

Jeniffer Gonçalves

**Estimação de efeito de exposição  
midiática negativa em campanhas  
eleitorais: o caso da eleição para prefeito  
de São Paulo em 2016**

Florianópolis

2018



Jeniffer Gonçalves

**ESTIMAÇÃO DE EFEITO DE EXPOSIÇÃO  
MIDIÁTICA NEGATIVA EM CAMPANHAS  
ELEITORAIS: O CASO DA ELEIÇÃO PARA  
PREFEITO DE SÃO PAULO EM 2016**

Dissertação submetida ao Programa de Pós-Graduação em Economia da Universidade Federal de Santa Catarina para a obtenção do Grau de Mestra em Economia.

**Orientador:** Prof. Dr. Francis Carlo Petterini Lourenço

Universidade Federal de Santa Catarina – UFSC  
Programa de Pós-Graduação em Economia – PPGECO

Florianópolis

2018

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Gonçalves, Jeniffer

Estimação de efeito de exposição midiática  
negativa em campanhas eleitorais : o caso da eleição  
para prefeito de São Paulo em 2016 / Jeniffer  
Gonçalves ; orientador, Francis Carlo Petterini  
Lourenço, 2018.

109 p.

Dissertação (mestrado) - Universidade Federal de  
Santa Catarina, Centro Sócio-Econômico, Programa de  
Pós-Graduação em Economia, Florianópolis, 2018.

Inclui referências.

1. Economia. 2. Teoria da escolha. 3. Modelo  
Bayesiano. 4. Eleições. I. Lourenço, Francis Carlo  
Petterini . II. Universidade Federal de Santa  
Catarina. Programa de Pós-Graduação em Economia. III.  
Título.

Jeniffer Gonçalves

**ESTIMAÇÃO DE EFEITO DE EXPOSIÇÃO  
MIDIÁTICA NEGATIVA EM CAMPANHAS  
ELEITORAIS: O CASO DA ELEIÇÃO PARA  
PREFEITO DE SÃO PAULO EM 2016**

Esta dissertação foi julgada adequada para obtenção do Título de Mestra em Economia, e aprovada em sua forma final pelo Programa de Pós-graduação em Economia da Universidade Federal de Santa Catarina.

Florianópolis, 5 de março de 2018

---

**Prof. Jaylson Jair da Silveira, Dr.**  
Coordenador do Curso

**Banca Examinadora:**

---

**Orientador – Prof. Francis Carlo Petterini  
Lourenço, Dr.**

Universidade Federal de Santa Catarina

---

**Prof. Guilherme Valle Moura, Dr.**

Universidade Federal de Santa Catarina

---

**Prof. Gilson Geraldino Silva Júnior, Dr.**

Universidade Federal de Santa Catarina

---

**Prof. Sérgio Aquino de Souza, Dr.**

Universidade Federal do Ceará

# AGRADECIMENTOS

Meus sinceros agradecimentos:

À minha família, pelo amor, carinho e especialmente pelo apoio psicológico.

À CAPES, pela manutenção da bolsa de auxílio.

Aos meus colegas de mestrado, em especial à Amanda Fernandes, pela infinidade de cafés, conversas e lágrimas que compartilhamos nesses dois anos, e ao Mateus Noriller e a Aisha Schmidt, pela inspiração e por sempre me motivarem a continuar.

À minha grande amiga Adriane Merten, por estar sempre me apoiando, não importa o que aconteça.

À Tatyane Macedo, por toda ajuda com a correção dessa dissertação e pelo lar que compartilhamos nos últimos dois anos.

À Evelise Elpo da Silveira, por manter o PPGEÇO sempre em ordem e por sempre arrancar um sorriso meu.

Aos professores e professoras do Programa, por compartilharem o conhecimento.

Por fim, ao meu orientador, Prof. Dr. Francis Carlo Petterini, não apenas pelas instruções para elaboração dessa dissertação, mas também por todo apoio moral nos momentos mais difíceis, sempre me incentivando a manter o jardim florido.



# RESUMO

Desde que a democracia deixou de ser direta e passou a ser representativa, a decisão do voto passou a exigir do eleitor conhecimentos não apenas sobre as questões socioeconômicas da sua região como também sobre a capacidade dos candidatos em representar as necessidades daqueles que os elegem. Considerando que, em uma sociedade moderna, os meios de comunicação de massa representam a principal fonte de informação para os indivíduos, o que acontece quando os candidatos ou partidos são expostos de forma excessiva?

Este trabalho buscou estimar o impacto causado pela exposição midiática sobre o Partido dos Trabalhadores na eleição de 2016, utilizando o município de São Paulo como cenário para o estudo. A análise empírica utilizou os resultados das eleições de 2012 e 2016 computados por zonas eleitorais na capital paulista, além das características dos candidatos como variáveis explicativas. Como a exposição da mídia não é uma variável facilmente mensurável, foi utilizado o método de diferença-em-diferenças para conseguir estimar o efeito causado pela exposição. Os parâmetros do modelo foram estimados a partir do método econométrico Bayesiano.

Foram estimados três modelos: um simples, com todos os coeficientes fixos entre as zonas eleitorais; um com efeito fixo para a exposição midiática e efeitos aleatórios para as demais; e um com todos os coeficientes apresentando efeitos aleatórios. Os resultados encontrados indicam que se o PT não tivesse passado por exposição midiática excessiva, seu percentual de votos na eleição para a prefeitura de São Paulo em 2016 teria sido em

torno de 4 pontos percentuais maior.

**Palavras-chave:** Teoria da escolha. Modelo Bayesiano. Eleições.

# ABSTRACT

Since democracy has ceased to be direct and has become representative, the decision to vote has required voters not only to know about the socioeconomic issues of their region but also about the candidates' ability to represent the needs of those who elect them. Whereas in a modern society the mass media are the main source of information for individuals, what happens when candidates or political parties are excessively exposed?

This study sought to estimate the impact of the media exposure on the Workers' Party (PT) in the 2016 election, using the city of São Paulo as a case study. The empirical analysis used the results of the elections of 2012 and 2016, computed by electoral zones in the city of São Paulo, besides the candidates characteristics as explanatory variables. The media exposure is not an easily measurable variable, so the difference-in-differences method was used to estimate the effect caused by the exposure. The model parameters were estimated from the Bayesian econometric method.

Three models were estimated: a simple one, with fixed effect between the electoral zones for all coefficients; one with a fixed effect for the media exposure coefficient and random effects for the others; and one with all coefficients showing random effects. The results indicate that, if the workers' party had not been exposed excessively by the media its percentage of votes in the election for the city of São Paulo in 2016 would have been around 4 percentage points higher.

**Keywords:** Choice theory. Bayesian Approach. Elections.



# LISTA DE ILUSTRAÇÕES

Figura 1 – Participação nos votos em relação ao eleitorado total nas eleições de 2012 e 2016 para a prefeitura de São paulo . . . . .	72
Figura 2 – Cenário contrafactual para o resultado da eleição de São Paulo em 2016, em termos de votos totais, a partir da Eq.(3.14) . . . . .	86
Figura 3 – Cenário contrafactual para o resultado da eleição de São Paulo em 2016, em termos de votos totais, a partir da Eq.(3.15) . . . . .	88
Figura 4 – Cenário contrafactual para o resultado da eleição de São Paulo em 2016, em termos de votos válidos. . . . .	89



# LISTA DE TABELAS

Tabela 1 – Características associadas aos candidatos à prefeitura de São Paulo na eleição de 2012. . .	74
Tabela 2 – Características associadas aos candidatos à prefeitura de São Paulo na eleição de 2016. . .	75
Tabela 3 – Parametrização . . . . .	79
Tabela 4 – Resultado a Posteriori para os Parâmetros Seleccionados . . . . .	82
Tabela 5 – Elasticidade da exposição midiática do PT na proporção de votos de cada partido na eleição de 2016 em São Paulo – Calculada a partir da Eq.(3.14) . . . . .	85
Tabela 6 – Elasticidade da exposição midiática do PT na proporção de votos de cada partido na eleição de 2016 em São Paulo – Calculada a partir da Eq.(3.15) . . . . .	87
Tabela 7 – Resultados a Posteriori para o Modelo $M_1$ .	101
Tabela 8 – Resultados a Posteriori para o Modelo $M_2$ .	102
Tabela 9 – Efeito de cada covariada específico a cada zona eleitoral no Modelo $M_2$ . . . . .	103
Tabela 10 – Resultados a Posteriori para o Modelo $M_3$ .	106
Tabela 11 – Efeito de cada covariada específico a cada zona eleitoral no Modelo $M_3$ . . . . .	107



# SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>17</b>
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>27</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>35</b>
<b>3.1</b>	<b>O problema da escolha</b>	<b>36</b>
<b>3.2</b>	<b>Método de contrafactuais</b>	<b>46</b>
<b>3.3</b>	<b>Abordagem Bayesiana</b>	<b>52</b>
3.3.1	Econometria Bayesiana	53
3.3.2	Modelo de Regressão Linear com Dados em Paineis	58
<b>4</b>	<b>DADOS</b>	<b>71</b>
<b>5</b>	<b>ANÁLISE EMPÍRICA</b>	<b>77</b>
<b>5.1</b>	<b>Parametrização</b>	<b>77</b>
<b>5.2</b>	<b>Resultados</b>	<b>81</b>
<b>6</b>	<b>CONCLUSÃO</b>	<b>91</b>
	<b>REFERÊNCIAS</b>	<b>95</b>
	<b>APÊNDICE A – RESULTADOS</b>	<b>101</b>



# 1 INTRODUÇÃO

Na Atenas do século V a.C., as principais decisões a respeito da cidade e de seus habitantes, como as relacionadas à economia, aos impostos e mesmo à defesa contra ataques externos, eram tomadas em conjunto pela assembleia de cidadãos. Essa, por sua vez, era integrada apenas por homens livres, maiores de 21 anos, filhos de pai e mãe atenienses. Assim, cada cidadão pertencente a assembleia tinha direito ao voto direto nas decisões de interesse coletivo. Esse sistema político, no qual as pessoas tomam decisões coletivamente, recebeu o nome de democracia, que no sentido clássico conota “governo do povo” — do grego: *demos* (povo), *krateîn* (governar) (ELLIOTT, 1994).

Desde sua origem na Grécia, o sistema democrático foi passando por diversas transformações, especialmente em relação à quem pode exercer o poder do voto. Na maioria das democracias modernas, todos os cidadãos adultos conquistaram o direito de votar. Por outro lado, devido, principalmente, ao crescimento populacional, a democracia com voto direto, como era na Atenas antiga, passou a ser impraticável. Atualmente, quando se fala em sistema democrático, se adiciona implicitamente o adjetivo “representativo”. Isto é, o povo não vota mais com a finalidade de tomar uma determinada decisão diretamente, agora o voto serve para eleger os representantes do povo, e serão esses representantes que tomarão as decisões (MIGUEL, 2004).

Dessa forma, a democracia é vista mais como um instrumento para produzir governos legítimos, em que, ao votar, o

eleitor estaria passando o seu direito de decisão ao representante escolhido pela maioria. Assim, os candidatos eleitos ficam encarregados de tomar as decisões sem a necessidade de consultar o povo, pois suas opiniões já foram expressas no voto. Com base nisso, seria mais apropriado abandonar o sentido de democracia como “governo do povo” e substituí-lo por “um governo aprovado pelo povo”, como sugere Schumpeter (1961).

Entretanto, garantir que a decisão dos governantes reflita a vontade dos governados é um dos problemas mais importantes dentro da democracia moderna. Miguel (2004) considera a existência de vários obstáculos a serem superados para o aprimoramento desse vínculo eleitoral, como, por exemplo: o desinteresse de boa parte da população pelas questões públicas; a tendência que os representantes têm em desenvolver interesses próprios ligados às posições de poder que ocupam; a influência superior que alguns grupos especiais possuem sobre os tomadores de decisão (como os financiadores de campanha); e, acima de tudo, o obstáculo encontrado no fluxo de informação, considerando que o eleitor precisa estar bem informado sobre os candidatos, sobre a situação social de sua região, e do país, tarefa que, nas sociedades modernas, recai sobre sistemas específicos que formam os meios de comunicação.

Singer (2001) sugere que, por ser o principal meio de obter informações políticas, a mídia tem um papel de extrema importância na determinação do comportamento do cidadão, sendo, portanto, uma variável que deve ser incluída ao fazer previsões do comportamento eleitoral. Avelar (1992) apresenta alguns casos da história política pós-redemocratização do Brasil que considera representarem fortes indícios de que a mídia tem

influenciado no curso das campanhas eleitorais, como nas eleições para a prefeitura de São Paulo em 1985 e 1988.

No início de novembro de 1985, a vitória para o candidato Fernando Henrique Cardoso parecia certa, até um jornalista perguntar, em uma transmissão em rede nacional, se o candidato acreditava em Deus. Entendendo que o assunto era delicado, Cardoso tentou evitar a resposta, mas acabou dizendo que não acreditava. O resultado da eleição deu a vitória a Jânio Quadros, o então segundo colocado nas pesquisas, com diferença de apenas 3 pontos percentuais.

Em 1988, Paulo Maluf e Luíza Erundina estavam em condição de empate técnico nos dias que precediam à eleição. Contudo, uma semana antes da eleição, começaram a ser exibidas na TV imagens de repressão policial contra operários e professores. Com isso, a candidata Luíza Erundina, que estava concorrendo pelo Partido dos Trabalhadores (PT), redobrou seus esforços de oposição ao atual governo, considerando que estas imagens lhe favorecia, e acabou sendo eleita com uma margem de 7% de diferença em relação a Paulo Maluf.

Para Singer (2001), o impeachment do presidente Fernando Collor de Mello, em 1992, também teve influência da mídia. Segundo o autor, houve uma corrida regida por mecanismos de mercado, em busca de notícias exclusivas que denunciavam os negócios ilícitos do governo, especialmente na mídia impressa, provocando o isolamento político do presidente. Mais tarde, essa disputa pela última notícia se alastrou para os telejornais, culminando na derrubada do então presidente. Entretanto, para o autor, o que começou com um papel benéfico de policiamento da atividade política, logo se transformou em uma atitude de

constantes ataques à classe política por parte da grande imprensa, com a consequente desmoralização tanto dos políticos quanto das instituições representativas.

A crescente parcialidade crítica e a alta concentração no controle dos meios de comunicação são fatores que, para Singer (2001), podem prejudicar a natureza da pluralidade das informações e conseqüentemente fortalecer a influência da mídia nas campanhas eleitorais. O autor já sugeria que deveriam ser feitas mudanças na legislação dos meios de comunicação para evitar a formação de monopólio, além da necessidade de discutir medidas para dispersar a propriedade dos meios de comunicação. Entretanto, segundo a pesquisa Monitoramento da Propriedade da Mídia, realizada pela ONG brasileira Intervezes e pela Repórteres Sem Fronteiras (RSF) da França, a propriedade dos meios de comunicação continua altamente concentrada. Apenas 4 redes de TV acumulam mais de 70% da audiência nacional, sendo que uma delas detém 36,9% do total (INTERVOZES e RSF, 2017).

Essa concentração dos meios de comunicação é impulsionada, em parte, pelo perfil dos consumidores de mídia brasileiros. Um estudo da Secretaria de Comunicação Social sobre o hábito de consumo de mídia da população brasileira, (PMB, 2016), mostrou que 63% dos entrevistados costumam obter informações sobre o que acontece no Brasil por meio de programas televisivos. Com 26%, a internet foi o segundo meio mais escolhido. A pesquisa ainda revelou que as informações obtidas pela TV têm mais credibilidade para os entrevistados do que aquelas provenientes da internet. Já 54% afirmaram que confiavam muitas vezes ou sempre nas informações apresentadas na TV, já com relação à internet, 20% confiavam muitas vezes ou sempre quando liam as

reportagens em sites. Esse percentual reduz para 14% quando as informações eram lidas nas redes sociais. Isso mostra que, mesmo com o avanço da popularização da internet, a televisão ainda se sobressai como principal meio informativo.

Os recentes acontecimentos na política brasileira reacenderam a discussão sobre a influência que a mídia tem no meio eleitoral. O impeachment de Dilma Rousseff, em 2016, também foi visto como resultado da forte exposição negativa nos meios de comunicação. Para Almeida e Lima (2016) e Silva e Hashiguti (2016), a grande mídia brasileira adotou posicionamento claro de oposição, formando uma imagem desfavorável tanto para a presidente quanto para o seu partido, o PT.

Esse efeito parece ter se propagado para as eleições municipais subsequentes. Segundo Singer (2016), o megaproceto de investigação criminal, Lava Jato, cujos efeitos seletivos concentraram-se durante os primeiros meses no PT, contribuíram para a desmoralização do partido. Assim, a repercussão tanto do impedimento do governo de Dilma quanto as denúncias de envolvimento dos principais líderes de seu partido em casos de corrupção ajudam a entender o fracasso do Partido dos Trabalhadores nas eleições municipais de 2016, especialmente na região metropolitana de São Paulo, considerada como cinturão histórico da legenda (SINGER, 2016).

Alguns acontecimentos inesperados na eleição na capital paulista reforçam essa hipótese de que alguma força externa tem influenciado na decisão dos votos dos eleitores. Primeiro, a eleição foi definida em um turno, o que não era esperado, dado que todas as pesquisas de intenção de voto indicavam dois turnos (CAPUTO, 2016). Segundo, o avanço no desalento dos

eleitores em relação a classe política, caracterizado pela elevada parcela de não votos (34,9%) — que inclui os votos brancos, os nulos e as abstenções. O que mais chama a atenção, nesse caso, é que o montante desses não-votos superou a quantidade de votos recebidos pelo candidato eleito, João Dória (PSDB), que recebeu 34,7% de votos em relação ao eleitorado total (53,29% se considerar apenas os votos válidos, que exclui da contagem os não votos). Tanto a decisão da eleição em primeiro turno quanto a vitória de um candidato que recebeu menos votos do que os considerados não votos são fatos inéditos no município de São Paulo.

Para Avelar (1992), apesar de parecer evidente que os meios de comunicação influenciam, de alguma forma, no curso das campanhas eleitorais, existe uma grande dificuldade em conseguir obter evidência empírica desse impacto. Os autores que se propuseram a investigar os efeitos da mídia nas eleições, como será apresentado mais a frente, chegaram a resultados indicando que o impacto seria mínimo. Insatisfeitos com os resultados, consideravam que mudanças na metodologia e/ou na base de dados utilizada poderiam proporcionar resultados mais satisfatórios.

Entendendo a importância que essa análise tem para compreender o comportamento eleitoral, e, dado que, pouco tem se progredido em termos de metodologia que possibilite estimar a magnitude do efeito da exposição midiática, este trabalho pretende contribuir para a literatura de estudos eleitorais com uma abordagem metodológica, aparentemente, ainda pouco explorada para este propósito.

Em teoria, a decisão do voto pode ser resumida em uma função como  $y = f(x, z)$ , em que o voto do eleitor seria

condicionado por dois elementos principais, um deles seria a convicções do eleitor ( $x$ ) e o outro poderia ser considerado como sugestões externas ( $z$ ). Isto é, uma parte da escolha do voto estaria relacionada com o que o eleitor espera de um governante, como por exemplo, o posicionamento político. A outra parte seria influenciada por opiniões externas recebidas, principalmente pela mídia, que podem tanto sustentar quanto abrandar as convicções do eleitor.

A proposta aqui é, a partir de um modelo logit multinomial (comum para investigar problemas de escolha discreta como eleições), utilizar um método de inferência causal dentro do arcabouço de contrafactuais desenvolvido por Neyman (1923) e Rubin (1974), muito explorado para avaliação do impacto de políticas públicas, com o intuito de separar o efeito causado pela exposição da mídia,  $z$ , dos efeitos causados por outras variáveis que também afetam na opção do voto,  $x$ , como as características dos candidatos, por exemplo. E fazer inferências sobre os parâmetros do modelo através de econometria Bayesiana, utilizado um método de estimação baseado em amostragem.

Como uma das maiores dificuldades que estudos como esse se deparam refere-se a impossibilidade de mensurar a exposição midiática, a metodologia de contrafactuais entra como principal recurso para conseguir estimar o efeito da exposição, pois dentro de algumas circunstâncias, permite fazer inferências sobre o efeito causal de variáveis não mensuráveis. Uma das exigências dessa abordagem é que, dentro do conjunto de indivíduos analisados, exista um grupo que tenha sofrido a exposição e outro que não, além de que devem existir dados para esses grupos em dois períodos: antes e depois da exposição.

Com base nisso, considerou-se que o município de São Paulo oferece um bom cenário para análise. Além dos fatos já mencionados, indicando uma possível influências da mídia no resultados das eleições, a cidade é de extrema importância tanto para a economia quanto para a política nacional. Para se ter uma ideia, o município apresenta o maior colegiado eleitoral do país, com aproximadamente 8,9 milhões de eleitores, e é responsável por um orçamento anual superior a muitos estados. Em termos de comparação, o limite de despesas aprovada para a cidade em 2017 foi de R\$ 54,69 bilhões, enquanto que para o estado de Santa Catarina inteiro, por exemplo, o orçamento para o mesmo ano foi de R\$ 26,07 bilhões (SÃO PAULO, Lei No 16.608/2016 e SANTA CATARINA, Lei No 17.063/2016).

Acima de tudo, o PT, considerado aqui como o partido que teria sido afetado pela exposição da mídia, lançou Fernando Haddad como candidato à prefeitura de São Paulo tanto na eleição de 2012, antes da exposição, quanto na eleição de 2016, após a exposição. Dessa forma, o município de São Paulo proporciona as condições necessárias para a aplicação do método de contrafactuais para a estimação dos efeitos da exposição.

Admite-se, portanto, como hipótese básica a ser testada, que a exposição midiática sofrida pelo PT, relacionada, principalmente, a assuntos de cunho negativo como o impeachment da ex-presidente Dilma Rousseff e às investigações da Lava Jato, teve efeitos negativos para o mesmo partido nas eleições para a prefeitura do município de São Paulo em 2016.

De modo geral, este trabalho busca verificar se houve influência das exposição midiática negativa do Partido dos Trabalhadores no resultado das eleições da capital paulista em 2016.

Mais especificamente, por meio da análise dos resultados das eleições para a prefeitura no município de São Paulo nos anos de 2012 e 2016, pretende-se:

- Estimar a magnitude do efeito que a exposição negativa tem no resultado das eleições;
- Verificar se, na ausência da superexposição da mídia contra o PT nos casos relacionados ao impeachment e à Lava Jato, o candidato do PSDB, João Dória, teria sido eleito em primeiro turno.

Além dessa introdução, este trabalho divide-se em mais quatro capítulos. O Capítulo 2 traz uma revisão dos principais estudos encontrados na literatura nacional e internacional abordando os efeitos da exposição midiática nas eleições. O Capítulo 3 aborda, de forma detalhada: os modelos de utilidade aleatória que representam a base do problema de escolha; a metodologia de contrafactuais, em especial o modelos de *diferença-em-diferenças*; e o método Bayesiano de inferência estatística e suas especificações para um modelo de regressão linear com dados em painel. No Capítulo 4 é feita uma análise descritiva dos dados referente às eleições para a prefeitura de São Paulo. A parametrização e a aplicação do modelo são apresentadas no Capítulo 5, juntamente com a análise dos principais resultados obtidos com a estimação. Por fim, no Capítulo 6 são feitas algumas considerações finais.



## 2 REFERENCIAL TEÓRICO

A hipótese de que a exposição midiática de candidatos a cargos políticos tem efeito nos resultados das eleições não é recente. Na década de 1940, quando a mídia era majoritariamente composta por jornais impressos e rádio, esse assunto já era objeto de pesquisa. Este capítulo traz algumas das principais referências literárias que analisaram os efeitos que a exposição dos candidatos pelos meios de comunicação têm na tomada de decisão dos eleitores.

Em *The People's Choice*, Lazarsfeld et al. (1948) buscavam entender como os eleitores decidiam o voto, e o papel que a mídia tinha nesse processo, utilizando um painel de dados coletados de uma série de entrevistas repetidas com os mesmos eleitores entre maio e novembro de 1940. Os autores partiam da hipótese de que o ato de votar era individual, influenciado tanto pela personalidade do eleitor quanto pela exposição aos meios de comunicação. Entretanto, os primeiros resultados da pesquisa já indicavam que votar é essencialmente uma experiência em grupo, isto é, existe uma homogeneidade política em certos grupos — sociais, religiosos, família, etc. — e os membros de um mesmo grupo tendem a votar para o mesmo candidato.

Continuando com a análise, Lazarsfeld e seus associados identificaram um padrão na forma como as informações são transmitidas dentro dos grupos, o que chamaram de “fluxo de comunicação em dois passos”: primeiro, algumas pessoas do grupo, geralmente aquelas mais engajadas em questões políticas,

absorvem as informações da mídia e formam sua opinião; depois esses “líderes de opinião” transmitem as informações para os demais integrantes do grupo. Assim, as ideias geralmente fluem da mídia para os líderes de opinião, e daí para a parte menos ativa da população.

Apesar das evidências encontradas sobre a participação da mídia na decisão dos votos, os autores concluem que as relações pessoais são potencialmente mais influentes quando comparadas com os meios de comunicação formais, principalmente, porque existe certa influência psicológica dos líderes sobre os grupos, o que, para a mídia, é mais difícil de ser alcançado. Essa conclusão é examinada e corroborada em Schmitt-Beck (2003) a partir de estudos de eleições nacionais na Grã-Bretanha, Espanha, Estados Unidos e Alemanha.

Para Schmitt-Beck (2003), quando os eleitores recebem as informações da mídia, eles frequentemente conversam com outras pessoas de seu meio social. Se os eleitores pertencerem a um grupo em que as preferências políticas concordam com a mensagem da mídia, eles apoiarão essas mensagens e as levarão em consideração na hora da decisão do voto. Por outro lado, se as informações passadas pela mídia não estiverem de acordo com a opinião a priori daquele grupo social, elas serão rejeitadas e, portanto, não serão levadas em consideração na hora do voto.

Por outro lado, Bartels (1993, p.267) faz uma grande crítica a falta de empenho por parte da academia em tentar obter resultados que suportem a tese sobre o impacto da mídia nas eleições. Para o autor, “a onipresença da massa midiática e o monopólio virtual sobre a apresentação de diversos tipos de informações devem sugerir a um observador mediano que o que

---

essa mídia fala e a forma como é dita tem enorme consequência política e social”.

Partindo da hipótese de que os eleitores combinam informações preexistentes com informações novas, adquiridas principalmente por meio de noticiários, para decidirem em quem irão votar, Bartels (1993) propõe um modelo de formação de opinião que busca estimar em qual medida as informações novas contribuem nas ações individuais. Utilizando um painel de dados resultante de três ondas de pesquisas de opinião, feitas durante a corrida eleitoral para a presidência dos EUA em 1980, o autor consegue separar a informação nova do conhecimento preexistente e utiliza esse último como variável explicativa do modelo.

Entretanto, o resultado do estudo indica que a magnitude do efeito da exposição midiática na mudança de opinião dos eleitores é modesta, pois as opiniões formadas antes do início da campanha eleitoral (primeira onda de pesquisa) já se mostravam definidas após o início da campanha (segunda e terceira pesquisa). Assim, apenas uma exposição forte e unidirecional pode ser capaz de produzir mudanças observáveis nas intenções de votos, e, quanto mais fraca essa opinião, maior será a mudança. Ou seja, indivíduos com opinião formada sobre questões políticas, são menos suscetíveis a mudar de opinião após a exposição.

Contudo, Bartels (1993) acredita que a opinião a priori também recebe influência da mídia, mas pela periodicidade dos dados disponíveis para a análise — oito meses de diferença entre a primeira e a terceira onda de pesquisa, e todas no mesmo ano da eleição — não foi possível fazer inferências sobre como a mídia contribui para a formação da opinião antes das campanhas eleitorais.

Segundo Zaller (2002, p.298), o que está impedindo a estimação de resultados satisfatórios para o impacto da mídia nas eleições é a própria elaboração das pesquisas de opinião, utilizadas como base de dados para os estudos. Para o autor, essas pesquisas não oferecem os melhores dados para fazer tal análise, principalmente, porque não foram feitas para esse objetivo, “é como se astrônomos construíssem telescópios e começassem a procurar por novos objetos celestiais, sem antes calcular o que pode ser visto através de telescópios com tal resolução”.

Zaller (2002) propõe calcular o efeito da exposição na mídia a partir de simulações, considerando diferentes tamanhos de amostra e formas de abrangência da exposição. O autor utiliza as declarações dos candidatos, o desempenho nos debates transmitidos em rede nacional e as principais notícias sobre os candidatos reportadas pelos meios de comunicação, para simular a variável referente à exposição, e sugere três formas diferentes para a abrangência da exposição, são elas: “linear”, quando a exposição é proporcional ao nível de educação dos eleitores, assim, quanto maior o nível de escolaridade, maior o acesso aos meios de comunicação; “não-monotônica”, quando a informação chega de forma diferente para cada grupo da sociedade; e “elite”, quando a informação não chega as classes mais baixas.

Além das diferentes formas de exposição, Zaller (2002) também utiliza diferentes tamanhos de amostras, variando de 500 a 5000 observações. O principal resultado encontrado é que os efeitos estimados para a exposição serão confiáveis apenas sob duas condições: se o efeito for grande e capturado a partir de uma pesquisa grande, de preferência acima de 3000 observações. Dentro dessas condições, o modelo que apresentou maior poder

estatístico entre os três propostos foi aquele que analisava a exposição do tipo “elite”, chegando a causar uma variação de 5 pontos percentuais nas intenções de votos. Isto é, o estudo mostrou que informação destinadas às classes mais altas da sociedade têm mais probabilidade de aumentar ou diminuir a intenção de voto para o candidato exposto.

Fridkin e Kenney (2011) estudaram os efeitos da exposição negativa, entendida como fatos negativos associados ao nome dos candidatos, nas impressões que os eleitores têm dos candidatos alvo, a partir de dados de pesquisa sobre as eleições para o senado norte americano em 2006. Os resultados do estudo sugerem que o impacto de informações negativas é multifacetado e, sob certas circunstâncias, é grande. O impacto varia em termos da relevância da informação para cada indivíduo, de forma que indivíduos menos tolerantes a certos tipos de mensagens negativas são mais sensíveis e, portanto, serão mais influenciados e impactados por tais mensagens.

Mas Lau (1982) mostra que todos os indivíduos são mais sensíveis a notícias negativas, quando comparadas com notícias positivas de mesma intensidade, ou seja, ter fatos positivos ligados ao nome de um candidato não é tão importante quanto ter fatos negativos associados ao mesmo candidato. Para o autor, está claro que quando os eleitores estão formando suas opiniões sobre os candidatos políticos, as informações negativas que eles têm são mais importantes que as informações positivas. O que sugere que, durante as campanhas eleitorais, apresentar os resultados positivos alcançados por um candidato ou partido não chama tanta a atenção dos eleitores quanto “denunciar” os seus maus resultados.

Ainda segundo Lau (1982), os eleitores votam, geralmente, contra um candidato ao invés de votar a favor do candidato escolhido, citando, como exemplo, que quando há um aumento na taxa de desemprego, a popularidade do presidente reduz, mas quando há uma redução na taxa de desemprego não há efeito positivo concomitante na popularidade. Essa teoria poderia levar a conclusão de que seria vantajoso para os candidatos utilizarem a maior parte do tempo de campanha mostrando as fraquezas de seus adversários em vez de suas próprias habilidades ou conquistas.

Entretanto, o autor ressalta que ser conhecido como um difamador é ter um fato negativo associado ao próprio nome, o que resultaria na perda de votos. Assim, o truque é ser afortunado o suficiente para ter muitas informações negativas conhecidas sobre o oponente sem ter que trazer esses fatos negativos à luz. Dessa maneira, quando a mídia se encarrega de expor negativamente um candidato, pode estar favorecendo duplamente seu oponente, já que este não precisaria se arriscar para expor o outro candidato e ainda receberia os votos destinado candidato exposto.

No Brasil, Telles, Lourenço e Storni (2009) propõem avaliar os determinantes do voto para a prefeitura da capital mineira nas eleições de 2008 por meio de dados coletados em duas pesquisas de opinião realizadas no primeiro e no segundo turno. Considerando a importância da mídia durante as campanhas eleitorais, os autores incluem no modelo uma variável que contém o tempo destinado a cada candidato no horário gratuito de propaganda eleitoral, para estudar o efeito causado pela exposição dos candidatos na televisão.

Os resultados mostram que o efeito da percepção que

os eleitores têm das propagandas televisionadas é maior quando os candidatos adotam uma estratégia mais ativa, no sentido de combater o oponente. Entretanto, Telles, Lourenço e Storni (2009) limitaram-se a verificar se a exposição teve ou não influência na preferência dos eleitores, não chegando a estimar a magnitude do efeito da exposição sobre os votos, o que só poderia ser feito analisando as preferências dos eleitores em períodos distintos: antes e depois da exposição.

Estes trabalhos mostraram que, apesar de o efeito da exposição midiática nas eleições vir sendo estudado há muito tempo, esse campo de análise parece estar longe de ser exaurido. Como visto, há uma grande dificuldade em obter dados para análise empírica, já que pesquisas de opinião costumam ser custosas e utilizar dados de pesquisas feitas para outros fins dificilmente apresentará resultados significativos. Com isso, muitos dos trabalhos limitam-se a uma análise teórica da exposição, sem chegar a mensurar a magnitude do impacto.

Assim, o presente trabalho se propõe a estimar o impacto da exposição midiática a partir de uma metodologia de contrafactuais, muito abordada em avaliações de impacto de políticas públicas, mas que, aparentemente, ainda não foi explorada no contexto de eleições. E pretende-se estimar esse impacto a partir dos próprios resultados das eleições. O capítulo a seguir descreve, em detalhes, o método de estimação a ser utilizado neste trabalho.



## 3 METODOLOGIA

Ao escolher em qual candidato votar, ou não votar, o eleitor estará maximizando sua satisfação em relação a todas as opções disponíveis. Essa afirmação parte da teoria do problema de escolha, que supõe que cada escolha fornece uma utilidade, bem estar, para quem estiver escolhendo, e a opção selecionada será aquela que maximiza a utilidade para o indivíduo. Desta forma, o desenvolvimento de um modelo para entender como os eleitores decidem o voto deve partir da metodologia do problema de escolha, mais especificamente, da escolha discreta.

Para avaliar a influência que a exposição midiática tem na decisão do eleitor, pode-se incluir no modelo de escolha discreta uma metodologia geralmente utilizada na análise de impacto de políticas públicas, que possibilita isolar o efeito causado por um determinado tratamento, que neste caso é representado pela exposição midiática, do efeito de todas as outras variáveis que também podem influenciar nos resultados das eleições. Isso é feito através da abordagem de contrafactuais — situação hipotética sobre o que teria acontecido no resultados da eleições caso a exposição não tivesse ocorrido — que indica que diferenças entre a situação observada, factual, e a situação não observada, contrafactual, possibilita mensurar o impacto da exposição.

Assim, este capítulo destina-se a apresentação das etapas metodológicas necessárias para desenvolver o modelo a fim de estimar o impacto que a exposição da mídia tem no resultado eleitoral. As seções 3.1 e 3.2 abordam a metodologia do problema

de escolha discreta e o método de contrafactuais, respectivamente. Além disso, a última seção do capítulo, 3.3, expõe o método Bayesiano de regressão com dados em painel utilizado para estimar os parâmetros do modelo.

### 3.1 O problema da escolha

A maioria das escolhas com as quais uma pessoa se depara ao longo da vida tem caráter discreto, como, por exemplo, qual profissão seguir, se casará ou não, em qual bairro morar, etc.. A decisão sobre qual candidato deverá votar em uma eleição, feita por um eleitor, também representa um problema de escolha discreta. Modelos de escolha discreta são derivados, geralmente, sob a hipótese de que o indivíduo tem comportamento maximizador de utilidade, isto é, os indivíduos buscam maximizar seu bem-estar em cada escolha.

Para se enquadrar em um modelo de escolha discreta, o conjunto de opções com o qual um indivíduo se defronta, conhecido como conjunto de escolha, deve exibir três características. Primeira, as alternativas devem ser mutuamente excludentes, isto é, escolher uma opção implica necessariamente em não escolher nenhuma outra opção. Segunda, o conjunto de escolha deve ser exaustivo, todas as alternativas possíveis deverão estar incluídas nele. Terceira, o número de alternativas deve ser finito (TRAIN, 2002). Assim, o conjunto de escolhas com o qual um eleitor se depara pode ser definido como um conjunto  $C$  contendo  $j$  opções de voto, em que cada opção corresponde a um dos candidatos aptos para receberem votos, e, além dos candidatos, o conjunto também apresenta a opção de não votar (que inclui votar em branco, votar nulo ou se abster), de forma que  $C = \{j \in \mathbb{N}; 0 \leq j \leq J\}$ .

Seguindo a teoria do problema de escolha, a preferência dos indivíduos pode ser modelada por meio de uma função de utilidade, assim, os modelos de escolha racional se apoiam na hipótese de que as decisões observadas baseiam-se na comparação das utilidades resultantes de cada alternativa (PUDNEY, 1989). Partindo disso, pressupõe-se que um eleitor  $i$ , que se depara com  $J$  opções de voto, vai alcançar um nível diferente de utilidade para cada  $j$  alternativa. Definindo a utilidade que o eleitor  $i$  obtém de uma opção qualquer de voto  $j$  como  $U_{ij}$ , o modelo comportamental para o voto do eleitor pode ser estabelecido mediante a comparação da utilidade proporcionada por cada opção de voto, assim, a alternativa  $m$  será escolhida pelo eleitor se, e somente se,  $U_{im} > U_{ij}$ ,  $\forall m \in C$  e  $m \neq j$ .

Um pressuposto importante em modelos de escolha é que a utilidade atribuída a cada opção de escolha depende de um número de características mensuráveis, ou atributos, da opção em si (CASCETTA, 2009). Assim, a utilidade que um eleitor atribui a cada opção de voto depende, em teoria, das características dos candidatos, de forma que  $U_{ij} = U(x_j; \beta_i)$ , onde  $x_j$  é o vetor de atributos relativos ao candidato  $j$  e  $\beta$  é um vetor de parâmetros desconhecidos que correspondem as mudanças de preferência do eleitor  $i$ . Entretanto, a utilidade atribuída pelo eleitor  $i$  à opção de voto  $j$  é de conhecimento apenas para o próprio eleitor  $i$ . Isto é, existem características não mensuráveis que influenciam no voto do eleitor, mas, para um observador externo, não é possível contabilizá-las.

Assim, a função de utilidade passa a incluir um elemento aleatório,  $\varepsilon$ , para representar os fatores que afetam a decisão do eleitor  $i$ , que podem ser observados pelo mesmo, mas não são

mensuráveis por qualquer outra pessoa. Dessa forma, tem-se que  $U_{ij} = V_{ij} + \varepsilon_{ij}$ , onde  $V_{ij} = V(x_j; \beta_i)$ , é conhecida como utilidade representativa e inclui as características mensuráveis, tanto para o eleitor quanto para um observador externo, dos candidatos, como gênero, escolaridade, estado civil, naturalidade, etc., e  $\varepsilon$  corresponde ao desvio da utilidade, percebido pelo eleitor, em relação a utilidade representativa.

Modelos que seguem essa estrutura são conhecidos como modelos de utilidade aleatória, ou RUM (*random utility models*), e formam um subconjunto da classe de modelos de escolha probabilística. Foram desenvolvidos primeiramente na área de psicologia, na tentativa de caracterizar as inconsistências observadas em padrões de comportamento individual. Mais tarde, economistas, começando com McFadden (1974), abraçaram tais modelos como uma representação econométrica de maximização do comportamento. Nessa formulação, segundo Manski (1977), a utilidade é tratada como aleatória, não para refletir a falta de racionalidade de quem está decidindo, mas para refletir a falta de informações sobre as características das alternativas e/ou do tomador de decisão por parte de um observador externo.

A presença de elementos aleatórios na função de utilidade torna impossível prever com exatidão a opção que um eleitor específico escolherá. Por outro lado, a partir de especificações sobre a densidade conjunta do vetor aleatório  $\varepsilon_i = (\varepsilon_{i1}, \dots, \varepsilon_{iJ})$ , denotada por  $f(\varepsilon_i)$ , é possível expressar a probabilidade com a qual um eleitor  $i$  escolherá a alternativa  $j$  condicional a seu conjunto de escolha  $C$ , e essa será a probabilidade com que a utilidade da opção de voto  $j$  é maior do que todas as outras

opções de voto disponíveis:

$$P_i(j|C) = \text{Prob}(U_{ij} > U_{im}, \quad \forall m \neq j, j \in C). \quad (3.1)$$

Substituindo a expressão da utilidade apresentada anteriormente, tem:

$$\begin{aligned} P_i(j|C) &= \text{Prob}(V_{ij} + \varepsilon_{ij} > V_{im} + \varepsilon_{im}, \quad \forall m \neq j, j \in C) \quad (3.2) \\ &= \text{Prob}(\varepsilon_{im} - \varepsilon_{ij} < V_{ij} - V_{im}, \quad \forall m \neq j, j \in C). \end{aligned}$$

Essa probabilidade é uma distribuição cumulativa, a probabilidade de que qualquer termo aleatório  $\varepsilon_{im} - \varepsilon_{ij}$  é menor do que as quantidades observadas  $V_j - V_m$ . Usando a densidade  $f(\varepsilon_i)$ , essa probabilidade cumulativa pode ser reescrita como:

$$\begin{aligned} P_i(j|C) &= \text{Prob}(\varepsilon_{im} - \varepsilon_{ij} < V_{ij} - V_{im}, \quad \forall m \neq j, j \in C) \quad (3.3) \\ &= \int_{\varepsilon} I(\varepsilon_{im} - \varepsilon_{ij} < V_{ij} - V_{im}) f(\varepsilon_i) d\varepsilon_i, \end{aligned}$$

onde  $I(\cdot)$  é uma função indicadora, igual a 1 quando o termo em parênteses é verdadeiro e 0 em caso contrário. Essa é uma integral multidimensional da densidade da porção não observada da utilidade,  $f(\varepsilon_i)$ . Diferentes especificações sobre esta densidade resultam em diferentes modelos de escolha discreta.

McFadden (1974) mostra que se a parte não observada da utilidade apresentar distribuição de Gumbel (valor extremo independente e identicamente distribuído), essa integral terá forma fechada, caracterizando a classe de modelos logit. Isto é, se a densidade de cada componente não observado for

$$f(\varepsilon_{ij}) = e^{-\varepsilon_{ij}} e^{-e^{\varepsilon_{ij}}}, \quad (3.4)$$

e a função de distribuição acumulada for

$$F(\varepsilon_{ij}) = e^{-e^{\varepsilon_{ij}}}, \quad (3.5)$$

então  $\varepsilon_{im} - \varepsilon_{ij}$  segue uma distribuição logística, dado que a diferença entre duas variáveis de valor extremo é distribuída logisticamente, logo

$$P(\varepsilon_{im} - \varepsilon_{ij}) = \frac{\exp(\varepsilon_{im} - \varepsilon_{ij})}{1 + \exp(\varepsilon_{im} - \varepsilon_{ij})}. \quad (3.6)$$

Assim, a probabilidade de que o eleitor  $i$  escolha votar no candidato  $j$  entre todas as opções de voto disponíveis,  $(1, 2, \dots, J) \in C$ , pode ser expressa na forma fechada.

$$P_{ij} = \frac{\exp(V_{ij})}{\sum_{m=1}^J \exp(V_{im})}. \quad (3.7)$$

A utilidade representativa geralmente é especificada para ser linear em seus parâmetros,  $V_{ij} = \beta'_i x_j + \xi_j$ , onde  $x_j$  é um vetor de variáveis observadas relacionadas as alternativas  $j$  e  $\xi_j$  corresponde a outras variáveis que influenciam na decisão do eleitor, mas não são observadas por um pesquisador e, por consequência, não foram incluídas em  $x$ . Com essa especificação, a probabilidade logit se torna

$$P_{ij} = \frac{\exp(\beta'_i x_j + \xi_j)}{\sum_{m=1}^J \exp(\beta'_i x_m + \xi_j)}. \quad (3.8)$$

Considerando o modelo para a decisão do voto, é importante destacar que, como o voto é secreto, não é possível obter a informação sobre qual foi a decisão de cada eleitor. Mas é possível apurar quantos votos cada opção recebeu por zona eleitoral. Dessa forma, a decisão do voto deve ser analisada a

partir de um conjunto de eleitores. Considerando um conjunto formado por todos os  $n$  eleitores, tem-se que a proporção dos que votaram na opção  $j$  será uma média da probabilidade de votar na opção  $j$  de todos os eleitores  $i$ .

$$s_j = \int_{i=1} \int_{i=2} \cdots \int_{i=n} \frac{\exp(\beta_i x_j + \xi_j)}{1 + \exp(\beta_i x_j + \xi_j)} dF(\beta_{i=n}) \cdots \quad (3.9)$$

$$\cdots dF(\beta_{i=2}) dF(\beta_{i=1})$$

Considerando, agora, um caso especial no qual todos os  $n$  eleitores votam de forma semelhante, então cada covariada apresentará o mesmo parâmetro para todas as zonas eleitorais (i.e.  $\beta_i = \beta, \forall i$ ). Com isso a Eq.(3.9) será simplificada para:

$$s_j = \frac{\exp(\beta x_j + \xi_j)}{1 + \exp(\beta x_j + \xi_j)} \quad (3.10)$$

A expressão (3.10) define um modelo logit multinomial, o modelo de utilidade aleatória mais simples e amplamente utilizado. Uma vez que o lado direito de (3.10) não é indexado por  $i$ , a probabilidade não condicional de voto é igual a probabilidade condicional,  $P_{ij} = P_j$ . Neste caso temos, então, que  $s_{ij} = s_j$ .

Além disso, o conjunto de escolha contém não só os candidatos cujo nome aparecem nas urnas, mas também a opção de não-voto, por convenção, denotada por opção 0, que contabiliza os votos brancos, votos nulos e as abstenções. Uma observação importante sobre o não-voto é que não existem características mensuráveis para explicar sua escolha,  $V_{i0} = 0$ , esta opção é conhecida nos modelo de utilidade aleatória como opção externa,

e resulta em utilidade nula para o indivíduo que escolhe-la<sup>1</sup>,  $U_{i0} = 0$ . Destarte, o percentual de votos recebidos pela opção externa é assim denotado:

$$s_0 = \frac{1}{\sum_{m=1}^J \exp(\beta x_m + \xi_j)}. \quad (3.11)$$

Domencich e McFadden (1975) mostram que em situações como esta, em que uma opção de escolha apresenta todas as observações do vetor de variável explanatória repetidas, a forma mais comum de tratamento é utilizar  $\ln(s_j) - \ln(s_0)$ ,  $\forall j \in C$ , como variável dependente — o que decorre do tratamento algébrico dado na razão entre as Equações (3.11) e (3.12). Aplicando o logaritmo nas funções  $s_{ij}$  e  $s_{i0}$ , chega-se a seguinte estrutura log-linear:

$$\ln(s_j) - \ln(s_0) = \beta' x_j + \varepsilon_j, \quad (3.12)$$

onde  $\varepsilon_{ij}$  é um termo de erro normalmente distribuído.

Uma vez que tem-se  $s_{ij} = s_j$  pode-se utilizar qualquer nível de agregação dos votos para estimar (3.12), quanto mais próximo do resultado por indivíduo, maiores serão os graus de liberdade para o modelo, pois aumenta o tamanho da amostra. Como o nível mais desagregado disponível para os resultados das eleições é a zona eleitoral, a Eq.(3.12) pode ser reescrita como:

$$\ln(s_{zj}) - \ln(s_{z0}) = \beta' x_j + \varepsilon_j, \quad (3.13)$$

onde o subscrito  $z$  identifica a zona eleitoral.

<sup>1</sup> Isso pode ser feito sem perda de generalidade em decorrência das propriedades de ordinalidade da função de utilidade.

De (3.13) é possível derivar as seguintes elasticidades:

$$\frac{\partial s_j}{\partial x_j} \frac{x_j}{s_j} = \beta x_j (1 - s_j) \quad [\text{Elasticidade própria}] \quad (3.14a)$$

$$\frac{\partial s_{j \neq m}}{\partial x_m} \frac{x_m}{s_j} = -\beta x_m s_m \quad [\text{Elasticidade cruzada}] \quad (3.14b)$$

Assim, tendo em mãos a estimativa para o parâmetro referente a variável de exposição midiática e a participação dos partidos no resultados das eleições, basta substituí-las nas equações em (3.14) para obter a variação causada pela exposição no percentual de votos para o PT (elasticidade própria) e a variação causada no percentual de todas as outras opções (elasticidade cruzada).

Entretanto existe um problema conceitual nessa abordagem. Ao utilizar a equação de elasticidade cruzada (3.14b), a variação no percentual de votos recebidos por um candidato  $j$  decorrente da variação da característica de outro candidato,  $x_m$ , dependerá apenas de  $m$ . Com isso a exposição midiática sobre o PT afetaria de forma idêntica a todas as outras opções de voto. Essa propriedade é comum aos modelos LOGIT e recebe o nome de *competição não localizada*, mas é pouco plausível em situações em que as opções de escolha têm características diferentes (PETTERINI; DESOUSA, 2009).

Além disso, as zonas eleitorais, que representam aqui os indivíduos, correspondem a parcelas do município que, assim como os bairros, não são homogêneas entre si. Havendo características próprias a cada zona — renda, escolaridade, etc. — influenciando na percepção que os eleitores têm em relação aos atributos dos candidatos a prefeitura. Ignorar essa heterogeneidade pode resultar em um modelo com coeficientes tendenciosos

e inconsistentes.

O emprego de coeficientes aleatório em modelos como este costuma mitigar o problema de heterogeneidade entre os indivíduos, já que permite que os parâmetros variem de indivíduo para indivíduo, captando as diferentes influências de cada zona eleitoral. Mas para estimar um modelo com coeficientes aleatórios será necessário resolver as integrais em (3.9), o que exige métodos econométricos mais avançados.

Petterini e DeSouza (2009) mostram que, se os parâmetros aleatórios forem estimados por método de simulação das escolhas, é possível utilizar os  $\beta$ 's e seus respectivos desvios padrões estimados para calcular as elasticidades de uma maneira mais precisa, por meio das seguintes derivadas:

$$\frac{\partial s_j}{\partial x_j} = \int_{S(x)} \left( \left[ \frac{\exp(V_{ij})}{1 + \sum_{r=1}^J \exp(V_{ir})} \right] \times \left[ 1 - \frac{\exp(V_{ij})}{1 + \sum_{r=1}^J \exp(V_{ir})} \right] \times \beta(x) \right) dF(x) \quad (3.15a)$$

e

$$\frac{\partial s_j}{\partial x_m} = \int_{S(x)} \left( \left[ \frac{\exp(V_{ij})}{1 + \sum_{r=1}^J \exp(V_{ir})} \right] \times \left[ \frac{\exp(V_{im})}{1 + \sum_{r=1}^J \exp(V_{ir})} \right] \times \beta(x) \right) dF(x). \quad (3.15b)$$

As derivadas em (3.15) representam os valores estimados para uma matriz de elasticidades, onde (3.15a) são os valores da diagonal principal e (3.15b) são os valores fora da diagonal. Dessa forma, os valores cruzados de elasticidade não serão idênticos (i.e.  $\frac{\partial s_j}{\partial x_m} \neq \frac{\partial s_m}{\partial x_j}$ ) como ocorre nos modelos LOGIT simples. Portanto,

espera-se que os valores encontrados a partir dessas derivadas estejam mais próximos dos verdadeiros valores.

Outra particularidade do modelo está diretamente relacionada ao objetivo da pesquisa: estimar o efeito de exposição midiática no resultado das eleições. Para ser possível estimar seu efeito, a variável que representa a exposição midiática deve pertencer ao vetor de variáveis explanatória  $x$ , que, como já explicado, apresenta valores observados tanto pelo eleitor quanto por um observador externo. Entretanto, apesar de haver estudos que defendem a existência da influência da mídia nas eleições, não existem meios disponíveis para mensurar a exposição midiática de um candidato ou partido, como quanto tempo ou quantas vezes um determinado candidato foi exposto nos noticiários.

Em geral, o termo de erro  $\varepsilon$  é o responsável por incluir no modelo todas as variáveis não mensuráveis, de forma que o efeito da exposição midiática, ou de qualquer outra variável que possa influenciar o voto, mas não foi incluída em  $x$ , seria uma parcela do termo de erro, portanto, impossível de ser determinada separadamente. Felizmente, já foram desenvolvidas metodologias que permitem, dentro de algumas circunstâncias, distinguir do termo de erro o efeito de variáveis como a exposição midiática, que a teoria afirma que tem efeito, mas que não existem dados mensuráveis para estimá-lo.

A metodologia de inferência causal desenvolvida dentro do arcabouço teórico de contrafactuais possibilita avaliar o impacto de uma variável não mensurável. Desse modo, antes de apresentar as especificações do modelo com coeficientes aleatórios, será abordada na Seção 3.2, a seguir, a metodologia de contrafactuais.

## 3.2 Método de contrafactuais

A impossibilidade de mensurar uma variável representa um grande obstáculo quando se pretende fazer inferências sobre a mesma. Talvez esse seja o motivo pelo qual não existem muitos estudos, especialmente na última década, investigando a influência da exposição midiática na eleições. Nos modelos apresentados nos trabalhos de Schmitt-Beck (2003), Bartels (1993) e Zaller (2002), o impacto da exposição foi estimado a partir de dados simulados. Consequentemente, se a metodologia para a simulação dos dados não for apropriada, a estimativa poderá apresentar viés.

Existe outra forma de estimar o impacto de uma variável não observada, sem a necessidade de simular dados para mensurá-la. A técnica consiste na análise de resultados potenciais, que compara os resultados efetivamente observados de uma eleição, sob a hipótese de que foram afetados pela exposição midiática, com os resultados que seriam observados se a exposição não tivesse ocorrido. Tal metodologia tornou-se muito comum na estimação do efeito causal de determinados tratamentos, como o caso do impacto de políticas públicas.

Neyman (1923) foi o primeiro a propor a ideia de resultados potenciais em um contexto de experimentos completamente aleatórios, por meio de um experimento agrícola hipotético. Em tal experimento, o autor considera que existe um terreno para plantio e divide-o em  $n$  parcelas iguais. Em cada trecho  $n_i$  de terra deverá ser cultivada uma variedade  $v_j$  de planta, de forma que a produção de cada parcela de terra é representada por  $U_{ij}$ , onde  $i$  indexa o trecho e  $j$  indexa a variedade. Assim, uma vez que só é possível plantar uma única variedade em cada trecho de terra, só será possível observar os resultados da variedade

plantada, isto é, se a variedade  $v_1$  for plantada no trecho  $n_1$ , a única produção que poderá ser observada nessa parcela de terra será  $U_{11}$ .

Para conseguir comparar os resultados da produção de diferentes variedades, o autor considerou que todos os trechos de terra teriam características semelhantes, de forma que as diferenças na produção viriam das variedades de plantio. Assim, a partir de uma amostragem sem repetição, primeiro, determinou de forma aleatória qual seria a variedade plantada em cada trecho de terra, depois, por meio da comparação dos resultados de duas variedades, por exemplo  $U_{11}$  e  $U_{22}$ , chegou a uma fórmula para a variabilidades da diferença entre as médias dos resultados observados do plantio das duas variedades. Isso possibilitou estimar o efeito que diferentes tipos de culturas têm no resultado da produção, pois, mesmo sendo possível observar apenas o resultado de uma variedade em cada trecho, o resultado  $U_{22}$  serve como base para um contrafactual de  $U_{11}$ , que seria  $U_{12}$ , não mensurável.

Considerando que experimentos aleatórios dificilmente são possíveis, Rubin (1974) aplicou a perspectiva de resultados potenciais de Neyman (1923) para inferência causal em estudos utilizando dados observacionais, não aleatórios. Mas o autor deixa claro que sempre que houver a possibilidade de escolha entre um experimento aleatório e um estudo não aleatório equivalente, é preferível optar pelo experimento, principalmente nas ciências sociais onde a maior parte da variabilidade geralmente não está associada a uma causa particular. Este modelo ficou conhecido como “modelo de inferência causal de Neyman-Rubin” (SEKHON, 2008).

Seguindo o arcabouço de Neyman-Rubin, se um can-

didato, digamos  $j$ , for exposto negativamente ele receberá um percentual  $s_{j1}$  de votos, e caso este candidato  $j$  não sofra nenhum tipo de exposição, ele receberá  $s_{j0}$ , então tem-se que  $s_{j1}$  e  $s_{j0}$  seriam os resultados potenciais do candidato  $j$ . Com isso, o efeito da exposição sobre este candidato poderia ser dado por  $\tau_j = s_{j1} - s_{j0}$ . Mas, segundo Neyman (1923), existe um grande problema nesta abordagem, chamado por Holland (1986) de “Problema Fundamental da Inferência Causal”, que é a impossibilidade de observar  $s_{j1}$  e  $s_{j0}$  simultaneamente, já que são eventos mutuamente exclusivos.

Como não é possível obter o resultado de um indivíduo para caso em que ele não tivesse passado por um tratamento, dado que passou pelo tratamento, pode ser construído um cenário contrafactual por meio da comparação entre o resultado factual desse indivíduo, ou grupo, que passou pelo tratamento e os resultados obtidos por outros indivíduos, com características semelhantes a ele, mas que não passaram pelo tratamento. Como consequência, é necessário fazer algumas hipóteses estatísticas para que se possa construir esse cenário e, assim, dar continuidade nas estimativas do efeito da exposição sobre os resultados das eleições.

Segundo Gertler et al. (2015), se a escolha do grupo que passará pelo tratamento for feita de forma que não afete os resultados do outro grupo, o impacto médio do tratamento é estimado sem viés por meio da diferença entre as médias dos resultados dos grupos tratados e não tratados. Com isso, a melhor forma de fazer essa escolha seria através de um processo aleatório, mas, como já mencionado, a aleatoriedade é uma hipótese pouco factível em estudos como este, pois é pouco provável que um

determinado meio de comunicação escolha de forma aleatória um partido e exponha seus pontos negativos, deixando todos os outros partidos isentos.

Portanto, dado que a hipótese de aleatoriedade não é satisfeita, o efeito médio da exposição midiática nos resultados do partido exposto (*Average Treatment Effect on Treated* – ATT) será estimado com viés. Mas, segundo Gertler et al. (2015), existe uma série de alternativas, ainda dentro do arcabouço de Neyman-Rubin, que mimetizam um experimento aleatório e podem, portanto, mitigar esse viés de seleção. Cada uma dessas alternativas é passível de ser implementada dependendo de como o programa foi executado e do tipo de dados que se dispõe.

Uma alternativa é o modelo de *diferença-em-diferenças* (também representado por *diff-in-diff*, DiD ou DD), que demanda uma base de dados contendo uma quantidade relativamente grande de unidades *crosssection*, o que não exige muitas observações temporais — dois períodos são suficientes. Isto é, uma metodologia adequada para a base de dados eleitorais disponível para este trabalho.

O método *diff-in-diff* ficou conhecido após os trabalhos de Ashenfelter (1978) e Ashenfelter e Card (1984) analisarem o impacto de um programa de treinamento nos salários dos participantes. O modelo mais simples consiste na comparação dos resultados de dois grupos em dois períodos diferentes. Um deles é exposto ao tratamento no segundo período, mas não no primeiro, o outro grupo não é exposto ao tratamento em nenhum dos dois períodos. O primeiro grupo é chamado de grupo de tratamento, enquanto o segundo é conhecido como grupo de controle.

Primeiro calcula-se, separadamente, a variação média

nos resultados de cada grupo entre os dois períodos. Depois, subtrai-se a variação encontrada para o grupo de tratamento da variação encontrada para o grupo de controle. Para Imbens e Wooldridge (2009), essa diferença dupla, que dá nome ao modelo, remove o viés na comparação dos resultados do segundo período que poderia ser causado por diferenças permanentes entre os dois grupos. Isto é, possíveis diferenças entre os dois grupos, que já existiam antes do tratamento, podem levar a variações diferentes na trajetória dos resultados de cada grupo, logo, se utilizar apenas os resultados do segundo período para comparar a diferença entre os grupos, essa variação que não tem necessariamente relação com o tratamento, seria erroneamente atribuída ao mesmo. Mas quando o efeito é estimado por meio da dupla diferença, esse viés é mitigado e o resultado apresenta apenas a variação causada pelo tratamento.

Para estimar o impacto da exposição midiática, considerou-se o PT como o grupo de tratamento, sob a hipótese de que apenas este partido teria passado pela exposição, e todos os outros partidos entram no grupo de controle. Definidos os grupos, o método de *diff-in-diff* resume-se em três passos: primeiro, toma-se a diferença da esperança condicionada do indicador de impacto na unidade exposta (PT) entre 2016 e 2012, medindo a variação dos votos sob a exposição; segundo, repete-se o procedimento com o grupo de controle, medindo a variação dos votos não influenciados diretamente pela exposição; e terceiro, toma-se a diferença entre as diferenças encontradas nos dois primeiros passos, é este terceiro passo que permite isolar o efeito causal da exposição.

O *diff-in-diff* apresentado nos três passos acima pode

ser representado da seguinte forma:

$$\begin{aligned}
 DD = & [E(y_{ijt}|dP_t = 1, dT_j = 1, x_{ijt}) \\
 & - E(y_{ijt}|dP_t = 0, dT_j = 1, x_{ijt})] \\
 & - [E(y_{ijt}|dP_t = 1, dT_j = 0, x_{ijt}) \\
 & - E(y_{ijt}|dP_t = 0, dT_j = 0, x_{ijt})],
 \end{aligned} \tag{3.16}$$

onde  $y_{ijt}$  é o resultado de interesse e representa a participação do candidato  $j$  nos votos contabilizados na zona  $i$  no período  $t$ ,  $dP_t$  é uma variável *dummy*, onde  $dP_t = 1$  indica o período após a exposição,  $dT_j$  é uma *dummy*, onde  $dT_j = 1$  indica sofrer a exposição, e  $x$  representa um vetor de covariadas observadas.

Em modelos com dados em painel, a metodologia de *diff-in-diff* pode ser utilizada como

$$\begin{aligned}
 y_{ijt} = & \beta_{0i} + \beta_{1i}^{(1)} dP_t + \beta_{1i}^{(2)} dT_j + \beta_{1i}^{(3)} dP_t \cdot dT_j \\
 & + \sum_{k=2}^K \beta_{ki} x_{kjt} + \varepsilon_{ijt} \quad ,
 \end{aligned} \tag{3.17}$$

onde o subscrito  $i$  indica a existência de heterogeneidade entre as zonas eleitorais. A variável *dummy* para o período,  $dP$ , captura os fatores agregados que causariam mudanças nos resultados da eleição mesmo na ausência da exposição da mídia, e a variável *dummy* para o tratamento,  $dT$ , captura as possíveis diferenças entre o PT e os outros partidos (grupo de controle) que já existiam antes da exposição.

O coeficiente de interesse, que representa o efeito da exposição, é o  $\beta_{1i}^{(3)}$ , que multiplica o termo de interação,  $dP_t \cdot dT_j$ , que seria o mesmo que uma variável *dummy* igual a 1 para as observações do grupo que passou pela exposição no segundo período. Assim, a média estimada para  $\beta_{1i}^{(3)}$  terá o efeito de semi-

elasticidade definido em (3.14), indicando a variação média na proporção de votos causada pela exposição negativa.

Os parâmetros de modelos com coeficientes aleatórios podem ser estimados tanto pela metodologia desenvolvida por Berry, Levinsohn e Pakes (1995), conhecida como BLP, quanto pela abordagem Bayesiana de inferência estatística. Em relação ao primeiro método, Dubé, Fox e Su (2012) e Knittel e Metaxoglou (2014) discutem vários desafios metodológicos que, no estado atual da literatura, só são resolvidos quando as distribuições de interesse são aproximadamente normais e o vetor de parâmetros for de pequena dimensão.

Em contrapartida, como será apresentado a seguir, a abordagem Bayesiana utiliza métodos de Monte Carlo via Cadeias de Markov para estimar as densidades marginais a posteriori dos parâmetros, o que não exige normalidade nas distribuições, nem limita a dimensão do vetor de parâmetros, apesar de demandar alta capacidade computacional devido ao grande número de iterações necessárias para a convergência do algoritmo, enquanto o modelo BLP exige um custo computacional menor. Assim, como não há certeza quanto forma completa da distribuição dos parâmetros de interesse, optou-se pela utilização da metodologia Bayesiana.

### 3.3 Abordagem Bayesiana

A metodologia estatística Bayesiana é fundamentalmente diferente do método clássico de inferência estatística. O método clássico pressupõe que os parâmetros  $\beta$  são quantidades desconhecidas, mas fixas. Em tal metodologia, uma amostra aleatória  $X_1, \dots, X_n$  é recolhida de uma população indexada por  $\beta$  e,

baseando-se nos valores observados na amostra, são feitas inferências sobre  $\beta$ . Dado que os parâmetros são fixos, não é possível calcular a probabilidade de um parâmetro ser maior ou menor do que um determinado valor.

Por outro lado, na metodologia Bayesiana,  $\beta$  é considerado como uma quantidade cuja variação pode ser descrita por uma distribuição de probabilidade, assim os parâmetros a serem estimados passam a ser tratados como variáveis aleatórias, enquanto a base de dados é fixa (CASELLA; BERGER, 2002; GREENBERG, 2008).

O método Bayesiano de inferência estatística será explicado mais detalhadamente na seção 3.3.1. A seção 3.3.2 apresenta uma especificação da metodologia Bayesiana para um modelo geral de regressão linear com dados em painel, seguindo, principalmente, os trabalhos de Koop (2003) e Greenberg (2008).

### 3.3.1 Econometria Bayesiana

A estatística Bayesiana foi desenvolvida pelo Rev. Thomas Bayes (1702-1761) e depois expandida por Pierre-Simon Laplace (1749–1827), que fez importantes contribuições para a teoria da chamada “probabilidade inversa”. Arnold Zellner foi um dos primeiros a aplicar a estatística Bayesiana à econometria ainda no início dos anos 1960, e apresentou seu trabalho pioneiro em Zellner (1971). Com o rápido avanço da capacidade computacional ao longo da década de 1980, as aplicações de simulações cadeia de Markov Monte Carlo nos modelos estatísticos e econométricos permitiram que a análise Bayesiana aumentasse sua influência na econometria (GREENBERG, 2008).

Zellner (1971), Koop (2003) e Greenberg (2008) fornecem

uma boa introdução à econometria, sendo que os dois últimos tem seção especial dedicada a modelos com dados em painel. Independentemente da estrutura apresentada pelo conjunto de dados, todos os modelos Bayesianos seguem o mesmo princípio, a regra de Bayes. A especificação do modelo começará pela apresentação da regra de Bayes, seguindo, principalmente, a notação utilizada em Koop (2003).

Segundo Koop (2003), a regra de Bayes, apesar de ser uma regra simples de probabilidade, permite que todos os desejos de um economista sejam satisfeitos, como: estimar parâmetros de um modelo, obter previsões e fazer comparações entre modelos diferentes, por exemplo. Para motivar a simplicidade dessa metodologia, pode-se considerar duas variáveis aleatórias,  $A$  e  $B$ . Pela regra da probabilidade, tem-se:

$$p(A, B) = p(A|B)p(B),$$

onde  $p(A, B)$  é a probabilidade conjunta de  $A$  e  $B$  ocorrerem,  $p(A|B)$  é a probabilidade de ocorrência de  $A$  dado que  $B$  ocorreu e  $p(B)$  é a probabilidade marginal de  $B$ . Alternativamente, é possível inverter os papéis de  $A$  e  $B$  e encontrar outra expressão para a probabilidade conjunta de  $A$  e  $B$ :

$$p(A, B) = p(B|A)p(A).$$

Igualando essas duas expressões e rearranjando os termos, chega-se a *regra de Bayes*, que se encontra no cerne da econometria Bayesiana:

$$p(B|A) = \frac{p(A|B)p(B)}{p(A)}. \quad (3.18)$$

Os modelos econômicos normalmente dependem de parâmetros, em se tratando de modelos de regressão, a atenção

direciona-se para os coeficientes do modelo, mais especificamente em como estimá-los. Assim, pode-se adaptar a regra de Bayes para os parâmetros de interesse, os coeficientes do modelo de regressão. Considerando  $y$  como um vetor ou matriz de dados e  $\theta$  como vetor ou matriz que contém os parâmetros do modelo que pretendem explicar  $y$ , a econometria Bayesiana busca fazer inferências sobre  $\theta$ , através da regra de Bayes, baseando-se nos dados,  $y$ . Isto é, substituindo  $B$  por  $\theta$  e  $A$  por  $y$  em (3.18), a regra de Bayes passa a ser:

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)}, \quad (3.19)$$

onde  $p(y) = \int p(y|\theta)p(\theta)d\theta$ . O efeito de dividir por  $p(y)$  é fazer  $p(\theta|y)$  uma distribuição de probabilidade normalizada, isto é, permite que ao integrar a equação (3.19) com respeito a  $\theta$  resulte em  $\int p(\theta|y)d\theta = 1$  (GREENBERG, 2008).

O termo  $p(y|\theta)$  representa a densidade condicional dos dados, dado  $\theta$ , também conhecido como processo gerador dos dados e tem papel central tanto na estatística clássica quanto na Bayesiana (KOOB, 2003). É importante destacar que, uma vez que os dados são conhecidos,  $p(y|\theta)$  é uma função de  $\theta$ , o que não pode ser confundido com a função densidade de probabilidade (FDP) para o  $\theta$ , pois integrando a função em relação a  $\theta$  não chegará a 1, por outro lado, integrando em relação a  $y$  chega-se a 1.

A *densidade a priori*,  $p(\theta)$ , incorpora o conhecimento sobre a distribuição de  $\theta$  anterior a observação dos dados. A priori dependerá do conhecimento de quem está pesquisando sobre o problema, podendo ser baseado em considerações teóricas ou em trabalhos empíricos anteriores (GREENBERG, 2008). Para

Koop (2003), o uso dos próprios dados para formular a priori, apesar de violar a premissa básica do método Bayesiano, tem se tornado mais popular, principalmente em estudos empíricos, e parecem trabalhar bem na prática. No mais, a distribuição a priori geralmente depende de parâmetros, conhecidos como *hiperparâmetros*, que também podem ser fornecidos por quem está pesquisando.

O termo restante,  $p(\theta|y)$  corresponde a *densidade a posteriori* e é de fundamental interesse na metodologia Bayesiana, pois combina em uma única expressão todas as informações que se tem sobre  $\theta$ . Nela estão incluídas as informação obtidas antes de observar os dados, fornecidas pela distribuição a priori, e as informações contidas nos dados, fornecidas pela função de verossimilhança.

Uma vez que o interesse está direcionado em obter informações sobre  $\theta$  apenas, o termo  $p(y)$  pode ser ignorado, já que ele apenas normaliza a distribuição a posteriori, não envolvendo  $\theta$ . Assim, a distribuição a posteriori é frequentemente escrita como:

$$p(\theta|y) \propto p(y|\theta)p(\theta), \quad (3.20)$$

isto é, a distribuição a posteriori é proporcional a função de verossimilhança multiplicada pela distribuição a priori. Isso implica que a integral do lado direito da equação não será igual a um, mas como uma função de  $\theta$ , apresentará a mesma forma de  $p(\theta|y)$ .

Adicionalmente, a equação (3.20) permite que a distribuição a posteriori seja atualizada sempre que houver informação nova disponível, pois a distribuição a posteriori baseada no conjunto de dados inicial ocupa o lugar destinado a distribuição a priori quando as informações forem atualizadas. Isto é, como

colocado em Greenberg (2008), a distribuição a posteriori para os parâmetros baseada nas novas evidências é proporcional a função de verossimilhança para os dados novos, dada a amostra da dados anteriores e parâmetros, multiplicada pela distribuição a posteriori para os parâmetros, dado os dados anteriores.

Apesar de partir de uma regra simples de probabilidade, até pouco tempo, o método Bayesiano restringia-se a uma pequena área dentro da econometria. Problemas relacionados sobretudo às informações a priori e à capacidade computacional restringiam o uso dessa metodologia em muito trabalhos. Koop (2003) cita como exemplo a frequente rejeição por parte de pesquisadores com relação ao uso de informações a priori subjetivas pra inferência dentro das ciências econômicas, supostamente, objetivas. Mas o autor ressalta a honestidade e o rigor com que os modelos Bayesianos permitem determinar o quanto uma informação a priori é útil em cada modelo, e se uma informação a priori está disponível, então ela deve ser utilizada, pois mais informação é sempre melhor do que menos. No mais, o modelo Bayesiano oferece a possibilidade de utilizar o que chama de “priori não informativa”. Isso implica que se uma informação a priori disponível não agrada a quem estiver fazendo a pesquisa, poderá ser atribuído um peso menor para a mesma.

Outro obstáculo que historicamente tem dificultado o desenvolvimento e aplicação da econometria Bayesiana está relacionado a capacidade computacional. Estes modelos envolvem de forma direta ou indiretamente integrais, logo, são raros os casos em que existe solução analítica. Portanto, a revolução computacional dos últimos 30 anos tem sido de extrema importância na superação deste problema. Ainda assim, a metodologia Bayesi-

ana exige conhecimento computacional superior ao utilizado no método clássico, e dada a grande variedade de distribuições a priori e funções de verossimilhança, é muito comum a pesquisadores criarem seus próprios códigos em programas de linguagem matricial. Mas existem pacotes computacionais disponíveis para modelos Bayesianos padrão, que podem ser facilmente adaptados às necessidades dos pesquisadores. Neste trabalho será utilizado um modelo de econometria Bayesina para dados em painel disponibilizado por Koop (2003).

A seção 3.3.2 apresenta um modelo geral de regressão linear com dados em painel dentro da metodologia Bayesiana, que considera a possibilidade de alguns parâmetros da regressão poderem variar entre as zonas eleitorais, enquanto outros poderão se manter fixos.

### 3.3.2 Modelo de Regressão Linear com Dados em Painel

A análise de dados em painel permite trabalhar com a existência de heterogeneidade entre os indivíduos, neste caso, zonas eleitorais. A abordagem Bayesiana para modelos de regressão linear para estes dados possibilita desenvolver tanto modelos de dados empilhados, que assume que os coeficientes de regressão são iguais para todas as zonas eleitorais, quanto modelos de efeitos individuais, onde o intercepto da regressão é diferente para cada zona eleitoral, e também, modelos com coeficientes aleatórios, no qual tanto o intercepto quanto os coeficientes angulares podem variar entre as zonas eleitorais.

Será apresentado aqui um modelo mais geral, no qual alguns coeficientes podem variar entre as zonas eleitorais e outros podem manterem-se fixos. Dessa forma, os modelos de dados

empilhados, de efeito individual e de efeito aleatório representarão casos particulares de tal modelo. Para especificar o modelo é necessário distinguir quais variáveis explicativas apresentarão cada tipo de coeficiente, assim, define-se  $X_i$  como uma matriz  $T \times k$  contendo  $T$  observações para cada  $k$  variável explicativa para o indivíduo  $i$ .  $W_i$  é uma matriz  $T \times k_w$  contendo as  $k_w$  variáveis que apresentarão coeficientes com efeito aleatório e  $Z_i$  é uma matriz  $T \times k_z$  contendo as  $k_z$  variáveis que terão coeficientes com efeito fixo, de modo que  $X_i = [W_i \quad Z_i]$ . A variável dependente e o termo de erro permanecem como já definidas anteriormente.

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \quad e \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_N \end{bmatrix}.$$

Similarmente, juntando todas as variáveis explanatórias tem-se a seguinte matriz  $TN \times k$ :

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_N \end{bmatrix}.$$

O modelo, em sua forma mais geral, pode ser definido da seguinte forma:

$$y_i = Z_i\alpha + W_i\beta_i + \varepsilon_i, \quad \text{para } i = 1, \dots, N \quad (3.21)$$

onde  $\beta_i$  é um vetor de ordem  $k_x$  de coeficientes de regressão diferentes para cada indivíduo,  $\alpha$  é um vetor  $k_z$  de coeficiente de regressão igual para todos os indivíduos.

Apresentada a forma geral do modelo, parte-se para a especificação dos termos característicos da regra de Bayes — função de verossimilhança e as distribuições a priori. A escolha das funções segue os trabalhos de Greenberg (2008) e Koop (2003) para modelos com dados em painel.

### Função de verossimilhança

A função de verossimilhança do modelo Bayesiano segue a mesma estrutura apresentada na metodologia clássica, ela é uma função dos valores dos parâmetros uma vez que os dados do modelo são conhecidos. Segundo Savage (1962), pelo princípio da verossimilhança, considerado como função de  $\theta$ ,  $p(y|\theta)$  contém todas as informações que o experimento tem a oferecer.

A forma da função de verossimilhança é determinada a partir de hipóteses sobre o termo de erro  $\varepsilon$  e a base de dados  $X$ . Para um modelo com dados em painel, assumindo  $i, j = 1, \dots, N$ , tem-se:

1.  $\varepsilon_i$  tem distribuição normal multivariada com média  $0_T$  e matriz de covariância  $h^{-1}I_T$ , onde  $h = \sigma^{-2}$  e refere-se à precisão do erro.
2.  $\varepsilon_i$  e  $\varepsilon_j$  são independentes entre si, para  $i, j = 1, \dots, N$  e  $i \neq j$ .
3. Todos os elementos de  $X_i$  são fixos.

A equação 3.21, junto com as hipóteses sobre o termo de

erro, implicam em uma função de verossimilhança com a forma:

$$p(y|\alpha, \beta, h) = \prod_{i=1}^N \frac{h^{\frac{T}{2}}}{(2\pi)^{\frac{T}{2}}} \left\{ \exp \left[ -\frac{h}{2} (y_i - Z_i\alpha - W_i\beta_i)' (y_i - Z_i\alpha - W_i\beta_i) \right] \right\}, \quad (3.22)$$

onde  $\beta = (\beta_1', \dots, \beta_N')'$ .

### A Priori

O modelo apresenta dois grupos de parâmetros, cada grupo utiliza um tipo diferente de priori. A análise Bayesiana com priori não hierárquica fornece um modelo semelhante ao que é conhecido na econometria frequentista como *efeito fixo*, isto é, cada covariada apresenta parâmetro igual para todos os indivíduos. Já a priori hierárquica tem sido utilizada para superar problemas de espaço de parâmetros de grande dimensão. Por exemplo, os modelos com *efeitos aleatórios* entre  $N$  indivíduos, apresentam um espaço de parâmetros que contém  $N^*k_w + k_z + 1$  parâmetros (i.e.  $N^*k_w$  parâmetros em  $\beta$ ,  $k_z$  parâmetros em  $\alpha$  e a precisão do erro,  $h$ ), se o número de períodos  $T$  for pequeno em relação a  $N$ , o número de parâmetros será grande em relação ao tamanho da amostra, então uma priori hierárquica pode ser apropriada.

Começando pelo conjunto de parâmetros com efeitos aleatórios, uma a priori hierárquica conveniente assume uma distribuição Normal tal como, para  $i = 1, \dots, N$ ,

$$\beta_i \sim \mathcal{N}(\beta_0, \Sigma), \quad (3.23)$$

com  $\beta_i$  e  $\beta_j$  sendo independentes entre si para  $i \neq j$ . A distribuição de  $\beta_i$  é uma função dos hiperparâmetros para a média  $\beta_0$  e para a variância  $\Sigma$ . Quando conhecidos, os valores destes hiperparâmetros são determinados pela pesquisadora. Caso contrário, é atribuída uma distribuição de probabilidade para cada um. Neste caso, os valores de  $\beta_0$  e  $\Sigma$  são desconhecidos, dado que  $\beta_i$  representa um valor diferente para cada zona eleitoral  $i$ . Este segundo nível com as distribuições para os hiperparâmetros é o que caracteriza a priori hierárquica.

A média  $\beta_0$  assume uma distribuição normal, mas agora com média e variância conhecidas e caracterizadas pelos hiperparâmetros  $\underline{\mu}_\beta$  e  $\underline{V}_\beta$ , respectivamente.

$$\beta_0 \sim \mathcal{N}(\underline{\mu}_\beta, \underline{V}_\beta) \quad (3.24)$$

A matriz de variância e covariância  $\Sigma$ , deve ser uma matriz simétrica de ordem  $k_w$  positiva definida, e a distribuição de Wishart atende a estas características. Então  $\Sigma$  assume uma distribuição de Wishart com  $\underline{\rho}$  graus de liberdade, onde  $\underline{\rho} = 0$  representa uma a priori não informativa, e matriz de escala simétrica definida positiva,  $\underline{R}$ , de ordem  $k_w$ .

$$\Sigma^{-1} \sim \mathcal{W}(\underline{\rho}, \underline{R}). \quad (3.25)$$

A distribuição de Wishart é uma generalização da distribuição Gamma, de modo que quando apenas uma covariada apresentar efeito aleatório (i.e.  $k_w = 1$ ) a variância,  $\Sigma$ , vai colapsar em uma distribuição Gamma.

Para os parâmetros restantes, é assumida a priori não hierárquica:

$$\alpha \sim \mathcal{N}(\underline{\mu}_\alpha, \underline{V}_\alpha). \quad (3.26)$$

Os coeficientes fixos,  $\alpha$ , são amostrados de uma distribuição Normal com média  $\underline{\mu}_\alpha$  e variância  $\underline{V}_\alpha$  conhecidas. A precisão do erro,  $h$ , vem de uma distribuição Gamma com média  $\underline{s}^2$  e  $\underline{\nu}$  graus de liberdade.

$$h \sim \mathcal{G}(\underline{s}^{-2}, \underline{\nu}). \quad (3.27)$$

Seguindo a regra de Bayes, a distribuição a posteriori será proporcional a função de verossimilhança (3.22) multiplicada pelas densidades a priori (3.23) - (3.27). Entretanto, esse resultado dificilmente é encontrado de forma analítica, devido, principalmente, a complexidade das integrais. É aqui que entra a importância da revolução na capacidade computacional dos últimos anos, que permitiu desenvolver meios mais rápidos para estimar a distribuição a posteriori, como a estimação por simulação de Monte Carlo via Cadeias de Markov.

### Simulação de Monte Carlo via Cadeias de Markov (MCMC)

O objetivo da estimação Bayesiana é encontrar a distribuição marginal a posteriori dos parâmetros. Com as distribuições a priori assumidas, este resultado não poderá ser alcançado de forma analítica. Mas existem diversas maneiras para lidar com essa limitação, uma delas é por integração numérica, tradicional para avaliar integrais que não tem solução analítica, mas este método é útil apenas em modelos com poucos parâmetros. Outro método é o de simulação, que utiliza números pseudo-aleatórios gerados por computador para amostrar das distribuições, cada

amostra é uma simulação. O método de simulação é uma alternativa ao de integração numérica que pode ser usado para modelos de grande dimensão (GREENBERG, 2008).

O método de simulação de Monte Carlo via Cadeias de Markov (MCMC) aproxima a distribuição exata dos parâmetros a partir da distribuição a posteriori conjunta, que é proporcional a função de verossimilhança (3.22) multiplicada pelas distribuições a priori (3.23) - (3.27). Entretanto, o núcleo da densidade conjunta a posteriori não tem uma forma conhecida e, portanto, não pode ser diretamente usado para fazer inferências sobre a posteriori.

Apesar de não conseguir identificar a distribuição conjunta a posteriori, ainda é possível determinar a forma do núcleo de algumas distribuições condicionais a posteriori. O amostrador de Gibbs, algoritmo comum para simulações de MCMC, requer apenas o núcleo destas distribuições condicionais completas. Assim, antes de partir para o amostrador de Gibbs, serão apresentadas as distribuições condicionais a posteriori derivadas da multiplicação da função de verossimilhança pelas distribuições a priori.

As distribuições a posteriori, identificadas por meio da multiplicação de (3.22) por (3.23) - (3.27), para os parâmetros fixos,  $\alpha$ , e para a precisão do erro,  $h$ , condicionais a  $\beta$ , são, respectivamente:

$$\alpha|y, h, \beta, \beta_0, \Sigma \sim \mathcal{N}(\bar{\mu}_\alpha, \bar{V}_\alpha) \quad (3.28)$$

e

$$h|y, \alpha, \beta, \beta_0, \Sigma \sim \mathcal{G}(\bar{s}^{-2}, \bar{v}) \quad (3.29)$$

onde

$$\begin{aligned}\bar{V}_\alpha &= (\underline{V}_\alpha^{-1} + h \sum_{i=1}^N Z_i' Z_i)^{-1}, \\ \bar{\mu}_\alpha &= \bar{V}_\alpha (\underline{V}_\alpha^{-1} \underline{\mu}_\alpha + h \sum_{i=1}^N Z_i' (y_i - W_i \beta_i)), \\ \bar{\nu} &= TN + \underline{\nu} \quad \text{e} \\ \bar{s}^2 &= \frac{\sum_{i=1}^N (y_i - W_i \beta_i - Z_i \alpha)' (y_i - W_i \beta_i - Z_i \alpha) + \underline{\nu} s^2}{\bar{\nu}}.\end{aligned}$$

A posteriori condicional para os  $\beta_i$ 's são independentes uma da outra, para  $i = 1, \dots, N$ , com distribuição

$$\beta_i | y, \alpha, h, \beta_0, \Sigma \sim \mathcal{N}(\bar{d}, \bar{D}) \quad (3.30)$$

onde

$$\begin{aligned}\bar{D} &= (hW_i' W_i + \Sigma^{-1})^{-1} \quad \text{e} \\ \bar{d} &= \bar{D} (hW_i' (y_i - Z_i \alpha) + \Sigma^{-1} \beta_0).\end{aligned}$$

Para os parâmetros hierárquicos,  $\beta_0$  e  $\Sigma$ , as distribuições a posteriori condicionais são

$$\beta_0 | y, \alpha, h, \beta, \Sigma \sim \mathcal{N}(\bar{\mu}_\beta, \bar{V}_\beta) \quad (3.31)$$

onde

$$\begin{aligned}\bar{V}_\beta &= (N\Sigma^{-1} + \underline{V}_\beta^{-1})^{-1} \quad \text{e} \\ \bar{\mu}_\beta &= \bar{V}_\beta \left( \Sigma^{-1} \sum_{i=1}^N \beta_i + \underline{V}_\beta^{-1} \underline{\mu}_\beta \right).\end{aligned}$$

E

$$\Sigma^{-1}|y, \alpha, h, \beta, \beta_0 \sim \mathcal{W}(\bar{\rho}, [\bar{\rho}\bar{R}]^{-1}) \quad (3.32)$$

onde

$$\begin{aligned} \bar{\rho} &= N + \underline{\rho} \quad \text{e} \\ \bar{R} &= \sum_{i=1}^N (\beta_i - \beta_0)(\beta_i - \beta_0)' + \underline{R}. \end{aligned}$$

O amostrador de Gibbs fará simulações a partir das distribuições condicionais a posteriori (3.28) - (3.32) sequencialmente. Primeiro, deve-se determinar os valores iniciais do parâmetros, identificados por  $\beta_0^{(0)}$  e  $\beta^{(0)}$ . O amostrador de Gibbs vai começar a simular valores para os parâmetros a partir das distribuições a posteriori condicionais a estes valores iniciais. Isto é, vai amostrar  $\Sigma$  de uma distribuição Wishart, condicional a  $\beta_0^{(0)}$  e  $\beta^{(0)}$ , gerando  $\Sigma^{(1)}$ , depois amostra  $\alpha^{(1)}$  de uma distribuição Normal, também condicionada a  $\beta_0^{(0)}$  e  $\beta^{(0)}$ . Dessa forma, amostra  $h^{(1)}$  de uma distribuição Gama condicional a  $\beta_0^{(0)}$ ,  $\beta^{(0)}$  e  $\alpha^{(1)}$ . Por fim, amostra novos valores  $\beta_0^{(1)}$  e  $\beta^{(1)}$  a partir de distribuição posteriori Normal condicionada a  $\beta_0^{(0)}$ ,  $\beta^{(0)}$ ,  $\alpha^{(1)}$  e  $\Sigma^{(1)}$ . Esta é a primeira sequência de simulação. O algoritmo repetirá esta sequência  $S$  vezes, das quais as primeiras  $S_0$  amostragens serão descartadas para eliminar o efeito dos valores escolhidos para  $\beta_0^{(0)}$  e  $\beta^{(0)}$ . As  $S_1$  simulações restantes são utilizadas para criar as estimações a posteriori desejadas (KOOP, 2003; GREENBERG, 2008).

De forma reduzida, o amostrador de Gibbs apresenta os seguintes passos:

1. Escolher os valores para  $\beta_0^{(0)}$  e  $\beta^{(0)}$ .  
Para  $s = 1, \dots, S$ :
2. Amostrar  $\Sigma^{(s)}$  de  $\Sigma|y, \alpha^{(s-1)}, h^{(s-1)}, \beta_i^{(s-1)}, \beta_0^{(s-1)}$ .
3. Amostrar  $\alpha^{(s)}$  de  $\alpha|y, \Sigma^{(s)}, h^{(s-1)}, \beta_i^{(s-1)}, \beta_0^{(s-1)}$
4. Amostrar  $h^{(s)}$  de  $h|y, \Sigma^{(s)}, \alpha^{(s)}, \beta_i^{(s-1)}, \beta_0^{(s-1)}$
5. Amostrar  $\beta_i^{(s)}$  de  $\beta_i|y, \Sigma^{(s)}, \alpha^{(s)}, h^{(s)}, \beta_0^{(s-1)}$
6. Amostrar  $\beta_0^{(s)}$  de  $\beta_0|y, \Sigma^{(s)}, \alpha^{(s)}, h^{(s)}, \beta_i^{(s)}$

Seguindo estes passos chegará a um conjunto com  $S$  amostragens,  $\theta^{(s)}$  (para  $\theta = (\Sigma, \alpha, h, \beta_0, \beta_i)$  e  $s = 1, \dots, S$ ). Depois de descartar as primeiras  $S_0$  amostragens para eliminar o efeito de  $\theta^{(0)}$ , a média das  $S_1$  amostragens restantes pode ser calculada para criar estimações dos valores a posteriori. Assim, escolhendo um  $S$  suficientemente grande o erro de aproximação será reduzido.

### Diagnóstico de Convergência

Para saber quando o  $S$  escolhido é apropriado, Geweke et al. (1991) sugere um diagnóstico de convergência. Partindo da hipótese de que, se for amostrada uma quantidade suficientemente grande, a estimação dos parâmetros baseada na primeira metade da amostragem deve ser essencialmente igual à uma estimação feita com a segunda metade. Caso os parâmetros estimados em cada amostragem forem diferentes, então o número de replicações foi muito baixo e/ou o efeito da primeira amostragem,  $\theta^{(0)}$ , está contaminando a estimação que utiliza as primeiras replicações.

Considerando apenas as  $S_1$  replicações do amostrador de Gibbs, o diagnóstico de convergência é, então, desenvolvido

através da divisão dessas replicações em três subconjuntos. O subconjunto  $S_A$  contém as primeiras amostragens de  $S_1$ , o subconjunto  $S_B$  contém as amostragens do meio e  $S_C$  contém as últimas amostragens. Segundo Koop (2003), se o tamanho de cada subconjunto foi definido como  $S_A = 0.1S_1$ ,  $S_B = 0.5S_1$  e  $S_C = 0.4S_1$ , o diagnóstico funcionará bem em muitas aplicações. Para calcular o diagnóstico são utilizados apenas os subconjuntos  $S_A$  e  $S_C$ , de forma que excluindo o subconjunto do meio, os outros dois conjuntos serão independentes entre si.

Assim, Geweke et al. (1991) mostra que, considerando  $\hat{\theta}_{S_A}$  e  $\hat{\theta}_{S_C}$  como as estimativas dos parâmetros usando as primeiras replicações,  $S_A$ , e as últimas,  $S_C$ , respectivamente, e definindo  $\hat{\sigma}_A/\sqrt{S_A}$  e  $\hat{\sigma}_C/\sqrt{S_C}$  como os erros padrões numéricos destas duas estimações, pelo teorema do limite central, o diagnóstico de convergência pode ser denotado como:

$$DC = \frac{\hat{\theta}_{S_A} - \hat{\theta}_{S_C}}{\frac{\hat{\sigma}_A}{\sqrt{S_A}} + \frac{\hat{\sigma}_C}{\sqrt{S_C}}}. \quad (3.33)$$

Onde valores altos para DC indicam que  $\hat{\theta}_{S_A}$  e  $\hat{\theta}_{S_C}$  são bem diferentes entre si, o que significa que não foram feitas replicações o suficiente para convergência. Se o diagnóstico de convergência indicar que um número suficientemente grande de replicações foram feitas, então o resultado final pode ser calculado baseado no conjunto completo de  $S_1$  replicações.

### Método de Chib para o cálculo da verossimilhança marginal

Uma das grandes vantagens da metodologia de inferência Bayesiana é o campo teórico sólido em comparação de modelos, que depende do cálculo preciso da verossimilhança marginal,  $p(y)$ .

Um dos métodos mais populares no cálculo de  $p(y)$  é o método de Chib, proposto em Chib (1995), que possibilita calcular  $p(y)$  quando a dimensionalidade do espaço de parâmetros é grande, e, além disso, trabalha com o amostrador de Gibbs.

Considerando a notação geral dos parâmetros,  $\theta$ , o método de Chib começa com o rearranjo da Regra de Bayes na seguinte identidade:

$$p(y) = \frac{p(y|\theta)p(\theta)}{p(\theta|y)} \quad (3.34)$$

Como  $p(y)$  não depende de  $\theta$ , a igualdade é válida para qualquer valor de  $\theta$ . Utilizando a média a posteriori denotada por  $\theta^*$  e aplicando o logaritmo, chega-se a seguinte estrutura:

$$\log p(y) = \log p(y|\theta^*) + \log p(\theta^*) - \log p(\theta^*|y). \quad (3.35)$$

Os termos  $p(y|\theta^*)$  e  $p(\theta^*)$  são facilmente computados, mas a densidade a posteriori,  $p(\theta^*|y)$ , gera complicações, já que sua distribuição não é conhecida completamente. Para calcular  $p(\theta^*|y)$ , Chib (1995) propõe a decomposição dos blocos de  $\theta$ , isto é,

$$p(\theta^*|y) = p(\beta^*|y)p(\Sigma^*|y, \beta^*)p(h^*|y, \beta^*, \Sigma^*), \quad (3.36)$$

onde cada termo do lado direito da equação pode ser calculado através de simulações de MCMC.

Por exigir que seja rodado um amostrador de Gibbs para cada bloco de parâmetros, o método de Chib demanda maior capacidade computacional e programação do que outros métodos. Apesar de mais difícil de implementar, segundo Liu e Liu (2012) o método de Chib para o cálculo da verossimilhança marginal é mais preciso.

Por fim, a comparação de modelos na metodologia Bayesiana é feita pelo Fator de Bayes. Considerando  $p(y|M_A)$  e  $p(y|M_B)$  como as verossimilhanças marginais dos modelos  $M_A$  e  $M_B$ , respectivamente, o Fator de Bayes pode ser definido pela seguinte estrutura:

$$FB_{AB} = \frac{p(y|M_A)}{p(y|M_B)}, \quad (3.37)$$

onde um valor para  $FB_{AB} > 1$  indica que  $M_A$  é mais fortemente suportado pelos dados, quando comparado a  $M_B$ . Kass e Raftery (1995, p.777 ) sugerem interpretar a evidência em favor de  $M_A$  como: “não vale mais do que uma simples menção” para  $1 \leq FB_{AB} < 3$  ; “positiva” para  $3 \leq FB_{AB} < 20$ ; “forte” para  $20 \leq FB_{12} < 150$ ; e “muito forte” para  $FB_{AB} > 150$ .

A aplicação da metodologia abordada ao longo deste capítulo será apresentada no Capítulo 5. Antes disso, no capítulo a seguir é feita uma análise descritiva dos dados que serão utilizados no modelo.

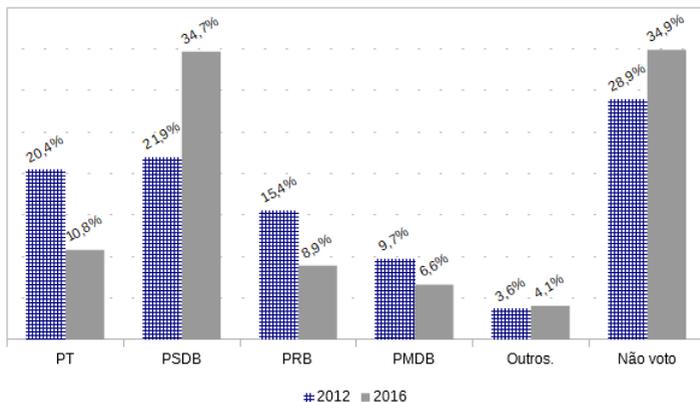
## 4 DADOS

Este trabalho utiliza um painel composto por dados das 58 zonas eleitorais do município de São Paulo nas eleições de 2012 e 2016. Em 2012 haviam 11 candidatos concorrendo ao cargo de prefeito e em 2016 eram 12, formando, assim, uma amostra com um total de 1334 observações para os dois períodos.

A variável dependente, resultado do tratamento algébrico sugerido por Domencich e McFadden (1975) e expresso na equação (3.12), utilizou o percentual de votos computados em cada uma das 58 zonas eleitorais da cidade, disponível no Repositório de Dados Eleitorais do Tribunal Superior Eleitoral (TSE). A Figura 1 apresenta os resultados dos partidos mais votados nas duas eleições, em termos de eleitorado total. A primeira coluna de cada partido representa os resultados de 2012, e a segunda de 2016. Importante destacar que foram utilizados apenas os resultados referentes ao primeiro turno da eleição de 2012.

Como pode ser visto, em 2016 o PT teve seus votos reduzidos a praticamente metade do que recebeu em 2012. O Partido Republicano Brasileiro (PRB) e o Partido do Movimento Democrático Brasileiro (PMDB) também reduziram os votos, mas em proporções menores. O PSDB, que em 2012 teve apenas 1,5 pontos percentuais a mais que o PT, em 2016 parece ter arrecadado grande parte dos votos perdidos pela oposição. O percentual de não votos (brancos, nulos e abstenções) foi superior ao de todos os candidatos nos dois anos, a diferença é que em 2012 o PT e o PSDB foram para o segundo turno e em 2016, João

Figura 1 – Participação nos votos em relação ao eleitorado total nas eleições de 2012 e 2016 para a prefeitura de São paulo



Fonte – Elaboração própria a partir dos dados do Tribunal Superior Eleitoral (TSE).

Dória (PSDB) conseguiu somar mais de 50% do votos válidos (não inclui os não votos: brancos, nulos e abstenções), sendo, portanto, eleito em primeiro turno.

Como a teoria do problema de escolha sugere, a decisão dos eleitores é influenciada pelas características das opções de escolha disponível. Dessa forma, além das variáveis *dummies* que compõem a metodologia de *diff-in-diff*, apresentadas na Seção 3.2 (dP, dT e dP·dT), o conjunto de variáveis explicativas do modelo inclui, também, algumas informações sobre os candidatos, que podem ser observadas tanto pelos eleitores quanto pelo pesquisador. As Tabelas 1 e 2 apresentam as informações referente aos candidatos nas eleições de 2012 e de 2016, respectivamente, na capital paulista.

Por se tratar, principalmente, de atributos, algumas das covariadas foram tabuladas como *dummies*. Para gênero, foi considerado homem = 1 e mulher = 0, a fim de verificar se o fato de um candidato ser homem aumenta, ou diminui, o percentual de votos recebidos. Com relação a naturalidade dos candidatos, foi escolhido São Paulo = 1 e outros = 0, para verificar se paulistanos tendem a receber mais votos. Para o estado civil foram utilizadas duas variáveis *dummies*, uma identificando os candidatos casados e outra para os divorciados, o mesmo foi feito para a escolaridade, definindo uma variável *dummy* para os candidatos com ensino superior completo e outra para aqueles com ensino superior incompleto. Por fim, foi utilizada uma variável *dummy* para identificar os candidatos que participaram de debate transmitido em rede nacional (S = 1, N = 0).

A covariada “Idade” está em nível e informa quantos anos o candidato tinha no ano da eleição. A covariada “Bens”, que relaciona o valor total de bens declarados por cada candidato, está representada aqui em termos de salários mínimos relativo ao ano da eleição, para facilitar a comparação entre os candidatos. Todos os dados das covariadas foram coletados no Repositório de Dados Eleitorais (TSE), com exceção da variável Debate, coletada no portal de notícias G1 (G1 SP). Assim, o modelo conta com 12 variáveis explicativas, mais o intercepto.

Tabela 1 – Características associadas aos candidatos à prefeitura de São Paulo na eleição de 2012.

2012									
Nº	Candidato	Partido	Gênero	Naturalidade	Estado Civil	Escolaridade	Debate	Idade	Bens
10	Celso Russomanno	PRB	M	São Paulo	Casado	Sup. Completo	S	56	3.802,5
12	Paulinho da Força	PDT	M	Porecatu	Casado	E. M. Completo	S	56	792,6
13	Fernando Haddad	PT	M	São Paulo	Casado	Sup. Completo	S	49	761,7
15	Gabriel Chalita	PMDB	M	Cachoeira Paulista	Solteiro	Sup. Completo	S	43	18.523,6
16	Ana Luiza	PSTU	F	Porto Alegre	Divorciada	Sup. Incompleto	N	53	0,0
23	Soninha	PPS	F	São Paulo	Divorciada	Sup. Completo	S	45	578,1
27	Eymael	PSDC	M	Porto Alegre	Casado	Sup. Completo	N	73	7.459,4
28	Levy Fidelix	PRTB	M	Mutum	Casado	Sup. Incompleto	S	61	660,3
29	Anaf Caproni	PCO	F	S. B. do Campo	Casada	Sup. Incompleto	N	46	0,0
45	José Serra	PSDB	M	São Paulo	Casado	Sup. Completo	S	70	2.365,6
50	Carlos Giannazi	PSOL	M	São Paulo	Divorciado	Sup. Completo	S	51	618,7
54	Miguel	PPL	M	São Paulo	Divorciado	Sup. Completo	N	56	845,7

Fonte – Elaboração própria a partir dos dados do Tribunal Superior Eleitoral (TSE).

Tabela 2 – Características associadas aos candidatos à prefeitura de São Paulo na eleição de 2016.

2016									
Nº	Candidato	Partido	Gênero	Naturalidade	Estado Civil	Escolaridade	Debate	Idade	Bens
10	Celso Russomanno	PRB	M	São Paulo	Casado	Sup. Completo	S	60	2.357,9
13	Fernando Haddad	PT	M	São Paulo	Casado	Sup. Completo	S	53	513,6
15	Marta	PMDB	F	São Paulo	Casada	Sup. Completo	S	71	1.5135,5
16	Altino	PSTU	M	São Paulo	Divorciado	Sup. Completo	N	49	0,0
18	Ricardo Young	REDE	M	São Paulo	Divorciado	Sup. Completo	N	59	10.017,5
27	João Bico	PSDC	M	São Paulo	Casado	Ens. Médio Completo	N	49	5.012,5
28	Levy Fidelix	PRTB	M	Mutum	Casado	Sup. Incompleto	N	65	877,2
29	Henrique Áreas	PCO	M	Ribeirão Preto	Solteiro	Sup. Completo	N	31	0,0
45	João Dória	PSDB	M	São Paulo	Casado	Sup. Completo	S	59	204.279,2
50	Luiza Erundina	PSOL	F	Uiraúna	Solteira	Sup. Completo	S	82	688,2
77	Major Olímpio	SD	M	P. Venceslau	Casado	Sup. Completo	S	54	625,6

Fonte – Elaboração própria a partir dos dados do Tribunal Superior Eleitoral (TSE).



## 5 ANÁLISE EMPÍRICA

Serão apresentadas aqui as aplicações do modelo proposto no Capítulo 3. Foram rodados três modelos que têm como única diferença entre si a especificação dos coeficientes relativos às covariadas. O primeiro modelo, chamado de  $M_1$ , corresponde a um modelo com dados empilhados, o mais simples da classe de modelos com dados em painel e assume que os coeficientes de regressão são iguais para todos os indivíduos. O  $M_2$  é um modelo misto, com coeficientes fixos para as covariadas da metodologia de *diff-in-diff* (dP, dT e dP·dT) e aleatórios para todas as outras. Por fim,  $M_3$ , apresenta todos os coeficientes de regressão com efeito aleatório.

### 5.1 Parametrização

O três modelos utilizam a mesma base de dados referente as eleições de 2012 e 2016 para a prefeitura da cidade de São Paulo, apresentada no Capítulo 4. Os resultados a posteriori para todos os modelos são baseado em 105.000 replicações, das quais as primeiras 5.000 foram descartadas.

Com relação às distribuições a priori, foram estabelecidos hiperparâmetros para a precisão do erro,  $h$ , iguais para os três modelos. Como a precisão representa o inverso da variância, quanto maior a precisão, menor será a variância do erro. Como motivação para determinar a informação a priori referente a precisão, foi utilizado a margem de erro apresentadas pelas

pesquisas eleitorais. Apesar de ser muito comum que as pesquisas considerem uma margem de erro de 2 pontos percentuais nos resultados, para um intervalo de confiança de 95%, Zylberstajn et al. (2016) mostra que essa margem foi de 7,5 pontos percentuais, para um intervalo de confiança de 95%, nas eleições brasileiras de 2012, 2014 e 2016.

Assim, para simplificar, será considerada uma possível variação de 8 pontos percentuais no erro de previsão, o que sugere um desvio padrão,  $\sigma$ , em torno de 4, isto é, como os erros são normalmente distribuídos com média 0, se  $\sigma = 4$ , então 95% dos erros serão menores do que  $1,96 \times 4 = 7,84$ . Como  $h = \frac{1}{\sigma^2}$ , uma a priori razoável para  $h$  seria então  $1/16$ , o que resultaria em  $\underline{s}^{-2} = 1/16$ . Entretanto, dada a subjetividade da análise, é possível atribuir um peso pequeno para essa a priori definindo um valor baixo para  $\underline{\nu}$  em relação ao tamanho da amostra,  $NT = 1334$ . Assim, definindo  $\underline{\nu} = 13$ , tem-se  $\frac{\underline{\nu}}{NT} \approx 0,01$ , representando uma priori pouco informativa, isto significa que a informação a priori sobre  $h$  deve ter 1% do peso em relação a informação dos dados.

Além disso, todos os modelos utilizam os resultados estimados por um modelo de Mínimos Quadrados Ordinários (MQO) como valores iniciais dos coeficientes para começar as simulações do Amostrador de Gibbs. Os outros hiperparâmetros foram definidos com relação a característica de cada modelo. O modelo  $M_1$  refere-se a um modelo com dados empilhados, com a hipótese de que os erros são independentes entre todos os indivíduos e todos os períodos, isso faz com que o modelo se reduza a um modelo de regressão linear com priori Normal-Gama, então:  $\alpha \sim \mathcal{N}(0_{13}, I_{13})$  e  $h \sim \mathcal{G}(\frac{1}{16}, 13)$ .

O modelo  $M_2$  corresponde a um modelo misto com coeficientes fixos para as covariadas da metodologia de contrafactuais —  $dP$ ,  $dT$  e  $dP \cdot dT$  — e coeficientes aleatórios para o restante das covariadas. Esse modelo utiliza priori hierárquica para os parâmetros com efeitos aleatórios e priori Normal-Gama independente para os coeficientes fixos. O modelo  $M_3$  apresenta todos os coeficientes com efeito aleatório, conseqüentemente, todos terão priori hierárquica.

Tabela 3 – Parametrização

Parâmetro	$M_1$	$M_2$	$M_3$
$k$	13	13	13
$k_w$	0	10	13
$k_z$	13	3	0
$\underline{\mu}_\beta$	-	$0_{10}$	$0_{13}$
$\underline{V}_\beta$	-	$I_{10}$	$I_{13}$
$\underline{R}$	-	$0,5I_{10}$	$0,5I_{13}$
$\underline{\rho}$	-	10	13
$\underline{\mu}_\alpha$	$0_{13}$	0	-
$\underline{V}_\alpha$	$I_{13}$	1	-
$\underline{s}^{-2}$	1/16	1/16	1/16
$\underline{\nu}$	13	13	13
	Distribuição		
$\beta_i$	-	Normal	Normal
$\beta_0$	-	Normal	Normal
$\Sigma$	-	Wishart	Wishart
$\alpha$	Normal	Normal	-
$h$	Gama	Gama	Gama

Fonte – Elaboração própria baseada em Koop (2003)

Todos os hiperparâmetros escolhidos para os três modelos são apresentados na Tabela (3). Os valores dos graus de

liberdade, baixos em relação ao tamanho da amostra, e a matrizes de variância e covariância revelam que as informações a priori escolhidas são pouco informativas, isto é, as estimações dos modelos darão mais peso para as informações obtidas com os dados do que com as informações fornecidas à priori.

## 5.2 Resultados

Cada modelo apresenta um elevado número de parâmetros como resultado da estimação da densidade a posteriori,  $p(\theta|y)$ . Dessa forma, serão apresentados aqui somente os resultados a posteriori de maior relevância para a análise do comportamento dos eleitores. Os resultados completos provenientes de cada modelo se encontram no Apêndice - A.

As colunas identificadas por  $E(\theta|y)$ , na Tabela 4, apresentam as estimativas das médias dos coeficientes angulares e dos interseptos. Para o modelo  $M_1$  são apresentadas as médias a posteriori dos  $\alpha$ 's. Com relação aos modelos  $M_2$  e  $M_3$ , foram consideradas as estimações das médias,  $\beta_0$ , dos coeficientes com efeitos variáveis, para facilitar a comparação entre os três modelos. Como  $\Sigma$  corresponde a variância de  $\beta_i$ , só há estimativas para esse parâmetro nos modelos com coeficientes aleatórios. Todos os valores entre parênteses correspondem aos respectivos desvios padrões dos parâmetros.

Além dos coeficientes de regressão, a tabela também mostra a estimação da verossimilhança marginal de cada modelo, calculada pelo método de Chib. Percebe-se que  $M_1$ , o modelo com efeitos fixos para todos os coeficientes, apresenta o maior valor para  $p(y)$ , o que indica que, entre os três modelos propostos,  $M_1$  é o mais fortemente suportado pelos dados. Calculando o fator de Bayes<sup>1</sup>, chega-se a  $FB_{12} = 1,31$ ,  $FB_{13} = 1,70$ ,  $FB_{23} = 1,30$ , que, segundo a escala de interpretação de Kass e Raftery (1995), significa que “não vale mais do que uma simples menção” de que o modelo  $M_2$  é melhor que o  $M_3$  e que  $M_1$  é melhor do que  $M_2$  e melhor do que  $M_3$ , mas a diferença entre os modelos não chega

<sup>1</sup>  $FB_{AB} = \exp[\log p(y|M_A) - \log p(y|M_B)]$ .

a ser forte.

Tabela 4 – Resultado a Posteriori para os Parâmetros Seleccionados

$\theta$	$M_1$		$M_2$		$M_3$	
	$E(\theta y)$	$\Sigma$	$E(\theta y)$	$\Sigma$	$E(\theta y)$	$\Sigma$
Intercepto	-8,9995 (0,2168)	-	-8,8625 (0,2462)	0,4639 (0,1657)	-8,8080 (0,2586)	0,5455 (0,1853)
dGênero	-0,3968 (0,0897)	-	-0,4195 (0,1140)	0,2287 (0,0568)	-0,4274 (0,1216)	0,2764 (0,0666)
dSão Paulo	1,0316 (0,1012)	-	1,0171 (0,1254)	0,2477 (0,0640)	1,0142 (0,1339)	0,3019 (0,0759)
dCasado	-1,5159 (0,1315)	-	-1,4921 (0,1577)	0,2897 (0,0820)	-1,4859 (0,1663)	0,3453 (0,0945)
dDivorciado	-1,4547 (0,1301)	-	-1,4498 (0,1556)	0,2806 (0,0780)	-1,4495 (0,1631)	0,3409 (0,0919)
dE.S. Comp.	0,8937 (0,1225)	-	0,8848 (0,1464)	0,2775 (0,0764)	0,8797 (0,1559)	0,3352 (0,0897)
dE.S. Incomp.	1,0279 (0,1377)	-	0,9981 (0,1655)	0,3173 (0,0924)	0,9868 (0,1741)	0,3782 (0,1063)
dDebate	3,0965 (0,0864)	-	3,0683 (0,1107)	0,2279 (0,0563)	3,0586 (0,1196)	0,2824 (0,0688)
ln_Bens	0,1332 (0,0166)	-	0,1360 (0,0452)	0,0990 (0,0192)	0,1366 (0,0505)	0,1270 (0,0246)
Idade	0,0482 (0,0040)	-	0,0464 (0,0397)	0,0896 (0,0173)	0,0454 (0,0452)	0,1164 (0,0223)
dP	0,2033 (0,0700)	-	0,1966 (0,0747)	-	0,1899 (0,1020)	0,2458 (0,0562)
dT	2,1776 (0,1593)	-	2,1612 (0,1696)	-	2,1360 (0,1952)	0,4065 (0,1189)
dP·dT	<b>-1,0412</b> (0,2057)	-	<b>-1,0194</b> (0,2193)	-	<b>-0,9937</b> (0,2449)	0,4927 (0,1589)
$\log[p(y)]$	-2.50		-2.77		-3.03	

Fonte – Elaboração própria a partir dos resultados dos modelos  $M_1$ ,  $M_2$  e  $M_3$ .  
Os valores entre parênteses são os desvios padrões dos parâmetros.

## O Efeito da Exposição Midiática Negativa nas Eleições

O efeito da exposição midiática foi calculado por intermédio dos parâmetros da metodologia de diferença em diferenças —  $dP$ ,  $dT$  e  $dP \cdot dT$ . Lembrando que  $dP = 1$  para os resultados das eleições de 2016,  $dT = 1$  para o candidato do PT (considerado como o alvo da exposição midiática) e  $dP \cdot dT$  é a multiplicação das duas variáveis anteriores, que seria o mesmo que uma *dummy* para o PT nas eleições de 2016, e é o coeficiente desta última variável que indica o efeito da exposição midiática negativa sobre os resultados das eleições na capital paulista.

Pela estrutura do modelo, não é possível interpretar os coeficientes estimados diretamente como a influência de cada covariada no percentual de votos de cada candidato  $j$ . Será necessário substituir os coeficientes estimados (Tabela 4) e a participação ( $s_j$ ) de cada partido nas eleições de 2016 (Figura 1) nas equações de elasticidade, apresentadas na Seção 3.1.

A derivada parcial da proporção de votos  $s_j$  com relação a uma covariada  $x_{jk}$ , resulta, por definição, no efeito causado em  $s_j$  pela variação dessa covariada quando as demais  $k - 1$  variáveis explicativas mantiverem-se constantes. Assim, tendo em mãos a participação do PT na eleição para a prefeitura de São Paulo em 2016 ( $s_{pt} = 0,108$ ), a equação de elasticidade do modelo LOGIT simples (3.14a) pode ser reescrita para cada covariada  $k$  como:

$$\frac{\partial s_{pt}}{\partial x_k} = \beta_k \times 0,108(1 - 0,108), \quad (5.1)$$

onde  $\beta_k$  representa o coeficiente estimado para cada covariada  $k$  do modelo.

Assim, substituindo o parâmetro de  $dP \cdot dT$  do modelo  $M_1$  na Eq. (5.1), encontra-se que, estando tudo o mais contante, o fato de o PT ter passado pela exposição midiática está associado a uma redução de 10 pontos percentuais, em média, na proporção de votos recebidos pelo partido na eleição. Isto é, caso o partido não tivesse sido exposto pela mídia, o seu percentual de votos na eleição para a prefeitura de São Paulo em 2016 teria sido em torno de 20,8%, valor até superior ao de 2012.

O *diff-in-diff* possibilita criar esse cenário contrafactual, sobre qual teria sido o resultado do partido na ausência de exposição, a partir das médias das variações de todos os outros partidos. Em trabalho como este é comum apresentar a estimação do cenário contrafactual em um gráfico junto com as tendências dos grupos exposto e o de controle, onde o contrafactual do grupo exposto seguiria a mesma tendência do grupo de controle. Mas, como se trata de um modelo com alternativas de escolha mutuamente excludentes, o aumento de votos para um partido implica, conseqüentemente, em redução nos votos para as outras opções.

Dessa forma, caso não tivesse ocorrido a exposição sobre o PT, não seria apenas o percentual de votos do PT que mudaria, mas o de todos os outros partidos também, inclusive os não votos. Isso dificulta a apresentação do cenário contrafactual na forma padrão. Por outro lado, a estrutura do modelo permite criar um cenário contrafactual para todas as opções de escolha utilizando a Eq. (3.14) de elasticidade cruzada. Assim, é possível mostrar, não somente como o PT estaria na ausência da exposição, mas também qual seria o percentual de votos recebidos por todos os outros partidos e dos não votos.

A Tabela 5 apresenta os resultados do cálculo das elasticidades própria e cruzadas em relação ao percentual de votos de cada partido após a exposição do PT. Os valores foram multiplicados por 100 para que as variações fossem interpretadas em termos de pontos percentuais. As estimações mostram que os partidos que absorveram a maioria dos votos, teoricamente, perdidos pelo PT, foram o PSDB, PRB, PMDB e os não votos.

Tabela 5 – Elasticidade da exposição midiática do PT na proporção de votos de cada partido na eleição de 2016 em São Paulo – Calculada a partir da Eq.(3.14)

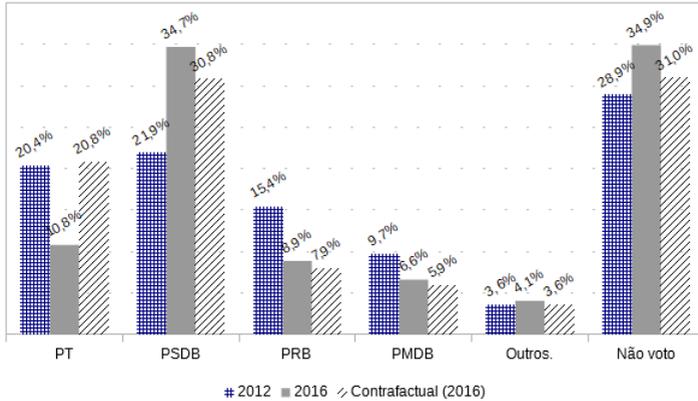
	PT	PSDB	PRB	PMDB	PSTU	REDE	PSDC	PRTB	PCO	PSOL	SD	Não-voto
PT	-10,03	3,95	0,99	0,70	0,01	0,03	0,01	0,03	0,00	0,23	0,15	3,93

Fonte – Elaboração própria a partir dos resultados do modelo  $M_1$ .

A Figura 2 apresenta a estimativa do cenário contrafactual para os resultados dos partidos selecionados nas eleições de 2016 para a prefeitura do município de São Paulo, criado a partir das estimações da Tabela 5. Como pode ser visto, na ausência da exposição midiática, o PT teria recebido praticamente o dobro do número de votos que efetivamente recebeu. Em contrapartida, o PSDB teria seu percentual reduzido em torno de 4 pontos, valor muito próximo ao da variação dos não votos, o PRB receberia em torno de 1 ponto percentual a menos, e o PMDB, 0,7.

Em geral, essa diferença entre o resultado do contrafactual e o resultado que realmente ocorreu para todos os partidos, com exceção do PT, foi de aproximadamente 12%. Assim, percebe-se claramente o problema causado pela utilização das

Figura 2 – Cenário contrafactual para o resultado da eleição de São Paulo em 2016, em termos de votos totais, a partir da Eq.(3.14)



Fonte – Elaboração própria a partir dos dados do Tribunal Superior Eleitoral (TSE) e dos resultados do modelo  $M_1$ .

elasticidades calculadas por meio de um modelo LOGIT simples. Quando aumenta a participação de uma opção, todas as outras reduzem de forma proporcional. Considerando pelo lado do eleitor, seria como se existissem apenas duas opções de voto: votar no PT ou não votar no PT.

Por outro lado, utilizando as estimativas do modelo com coeficientes aleatórios ( $M_3$ ) e seus respectivos desvios padrões para estimar as elasticidades a partir de (3.15), como sugere Petterini e DeSouza (2009), espera-se que os valores encontrados estejam mais próximos dos verdadeiros valores, já que, dessa forma, as diferenças entre os indivíduos estarão sendo levadas em conta.

A Tabela 6 apresenta apenas a primeira linha da ma-

triz de elasticidades estimada a partir de (3.15) em relação a exposição midiática sobre o PT. Como esperado, os valores são diferentes daqueles apresentados na Tabela 5. Em relação a variação nos votos do PT após a exposição, estima-se que houve uma redução de 4,06 pontos percentuais. As estimativas mostram que o PMDB recebeu a maior parte desses votos perdidos pelo PT, tendo um aumento de 1,58 pontos percentuais após a exposição, e a opção de não votos aumentou aproximadamente 1,15 pontos percentuais.

Tabela 6 – Elasticidade da exposição midiática do PT na proporção de votos de cada partido na eleição de 2016 em São Paulo – Calculada a partir da Eq.(3.15)

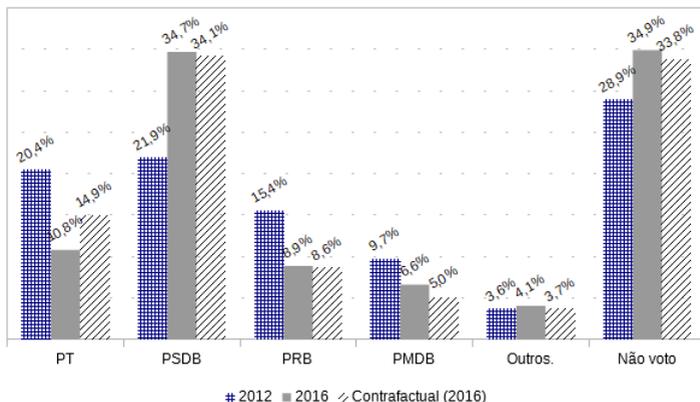
	PT	PSDB	PRB	PMDB	PSTU	REDE	PSDC	PRTB	PCO	PSOL	SD	Não-voto
PT	-4,06	0,60	0,32	1,58	0,00	0,02	0,00	0,01	0,00	0,31	0,06	1,15

Fonte – Elaboração própria a partir dos resultados do modelo  $M_3$ .

A Figura 3 ilustra o cenário contrafactual construído a partir da Tabela 6. Percebe-se, agora, como os votos perdidos pelo PT com a exposição, não são mais distribuídos proporcionalmente entre todas as outras opções.

Pelos resultados do contrafactual da Figura 3, na ausência de exposição midiática o PT teria recebido em torno de 14,9% dos votos totais na eleição, aproximadamente 6 pontos percentuais a menos do que era estimado no modelo com coeficientes fixos. Como contrapartida, o contrafactual estimado para as outras opções de votos ficaram relativamente próximos dos resultados de fato ocorrido.

Figura 3 – Cenário contrafactual para o resultado da eleição de São Paulo em 2016, em termos de votos totais, a partir da Eq.(3.15)



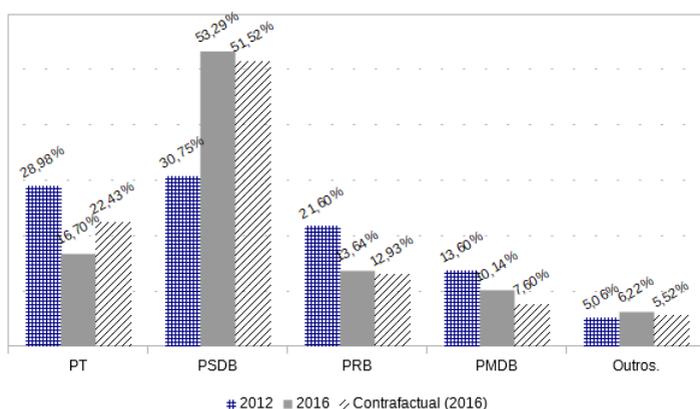
Fonte – Elaboração própria a partir dos dados do Tribunal Superior Eleitoral (TSE) e dos resultados do modelo  $M_3$ .

A estimativa da variação nos percentual de votos para o PSDB, resultante da exposição sobre o PT, é pequena, em torno de 0,6 ponto percentual. Mas para verificar se, caso a exposição não tivesse ocorrido, João Dória ainda teria sido eleito em primeiro turno é necessário excluir da contagem os não votos, já que o resultado oficial das eleições considera apenas os votos válidos.

A Figura 4 apresenta o percentual, em relação ao total de votos válidos, dos votos recebidos pelos candidatos nas eleições de 2012 e 2016 e o contrafactual para a eleição de 2016. Como pode ser visto, na situação hipotética de não exposição midiática sobre o PT, o candidato do PSDB receberia em torno de 51,5% dos votos válidos. Isto é, pelos resultados do modelo, mesmo se o

PT não tivesse sofrido a exposição, ainda assim a eleição teria se decidido no primeiro turno, com João Dória sendo eleito a prefeito. Fernando Haddad permaneceria como segundo colocado, com 22,4% dos votos válidos.

Figura 4 – Cenário contrafactual para o resultado da eleição de São Paulo em 2016, em termos de votos válidos.



Fonte – Elaboração própria a partir dos dados do Tribunal Superior Eleitoral (TSE) e dos resultados do modelo  $M_3$ .

Os resultados levam à conclusão de que a exposição excessiva da mídia sobre o PT influenciou na decisão dos votos, ocasionando uma redução de aproximadamente 4 pontos percentuais na disposição dos eleitores em votar para o partido exposto. Como consequência, os eleitores que votariam no PT, mas devido à exposição acabaram mudando o voto, escolheram principalmente o PMDB ou não votar. Assim, a partir dessa análise, não é possível atribuir à exposição midiática sobre o PT, como a causa da vitória do PSDB na eleição para a prefeitura de São Paulo em 2016.



## 6 CONCLUSÃO

O presente estudo se propôs a avaliar o efeito da exposição midiática nas eleições, utilizando o município de São Paulo como cenário para a análise. Partiu-se da hipótese de que a exposição do PT nos meios de comunicação de massa, relacionada principalmente com o impeachment da ex-presidente Dilma Rousseff e às investigações da Lava Jato, teve efeitos negativos para o mesmo partido nas eleições de 2016. Para construção da variável de impacto, foram considerados os resultados das eleições de 2012 e 2016, computados em cada umas das 58 zonas eleitorais do município.

A primeira dificuldade encontrada foi a impossibilidade de mensurar a exposição da mídia, pois não existem dados disponíveis para esta variável, como, por exemplo, por quanto tempo ou quantas vezes um telejornal expôs cada partido, ou candidato. A estratégia empírica para contornar esse obstáculo foi a utilização da metodologia de diferença em diferenças, que permite inferir sobre como teria sido o resultado da eleição, caso a exposição não tivesse ocorrido, isto é, permite construir um contrafactual.

As covariadas resultantes do método de contrafactual foram inseridas, junto com outras variáveis explicativas referentes às características dos candidatos, em um modelo de escolha discreta. Foram propostas duas abordagens para estimar os parâmetros de regressão. A primeira, mais ingênua, considerou que todas as zonas eleitorais votavam de forma semelhante, resultando em um modelo com coeficientes fixos. A segunda, mais

robusta, levou em conta as diferenças que existem entre as zonas eleitorais, permitindo assim que os coeficientes variassem entre as zonas.

Apesar de ser possível utilizar um método econômico tradicional para estimar os parâmetros do modelo simples, optou-se por estimar os dois modelos por meio de um método de regressão Bayesiano, já que este pareceu ser o método mais apropriado para estimar os parâmetros do modelo com coeficientes aleatórios. Além disso, foi estimado um terceiro modelo com coeficientes fixos para as covariadas do *diff-in-diff* e coeficientes aleatórios para todas as outras.

Os três modelos apresentaram resultados muito próximos para as médias dos coeficientes de regressão. Entretanto, quando calculadas as elasticidades do efeito da exposição midiática nos resultados dos partidos, característica dos modelos de escolha, percebe-se uma grande diferença nos métodos de estimação. No modelo com coeficientes fixos, o mais simples, o impacto estimado indicam uma variação de aproximadamente 10 pontos percentuais, para menos, nos votos para o PT após a exposição do partido. Já no modelo com coeficientes aleatórios essa variação se reduz para aproximadamente 4 pontos percentuais.

Essa diferença entre as abordagens ocorre porque os modelos com coeficientes fixos costumam mascarar valores que poderiam ser estimados por meio de metodologias alternativas, o que resultou, aqui, na atribuição de toda a variação no percentual de votos do PT entre as duas eleições à exposição midiática. Enquanto que o modelo com coeficientes aleatórios consegue captar de forma mais acurada o efeito da exposição, considerando que apenas 4 p.p. da variação nos votos do PT foi causada pela

mídia, os outros 6 p.p. da variação foram causados por outros fatores, possivelmente relacionados às zonas eleitorais.

Assim, um desdobramento para este estudo, é considerar as questões geográficas, já que existem grandes diferenças entre as zonas eleitorais. Pode-se, em trabalhos futuros, analisar como a exposição midiática afeta os votos em cada zona eleitoral, ou em subgrupos como zonas ricas e pobres, centro e periferia, entre outros. Ainda há muito a ser explorado com relação aos efeitos dos meios de comunicação no campo eleitoral.



# REFERÊNCIAS

ALMEIDA, A. T. S. de; LIMA, V. T. da S. Dilma rousseff na imprensa brasileira: Da reeleição ao processo de impeachment. **Revista Encontros**, v. 14, n. 26, p. 102–113, 2016.

ASHENFELTER, O. Estimating the effect of training programs on earnings. **The Review of Economics and Statistics**, p. 47–57, 1978.

ASHENFELTER, O. C.; CARD, D. Using the longitudinal structure of earnings to estimate the effect of training programs. **National Bureau of Economic Research**, Cambridge, Mass., USA, 1984.

AVELAR, L. As eleições na era da televisão. **Revista de Administração de Empresas**, SciELO Brasil, v. 32, n. 4, p. 42–57, 1992.

BARTELS, L. M. Messages received: the political impact of media exposure. **American Political Science Review**, Cambridge Univ Press, v. 87, n. 02, p. 267–285, 1993.

BERRY, S.; LEVINSOHN, J.; PAKES, A. Automobile prices in market equilibrium. **Econometrica: Journal of the Econometric Society**, p. 841–890, 1995.

CAPUTO, V. As surpresas do primeiro turno nas eleições 2016. **Exame**, 3 de outubro de 2016. Disponível em: <<https://exame.abril.com.br/brasil/as-surpresas-do-primeiro-turno-nas-eleicoes-2016/>>.

CASCETTA, E. **Transportation systems analysis: models and applications**. [S.l.]: Springer Science & Business Media, 2009. v. 29.

CASELLA, G.; BERGER, R. L. **Statistical inference**. [S.l.]: Duxbury Pacific Grove, CA, 2002. v. 2.

CHIB, S. Marginal likelihood from the gibbs output. **Journal of the american statistical association**, Taylor & Francis Group, v. 90, n. 432, p. 1313–1321, 1995.

DOMENCICH, T. A.; MCFADDEN, D. **Urban travel demand-a behavioral analysis**. [S.l.], 1975.

DUBÉ, J.-P.; FOX, J. T.; SU, C.-L. Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation. **Econometrica**, Wiley Online Library, v. 80, n. 5, p. 2231–2267, 2012.

ELLIOTT, J. E. Joseph a. schumpeter and the theory of democracy. **Review of Social Economy**, Taylor & Francis, v. 52, n. 4, p. 280–300, 1994.

FRIDKIN, K. L.; KENNEY, P. Variability in citizens' reactions to different types of negative campaigns. **American Journal of Political Science**, Wiley Online Library, v. 55, n. 2, p. 307–325, 2011.

G1 SP. Candidatos à prefeitura de são paulo discutem propostas em debate na tv. **G1**, 25 de setembro de 2012. [Online; acessado em 9-fevereiro-2017]. Disponível em: <<http://g1.globo.com/sao-paulo/eleicoes/2012/noticia/2012/09/candidatos-prefeitura-de-sao-paulo-discutem-propostas-em-debate-na-tv.html>>.

G1 SP. Globo faz debate com candidatos à prefeitura de são paulo nesta quinta. **G1**, 29 de setembro de 2016. [Online; acessado em 9-fevereiro-2017]. Disponível em: <<http://g1.globo.com/sao-paulo/eleicoes/2016/noticia/2016/09/globo-faz-debate-com-candidatos-prefeitura-de-sao-paulo-nesta-quinta.html>>.

GERTLER, P. J. et al. **Avaliação de Impacto na Prática**. [S.l.]: World Bank Publications, 2015.

GEWEKE, J. et al. **Evaluating the accuracy of sampling-based approaches to the calculation of posterior**

**moments.** [S.l.]: Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA, 1991. v. 196.

GREENBERG, E. **Introduction to Bayesian Econometrics.** [S.l.]: Cambridge University Press, 2008.

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

IMBENS, G. W.; WOOLDRIDGE, J. M. Recent developments in the econometrics of program evaluation. **Journal of economic literature**, American Economic Association, v. 47, n. 1, p. 5–86, 2009.

INTERVOZES e RSF. **Monitorados de Propriedade de Mídia - Brasil.** 2017. Disponível em: <<http://brazil.mom-rsf.org/br/>>.

KASS, R. E.; RAFTERY, A. E. Bayes factors. **Journal of the american statistical association**, Taylor & Francis Group, v. 90, n. 430, p. 773–795, 1995.

KNITTEL, C. R.; METAXOGLU, K. Estimation of random-coefficient demand models: Two empiricists' perspective. **Review of Economics and Statistics**, MIT Press, v. 96, n. 1, p. 34–59, 2014.

KOOP, G. **Bayesian Econometrics.** [S.l.]: Wiley, 2003.

LAU, R. R. Negativity in political perception. **Political Behavior**, Springer, v. 4, n. 4, p. 353–377, 1982.

LAZARSELD, P. F. et al. **The peoples choice: how the voter makes up his mind in a presidential campaign.** [S.l.: s.n.], 1948.

LIU, C.; LIU, Q. Marginal likelihood calculation for the gelfand–dey and chib methods. **Economics Letters**, Elsevier, v. 115, n. 2, p. 200–203, 2012.

- MANSKI, C. F. The structure of random utility models. **Theory and decision**, Springer, v. 8, n. 3, p. 229–254, 1977.
- MCFADDEN, D. The measurement of urban travel demand. **Journal of public economics**, Elsevier, v. 3, n. 4, p. 303–328, 1974.
- MIGUEL, L. F. Mídia e vínculo eleitoral: a literatura internacional e o caso brasileiro. **Opinião Pública**, SciELO Brasil, v. 10, n. 1, p. 91–111, 2004.
- NEYMAN, J. Sur les applications de la theorie des probabilités aux expériences agricoles: Essai des principes. **Statistical Science**, Institute of Mathematical Statistics, v. 5, p. 465–472, 1923.
- PETTERINI, F.; DESOUZA, S. Elasticidades e markups do mercado brasileiro de automóveis: uma análise mixed logit. In: **31º Meeting of the Brazilian Econometric Society**. [S.l.: s.n.], 2009.
- PMB. **Relatório Final Pesquisa Brasileira de Mídia**. Brasília - DF: Secretaria de Comunicação Social, 2016.
- PUDNEY, S. Modelling individual choice. **The econometrics of corners, kinks and holes**, 1989.
- RUBIN, D. B. Estimating causal effects of treatments in randomized and nonrandomized studies. **Journal of Educational Psychology**, American Psychological Association, v. 66, n. 5, p. 688, 1974.
- SAVAGE, L. J. **The foundations of statistical inference**. [S.l.]: Methuen, 1962.
- SCHMITT-BECK, R. Mass communication, personal communication and vote choice: The filter hypothesis of media influence in comparative perspective. **British Journal of Political Science**, Cambridge University Press, v. 33, n. 2, p. 233–259, 2003.

SCHUMPETER, J. A. Capitalismo, socialismo e democracia. **Rio de Janeiro: Editora Fundo de Cultura**, 1961.

SEKHON, J. S. The neyman-rubin model of causal inference and estimation via matching methods. **The Oxford handbook of political methodology**, Citeseer, v. 2, 2008.

SILVA, P.; HASHIGUTI, S. T. Uma espada através do corpo de Dilma Rousseff: mídia, discurso e imagem. **MOARA—Revista Eletrônica do Programa de Pós-Graduação em Letras ISSN: 0104-0944**, v. 2, n. 40, p. 35–46, 2016.

SINGER, A. Mídia e democracia no Brasil. **Revista usp**, n. 48, p. 58–67, 2001.

SINGER, A. Derrota da esquerda não significa avanço de valores liberais. **Folha de S.Paulo**, 2016. [Online; acessado em 20-novembro-2016]. Disponível em: <<http://www1.folha.uol.com.br/colunas/andresinger/2016/11/1829538-derrota-da-esquerda-nao-significa-avanco-de-valores-liberais.shtml>>.

TELLES, H. d. S.; LOURENÇO, L. C.; STORNI, T. P. L. Partidos, campanhas e voto: como o eleitor decide nas municipais. **Sociedade e Cultura**, Universidade Federal de Goiás, v. 12, n. 1, 2009.

TRAIN, K. E. Discrete choice methods with simulation. Citeseer, 2002.

TSE. Repositório de dados eleitorais. **Tribunal Superior Eleitoral**. Disponível em: <<http://www.tse.jus.br/eleicoes/estatisticas/repositorio-de-dadoseleitorais>>.

ZALLER, J. The statistical power of election studies to detect media exposure effects in political campaigns. **Electoral Studies**, Elsevier, v. 21, n. 2, p. 297–329, 2002.

ZELLNER, A. **An introduction to Bayesian inference in econometrics**. [S.l.], 1971.

ZYLBERSTAJN, E. et al. A (im)precisão das pesquisas eleitorais. **O Estado de S. Paulo**, Novembro 2016. Disponível em: <<http://economia.estadao.com.br/blogs/almoco-gratis/a-impresicao-das-pesquisas-eleitorais/>>.

# APÊNDICE A – RESULTADOS

Tabela 7 – Resultados a Posteriori para o Modelo  $M_1$

Parâmetro	Média a Posteriori	DP	n.s.e
$\alpha(\text{Intercepto})$	-8,9995	0,2168	0,0007
$\alpha(\text{Gênero})$	-0,3968	0,0897	0,0003
$\alpha(\text{S.Paulo})$	1,0316	0,1012	0,0003
$\alpha(\text{Casado})$	-1,5159	0,1315	0,0004
$\alpha(\text{Divorciado})$	-1,4547	0,1303	0,0004
$\alpha(\text{Sup.Completo})$	0,8937	0,1225	0,0004
$\alpha(\text{Sup.Incompleto})$	1,0279	0,1377	0,0004
$\alpha(\text{Debate})$	3,0965	0,0864	0,0003
$\alpha(\text{Bens})$	0,1332	0,0166	0,0001
$\alpha(\text{Idade})$	0,0482	0,0040	0,0000
$\alpha(dP)$	0,2033	0,0700	0,0002
$\alpha(dT)$	2,1776	0,1593	0,0005
$\alpha(dD \cdot dT)$	-1,0412	0,2057	0,0007
$h$	0,8800	0,0343	0,0001

Fonte – Elaboração própria a partir dos resultados do modelo  $M_1$

Tabela 8 – Resultados a Posteriori para o Modelo  $M_2$ 

Parâmetro	Média a Posteriori	DP	n.s.e
$\beta_0(\text{Intercepto})$	-8,8625	0,2462	0,0008
$\beta_0(\text{Gênero})$	-0,4195	0,1140	0,0004
$\beta_0(\text{S.Paulo})$	1,0171	0,1254	0,0004
$\beta_0(\text{Casado})$	-1,4921	0,1577	0,0005
$\beta_0(\text{Divorciado})$	-1,4498	0,1556	0,0005
$\beta_0(\text{Sup.Completo})$	0,8848	0,1464	0,0005
$\beta_0(\text{Sup.Incompleto})$	0,9981	0,1655	0,0005
$\beta_0(\text{Debate})$	3,0683	0,1107	0,0004
$\beta_0(\text{Bens})$	0,1360	0,0452	0,0001
$\beta_0(\text{Idade})$	0,0464	0,0397	0,0001
$\alpha(dP)$	0,1966	0,0747	0,0002
$\alpha(dT)$	2,1612	0,1696	0,0005
$\alpha(dD \cdot dT)$	-1,0194	0,2193	0,0007
$h$	0,7659	0,0324	0,0001
$\Sigma_{1,1}$	0,4639	0,1657	0,0005
$\Sigma_{2,2}$	0,2287	0,0568	0,0002
$\Sigma_{3,3}$	0,2477	0,0640	0,0002
$\Sigma_{4,4}$	0,2897	0,0820	0,0003
$\Sigma_{5,5}$	0,2806	0,0780	0,0002
$\Sigma_{6,6}$	0,2775	0,0764	0,0002
$\Sigma_{7,7}$	0,3173	0,0924	0,0003
$\Sigma_{8,8}$	0,2279	0,0563	0,0002
$\Sigma_{9,9}$	0,0990	0,0192	0,0001
$\Sigma_{10,10}$	0,0896	0,0173	0,0001

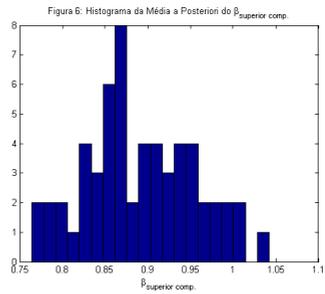
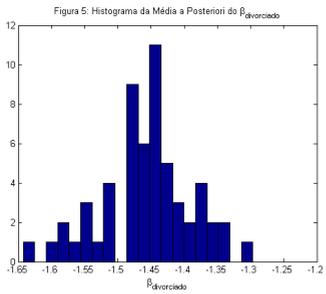
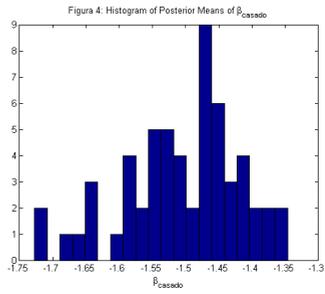
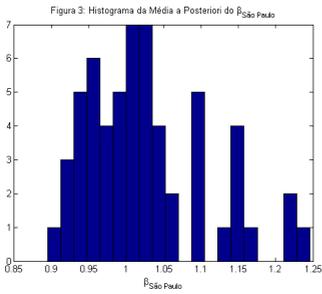
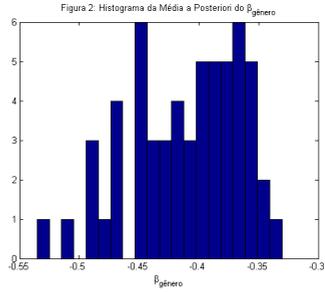
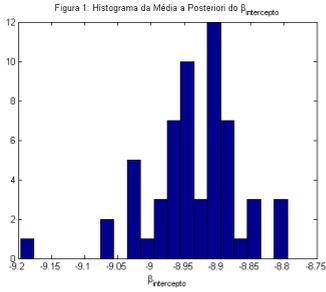
Fonte – Elaboração própria a partir dos resultados do modelo  $M_2$

Tabela 9 – Efeito de cada covariada específico a cada zona eleitoral no Modelo  $M_2$ 

Z.E.	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	$\beta_9$	$\beta_{10}$
1	-9,0164	-0,4514	1,1239	-1,5269	-1,3388	0,8462	0,9702	2,9892	0,1377	0,0464
2	-8,9704	-0,4857	1,1528	-1,5850	-1,3694	0,9506	0,9797	2,9499	0,1438	0,0450
3	-8,9088	-0,3481	1,0883	-1,4784	-1,4630	0,8671	0,9717	3,0016	0,1346	0,0430
4	-8,9283	-0,3724	1,0237	-1,5750	-1,4492	0,9766	1,1144	3,0641	0,1230	0,0450
5	-9,1966	-0,4901	1,2157	-1,7275	-1,2965	0,9549	0,9576	3,0049	0,2136	0,0349
6	-8,9756	-0,4227	1,1721	-1,6567	-1,3427	0,9115	1,0252	2,9684	0,1638	0,0428
20	-9,0162	-0,4432	1,0343	-1,4708	-1,4419	0,8615	1,0006	3,1924	0,1208	0,0494
246	-8,9801	-0,4521	1,1516	-1,6437	-1,3642	0,9811	1,0930	3,0418	0,1536	0,0432
247	-8,9378	-0,3831	0,8947	-1,3995	-1,4597	0,7882	0,8945	3,1021	0,1257	0,0523
248	-8,8905	-0,3588	0,9404	-1,4644	-1,4745	0,8690	1,0093	3,1216	0,1127	0,0508
249	-8,9023	-0,3856	0,9504	-1,5530	-1,4115	0,9892	1,0226	2,9722	0,1576	0,0438
250	-8,8994	-0,4654	1,0979	-1,5574	-1,4391	0,9094	1,0557	3,0179	0,1481	0,0447
251	-9,0217	-0,4859	1,2227	-1,7149	-1,3543	1,0424	0,9918	2,9406	0,1410	0,0447
252	-8,9489	-0,3642	0,9766	-1,5676	-1,4228	0,9252	1,0473	3,0631	0,1409	0,0461
253	-8,9463	-0,3943	1,0138	-1,5893	-1,4602	0,9545	1,1084	3,0925	0,1465	0,0448
254	-8,8984	-0,3587	1,0099	-1,4185	-1,4212	0,7685	0,9657	3,0084	0,1702	0,0417
255	-8,8955	-0,3672	0,9935	-1,5194	-1,4487	0,8299	1,0134	3,0508	0,1604	0,0453
256	-8,9034	-0,3889	0,9868	-1,4754	-1,4380	0,8729	0,9947	3,0683	0,1561	0,0443
257	-8,9403	-0,3791	0,9753	-1,5384	-1,4839	0,9387	1,1234	3,1164	0,1251	0,0462
258	-9,0736	-0,4459	1,1448	-1,6770	-1,3681	1,0027	1,0712	3,0232	0,1866	0,0382
259	-8,9568	-0,4506	1,1027	-1,6364	-1,3668	0,9944	1,0995	2,9991	0,1363	0,0463
260	-8,9414	-0,3680	1,0970	-1,4928	-1,4316	0,9018	1,0589	3,0332	0,1453	0,0439
280	-9,0108	-0,4068	1,0597	-1,5973	-1,4028	1,0013	1,0112	3,2201	0,1247	0,0493
320	-8,9234	-0,3800	1,0379	-1,5423	-1,4379	0,9395	1,0726	3,1120	0,1137	0,0487
325	-8,9529	-0,4333	1,0040	-1,5276	-1,3955	0,9026	0,9974	3,0738	0,1375	0,0484
326	-8,9741	-0,3995	0,9455	-1,4318	-1,4826	0,8330	0,9707	3,0542	0,1499	0,0489
327	-8,9071	-0,4294	1,0030	-1,5178	-1,4328	0,8864	1,0127	3,1082	0,1337	0,0484
328	-8,9680	-0,3747	1,0557	-1,5557	-1,4422	0,9577	1,0010	3,1513	0,1191	0,0484
346	-8,9587	-0,3596	1,2455	-1,6510	-1,3673	0,9605	1,0056	2,9409	0,1603	0,0430
347	-8,9149	-0,4056	0,9190	-1,4579	-1,4846	0,9054	1,0666	3,0417	0,1370	0,0490
348	-9,0185	-0,4675	0,9967	-1,5108	-1,4049	0,8597	1,1272	3,1434	0,1573	0,0452
349	-8,8404	-0,3544	0,9827	-1,4752	-1,4553	0,8357	1,0007	3,0620	0,1291	0,0473
350	-8,9802	-0,3970	0,9210	-1,5005	-1,5087	0,8653	1,0373	3,1662	0,1248	0,0503
351	-8,9426	-0,4164	1,0237	-1,5006	-1,4618	0,8936	1,0174	3,1403	0,1114	0,0497
352	-8,8872	-0,3947	1,0009	-1,3824	-1,5260	0,7644	0,9439	3,0877	0,1334	0,0488
353	-8,9450	-0,4806	0,9617	-1,3601	-1,5551	0,8261	0,9855	3,1382	0,1285	0,0496
371	-8,9755	-0,4432	1,0525	-1,4476	-1,4447	0,9055	0,9686	3,1785	0,1212	0,0497
372	-8,9285	-0,4416	1,0451	-1,4189	-1,5202	0,8643	1,0141	3,1730	0,1233	0,0478
373	-8,9503	-0,4031	1,0370	-1,4915	-1,4756	0,9413	1,0311	3,1363	0,1047	0,0500
374	-8,8945	-0,4227	1,0981	-1,5423	-1,4575	0,9239	1,0047	3,0386	0,1240	0,0488
375	-8,9122	-0,5354	0,9710	-1,3890	-1,6426	0,8593	1,0347	3,1394	0,1127	0,0515
376	-8,8576	-0,3780	1,0180	-1,4412	-1,4249	0,7883	0,9057	3,1173	0,1157	0,0510
381	-8,8043	-0,5050	1,0039	-1,3463	-1,5695	0,9385	0,9700	3,1531	0,0845	0,0515
389	-8,8998	-0,3846	1,0329	-1,4695	-1,5165	0,8320	0,8988	3,1174	0,1184	0,0514
390	-8,9475	-0,3868	0,9543	-1,4462	-1,4354	0,8668	1,0232	3,0998	0,1371	0,0477
392	-8,8782	-0,3611	0,9458	-1,4583	-1,4507	0,8602	1,0341	3,0511	0,1286	0,0504
397	-8,8932	-0,4285	0,9706	-1,4156	-1,5790	0,8593	0,9889	3,1308	0,1042	0,0522
398	-8,8544	-0,3305	0,9521	-1,4580	-1,5494	0,8056	0,9232	3,1359	0,1270	0,0494
403	-8,7950	-0,4363	0,9415	-1,4089	-1,5796	0,8100	0,9145	3,1805	0,1252	0,0491
404	-8,9094	-0,4659	0,9869	-1,3765	-1,5967	0,8691	1,0243	3,1767	0,1016	0,0517
405	-8,9115	-0,4165	0,9576	-1,4630	-1,4485	0,8819	1,0645	3,1235	0,1216	0,0512
408	-9,0331	-0,4299	1,1502	-1,5806	-1,3843	0,9031	1,0188	3,1364	0,1448	0,0470
413	-8,8948	-0,3635	1,0236	-1,5211	-1,4755	0,9714	1,0988	3,0694	0,1184	0,0478
417	-8,8469	-0,3653	0,9221	-1,4722	-1,5174	0,9196	1,0038	3,1089	0,1099	0,0504
418	-9,0563	-0,4668	1,0078	-1,5330	-1,4744	0,8669	0,9588	3,1881	0,1504	0,0475
420	-8,8839	-0,3467	0,9473	-1,4404	-1,4774	0,8354	1,0286	3,0736	0,1361	0,0464
421	-8,9680	-0,3975	0,9534	-1,4731	-1,5401	0,8589	1,0113	3,1573	0,1209	0,0510
422	-8,7945	-0,3690	1,0233	-1,4512	-1,4700	0,8037	0,9811	2,9622	0,1706	0,0449

Fonte – Elaboração própria a partir dos resultados do modelo  $M_2$

## Resultados a posteriori para os $\beta_i^t$ s do modelo $M_2$



[Cont.] Resultados a posteriori para os  $\beta'_i$ s do modelo  $M_2$

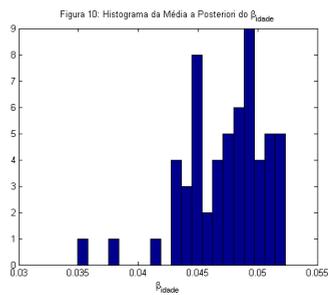
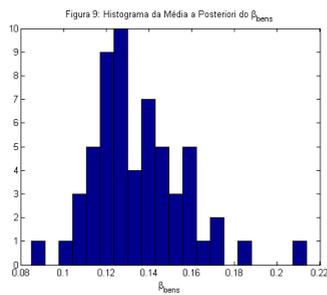
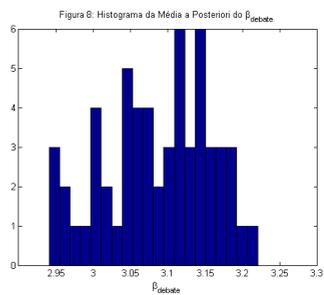
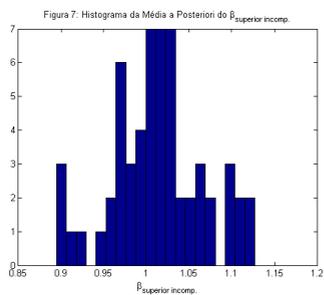


Tabela 10 – Resultados a Posteriori para o Modelo  $M_3$ 

Parâmetro	Média a Posteriori	DP	n.s.e
$\beta_0(\text{Intercepto})$	-8,8080	0,2583	0,0008
$\beta_0(\text{Gênero})$	-0,4274	0,1216	0,0004
$\beta_0(\text{S.Paulo})$	1,0142	0,1339	0,0004
$\beta_0(\text{Casado})$	-1,4859	0,1663	0,0005
$\beta_0(\text{Divorciado})$	-1,4495	0,1631	0,0005
$\beta_0(\text{Sup.Complete})$	0,8797	0,1559	0,0005
$\beta_0(\text{Sup.Incomplete})$	0,9868	0,1741	0,0006
$\beta_0(\text{Debate})$	3,0586	0,1196	0,0004
$\beta_0(\text{Bens})$	0,1366	0,0505	0,0002
$\beta_0(\text{Idade})$	0,0454	0,0452	0,0001
$\beta_0(dP)$	0,1899	0,1020	0,0003
$\beta_0(dT)$	2,1360	0,1952	0,0006
$\beta_0(dP \cdot dT)$	-0,9937	0,2449	0,0008
$h$	0,7095	0,0305	0,0001
$\Sigma_{1,1}$	0,5455	0,1853	0,0006
$\Sigma_{2,2}$	0,2764	0,0666	0,0002
$\Sigma_{3,3}$	0,3019	0,0759	0,0002
$\Sigma_{4,4}$	0,3453	0,0945	0,0003
$\Sigma_{5,5}$	0,3409	0,0919	0,0003
$\Sigma_{6,6}$	0,3352	0,0897	0,0003
$\Sigma_{7,7}$	0,3782	0,1063	0,0003
$\Sigma_{8,8}$	0,2824	0,0688	0,0002
$\Sigma_{9,9}$	0,1270	0,0246	0,0001
$\Sigma_{10,10}$	0,1164	0,0223	0,0001
$\Sigma_{11,11}$	0,2458	0,0562	0,0002
$\Sigma_{12,12}$	0,4065	0,1189	0,0004
$\Sigma_{13,13}$	0,4927	0,1589	0,0005

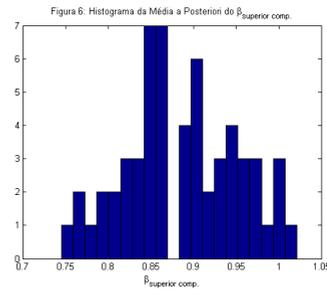
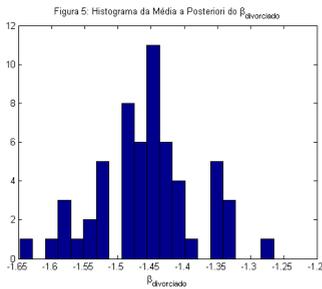
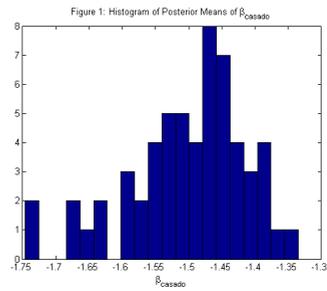
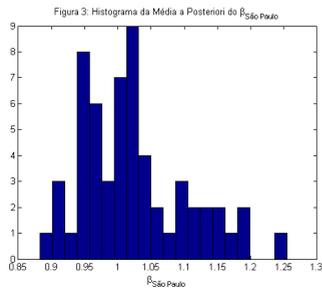
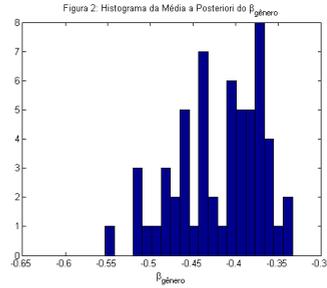
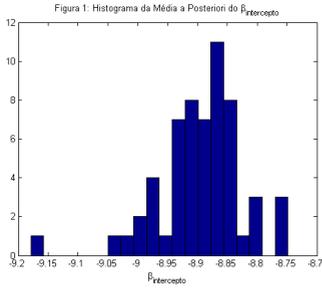
Fonte – Elaboração própria a partir dos resultados do modelo  $M_3$

Tabela 11 – Efeito de cada covariada específico a cada zona eleitoral no Modelo  $M_3$ 

Z.E.	$\beta_1$	$\beta_2$	$\beta_3$	$\beta_4$	$\beta_5$	$\beta_6$	$\beta_7$	$\beta_8$	$\beta_9$	$\beta_{10}$	$\beta_{11}$	$\beta_{12}$	$\beta_{13}$
1	-8.9791	-0.4660	1.1128	-1.5328	-1.3259	0.8301	0.9555	2.9697	0.1416	0.0460	0.1772	2.1997	-0.9172
2	-8.9330	-0.5073	1.1403	-1.5946	-1.3543	0.9356	0.9716	2.9351	0.1483	0.0443	0.2285	2.1967	-0.8883
3	-8.8733	-0.3593	1.0831	-1.4762	-1.4534	0.8515	0.9682	3.0023	0.1359	0.0419	0.2622	2.1484	-0.9912
4	-8.8829	-0.3755	1.0230	-1.5686	-1.4490	0.9777	1.1159	3.0667	0.1231	0.0439	0.2278	2.0764	-0.9815
5	-9.1792	-0.5172	1.1975	-1.7459	-1.2650	0.9294	0.9541	3.0066	0.2198	0.0339	0.2833	2.2068	-0.8782
6	-8.9361	-0.4379	1.1649	-1.6697	-1.3302	0.8942	1.0198	2.9610	0.1682	0.0418	0.2331	2.1469	-0.9028
20	-8.9807	-0.4536	1.0214	-1.4735	-1.4398	0.8507	0.9841	3.1777	0.1230	0.0495	0.1418	2.2520	-1.0095
246	-8.9369	-0.4641	1.1521	-1.6428	-1.3554	0.9786	1.0945	3.0387	0.1555	0.0422	0.2191	2.1189	-0.9428
247	-8.8994	-0.3889	0.8827	-1.3899	-1.4615	0.7834	0.8722	3.0925	0.1265	0.0522	0.1665	2.1839	-1.0421
248	-8.8483	-0.3617	0.9388	-1.4612	-1.4803	0.8642	0.9978	3.1225	0.1123	0.0502	0.1978	2.1548	-1.0584
249	-8.8533	-0.3875	0.9548	-1.5451	-1.4165	0.9978	1.0163	2.9612	0.1577	0.0429	0.1991	2.0586	-0.9512
250	-8.8531	-0.4202	1.1023	-1.5553	-1.4385	0.9005	1.0471	3.0145	0.1490	0.0439	0.2186	2.1179	-0.9654
251	-8.9896	-0.5157	1.1969	-1.7336	-1.3243	1.0212	0.9859	2.9302	0.1473	0.0436	0.2726	2.2526	-0.8553
252	-8.9045	-0.3695	0.9750	-1.5600	-1.4234	0.9265	1.0427	3.0605	0.1422	0.0452	0.2092	2.0978	-0.9999
253	-8.8991	-0.3996	1.0179	-1.5872	-1.4645	0.9546	1.1032	3.0899	0.1463	0.0441	0.1934	2.1010	-0.9892
254	-8.8534	-0.3595	1.0171	-1.4034	-1.4263	0.7640	0.9599	3.0116	0.1695	0.0409	0.2083	2.0536	-1.0252
255	-8.8591	-0.3710	0.9947	-1.5156	-1.4459	0.8164	1.0087	3.0599	0.1604	0.0444	0.2512	2.1058	-1.0195
256	-8.8669	-0.3968	0.9886	-1.4684	-1.4370	0.8693	0.9933	3.0799	0.1556	0.0434	0.2438	2.0816	-1.0298
257	-8.8949	-0.3808	0.9770	-1.5273	-1.4906	0.9474	1.1211	3.1074	0.1251	0.0457	0.1600	2.1146	-0.9981
258	-9.0446	-0.4579	1.1393	-1.6826	-1.3576	0.9938	1.0694	3.0205	0.1890	0.0374	0.2428	2.1229	-0.9216
259	-8.9142	-0.4646	1.0981	-1.6430	-1.3597	0.9885	1.0992	2.9978	0.1387	0.0454	0.2253	2.1182	-0.9173
260	-8.8937	-0.3683	1.1089	-1.4854	-1.4394	0.9030	1.0543	3.0276	0.1438	0.0433	0.1887	2.1040	-1.0145
280	-8.9725	-0.4135	1.0572	-1.5953	-1.4019	1.0052	1.0006	3.2214	0.1251	0.0488	0.1676	2.1566	-1.0302
320	-8.8780	-0.3852	1.0399	-1.5340	-1.4365	0.9388	1.0669	3.1122	0.1133	0.0479	0.2046	2.1116	-1.0225
325	-8.9041	-0.4364	1.0095	-1.5180	-1.4028	0.9072	0.9835	3.0596	0.1368	0.0481	0.1448	2.1241	-1.0250
326	-8.9294	-0.4072	0.9370	-1.4220	-1.4870	0.8291	0.9532	3.0331	0.1520	0.0487	0.1476	2.1805	-1.0260
327	-8.8598	-0.4347	1.0034	-1.5123	-1.4378	0.8871	1.0016	3.1019	0.1337	0.0480	0.1711	2.1246	-1.0163
328	-8.9288	-0.3795	1.0513	-1.5564	-1.4442	0.9551	0.9896	3.1435	0.1197	0.0481	0.1687	2.1850	-1.0204
346	-8.9154	-0.3668	1.2545	-1.6492	-1.3560	0.9563	0.9972	2.9318	0.1611	0.0421	0.2252	2.1233	-0.9571
347	-8.8663	-0.4074	0.9164	-1.4505	-1.4942	0.9092	1.0576	3.0298	0.1372	0.0485	0.1593	2.1144	-1.0217
348	-8.9762	-0.4741	0.9985	-1.5010	-1.4075	0.8606	1.1214	3.1337	0.1580	0.0449	0.1453	2.1120	-1.0025
349	-8.7963	-0.3531	0.9942	-1.4619	-1.4575	0.8275	0.9964	3.0755	0.1269	0.0563	0.2488	2.0540	-1.0479
350	-8.9434	-0.4014	0.9148	-1.4944	-1.5179	0.8673	1.0256	3.1543	0.1257	0.0403	0.1328	2.1684	-1.0334
351	-8.8955	-0.4224	1.0255	-1.4959	-1.4667	0.8904	1.0028	3.1353	0.1114	0.0494	0.1673	2.1422	-1.0314
352	-8.8545	-0.4000	0.9966	-1.3776	-1.5266	0.7457	0.9353	3.0934	0.1340	0.0481	0.2384	2.1969	-1.0800
353	-8.9099	-0.4946	0.9478	-1.3594	-1.5581	0.8140	0.9721	3.1303	0.1313	0.0493	0.1842	2.2425	-1.0534
371	-8.9350	-0.4549	1.0412	-1.4460	-1.4501	0.9030	0.9479	3.1541	0.1238	0.0499	0.1102	2.2572	-1.0252
372	-8.8926	-0.4548	1.0327	-1.4225	-1.5213	0.8540	1.0022	3.1597	0.1260	0.0477	0.1569	2.2522	-0.9951
373	-8.9024	-0.4087	1.0352	-1.4871	-1.4810	0.9431	1.0176	3.1145	0.1063	0.0498	0.1303	2.1970	-1.0105
374	-8.8568	-0.4329	1.0935	-1.5370	-1.4512	0.9179	0.9950	3.0407	0.1238	0.0482	0.2265	2.1432	-1.0028
375	-8.8767	-0.5537	0.9572	-1.3897	-1.6474	0.8451	1.0271	3.1318	0.1145	0.0513	0.1972	2.2466	-1.0607
376	-8.8172	-0.3833	1.0163	-1.4352	-1.4205	0.7701	0.8941	3.1220	0.1159	0.0503	0.2223	2.1658	-1.0554
381	-8.7520	-0.5187	1.0011	-1.3346	-1.5848	0.9440	0.9520	3.1342	0.0848	0.0514	0.1309	2.2214	-1.0579
389	-8.8624	-0.3943	1.0267	-1.4609	-1.5146	0.8233	0.8838	3.1170	0.1184	0.0509	0.2067	2.1977	-1.0709
390	-8.9087	-0.3913	0.9531	-1.4399	-1.4371	0.8612	1.0209	3.1050	0.1370	0.0470	0.2166	2.1213	-1.0350
392	-8.8349	-0.3601	0.9490	-1.4445	-1.4551	0.8617	1.0272	3.0475	0.1275	0.0498	0.1866	2.1084	-1.0465
397	-8.8566	-0.4410	0.9644	-1.4103	-1.5788	0.8496	0.9794	3.1338	0.1051	0.0515	0.2290	2.2018	-1.0825
398	-8.8117	-0.3336	0.9463	-1.4508	-1.5515	0.7924	0.9105	3.1454	0.1261	0.0486	0.2401	2.1585	-1.0752
403	-8.7555	-0.4418	0.9404	-1.3994	-1.5830	0.8022	0.9058	3.1872	0.1243	0.0484	0.2141	2.1572	-1.0851
404	-8.8734	-0.4846	0.9709	-1.3765	-1.5961	0.8551	1.0182	3.1712	0.1036	0.0513	0.2035	2.2683	-1.0431
405	-8.8629	-0.4217	0.9575	-1.4568	-1.4595	0.8847	1.0487	3.1046	0.1216	0.0512	0.1243	2.1773	-1.0314
408	-8.9860	-0.4381	1.1489	-1.5771	-1.3845	0.9028	0.9980	3.1114	0.1468	0.0472	0.0962	2.2004	-0.9784
413	-8.8415	-0.3682	1.0277	-1.5159	-1.4830	0.9732	1.0915	3.0646	0.1188	0.0470	0.1930	2.1081	-1.0117
417	-8.8046	-0.3671	0.9185	-1.4646	-1.5217	0.9190	0.9604	3.1082	0.1094	0.0498	0.1991	2.1417	-1.0476
418	-9.0236	-0.4797	0.9953	-1.5360	-1.4773	0.8578	0.9397	3.1792	0.1526	0.0476	0.1538	2.2209	-1.0241
420	-8.8362	-0.3412	0.9583	-1.4241	-1.4878	0.8396	1.0219	3.0748	0.1338	0.0457	0.1920	2.0635	-1.0456
421	-8.9339	-0.4049	0.9395	-1.4744	-1.5452	0.8533	0.9985	3.1471	0.1226	0.0509	0.1635	2.2317	-1.0346
422	-8.7485	-0.3688	1.0300	-1.4380	-1.4718	0.7974	0.9752	2.9612	0.1695	0.0440	0.2291	2.0730	-1.0253

Fonte – Elaboração própria a partir dos resultados do modelo  $M_3$

### Resultados a posteriori para os $\beta'_i$ s do modelo $M_3$



[Cont.] Resultados a posteriori para os  $\beta_i$ 's do modelo  $M_3$ 