

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

Desempenho educacional antes e após a eleição do ex-presidente Lula

Clinton Justino Mallet

Matrícula: 1212992

Orientador: Carlos Viana de Carvalho

Co-orientador: Eduardo Zilberman

Rio de Janeiro

Dezembro de 2016

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO RIO DE JANEIRO

DEPARTAMENTO DE ECONOMIA

MONOGRAFIA DE FINAL DE CURSO

Desempenho educacional antes e após a eleição do ex-presidente Lula

Clinton Justino Mallet

Matrícula: 1212992

Orientador: Carlos Viana de Carvalho

Co-orientador: Eduardo Zilberman

Declaro que o presente trabalho é de minha autoria e que não recorri para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor

---

Clinton Justino Mallet

Rio de Janeiro

Dezembro de 2016

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do  
autor

“Gratidão é o sentimento que mais depressa envelhece.”

Aristóteles

## Agradecimentos

Gostaria de agradecer, especialmente, as pessoas que foram fundamentais para que essa etapa da minha vida pudesse ser concluída.

Primeiramente, agradeço a Deus. Sem Ele nada é possível. Obrigado por todas as bênçãos e pelo maior presente que alguém pode ganhar, o dom da vida.

Ao meu orientador Carlos Viana de Carvalho, pelo aprendizado, pela atenção, pela paciência, pela motivação e por tratar as pessoas como um verdadeiro gentleman.

Ao meu co-orientador Eduardo Zilberman, por todo apoio dado durante a pesquisa.

A minha família, que me deu suporte durante todos esses anos, instigando-me a continuar os estudos. Em especial, devoto o meu maior agradecimento a minha vó, que é sinônimo de amor incondicional e sem fronteiras. O que ela fez e faz por mim não há valor que sequer possa ser comparado. Nada seria possível sem ela.

As secretárias do departamento, que sempre me trataram muito bem. Em especial, agradeço a Graça e a Roseli, pelos momentos divertidos e pelas conversas agradáveis.

A Bernardo Ribeiro, pelas sugestões e pela solidariedade. Suporte imprescindível para concluir a pesquisa.

A Pedro Pessoa e a Marina Dias, que me ajudaram a pensar na melhor forma de expor a pesquisa.

Ao Dr. Marcos, que foi imprescindível para eu pudesse concluir o curso. A minha gratidão será eterna.

A professora Helena, que me acompanhou durante esses anos e orientou-me em diversas questões.

A tia Leonôra, pelo carinho, amor e por todo ensinamento.

A Lucas Dourado, que me influenciou a fazer economia.

Aos meus amigos da PUC-RIO que, sem dúvida, eu quero carregar para o resto da vida. Além deles terem tornado os dias mais alegres, eles foram fiéis companheiros nos momentos mais difíceis.

# Sumário

<b>1</b>	<b>Introdução</b>	<b>8</b>
<b>2</b>	<b>Revisão da Literatura</b>	<b>10</b>
<b>3</b>	<b>Dados</b>	<b>14</b>
3.1	Tribunal Superior Eleitoral . . . . .	14
3.2	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep) . . . . .	15
3.2.1	Metodologia do cálculo das taxas de abandono . . . . .	16
3.2.1.1	Período 1: 1995-2006 . . . . .	16
3.2.1.2	Período 2: 2007-2015 . . . . .	17
<b>4</b>	<b>Estratégia empírica</b>	<b>20</b>
4.1	Especificação 1 . . . . .	20
4.2	Especificação 2 . . . . .	21
<b>5</b>	<b>Resultados</b>	<b>23</b>
5.1	Modelo 1: equações 4.1 e 4.2 . . . . .	23
5.2	Modelo 2: equações 4.3, 4.4 e 4.5 . . . . .	24
<b>6</b>	<b>Conclusão</b>	<b>26</b>
	<b>Referências Bibliográficas</b>	<b>28</b>
<b>A</b>	<b>Apêndice</b>	<b>29</b>

# Lista de Figuras

A.1	Diferença entre as duas taxas para o município de São Paulo . . . . .	31
A.2	Diferença entre as duas taxas para o município de Belo Horizonte . . .	31
A.3	Diferença entre as duas taxas para o município do Rio de Janeiro . . .	32
A.4	Diferença entre as duas taxas para o município de União da Serra . . .	32
A.5	Diferença entre as duas taxas para o município de Chapada de Areia . . .	33
A.6	Diferença entre as duas taxas para o município de Aroeira do Itaim . . .	33
A.7	Dispersão da variação do percentual de votos para seis municípios . . .	34

# Lista de Tabelas

A.1	Especificação 1: estimado por MQO . . . . .	29
A.2	Especificação 2: estimado por MQO . . . . .	30

# Capítulo 1

## Introdução

A literatura sobre role model é incipiente, entretanto há alguns autores como Chung (2000), que faz uma formalização matemática sobre o modelo. Ele sugere que argumentos de ação afirmativa pode aumentar o número de alunas negras bem-sucedidas, a fim de servirem como exemplo para outras mulheres estudantes negras. Já em Grossman (2015) é avaliado como o papel do gênero influencia o comportamento no uso da liderança em ações coletivas. Basicamente, estes modelos, denominados de role model, avaliam como que ações de agentes individuais podem influenciar outras pessoas.

A fim de explorar este universo de role model, esta pesquisa visa avaliar se o ex-presidente da república do Brasil, Luiz Inácio Lula da Silva, gerou alguma externalidade negativa sobre o número de alunos que abandonaram os estudos. Em diversos discursos pelo país e em entrevistas concedidas aos meios de comunicação, ele expôs, sem nenhuma vergonha, a sua trajetória e a sua baixa escolaridade e que isso o fez alcançar um posto de destaque, a presidência da república. Nesse sentido, a pesquisa busca identificar se há efeito causal entre o apoio/simpatia obtido, ao longo das cinco eleições presidenciais que concorreu, e a taxa de abandono, antes e depois que ele foi presidente. Esta será a nossa premissa. A ideia por de trás é que se ele conseguiu um emprego de prestígio e não precisou de muito estudo, talvez alguns jovens podem ter sido motivados a abandonar os estudos, ao considerar o seu exemplo e procurar seguir trajetória similar. O Lula passou de operário e sindicalista para presidente. Isso, sem dúvida, causa influência. Não estudar poderia ser interpretado por eles como um meio de ascender socialmente,

visto que o ex-presidente assim o fez.

A variável dependente a ser estimada é a variação da taxa de abandono. Ela foi construída com os dados do INEP. Montou-se um painel de dados, não balanceado, com as taxas de abandono entre 1995-2015. Posteriormente, tomou-se a variação entre dois anos para três períodos (1998-2002, 2002-2006 e 2006-2010). Os modelos foram controlados para a variação do percentual de votos entre duas eleições para três períodos (1994-1998, 1998-2002 e 2002-2006), para dummies dos percentis mais altas da distribuição da variação percentual de votos, de modo a refletir os municípios de intensa popularidade, e para a variação de renda, para os três período das eleições mencionados, medida pelo PIB a preços correntes das séries do IPEA DATA.

Os capítulos foram estruturados numa sequência, de modo a discorrer sobre a premissa inicial. No capítulo 2 é feita a revisão da literatura. Os principais estudos que serviram de base teórica para desenvolver a pesquisa. No capítulo 3 é feita uma descrição dos dados utilizados. Os dados principais são de votação, base do TSE e os censos escolares, base do INEP. No capítulo 4 é definida a estratégia empírica, a fim de manter ou refutar a tese. No capítulo 5 é apresentado os principais resultados deste estudo. Já o capítulo 6 apresenta as principais conclusões obtidas, de acordo com os modelos propostos.

## Capítulo 2

# Revisão da Literatura

No trabalho de Bettinger e Long (2005) procura-se avaliar se a presença de professores do mesmo sexo (masculino ou feminino) impacta o interesse dos alunos em uma disciplina. A teoria e a evidência sugerem que professoras podem ser instrumentos para encorajar mulheres a se inscrever e se sobressair em disciplinas que elas são sub-representadas. Estudantes do sexo feminino podem evitar áreas dominadas pelos homens devido a preconceitos que ocorrem. Assim, este paper responde a esse problema estimando como ter professoras em cursos iniciais afeta a probabilidade que as alunas pegarão créditos adicionais ou se formarão numa determinada disciplina.

De acordo com os padrões de inscrição de mulheres nas disciplinas, utilizando algumas estatísticas descritivas, fica claro, por exemplo, que 55% de aproximadamente 54000 estudantes, que pegaram aulas de inglês eram mulheres. Ainda em física, matemática, ciência da computação e business, entre 44%-51% também eram mulheres. Entretanto, dos alunos que eventualmente se formaram nessas áreas, as mulheres eram claramente sub-representadas. Por exemplo, na base de dados somente 14% que se formaram eram mulheres. Como é apresentado, menos de um terço dos professores de cursos introdutórios em história, economia, ciência política, química, física, matemática, business, ciência da computação e engenharia eram mulheres. A representação feminina entre professores de tempo integral foi ainda menor.

Para determinar o impacto de professoras em alunas, eles examinaram três resultados: se o estudante pegou disciplinas adicionais sobre o assunto, o número total de créditos subsequentes escolhido e a escolha do curso. As estimativas sugerem

que as professoras têm efeitos mistos sobre os interesses das alunas. Nas ciências, as alunas que inicialmente tinham professoras eram menos propensas a fazer cursos adicionais em biologia e física comparadas com aquelas que tiveram professores homens em seu primeiro curso. Por outro lado, as professoras afetaram positivamente a probabilidade de pegar um curso adicional e do número total de créditos subsequentes escolhido em geologia, matemática e estatística. Particularmente, na maioria dos cursos, alunas que inicialmente tinham professoras foram quase duas vezes mais propensas a pegar um curso adicional e, em média, pegaram 5,2 créditos a mais do que outras estudantes. Em termos da escolha do curso, os autores não encontraram efeitos positivos ou negativos.

Por sua vez, os estudantes homens que tiveram professores homens em cursos iniciais de educação pegaram 12,9 créditos a mais subsequentes e eram mais propensos a se formar na matéria. Esses resultados corroboram a ideia de que o mesmo sexo do professor pode afetar positivamente o interesse do estudante na disciplina.

Seguindo por esses modelos (role models), o estudo feito por Chung (2000) faz uma formalização matemática desse tipo de modelo e examina a validade dele para argumentos de ações afirmativas. É válido focar na intuição por de trás dele, uma vez que reproduzi-lo aqui não é o objetivo. Todavia, é importante perceber como ele estabelece essa conexão (modelo x argumentos de ações afirmativas). Primeiramente, o autor faz três definições, de acordo com Allen: (i) “modelos morais” que afeta as preferências de outras pessoas, talvez através de efeitos de conformidade; (ii) “modelos informacionais” que providencia informação sobre o valor presente de decisões atuais e (iii) “mentores” que representam recursos através dos quais o capital humano pode ser aumentado. Assim, ele visa tornar a segunda definição mais precisa com o modelo formal.

Assim, ele sugere que a ação afirmativa pode aumentar o número de alunas negras bem-sucedidas e elas podem servir como exemplo para outras mulheres estudantes negras. A existência destes efeitos particulares do modelo depende da incerteza do estado do mercado de trabalho, uma vez que a ação afirmativa é anunciada, esta incerteza desaparece. Essas alunas, no âmbito de ações afirmativas, não transmitem qualquer informação útil aos estudantes subsequentes.

Em Grossman (2015) é utilizado um laboratório experimental para investigar

as diferenças de gênero, masculino e feminino, no uso da liderança. Isto é, como o papel do gênero influencia o comportamento do líder e seguidores em ações coletivas. Deste modo, eles tentam explicar porquê as mulheres são pouco representadas em papéis de liderança. Para isso, o autor cria um jogo de uma única rodada em que se busca avaliar a ação coletiva em que a “carona” (free riding) e as falhas de coordenação podem evitar a eficiência da cooperação em grupo. Os líderes, por exemplo, são informados sobre o valor do projeto que têm em mãos, mas os seguidores não. Estes não são informados sobre o retorno dos projetos, porém são capazes de observar a decisão dos líderes de se investir neles.

O jogo que ele criou é de uma única rodada em que cada grupo há três jogadores iguais. Cada jogador  $i$  possui uma dotação de \$10 e tem que decidir se vai ou não participar (investir na sua dotação) de um projeto de investimento conjunto. Cada jogador  $i$  escolhe uma ação  $a_i \in A = \{N, P\}$ , onde  $N$  denota a não-participação e  $P$  a participação no projeto. Há três recompensas aleatórias que são equiprováveis: cenário 1 que tem o maior payoff, cenário 2 que tem payoff médio e o cenário 3 que tem menor payoff. Os payoffs apresentam duas características importantes. Em primeiro lugar, ninguém está disposto a participar sozinho, mas existem retornos crescentes de participação, no sentido de que qualquer ganho de participação do jogador  $i$  é uma função crescente da participação do outro. Este efeito de rede pode causar ineficiência através de falhas de coordenação. Em segundo lugar, nos cenários 1 e 2, não participantes beneficiam-se da participação dos outros. Isso tende a causar ineficiência através da fuga.

Este experimento criado com os alunos de Saint Cloud State University considerou dois gêneros únicos de tratamento em que todos os líderes e seguidores são ou homens ou mulheres. O autor considerou também dois gêneros mistos de tratamento, um sem sinalização do sexo do líder e outro com a sinalização explícita do sexo do líder.

Os resultados dentro do experimento mostram mais cooperação de seguidoras mulheres comparado com homens no gênero único de tratamento e no gênero misto de tratamento sem sinalização do sexo do líder. No mesmo tipo de grupo, mas que há a sinalização, as seguidoras mulheres não se comportam diferentemente dos homens.

Além disso, os resultados sugerem que as políticas destinadas a aumentar as oportunidades para as mulheres nos níveis superiores de organizações podem não ter os resultados desejados. Só ter oportunidades iguais para papéis de liderança pode não levar a um maior equilíbrio de gênero em cargos de liderança. Os resultados sugerem que os líderes do sexo feminino interpretaram mal as intenções dos seus seguidores; os seguidores não apresentam um comportamento consistente com a hipótese de que as mulheres são vistas como líderes menos legítimas. Esta interpretação errada pode ser devido à falta de experiência em cargos de liderança, o que sugere um problema da galinha e do ovo. Mais experiência por parte de líderes do sexo feminino pode servir para atenuar a relutância dos líderes do sexo feminino, mas essa relutância pode impedir as mulheres de receber experiência de liderança. Isto sugere a necessidade de uma orientação ou preparação de programas que aumentem a exposição das mulheres a cargos de liderança e dar-lhes tempo para adquirir experiência.

Com base nos artigos apresentados, esta pesquisa procurará verificar qual impacto que uma personalidade, no caso o ex-presidente Luiz Inácio Lula da Silva, o Lula, influencia a taxa de abandono escolar. O fato do ex-presidente ter o ensino fundamental incompleto e ter ascendido à presidência da República serviu de parâmetro para que os jovens, principalmente, das regiões mais pobres, abandonassem a escola? Tendo ocorrido uma identificação com o presidente, numa espécie de líder e seguidor, esse estudante sentiu-se desmotivado em continuar os estudos?

Desta forma, estas referências serão de grande valor para orientar o início da pesquisa e pretende-se contribuir com essa literatura com resultados ainda não estudados.

# Capítulo 3

## Dados

### 3.1 Tribunal Superior Eleitoral

Os dados de votação foram obtidos no Tribunal Superior Eleitoral (TSE). A partir dos microdados foi possível criar um painel de dados, não balanceado, para dados de votação com frequência quadrienal, só considerando os votos para presidente em todo o país. Os dados são apresentados para nível municipal, para as quatro eleições que o ex-presidente Lula concorreu (1994, 1998, 2002 e 2006). Embora 1989 ele também tenha concorrido, só há dados disponíveis para nível estadual. Conseqüentemente, isso impossibilitaria a análise, uma vez que se busca o nível mais granular possível, de modo, posteriormente, a fazer um cruzamento, a nível municipal, com os dados de educação e de renda.

Primeiramente, foi calculado o total de votos para cada município em cada eleição. Depois, o total de votos para cada candidato, que concorreu as eleições, em cada município nas quatro eleições. Posteriormente, excluiu-se todos os outros candidatos e só ficou mantido o ex-presidente Lula. Esses dados permitem analisar o padrão de votos que o Lula obteve em todas as eleições que concorreu. Previamente, já é sabido que ele aumentou o seu número de votos, desde a primeira eleição que concorreu (1989), ainda que avaliado em nível mais agregado, estadual. Isso é evidente, uma vez que foi eleito em 2002.

Os dados foram separados em dois turnos a fim de que se possa analisá-los com maior precisão. Para identificar os locais onde o Lula obteve mais votos foi construído algumas variáveis de apoio/simpatia: (1) percentual de votos para cada

município nas quatro eleições; (2) a média de votos para cada município; (3) a diferença percentual entre a primeira e a quarta eleição; (4) a diferença percentual entre a segunda e a quarta eleição; (5) ranks para todas as variáveis anteriores. Convém justificar a construção da penúltima variável construída. Os dados de 1994 estão incompletos. Alguns estados não estão na base, pois requer que os tribunais eleitorais regionais reportem os dados para o TSE e que, conseqüentemente, tal órgão possa validá-los. Este procedimento ainda está em curso e não foi concluído em tempo hábil de ser incluído nesta pesquisa.

Uma vez tendo criado tais variáveis, é possível corroborar o fato do Lula ter aumentado os votos ao longo do tempo, mas também identificar onde ele teve mais votos. Por conseguinte, são calculados os percentis da variação do percentual de votos, a fim de saber como se dá a distribuição de votos, sobretudo onde ele é mais preferido.

Como já foi evocado anteriormente, com essa base procura-se identificar municípios pró-Lula, i.e., aqueles que aumentaram ou mantiveram constante o seu percentual de votos para o ex-presidente. Desta forma, avaliar-se-á se nesses municípios a taxa de abandono aumentou, quando jovens abandonaram os estudos, pois tiveram uma afinidade com a história do ex-presidente, uma vez que este ascendeu à presidência com pouca escolaridade. Por conseguinte, aqueles poderiam querer replicar tal ação em suas vidas, ou seja, por qualquer que seja a razão, não estudar ou estudar pouco poderia leva-los a níveis similares, tanto de renda quanto de prestígio obtidos pelo ex-presidente.

## **3.2 Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep)**

Os dados de educação foram obtidos do Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Inep). A partir dos microdados foi possível criar um painel de dados, não balanceado, para dados de educação com frequência anual dos censos escolares de 1995 a 2015 a nível municipal. Foram excluídos os dados de educação privada, indígena, profissional, infantil e alunos da pré-escola e primeira série do ensino fundamental, por acreditar que eles não influenciam a análise. Além

disso, foi excluído também os dados referentes ao terceiro ano do ensino médio, porque não é possível verificar se o aluno se formou ou se abandonou, visto que essa informação não está disponível.

O painel é constituído de dados do ensino fundamental e médio para escolas do ensino público: federal, estadual e municipal. Desta forma, será possível calcular uma taxa de abandono a nível municipal.

Assim, a base de educação será unida a base de votação pelo código do município. Como anteriormente mencionado, as variáveis construídas com a base de votação por município são uma proxy de apoio/simpatia para determinar se, de fato, o ex-presidente teve alguma influência sobre a taxa de abandono, controlando também para efeito fixo de tempo e de município, de modo a verificar se há causalidade, questão fundamental para este estudo.

### **3.2.1 Metodologia do cálculo das taxas de abandono**

Para que houvesse uma metodologia de cálculo que fosse transversal para todos os anos, foi-se adotado alguns critérios, de modo a permitir calcular a taxa de abandono por município. Dado que o INEP divulgou os dados de educação para alguns anos de formas diferentes, é conveniente separar em dois períodos para se analisar os cálculos com mais detalhes. Portanto, a abordagem também deve refletir isso. Antes de adentrar na metodologia, cabe o último destaque. Os dados para os anos de 2000, 2001 e 2002 seguem o mesmo cálculo para o período 2, uma vez que eles foram divulgados de modo similar.

#### **3.2.1.1 Período 1: 1995-2006**

O INEP divulgou os dados do censo escolar, para este período, agregado por escola, como o total de alunos por série para cada escola por município. De igual forma, ele dispôs o total de alunos que abandonaram. Então, o total de alunos para cada escola é o somatório do número de alunos por série. Consequentemente, o total de alunos por município é somatório de alunos por escola. Logo, este é o total de alunos por município. Analogamente, foi calculado o total de alunos que abandonaram por município. Assim, a taxa de abandono por município é a razão do total de alunos que abandonaram pelo total de alunos. Matematicamente, tem-se o seguinte:

$$\text{tot\_alunos\_munic}_j = \sum_i \sum_s \text{tot\_alunos\_esc}_{s,i}$$

$$\text{tot\_alunos\_aban\_munic}_j = \sum_i \sum_s \text{tot\_alunos\_aban\_esc}_{s,i}$$

$$\text{Taxa de abandono}_j = \frac{\text{tot\_alunos\_aban\_munic}_j}{\text{tot\_alunos\_munic}_j}$$

Onde:

$j$  = Município

$i$  = Escola

$s$  = Série

tot\_alunos\_munic = Total de alunos por município

tot\_alunos\_esc = Total de alunos por escola

tot\_alunos\_aban\_munic = Total de alunos que abandonaram por município

tot\_alunos\_aban\_esc = Total de alunos que abandonaram por escola

### 3.2.1.2 Período 2: 2007-2015

Primeiramente, é preciso dispor a estrutura de dados para os anos 2000, 2001 e 2002, uma vez que o INEP não reportou o número de alunos que abandonaram por série na mesma base do censo escolar, mas a taxa de abandono por série por escola noutra, que é a base de indicadores da educação básica por escola. Para calcular o número de alunos por escola foi utilizado o censo. Enquanto, que as taxas de abandono foram os indicadores de rendimento escolar.

Segundamente, os dados para este subperíodo foram obtidos do censo escolar por matrícula, que a partir de 2007 consegue acompanhar o aluno ao longo do tempo. Portanto, foi calculado o total de alunos por série por escola utilizando o código de identificador do INEP. Desta forma, analogamente ao que foi feito no período 1, tem-se o total de alunos por município. Por sua vez, as taxas de abandono foram obtidas da base de taxas de rendimento.

Assim, para os três anos e para este subperíodo, tem-se o seguinte: (1) número de alunos por série por escola; (2) taxas de abandono por série por escola. A fim de compatibilizar os cálculos com o período 1, as taxas de abandono por município é o produto da taxa de abandono por série vezes o número de alunos por série, que

resulta no número de alunos que abandonaram por escola. Depois este resultado é agregado por município. E por fim, tem-se a taxa de abandono por município, que é a razão entre o total de alunos que abandonaram pelo total de alunos. Matematicamente, tem-se o seguinte:

1. *Base de dados do censo escolar*

Total de alunos por município:

$$\text{tot\_alunos\_munic}_j = \sum_i \sum_s \text{tot\_alunos\_esc}_{s,i}$$

Onde:

$j$  = Município

$i$  = Escola

$s$  = Série

tot\_alunos\_munic = Total de alunos por município

tot\_alunos\_esc = Total de alunos por escola

2. *Base de dados dos indicadores da educação básica por escola e base de dados das taxas de rendimento (2007-2015)*

Total de alunos que abandonaram por município:

$$\text{tot\_alunos\_aban\_munic}_j = \sum_i \prod_s \text{tx\_aban\_esc}_{s,i} \cdot \text{tot\_alunos\_esc}_{s,i}$$

Onde:

$j$  = Município

$i$  = Escola

$s$  = Série

tot\_alunos\_aban\_munic = Total de alunos que abandonaram por município

tx\_aban\_esc = Taxa de abandono por série

tot\_alunos\_esc = Total de alunos por escola

3. Taxa de abandono por município:

$$\text{Taxa de abandono}_j = \frac{\text{tot\_alunos\_aban\_munic}_j}{\text{tot\_alunos\_munic}_j}$$

Evidentemente, calcular as taxas de abandono por município utilizando o código de identificador do aluno, o que permite saber se no ano seguinte ele se encontra em alguma outra escola ou noutro município, seria mais fácil e mais preciso. Todavia, esse exercício não é replicável para o período 1, pois essa estrutura de dados só passou a ser disponível a partir de 2007. Portanto, para não se por em xeque a validade da pesquisa, em primeira instância, adotou-se uma metodologia capaz de ser transversal para todos os anos. Nos apêndices há alguns gráficos que permitem verificar a diferença entre as duas taxas calculadas: (1) a que utiliza o identificador do aluno; (2) a que não utiliza o identificador, que está sendo usada para este estudo. É fundamental registrar que o código de identificação do aluno utilizado para o período 2 é a melhor métrica para calcular o total de alunos por escola. Além disso, dado que o INEP reporta os dados no nível do aluno, não parece convincente que ele regrida a forma como dispõe os dados, voltar para um modelo com menos informações. No mais, esse dado também não está disponível, a menos que fosse compilada a nível mais granular, por exemplo, municipal. Este meio fatalmente seria pouco eficiente e bem mais exaustivo que o adotado. Embora isso não seja garantia de qualidade da série, em linhas gerais, este foi o melhor método disponível encontrado que permitiu compatibilizar os cálculos ao longo dos anos. De antemão como verificado pelos gráficos no apêndice, para três municípios grandes e três pequenos, há uma diferença considerável entre a taxa de abandono que acompanha o aluno da que não acompanha, a qual foi utilizada.

# Capítulo 4

## Estratégia empírica

Para estimar o efeito da variação da taxa de abandono a nível municipal foram utilizados dois modelos a fim de traçar uma estratégia que pudesse avaliar o efeito da variação do percentual de votos sobre a variação da taxa de abandono. Em termos gerais, procura-se verificar se a variável explicativa principal causa a variável dependente. Para os dois modelos utilizou-se efeito fixo de tempo e efeito fixo de município, a fim de medir se fatores que não mudam entre os municípios, mas são mantidos ao longo do tempo ou se fatores que mudam entre os municípios, mas são constantes no tempo se impactam o variação da taxa de abandono, uma vez que se tem características regionais bastante distintas e dado as inúmeras transformações socioeconômicas vividas pelo país.

### 4.1 Especificação 1

Sobre esta especificação foi montada uma estrutura em painel com a variação da taxa de abandono entre dois anos para três períodos (1998-2002, 2002-2006 e 2006-2010), controlando para a variação do percentual de votos que o Lula obteve entre duas eleições para três períodos (1994-1998, 1998-2002 e 2002-2006), para um efeito fixo de tempo e um efeito fixo de município. O modelo permite incluir regressores adicionais para nível do município, como a diferença de renda, captada pela variação do PIB entre dois anos, também para três períodos.

Sobre esta estratégia de especificação, temos o seguinte modelo:

$$\Delta Y_{it} = \beta_0 + \lambda_t + \eta_i + \beta_1 Z_{it} + \varepsilon_{it} \quad (4.1)$$

$$\Delta Y_{it} = \beta_0 + \lambda_t + \eta_i + \beta_1 Z_{it} + \beta_2 R_{it} + \varepsilon_{it} \quad (4.2)$$

onde  $\Delta Y_{it}$  é a variação da taxa de abandono entre dois anos para o município  $i$  no tempo  $t$ ,  $\lambda_t$  é uma dummy que captura características específicas que não variam entre os municípios, mas sim ao longo do tempo,  $\eta_i$  é uma dummy que captura características específicas que variam por município, mas não ao longo do tempo,  $Z_{it}$  é uma proxy de apoio/simpatia que o município tem pelo ex-presidente Lula, capturada pela variação do percentual de votos que ele obteve entre duas eleições e  $\varepsilon_{it}$  é um termo de erro estocástico.  $\beta_1$  nas equações 4.1 e 4.2 é o coeficiente de interesse e mede o quanto de apoio/simpatia dado ao ex-presidente pelo município  $i$  influenciou a taxa de abandono.

## 4.2 Especificação 2

Dado o modelo de especificação proposto na equação 4.1, há um conjunto de fatores que provavelmente afetam a habilidade em interpretar os possíveis efeitos causais do apoio dado ao Lula sobre a taxa de abandono. Em virtude disto, procurou-se incluir mais controles como a variação da renda municipal, usando o PIB municipal a preços correntes, ao longo do tempo, conforme descrito pela equação 4.2. Além disso, a fim de refinar a especificação proposta foi incluído três dummies separadamente para os percentis da variação do percentual de votos. Quando a dummy for igual a 1, são regiões com maior afinidade pelo ex-presidente, i.e., locais que ele tem intensa popularidade, portanto se encontram na cauda superior da distribuição e 0 para caso contrário. Assim, o modelo foi reescrito para captar os percentis (50%, 75% e 90%) da variação do percentual de votos. Isto implica dizer que quando a dummy for igual a 1, o município teve uma variação percentual de votos acima da mediana da distribuição. Embora isso não determine se ele teve a maioria de votos, porque o que conta é o número absoluto de votos válidos conquistados, fator crucial para dizer se um candidato ganhou num determinado município, é sintomático a adesão à sua pessoa. Tomando as dummies para os percentis mais altos revela os municípios

que ele obteve alto ganho de popularidade. Logo, estimou-se as seguintes equações:

$$\Delta Y_{it} = \beta_0 + \lambda_t + \eta_i + \phi_1 D1_{it} + \beta_2 R_{it} + \varepsilon_{it} \quad (4.3)$$

$$\Delta Y_{it} = \beta_0 + \lambda_t + \eta_i + \phi_2 D2_{it} + \beta_2 R_{it} + \varepsilon_{it} \quad (4.4)$$

$$\Delta Y_{it} = \beta_0 + \lambda_t + \eta_i + \phi_3 D3_{it} + \beta_2 R_{it} + \varepsilon_{it} \quad (4.5)$$

onde os efeitos fixos de tempo e município continuam mantidos,  $D1_{it}$  é uma dummy para o 50º percentil, que é igual a 1 para os municípios em que o Lula esteve acima da mediana da distribuição da variação do percentual de votos e 0 caso contrário. Analogamente,  $D2_{it}$  é uma dummy para o 75º percentil, ou seja, o terceiro quartil e  $D3_{it}$  é uma dummy para o 90º percentil, ou seja, o nono decil da distribuição.  $R_{it}$  é um vetor para a variação da renda municipal para quatro períodos (1994-1998, 1998-2002, 2002-2006 e 2002-2010). Esta última variável foi incluída, porque se acredita que indivíduos com menor poder aquisitivo estejam mais propensos a abandonar os estudos, a fim de seguir o exemplo do ex-presidente Lula.

Poder-se-ia perguntar qual é o propósito de incluir dummies, que é uma variável discreta, ao invés da variável contínua, já em mãos. O pressuposto inicial é que regiões de maior aderência, o efeito sobre a variação da taxa de abandono tende a ser potencializado. Parece razoável pensar que quando pessoas têm maior afinidade pelo líder, as ações deste exercem maior influência sobre seus liderados. Nesse sentido, em linha com o exercício que é proposto com este estudo.

# Capítulo 5

## Resultados

Neste capítulo é apresentado os principais resultados empíricos da pesquisa. Na seção 1 é feita a análise da estimativa da relação entre a variação da taxa de abandono e a variação do percentual de votos. Posteriormente, é controlado para renda também, de acordo com a equação 4.2. Já a seção 2 estima a relação entre a variação da taxa de abandono e as dummies junto com a variação da renda. Os principais resultados se encontram nas duas tabelas do apêndice.

### 5.1 Modelo 1: equações 4.1 e 4.2

Sobre esta estratégia de especificação, controlando para a variação do percentual de votos e para efeito fixo de tempo e de município, disposta no primeiro resultado da tabela *A.1*, é possível perceber que variação da taxa de abandono por município caiu. Embora o coeficiente da variação do percentual de votos não seja significativo para os níveis de confiança usuais (0,1%, 1% e 5%) ele exerce, ainda que pequeno, efeito negativo sobre a variável que se busca explicar. Um aumento de 1% da variação do percentual de votos diminui a variação da taxa de abandono em 0,0296%.

Com os resultados obtidos com a regressões feitas por MQO para a equação 4.1, esta especificação não explica porquê a variação da taxa de abandono caiu quando o ex-presidente ascendeu à presidência, de modo a orientar, ainda que via indiretamente, mais estudante a abandonarem os estudos. A literatura sobre role model é incipiente, entretanto seus pressupostos básicos estabelecem o quanto que o exemplo pessoal ou profissional de um agente consegue exercer influência sobre as

ações de outros indivíduos. Como ficou claro, a estratégia não permitiu verificar isso. Há alguns fatores que podem explicar porquê isso ocorreu. Nesse sentido, é intuitivo pensar em colocar mais controles, a fim que a principal variável independente possa causar a variável de interesse.

Indubitavelmente, é preciso considerar que as taxas de abandono já vinham caindo ao longo do tempo. Em parte, convém notar que o montante de recursos disponíveis para educação aumentou. Os dados apontam que depois que o ex-presidente Lula assumiu a presidência as taxas continuaram declinantes. Provavelmente, os programas de transferência de renda, o bolsa família, por exemplo, sobrepujam o efeito da influência que o Lula teve sobre a variável que se busca explicar. Antes dele assumir a presidência poder-se-ia pensar que a taxa de abandono já vinham numa trajetória de queda, por conta de políticas públicas destinadas a resolver essa questão nos três níveis: federal, estadual e municipal. Em linhas gerais, quando há um maior provisionamento de bem público, como mais educação para as pessoas, através de maior acessibilidade ao ensino, por exemplo, isso contribui para diminuir a taxa de abandono. Ou talvez se pode ir além. Uma quebra de paradigma, em que as pessoas passam a endogeneizar a ideia de que mais educação seja um fator crucial para aumentar sua renda no futuro e, portanto, abandonar os estudos não seja a melhor alternativa.

## 5.2 Modelo 2: equações 4.3, 4.4 e 4.5

Sobre esta especificação rodou-se algumas regressões, de modo a controlar para as dummies dos percentis (50%, 75% e 90%) da variação do percentual de votos e a variação da renda a nível municipal. Ao fazer o exercício, rodou-se as regressões para cada dummy e a variação da renda. Não faz mais sentido incluir neste bloco a variação do percentual de votos. Ainda que sejam tipos diferentes, uma é discreta e a outra é contínua, elas apontam para a mesma direção, porém de uma forma um pouco mais sofisticada, uma vez que as dummies conseguem focar nos locais onde o Lula é muito querido. Para todas as três regressões a variação da renda afeta positivamente a variação da taxa de abandono, porém os valores dos coeficientes não são estatisticamente significativos, o que faz com que estes resultados sejam

inconclusivos.

Para a dummy do percentil 50%, isto é, os municípios em que o Lula obteve mais do que a mediana da variação do percentual de votos, o aumento de 1% acima da mediana da distribuição faz com que a taxa de abandono caia em 0,147%. O coeficiente é estatisticamente significativo ao nível de 1% e revela que a variação da taxa de abandono caiu nos locais em que ele foi mais querido. Diante disto, a primeira pergunta que poderia ser feita é: mas se ele está numa região em que ele é mais querido, por que a taxa de abandono, mantido tudo mais constante, caiu e não o contrário? O racional é análogo ao que foi mencionado para o modelo 1. Outros fatores podem estar explicar melhor porquê a taxa de abandono caiu. Além disso, pode-se considerar, por exemplo, que a medida de popularidade pode não sobrepujar o efeito de políticas públicas que vinham ocorrendo ao longo do tempo, a fim de corrigir as distorções da taxa de abandono.

Para as duas regressões que controla para os percentis mais altos (75% e 90%), localizados na cauda superior da distribuição, o efeito se reverte. A variação na taxa de abandono aumenta quando ele aumenta em 1% acima dos percentis correspondentes. Para o percentil do terceiro quartil da distribuição o efeito é de um aumento da variação da taxa de abandono em 0.006%. Já para a dummy do nono decil o efeito é de um aumento de 0,137%. Embora para percentis elevados a variação da taxa de abandono aumente, isso não permite inferir que locais onde o Lula é muito querido a variação da taxa de abandono aumentou, uma vez que os coeficientes não são significativos aos níveis usuais. Logo, é inconclusivo.

# Capítulo 6

## Conclusão

Existem trabalhos empíricos que são focados em analisar os determinantes da taxa de abandono. Comumente, evoca-se os principais como renda, características familiares ou pessoais, questões demográficas, em que algumas são capturas pelos dados da PNAD, que visa produzir informações básicas para o estudo do desenvolvimento socioeconômico do país e permitir a investigação contínua de indicadores sobre trabalho e rendimento. Por outro lado, há outros que buscam compreender como fatores não usuais podem influenciar a taxa de abandono. Desta forma, este estudo lançou-se a verificar se a influência de uma pessoa, no caso o ex-presidente da república, com o seu exemplo de vida, em que muitas vezes externou a baixa escolaridade que tinha e o posto que alcançou, pode alterar as ações dos agentes, numa espécie de líder-seguidor.

Utilizando dois modelos de especificação foi possível encontrar resultados convergentes. Para o primeiro que só controla para a variação do percentual de votos verificou-se um efeito negativo de 0,0296% sobre a variação da taxa de abandono. Já controlando para a variação da renda, o coeficiente da variação do percentual de votos exerce efeito negativo de 0,0312%. Todavia, para o modelo 1 os resultados são inconclusivos, visto que os coeficiente não são estatisticamente significativos aos níveis usuais. Para o modelo 2 controlando para as dummies dos percentis mais altos da distribuição da variação do percentual de votos, somente a dummy para o percentil da mediana é estatisticamente significativo ao nível de significância de 1%. As possíveis razões disso já foram supracitadas no capítulo passado e no início deste, porém de modo mais amplo outros fatores deveriam ser controlados como,

por exemplo, as transferências do governo federal que são destinadas a educação. Para as dummies dos percentis 75% e 90%, o aumento de 1% acima da distribuição faz com que as variações da taxa de abandono aumentem 0,006% e 0,137%, respectivamente.

Por último, é importante compreender porquê os resultados frustraram a premissa inicial. Seria imprudente maltratar os dados, a qualquer custo, a fim de atingir os resultados pretendidos. Fica claro que outras variáveis poderiam ser incluídas nas regressões de modo a avaliar melhor o efeito da variação do percentual de votos que o Lula obteve sobre a variação da taxa de abandono. Além disso, a principal variável de interesse pode conter erros de medida, conforme pode se verificar nos gráficos contidos no apêndice, cuja discrepância entre a taxa de abandono por município que acompanha o aluno da que não acompanha, a qual foi utilizada, é grande, ainda que aparentem uma tendência similar entre as duas.

# Referências Bibliográficas

- [1] BETTINGER, ERIC P.; LONG, BRIDGET TERRY. Do Faculty Serve as Role Models? The Impact of Instructor Gender on Female Students, *American Economic Review*, Vol. 95(2), p. 152-157, 2005.
- [2] DAVID NEUMARK; ROSELLA GARDECKI. Women Helping Women? Role-Model and Mentoring Effects on Female Ph.D. Student in Economics, *National Bureau of Economic Research, Inc, NBER Working Papers*, Vol. 5733, 1996.
- [3] KIM-SAU CHUNG. Role Models and Arguments for Affirmative Action, *American Economic Review*, Vol. 90(3), p. 640-648, 2000.
- [4] NGUYEN, TRANG. Information, role models and perceived returns to education: Experimental evidence from Madagascar, *Unpublished manuscript*, Vol. 6, 2008.
- [5] PHILIP J. GROSSMAN; MANA KOMAI; JAMES E. JENSEN. Leadership and Gender in Groups: An Experiment, *Monash University, Department of Economics, Monash Economics Working Papers*, p. 42-12, 2012.

# Apêndice A

## Apêndice

Tabela A.1: Especificação 1: estimado por MQO

	(1)	(2)
	$\Delta\_tx\_abandono$	$\Delta\_tx\_abandono$
$\Delta\_percentual\_votos$	-0.0296 (-0.88)	-0.0312 (-0.93)
$\Delta\_pib$		0.102 (1.12)
rep dummies	Yes	Yes
$N$	10947	10947
$R^2$	0.051	0.052

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Tabela A.2: Especificação 2: estimado por MQO

	(1)	(2)	(3)
	$\Delta\_tx\_abandono$	$\Delta\_tx\_abandono$	$\Delta\_tx\_abandono$
$\Delta\_pib$	0.0793 (1.48)	0.0740 (1.38)	0.0734 (1.37)
d50_ $\Delta\_percentual$	-0.147** (-2.64)		
d75_ $\Delta\_percentual$		0.00623 (0.11)	
d90_ $\Delta\_percentual$			0.137 (1.78)
rep dummies	Yes	Yes	Yes
$N$	16438	16438	16438
$R^2$	0.045	0.044	0.045

*t* statistics in parentheses

\*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

Figura A.1: Diferença entre as duas taxas para o município de São Paulo

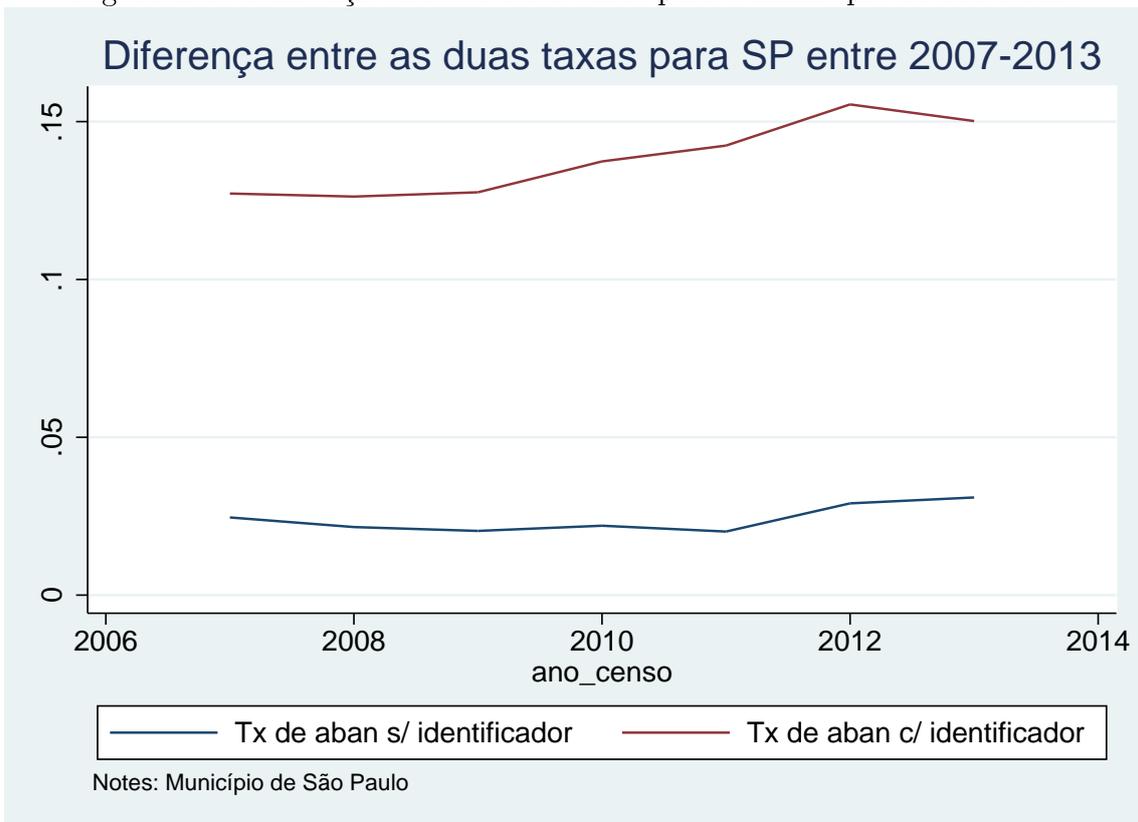


Figura A.2: Diferença entre as duas taxas para o município de Belo Horizonte

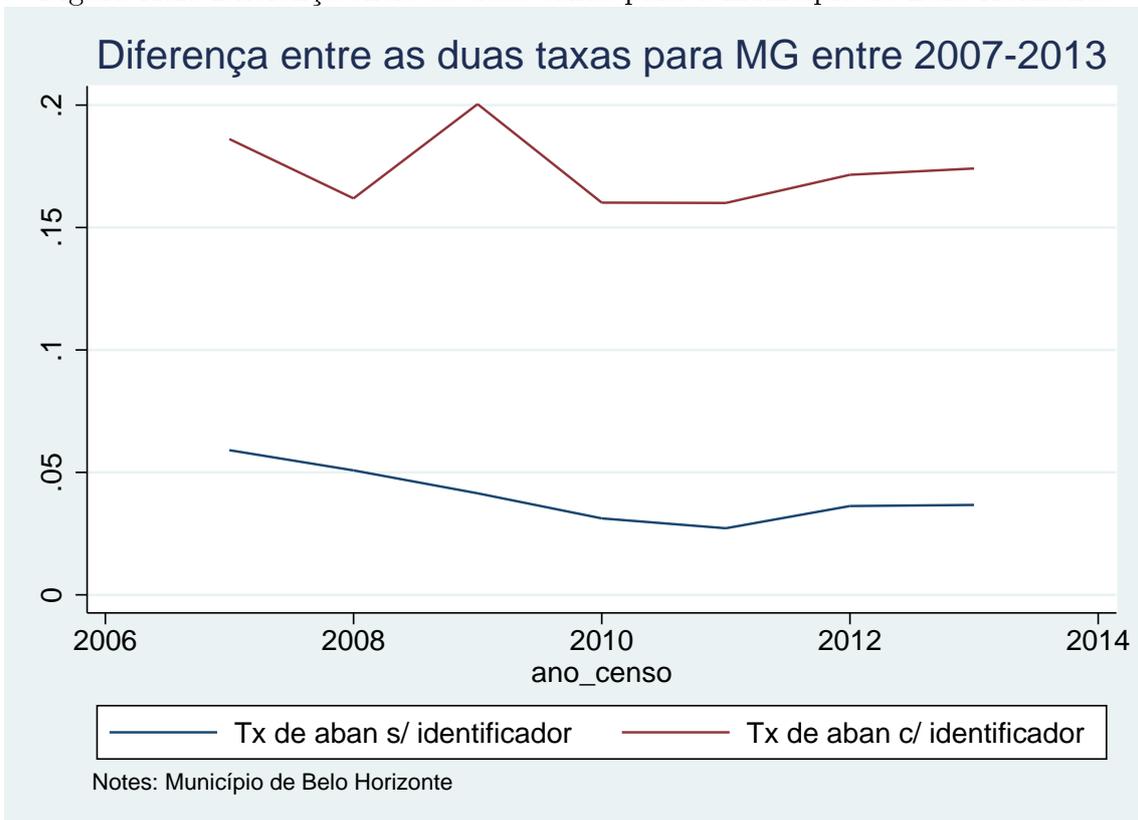


Figura A.3: Diferença entre as duas taxas para o município do Rio de Janeiro

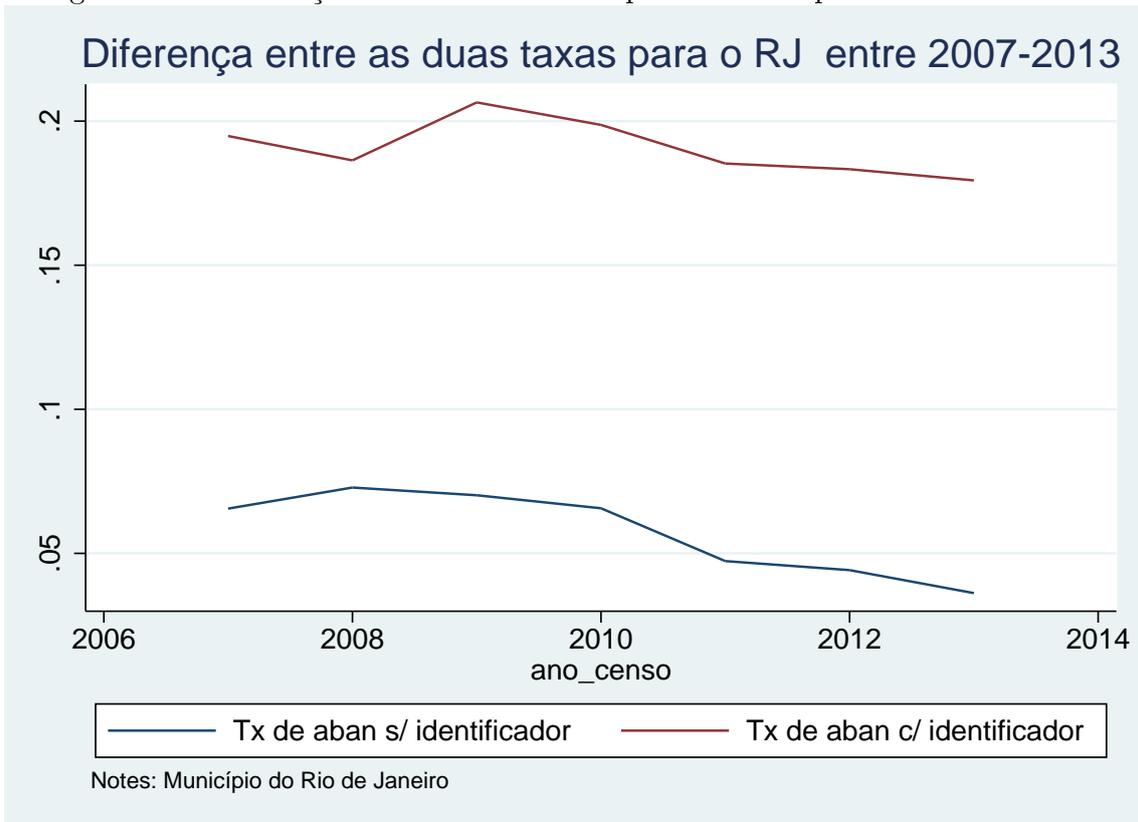


Figura A.4: Diferença entre as duas taxas para o município de União da Serra

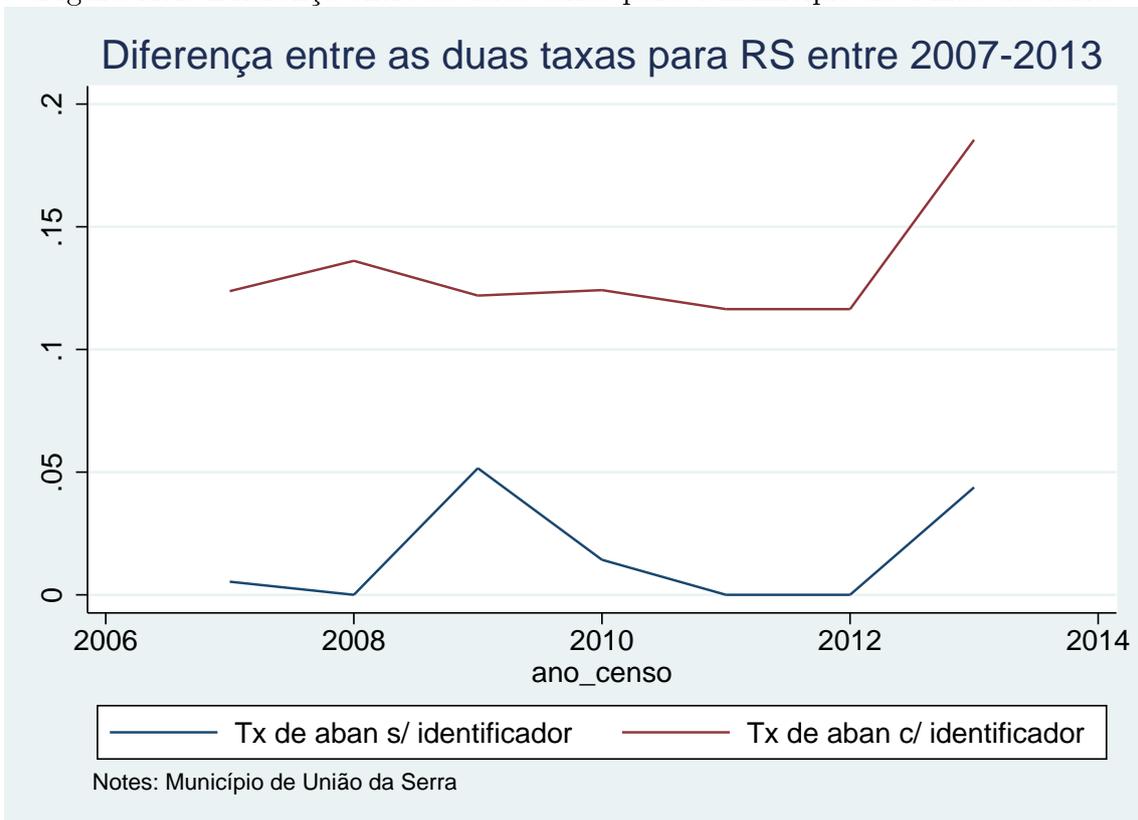


Figura A.5: Diferença entre as duas taxas para o município de Chapada de Areia

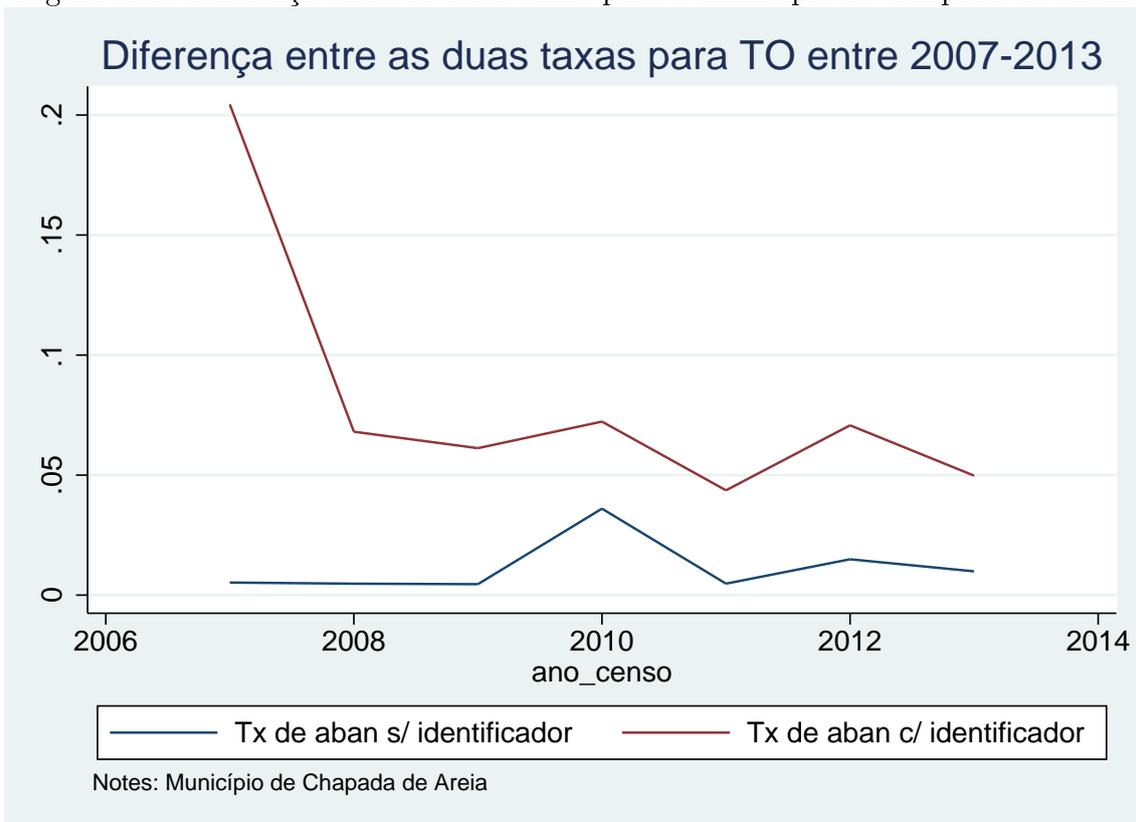


Figura A.6: Diferença entre as duas taxas para o município de Aroeira do Itaim

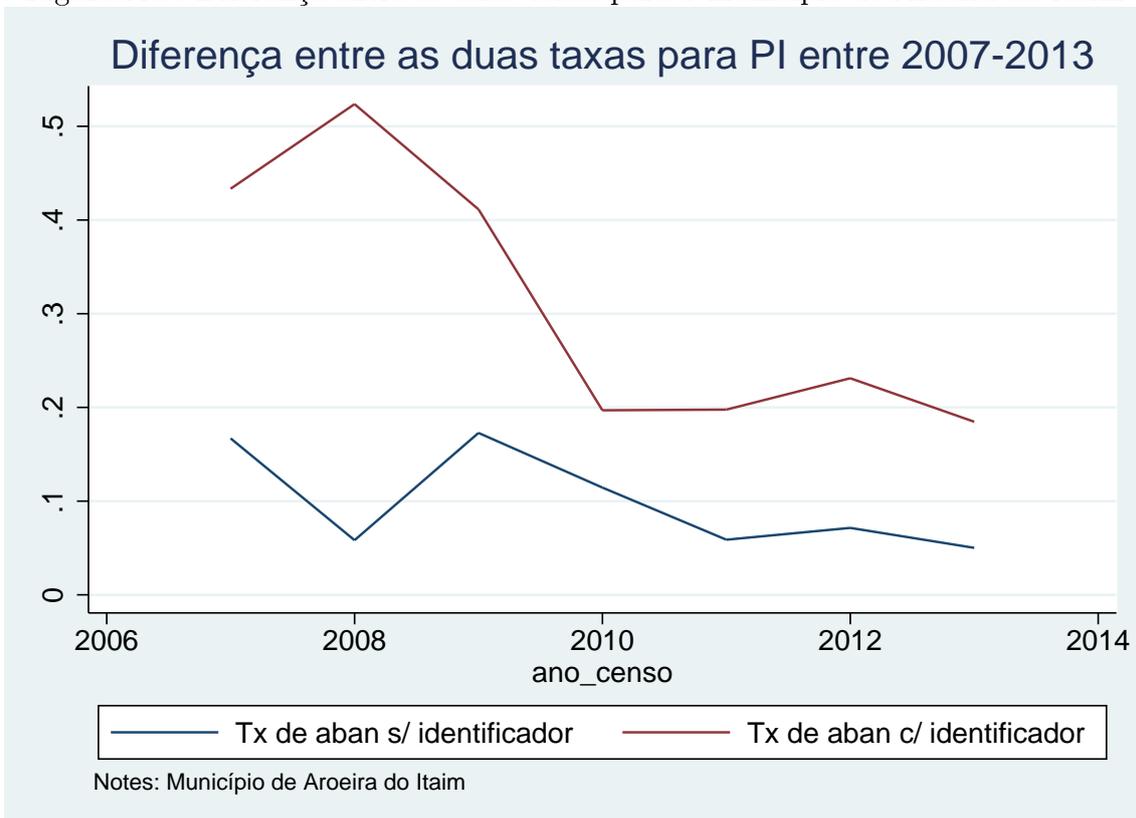


Figura A.7: Dispersão da variação do percentual de votos para seis municípios

