrou a mais relevante na explicação do comportamento do voto, tanto para a proporção de votos em Fernando Haddad (PT) como em Jair Bolsonaro (PSL), induzindo a ideia de que a associação à má condição econômica aos governos anteriores do PT, em especial ao último mandato da Dilma Rousseff, foi o canal mais relevante na determinação dos votos, e não necessariamente a expansão do programa bolsa família ou mesmo os redutos originais do ex-presidente Lula. Esse resultado é convergente, em certa medida, com os trabalhos de Amaral e Ribeiro (2015) e Fernandes e Fernandes (2019), que analisaram a importância do voto econômico para a eleição de 2014.

Esse ensaio encontra-se conectado à literatura que analisa a relevância de programas de transferência de renda no contexto eleitoral. Especificamente, reitera a importância da questão espacial nesse tipo de abordagem, enfatizando a necessidade de corrigir o viés espacial. O artigo também traz como contribuição a avaliação empírica para a eleição presidencial de 2018, que ainda não foi abordada sob essa perspectiva.

A próxima seção propõe uma revisão da literatura. A estratégia empírica e os métodos de pesquisa estão descritos na seção 3.2. A seção 3.3 descreve a base de dados utilizada, e as variáveis de trabalho. Os resultados e as conclusões estão nas seções 3.5 e 3.6, respectivamente.

3.2 Revisão da Literatura

Desde a reeleição do ex-presidente Lula em 2006, diversas pesquisas se dedicaram a estudar a relação entre o voto em candidatos e o acesso ao Bolsa Família. Dentre vários trabalhos, referencia-se Zucco (2013), Marzagão (2013), Nicolau (2014), Magalhães, Silva e Dias (2015), Amaral e Ribeiro (2015), Pereira et al. (2015), Lins et al. (2016), Schuhli (2018), Pinho Neto (2019), Zucco, Luna e Baykal (2020), Gerard, Naritomi e Silva (2021). E, ao considerar o pleito presidencial, os resultados convergem para uma associação positiva entre o percentual de beneficiários do bolsa família e o voto em candidatos vinculados a origem do programa. Contudo, a polarização eleitoral observada após as eleições de 2018, bem como a implementação do Auxílio Emergencial, em face ao enfrentamento da pandemia do Coronavírus em 2020, bem como a implementação do Auxílio Brasil, fez ressurgir o debate sobre as consequências de programas de transferência de renda e o comportamento de voto.

Zucco (2013) analisou o efeito de programas de transferência considerando as três eleições presidenciais consecutivas, desde a unificação de programas no Bolsa Família em 2004. O autor propôs um modelo via Generalized Propensity Score, a partir de duas bases de dados, um survey e outra base de dados a partir dos resultados eleitorais, disponibilizados pelo TSE. A variável de interesse foi o desempenho do incumbente associado à expansão do PBF. Os resultados mostraram relação positiva no curto prazo. Contudo, a longo prazo, esses efeitos são dissipados e perdem significância. Conectado a esses resultados.

Em outro estudo, Bohn (2013) mostrou que analisar apenas o desempenho eleitoral via a existência agregada de beneficiários do programa, pode causar viés na estimação, tendo em vista o transbordamento socioeconômico desproporcional, promovido pela expansão do Bolsa Família nas regiões do país.

Magalhães, Silva e Dias (2015) buscaram responder qual foi a importância do Bolsa Família e o peso do ex-presidente Lula da Silva, representado pela proporção de votos no ex-presidente em 2006, na reeleição da Dilma. Os autores analisam considerando uma perspectiva espacial para o corrigir o viés espacial, tal como Hunter e Power (2007), Shikida et al. (2009) e Marzagão (2013). Tal correção mostrou-se importante devido a dependência do voto e a posição geográfica dos municípios, haja vista as relações de transbordamentos (efeito *spillover*) entre os municípios vizinhos, em decorrência dos efeitos econômicos e sociais na economia local, e que foram modificados em virtude da distribuição de beneficiários do PBF. Via um modelo de correção espacial encontraram uma associação positiva e significante entre os votos na Dilma Rousseff e o Programa Bolsa Família, assim como pela proporção de votos no ex-presidente Luís Inácio Lula da Silva em 2006, resultados convergentes com o documento por Zucco (2013). Dilma Rousseff obteve desempenho melhor em municípios menores, mais rurais, e nos quais o prefeito incumbente não era do PT. Os autores tratam a eleição da Dilma como uma reeleição do "Lula".

A reeleição da Dilma em 2014 também já foi alvo de investigação por Amaral e Ribeiro (2015), que propuseram uma análise exploratória para analisar os determinantes do voto nos dois turnos das eleições para presidente de 2014. Os autores usaram como estratégia empírica um modelo de regressão logística, tendo como variáveis explicativas as condições sociais, a identidade política e uma medida de avaliação retrospectiva e prospectiva do próprio governo da ex-presidente Dilma Rousseff. Os principais resultados reportaram o Bolsa Família com efeitos exclusivamente no segundo turno, devido ao acirramento da disputa eleitoral e a constituição de "ameaças vazias" de cancelamento do programa Bolsa Família em caso de derrota da então presidente Dilma, perspectiva clientelista também observada por Stokes et al. (2013) e Hagene (2015). Outro ponto de destaque evidenciado pelos autores foi a existência de um julgamento do governo dissociado da situação econômica do país, pois apesar dos escândalos já noticiados, variáveis reais ainda haviam sido afetadas, tais como a taxa de desemprego, a inflação e a própria atividade econômica do país. Na visão dos autores, isso foi determinante para a reeleição

da ex-presidente Dilma Rousseff em 2014. Contudo, os autores não consideram a questão da dependência espacial na estratégica empírica, fator relevante e já identificado em trabalhos anteriores, tais como Bucheli (2015) e Magalhães, Silva e Dias (2015).

Ainda no contexto da importância da expansão do bolsa família e a associação entre a proporção de beneficiários do programa e o resultado eleitoral, Lins et al. (2016) construíram um modelo de previsão que utilizou a dependência do município ao bolsa família como variável preditora para avaliar o resultado das eleições presidenciais entre 2006 e 2014, argumentando que quanto maior for a dependência do município, maior a chance de eleição do candidato incumbente. Os resultados mostraram um erro médio de 2,15%, sendo bem ajustado comparativamente à literatura da área. Os autores também incluem testes espaciais para ajustar o modelo mais adequado.

O trabalho de Schuhli (2018) explorou as eleições de 2014, com foco mais específico na avaliação da importância da religiosidade na decisão de voto. Nesse caso, os resultados foram estimados considerando a necessidade de correção da dependência espacial, e ratificaram a importância do voto católico no desempenho eleitoral da Dilma Rousseff, mesmo a candidata sendo de origem de um partido mais progressista e com pautas menos conservadoras. Os resultados foram similares ao documentado por Magalhães, Silva e Dias (2015), que observaram a relação entre a proporção de votos na Dilma Rousseff e a proporção de beneficiários do BF na eleição de 2010, controlando pela proporção de eleitores católicos e evangélicos no município. Adicionalmente, o autor demonstra que o eleitor católico pode levar mais em consideração variáveis econômicas ou não associa o Partido dos Trabalhadores (Dilma) como sendo moralmente oposto aos seus valores. Por outro lado, o autor não considera o efeito direto de transferência de votos da própria Dilma de 2010, deixando de fora uma variável importante na explicação do comportamento do eleitor.

Mais recentemente trabalhos começam a avaliar o contexto das eleições de 2018. Amaral (2020) buscou identificar os determinantes da vitória do atual Presidente, Jair Bolsonaro. A estimação empírica foi realizada via regressão logística, considerando o efeito do desempenho retrospectivo da economia e a identificação política. Para tal, o autor usou uma base de dados de pesquisas pré-eleição. Os resultados apontam uma relevância maior da identificação política do que da análise retrospectiva, inclusive maior associação ao candidato Fernando Haddad ao Bolsa Família. Diferentemente de Schuhli (2018), o autor encontrou mais rejeição do voto católico ao candidato do PT (Fernando Haddad), vinculado as questões morais, nas quais Bolsonaro teve melhor desempenho.

A importância dos aspectos econômicos foi diretamente abordada por Fernandes e Fernandes (2019), mostrando a importância do voto econômico, isto é, de variáveis econômicas para a estimação do resultado eleitoral. Os autores mostraram, a partir de uma amostra de dados eleitorais em painel, para os municípios de São Paulo, que a performance da economia local exerce um papel importante no resultado eleitoral. Rennó

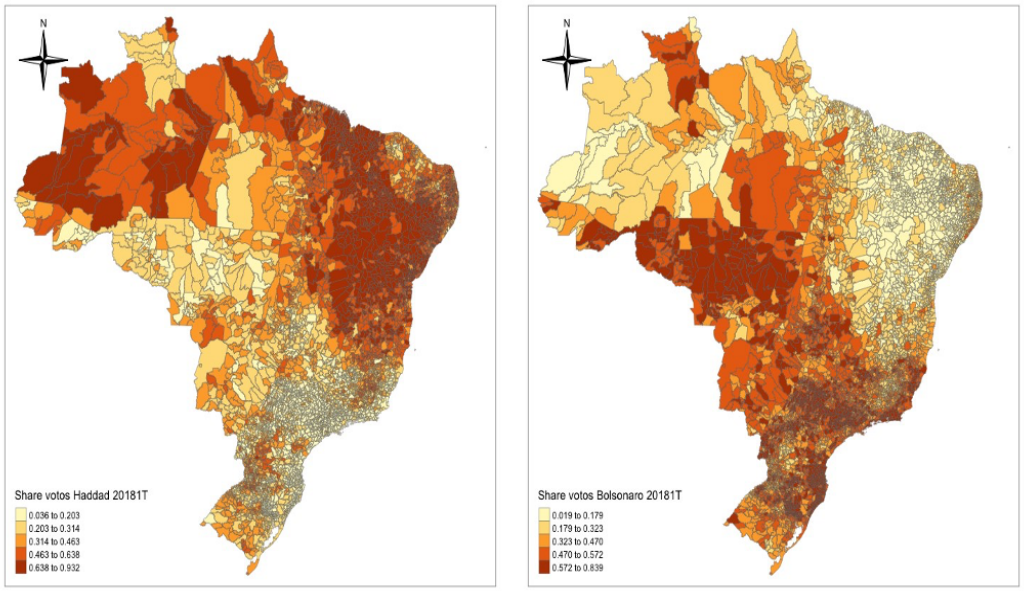
(2020) analisou as eleições de 2018, observando os fundamentos ideológicos e as variáveis motivadoras para a criação do movimento "bolsonarista", nas palavras do próprio autor. O ressentimento para com os governos do anteriores do PT, em função dos escândalos de corrupção, revelados pela Operação Lava Jato, assim como pelos resultados econômicos ruins, observados após o impeachment da ex-presidente Dilma Rousseff e o ressurgimento do pensamento liberal associado a uma onda de conservadorismo, que rejeitava políticas sociais, foram as variáveis mais relevantes. O artigo também destaca o movimento em escala mundial do crescimento de uma ideologia política nacionalista e do voto na extrema-direita, como propulsores o crescimento da popularidade do candidato Jair Bolsonaro no Brasil. A análise foi constituída a partir de dados de pesquisas eleitorais, porém sem nenhuma modelagem empírica. Braga e Zolnerkevic (2020) analisam o pleito de 2018 sob outra perspectiva, ao classificar as eleições presidenciais desde 1989 até 2018, a luz de uma base conceitual. Os autores documentam alta correlação com a eleição imediatamente anterior, fato intensificado a partir de 2006, sendo mais forte em 2010 e 2014. O movimento de transferência de voto ocorre tanto para o voto pró-PT, quanto para o movimento anti-PT, justificando a polarização eleitoral experimentada em 2018. No entanto, ambos os trabalhos desconsideram a importância da necessidade de correção do viés espacial, além de não possuir uma estratégia empírica que mensure a associação entre o voto presidencial e a expansão do Bolsa Família.

Mundim et al. (2020) mostrou que, a despeito de diversos resultados positivos observados a partir da expansão do programa bolsa família, e em diferentes áreas, tais como saúde, educação, pobreza e criminalidade, a população brasileira não é unânime em relação ao bolsa família. Os autores analisaram os dados de uma pesquisa nacional de opinião pública, e a partir de um modelo logístico multinomial documentaram que, apesar de haver associações positivas, em especial no combate à pobreza e fome, há forte conexão que o programa possui problemas no desenho, em relação ao critério de seleção dos beneficiários, bem como o fato de que o benefício estimula os participantes do programa a se acomodaram ou mesmo a ter mais filhos para aumentar a cota de recebimento. A exploração de fatores indiretos, como mecanismos importantes na decisão do eleitor foi abordada por Simoni (2021). O autor mostra a relevância não só dos efeitos diretos, caracterizado como o impacto apenas nos próprios beneficiários, mas também os efeitos indiretos. O argumento do autor é os transbordamentos econômicos também afetam o comportamento do voto do eleitorado.

Em adição às evidências levantadas, no pleito de 2018 se observou um país com um padrão geográfico de eleitores marcadamente polarizado. A Figura 10 expõe a distribuição de votos dos dois candidatos mais bem votados, que disputaram o pleito eleitoral de 2018 no segundo turno, explicitando a divisão espacial e a existência de marcante uniformidade nos votos a nível de município. Fica evidente a associação dos votos no candidato do PT, Fernando Haddad, nas regiões mais beneficiadas, em proporção à população, com o programa Bolsa Família, e naturalmente economicamente menos favorecidas. Por outro

lado, os votos no, à época candidato, Jair Bolsonaro foram diametralmente opostos, sendo mais presente nas Regiões Centro-Oeste e Sul do país, municípios com indicadores socioeconômicas melhores e, por regra, menos dependentes do programa Bolsa Família (ZUCCO; LUNA; BAYKAL, 2020).

Figura 10 – Proporção de votos presidenciais - 2018 (1 turno)



Fonte: TSE. Elaboração própria. Figura à esquerda ilustra a densidade de votos como proporção do total dos votos válidos no candidato Fernando Haddad (PT) para o primeiro turno das eleições presidenciais de 2018. A figura à direita, a densidade de votos no candidato Jair Bolsonaro (PSL).

Sob essa perspectiva, esse artigo visa contribuir com a literatura que estuda a relação entre programas de transferência de renda e comportamento de voto, analisando o pleito eleitoral de 2018. Observar a dinâmica de voto em tais configurações requer uma especificação que lide com a correlação espacial na estimação, haja vista a marcante diferença tanto nas condições sociais e econômicas, como no próprio perfil do eleitor, como enfatizam LeSage e Pace (2010), Marzagão (2013), Magalhães, Silva e Dias (2015), associando o melhor desempenho dos candidatos do PT em municípios cujos indicadores sociais são piores. A próxima seção se dedica a descrever não somente o modelo empírico utilizado, bem como os critérios para escolha da melhor especificação para lidar com a dependência espacial observada.

3.3 Estratégia Empírica

O modelo empírico desse artigo tem como variáveis dependentes a proporção de votos, na eleição presidencial de 2018, dos candidatos Fernando Haddad (PT) e Jair Bolsonaro (PSL), no primeiro turno. A escolha em não observar os resultados das eleições no segundo turno é devido a análise para essa fase do pleito eleitoral possuir limitações analíticas, em função da escolha entre os dois candidatos mais bem votadas

não necessariamente estar associada à avaliação da condução de propostas ou da política econômica adotada (MAGALHÃES; SILVA; DIAS, 2015).

A equação (3.1) representa a estrutura empírica do modelo utilizado nas estimações, onde a variável Y_{im} é a variável dependente, e equivale a proporção de votos no primeiro turno das eleições presidenciais de 2018, no município m para o candidato $i = \{\text{Haddad ou Bolsonaro}\}$. No exercício empírico também é considerado um conjunto de variáveis de controle X', no intuito de melhorar o ajuste da estimação, e prover mais robustez ao estimador do modelo. Por fim, β é um vetor de coeficientes para as variáveis representadas pela matriz X e u_m é o termo de erro clusterizado a nível de município.

$$Y_{im} = X'\beta + u_m \tag{3.1}$$

Na estratégia empírica adotada nesse artigo o modelo apresentado na eq. (3.1) possui quatro estimações distintas, nas quais as variáveis de controle são inseridas gradualmente. No modelo 1, a variável dependente Y_{mi} é regredida apenas contra a proporção de beneficiários do bolsa família no município. Depois, é incluído a proporção de votos obtidos pelo Lula em 2006 (modelo 2). No modelo 3 é incorporada a proporção de votos obtidos pela Dilma em 2014. E, finalmente, as outras variáveis socioeconômicas são adicionadas no modelo 4.

É digno de nota explicar que não foi utilizado, como variável dependente, a proporção de votos em Dilma nas eleições de 2010, em função da correlação dessa variável com a proporção de votos em Dilma em 2014. Então, optou-se por manter a variável mais próxima do pleito eleitoral de 2018.

No entanto, o modelo de regressão estimado via Ordinary Least Square - OLS, apesar de ser muito utilizado na análise da associação entre programa de transferência de renda e o comportamento de voto (NICOLAU, 2006; ZUCCO, 2011; MARZAGÃO, 2013; MAGALHÃES; SILVA; DIAS, 2015; AMARAL; RIBEIRO, 2015; SCHUHLI, 2018), na presença de dependência espacial pode ocasionar estimadores inconsistentes e viesados (ANSELIN et al., 2001; LESAGE; PACE, 2009). Essa ineficiência pode ocorrer tanto via erro padrão, como na variável dependente (lag). Os modelos mais comuns expostos pela literatura que lidam com essa correção são spatial Durbin model - SDM, o spatial autoregressive model - SAR e o spatial error model - SEM (ELHORST, 2010; CHEN, 2013; VEGA; ELHORST, 2015; FISCHER; WANG, 2011).

Com já observado por Bohn (2013), Marzagão (2013), Labonne (2013), Magalhães, Silva e Dias (2015), há avaliação da relação entre o comportamento de voto e programas de transferência de renda, a dependência espacial se mostrou relevante, requerendo uma estimação que lide com essa questão. Assim, o segundo passo metodológico desse artigo foi obter a Estatística I de Moran, para avaliar o grau de dependência espacial na variável de interesse. Conforme LeSage e Pace (2010), é necessário testar a significância da

dependência espacial e, com isso, corrigir a limitação estrutural contida nos modelos via OLS sob a presença de autocorrelação espacial. O Índice de Moran (I) verifica a presença de autocorrelação espacial, a estatística I de Moran pertence a um intervalo $-1e1$. A Equação (3.2) demonstra como esse indicador é obtido numericamente:

$$I = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{i,j}(x_i - \bar{X})(x_j - \bar{X})}{\sum_{i=1}^{n}(x_i - \bar{X})^2} \tag{3.2}$$

Onde n é o número de unidades espaciais, indexadas por i e j, x equivale à variável de interesse, $\bar{X}$ é o primeiro momento da variável, e $x_{i,j}$ indica a relação espacial entre i e j.

A hipótese nula da Estatística de Moran (I) é de que a variável de interesse esteja aleatoriamente distribuída no espaço, isto é, inexista autocorrelação espacial. Caso contrário, e $I > 0$, significa que há presença de um cluster espacial; se $I < 0$, há dispersão espacial. Além disso, se faz necessário realizar os testes de robustez, tanto em relação a defasagem da variável estimada (lag) bem como do erro, para que se possa escolher a melhor especificação do modelo de correção espacial.

Para avaliar a relevância da dependência espacial na estimação empírica, Elhorst (2010) sugere além do teste via a estatística I de Moran, a análise dos testes robustez. A importância dos testes robustos se dá pelo fato do LM erro impor a inexistência autocorrelação espacial na variável defasada (lag), assim como o LM lag, supor a inexistência autocorrelação espacial no erro (LESAGE; PACE, 2010). Os testes robustos corrigem a imposição dessas hipóteses e auxiliam na determinação do modelo mais ajustado caso a presença de dependência espacial for significante.

O passo seguinte foi identificar qual modelo é mais adequado. Segundo Anselin et al. (2001), Elhorst (2010), Chen (2013) a estratégia consiste em analisar todos os testes de robustez, verificando seu nível de significância para a dependência espacial. Como poderá ser visto na seção de Resultados, conforme as Tabelas 4 e 7, todos os testes foram significantes, tanto para o erro quanto para a defasagem, indicando a presença de autocorrelação espacial no erro e na variável dependente. A partir da evidência de que o LM robusto para o erro é maior do que o LM robusto com defasagem, a recomendação é seguir com o spatial error model como especificação para corrigir a dependência espacial identificada LeSage e Pace (2010), Elhorst (2010), Vega e Elhorst (2015). Essa, inclusive, foi a estratégia adotada por Magalhães, Silva e Dias (2015).

O spatial error model (SEM) é um modelo de correção espacial utilizado quando a dependência ocorre no erro. Ou seja, existe correlação entre o termo de perturbação estocástica e a variável de interesse. E a partir dessa evidência o uso do modelo via OLS se torna inadequado, haja visto que os estimadores obtidos na equação 1 são ineficientes.

Conforme LeSage e Pace (2010), LeSage (2014), a forma usual de especificação do SEM é descrita pela equação (3.3):

$$\epsilon_i = \lambda \sum_{j=1}^{n} W_{ij}\epsilon_j + u_i \tag{3.3}$$

Na equação (3.3), λ é o parâmetro autoregressivo e u_i o termo de erro aleatório. Conforme LeSage (2014), normalmente assume-se que essas variáveis são independentes e identicamente distribuídas. Reescrita de forma matricial tem-se:

$$\epsilon = \lambda \sum_{j=1}^{n} W\epsilon + u \tag{3.4}$$

Para obter a forma funcional da equação (3.4), assume-se que $|\lambda| < 1$ e então resolve para ϵ. Isso provê:

$$\epsilon = (I - \lambda W)^{-1}u \tag{3.5}$$

Ao inserir o resultado da equação (3.4) na forma matricial de um modelo estimado via OLS, tal como $y = X\beta + \epsilon$, obtém-se a especificação do *spatial error model*:

$$y = X\beta + (I - \lambda W)^{-1}u \tag{3.6}$$

Por fim, Chen (2013) enfatiza que o uso do SEM é entendido como uma combinação de regressão padrão via OLS, com um modelo autoregressivo espacial no termo de erro ϵ, mas com o ganho de eficiência dado a correção espacial.

Adicionalmente, segundo Elhorst (2010), um teste comumente utilizado para a escolha do modelo mais ajustado para corrigir a dependência espacial é o critério de informação de Akaike (AIC), que expressa a perda de informação, induzindo a um critério de escolha. O modelo escolhido deve ser aquele com menor perda, ou seja, menor distância do processo gerador de dados. Então, quanto menor for o AIC mais bem ajustado será (AKAIKE, 1974).

A escolha da matriz de pesos espaciais foi baseada via critério de vizinhança *Rook*. Contudo, foi testada a estimação dos modelos via matriz *Queen* e não houve alterações significativas nos resultados. Como mostram Anselin et al. (2001), LeSage (2014), Elhorst (2010) não existe um teste específico para a escolha da matriz de pesos espaciais, mesmo existindo forte sensibilidade dos resultados à arbitração da matriz W. Assim, como sugerido por Anselin et al. (2001) e Baumont, Ertur e Gallo (2004), a regra estabelecida neste artigo foi determinar a matriz de pesos, a partir do I de Moran. Ou seja, estimou-se o modelo 1, foi testada a presença de autocorrelação espacial via I de Moran para um conjunto de matrizes W e selecionou-se a matriz que gerou o maior valor para o teste I de Moran.

3.4 Base de Dados e Variáveis

A base de dados utilizadas na avaliação das eleições de 2018 foi construída a partir de três fontes de dados: i) estatísticas eleitorais obtidas no Tribunal Superior Eleitoral (TSE), que disponibiliza a proporção de votos nos candidatos por município e de forma agregada; ii) Dados do IBGE, construídos a partir do Censo 2010; e iii) A proporção de pessoas beneficiadas pelo Programa Bolsa Família no município, via Ministério do Desenvolvimento Social (MDS). A Tabela 9 resume e descreve as variáveis de trabalho.

As variáveis de interesse do artigo são a proporção de votos na eleição presidencial de 2018, nos candidatos Fernando Haddad (PT) e Jair Bolsonaro (PSL). A principal variável explicativa do modelo empírico é a proporção de beneficiários do Bolsa Família no município em 2017, ano imediatamente anterior ao pleito eleitoral. A justificativa para a adoção dessa variável é pelo fato de o programa ter se consolidado como a principal política social de transferência de renda no país e que, conforme Abensur et al. (2007), Marzagão (2013), Nicolau (2006), Amaral e Ribeiro (2015), está fortemente associada ao desempenho eleitoral dos candidatos (LINS et al., 2016). De forma complementar, avalia-se o efeito da proporção de votos no Lula em 2006, como visto em Zucco (2013) e Magalhães, Silva e Dias (2015), para tentar responder se os redutos eleitorais construídos ainda se mantêm importantes na explicação do voto presidencial, tais como visto na eleição da Dilma em 2014 (AMARAL; RIBEIRO, 2015; PINHO NETO, 2019). Adicionalmente, ao observado na literatura dedicada ao tema, construiu-se uma variável defasada para captar o poder de transferência de votos na ex-presidente Dilma Rousseff nas eleições de 2014.

Tabela 9 – Descrição das variáveis de trabalho

Variáveis dependentes	**Definição**	**Fonte**
sh_had_18	Percentual de votos em Haddad (2018)	TSE
sh_jb_18	Percentual de votos em Bolsonaro (2018)	TSE
Variáveis independentes	**Definição**	**Fonte**
sh_BF_17	Percentual de beneficiados pelo BF no município em 2017	MDS
sh_Lula_06	Percentual de votos em Lula (2006)	TSE
sh_dil_06	Percentual de votos em Dilma (2014)	TSE
POP_2017	Log da População residente (2017)	IBGE
PIB_pc	PIB (log) municipal per capita 2017	IBGE
sh_cato	Percentual de católicos no município em 2010	IBGE
sh_evang	Percentual de evangélicos no município em 2010	IBGE
Pref_PT	Dummy para prefeitura do PT em 2016	TSE

Fonte: Elaboração própria. A Tabela descreve as variáveis de trabalho utilizadas no artigo, bem como a fonte para as quais foram obtidas.

Além das variáveis eleitorais descritas acima, foi incluído no modelo empírico um conjunto de variáveis explicativas, conforme já evidenciado na literatura que estuda a associação destes componentes com a dinâmica de voto (NICOLAU, 2014). Por exemplo, Zucco (2015) sugere controlar pelo tamanho da população, haja vista que existe uma associação entre o tamanho da população e a alocação de beneficiários, assim como o peso dessa política social é influenciada pelo porte municipal.

Para captar a condição econômica do município, adotou-se o PIB (em log) per capita de 2017, como preconiza Magalhães, Silva e Dias (2015). Além disso, baseado em Schuhli (2018) também foi inserido variáveis de religião, para captar o voto católico e evangélico, mas sem correr o risco de multicolinearidade. Por fim, foi adicionado uma variável dummy para analisar o efeito do partido que administra a prefeitura em 2016, ano da última eleição municipal, ser da sigla do Partido dos Trabalhadores. E, com isso, verificar a influência da captação de votos pela gestão local (BOHN, 2013; ZUCCO, 2015; CANÊDO-PINHEIRO, 2015; FREY, 2019; BROLLO; KAUFMANN; FERRARA, 2020).

Na Tabela 10 expõem-se um resumo descritivo das variáveis de trabalho para se ter uma dimensão da distribuição dos votos e das condições sócio-econômicas da amostra de trabalho.

Tabela 10 – Estatísticas descritivas das variáveis

Variável	Média	Máximo	Mínimo	Desvio-padrão
Prop. votos Haddad 20181T	0.4121	0.9324	0.037	0.214
Prop. votos Bolsonaro 20181T	0.3871	0.8389	0.019	0.189
Prop. Beneficiários BF	0.094	0.2381	0.001	0.05
Prop. votos Lula 20161T	0.5162	0.9337	0.1141	0.175
Prop. votos Dilma 20141T	0.5226	0.9213	0.1020	0.127
População 2017	37.312	12.106.920	812	21.847
PIB pc 2017	21.982	34.847	3.285	20.917
Prop. Católicos 2010	0.76	0.99	0.09	0.13
Prop. Evangélicos 2010	0.16	0.85	0.00	0.09

Fonte: Elaboração própria. A Tabela expõe algumas estatísticas descritivas das variáveis de trabalho utilizadas no artigo.

A próxima seção expõe os resultados desse artigo. Além da estimação via spatial error model, também estimou-se via OLS padrão para comparação dos resultados. A análise ocorreu por ano eleitoral, nos quais analisou a respectiva variável dependente, ou seja, a proporção de votos em Fernando Haddad (PT) e Jair Bolsonaro (PSL) para o ano de 2018.

3.5 Resultados

O artigo se concentrou em analisar os resultados das eleições presidenciais de 2018, e se restringiu aos resultados do primeiro turno, uma vez que os votos do segundo turno apresentam ruídos devido ao voto nulo (MAGALHÃES; SILVA; DIAS, 2015). A análise ocorre para o desempenho dos dois candidatos mais bem votados em 2018, Fernando Haddad e Jair Bolsonaro, este último se tornou o vencedor das eleições. Analisar o desempenho do candidato do PT, Fernando Haddad, também se torna interessante dado que, diferentemente do que ocorreu em 2010 e 2014, o Partido dos Trabalhadores foi

derrotado nas eleições presidenciais. Então, abre-se a oportunidade para compreender como foi o efeito, não somente do Bolsa Família, mas também do mecanismo de transferência de votos nas bases municipais. Em complemento, é possível observar tal impacto tanto no longo prazo, dado que o efeito da influência do ex-presidente Lula se refere a proporção de votos em 2006, quanto no curto prazo, dado a reeleição da Dilma em 2014. Outra motivação está associada ao ambiente político, devido aos diversos choques ocorridos entre o pleito de 2014 e 2018, tais como o impeachment da própria Dilma, bem como o movimento contrário ao PT, que surgiu após 2014.

Para realizar a análise do modelo de regressão desse artigo, primeiro foi estimado o modelo via OLS, conforme exposto na equação (3.1). A Tabela 3 expõe os resultados para a proporção de votos no candidato Fernando Haddad (PT). A análise foi realizada em quatro especificações distintas. A primeira estimação considera apenas como variável explicativa a proporção de beneficiários do bolsa família no município em 2017. O modelo 2 inclui a proporção de votos no Lula em 2006. Em seguida, é adicionada como variável a proporção de votos na Dilma em 2014. Finalmente, são inseridas todas as demais variáveis de controle descritas na seção 3.4.

Os coeficientes estimados mostram que, para a eleição presidencial de 2018, existe uma forte e significativa associação entre a proporção de votos no candidato do PT, Fernando Haddad, e a proporção de beneficiários do bolsa família, resultado convergente com as eleições de 2006, 2010 e 2014 e observados por Zucco (2013), Magalhães, Silva e Dias (2015), Amaral e Ribeiro (2015), respectivamente. A inclusão do percentual de votos no Lula em 2006 reduz drasticamente o coeficiente obtido no modelo 1. Contudo, a inclusão do percentual de votos obtidos da Dilma em 2014 mostrou-se ser mais relevante na explicação do comportamento de voto no Haddad do que o próprio percentual de voto no Lula em 2006.

A análise das variáveis de controle também traz inferências interessantes. O fato da Prefeitura ser administrada por um prefeito do PT contribui positivamente para a performance do candidato Fernando Haddad, fundamentalmente explicada pelo poder de disseminação de voto, além do voto por empatia à legenda, tal resultado foi documentado por Amaral e Ribeiro (2015), Amaral (2020) e Zucco, Luna e Baykal (2020). As variáveis que tratam do aspecto religioso, tanto a proporção de católicos quanto de evangélicos mostraram uma associação negativa com a performance de Haddad. Possivelmente, os escândalos de corrupção, a associação uma vice do PCdoB (esquerda mais extrema), que defende pautas menos conservadoras, tais como legalização do aborto e empoderamento feminino explicam esse resultado. Além dessas características, observa-se que o candidato Haddad teve pior desempenho em municípios maiores e mais ricos, corroborando os resultados visualizados por Magalhães, Silva e Dias (2015) para a eleição de 2010.

Tabela 11 – Proporção de votos no Haddad 2018 via OLS

	Proporção de votos no Haddad 2018			
	(1)	(2)	(3)	(4)
Proporção de BBF 2016	2.822***	1.811***	1.113***	0.994***
	(0.025)	(0.031)	(0.029)	(0.031)
Proporção votos no Lula 2006		0.491***	0.178***	0.200***
		(0.011)	(0.011)	(0.011)
Proporção votos na Dilma 2014			0.596***	0.560***
			(0.011)	(0.012)
População 2010 (log)				−0.010***
				(0.001)
PIB (log) per capita 2017				−0.011***
				(0.002)
Perc. Católicos				−0.151***
				(0.021)
Perc. Evangélicos				−0.281***
				(0.029)
Dummy Pref. PT				0.010***
				(0.004)
Constante	0.141***	−0.015***	−0.098***	0.289***
	(0.003)	(0.004)	(0.004)	(0.036)
N	5,564	5,564	5,564	5,564
R^2	0.700	0.779	0.854	0.860
Adj. R^2	0.700	0.778	0.854	0.859
Note:				*p<0.1; **p<0.05; ***p<0.01

Mesmo após a inclusão de todas as variáveis de controle, os dados obtidos modelo 4 da equação 1 mostram que para cada 1% de aumento na cobertura do programa bolsa família no município, a performance do candidato Fernando Haddad teria aumentado em 0,994 ponto percentual, já a variável que representa o efeito da Dilma 2014, teria sido de 0,56 ponto percentual. O incremento de 1% na proporção de votos no Lula em 2006 (efeito Lula) responderia por um aumento no desempenho de 0,2 ponto percentual nos votos do candidato no pleito de 2018, denotando distanciamento dos votos à imagem do ex-presidente Lula e explicado, possivelmente, pelos escândalos que envolviam, à época, o ex-presidente em função dos desdobramentos da Operação Lava-Jato. Embora, as variáveis

aqui trabalhadas não façam alusão direta à estas questões, se configurando apenas como uma hipótese inferencial.

A Tabela 12 expõe os resultados para a Estatística I de Moran e para os testes de dependência espacial. Os resultados mostram a presença de autocorrelação espacial tanto na variável dependente como no erro e a necessidade de uma especificação que corrija o viés obtido por uma estimação via OLS. Ao analisar o índice de Moran, verifica-se a existência de elevada correlação espacial para a proporção de votos no Haddad no primeiro turno de 2018. Os valores obtidos induzem a existência de padrões geográficos de voto com a formação de clusters espaciais, haja vista associação regional no padrão de voto.

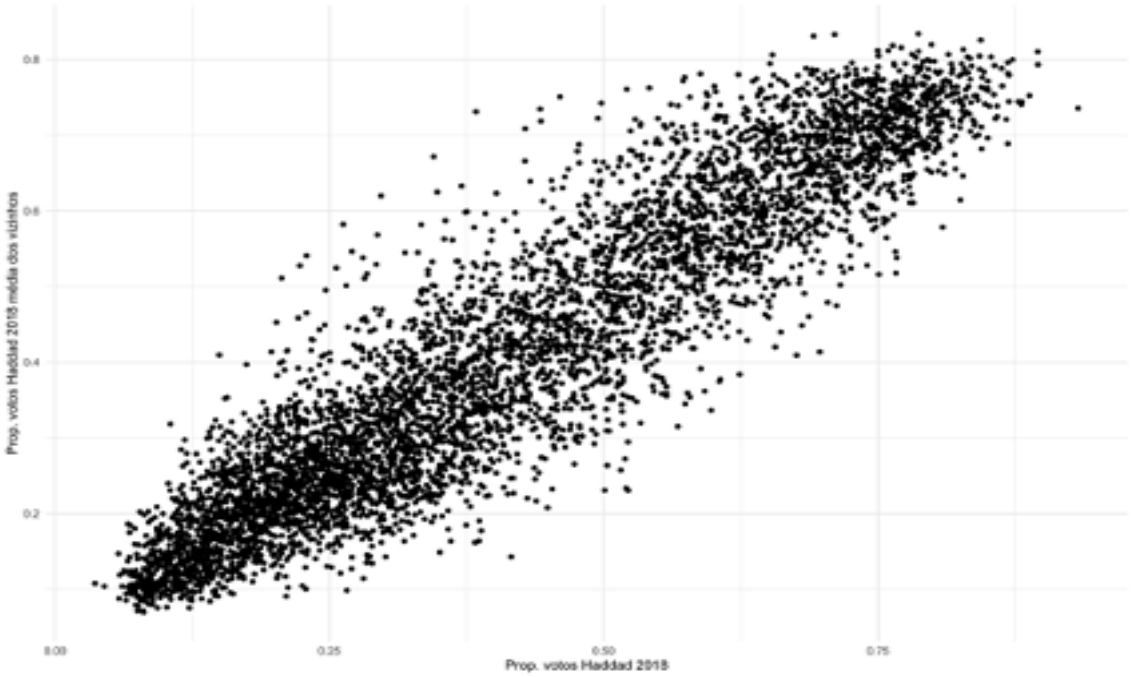
Tabela 12 – Resultados dos testes de dependência espacial – Haddad 2018_1

	Haddad 2018	
Testes	Valor	p-valor
I de Moran	87.08	0.001
LM (erro)	6332.7	0.000
LM (lag)	3238.5	0.000
LM Robusto (erro)	3232	0.000
LM Robusto (lag)	137.9	0.000

A tabela expõe o teste de dependência espacial de Moran para as duas variáveis dependentes, a proporção de votos no Haddad em 2018, assim como os testes LM e LM Robusto tanto para o *lag* como para o erro. Na primeira coluna está o valor do teste e na segunda o p-valor associado.

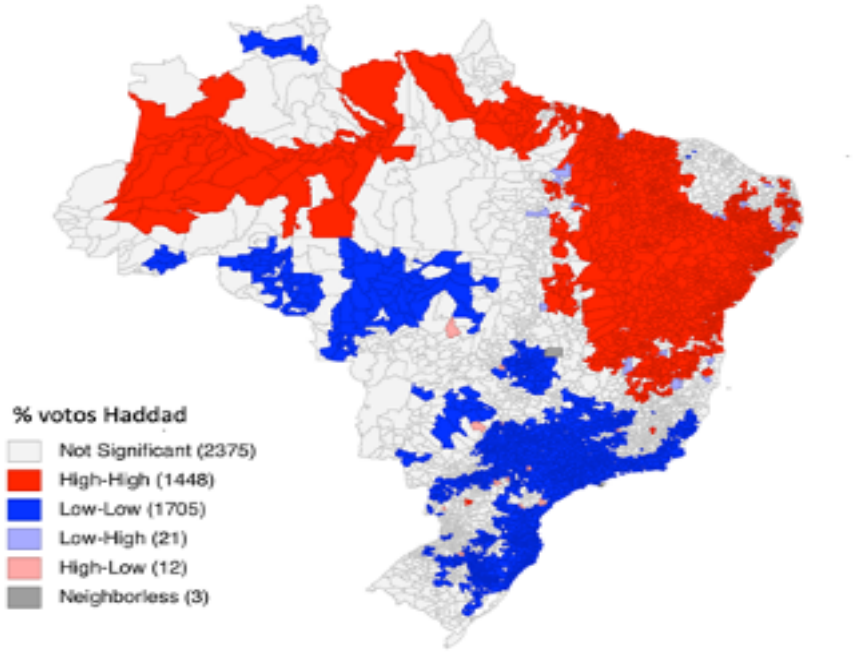
Adicionalmente, a Figura 11 ilustra o gráfico de correlação espacial de Moran para a variável dependente, sendo possível ver tanto a direção do padrão espacial identificado, como a existência e a intensidade da formação dos clusters. Já a Figura 12 confirma o padrão espacial identificado no gráfico de Moran e também expõe a dependência espacial e a formação de um cluster de alto valor na região Norte e Nordeste, com exceção apenas para o Estado do Ceará, que se configura como reduto político do candidato Ciro Gomes, que também estava na disputa eleitoral para a presidência da República em 2018. A análise da Figura 12 expõe uma evidente separação geográfica no padrão de voto do candidato, nas quais é possível ver regiões com alta iteração (hot spots) e localizadas no terceiro quadrante do gráfico e áreas com cluster de baixo valor (cold spots), localizados no segundo quadrante do gráfico.

Figura 11 – Moran Scatterplot para (%) Votos no Haddad 2018_1



Fonte: Elaboração própria. Figura ilustra o gráfico de dependência espacial de Moran. A estatística I de Moran é de 0.87, o que representa dependência espacial e presença cluster espaciais.

Figura 12 – Local Moran (LISA) para (%) Votos no Haddad 2018_1



Fonte: Elaboração própria. Figura expõe o Índice de dependência espacial local (LISA). As áreas em vermelho representam os *clusters* de alto valor e as áreas em azul, os *clusters* de baixo valor.

Elhorst (2010) sugere que em situações nas quais os resultados dos testes de robustez forem significantes tanto para a variável defasada como para o erro, a regra é estimar o

modelo pelo maior valor do teste robusto. Nesse caso, a indicação é utilizar o spatial error model, pois, o teste possui mais intensidade para a correção nos resíduos (LESAGE, 2014).

A Tabela 13 apresenta o resultado do modelo com correção espacial (coluna 2), em comparação com o modelo estimado via OLS, para a proporção de votos no Haddad em 2018. Os resultados foram similares em termos de direção e magnitude do modelo descrito pela equação (3.1). Diferentemente do observado em Magalhães, Silva e Dias (2015), a proporção de beneficiários do Bolsa Família se mostrou mais relevante em comparação com a proporção de votos no Lula em 2006 para a performance do candidato Haddad. Essa evidência acentua a importância que o programa ainda possui nas bases municipais, demonstrando ainda ser um canal importante para a obtenção de votos (FERNANDES; FERNANDES, 2019; BRAGA; ZOLNERKEVIC, 2020). Os resultados, após a correção para a dependência espacial, mostram que para cada aumento de 1% na proporção de beneficiários do bolsa família, o aumento de na votação do Haddad teria sido de 0,424 ponto percentual. Contudo, o aumento na proporção de votos no Lula teria impacto apenas em 0.184 ponto percentual no desempenho eleitoral.

Por outro lado, ao incluir como controle a proporção de votos obtidos por Dilma Rousseff em 2014, percebe-se não só que essa variável foi de maior magnitude em comparação com a proporção de beneficiários do bolsa família, mas também que houve um aumento no valor do estimador ao corrigir pela dependência espacial. Isso denota que os municípios com maior percentual de votos na Dilma em 2014 continuavam mais alinhados com o voto na legenda do PT, e foram menos contaminados com o desempenho econômico ruim observado a partir de 2015. Esse resultado se mantém em linha com Rennó (2020) e Mundim et al. (2020).

Tabela 13 – Modelo com correção espacial

	Proporção de votos no Haddad	
	OLS	*SEM*
	(1)	(2)
Proporção de BBF 2016	0.994***	0.424***
	(0.031)	(0.024)
Proporção votos no Lula 2006	0.200***	0.184***
	(0.011)	(0.009)
Proporção votos na Dilma 2014	0.560***	0.582***
	(0.012)	(0.010)
População 2010 (log)	−0.010***	−0.011***
	(0.001)	(0.001)
PIB (log) per capita 2017	−0.011***	−0.013***
	(0.002)	(0.002)
Perc. Católicos	−0.151***	−0.013
	(0.021)	(0.017)
Perc. Evangélicos	−0.281***	−0.129***
	(0.029)	(0.021)
Dummy Pref. PT	0.010***	0.003
	(0.004)	(0.002)
Constante	0.289***	0.233***
	(0.036)	(0.025)
N	5,564	5,564
R^2	0.860	
Adj. R^2	0.859	
λ		0.861***
Akaike Inf. Crit.	−12,745	−17,411
Wald Test		12,988*** (df = 1)
LR Test		5,135*** (df = 1)
Note:		*p<0.1; **p<0.05; ***p<0.01

As eleições de 2018 também permitem a análise do desempenho do atual presidente, Jair Bolsonaro. A Tabela 14 expõe os resultados das estimações via OLS. No entanto, antes da discussão dos resultados em si, vale repercutir a mudança na distribuição de votos pelos municípios (Figura 10). Configuração diametralmente oposta ao padrão de votos obtidos pelo PT nas últimas três eleições, com as regiões Sul, Sudeste e principalmente

Centro-Oeste fortemente marcada pelo voto no candidato.

Os resultados mostraram, como a hipótese desse trabalho previa, uma associação negativa entre os votos no candidato e as variáveis de percentual de beneficiários do Bolsa Família, da proporção de votos no Lula em 2006 e da proporção de votos na Dilma em 2014. Na especificação 1, a associação entre os votos entre Jair Bolsonaro e a proporção de beneficiários do BF no município foi negativa e significante. O resultado mostrou que a cada 1% de expansão na proporção de beneficiários, o voto no município caiu 2,521 ponto percentual. Ao incluir, como variável dependente, a proporção de votos no Lula em 2006, a magnitude na relevância da cobertura do bolsa família é reduzida, mas ainda se mantendo positiva e significante. Assim como o observado para a proporção de votos em Haddad, a influência da proporção de votos na Dilma em 2014 é mais relevante do que a proporção de votos em Lula (2006), acentuando que a memória negativa do eleitor em relação ao mal momento econômico da época foi um canal significativo na formação do voto presidencial (AMARAL; RIBEIRO, 2015; RENNÓ, 2020; BRAGA; ZOLNERKEVIC, 2020; MUNDIM et al., 2020).

Analisando as variáveis de controle, é possível ver que municípios menos pobres contribuíram positivamente para o desempenho do atual presidente. O candidato também alcançou desempenho positivo no voto católico e evangélico, especialmente nesse último grupo, muito em função de sua estratégia de campanha e sua associação a diversas lideranças religiosas.

Tabela 14 – Proporção de votos Bolsonaro 2018.1 via OLS

	Proporção de votos Bolsonaro 2018			
	(1)	(2)	(3)	(4)
Proporção de BBF 2016	−2.521***	−1.591***	−0.984***	−0.834***
	(0.022)	(0.027)	(0.024)	(0.025)
Proporção votos no Lula 2006		−0.451***	−0.179***	−0.185***
		(0.010)	(0.009)	(0.009)
Proporção votos na Dilma 2014			−0.519***	−0.454***
			(0.010)	(0.010)
População 2010 (log)				−0.001
				(0.001)
PIB (log) per capita 2017				0.020***
				(0.002)
Perc. Católicos				0.079***
				(0.017)
Perc. Evangélicos				0.378***
				(0.023)
Dummy Pref. PT				−0.003
				(0.003)
Constante	0.629***	0.773***	0.845***	0.488***
	(0.002)	(0.004)	(0.003)	(0.029)
N	5,564	5,564	5,564	5,564
R^2	0.709	0.793	0.865	0.884
Adj. R^2	0.709	0.793	0.865	0.884
Note:				*p<0.1; **p<0.05; ***p<0.01

A Tabela 15 expõe a estatística I de Moran e o resultado dos testes de dependência espacial para a proporção de votos no Bolsonaro em 2018. O índice de Moran a existência de positiva e elevada correlação espacial, e a indicação de formação de clusters espaciais. Essa inferência é, inclusive, corroborada pela Figura **??**, que acentua um grau de municípios espacialmente correlacionados na determinação do voto (quadrante 3).

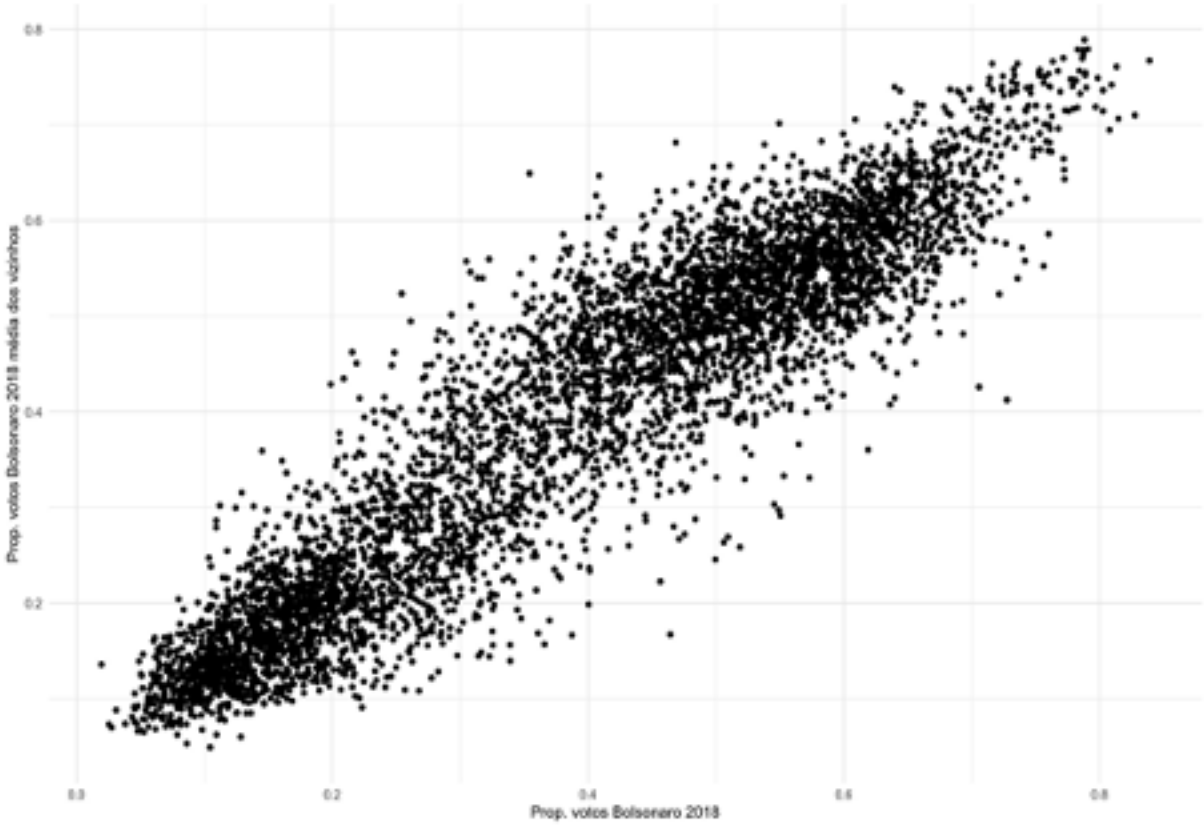
Tabela 15 – Resultados dos testes de dependência espacial – Bolsonaro 2018_1

	Bolsonaro 2018	
Testes	Valor	p-valor
I de Moran	88.09	0.001
LM (erro)	3515.2	0.000
LM (lag)	2246.6	0.000
LM Robusto (erro)	1552.8	0.000
LM Robusto (lag)	284.24	0.000

A tabela expõe o teste de dependência espacial de Moran para as duas variáveis dependentes, a proporção de votos no Haddad e no Bolsonaro em 2018, assim como os testes LM e LM Robusto tanto para o *lag* como para o erro. Na primeira coluna está o valor do teste e na segunda o p-valor associado.

A Figura 13 reforça a importância de corrigir o viés espacial. Já a Figura 14, ilustra a sinergia de votos nas regiões opostas aos redutos do PT, com cluster de alto valor em parte das regiões Centro-oeste, Sudeste e Sul.

Figura 13 – Moran Scatterplot para (%) Votos no Haddad 2018_1

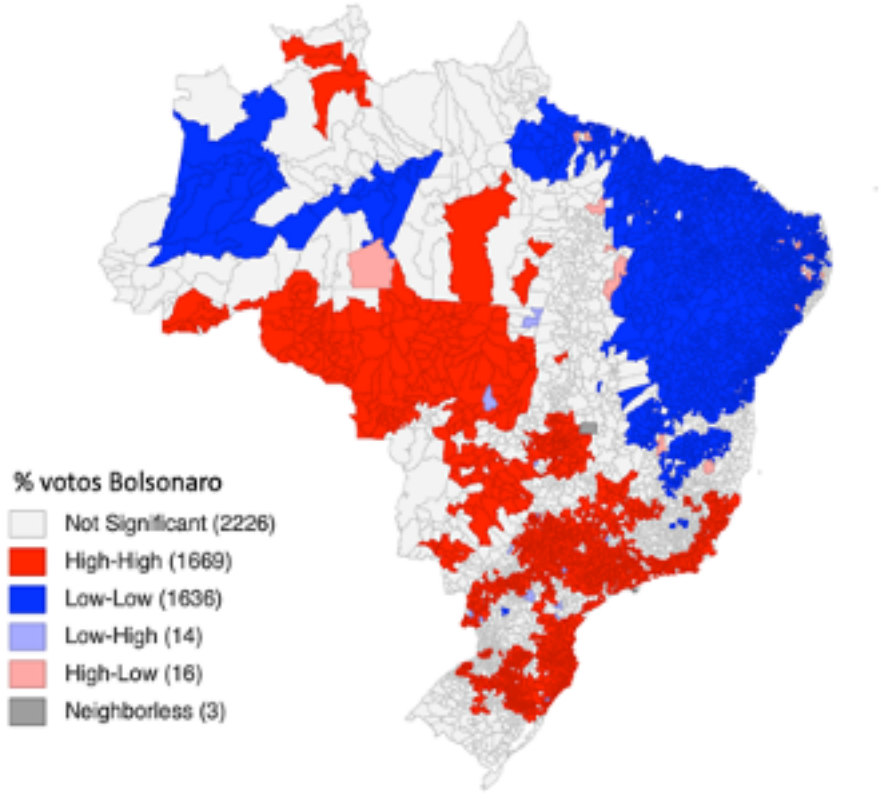


Fonte: Elaboração própria. Figura ilustra o gráfico de dependência espacial de Moran. A estatística I de Moran é de 0.87, o que representa dependência espacial e presença cluster espaciais.

O modelo de correção espacial (Tabela 16, coluna 2) mantém os resultados na estimação sem correção (coluna 1), em termos de direção, contudo o peso da cobertura do programa bolsa família, que no modelo via OLS é o fator mais relevante na explicação do voto, tem seu peso diminuído, ao corrigir pela dependência espacial. É possível ver que a

proporção de votos na Dilma em 2014 é mais relevante na explicação da performance do, então candidato, Jair Bolsonaro. Para cada 1% de aumento na cobertura do Bolsa Família, a perda de voto teria sido 0,37 ponto percentual. Já o aumento em 1% na proporção de votos no Lula teria reduzido o desempenho do Bolsonaro em 0,137 ponto percentual. No caso da proporção de votos na Dilma em 2014, a redução proporcional do desempenho de votos teria sido de 0,47 ponto percentual. Esses resultados enfatizam que o programa bolsa família ainda exerce um papel importante na dinâmica de voto presidencial, contudo, para o caso da eleição presidencial de 2018, a memória do eleitor médio, devidamente associada ao desempenho econômico do segundo mandato da Dilma Rousseff, bem como os escândalos de corrupção em vigor à época do pleito eleitoral parecem ter sido mais relevantes na formação do voto.

Figura 14 – Local Moran (LISA) para (%) Votos no Bolsonaro 2018_1



Fonte: Elaboração própria. Figura o Índice de dependência espacial local (LISA). As áreas em vermelho representam os clusters de alto valor e as áreas em azul, os clusters de baixo valor.

Os resultados discutidos se mantiveram significativos mesmo após a inclusão das variáveis de controle, demonstrando que o desempenho do Bolsonaro foi positivamente influenciado por municípios mais populosos, mais ricos e mais religiosos, em especial o voto evangélico.

Tabela 16 – Modelo com correção espacial

	Proporção de votos Bolsonaro 2018	
	OLS	SEM
Proporção de BBF 2016	-0.834^{***}	-0.372^{***}
	(0.025)	(0.023)
Proporção votos no Lula 2006	-0.185^{***}	-0.137^{***}
	(0.009)	(0.009)
Proporção votos na Dilma 2014	-0.454^{***}	-0.473^{***}
	(0.010)	(0.010)
População 2010 (log)	-0.001	0.007^{***}
	(0.001)	(0.001)
PIB (log) per capita 2017	0.020^{***}	0.013^{***}
	(0.002)	(0.002)
Perc. Católicos	0.079^{***}	0.034^{**}
	(0.017)	(0.016)
Perc. Evangélicos	0.378^{***}	0.217^{***}
	(0.023)	(0.021)
Dummy Pref. PT	-0.003	-0.004^{*}
	(0.003)	(0.002)
Constante	0.488^{***}	0.482^{***}
	(0.029)	(0.024)
N	5,564	5,564
R^2	0.884	
Adj. R^2	0.884	
λ		0.8214^{***}
Akaike Inf. Crit.	$-14{,}684$	$-17{,}761$
Wald Test		$8{,}717^{***}$ (df = 1)
LR Test		$3{,}079^{***}$ (df = 1)
Note:		*p<0.1; **p<0.05; ***p<0.01

3.6 Conclusões

Esse trabalho buscou verificar duas questões primordiais. Em primeiro plano, se o programa bolsa família ainda exerce relevância no processo eleitoral, ao passo de que o pleito de 2018 foi concebido sob fortes ameaças em relação à continuidade do programa. Além de criar uma agenda em perspectiva, que consiste em avaliar a substituição do bolsa

família pelo Auxílio Brasil nas eleições de 2022. Segundo, se a força do ex-presidente Lula continua a ter influência na performance dos candidatos patrocinados pela sigla do PT, mesmo o ex-presidente não sendo mais o predecessor imediato, como foi o caso da eleição de 2010, em que a ex-presidente Dilma Rousseff foi a candidata eleita. Para superar a perda de eficiência dos modelos estimados via OLS, dado a dependência espacial diagnosticada, utilizou-se o *spatial error model*, que se mostrou o mais indicado nos testes de robustez. Foi analisado a eleição presidencial de 2018, cujas variáveis de investigação foram a proporção de votos nos candidatos Fernando Haddad e Jair Bolsonaro.

O contexto de 2018 era bem distinto das eleições de 2010 e 2014. A ex-presidente Dilma Rousseff havia sido deposta da presidência após condenação no processo de impeachment, o país experimentava uma crise fiscal e econômica, além de observar o crescimento do movimento, como pontuado por Amaral e Ribeiro (2015) e Rennó (2020), de ressentimento ao PT, provocando o surgimento de uma onda mais conservadora e que marcou a corrida eleitoral. Tais eventos potencializaram a candidatura de Jair Bolsonaro, sendo determinantes para sua vitória, em segundo turno, frente ao candidato do PT, Fernando Haddad.

Os resultados reforçam a hipótese teórica que motivou o trabalho de investigação desse ensaio. O candidato Fernando Haddad (PT) teve sua performance melhorada em municípios com maior percentual de beneficiários do Bolsa Família, inclusive sendo uma variável mais relevante que o "efeito Lula", fato divergente do obtido por Magalhães, Silva e Dias (2015) na análise para as eleições de 2010. Por outro lado, ao incluir a proporção de votos obtidos por Dilma Rousseff em 2014, se observa que o efeito de transferência de votos foi bastante significante, induzindo que as bases locais ainda se mantém fortes. Esse resultado enfatiza a manutenção da força da associação do programa aos candidatos do PT, mas expõe o argumento da polarização e do voto por legenda (LINS et al., 2016; COELHO, 2017; FERNANDES; FERNANDES, 2019; AMARAL, 2020; BRAGA; ZOLNERKEVIC, 2020; MUNDIM et al., 2020).

Os resultados referentes ao então candidato, Jair Bolsonaro, foram diametralmente opostos. Ou seja, nos municípios cuja cobertura do programa Bolsa Família era maior o resultado foi negativo, contribuindo para a perda de voto no candidato. Contudo, os resultados do modelo empírico revelaram que o efeito dos votos obtidos por Dilma Rousseff em 2014 foi o mais significativo para perda de votos, superando tanto o efeito do Lula (proporção de votos em 2006) como a cobertura do programa bolsa família em 2017. Importante ressaltar que, o perfil identificado do eleitor de Bolsonaro se mostrou bem distinto daquele associado ao PT. As variáveis de controle que se mostraram significativas foram das cidades mais ricas, mais escolarizadas, mais populosas e mais religiosas, em especial, aqueles cuja proporção de evangélicos era maior.

Outra contribuição do artigo é a possibilidade de comparar o perfil do eleitor de dois candidatos antagônicos, e com a recente anulação dos processos que impediam

a candidatura do Lula à corrida presidencial, abre-se um leque de possibilidades de investigação para o certame de 2022, que em função da crise do SARS-Covid19, tem sua definição sob muita incerteza.

Os resultados observados também destacam alguns pontos que podem ser expandidos em pesquisas futuras. Primeiro, toda a massa de resultados expostos trata da correlação entre a proporção de votos de candidatos e variáveis eleitorais ou econômicas. É fato que este tipo de associação sofre tanto de viés de agregação como do problema de variáveis omitidas, então buscar estratégias para lidar com essa potencial endogeneidade deve ser conduzida na busca por inferências causais. Em segundo lugar, uma peça nova do pleito eleitoral de 2018 foi a influência de *fake news* da formação de opinião do eleitor. É fácil lembrar eventos conectados a disseminação desses conteúdos. Então, avançar na avaliação dessa variável e seu efeito no resultado pode traçar evidências interessantes para avaliação da dinâmica de voto.

4 Bolsa Família afeta voto presidencial? Evidências para o caso Brasil

4.1 Introdução

Os programas de transferência de renda surgiram com intuito de minimizar, de maneira imediata, a quantidade de pessoas em situação de pobreza extrema através de uma transferência monetária para cidadãos/famílias pre-selecionadas. Conforme estudos realizados constatou-se que tais programas mostraram-se como estratégias eficazes no combate à desigualdade e à pobreza, melhorando indicadores sociais, educacionais e de saúde. Os países pertencentes à América Latina apresentaram mais intensidade no uso desses programas, inseridos em um arcabouço de políticas de proteção social (ANGELUCCI; GIORGI, 2009; CORRÊA, 2015; CHIODA; MELLO; SOARES, 2016).

O desenho de políticas de transferência de renda estabelece um conjunto de condicionantes que vinculam a participação e a permanência no programa, a exemplo da comprovação em relação à matrícula escolar de crianças, a realização de pré-natal por gestantes, assim como a manutenção em dia da vacinação de crianças, visando, principalmente, a supressão da pobreza inter-geracional (CHIODA; MELLO; SOARES, 2016). Como consequência, um conjunto de evidências do impacto da expansão desses programas foram documentadas em diversas áreas, tais como saúde (SHEI et al., 2014; MILLER; JOHNSON; WHERRY, 2021), educação (ATTANASIO; KAUFMANN, 2014; BARRERA-OSORIO et al., 2011; FILMER; SCHADY, 2011), fertilidade (STECKLOV et al., 2006), bem-estar (BARROS; FOGUEL; ULYSSEA, 2006), criminalidade (BUCHELI, 2015; CHIODA; MELLO; SOARES, 2016), mercado de trabalho (GERARD; NARITOMI; SILVA, 2021; ANGELUCCI; GIORGI, 2009), consumo (ANGELUCCI; ATTANASIO, 2009; DUARTE; SAMPAIO; SAMPAIO, 2009), dentre outros.

Não obstante, ainda existe uma exploração em curso no que se refere ao impacto desses programas em variáveis eleitorais. Duch e Stevenson (2006) e Margalit (2019) documentaram que, apesar da existência de evidência empírica de que a condição econômica contemporânea afeta os resultados eleitorais e que mudanças no nível de renda estão associados ao realinhamento de votos no longo prazo, não há clareza de que essa transferência de renda afeta positivamente o desempenho eleitoral. Em busca de evidências nesse sentido, Pinho Neto (2019) documenta, por exemplo, o aumento da participação política, via redução do *turnout*. O autor estudou os resultados das eleições presidenciais entre 2006 e 2014, e mostra também uma correlação positiva com o suporte ao partido incumbente.

Um questionamento pertinente é sobre a potencial endogeneidade existente entre

as variáveis eleitorais e as variáveis que mensuram os programas assistencialistas (PINHO NETO, 2019). Para lidar com essa problemática, a literatura dedicada aos métodos de inferência causal expõem um leque de estratégias que objetivam superar tanto o viés de seleção, devido ao programa ser universal, quanto o efeito das características não observáveis (ANGRIST; PISCHKE, 2008; ABADIE; DIAMOND; HAINMUELLER, 2010). Em exemplos mais recentes, é possível ver que as conexões entre políticas de transferência de renda e o comportamento do voto do eleitor médio potencialmente premiam incumbentes, explicado pelo receio da potencial piora na condição econômica em caso de mudança na gestão (MANACORDA; MIGUEL; VIGORITO, 2011; JANVRY; FINAN; SADOULET, 2012).

Observa-se que a continuidade desses programas pode modelar o voto e a orientação política, em função da incorporação do benefício como renda permanente (AKEE et al., 2020). A associação entre políticas de renda e comportamento de voto também molda as estratégias de políticos, em especial aos postulantes à reeleição ou expostos à medidas de *accountability* mais aleatórias, como expõe (FERRAZ; FINAN, 2011; JANVRY; FINAN; SADOULET, 2012). Por conseguinte, é possível ver efeitos mais claros dessas políticas em variáveis eleitorais. Pode-se exemplificar tal reflexo nas famílias beneficiadas por programas de renda aumentam a participação nas eleições (BAEZ et al., 2012; PINHO NETO, 2019), visto que aumentam o suporte aos partidos incumbentes (MANACORDA; MIGUEL; VIGORITO, 2011), bem como melhoram o ambiente de competição e reduzem o apoio a partidos com foco clientelista, pois reduz a vulnerabilidade de eleitores mais pobres (FREY, 2019).

Em busca de contribuir na problemática, esse ensaio visa investigar, para o caso Brasil, o efeito da expansão do principal programa de transferência de renda, o Programa Bolsa Família (PBF) no resultado das eleições presidenciais. O estudo está concebido de forma longitudinal, no qual se analisa as eleições presidenciais de 2002 a 2018. A variável de interesse é a proporção de votos recebidos pelo candidato do Partido dos Trabalhadores (PT). A escolha por essa variável se dá por duas razões. Primeiro, foi exatamente no primeiro mandato de Luís Inácio Lula da Silva (Lula) que houve a unificação do PBF como principal programa social do governo. Segundo, por mais quatro eleições consecutivas (2006 a 2014), foi o candidato do PT o vencedor do pleito eleitoral. E como demonstrado por Baez et al. (2012), Zucco (2011), Magalhães, Silva e Dias (2015) e Amaral e Ribeiro (2015), existe uma relação positiva e significante entre a proporção beneficiários do Bolsa Família no município e voto nos candidatos vinculados ao PT.

Para superar a relação endógena entre a cobertura do programa e a proporção de votos, explora-se a uma reforma metodológica implementada pelo Governo Federal em 2009, e exposta em Gerard, Naritomi e Silva (2021). Os autores mostraram uma fonte de variação exógena ocorrida nas cotas de beneficiários a nível municipal, devido a uma mudança institucional na definição do critério de pobreza no município, o que provocou uma realocação desproporcional no número de beneficiários. O novo critério mudou a

forma de definição de famílias pobres e potenciais elegíveis ao benefício, baseando-se no Mapa de Pobreza do Banco Mundial, ver Elbers, Lanjouw e Lanjouw (2003).

O efeito causal é avaliado através de um modelo baseado na estratégia de diferenças em diferenças, o qual especifica-se a reforma de 2009 como base para constituição do efeito de tratamento na variável de interesse. O grupo de municípios tratados foi definido considerando duas regras de classificação. A primeira, seguiu-se a métrica exposta por Gerard, Naritomi e Silva (2021), a qual se refere aos municípios que ganharam mais beneficiários a partir da Reforma de 2009. A segunda, para isolar esse efeito de 2009, também foi filtrado o efeito da alocação ocorrida em 2006, que normatizou a inclusão de diversos beneficiários desde a criação do programa em 2004, conforme contextualiza Zucco (2013).

O impacto da expansão do programa será observado sob dois mecanismos. O primeiro, enfatiza o efeito demanda provocado pela inserção desses recursos às famílias mais vulneráveis, que estimula a atividade econômica regional e afeta o comportamento de voto devido às melhores condições econômicas no município (ANGELUCCI; GIORGI, 2009). Esse canal também produz um efeito indireto ao estimular a dinâmica local, via emprego e salário, como demonstra Bucheli (2015). O segundo mecanismo é delinea-se pelo efeito recompensa ao incumbente, que conforme Nicolau (2014), premia o atual gestor em função da inclusão no benefício, aumentando o suporte em especial aos mais recentemente incluídos (MANACORDA; MIGUEL; VIGORITO, 2011; PINHO NETO, 2019). Ainda nessa perspectiva Janvry, Finan e Sadoulet (2012) mostraram que programas de renda afetam também o comportamento de políticos, sendo os incumbentes mais propícios a patrocinar e estimular tais programas.

Os resultados mostraram que a expansão no número de beneficiários, em função da reforma de 2009, impactou positivamente no número de votos no PT. Em média, os municípios pertencentes ao grupo que experimentou crescimento mais intenso nos beneficiários, aumentaram em 3,7% a proporção de votos na legenda, em comparação com o período eleitoral pré-reforma (2002 e 2006). Na busca de uma validação interna do método, utilizou-se três especificações distintas como robustez, na qual variou-se a regra de classificação dos municípios como tratados e de controle. Em concomitância, realizou-se um exercício de falsificação, modificando a variável dependente para avaliar o efeito do tratamento. O uso de estratégia empírica baseada em diferenças em diferenças, corrigiu o viés de seleção e, como mostram os testes de robustez implementados, os resultados permaneceram válidos.

Outra contribuição peculiar dessa pesquisa é que implementa-se um modelo de diferenças em diferenças, a partir do choque na mudança exógena ocorrida nas cotas de beneficiários em 2009. Tal estratégia empírica supera as possíveis relações endógenas observadas em pesquisas anteriores, tais como: Zucco (2011), Magalhães, Silva e Dias (2015) e Amaral e Ribeiro (2015). Por fim, apesar de utilizar uma base de dados agregada,

tal como os trabalhos de Baez et al. (2012) e De La O (2013), esse ensaio implementa uma avaliação que utiliza uma fone de variação exógena já documentada, e que mostrou efeito positivo e significante em variáveis econômicas e de mercado de trabalho, como o observado em Gerard, Naritomi e Silva (2021), reforçando a validação externa do método. Essa estratégia ainda não foi aplicada na relação entre programas de transferência de renda e comportamento de voto.

A próxima seção descreve a estrutura institucional do sistema eleitoral, do Bolsa Família e da reforma metodológica ocorrida em 2009. A seção 4.3 apresentará a estratégica empírica implementada. E a seção 4.4 descreve a base de dados utilizada no estudo, bem como algumas estatísticas descritivas construídas. A seção 4.5 é dedicada aos Resultados e complementada com um conjunto de testes de robustez, que foram implementados para validar a estratégia de identificação utilizada. Por fim, expõe-se uma as conclusões na seção 4.7.

4.2 O Programa Bolsa Família e a Reforma Metodológica de 2009

O Programa Bolsa Família, como demonstram Barros, Foguel e Ulyssea (2006), Abensur et al. (2007), Canêdo-Pinheiro (2015), Chioda, Mello e Soares (2016) e Pinho Neto (2019), se configura como um dos maiores e bem sucedidos programas de transferência de renda, inclusive servindo como fonte de inspiração à diversas outras experiências (MANACORDA; MIGUEL; VIGORITO, 2011). Em sua formatação original, o PBF concatenou uma série de outras ações assistencialistas em um único programa centralizado e com um conjunto de condicionantes para as famílias beneficiárias (ZUCCO, 2011). Desde sua criação, em meados de 2004, a metodologia utilizada pelo Programa Bolsa Família para a alocação de novos beneficiários consistia em calcular o quantitativo de famílias pobres, a partir de dados obtidos via Pesquisa Nacional por Amostra de Domicílios (PNAD) e, com isso, definir o quantitativo de alocações por município. A regra de corte era estabelecida nas famílias cuja renda per capita não fosse superior a dois salários mínimos. Gerard, Naritomi e Silva (2021) expressam que as cotas de 2003 e 2006 foram definidas em dois momentos, conforme as equações (4.1) e (4.2):

$$Cota_{ms}^{2003} = \frac{Poor_{ms}^{2000}}{\sum_{k \in s} Poor_{ks}^{2000}} Poor_{s}^{2001} \tag{4.1}$$

$$Cota_{ms}^{2006} = \frac{Poor_{ms}^{2000} * n_{ms}^{[2000,2003]}}{\sum_{k \in s} (Poor_{ks}^{2000} * n_{ks}^{[2000,2003]})} Poor_{s}^{2004} \tag{4.2}$$

Onde $Poor_{ms}^{2000}$ representa o número de famílias pobres no município m no estado s obtidas no Censo 2000. $Poor_{s}^{2001}$ e $Poor_{s}^{2004}$ o número de famílias pobres no estado s obtidas via PNAD 2001 e PNAD 2004, respectivamente.

Em 2009, o Governo Federal modificou a metodologia de critério para alocação de novos beneficiários no Programa. Gerard, Naritomi e Silva (2021) mostraram que o primeiro passo se manteve igual (eq. (4.1)). No entanto, como pode ser visto em Elbers, Lanjouw e Lanjouw (2003), houve alteração na regra de classificação para uma família elegível ao benefício. Essa mudança afetou de maneira distinta a alocação de novos beneficiários nos anos subsequentes a 2009. Com a reforma institucional de 2009, o IBGE estabeleceu como critério de seleção de famílias elegíveis pelo "mapa de pobreza", cuja metodologia pode ser expressa pela equação (4.3).

$$Cota_{ms}^{2009} = \frac{\hat{Poor}_{ms}^{2006}}{\sum_{k \in s} \hat{Poor}_{ks}^{2006}} * 1,18 * Poor_{s}^{2006} \tag{4.3}$$

A motivação principal da mudança foi buscar captar uma medida de bem-estar mais ampla, incluindo indicadores sobre educação, saúde, segurança alimentar para estimar a proporção de famílias pobres a nível municipal. É válido ressaltar que estratégia foi usada unicamente em 2009, pois em 2012 o IBGE voltou a usar o Censo para definir a alocação de cotas (GERARD; NARITOMI; SILVA, 2021).

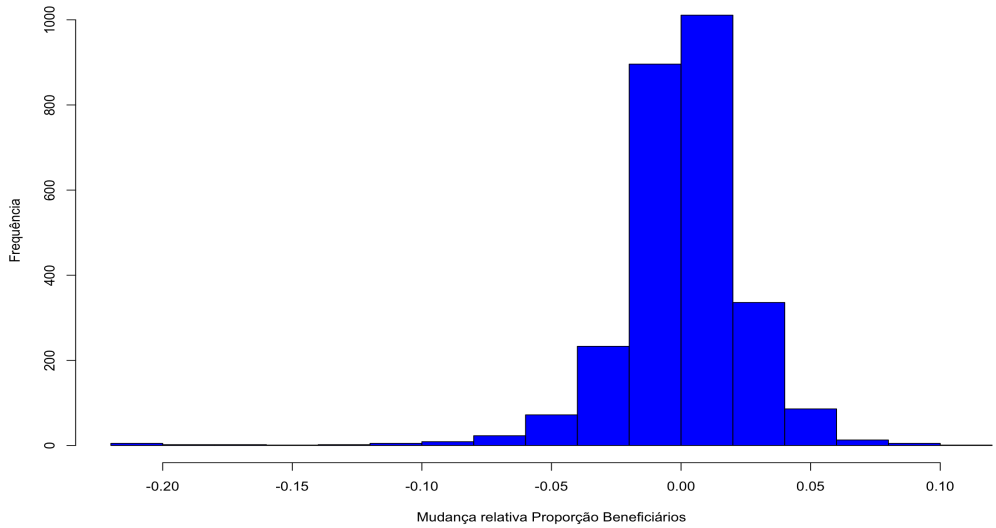
Gerard, Naritomi e Silva (2021) analisaram empiricamente o choque promovido pela mudança metodológica de 2009 e documentaram efeitos positivos no mercado de trabalho local. Para capturar a variação entre os municípios no número de beneficiários, estimaram a diferença entre a alocação em 2009 e uma alocação contrafactual, computada para dimensionar o número de beneficiários se a metodologia não fosse implementada. A partir da intensidade observada na variabilidade do choque provocado pela reforma de 2009, e dado sua fonte de variação exógena, foi possível separar os municípios em dois grupos de avaliação. Baseado em Gerard, Naritomi e Silva (2021), separou-se os municípios em função da variação no número de novos beneficiários a partir de 2009, conforme descrito na eq. (4.4):

$$Var_{BFB} = \frac{(BFB_{g}^{2009} - BFB_{contrafactual}^{2009})}{Pop_{2006}} \tag{4.4}$$

A variável construída Var_{BFB} representa o choque promovido pela Reforma de 2009. A partir da constituição dessa variável, foi possível classificar os municípios em dois grupos, em função da intensidade do choque, considerando a mediana como medida de separação. Como resultado, após a reforma metodológica de 2009, foi possível observar uma variação desproporcional no número de beneficiários entre os municípios. Essa mudança de magnitude do tamanho do programa é vista como exógena, dado que sua origem foi meramente institucional e não foi condicionada a nenhum critério anterior. A Figura 15

mostra que a variabilidade da reforma não foi homogênea entre os municípios, sendo mais impactante para um grupo de municípios do que para outros, sem haver um critério de seleção formal para novas alocações.

Figura 15 – Dispersão do choque da Reforma 2009



Fonte: Elaboração própria, Histograma da dispersão do choque provocado pela reforma metodológica no número de beneficiários do BF. O choque é representado pela variação relativa no número de beneficiários em 2009 (log) em relação à média do número de beneficiários entre 2006 e 2008 (log).

Contudo, para garantir que fosse possível observar o efeito isolado da Reforma de 2009, era preciso filtrar os municípios que já haviam sido impactados pela alocação inicialmente realizada até 2006, a qual houve o movimento de introdução de vários beneficiários ao programa dado a sua criação em 2004. A estratégia foi separar a base de municípios em grupos conforme a exposição às modificações ocorridas em 2006 e 2009. E assim, classificar os municípios nos grupos de controle e tratamento, de forma a considerar a varição do número de beneficiários como medida de separação.

O primeiro passo foi separar os municípios em função do ajuste ocorrido em 2006, em seguida, os municípios que estavam acima e abaixo da mediana do choque da Reforma de 2009, conforme a equação (4.4). O intuito foi definir o grupo de municípios tratados de forma a isolar o efeito da reforma de 2009. Consequentemente, nesse grupo estão apenas os municípios que ganharam mais beneficiários em proporção ao tamanho da população (acima da mediana da equação (4.4)), e já haviam sido beneficiados no ajuste de 2006 (Figura 16, quadrante D). Já o grupo controle (Figura 16, quadrante A), estão os municípios que haviam sido beneficiados em 2006, porém tiveram perda de cota de beneficiários em 2009. A aplicação desses filtros resultou em uma amostra de trabalho de 2.702 municípios. A Figura 16 reporta a distribuição conforme o critério destacado.

Figura 16 – Filtros aplicados na base de Municípios

Reforma 2006 Mediana = 0.2374	Reforma 2009 Mediana = 0.1624: 0	1
1	A (1.351)	B (1.431)
0	C (1.432)	D (1.351)

Fonte: Elaboração própria. A figura representa a distribuição dos municípios conforme filtros estabelecidos na base. Municípios pertencentes ao grupo tratamento hachurados em verde, e ao grupo controle em amarelo.

Por critério de robustez, outras combinações desses grupos foram testadas na estratégia empírica e os resultados, como podem ser vistos na seção 4.6, se mantiveram significativos. Além disso, como sugere Hainmueller (2012), também foi realizado um teste de falsificação.

4.3 Estratégia Empírica

A técnica utilizada para avaliar o efeito da expansão do Programa Bolsa Família no comportamento de voto presidencial foi o método de diferenças em diferenças, que além de ser uma das estratégias de análise utilizadas na avaliação de impacto e que visa corrigir problemas relacionados a possível relação endógena entre a definição dos grupos de comparação e a variável de interesse, tais como a existência de um viés de seleção e o efeito de variáveis não observáveis (ANGRIST; PISCHKE, 2008). Adicionalmente, também permite comparar uma mudança na variável de interesse ao longo do tempo, entre um grupo que recebeu algum tratamento em relação a um grupo de controle e que, por sua vez, não recebeu o tratamento em questão (ABADIE; DIAMOND; HAINMUELLER, 2010).

A validade desse método depende da hipótese de tendência paralela entre os grupos, pois se houver algum outro choque concomitante ao evento ou programa em análise, o impacto do estimador de diferenças em diferenças não será capaz de separar a fonte de variação exógena, o que torna o a estimativa viesada (DUFLO, 2001; GALIANI; GERTLER; SCHARGRODSKY, 2005). Não obstante, existem algumas formas de verificar a sua validade. Por exemplo, é possível comparar as alterações nos resultados dos grupos em momentos de tempo antes da implementação. Outra forma bastante usual é realizar um teste placebo, via a criação de um grupo de tratamento falso, isto é, que não tenha sido afetado pela política ou programa de análise. De maneira incremental, também é possível testar o efeito em uma variável diferente, e que por definição não seja correlacionada com a política ou programa, logo não deverá ter sido afetada. Caso haja algum efeito observado no estimador de diferenças em diferenças, é possível que o grupo de comparação seja falho. Por fim, outra alternativa seja a definição de mais de um grupo de controle, e se o impacto for significativo, os resultados devem ser similares (JANVRY; FINAN; SADOULET, 2012;

LIMA; SILVEIRA NETO, 2018; ARAUJO et al., 2020).

De forma estrita, para obter o efeito causal da expansão do BF na variável de interesse é desejável comparar os municípios tratados com um contrafactual. Idealmente, implicaria ser capaz de observar os mesmos municípios na ausência do tratamento. Segundo Angrist e Pischke (2008), considera-se $T = \{1, 0\}$ para denotar a presença de tratamento e define-se Y_{it} como um potencial *outcome*, por exemplo, para o município i no tempo t. Logo, temos $Y_{it}(1)$ como o *outcome* do município i, no ano t, no caso em que este foi tratado. E $Y_{it}(0)$ representa o potencial *outcome* do município i, no ano t, no caso em que o município não foi tratado.

Desse modo, o interesse consiste em verificar o *Average Treatment Effect on Treated* (ATT), que, conforme Hainmueller (2012), pode ser expresso como:

$$ATT = \mathbb{E}[Y_{i,t}(1)|T_i = 1] - \mathbb{E}[Y_{i,t}(0)|T_i = 1] \tag{4.5}$$

Contudo, somente é possível observar $\mathbb{E}[Y_{i,t}(1)|T = 1]$ e $\mathbb{E}[Y_{i,t}(0)|T = 0]$. E a partir dessa simples diferença os resultados serão viesados (ANGRIST; PISCHKE, 2008; ARAUJO et al., 2020). Para superar essa questão é necessário obter que:

$$\begin{aligned} &\mathbb{E}[Y_{i,t}(1)|T_i = 1] - \mathbb{E}[Y_{i,t}(0)|T_i = 0] \\ &= (\mathbb{E}[Y_{i,t}(1)|T_i = 1] - \mathbb{E}[Y_{i,t}(0)|T_i = 1]) \\ &+(\mathbb{E}[Y_{i,t}(0)|T_i = 1] - \mathbb{E}[Y_{i,t}(0)|T_i = 0]) \end{aligned} \tag{4.6}$$

Onde $\mathbb{E}[Y_{i,t}(1)|T_i = 1] - \mathbb{E}[Y_{i,t}(0)|T_i = 1]$ equivale ao ATT de interesse.

A mensuração dessa diferença via simples estimação OLS se mostra inconsistente devido ao viés de seleção, mesmo sob a inclusão de variáveis de controle e pelos efeitos fixos a nível de município (BERTRAND; DUFLO; MULLAINATHAN, 2004; ABADIE; DIAMOND; HAINMUELLER, 2010; IMBERT; PAPP, 2015). Holland (1986), Azuaga e Sampaio (2017) e Pinho Neto e Machado (2017) ainda argumentam que haverá comprometimento da validade interna dos resultados e potencial *overcontrolling*, dado que os coeficientes serão também viesados quando o controle é causado pela variável explicativa.

Tella e Schargrodsky (2004) e Janvry, Finan e Sadoulet (2012) expõe que o uso do método de diferenças em diferenças permite identificar o ATT sob certas hipóteses. Nessa especificação, continua-se a ter que $T = \{1, 0\}$ expressando a exposição ao tratamento, e que $t = \{1, 0\}$, denotando os período pre- e post-tratamento. Assim, o estimador de Diferenças em Diferenças é expresso por:

$$\begin{aligned}\beta_{DD} = (\mathbb{E}[Y_{i,t}|T_i = 1, t = 1] - \mathbb{E}[Y_{i,t}|T_i = 1, t = 0]) \\ -(\mathbb{E}[Y_{i,t}|T_i = 0, t = 1] - \mathbb{E}[Y_{i,t}|T_i = 0, t = 0])\end{aligned} \tag{4.7}$$

Em tal estrutura de estimação o estimador β_{DD} representa exatamente o efeito causal de interesse (ATT).

Nesse ensaio optou-se, para captar o efeito da expansão das cotas de beneficiários do Bolsa Família na proporção de votos no PT, e superar o problema de potencial endogeneidade, utilizar um modelo padrão de diferenças em diferenças, o qual é expresso por:

$$\begin{aligned}Y_{gt} = \beta_0 + \beta_1 Tempo + \beta_2 D_g + \beta_3 RF2009_{gt} \\ +\beta_4 Xg + \alpha_g + \lambda_t + \varepsilon_{gt}\end{aligned} \tag{4.8}$$

Onde Y_{gt} é a variável de interesse e representa a proporção de votos no candidato do PT para presidente no município g ocorrida no ano eleitoral t. Nesse caso o ano equivale a ocorrência da eleição presidencial entre os anos 2002 e 2018. β_3 é o estimador de diferenças em diferenças e capta o efeito causal do programa. $RF2009_{gt}$ é uma variável *dummy*, e assume valor 1 se o município g foi tratado, e 0, caso contrário. No modelo ainda existem dois efeitos α_g e λ_t que representam os efeitos fixo de município e tempo, respectivamente. X_g é um vetor de variáveis de controle. O erro aleatório será clusterizado a nível de município, como sugere Bertrand, Duflo e Mullainathan (2004).

No conjunto pertencente a X_g foram selecionadas variáveis de controle que variam no tempo e que representam características sociodemográficas dos municípios. Utilizou-se a população (log), para captar o tamanho de um município, dado o efeito variado do PBF na economia local, em face o tamanho populacional; o pib per capita (log), para obter a dinâmica econômica do município; a proporção de votos femininos e a proporção de votos de eleitores com ensino superior, para captar a estrutura de voto local; o número de empregados formais (log), com o intuito de observar a dinâmica produtiva; bem como uma variável *dummy* para municípios que pertencem à Região Metropolitana, para controlar o efeito de municípios pequenos que são afetados por metrópoles adjacentes e, com isso, sofrem o efeito *spillover*. Além dessas, incluiu-se como regressor a média de beneficiários no município g em 2006, que absorve a tendência na variação entre os municípios com condições iniciais distintas (IMBERT; PAPP, 2015; CARAZZA; SILVEIRA NETO; EMANUEL, 2021).

A eq. (4.8) será estimada, a critério de comparação, com e sem as variáveis de controle e os efeitos fixos. Assim como, será estimado via OLS puro. O intuito é verificar se

a omissão desses controles modificam (potencialmente) no tempo o estimador de diferenças em diferenças, o que pode caracterizar um viés. E a comparação com uma estimação simples via OLS, permitirá como ignorar o problema de inferência causal pode levar a estimadores inconsistentes alterando a interpretação do efeito de algum programa ou política (HOLLAND, 1986; ANGRIST; PISCHKE, 2008).

4.4 Dados

A base de dados constituída nesse trabalho é configurada como um painel que observa os resultados eleitorais para as eleições presidenciais entre 2002 e 2018. Foram utilizadas três fontes de dados principais: i) dados eleitorais, oriundos do Tribunal Superior Eleitoral (TSE); ii) dados socioeconômicos e demográficos dos municípios, obtidos no Instituto Brasileiro de Geografia e Estatística (IBGE); e iii) Folha de pagamento de beneficiários e o número de beneficiários do BF, disponibilizado pelo Ministério do Desenvolvimento Social (MDS).

A partir dos resultados eleitorais (TSE) foi possível construir as seguintes variáveis: proporção de votos presidenciais no Partido dos Trabalhadores (variável dependente), que representa o número de votos no PT em relação ao total de votos válidos no primeiro turno do ano eleitoral. As características socioeconômicas e demográficas dos municípios foram obtidas no IBGE. Essas variáveis foram: a população estimada e o PIB per capita (PIBpc) em cada ano eleitoral, a Região político-administrativa e o porte dos municípios, em termos populacionais. Os dados de pagamentos do Bolsa Família permitiram observar, por município, o número de beneficiários e o valor (em R$ nominal) recebido pelo beneficiário. Essas variáveis foram úteis para controlar o efeito do programa nos resultados eleitorais.

A Tabela 17 ilustra as estatísticas descritivas das variáveis trabalhadas neste estudo. Os dados estão separadas por dois grupos: controle e tratamento conforme descrito na seção anterior. É possível notar que, por exemplo, para os municípios que foram afetados positivamente (grupo de tratados) pela Reforma de 2009, a proporção de votos no PT subiu de 41,7% para 45,6%. Já nos municípios do grupo controle, esse incremento de voto foi menos intenso (de 41,2% para 44,2%). Outro destaque relevante é que os municípios tratados experimentaram uma mudança relativa na proporção de beneficiários do BF em torno 3.7 pontos percentuais; já o grupo controle de 1.4 pontos percentuais.

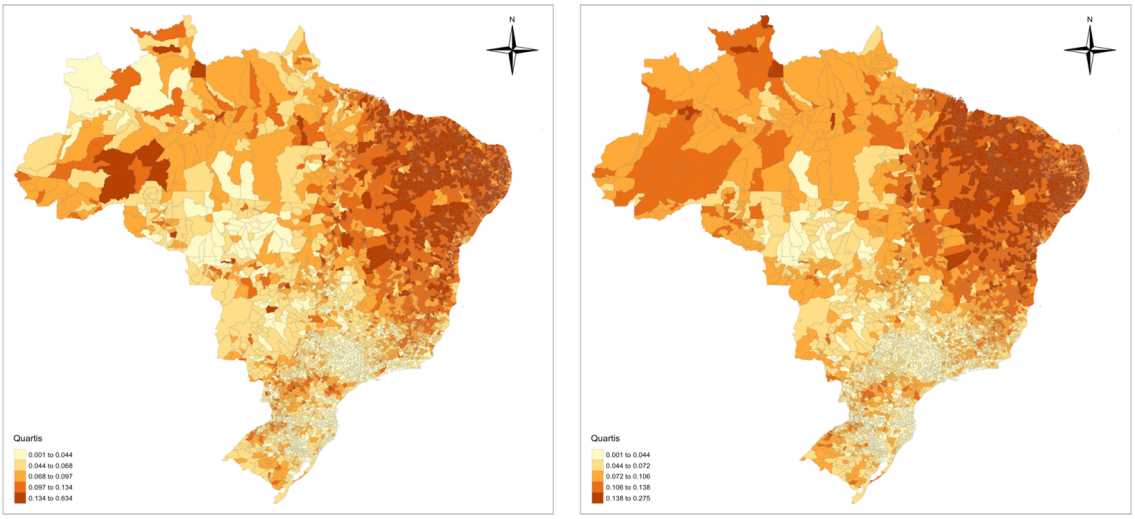
Tabela 17 – Estatísticas Descritivas por Grupo de Municípios

	Controle		Tratados	
Variáveis	Antes	Depois	Antes	Depois
Prop. voto PT (1turno)	41,2%	44,2%	41,7%	45,6%
Prop. Beneficiários BF	7,7%	9,1%	6,6%	10,3%
Prop. voto feminino	49,2%	50,1%	48,9%	50,0%
Prop. voto Ensino Superior	2,8%	2,6%	2,8%	2,4%
PIB pc	7.004,09	19.277,07	6.330,26	17.769,76
População	22.591	25.336	36.764	40.252
Emprego formal	4.630	7.459	3.613	9.591
Rendimento médio	1.096,97	1.122,72	1.060,12	1.076,93
Amostra	N	%	N	%
Norte	54	4,0	143	10,6
Nordeste	432	32,0	459	34,0
Sul	318	23,5	280	20,7
Centro-Oeste	108	8,0	130	9,6
Sudeste	439	32,5	339	25,1
Total Municípios	1.351		1.351	

A tabela expressa algumas características sócio-econômicas da amostra de municípios utilizados na estimação (2.702 de 5.565 unidades). Coluna (1) e (2) expõe informações a cerca dos municípios do grupo de controle e as colunas (3) e (4) do grupo de tratamento.

A Figura 17 expressa que a distribuição dos beneficiários em termos populacionais não mudou de forma aparente após a Reforma de 2009. É possível ver que a predominância nas regiões Nordeste e Norte do país se mantiveram. A distribuição expressa a existência de um padrão de concentração geográfico na alocação de beneficiários, sendo este mais intenso nas regiões economicamente mais pobres, pertencentes a essas regiões do país.

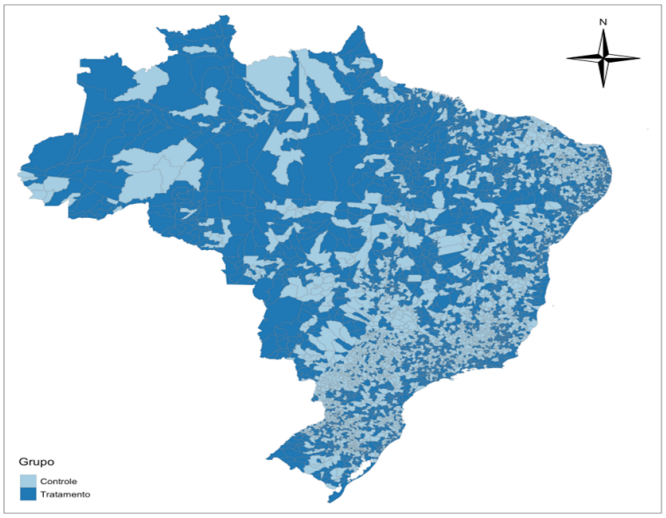
Figura 17 – Proporção de beneficiários do BF - 2006/2010



Fonte: TSE. Elaboração própria. A figura ilustra a proporção de beneficiários relativos à população em 2006 e 2010, representando a distribuição dos beneficiários antes e depois da reforma metodológica de 2009.

Por outro lado, ao se avaliar a distribuição dos municípios que compõe o grupo de tratados e de controle (Figura 18) da variação do choque provocado pela reforma de 2009, fica evidente a não uniformidade em termos de distribuição geográfica, revelando uma desconcentração a cerca dos municípios que recebam mais (menos) novas cotas para alocação de beneficiários.

Figura 18 – Variação geográfica da Reforma de 2009



Fonte: Elaboração própria. A figura ilustra distribuição geográfica da alteração no número de beneficiários intra-municípios. É possível ver que não houve regional no choque.

4.5 Resultados

Na tabela 18 são apresentados os resultados da estimação via Ordinary Least Square (OLS), com e sem variáveis socioeconômicas dos municípios, com o intuito de analisar se participar do grupo que mais receberam beneficiários devido a reforma de 2009, afetou o desempenho de votos no PT. Empiricamente, e considerando a eq. (4.8), avaliou-se o coeficiente β_2. Os resultados demonstram que pertencer ao grupo de municípios com choque positivo no crescimento na proporção de beneficiários no Bolsa Família após 2009, aumentou o voto no PT em 1%. Ao introduzirmos as variáveis de controle, verifica-se que o coeficiente estimado continua significante (0,8%). No entanto, a inclusão do efeito fixo de tempo e município altera a essência do estimador obtido, tornando-o não significante.

Adicionalmente, é possível observar que os resultados expostos na Tabela 18 são negativamente viesados, mesmo ao controlar pelas variáveis sociodemográficas e para o efeito fixo de tempo. Isso porque somente controlar pelo fato de pertencer ao grupo de controle ou tratamento, permite realizar uma comparação corrigindo apenas variáveis que não variam no tempo. Essa medida equivale apenas a primeira diferença do método de diferenças em diferenças. É preciso avaliar o efeito de mudança no tempo.

Tabela 18 – Estimação via OLS

	Proporção de votos no PT			
	OLS		*Efeito Fixo*	
	(1)	(2)	(3)	(4)
Tratamento	0.010***	0.008***	0.010***	0.003
	(0.003)	(0.002)	(0.003)	(0.002)
log(pop)		−0.004***		0.002**
		(0.001)		(0.001)
log(pibpc)		−0.034***		−0.070***
		(0.001)		(0.002)
sh_votofem		1.258***		0.495***
		(0.064)		(0.065)
sh_voto_sup		−3.195***		−2.523***
		(0.042)		(0.043)
log(emp_form)		0.003***		0.002**
		(0.001)		(0.001)
log(rend_med)		−0.002		−0.001
		(0.002)		(0.002)
Constant	0.430***	0.235***		
	(0.002)	(0.031)		
N	13,510	13,510	13,510	13,510
R^2	0.001	0.375	0.001	0.381
Note:	*p<0.1; **p<0.05; ***p<0.01			

Fonte: Elaboração própria. A Tabela resume o estimador OLS para analisar o efeito de pertencer ao grupo de municípios que aumentaram a proporção de beneficiários após 2009 no comportamento do voto no PT. Coluna (1) estima sem controles e sem efeito fixo. Coluna (2) inclui grupo de variáveis de controle e, por fim, a coluna (3) inclui controles e efeito fixo.

A Tabela 19 coloca em síntese os resultados obtidos pela estimação via OLS, comparando-os com a estimação via modelo de diferenças em diferenças. É possível ver que modelo simples via OLS produz estimativas negativamente viesadas, subestimando o efeito do programa bolsa família no comportamento de voto presidencial no PT. Por outro lado, as estimativas baseadas na eq. (4.8), se mostraram positivas e significantes, além de robustas às variáveis de controle e ao efeito fixo de tempo e município. Verifica-se que a expansão no número de beneficiários é positivo e significativo para o incremento de voto no

PT. Houve um impacto médio de 3,7%, sugerindo que a expansão do programa impactou no crescimento de voto no candidato presidencial vinculado ao PT. Esse resultado é robusto tanto à inclusão das variáveis de controle sociodemográficas quanto ao controle pelo efeito fixo de tempo e município. Os resultados sugerem que ser um município afetado pela Reforma de 2009 impactou positivamente a proporção de votos no PT.

Tabela 19 – Estimação do efeito da expansão do BF no voto presidencial

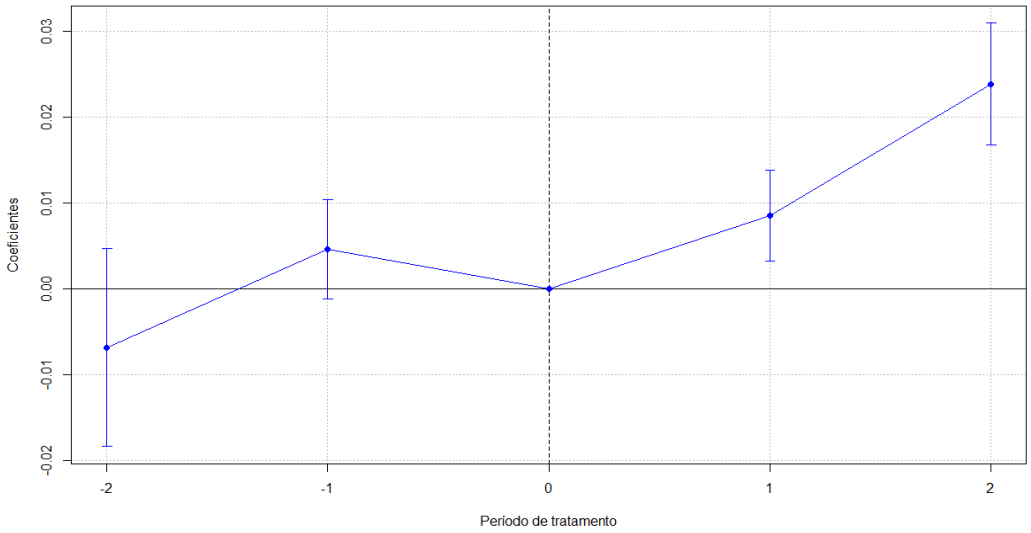
	Proporção de votos no PT					
	OLS	OLS	OLS	DID	DID	DID
	(1)	(2)	(3)	(4)	(5)	(6)
Reforma_2009	0.010***	0.008***	0.003	0.030***	0.041***	0.037***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.003)
N	13,510	13,510	13,510	13,510	13,510	13,510
R^2	0.005	0.363	0.371	0.007	0.396	0.387
Controles	N	S	S	N	S	S
EF Mun/Tempo	N	N	S	N	N	S
Note:						*p<0.1; **p<0.05; ***p<0.01

Fonte: Elaboração própria. A Tabela resume os resultados para o modelo estimado via OLS (Colunas (1) a (3)) e o estimador de diferenças em diferenças (Colunas (4) a (6)) para analisar o efeito de pertencer ao grupo de municípios que aumentaram a proporção de beneficiários após 2009 no comportamento do voto no PT.

Na Figura 19 mostra-se que o efeito da reforma de 2009 é significante nos períodos 1 e 2, equivalentes às eleições de 2014 e 2018 e expõe a tendência paralela pré-tratamento, expondo que ambos os grupos, controle e tratamento, não possuem diferenças significativas, fortalecendo a confiança na hipótese de que os municípios manteriam tendência similar, se o choque promovido pela Reforma de 2009 não tivesse ocorrido. Tal resultado é um peça de evidência de que a associação, já documentada entre o PBF e o comportamento de voto, tanto para a legenda do PT como para candidatos incumbentes, pode ser observada também como um efeito causal da expansão do programa.

Os resultados visualizados na Figura 19 mostram que o efeito do programa já pode ser observado na eleição de 2014, e é mais em 2018. Portanto, constrói-se uma evidência causal de que mais beneficiários do programa bolsa família proporcionou mais votos no candidato vinculado ao PT.

Figura 19 – Efeito da Reforma 2009 na proporção de beneficiários



Fonte: Elaboração própria. A figura ilustra a estimação do efeito causal da Reforma de 2009 na proporção de votos no PT.

4.6 Robustez

Uma ameaça à validade interna da estratégia de identificação utilizada nesse ensaio está associada à seleção dos municípios que pertencem a cada grupo. Então, como teste de robustez aos resultados encontrados na seção anterior, alterou-se a regra de definição das observações em relação a qual grupo pertencem. Os tratados foram definidos considerando os municípios afetados apenas pela Reforma de 2009, isto é, tiveram incremento na proporção de beneficiários em função da mudança metodológica apenas nesse respectivo ano. Já o controle foi definido considerando o grupo de municípios afetados em 2006 e não afetados em 2009. A Figura 20 especifica os grupos de municípios selecionados, conforme os quadrantes demarcados. Com isto, a amostra de observações ficou com 2.782 municípios pertencentes ao grupo de tratamento e 1.351 ao grupo de controle, totalizando uma total de 4.133 observações.

Figura 20 – Filtros aplicados na base de Municípios - Robustez 1

	Reforma 2009 Mediana = 0.1624	
Reforma 2006 Mediana = 0.2374	0	1
1	A (1.351)	B (1.431)
0	C (1.432)	D (1.351)

Fonte: Elaboração própria. A figura representa a distribuição dos municípios conforme filtros estabelecidos na base. Municípios tratados hachurados em verde e municípios do grupo de controle, em amarelo.

A Tabela 20 mostra que os coeficientes estimados continuam a demonstrar resulta-

dos positivos e significantes, corrigindo o viés de subestimação de uma estimação simples via OLS. Como é possível notar, a expansão do programa causou um impacto positivo de 1,6% no voto, e se mostrou robusto tanto à inclusão de variáveis de controle, como para o controle no efeito fixo de tempo e município.

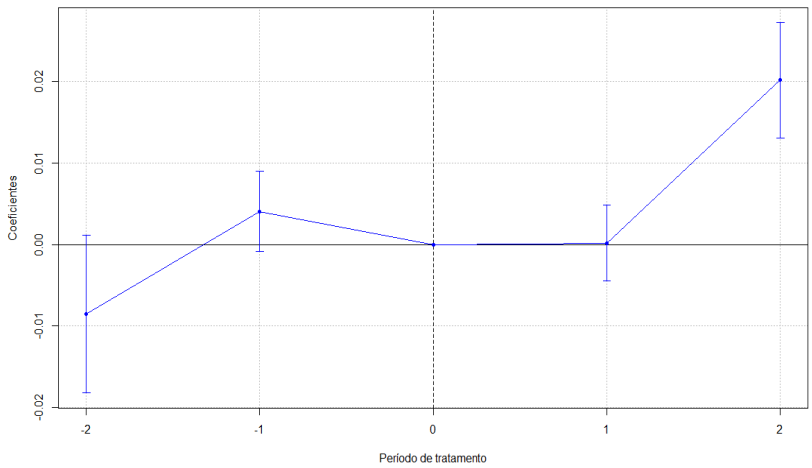
Tabela 20 – Estimação do efeito da expansão do BF no voto presidencial

	Proporção de votos no PT					
	OLS	OLS	OLS	DID	DID	DID
	(1)	(2)	(3)	(4)	(5)	(6)
Reforma_2009	0.012***	0.010***	0.001***	0.045***	0.042***	0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
N	20,865	20,865	20,865	20,865	20,865	20,865
R^2	0.001	0.365	0.366	0.015	0.37	0.74
Controles	N	S	S	N	S	S
EF Mun/Tempo	N	N	S	N	N	S
Note:						*p<0.1; **p<0.05; ***p<0.01

Fonte: Elaboração própria. A Tabela resume os resultados para o modelo estimado via OLS (Colunas (1) a (3)) e o estimador de diferenças em diferenças (Colunas (4) a (6)) para analisar o efeito de pertencer ao grupo de municípios que aumentaram a proporção de beneficiários após 2009 no comportamento do voto no PT.

A Figura 21 expõe o efeito da reforma de 2009, e mostra um impacto positivo no comportamento de voto no PT, contudo, diferentemente do observado na estimação principal (equação (4.8)), e efeito passa a ser percebido apenas no período 2 após o evento. Uma justificativa possível está associada a possibilidade do grupo tratamento conter municípios que já haviam sido beneficiados pelo aumento da proporção de beneficiários em 2006, o que por sua vez, pode ter afetado o resultado do voto na eleição subsequente.

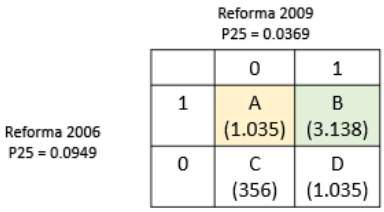
Figura 21 – Efeito da Reforma 2009 na proporção de beneficiários - Robustez 1



Fonte: Elaboração própria. A figura ilustra a estimação do efeito causal da Reforma de 2009 na proporção de votos no PT.

O segundo exercício de robustez implementado nesse ensaio via observar o efeito da reforma de 2009 dentre os municípios pertencentes apenas ao grupo que já havia observado um crescimento em 2006. Neste recorte a medida de separação entre controle e tratamento foi o percentil 25, para que o incremento de 2006 tivesse uma menor variabilidade e ainda permitisse um número de observações maiores. A amostra foi de 3.138 municípios no grupo tratamento (quadrante B), 1.035 de controle (quadrante A), totalizando 4.173 municípios para a construção do painel de avaliação. A Figura 22 ilustra a distribuição da amostra.

Figura 22 – Filtros aplicados na base de Municípios - Robustez 1



Fonte: Elaboração própria. A figura representa a distribuição dos municípios conforme filtros estabelecidos na base. Municípios tratados hachurados em verde e municípios do grupo de controle, em amarelo.

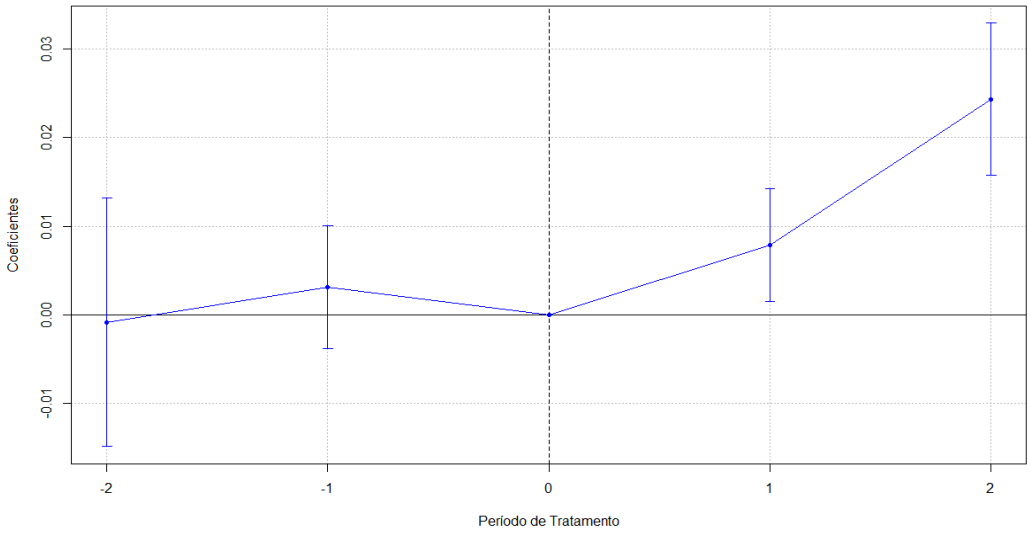
Os resultados dessa estimação são apresentados na Tabela 21 e mostram que o efeito da reforma de 2009 ainda se manteve significativo, reforçando a validade do estimador. A Figura 23 mostra o impacto positivo e valida a hipótese de tendências paralelas no modelo de diferenças em diferenças.

Tabela 21 – Estimação do efeito da expansão do BF no voto presidencial

	Proporção de votos no PT					
	OLS	OLS	OLS	DID	DID	DID
	(1)	(2)	(3)	(4)	(5)	(6)
Reforma_2009	0.010***	0.010***	0.001	0.055***	0.039***	0.012***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.002)
N	12,235	12,235	12,235	12,235	12,235	12,235
R^2	0.001	0.36	0.367	0.002	0.363	0.25
Controles	N	S	S	N	S	S
EF Mun/Tempo	N	N	S	N	N	S
Note:						*p<0.1; **p<0.05; ***p<0.01

Fonte: Elaboração própria. A Tabela resume os resultados para o modelo estimado via OLS (Colunas (1) a (3)) e o estimador de diferenças em diferenças (Colunas (4) a (6)) para analisar o efeito de pertencer ao grupo de municípios que aumentaram a proporção de beneficiários após 2009 no comportamento do voto no PT.

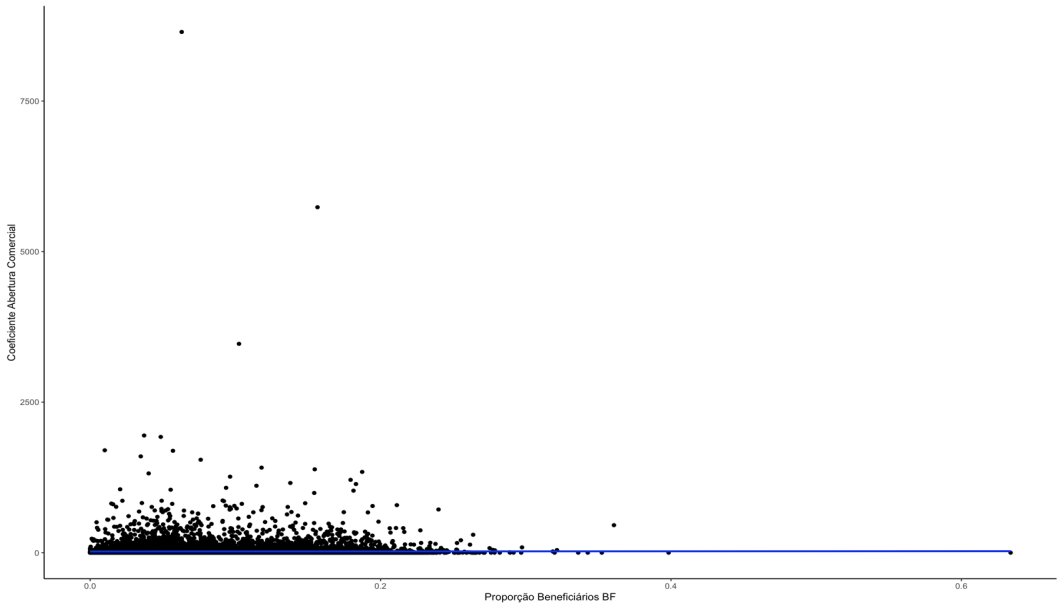
Figura 23 – Efeito da Reforma 2009 na proporção de beneficiários - Robustez 2



Fonte: Elaboração própria. A figura ilustra a estimação do efeito causal da Reforma de 2009 na proporção de votos no PT.

Em complemento aos exercícios de robustez, e verificar que nossos resultados não possam ser facilmente reproduzidos aleatoriamente, realizou-se dois testes de falsificação à nossa estimação. O primeiro atribui como variável de interesse o coeficiente de abertura comercial do município. Como pode ser visto pela Figura 24, a correlação entre a variável e o percentual de beneficiários é inexiste.

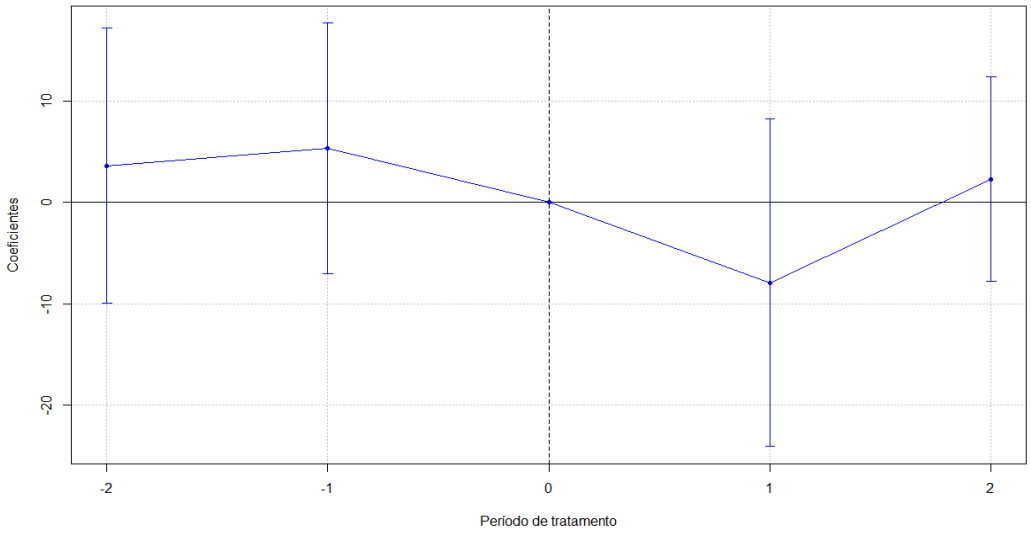
Figura 24 – Dispersão entre o percentual de beneficiários do BF e o coeficiente de abertura comercial



Fonte: Elaboração própria. A figura ilustra o gráfico de dispersão entre o percentual de beneficiários do BF e o coeficiente de abertura comercial do município

Variável: Coeficiente de Abertura Comercial

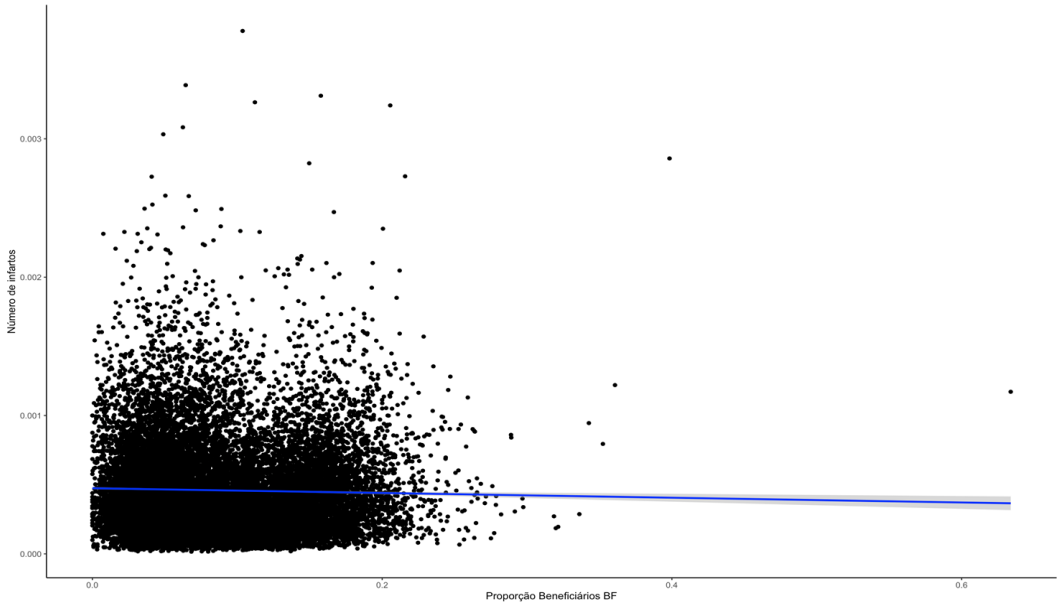
Figura 25 – Efeito da Reforma 2009 na proporção de beneficiários - Falsificação 1



Fonte: Elaboração própria. A figura ilustra a estimação do efeito causal da Reforma de 2009 na proporção de votos no PT.

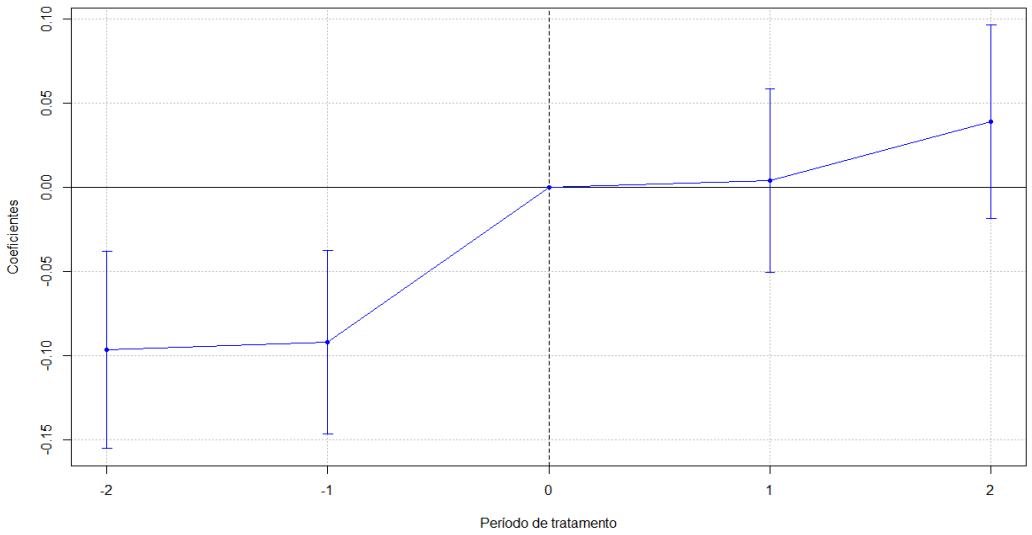
O segundo exercício de falsificação considerou como variável de interesse a proporção de ataques cardíacos, em relação à população, no município. As Figuras 26 e 23 documentam que a correlação entre a variável e o percentual de beneficiários não é significante, criando mais confiança na validade do estimador construído nesse ensaio.

Figura 26 – Dispersão entre o percentual de beneficiários do BF e o coeficiente de abertura comercial



Fonte: Elaboração própria. A figura ilustra o gráfico de dispersão entre o percentual de beneficiários do BF e o coeficiente de abertura comercial do município

Figura 27 – Efeito da Reforma 2009 na proporção de beneficiários - Falsificação 2



Fonte: Elaboração própria. A figura ilustra a estimação do efeito causal da Reforma de 2009 na proporção de votos no PT.

4.7 Conclusões

Neste artigo buscou-se encontrar evidências sobre o impacto da expansão no programa bolsa família no comportamento do voto presidencial. Avaliou-se o efeito de uma reforma administrativa no Programa Bolsa Família, ocorrida em 2009 e que promoveu uma mudança metodológica alterando, de forma expressiva e desproporcional, o número de beneficiários do programa bolsa família a nível municipal. Para lidar com a potencial correlação espúria entre o comportamento de voto e o percentual de beneficiários do BF no município, adotou-se como estratégia empírica o modelo de diferenças em diferenças, considerando como fonte variação exógena o choque da reforma de 2009, descrito em Gerard, Naritomi e Silva (2021).

Demonstrou-se que ter mais beneficiários pertencentes ao programa bolsa família têm um efeito positivo e significante no voto aos candidatos à presidente vinculados ao PT. Tal resultado é convergente com o observado por Pinho Neto (2019), contudo nesse ensaio se utilizou uma estratégia de identificação distinta, bem como a agregação de votos foi diferente. Além disso, o painel de anos considerado aqui é maior, pois a base das eleições foi entre 2002 e 2018. Outro ponto de adição à literatura é referente à estratégia de seleção dos municípios tratados e de controle, na qual se usou diversos cruzamentos para buscar isolar o efeito da Reforma de 2009, considerando a dinâmica de expansão sem aleatoriedade do programa nos municípios.

Por fim, esse ensaio constrói mais uma peça de evidência na análise dos efeitos de programas de transferência de renda no comportamento de voto. As consequências de se observar efeitos causais em políticas de transferência de renda devem ser tratadas com

cautela, dado que seus objetivos são focalizados em redução de pobreza e desigualdades e conectar seus efeitos ao comportamento de voto pode levar a um perigoso caminho em termos de desenho da política pública.

Como pontos de expansão para pesquisas futuras, tem-se que, avançar essa análise para resultados de votos municipais (prefeituras e governo) é um caminho para complementar as inferências aqui observadas, haja vista as características de voto local, cuja relação é mais individualista. Segundo, inserir na análise do Auxílio Emergencial, novo programa de governo, e substituto ao Bolsa Família, para entender se a mudança no tipo de benefício trouxe alguma alteração nas variáveis eleitorais.

Referências

ABADIE, A.; DIAMOND, A.; HAINMUELLER, J. Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, Taylor & Francis, v. 105, n. 490, p. 493–505, 2010.

ABE, Y. The equal employment opportunity law and labor force behavior of women in japan. *Journal of the Japanese and International Economies*, Elsevier, v. 25, n. 1, p. 39–55, 2011.

ABENSUR, T. C.; CRIBARI-NETO, F.; MENEZES, T. A. et al. Impactos do programa bolsa família nos resultados das eleições presidenciais no brasil em 2006. *Anais do XXXV Encontro Nacional de Economia*, Anpec Recife, v. 51, 2007.

AKAIKE, H. A new look at the statistical model identification. *IEEE transactions on automatic control*, Ieee, v. 19, n. 6, p. 716–723, 1974.

AKEE, R.; COPELAND, W.; HOLBEIN, J. B.; SIMEONOVA, E. Human capital and voting behavior across generations: Evidence from an income intervention. *American Political Science Review*, Cambridge University Press, v. 114, n. 2, p. 609–616, 2020.

ALESINA, A.; GIULIANO, P.; NUNN, N. On the origins of gender roles: Women and the plough. *The quarterly journal of economics*, Oxford University Press, v. 128, n. 2, p. 469–530, 2013.

ALONSO, W. et al. Location and land use. toward a general theory of land rent. *Location and land use. Toward a general theory of land rent.*, Cambridge, Mass.: Harvard Univ. Pr., 1964.

AMARAL, O. E. d. The victory of jair bolsonaro according to the brazilian electoral study of 2018. *Brazilian Political Science Review*, SciELO Brasil, v. 14, n. 1, 2020.

AMARAL, O. E. d.; RIBEIRO, P. F. Por que dilma de novo? uma análise exploratória do estudo eleitoral brasileiro de 2014. *Revista de Sociologia e Política*, SciELO Brasil, v. 23, n. 56, p. 107–123, 2015.

ANDERSSON, F.; HALTIWANGER, J. C.; KUTZBACH, M. J.; POLLAKOWSKI, H. O.; WEINBERG, D. H. Job displacement and the duration of joblessness: The role of spatial mismatch. *Review of Economics and Statistics*, MIT Press, v. 100, n. 2, p. 203–218, 2018.

ANGELUCCI, M.; ATTANASIO, O. Oportunidades: program effect on consumption, low participation, and methodological issues. *Economic development and cultural change*, The University of Chicago Press, v. 57, n. 3, p. 479–506, 2009.

ANGELUCCI, M.; GIORGI, G. D. Indirect effects of an aid program: how do cash transfers affect ineligibles' consumption? *American economic review*, v. 99, n. 1, p. 486–508, 2009.

ANGRIST, J. D.; PISCHKE, J.-S. *Mostly harmless econometrics: An empiricist's companion*. [S.l.]: Princeton university press, 2008.

ANSELIN, L. et al. Spatial econometrics. *A companion to theoretical econometrics*, v. 310330, 2001.

ARAUJO, D.; BAYMA, G.; MELO, C.; MENDONCA, M.; SAMPAIO, L. Do extended school day programs help students get into college? 2020.

ÅSLUND, O.; ÖSTH, J.; ZENOU, Y. How important is access to jobs? old question—improved answer. *Journal of Economic Geography*, Oxford University Press, v. 10, n. 3, p. 389–422, 2010.

ATTANASIO, O. P.; KAUFMANN, K. M. Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics*, Elsevier, v. 109, p. 203–216, 2014.

AZUAGA, F.; SAMPAIO, B. Violência contra mulher: O impacto da lei maria da penha sobre o feminicídio no brasil. *ANPEC. Anais do XX Encontro de Economia da Região Sul*, 2017.

BAEZ, J. E.; CAMACHO, A.; CONOVER, E.; ZÁRATE, R. A. *Conditional cash transfers, political participation, and voting behavior*. [S.l.]: The World Bank, 2012.

BARBANCHON, T. L.; RATHELOT, R.; ROULET, A. Gender differences in job search: Trading off commute against wage. *The Quarterly Journal of Economics*, Oxford University Press, v. 136, n. 1, p. 381–426, 2021.

BARBOSA, M. R. d. M.; SILVEIRA NETO, R. d. M. Adensamento urbano como condicionante da mobilidade: o caso da região metropolitana do recife. *Revista Brasileira de Estudos Regionais e Urbanos*, v. 11, n. 2, p. 233–250, 2017.

BARRERA-OSORIO, F.; BERTRAND, M.; LINDEN, L. L.; PEREZ-CALLE, F. Improving the design of conditional transfer programs: Evidence from a randomized education experiment in colombia. *American Economic Journal: Applied Economics*, v. 3, n. 2, p. 167–95, 2011.

BARROS, R. P. d. O.; FOGUEL, M. N. O.; ULYSSEA, G. O. Desigualdade de renda no brasil: uma análise da queda recente. Instituto de Pesquisa Econômica Aplicada (IPEA), 2006.

BARUFI, A. M. B.; HADDAD, E. A. Spatial mismatch, wages and unemployment in metropolitan areas in brazil. *REGION*, v. 4, n. 3, p. 175–200, 2017.

BAUMONT, C.; ERTUR, C.; GALLO, J. L. Spatial analysis of employment and population density: the case of the agglomeration of dijon 1999. *Geographical analysis*, Wiley Online Library, v. 36, n. 2, p. 146–176, 2004.

BERTRAND, M.; DUFLO, E.; MULLAINATHAN, S. How much should we trust differences-in-differences estimates? *The Quarterly journal of economics*, MIT Press, v. 119, n. 1, p. 249–275, 2004.

BETZ, T.; COOK, S. J.; HOLLENBACH, F. M. Spatial interdependence and instrumental variable models. *Political Science Research and Methods*, Cambridge University Press, v. 8, n. 4, p. 646–661, 2020.

BIVAND, R.; MILLO, G.; PIRAS, G. A review of software for spatial econometrics in r. *Mathematics*, Multidisciplinary Digital Publishing Institute, v. 9, n. 11, p. 1276, 2021.

BIVAND, R.; PIRAS, G. Comparing implementations of estimation methods for spatial econometrics. *Journal of Statistical Software*, v. 63, p. 1–36, 2015.

BLACK, D. A.; KOLESNIKOVA, N.; TAYLOR, L. J. Why do so few women work in new york (and so many in minneapolis)? labor supply of married women across us cities. *Journal of Urban Economics*, Elsevier, v. 79, p. 59–71, 2014.

BOHN, S. The electoral behavior of the poor in brazil: A research agenda. *Latin American Research Review*, Latin American Studies Association, v. 48, n. 2, p. 25–31, 2013.

BRAGA, M. d. S.; ZOLNERKEVIC, A. Padrões de votação no tempo e no espaço: classificando as eleições presidenciais brasileiras. *Opinião Pública*, SciELO Brasil, v. 26, n. 1, p. 1–33, 2020.

BROLLO, F.; KAUFMANN, K.; FERRARA, E. L. The political economy of program enforcement: Evidence from brazil. *Journal of the European Economic Association*, Oxford University Press, v. 18, n. 2, p. 750–791, 2020.

BUCHELI, J. F. Conditional cash transfer schemes and the politicisation of poverty reduction strategies schemes and politicisation of poverty reduction strategies. *Análisis político*, Instituto de Estudios Políticos y Relaciones Internacionales (IEPRI . . . , v. 28, n. 83, p. 19–31, 2015.

CANÊDO-PINHEIRO, M. Bolsa família ou desempenho da economia? determinantes da reeleição de lula em 2006. *Economia Aplicada*, SciELO Brasil, v. 19, p. 31–61, 2015.

CARAZZA, L.; SILVEIRA NETO, R. da; EMANUEL, L. Juvenile curfew and crime reduction: Evidence from brazil. *Papers in Regional Science*, Wiley Online Library, v. 100, n. 2, p. 561–579, 2021.

CENSO. Disponível em: http://www.censo2010.ibge.gov.br. *Acesso em 26 de maio de 2020*, v. 23, 2010.

CHAUVIN, J. P.; GLAESER, E.; MA, Y.; TOBIO, K. What is different about urbanization in rich and poor countries? cities in brazil, china, india and the united states. *Journal of Urban Economics*, Elsevier, v. 98, p. 17–49, 2017.

CHEN, Y. New approaches for calculating moran's index of spatial autocorrelation. *PloS one*, Public Library of Science, v. 8, n. 7, 2013.

CHIDAMBARAM, B.; SCHEINER, J. Understanding relative commuting within dual-earner couples in germany. *Transportation Research Part A: Policy and Practice*, Elsevier, v. 134, p. 113–129, 2020.

CHIODA, L.; MELLO, J. M. D.; SOARES, R. R. Spillovers from conditional cash transfer programs: Bolsa família and crime in urban brazil. *Economics of Education Review*, Elsevier, v. 54, p. 306–320, 2016.

COELHO, B. D. P. *Essays in applied economics: inequality and voting decision in Brazil*. Tese (Doutorado), 2017.

CORRÊA, D. S. Conditional cash transfer programs, the economy, and presidential elections in latin america. *Latin American Research Review*, JSTOR, p. 63–85, 2015.

COSTA, D. L.; KAHN, M. E. Power couples: changes in the locational choice of the college educated, 1940–1990. *The Quarterly Journal of Economics*, MIT Press, v. 115, n. 4, p. 1287–1315, 2000.

CRANE, R. Is there a quiet revolution in women's travel? revisiting the gender gap in commuting. *Journal of the American planning association*, Taylor & Francis, v. 73, n. 3, p. 298–316, 2007.

DE LA O, A. Do conditional cash transfers affect electoral behavior? evidence from a randomized experiment in mexico. *American Journal of Political Science*, Wiley Online Library, v. 57, n. 1, p. 1–14, 2013.

DUARTE, G. B.; SAMPAIO, B.; SAMPAIO, Y. Programa bolsa família: impacto das transferências sobre os gastos com alimentos em famílias rurais. *Revista de economia e sociologia rural*, SciELO Brasil, v. 47, p. 903–918, 2009.

DUARTE, L. B. *Acessibilidade ao emprego e resultados no mercado de trabalho.* Dissertação (Mestrado), 2020.

DUCH, R. M.; STEVENSON, R. Assessing the magnitude of the economic vote over time and across nations. *Electoral Studies*, Elsevier, v. 25, n. 3, p. 528–547, 2006.

DUFLO, E. Schooling and labor market consequences of school construction in indonesia: Evidence from an unusual policy experiment. *American economic review*, v. 91, n. 4, p. 795–813, 2001.

DURANTON, G.; TURNER, M. A. Urban growth and transportation. *Review of Economic Studies*, Oxford University Press, v. 79, n. 4, p. 1407–1440, 2012.

ELBERS, C.; LANJOUW, J. O.; LANJOUW, P. Micro-level estimation of poverty and inequality. *Econometrica*, JSTOR, v. 71, n. 1, p. 355–364, 2003.

ELHORST, J. P. Applied spatial econometrics: raising the bar. *Spatial economic analysis*, Taylor & Francis, v. 5, n. 1, p. 9–28, 2010.

FARRÉ, L.; JOFRE-MONSENY, J.; TORRECILLAS, J. Commuting time and the gender gap in labor market participation. IZA Discussion Paper, 2020.

FERNANDES, I. F. d. A. L.; FERNANDES, G. A. d. A. L. Economic voting in brazilian presidential elections: evidence with panel data from municipalities in são paulo. *Organizações & Sociedade*, SciELO Brasil, v. 26, p. 114–137, 2019.

FERRAZ, C.; FINAN, F. Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, v. 101, n. 4, p. 1274–1311, 2011.

FILMER, D.; SCHADY, N. Does more cash in conditional cash transfer programs always lead to larger impacts on school attendance? *Journal of development Economics*, Elsevier, v. 96, n. 1, p. 150–157, 2011.

FISCHER, M. M.; WANG, J. *Spatial data analysis: models, methods and techniques.* [S.l.]: Springer Science & Business Media, 2011.

FOGLI, A.; VELDKAMP, L. Nature or nurture? learning and the geography of female labor force participation. *Econometrica*, Wiley Online Library, v. 79, n. 4, p. 1103–1138, 2011.

FREY, A. Cash transfers, clientelism, and political enfranchisement: Evidence from brazil. *Journal of Public Economics*, Elsevier, v. 176, p. 1–17, 2019.

GALIANI, S.; GERTLER, P.; SCHARGRODSKY, E. Water for life: The impact of the privatization of water services on child mortality. *Journal of political economy*, The University of Chicago Press, v. 113, n. 1, p. 83–120, 2005.

GERARD, F.; NARITOMI, J.; SILVA, J. Cash transfers and formal labor markets: Evidence from brazil. CEPR Discussion Paper No. DP16286, 2021.

HADDAD, E. A.; BARUFI, A. M. B. From rivers to roads: Spatial mismatch and inequality of opportunity in urban labor markets of a megacity. *Habitat International*, Elsevier, v. 68, p. 3–14, 2017.

HADDAD, E. A.; VIEIRA, R. S. Mobilidade, acessibilidade e produtividade: nota sobre a valoração econômica do tempo de viagem na região metropolitana de são paulo. *Revista de Economia Contemporânea*, v. 19, n. 3, 2015.

HAGENE, T. Political clientelism in mexico: Bridging the gap between citizens and the state. *Latin American Politics and Society*, Wiley Online Library, v. 57, n. 1, p. 139–162, 2015.

HAINMUELLER, J. Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. *Political analysis*, JSTOR, p. 25–46, 2012.

HOLLAND, P. W. Statistics and causal inference. *Journal of the American statistical Association*, Taylor & Francis, v. 81, n. 396, p. 945–960, 1986.

HUNTER, W.; POWER, T. J. Rewarding lula: Executive power, social policy, and the brazilian elections of 2006. *Latin American Politics and Society*, Cambridge University Press, v. 49, n. 1, p. 1–30, 2007.

IMBERT, C.; PAPP, J. Labor market effects of social programs: Evidence from india's employment guarantee. *American Economic Journal: Applied Economics*, v. 7, n. 2, p. 233–63, 2015.

IPEA, I. d. P. E. A. *Atlas de vulnerabilidade social dos municípios brasileiros*. [S.l.]: IPEA Brasília, Brasil, 2018.

JANVRY, A. D.; FINAN, F.; SADOULET, E. Local electoral incentives and decentralized program performance. *Review of Economics and Statistics*, The MIT Press, v. 94, n. 3, p. 672–685, 2012.

KAWABATA, M.; ABE, Y. Intra-metropolitan spatial patterns of female labor force participation and commute times in tokyo. *Regional Science and Urban Economics*, Elsevier, v. 68, p. 291–303, 2018.

KELEJIAN, H. H.; PRUCHA, I. R. Estimation of spatial regression models with autoregressive errors by two-stage least squares procedures: a serious problem. *International regional science review*, Sage Publications Sage CA: Thousand Oaks, CA, v. 20, n. 1-2, p. 103–111, 1997.

LABONNE, J. The local electoral impacts of conditional cash transfers: Evidence from a field experiment. *Journal of Development Economics*, Elsevier, v. 104, p. 73–88, 2013.

LAMEIRA, V.; GOLGHER, A. Mobilidade para trabalho na região metropolitana de são paulo a partir dos dados do censo demográfico de 2010. *Revista Brasileira de Estudos de População*, SciELO Brasil, v. 38, 2021.

LESAGE, J. What regional scientists need to know about spatial econometrics. *Rev. Reg. Stud*, p. 44, 13–32, 2014.

LESAGE, J.; PACE, R. K. *Introduction to spatial econometrics*. [S.l.]: Chapman and Hall/CRC, 2009.

LESAGE, J. P.; PACE, R. K. Spatial econometric models. In: *Handbook of applied spatial analysis*. [S.l.]: Springer, 2010. p. 355–376.

LIMA, R. C. d. A.; SILVEIRA NETO, R. d. M. Secession of municipalities and economies of scale: Evidence from brazil. *Journal of Regional Science*, Wiley Online Library, v. 58, n. 1, p. 159–180, 2018.

LINS, R.; SILVA, L.; FILHO, D. F.; ROCHA, E. O bolsa família e as eleições presidenciais no brasil: um modelo de predição eleitoral. *Revista Eletrônica de Ciência Política*, v. 7, n. 1, 2016.

MAGALHÃES, A. M.; SILVA, M. E. A. d.; DIAS, F. d. M. Eleição de dilma ou segunda reeleição de lula? uma análise espacial do pleito de 2010. *Opinião Pública*, SciELO Brasil, v. 21, n. 3, p. 535–573, 2015.

MANACORDA, M.; MIGUEL, E.; VIGORITO, A. Government transfers and political support. *American Economic Journal: Applied Economics*, v. 3, n. 3, p. 1–28, 2011.

MARGALIT, Y. Political responses to economic shocks. *Annual Review of Political Science*, Annual Reviews, v. 22, p. 277–295, 2019.

MARZAGÃO, T. A dimensão geográfica das eleições brasileiras. *Opinião Pública*, SciELO Brasil, v. 19, n. 2, p. 270–290, 2013.

MILLER, S.; JOHNSON, N.; WHERRY, L. R. Medicaid and mortality: new evidence from linked survey and administrative data. *The Quarterly Journal of Economics*, Oxford University Press, v. 136, n. 3, p. 1783–1829, 2021.

MILLS, E. S. An aggregative model of resource allocation in a metropolitan area. *The American Economic Review*, JSTOR, v. 57, n. 2, p. 197–210, 1967.

MORAN, P. Notes on continuous stochastic phenomena. *Biometrika*, p. 37 (1), 17–23., 1950.

MUNDIM, P. S.; VIDIGAL, R.; MICHELOTTI, F. C.; PINTO, A. J. d. P. O programa bolsa família e seus beneficiários na opinião pública brasileira. *Opinião Pública*, SciELO Brasil, v. 25, p. 556–576, 2020.

MUTH, R. F. Cities and housing: the spatial pattern of urban residential land use. 1969.

NICOLAU, J. O sistema eleitoral de lista aberta no brasil. *Dados*, SciELO Brasil, v. 49, p. 689–720, 2006.

NICOLAU, J. Determinantes do voto no primeiro turno das eleições presidenciais brasileiras de 2010: uma análise exploratória. *Opinião Pública*, SciELO Brasil, v. 20, n. 3, p. 311–325, 2014.

OMMEREN, J. N. V.; PUIGARNAU, E. Gutiérrez-i. Are workers with a long commute less productive? an empirical analysis of absenteeism. *Regional Science and Urban Economics*, Elsevier, v. 41, n. 1, p. 1–8, 2011.

ONG, P. M.; MILLER, D. Spatial and transportation mismatch in los angeles. *Journal of Planning Education and Research*, Sage Publications Sage CA: Thousand Oaks, CA, v. 25, n. 1, p. 43–56, 2005.

PEREIRA, A. E. G.; SHIKIDA, C.; RIBEIRO, F. G.; NAKABASHI, L. A eleição de dilma em 2010 e seus determinantes: evidências empíricas do programa bolsa família. *Análise Econômica*, v. 33, n. 64, 2015.

PEREIRA, R. H. M.; SCHWANEN, T. *Commute time in Brazil (1992-2009): differences between metropolitan areas, by income levels and gender*. [S.l.], 2015.

PINHO NETO, V. R.; BERRIEL, C. M. Conditional cash transfers and nutrition: An evaluation of the bolsa familia program in rural and urban areas of brazil. *Economia Aplicada*, v. 21, n. 2, p. 185–205, 2017.

PINHO NETO, V. R.; MACHADO, C. Conditional cash transfers and nutrition: An evaluation of the bolsa familia program in rural and urban areas of brazil. *Economia Aplicada*, v. 21, n. 2, p. 185–205, 2017.

PINHO NETO, V. R. d. *Economic and behavioral effects of safety net programs*. Tese (Doutorado) — Fundação Getúlio Vargas, Escola de Pós-graduação em Economia, Doutorado em Economia, São Paulo, 2019.

RAPINO, M. A.; COOKE, T. J. Commuting, gender roles, and entrapment: A national study utilizing spatial fixed effects and control groups. *The Professional Geographer*, Taylor & Francis, v. 63, n. 2, p. 277–294, 2011.

RENNÓ, L. The bolsonaro voter: Issue positions and vote choice in the 2018 brazilian presidential elections. *Latin American Politics and Society*, Cambridge University Press, v. 62, n. 3, p. 1, 2020.

SÁNCHEZ, M. I. O.; GONZÁLEZ, E. M. Gender differences in commuting behavior: Women's greater sensitivity. *Transportation research procedia*, Elsevier, v. 18, p. 66–72, 2016.

SÁNCHEZ, O.; ISABEL, M.; GONZÁLEZ, E. M. Travel patterns, regarding different activities: work, studies, household responsibilities and leisure. *Transportation Research Procedia*, Elsevier, v. 3, p. 119–128, 2014.

SCHUHLI, G. T. O partido dos trabalhadores e o voto católico no segundo turno da eleição presidencial de 2010: uma análise espacial a nível municipal. *Revista da FAE*, v. 21, n. 1, p. 156–167, 2018.

SEABRA, D. M. d. S.; SILVEIRA NETO, R. d. M.; MENEZES, T. A. de. Urban amenities and real state value: An empirical analysis for recife city. *Economia Aplicada*, v. 20, n. 1, p. 143–169, 2016.

SHEI, A.; COSTA, F.; REIS, M. G.; KO, A. I. The impact of brazil's bolsa família conditional cash transfer program on children's health care utilization and health outcomes. *BMC international health and human rights*, Springer, v. 14, n. 1, p. 1–9, 2014.

SHIKIDA, C. D.; MONASTERIO, L. M.; JR, A. F. d. A.; CARRARO, A.; DAMÉ, O. M. 'it is the economy, companheiro!': an empirical analysis of lula's re-election based on municipal data. *Economics Bulletin*, AccessEcon, v. 29, n. 2, p. 976–991, 2009.

SILVEIRA NETO, R.; MOURA, K. Commuting time, public exposure and victimization: Evidence from brazilian metropolitan regions. *Papers in Regional Science*, Wiley Online Library, v. 98, n. 2, p. 1159–1175, 2019.

SILVEIRA NETO, R. d. M.; DUARTE, G.; PÁEZ, A. Gender and commuting time in são paulo metropolitan region. *Urban Studies*, Sage Publications Sage UK: London, England, v. 52, n. 2, p. 298–313, 2015.

SIMONI, S. Efeitos diretos e indiretos do programa bolsa família nas eleições presidenciais brasileiras. *Opinião Pública*, SciELO Brasil, v. 27, p. 230–260, 2021.

STECKLOV, G.; WINTERS, P.; TODD, J.; REGALIA, F. et al. Demographic externalities from poverty programs in developing countries: experimental evidence from latin america. *American University: Washington DC Department of Economics Working Paper Series*, n. 2006-1, 2006.

STOKES, S. C.; DUNNING, T.; NAZARENO, M.; BRUSCO, V. *Brokers, voters, and clientelism: The puzzle of distributive politics*. [S.l.]: Cambridge University Press, 2013.

TEDESCO, A. C. F.; SOUZA, K. B. Ser mulher importa? determinantes, evidências e estimativas da participação feminina no mercado de trabalho brasileiro. *Textos de Economia*, v. 23, n. 1, p. 1–21, 2020.

TELLA, R. D.; SCHARGRODSKY, E. Do police reduce crime? estimates using the allocation of police forces after a terrorist attack. *American Economic Review*, v. 94, n. 1, p. 115–133, 2004.

VEGA, S.; ELHORST, J. P. Regional labour force participation across the european union: A time–space recursive modelling approach with endogenous regressors. *Spatial Economic Analysis*, Routledge, v. 12, n. 2-3, p. 138–160, 2017.

VEGA, S. H.; ELHORST, J. P. The slx model. *Journal of Regional Science*, Wiley Online Library, v. 55, n. 3, p. 339–363, 2015.

ZUCCO, C. Conditional cash transfers and voting behavior: Redistribution and clientelism in developing democracies. *Unpublished manuscript. Princeton University*, 2011.

ZUCCO, C. When payouts pay off: Conditional cash transfers and voting behavior in brazil 2002–10. *American journal of political science*, Wiley Online Library, v. 57, n. 4, p. 810–822, 2013.

ZUCCO, C. The impacts of conditional cash transfers in four presidential elections (2002–2014). *Brazilian Political Science Review*, SciELO Brasil, v. 9, p. 135–149, 2015.

ZUCCO, C.; LUNA, J. P.; BAYKAL, O. G. Do conditionalities increase support for government transfers? *The Journal of Development Studies*, Taylor & Francis, v. 56, n. 3, p. 527–544, 2020.