Essays on Empirical Political Economy

Aline Maria Souza Menezes

A thesis submitted for the degree of Doctor of Philosophy

Department of Economics

University of Essex

April 2017

Abstract

This thesis studies three channels through which elections and, ultimately, public policy may be interrelated: new media, electoral systems and vote motivation.

The media has the fundamental role of providing political information to voters. New media such as the Internet brought about an enormous shift in the availability of political information during elections. Exploiting the timing and geographic variation in the introduction of Internet in Brazil, in the first chapter, I show that municipalities with higher Internet penetration voted more often in candidates who faced legal restrictions for advertising in traditional media.

Electoral systems, in turn, have specific features that, in theory, may allow voters to select better politicians by providing more information about candidates and other voters' preferences. In the second chapter, using the discontinuous allocation of single- and dual-ballot electoral rules across mayoral elections in Brazil, I compare the quality of politicians fielded and elected in these systems. In general, dual-ballot candidates from major parties are more politically experienced. This experience may be translated into unobserved political skills that are required to deal with the more competitive electoral process, that, by itself, punishes female candidates, to the extent to which women's participation in politics has been historically low. No differences in performance are observed, except in the attraction of discretionary resources by dual-ballot mayors eligible for reelection, but only in election years.

Finally, in the third chapter, I use a quasi-naturally generated group of voters with differential political information and voting motivations to show that politicians extract more rents in municipalities where they know a number of voters is not directly interested in public goods and do not have readily access to local sources of information.

Acknowledgements

This work would not have been possible without the contribution of a number of people.

First, I would like to thank the Capes Foundation in Brazil for funding my PhD.

I am grateful to Eric Smith for the support and feedback provided over the course of this thesis and Santiago Oliveros for the insightful discussions, guidance and invaluable help.

I gratefully acknowledge the members of my committee, Giovanni Mastrobuoni and Ruben Durante, for their time and valuable feedback on a preliminary version of this thesis. The first chapter has also benefited from comments of Joao Santos Silva, Jayant Ganguli and participants of the 2014 RES meetings. However, all remaining errors are solely mine. Centro de Estudos de Opinião Pública - CESOP-UNICAMP and Marcelo Motta kindly provided me with the survey data and backbone maps used in chapter one, respectively.

Finally, I am indebted to my undergraduate teacher and supervisor Jose Raimundo Carvalho for incentivizing me to pursue a PhD; to my family for their infinite patience during the writing of this work; to the colleagues who shared with me the task of completing a PhD; to my best friend, Davinder, for his exceptional support during these years (If this thesis is finally completed it is also your credit, thank you!); and God who has managed all this.

This thesis is dedicated to the memory of my grandmother, Anair Colares (1931-2014).

Contents

List of Tables

List of Figures

"O que mais dói não é sofrer saudade
Do amor querido que se encontra ausente
Nem a lembrança que o coração sente
Dos belos sonhos da primeira idade.
Não é também a dura crueldade
Do falso amigo, quando engana a gente,
Nem os martírios de uma dor latente,
Quando a moléstia o nosso corpo invade.
O que mais dói e o peito nos oprime,
E nos revolta mais que o próprio crime,
Não é perder da posição um grau.
É ver os votos de um país inteiro,
Desde o praciano ao camponês roceiro,
Pra eleger um presidente mau."

Patativa do Assaré

Chapter 1

Internet availability, political information, and voting: Evidence from Brazil

1.1 Introduction

Recent literature in political economy has shown information plays an important role in voters' decisions. Objectively, voters need to be able to predict and evaluate politicians' performance in office, as well as what the best policy platforms would be contingent to different states of the world. Most of this information is provided exclusively by the media.

The Internet has considerably changed the supply of information available to voters. It gives cheap access to a number of sources in a wide range of political ideologies and preferences. It is also a new medium candidates can use for their campaigns and fund raising. A small but growing literature has been interested on the effects of these additional information sources on people's voting decisions. Specially because this information structure facilitates ideological segregation and the simultaneous growth in the supply of online entertainment can affect the opportunity

cost of news consumption and crowd out traditional media. This literature has focused on the effects of Internet availability, as a source of political news, on turnout and parties' vote shares in parliamentary elections of established democracies.

In this paper, I investigate Internet effects on candidates' outcomes in the 2010 Brazilian presidential election. I find that parties with restricted exposure on traditional media had systematically more votes in municipalities with higher Internet availability.

The expansion of Internet availability in Brazil occurred unequally across municipalities in the late 2000s and the use of online tools in politics was intensified only in 2010, as a result of their successful application in Barack Obama's political campaign in 2008. This scenario allows me to exploit variation in Internet availability across time and space, and assess the effects of the interaction between political advertisement and news consumption. Brazil's presidential elections have distinctive features that create a unique set up able to remove some of the elements likely to reduce the information effects Internet might have that cannot be captured in electoral systems considered in previous research.

First, the elections follow a majority runoff system. First round elections are less plagued by strategic considerations (Martinelli, 2002; Bouton, 2013; Kawai and Watanabe, 2013); when voters know a second round is likely to occur, they may choose to vote for their preferred candidate in the first round, instead of voting strategically. This mechanism would allow voters to reveal preferences generated by information collected online even if the preferred candidate had no chance of victory.

Second, campaign advertisement on TV and radio is restricted to the time made available to candidates by the electoral authority, which is free of charge. Political advertisement is aired simultaneously in all free-to-air channels such that a citizen in a small city in the poor north of the country will watch the same advertisements a citizen in a large city in the rich south does, conditional on watching TV. The

same is true for radio stations.[1] The appearance in newspapers is also restricted, even though newspaper penetration is not significant in the country.[2] This implies that candidates cannot target advertisement, as it is normally done in the U.S. elections. That is, they are not free to adjust the use of traditional media according to the level of Internet penetration municipalities have. The main advantage here is that candidates' adjustments in the use of traditional media[3] would produce baseline municipalities with low Internet penetration that are essentially different from those with high penetration, in terms of intensity of treatment. This could potentially underestimate online information effects.[4] This set up guarantees the homogeneity of what is produced in traditional media across municipalities.

Third, the country's compulsory voting system has been producing historically high turnout rates. In this scenario, voters are more likely to turn up to vote independently of the information they have. That is, it could be the case that online information affects vote decision but not turnout, and in a voluntary voting system we would not observe the choice of this type of voters.

My hypothesis is that the additional political information made available online affected decisions of voters for whom Internet was available, especially with respect to voting for candidates with low share of advertisement in traditional media. I conduct the analysis at the candidate level, but considering the party label as a feature that allows me to compare vote shares over time, so I can use a first differences approach.[5]

[1] Even though voters do not have access to the same radio stations everywhere, every radio station has to air the same advertisements.

[2] I include controls for possible variations in information at the local level, such as mayor's endorsement and candidate's visits to municipalities for public events.

[3] Indeed, Da Silveira and De Mello (2011) have shown TV advertising time affects election outcomes.

[4] Targeted online advertisement is also possible, but there are reasons to believe in the homogeneity of the online political information available to all voters due to the fact that advertisement on the Internet is also regulated, in the sense that paid advertisement is not allowed (such as advertising in websites with high traffic or paying for Facebook or Google to display ads to certain types of voters, etc.)

[5] As it will become clear, it is the party label that determines exposure on the media. And in the models considered here I will use only the variation in the vote shares a candidate receives across municipalities. I am not making comparisons across candidates.

The measure of Internet availability is the number of Internet Service Providers (ISPs) in the municipality.

To deal with the endogeneity of Internet availability – ISPs may serve highly dense and wealthier municipalities first and therefore those with potentially different political preferences – I use an instrumental variable approach. Similar to previous literature, I exploit geographic components of the cost of deploying cables for the Internet infrastructure across municipalities. To reduce costs, telecommunication companies chose to deploy cables through existing power transmission lines, gas and oil ducts, railways, and roads. I use distance from municipalities' centroids to the main routes of fiber optic cables (backbones) as an instrument for Internet availability in a first differences model. The identifying assumption is that, conditional on a set of demographics and infrastructure controls, distances only affected the change in political outcomes through their effect on the costs of infrastructure deployment.

I use official election results at the municipality level to investigate the aggregate effects of Internet availability, and a large and representative survey of the electorate applied at the beginning and at the end of the three-month-long electoral campaign to analyze the same effects at the individual level.

The aggregate data show a positive and large effect for small candidates, remarkably for the third-placed candidate, with a 6.3% increase (on average) in her votes share. For the aggregate of other minor candidates, there was a 7.8% increase (on average). For the main candidates, the Internet effect is negative and significant for the winner candidate, -0.5 percentage point (-25% on average), whereas for the runner-up the effects are not statistically significant. I also find that Internet availability did not affect turnout or the share of blank votes.[6]

Using the individual-level data, I find the same overall effects: positive and significant for the third-placed candidate and negative and significant for the winner. These

[6]Brazil has an electronic voting system. The machine has a button which displays "blank", so voters can choose this option when they do not want to vote for any candidate.

results are consistent with voters switching party lines as a result of the consumption of online information. Especially those unhappy with the incumbent party.

To investigate about the validity of the instruments, I run a placebo experiment where I regress the measure of Internet availability on the vote shares of the presidential election in 2002, when online campaigning was barely used by candidates and Internet was still not widely available to the population. The results suggest that unobserved characteristics of municipalities that could be correlated with political preferences are not driving our main results.

These results reinforce the importance of information in elections, in particular, that obtained online, and presents new evidence that this influence is perhaps augmented by online political campaigns. Moreover, they suggest that the lack of effects found in previous literature for small candidates, attributed to the crowd out of more informative media, may be partially due to attenuating factors stemming from the institutional setting in which they have been evaluated.

The main result that small candidates benefit from the possibilities of advertisement allowed by the Internet, prompts to more specific issues regarding the disclosure of information in elections. As Castanheira (2003) and Piketty (2000) point out, the possibility of voting for candidates that have no chance of winning increases the informational content of elections when voters signalize their policy positions to influence mainstream parties. My results suggest that a possible substitution effect between communicating policy preferences using online tools and via casting a vote may not outweigh the effects of minor candidates reaching more voters through the Internet.

This paper contributes to the literature on media economics that identifies effects of new media on political behavior (Gentzkow, 2006; Strömberg, 2004; Enikolopov, Petrova and Zhuravskaya, 2011; DellaVigna and Kaplan, 2007). It closely relates to Miner (2015), Campante, Durante and Sobbrio (2014), and Falck, Gold and Heblich (2014) that identify Internet effects on turnout and parties' vote shares for parliamen-

tary elections in Malaysia, Italy, and Germany, respectively. Unlike these studies, I investigate these effects in presidential elections in a different institutional environment. The problem about measuring the effects of campaigns at the candidate level in general is the selection bias in advertisement. For example, candidates might target cities where they already have some advantage (or not). Looking at the data at a more macro level, parties at parliamentary elections, takes averages across candidates and reduces the bias. Here, I have the framework described above that accounts for this fact. This level of analysis here allows for a more precise inference on the effects of the interaction between online and offline news content and advertisement.

The paper also adds to the literature on campaign advertisement and voting behavior (Da Silveira and De Mello, 2011; Levitt, 1994; Ansolabehere and Iyengar, 1994*a*,*b*). The variation in the use of different media created by the institutional framework of these elections allows me to show advertisement has indeed an important effect on votes, and that the use of Internet as an advertising medium can compensate for the lack of advertisement in traditional media, which might be an important factor for candidates with small budgets, for example.

The rest of the paper is organized as follows. Section 1.2 describes the background of the Internet roll-out, the electoral system and the media market in Brazil. Section 3.2 describes the data and construction of variables. Section 1.4 discusses the identification strategy. Section 1.5 presents the main results. Section 1.6 concludes the paper.

1.2 Background Information

1.2.1 Electoral system

Brazil is a democratic federative republic with a presidential system. With approximately 136 million registered voters, it is one of the world's biggest democracies.

General elections are held each four years: president and parliament members are chosen at a national level, and governor and local parliament members at a state level. These elections occur simultaneously, so each voter has five choices to make: president, senator, federal deputy, governor and state deputy.

The country has a compulsory voting system where citizens between 18 and 70 years of age are required to vote by law. Voting is voluntary only for citizens aged 16 or 17, older than 70 years, and illiterates. Because of the compulsoriness, election dates are national holidays to facilitate voters' attendance. This requirement is enforced by state agencies who will demand a proof of electoral participation in any type of service citizens may need, such as the issuing of a passport or a driving license. Non-attendees have the option of justifying their absence with the electoral authority and use it as a proof of electoral participation. Though, this option is often as costly as voting itself, since the voter has to visit a polling station or a local post office to do the paperwork.

The average turnout in the last four presidential elections was about 80% of the total population in voting age, with 77% in the presidential election of 2010. To give a sense of magnitude, in the last four U.S. presidential elections, which follow a voluntary voting system, the turnout rate was about 55% of the population in voting age.

Presidential and gubernatorial elections follow a two-round majority system, where a second round is held, after a month of the first round, with the two candidates with the majority of votes, when none of them has more than 50% of the valid

votes in the first round. Terms last four years and elected candidates can run for re-election for one consecutive term.

Senate elections follow a first-past-the-post system. Terms last eight years and each state has three representatives. The election for federal deputies follows a proportional system, where 513 seats are allocated to the parties according to an open list. The same is valid for state deputies, but the number of seats varies across states according to the population size.

Candidates must be supported by a party or a coalition of parties. There are about 29 parties in the country, but not all of them are relevant at a national level. In fact, only about six of them are politically strong nationally. The party system is extremely unstable and the majority of voters cannot recognize what their ideological positions are (CSES, 2010). Politicians, in turn, change party identification very often from election to election. This lack of identity makes political campaigns even more relevant to voters' decision making (Samuels, 2001).

As in the U.S., candidates, rather than parties, are responsible for raising and spending the most of campaign funds. There are no *a priori* limits to campaign spending. At the time of registration for the race, candidates declare their expected spending and that becomes the limit of what they can spend.

Political parties are funded by public and private sources. Direct public funding is allocated proportionally to the number of votes the party obtained in the last election for the chamber of deputies. Companies can donate up to 2% of their gross revenue on the previous year and ordinary people up to 10% of their total income on the previous year. There are no limits when candidates spend their own resources in the campaign.

Advertisement slots for political campaigns on TV and radio are free of charge by law. They are allocated to each party according to an specific division rule: (i) one-third of the total time available is proportionally divided among candidates; and

(ii) two-thirds are proportional to the number of representatives each party has in the chamber of deputies. Therefore, major parties get more time of advertisement in the mass media.

Table 1.1 displays the distribution of the time each candidate had to broadcast their adverts on TV and radio in the 2010 election. While the first ranked candidate had 10 minutes of advertisement time on TV, minor candidates had only 55 seconds. Additional time cannot be purchased so budget does not determine media exposure, at least not in the short run.

The media coverage of candidates' activities during the campaign period is fairly regulated in terms of time devoted to each candidate on TV and radio's news bulletins or space in newspapers. Negative advertisement is forbidden and the law is fairly enforced because of the monitoring parties themselves do against opponents.

The use of Internet for political campaigns in previous elections was restricted (by law) to candidates' and parties' web pages. The Bill 12034 of 2009 relaxed these restrictions allowing any citizen to make use of online tools to express political positions, as long as they were not made anonymously or through paid advertisement. Note that this condition made targeted advertisement in websites such as Facebook or Google illegal. The electoral authority would take down any website breaking these rules for a time determined on a case-by-case basis and charge a penalty from the implicated candidate. Even though the magnitude of online content available imposes restrictions in identifying and punishing such cases, the candidates themselves make sure to sue opponents whenever they can identify irregularities. This mechanism hold candidates accountable.

In fact, according to data from Google Transparency Report[7], from June to December 2010 about 98 court orders (more than 12000 items) for defamation and 19 for

[7]The data can be found on this address `http://www.google.com/transparencyreport/removals/government/data/?hl=en`. Google has been publishing in this website all worldwide government requests for content removal since 2010.

noncompliance with the electoral law (35 items) were filed to remove content across all Google's platforms. For comparison purposes, in 2011, a non-electoral year, there were zero orders for noncompliance with the electoral law and 136 for defamation (including only 683 items to be removed). This number went up in 2012 again, when mayoral elections were held.

The Internet itself also provides tools that facilitate monitoring illegal practices. For example, one information crowdsourcing website was used to collect reports from all over the country about abuses of the electoral process.[8] The website eventually attracted attention of traditional media and was featured on Google's dedicated page to the elections of that year. The project received over 1200 repor ts of irregularities during the electoral period.

1.2.2 Media and Elections

Free-to-Air (FTA) television is the most popular media in Brazil. In 2010, TV penetration was about 97%, with FTA accounting for 80% of the audience. Pay-TV has still a small market in the country, despite its growth in the last decade. Approximately 73% of the population listens to FM radio stations and 31% AM radio stations. From these, 10% using the Internet as a medium, and 18% a mobile.[9] With respect to news consumption, survey data from 2010 shows that 45% of the population preferred television as their main source of information, and 40% the Internet. Newspapers penetration is one of the lowest in the world, with a rate of 0.04 newspapers for each 1000 inhabitants.[10]

[8]The website was called *Eleitor 2010* and used the crowdsourcing online service *Ushahidi* to map reports of irregularities coming from the population. The website is not available anymore, but more details about it can be found in this address `http://transparency.globalvoicesonline.org/project/eleitor-2010`.

[9]Data from *Panorama da Comunicação e das Telecomincações no Brasil*(Panorama of Communications and Telecommunications in Brazil), 2011-2012.

[10]Instituto Verificador de Circulação (IVC), in *Panorama da Comunicação e das Telecomunicações no Brasil* (Panorama of Communications and Telecommunications in Brazil), 2011-2012.

The news market in FTA television is dominated by one channel, TV Globo, which registers the highest audiences among news bulletins in television. Nonetheless, it has presented a decline in audiences since 2005, registering a minimum in 2010, as viewers migrate to other platforms and media outlets. According to IBOPE (Brazilian Institute of Public Opinion and Statistics), between 2000 and 2009 the average percentage of television sets turned on during prime time fell from 66 to 59 percent, and the five biggest broadcasters of the country lost, as a group, 4.3 percentage points in the ratings. Most importantly, the growth of the Internet appears to be a pivotal cause of the decline in FTA television. According to research carried out by *F/Nazca*, a significant proportion of Brazilian Internet users say they have abandoned some forms of traditional media and focused attention on the Internet.

Survey data from 2010 shows that 55% of Brazilian Internet users used Google as the primary tool to search for news, and 51% social networks. Websites of traditional news outlets appeared in 5th place only. In June 2010, Brazil was the second biggest market in the world in terms of Twitter penetration, with 20.5 percent of its Internet users having an account, reaching the first position in the last month of the presidential campaign in October 2010. The country at the time was also the largest market for Google's social network Orkut.

Candidates use television, radio, newspapers and Internet for campaigning. As previously discussed, radio and TV time are determined by the electoral authority based on the number of representatives a party has in the Lower Chamber. Each presidential candidate produces television and radio "shows", normally with new editions every week, that are aired everyday at lunch time and in evening prime time across all free-to-air channels and radio stations. That is, a voter that has access to FTA TV only, will have to watch political advertising during the three months of the electoral campaign during this time.

Television newscasts are enforced by law to allocate the same time of coverage of daily activities to each of the main contestants in the election, normally the top three. For the remaining candidates they just make rapid mentions about what the candidate has done on that day. Apart from the advertisement in the media, candidates travel to municipalities for rallies, inaugurations, etc.

A key aspect of the Internet's role in elections may be the transformation of how political campaigns are organized and conducted. Now, candidates make use of a wide range of unprecedented tools, available online, to track and reach voters in a targeted manner. In Brazil, this new possibility was particularly important for candidates with negligible time of advertisement in traditional platforms. Anecdotal evidence says that the third-placed candidate, Marina Silva, made "efficient use of the Internet to receive donations, publicize her political agenda, and interact with voters. [She also] made systematic use of social media, mainly through Twitter"(Mizukami, Reia and Varon, 2013), and this was an important determinant of her final share of votes.

Each candidate had a website where voters could find details about policy propositions, biography, agenda, and so forth. All of them were also in social networks, with a constant stream of posts, videos and chats. Nonetheless, small candidates were the ones who tried to use online tools in the most creative ways. For example, the debates on TV most of the time include only candidates that polls indicate are the top three. In a similar debate held online, with just the top three, the forth-placed candidate, Plinio Sampaio, independently held a parallel live section on *Twitcam* to answer the questions asked in the debate as well as to comment on the other candidates' answers. This led to the candidate being featured on Twitter's trending topics on that day, above all the other candidates.

An important aspect of this competition for attention online was that small candidates would come up with innovative ways of using existing online platforms and major parties' candidates would follow up. In fact, most of the online services such

as blogs, Wikipedia and Google search engine experienced peaks in terms of political themes in the three months of presidential campaign. Because Google is the leading search engine in the market, its volume of searches for a given topic serves as a good snapshot of interests, concerns, and intentions of the population over time. Figure 1.6 displays weekly Web search volume on Google Trends for terms related to politics – as defined by Google itself– in 2010, the election year, and in 2009-2011, non-election years. The grey area highlights the period of electoral campaign. Interest in politics was higher than in the previous and subsequent year of the election for the entire campaign period, reaching a first peak in the week of the first round, and a second peak in the runoff, going back to normal levels just after.

To further illustrate this online activity, Figure 1.1 breaks down the search volume by candidate and displays search intensity by states. Darkest areas in the map depict those states with the highest search volume in the period. As it can be seen, these queries were fairly distributed across the country, with surprisingly blank areas in states where campaigns are produced and concentrated – South and Southeastern regions. This perhaps illustrates how important alternative sources of information were to voters in relatively more peripheral areas in a country with continental dimensions such as Brazil.

Since Obama's 2008 campaign, where it was first used for political purposes, the microblog Twitter has become one of the most important platforms for political debate (Enli and Skogerbø, 2013; Petrova, Sen and Yildirim, n.d.). It is therefore informative to verify the dynamics of Twitter usage across candidates. Figures 1.2 and 1.3 illustrate Twitter activity for the top four candidates during 2010. Except for the second-placed candidate, Jose Serra, who created a Twitter account already in 2009, the other three candidates did so in 2010 only. The data are displayed from the time they created their accounts until later after the elections. Perhaps as a result of this earlier creation, PSDB's candidate was the one with the highest number of

followers throughout the year, but all candidates experienced increases in the count of followers in the electoral period and became stagnant afterwards (Panel (a) of Figure 1.2). Interestingly, the highest monthly growth was verified by the fourth-placed candidate, Plinio de Arruda (PSOL), when the other candidates presented similar rates (Panel (b) of Figure 1.2). He was exactly the candidate who posted the most during election time (Figure 1.3), followed by the third-placed candidate, Marina Silva (PV).

comScore data, which measure online audience at the user level through a tracking software installed in the computers of a set of panelists, also show that the audience of news websites and blogs were the highest on the months leading up to the election.[11] This is anecdotally attributed to the use of the Internet for campaigning purposes.

It is then important in its own right that we understand how this availability of information affected voters decisions. I claim that the Internet introduction might have had both an information and a propaganda effect. Information because more news sources were made available online and propaganda because candidates had a supposedly cheaper and not as regulated medium to run campaigns. Minor candidates, normally with very low budgets and negligible broadcast time on TV and radio, were able to reach voters more easily. My hypothesis is that these two effects raised the number of votes these candidates received.

1.3 Data

I estimate Internet effects on voting outcomes both at the aggregate level (municipality) and individual level. The subsections below describe in detail the data sources used and the construction of each relevant variable for the two levels of analysis.

[11] These data are available on `http://www.slideshare.net/DigiTalks/alex-banks-a-audincia-brasileira-em-2010?related=1`.

1.3.1 Aggregate-level data

To be able to estimate the effects of Internet on voting outcomes at the aggregate level, I match voting information to a series of variables, aggregated at the municipality level.

Election data

The election data are from Brazil's federal electoral authority, the *Tribunal Superior Eleitoral (TSE)*, for 2002, 2006, 2008 and 2010 elections. The data are provided at the precinct level and includes candidates names, parties, and votes broken down by those received by candidates, blanks, and invalid[12]. I aggregate them up to the municipality level to match the level of the other variables considered. To construct the shares for the candidates, I divide the number of votes received by the total number of registered voters in the municipality.

The number of municipalities in the country goes from 5507 in 2000 to 5565 in 2010 as a result of the creation of new municipalities over the decade. I exclude from the sample all the 119 municipalities involved in this process due to incompatibility of data levels of aggregation or disaggregation. So I am left with 5446 municipalities in the sample. The majority of these excluded municipalities (78%) are from the south of the country, where the penetration of Internet is higher on average. Hence, at most this would insert a downward bias in the Internet availability effects. The final number of municipalities is 5445 though, because I loose Fernando de Noronha, a small island in the northeast cost of the country for which I have missing observations.

[12]Invalid votes are those in which the voter types a number that does not correspond to any candidate. It can be done either intentionally, to spoil the vote, or by mistake.

Controls

I split the variables used as controls into four groups: controls, other media, local support, and state dummies.

Controls. This group includes demographics relevant to election outcomes. I use a 10% sample of the 2010[13] and 2000 censuses from the *Instituto Brasileiro de Geografia e Estatística - IBGE*, about 20 million people, which is representative at the municipality level. I use the frequency weights provided to expand this sample to the whole population and then create the proportions for each variable used. All variables extracted are conditional to people being in voting age, that is, above 15. I use the following variables (in shares): unemployed, college education, females, catholics, migrants, urban population, receives some social benefit from the government, blacks, and age groups (16-17, 70-plus). I also extract the median household monthly income (in terms of minimum wage) in each municipality. The variables GDP growth in 2009, and GDP per capita in 2006, and the (log of) population density, at the municipality level, are also in this group. The data are also from IBGE.

I include density of roads, transmission and distribution lines and railways at the municipality level to account for the development effects this pre-existing infrastructure might have entailed. They are calculated as the ratio between the length of all roads/lines/railways within a municipality and its area. The data are for 2005 and provided by IBGE in shape files. I then used GIS software to calculate these densities. The development index used, the Gini index, captures income inequality and development levels. The data are from the United Nations Development Program (UNDP) and *Instituto de Pesquisa Econômica Aplicada (IPEA).* I also use per capita GDP and total population as controls with data also provided by IBGE. *Other*

[13] According to IBGE, census interviews were conducted from August 1st to October 30th, so they overlap the electoral period. For variables that are potentially bad controls, I interpolate the values for 2000 and 2010.

Media. These are controls for media availability other than the Internet. There are dummies indicating whether the municipality had a local newspaper, magazine, radio station, local TV channels and the number of FTA TV channels they had access to. This group also includes the share of voters with TV and radio sets at home, extracted from Census data. The data on the structure of media outlets available in each municipality comes from the *Perfil dos Municípios Brasileiros 2006 - IBGE* survey.

Local Support. This group has to do with all candidate-level variables that might affect election outcomes. First, I include an indicator for those cities visited by a given candidate during the campaign. I denote this variable $I(VisitedCand_i)$, where i is the candidate. It accounts for candidates' direct activities at the municipality level. To construct this variable, I use data from a magazine of national coverage, *Veja*, on the agenda of the first three ranked candidates in the election. The magazine details the place each candidate visited, by dates, and qualify their activities there. These variables for the three candidates are highly correlated, indicating that the candidates concentrate this type of campaign on the same places, on average. Because of that, even tough we do not have the same data available for the other candidates, we suspect that they have focused their activities on some of the other candidates' visited cities, such as capitals. The next level of controls include the share of voters who hold a membership with a party that supports a given candidate. I assume these voters are more likely to vote for these parties independently of the level of information they have. I call this variable *Members* $Cand_i$. It is built using data from TSE, which details (i) the name of voters affiliated to each party by municipality, (ii) when they have registered with the party, and (iii) whether they have already left the party or not. To construct the variable, I count the number of voters registered with a party before 2010 and have not cancelled their membership. I then divide these counts by the total number of voters affiliated to some party at the municipality.

Finally, in $Mayor\ Cand_i$ ($Gov.\ Cand_i$), I include those municipalities with a mayor (governor), at the 2010 legislature, that was of the same party of the presidential candidate in question so as to control for the networks of support they can create locally. These indicator variables were constructed using election results of the 2008 mayoral elections and 2006 gubernatorial elections, also provided by TSE.

State Dummies. Here I include (i) a dummy for each state, to control for state fixed-effects; (ii) a dummy for capitals, where politicians normally concentrate their campaigns; and (iii) a dummy for municipalities belonging to the Amazon region, where both population density and Internet penetration are below the average.

1.3.2 Individual-level data

To evaluate whether the aggregate-level results can be replicated at the individual level, and are not just a feature of ecological inference, I use individual-level data from a survey applied both at the beginning, in July, and at the end, in September, of the electoral campaign. Even though these data are representative at a national level only, each survey has about ten thousand observations and asked voters about their candidates' choice and the sources of information they used to get informed about the election. The data are from *Datafolha*, one of the main polling institutes in Brazil.

July's survey interviewed 10905 voters. It asked voters to rank in order of importance which type of media they used the most to get information about candidates in the election. I create a variable that is 1 for those who ranked Internet as first or second media used, and zero otherwise. From the same question I also create two other variables indicating voters who ranked newspapers and talk with friends as their first or second most preferred means to acquire electoral information.

As controls, I am able to use age, gender, university degree, income, employment status, partisanship, and ownership of radio and television sets. The survey also asked

whether voters had Internet access at home or not. With this information I create the indicator variable *Home internet*. From the survey I also know the party to which the respondent reported being closer to their political views.

September's survey had a different set of respondents. It interviewed 11660 voters nationwide. And, unfortunately, they did not respond to the same questions. But I am able to construct the variable for Internet use from a slightly different question: Which media have you been using to get information about the candidates running for the national parliament? I assume these voters are also collecting information about candidates running for other positions in the election from the same media. Additionally, I construct a variable that accounts for voters' level of interest in politics. I use a question that asks whether the voter could remember for which candidate they had voted for the lower chamber in the previous election. This is a concurrent election and there is always a sizable number of candidates to choose from. So, if the voter can remember the name of the candidate chosen in that election, it is an indication of an interest in politics that is above average. The voters were prompted to tell the name of the candidate, so lies are less likely to occur.

1.3.3 Measuring Internet availability

The 2010 Census asked a random sample of 10% of the population, approximately 20 million people, whether they had a computer and/or Internet access at home.

The sequence of questions in the survey was as follows: (i) Do you have a computer at home? In case the person said yes, the interviewer would further ask (ii) Do you have Internet access at home? That is, only people who had computers answered the question about Internet. Table 1.2 displays descriptive statistics[14] of the distribution

[14]The missing values found in the data correspond to the places where these questions were not applied: hotels, accommodations, hospitals, military bases, etc. That is, any place that cannot be characterized as a household.

of computer and Internet availability using the data at the individual-level.[15] Note that those who had a computer at home but not Internet are 8 percent of the total population.

A survey held in 2010 by Brazil's Ministry of Communication[16] asked voters the place where they were able to use the Internet more often. The majority of them reported having access to it at home, but 32% were able to connect from other places such as cybercafes (18.4%) and workplace (7.5%). Hence the availability of Internet at home is likely to be a lower bound for the actual Internet availability.

Comparing these data with other measures of Internet penetration further reveals that the share of people who reported having both computer and Internet at home in the census is a noisy measure of Internet availability. I arrive at this conclusion by comparing the two measures with a third measure of Internet availability: the number of Internet subscriptions in each municipality. The data are provided by Brazil's telecommunications authority (Anatel). Each year, Internet Service Providers (ISPs) are required by law to report to the agency the number of subscriptions they have in each municipality. I create a variable that measures the number of subscriptions per 10 people and check the correlation with the other measures (Table 1.4). While the correlation between Internet and number of subscriptions is 0.37, this number goes up to 0.75 with computer ownership. It is true that the measure of access points also includes business addresses, but this is not necessarily a measurement error if we consider the distribution of places where people do have Internet access. Importantly, this measure also accounts for cybercafes, for example, an important medium used by people in poorer areas.

Apart from that, its true that the measure of Internet take-up is more plagued by endogeneity concerns: once the Internet service is available in a location, the consumer

[15] I expand the 10% sample using the probability weights that extends it to the entire population.
[16] *Hábitos de Informação e Formação de Opinião da População Brasileira.*

decides whether to contract the service or not. It is also much more difficult to find an instrument that can identify compliance to the treatment at this level of decision.

So I follow the previous literature on media economics that has used availability of treatment to identify LATE effects of the introduction of a new media on voter behavior. Enikolopov, Petrova and Zhuravskaya (2011) have used signal availability of a radio station to show that opposition parties benefited from voters' access to a source of information not controlled by the government. Gentzkow (2006) uses data on availability of television in U.S. cities to show that television introduction had a negative impact on voter turnout. DellaVigna and Kaplan (2007) uses Fox News availability to show there was a positive effect of its introduction on the vote shares of Republicans in U.S. presidential elections. Specifically with respect to the Internet, Campante, Durante and Sobbrio (2014) have used data on the availability of ADSL technology, measured as the percentage of households with access to ADSL-based services, and Falck, Gold and Heblich (2014) uses the number of households that are covered by networks to show the effects of Internet in voter turnout and election outcomes.

Based on these considerations, I use the number of ISPs in operation as my measure of Internet availability in a municipality.[17] A measure of availability should capture the number of households that could be served by an Internet connection if they all could contract the service, instead of actual take up. Tables 1.3 and 1.4 displays correlations between Internet take-up and number of ISPs in levels and differences between 2006 and 2010. As it can be seen, there is a high correlation between these measures. Figure 1.5 presents graphically the estimations in Table 1.3's columns (2) and (3).

Due to the fact that telephone companies had the advantage of holding infrastructure that could be adapted to provide Internet, and that this market was already

[17]The number of ISPs as a measure of Internet availability has also been used by Jaber (2013).

highly concentrated, the Internet market in Brazil is dominated by six large ISPs that provide services to 78% of the market (Cetic, 2011). The largest number of ISPs fall within the small and medium categories (99.7%) accounting for 22% of the market. The important factor about the composition of this market is that small and medium ISPs enter smaller municipalities, which are normally unattractive to large providers. Therefore, using number of ISPs as a measure of Internet availability will probably allocate more weights in the variation of the service in smaller cities, which are 98% of all municipalities in Brazil. This is an important feature in favor of this measure, as the large urban centers are the first to be served with Internet connectivity and have greater exposure to all types of political campaign, factors that interact to produce stronger endogeneity biases.

Even though the market for large ISPs does not seem to have competition, in the market for smaller ISPs, competition over price and quality seem to be relevant since their number present stead growth over time within municipalities. The delay in auctioning the spectrum to be used for providing 3G technology, which involved a different structure of costs, meant that this type of Internet access, that is used especially in smartphones, was not available in the country until 2009. In fact, the auction of this frequency band was held only in 2008. Internet provision through cable technologies, are commercialized bundled with TV services, what makes it more expensive. In addition, cable television pre-existing infrastructure serves only affluent areas of larger cities, restricting its reach. All in all, most of the Internet in Brazil is provided through copper cables, the pre-existing telephone infrastructure already available at the time of introduction of the Internet. A detailed discussion about the composition about the representativeness of each type of different technologies involved in the provision of the service is left for the Appendix 1.7.3.

1.4 Empirical Framework

1.4.1 Identification Strategy

The cost of Internet deployment depend on geography, topological aspects such as the type of terrain or altitude, and predicted demand. The higher the geographical cost of installation, the higher the predicted demand for the service has to be, which depends primarily on the population density of the area. There is a series of costs involved in building the many layers of the Internet infrastructure. To understand this structure of costs, we have to understand more broadly the composition of this infrastructure.

The provision of Internet involves many network elements: international network links, domestic backbones, backhauls, and access points. At the top of the chain is the international network of links connecting countries via submarine cabling systems or satellite, the international backbones. At the second level, each country has a domestic system of fibre-optics (in some cases it can also use satellite or microwaves) that takes over the signal from the submarine cable/satellite and distributes it across the country; these are designated regional backbones, they aggregate and transport high volumes of communications traffic. At the third level, there are the backhauls, the switching and routing nodes. They are the intermediate links (also in fiber-optics) between the backbone, the core of the network, and the hierarchically smaller networks. Finally, the access points constitute the so called "last mile" of the system, that is, the links connecting final consumers to the rest of the network. The technology used at this part of the network is normally xDSL, which requires telephone infrastructure, or cable technology (bundled with the cable-TV infra-structure). Internet Service Providers (ISPs) coordinate the distribution of the service to end users.

The company with the largest coverage in Brazil has 628,000 km of fiber backbone networks all over the country. The market includes 14 different networks at national

level, from which 95% is owned by the five largest telephone companies, even though there were 2500 ISPs in the market in 2010.

Work on the foundations of the Internet infrastructure in Brazil started in the 90's, with the construction of a network connecting universities, but only with the privatization of the state-owned telecommunications system in the late 90's and the approval of regulatory measures that prepared the ground for a system of shared infrastructure among companies, Internet services were extended to private users. The privatization fostered competitiveness in the market, bringing down prices and democratizing access to the technology. Since mid-2003, the Brazilian telecommunications market has enjoyed a period of remarkable growth, led by mobile telephone and broadband. As a result, the number of households with internet access raised to 22 million (38%) in 2010, almost three times the numbers registered in 2000. A large share of the low income population has been able to have access to the Internet due to the combined effects of lowered prices of computers and Internet services and the growth of households' income that brought 30% of the population to the middle class in the same period. Those who are still unable to afford computers or Internet access can take advantage of thousands of cybercafes spread throughout the country, built as a consequence of the information boom.

To meet the ever growing demand for Internet services, many companies have built their own networks (backbones). The placement of fibre-optic routes are normally defined on the basis of pre-existing infrastructure. The electricity, gas and oil transportation and distribution systems, that also follow paths defined by roads and railways are normally used to host fibre-optic cables. In Brazil, the large part of the infra-structure is composed by Optical Ground Wires (OPWG) hung on pylons that are also used for electricity distribution (Knight, 2014).

According to specialized literature, the advantages of using overhead electricity transmission and distribution lines to carry fibre-optic cables are related to the speed

of installation, security in terms of the occurrence of cable damages, and the reduction of deployment costs. Underground cable requires the digging of trenches, which in turn need environmental permits. These requirements increase considerably the time for installation. Attaching them to poles and pylons is considerably faster, and depend less of favorable natural conditions for its execution. About security, rates of theft and vandalism are lower, given the natural protection the proximity to power conductors offers, and not subjected to damage by excavations.

The cost of aerial cables is higher, but the easy with which they can be deployed and the savings from not building additional infrastructure, imply that the installation of aerial cables have the total lowest cost, when compared to underground alternatives Ostendorp and Gela (1997). Another advantage of the use of these lines is that their path of distribution eventually reaches all the costumers of interest, enabling the connection of the fiber to the broadband service providers' infrastructure, usually the telephone infrastructure, and also reaches the end user. Part of the fiber-optic routes were also laid underground following the path of roads and railways. Deploying underground cabling involves getting rights of way and negotiating leasing prices. Deployer companies are normally liable for any damage or interference with existing networks and have to preserve the integrity of the road during earthworks (Kelly and Rossotto, 2012; Knight, 2014).

Even though most of the Brazilian cities are covered by telephone access, technically not all of them can have Internet access. The problem is related to the distance of the consumer premises from the so called central office structure of the telephone company. The further the consumer is from the premises, the higher is the attenuation the signal suffers, and the slower the connections are. With respect to the quality of the call the problem can be corrected by inserting loading coals to the cables to increase the inductance of the circuit. This is good for calls, but terrible for the data transmission at the frequency used in the Internet data. And to correct

this issue is not enough to have the cables replaced since the distance itself will be an impeding factor. The provision of Internet to these areas depend on the availability of fiber-optics signals nearby, via backhaul, or the construction of a new central office, or the use of another type of technology, such as FTTH (Fiber to the home), but the cost might be prohibitive. Since the backhaul installation costs are also high, this infrastructure is not normally immediately available.

In Brazil, most of the infrastructure already built for electricity and gas transportation was used to lay fiber-optics cables. The costs involved in this case were the negotiations with power companies to allow the inclusion of the cables in their structure. Backhaul owner companies with bandwidth capacity can both operate on the wholesale market, leasing bandwidth to Internet Service Providers (ISPs), or provide the final service themselves.

I use the networks of backbones dataset built by Motta (2012) to calculate distances from the municipalities' population center to the closest probable location of a network of backbones. The rationale is that distances[18] increase costs of deployment, since the network will have longer paths to follow, there will be an increase in the quantity and quality of cables used along with the amount of hosting infrastructure that needs to be leased. These data display only the nodes of the network without the details of the path of the fiber-optic cables along the nodes. Technically, the ISPs can construct a backhaul from any section of these paths.

Using Geographic Information Systems (hereafter, GIS), I put together four layers of networks that cover the whole area of the country: federal and state roads, railways, gas and oil ducts, and transmission lines. These are geocoded data provided by IBGE. I combine these layers into a unique layer that connects all the municipalities Motta (2012)'s dataset reports to have a link between backbones' nodes. Using this new layer, I calculate the minimum distances from the geographic mean population centre

[18]Distances to backbones have also been used as instruments for Internet availability in Miner (2015).

of each municipality to the closest complete link between two cities, the probable path of the backbones structure.

The assumption here is that the cheapest way of building this network would be using the pre-existing infrastructure optimally. Proxy cost models of network deployment in developing countries, see Benitez et al. (2000), point that the best way of lowering Internet infrastructure costs is making use of overhead power line structure used in electric power transmission and distribution of electrical energy, deployed along roads, railways and pipelines networks for oil and gas transportation, to avoid going into rain forests and/or digging trenches that would considerably increase the time and costs of provision, because apart from the higher building costs they also require environmental permits.

This identification strategy is in line with Duflo, Glennerster and Kremer (2007) and Lipscomb, Mobarak and Barham (2013) where counterfactual costs are used as instruments. The idea is to calculate what the cost of building infrastructure would be if the providers were taking into consideration only factors that are not directly related to the demand for the service, such as geographic distances, land gradient, etc.

The necessary identification assumption here is that pre-existing infrastructure does not have a long lasting effect on municipalities characteristics, conditional on a set of sociodemographic controls. If we also consider that the position of pre-existing infra-structure was partly based on geographic suitability – for example, it is easier to build roads or deploy pylons in places with suitable terrain properties such as flat slopes and good drainage – once we control for long lasting developmental effects infrastructure might have, the variation left is plausibly exogenous.

As Rodrik, Subramanian and Trebbi (2004) point out, geography may have a direct effect on incomes through agricultural productivity and indirect effects through many channels: distances from markets, quality of institutions and trade integration.

Figure 1.4 displays the chain of these effects. For geographic instruments to be valid, the potential biases coming from the effects indicated by arrows (3) and (5) need to be accounted for. With respect to (3), it is unlikely that the instruments will have a direct effect on political preferences other than via demographics. For example, it could be the case that places with unfavorable geography are underserved in terms of infrastructure and voters tend to dislike politics above average (It could as well be the other way around). However, these aspects should be reasonably captured by the characteristics of the place – such as development index, income, percentage of people in urban areas, etc – and voting behavior in previous elections. Hence, it suffices to show that I am able to control for the channel (5) of correlation.

Assuming my instrument has a long-lasting effect on development that in turn affect political preferences systematically, it is necessary to find a control that can further account for these effects up to the 2010 election. It seems reasonable that the political preferences induced by the development effects coming from geography would already show up in a previous election such that the vote shares obtained by a given party in a previous election would be a good candidate as a control. Hence, the key identification assumption here would be that conditional on previous election outcomes, geographic instruments are plausibly exogenous. I use a first differences model to account for that.

Another possible source of bias is the effect displayed by arrow (6), which may not be accounted for by the control proposed above if there are significant developmental effects resulting from Internet availability contemporaneous to the 2010 election. It means that it could be the case that voters choose a candidate because they are happier or there is economic growth, etc. To take this into consideration, I also apply first-differences in some contemporaneous demographics: income, population size, social benefits, and GDP. These variables already capture contemporaneous Internet

effects on development/growth and, thus, account for voting with economic beliefs, avoiding confoundedness with the direct Internet effects via information.

With these assumptions in place and conditional on the controls proposed above, I can plausibly estimate the effect displayed by arrow (8).

1.4.2 Model Specification

My baseline model[19] is given by:

$$y_{it} = \beta_0 + \beta_2 ISP_{it} + \rho_i + \delta_i t + \varepsilon_{it} \tag{1.1}$$

where y_{it} is the outcome of interest, namely the vote shares of each candidate/party considered. I omit the index that characterizes candidate/party for simplicity, and use i to index municipalities and t the time. ρ_i is the municipality fixed effect and $\delta_i t$ the municipality trend.

In order to control for the time invariant elements I use a first-differences specification:

$$\Delta y_{it} = \beta_1 + \beta_2 \Delta ISP_{it} + (\delta_i + \Delta\varepsilon_{it}) \tag{1.2}$$

With just two periods of data, δ_i is not identified. I therefore approximate municipalities trends by using state trends. That is, $\delta_i = \widetilde{\delta}_i + \lambda_{is}$. The assumption is that the municipality's trend can be decomposed into two components: one that is idiosyncratic to the municipality itself $\widetilde{\delta}_i$, and one that is common across municipalities within the same state λ_{is}. Note that these municipalities are in the same gubernatorial constituency. I also use robust standard errors.

[19]Since there is strategic interaction between the candidates, the errors across the equations for each candidate are correlated. Estimating a SUR model is an option, but since the controls used across equations are the same, there would not be any gains in efficiency by doing so.

I subsequently add controls at the baseline period, X_{it-1}, and time variant controls, ΔW_{it}, that include the four groups of controls described in Subsection 1.3.1: controls, other media, local support, and state dummies.

$$\Delta y_{it} = \beta_1 + \beta_2 \Delta ISP_{it} + \beta_3 X_{it-1} + \beta_4 \Delta W_{it} + \lambda_{is} + (\widetilde{\delta}_i + \Delta\varepsilon_{it}) \tag{1.3}$$

To account for unobservables that correlate both with ΔISP_{it} and the error term $(\widetilde{\delta}_i + \Delta\varepsilon_{it})$, I use Z_{it}, the distances, as an instrument in the following two-stages least squares specification:

$$\Delta y_{it} = \beta_1 + \beta_2 \Delta ISP_{it} + \beta_3 X_{it-1} + \beta_4 \Delta W_{is} + \lambda_{is} + (\widetilde{\delta}_i + \Delta\varepsilon_{ist}) \tag{1.4}$$

$$\Delta ISP_{it} = \widetilde{\beta_1} + \widetilde{\beta_3} X_{it-1} + \widetilde{\beta_4} \Delta W_{it} + \widetilde{\eta} Z_{it} + \widetilde{\lambda}_{is} + v_{ist} \tag{1.5}$$

For example, the number of ISPs tend to be higher in densely populated areas. Since densely populated areas tend to have a whole set of characteristics that correlate with political preferences, they might be substantially different from others in many unobservable time variant dimensions. The two-stages least squares specification should account for that as long as my identification assumption holds.

1.5 Results

1.5.1 Outcomes of Interest and Descriptive Statistics

I define vote shares of each candidate as outcomes of interest. In the 2010 election, nine candidates were fielded under different party denominations. I will refer to the candidates by their party acronym. I do so because when I introduce the first-

differences model, I am comparing vote shares of the same party, but possibly different candidates over time.

The two largest parties in the country, PSDB (center-right) and PT (center-left) have been consistently fielding candidates in all presidential elections since the country's redemocratization in 1989. In 2010, PSDB's candidate finished the race in second place and PT's won the election. The details of the election can be found on Table 1.5. The other parties, either support one of these two candidates, field a candidate themselves or do not compete at the presidential level. PMDB is a large (centre) party that has been conveniently supporting one of the candidates of the two parties mentioned. The candidate of *Partido Verde* (Green Party), henceforth PV, a relatively small party at that year, as measured by the number of seats in the lower chamber, fielded a candidate who finished the first round in the third position.

I classify as small/minor candidates, candidates/parties ranked 4th to 9th in this election and that had the same time of advertisement on television and radio. The sum of the vote shares of these candidates is one of the outcomes of interest. As for the three first ranked candidates, I refer to them as PT, PSDB and PV, respectively.

Descriptive statistics are displayed in Tables 1.6-1.9. In Table 1.6 it can be seen that the average differences of vote shares between 2010 and 2006 for the PT is greater for the bottom 50% municipalities, as per number of ISPs, whereas the vote for the PSDB is about the same. When we look at the vote shares of PV and small candidates, however, the reverse happens. The average differences in the vote shares between 2010 and 2006 is greater in the top 50% municipalities, as per number of ISPs.

This table also displays the distances to backbones, my instrument, which are considerably greater, on average, for municipalities at the bottom 50%.

I also look at each parties' aggregate vote shares for the Lower Chamber, to have a sense of the voting patterns of each party in a setting where candidates' quality is

averaged and campaigns were not so focused on the Internet. The same pattern is verified for the PT and PV, bottom 50% municipalities vote more often for PT and top 50% municipalities vote more often for PV. On the other hand, PSDB received on average more votes in the top 50% municipalities, whereas small candidates had more votes in the bottom 50% municipalities, getting even a negative average in the top 50% municipalities.

When we also take into account the demographics presented in Tables 1.7, these results indicate that it is probably the case that municipalities with above median number of ISPs are wealthier and have preferences for center-right political parties and poor, underdeveloped areas have preferences for center-left, populist political parties. However, the PV, as a center-left party would not have its vote shares patterns explained by this theory. Table 1.9 that describes the variables indicating local support for these parties seem to present a balance between the two groups of municipalities. Moreover, this is also verified in the vote shares of PT, PSDB and PV in the lower chamber, but for small parties, their average vote shares was negative as compared to the previous election in the municipalities above the median number of ISPs.

Interestingly, the voting patterns from the election in 2012 relative the ones in 1998 do not seem to differ much across the two groups, on average, at least for PSDB, PT and small parties. Some difference is verified only when we look at the PPS[20], the party which came in third in that election, mostly favored by bottom 50% municipalities.

Finally, municipalities' access to traditional media, namely newspapers, magazines, radio stations and TV channels that are produced locally, as well as the number of channels they have access to, are unbalanced towards the above median ISPs group (see Table 1.8). This means that differential access to other types of media is

[20] *Partido Popular Socialista*(PPS). A party that was created in 1992 by former members of the Brazilian communist party (*Partido Comunista Brasileiro-PCB*).

a potential confounder for the Internet effects. Therefore, I include all these controls for other media in my regressions.

In the next section, I try to account for the factors that may be inducing these higher vote shares for PV and small parties in municipalities with higher Internet availability: trends and heterogeneity.

1.5.2 Preliminary Analysis

In Table 1.10 I estimate the model in equation 1.3. The results indicate that vote shares for PV and small parties were significantly larger in municipalities where the penetration of Internet was higher. Interestingly, this increase seem to come from the reduction in the PT's share of votes.

Since PV and some of the small parties had no candidates running for the 2006 elections, I assume that the vote shares of similar parties in 2006 are a good proxy for the vote shares these parties would have got, had they fielded candidates. For example, I consider the vote shares that the third-placed party got in the previous election as the baseline for PV votes.

Given that there is a large proportion of zeros in the sample of vote shares for small candidates and because of the problems caused by assuming a linear conditional mean for fractional dependent variables (see Papke and Wooldridge (1996) for details) I also estimate non-linear models: Poisson[21] and Probit (not reported). There is no variation in the magnitude of the coefficients for Internet availability. The linear estimates are visibly not different from the non-linear estimates, so I use only the linear models for simplicity.

These results are robust to a number of controls, time-invariant unobservables and state trends. However, there can be remaining factors that vary over time fol-

[21] A Poisson model in this context is equivalent to assuming the conditional mean to be $E(y|x) = exp(x\beta)$ and estimating the parameters via non-linear least squares or GMM (see Windmeijer and Silva (1997) for more details).

lowing Internet availability growth that can confound our results. For example, if municipalities where Internet growth was higher enjoyed better economic growth, as a consequence or not of Internet penetration (see Czernich and Falck (2011) for the effects of Internet on growth), and considering that this growth might involve many dimensions that are not captured by GDP, voters could attribute these economic effects to the performance of the incumbent party (PT) in which case the coefficient of the effect of Internet on the PT votes would be upward biased, that is, less negative. And the coefficient for PV would probably be downward biased.

To tackle this sort of confounding mechanisms, I use an instrumental variables approach. The identification assumption is that, distances to backbones, the system of main network cables to which municipalities need to be connected in order to have access to the Internet, are plausibly exogenous, conditional on a set of demographic controls. Since the construction of backbones was facilitated in places where pre-existing infrastructure was available, we can assume this path was pre-determined, making sure to control for long lasting effects of infrastructure by adding socioeconomic controls for the baseline year.

1.5.3 Two-Stage Least Squares

Now, I turn my attention to the IV regressions in Table 1.12. As discussed previously, the estimates are likely to be biased in OLS. We can find stories that rationalize both downward and upward bias. For example, it could be the case that voters more interested in politics self select into more densely populated areas in which the provision of Internet is higher because companies exploit economies of scale in these areas. Conversely, these voters could experience a crowding out of information in these places and/or the availability of more amenities could be incentives to devoting less time to acquiring political information.

My key assumption in the IV specification is that my instruments are conditionally correlated with Internet availability, via increasing costs of provision and the existence of complementary infrastructure, but not correlated with unobserved municipality level characteristics that are related to voting behavior. The major concern about the validity of the exclusion restriction is that even though the deployment of the Internet cables network used pre-existing infrastructure, providers used only those which were closer to more dense areas with favorable demographic characteristics for high demand. So, I include the measures of infrastructure themselves to account for this. I also use further controls such as urban population and income.

Table 1.11 shows the first-stage and reduced form regressions. In column (5) the reduced form regression is displayed. As expected, log distances are significantly negatively correlated with the change in the number of ISPs. Columns (1) through (4) display the first-stage estimates for each of the parties or group of parties analyzed. Log distances correlates positively with PT and PSDB vote and negatively with both PV and small parties. That is, it displays the same pattern of correlations with Internet availability. The F-statistic for excluded instruments, 17.6, indicate log distances are a relevant instrument, relieving concerns about finite sample bias in the IV estimates.

The results of the IV regressions are presented in Tables 1.12 and 1.13. In Table 1.12 I assume $PV_{t-1} = 0$, that is, I consider the PV vote shares in levels and in 1.13 I assume $PV_{t-1} = PPS_{t-1}$, that is, vote shares of the third-placed party in 2006 is a proxy for the vote shares PV would have received had it fielded a candidate. As in the OLS estimations, they show both minor and third-placed candidates benefited from Internet availability. It also had a comparatively high and significant negative effect on the vote shares of the pro-government candidate.

The coefficient for the PV vote in the more conservative specification, Table 1.13, is approximately 0.5 percentage point, a 6.3% increase (on average) in her vote shares.

For the aggregate of small candidates, there was a 7.8% increase (on average) in vote shares. In absolute terms, one additional ISP increased their vote shares by about 0.04 percentage point.

For the main candidates, it is a different picture. While the Internet effect is negative and significant for the winner, approximately -0.5 percentage point for each additional ISP (-25% on average), they are not statistically significant for the runner-up. Its relevant to point out here that my instrument does not predict well the PSDB vote shares as it can be seen in Table 1.12, column (2), so I cannot test very precisely the assumption that Internet availability affected the vote shares of this party.

The differences found in magnitudes of the coefficients in OLS and 2SLS call for a discussion here. It is probably the case that OLS gives more weight to municipalities where the rise of the so called class C in Brazil – a group of households that transitioned from lower to middle class – and their endogenous demand for services contributed significantly to the growth in Internet availability, as measured by the number of ISPs. The victories of PT's candidates in Brazil are anecdotally attributed to the vote of this new middle class that related their better economic conditions to the incumbent government. So the effect of Internet on vote shares would be upward biased for the PT's candidate and, by complementarity, downward biased for PV's and small candidates. The IV approach allocates more weights to the vote in those municipalities in which Internet availability was more affected by technical factors than lack of demand. As a result, PT's vote share is more negative, that is, it is net of the vote shares of those municipalities who had high Internet availability but were already voting for PT due to other reasons. The effects on PV's and small candidates' votes follow by complementarity.

Previous literature has also found that OLS leads to downward bias in the effects of Internet on election outcomes, and normally the size of this bias is very large, even when there are many periods of time available for inference and cities' trends

are accounted for more precisely (see for example Campante, Durante and Sobbrio (2013)). To further compare the results found with the previous literature, I calculate persuasion rates as in DellaVigna and Gentzkow (2010). Basically, in order to compare these effects we need to take into account the differences in the size of the group for which the treatment is applied and the intensity of treatment. That's what this measure tries to do. The upper bound of the persuasion rate (where I consider an Internet penetration of 42%, a lower bound would consider full penetration.) for the positive message is given by $f_+^u = \frac{7.3}{1-0.04*0.81} * \frac{0.0053}{0.42} * 100 = 9.5\%$ and for the negative message $f_-^u = \frac{-7.3}{0.38*0.81} * \frac{0.0047}{0.42} * 100 = -26.4\%$. This indicates that about 9.5% of the treated voters were persuaded to vote for small parties and 26.4% to not vote for the incumbent party (PT). In the literature, the range for positive measures goes from 4% to 20% and for negative messages from 14% to 65.4%.

Placebo Test

To provide evidence of the validity of the exclusion restriction assumed with respect to the instrument, I conduct a placebo experiment using the outcomes of the presidential elections of 2002 and 2006, a period where we do not expect Internet availability in 2010 would play an important role, as most of the expansion in its use in politics happened in the 2010 election only. I use the same specification of the previous models and just plug in the equivalent outcomes for the previous elections.

This placebo test serve two purposes. First, to demonstrate that (log) distances are not correlated with unaccounted confounding factors that would be also present in the 2006 election and produce the same correlations we see in the 2010 election. Second, even though there was already a reasonable level of Internet penetration in 2006, its effect in candidates' vote shares was not significant, probably due to the fact that candidates were not using it for political campaigns as intensely as they did in 2010.

To be more conservative, I use the levels of vote shares for those parties who did not field a candidate in 2002 (the coefficient is insignificant if we consider differences as before too). Results are shown in Table 1.14. The variation in the number of IPS cannot explain the variation in vote shares for PT, PSDB and small parties. As for the third-placed party, the coefficient on ISPs is still significant at the 10% level. So this significance could be an indication that either there is some remaining endogeneity or that Internet had already some effect in 2006, even when not being used as intensely for political campaigns, but as a source of news content only.

To test this hypothesis, I also conduct a placebo using the 2002 and 1998 elections' outcomes. The results are presented in Table 1.15. Parties' are displayed in columns following their ranking in the election. Now, all the coefficients are statistically insignificant, providing evidence in favor of the hypothesis that Internet had some effect on the votes of the third-placed candidate in 2006.

To sum up, these results provide evidence that my specification might not account for all the relevant covariates that relate to the voting outcomes, but the instruments are plausibly valid and account for these omitted variables.

1.5.4 Turnout and Blank Votes

Now I turn my attention to the Internet effects on turnout and blank voting. As previously described, voting is compulsory for citizens aged 18-70 years. For 16 and 17 year olds, and those above 70 years of age the vote is voluntary. This compulsoriness seems to hold well, since turnout has been around 80% on average in presidential elections. So we should expect that if anyone's decision on whether to vote or not is affected by Internet, it should especially be 16 and 17 year olds. Table 1.16, Column(2), displays the results for overall turnout. It shows Internet availability did not significantly affect aggregate turnout. The age composition of the voting population across municipalities is very similar, especially for the group of

16-17 year olds, according to the census of 2010. When I include the share of voters of this age in the regression and interact with Internet availability, I still find no statistically significant effects (not reported). This provides evidence that this group of voters already had a high turnout. Probably due to the fact that their parents are required to vote and by some imitation mechanism they also vote, even when they are not required to do so.

In any case, these results indicate that the effects of Internet availability verified on voting are likely to be coming from voters switching party lines, rather than from a recomposition of the electorate that turns out to vote.

Table 1.16, Column(1), shows that Internet availability did not have any statistically significant impact on the variation of the share of blank votes across municipalities. This is an important result to the extent that it indicates Internet availability had no effect on what could potentially be protest voting. If it had any, it was through voting for other candidates, perhaps small candidates. However, note that my identification strategy takes care of any fixed effects in protest voting, and the IV should account for whatever factors that are correlated with both Internet availability and the presence of voters that are likely to protest, that is independent of Internet use itself. Thus, if those who vote for small candidates are doing so as a form of protest, it was induced by the Internet.

1.5.5 Individual-Level Results

The analysis of individual-level data allows me to (i) evaluate whether the results obtained at the aggregate level can be replicated at the individual level; (ii) infer about voters' usage of Internet for political purposes; and (iii) have an idea of what voters' level of information was before and after the period of political campaigns in the media.

Verifying the existence of similar effects at the individual level provides evidence that the aggregate variation in voting shares across municipalities can be used for inference about voting behavior. More than that, it provides evidence that the model specification used is able to reasonably account for heterogeneity across municipalities. Since I have used the concept of Internet availability at the aggregate level, which is important only to the extent that it correlates to effective Internet usage, it is important to evaluate the magnitude of this correlation at the individual level.

Having surveys conducted at the beginning and at the end of the campaign period may suggest whether Internet effects are primarily driven by the information effects, in which case it is likely we will find similar effects both before and after the campaign period, or propaganda effects, a result of the campaigning conducted by the candidates.

The period when candidates run their political campaigns goes from July to September. I analyze the results of a survey conducted at the beginning of the campaign period, July, and at the end, September. I start by the results of September's survey. In this survey, voters were asked to rank a list of sources of political information in a scale from "very important" to "unimportant" in terms of providing information about the candidates running for the Lower Chamber. From this question I create the dummy "Internet", that takes 1 for those who ranked Internet as being either a "very important" or an "important" source. The assumption made here is that voters who use the Internet to obtain information about candidates in the Lower Chamber election, which occurs simultaneously with the presidential election, are more likely to do the same for the presidential election. The other variables that have to do with media are created in the same fashion.

I also add controls for affiliation with the three main parties (PT, PSDB, and PV), and demographics: income, unemployment, employment in the public sector, age, gender, education and income. The Linear Probability Model in Table 1.17 replicates

the qualitative results found in the aggregate data for the three main parties: positive for PV, negative for PT and insignificant for PSDB. On the other hand, the results for small candidates are not significant. Perhaps because the effect for these candidates is undetectable when looked at a sample of this size. Importantly, note that the traditional media is relevant in explaining the likelihood of voting for PT only. A different pattern is verified in the dummy variable that is one when the voter watches the campaigns broadcast on TV or radio: it is significantly positive for PT vote shares and significantly negative for PV vote shares.

As verified earlier in the aggregate results, OLS provides downward biased estimates for the Internet coefficient. In order to use the same identification strategy applied at the aggregate level analysis, I match voters' municipality to the number of existing ISPs to obtain the effects of Internet availability at the individual level. The results of a linear probability model (LPM), where the outcome variable is a dummy indicating whether a voter reports voting for a given party, is presented in Table 1.18. With this specification, the coefficients are also qualitatively similar to the ones obtained at the aggregate level, except for the significantly positive effects for PSDB.

In Table 1.19 I report the results of a two-stages least squares model where I also instrument Internet availability by distances. Essentially, each additional ISP in a municipality implies that a voter is on average less likely to vote for the PT (-0.142) and more likely to vote for the PV (0.028) and small parties (0.009).

When we look at the July's survey, in which I also have information as of whether voters had an Internet subscription at home, 13% of voters who reported using the Internet for political purposes did not have Internet at home, whereas 49% of those who had Internet at home used it for the acquisition of political information. This sheds light on the degree of compliance to the treatment when we consider Internet availability. It is important to highlight that at the time at which this survey was

applied, July 7-23, 2010, campaigns on TV and radio had yet to start (August 17, 2010). The first debate was held only on August 5, 2010. This allows us to infer the effects of Internet before the campaign started, that is, effects that are mostly free of online advertisement by candidates.

Table 1.20 displays the IV-2SLS estimations for these data. As before, an additional ISP would lower the probability of voting for PT and increase the probability of voting for PSDB. Different from September's results, the coefficients here are smaller and PSDB vote is significantly affected by Internet. Obviously, part of the differences found here may be just given to sampling variation. But the differences in the results are too large to be attributed just to this factor, as a two sample test reveals. PV's coefficient is also smaller and not significant. Another important fact is that, an additional ISP also increases the probability of voting blank, for which the coefficient has about the same magnitude. These results suggest that the political campaign period in the media, especially online advertising, seem to have changed vote choices in favor of PV's and against PT and PSDB's.

1.6 Conclusion

I analyze the effects of Internet availability on voting outcomes of the 2010 presidential election in Brazil. Identification of causal effects is possible by exploiting the variation in Internet availability across municipalities and using distances to the core section of fiber-optic cables (backbones) that connects municipalities to the regional and international network of cables that constitute the Internet. The identification assumption is that conditional on economic and demographic variables, infrastructure is a pre-determined variable and thus plausibly independent of unobserved factors that jointly determine political preferences and Internet availability.

With aggregate data, I find that the effect of Internet availability was positive and significant for both small parties and the third-placed party. One additional ISP in a municipality increases the vote for small parties by 0.04 percentage point and the vote for the third-placed party by 0.48 percentage point. This result is robust to a series of important controls: demographics, geography, information availability, candidates' support networks and candidates' visits. In addition, a placebo experiment provides evidence in favor of the exogeneity of the instrument used and that the Internet effect was potentialized by online political campaigns.

Individual level data confirm the direction of these effects and provide evidence of some of the mechanisms composing them. (1) Internet availability is a reasonable proxy for Internet use. (2) The effects verified at the beginning of the campaign period reveal an informational effect of Internet availability that is to a certain extent free from interactions with online and offline political campaign. (3) Though, a stronger effect of Internet availability on vote shares is found at the end of the three-month-long campaign period and (4) Voters who report Internet as an important source of information on average have reduced their consumption of other more traditional media.

Further checks show Internet had no causal effect on blank voting or turnout. Hence, it is likely that the persuaded voters switched parties as a result of the additional information available (if we hold constant the flow of new voters).

A small theoretical literature has called attention to the importance of elections that allow for voters to communicate their political preferences that is independent of the choice of the candidate that they actually vote for (see Piketty (2000) and Castanheira (2003)). That is, the vote in the first round in runoff elections would allow for communication and in the second round the voter would actually vote for the candidate who is closer to her preferences. Voters normally vote for small, more extreme candidates to send a message to mainstream politicians about policy posi-

tions. There are two effects that the Internet use would produce in this respect: (1) the possibility of communication and mobilization voters themselves have through social networks would reduce the necessity of communicating preferences though voting; and (2) more knowledge about small, more extreme candidates, would persuade more voters to vote for them as a form of communicating. The results found here point to a stronger second effect.

This suggests that the Internet, both as an additional source of information or as a new advertising medium for candidates, had an important causal effect on the votes of candidates that did not advertise as much in the traditional media. A playing field that fostered competitiveness among candidates. The mechanisms through which this might happen have important policy implications in terms of media related regulatory policies for campaign advertisement, even though not providing clear-cut recommendations.

Internet informational effects and low costs (at least a priori) might eliminate barriers to entry for new candidates, but we cannot say much about voters' implied welfare effects. It could be the case that now candidates of higher quality are allowed to run for office, or that voters can identify more easily the candidates that are closer to their preferences. Alternatively, having more candidates running for office and proportionally more information about them, both positive, negative or false, may just reduce the signal to noise ratio on information and increase the amount of effort required in order to make an informed choice. These are questions for further consideration.

It is fair to consider that the campaigning effects are also bounded by the persuasive abilities of the candidates, as well as their valence and political platforms. Since this was an election where this media was first used, there might exist some learning curve and its effects fade with time. The dynamics of this type of effect will only be verified as more data become available.

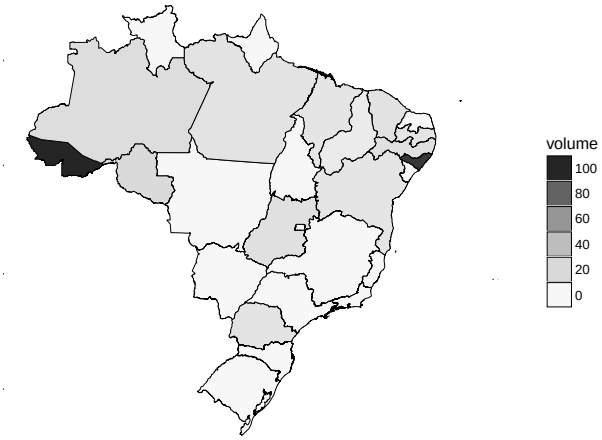
1.7 Appendix

1.7.1 Figures

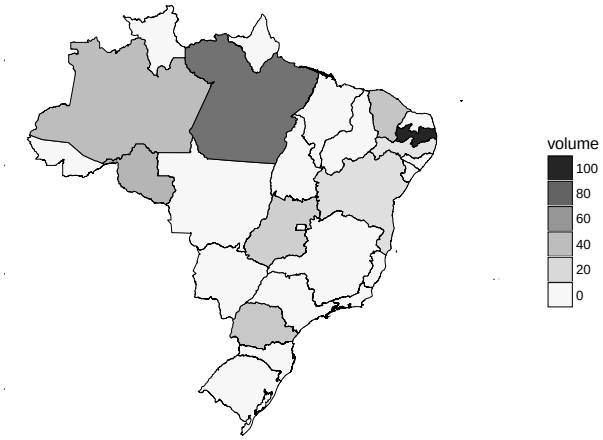
Figure 1.1: Web search volume (by state) of content in politics related to each candidate.

(a) Search volume for content in politics related to PT's candidate (Dilma Roussef).

(b) Web search volume for content in politics related to PSDB's candidate (Jose Serra).



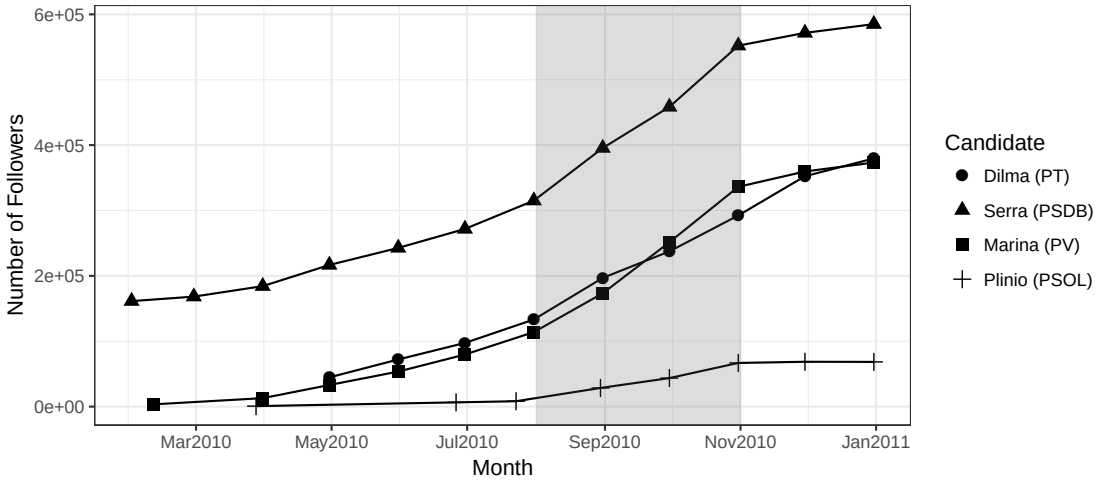
(c) Web search volume for content in politics related to PV's candidate (Marina Silva).



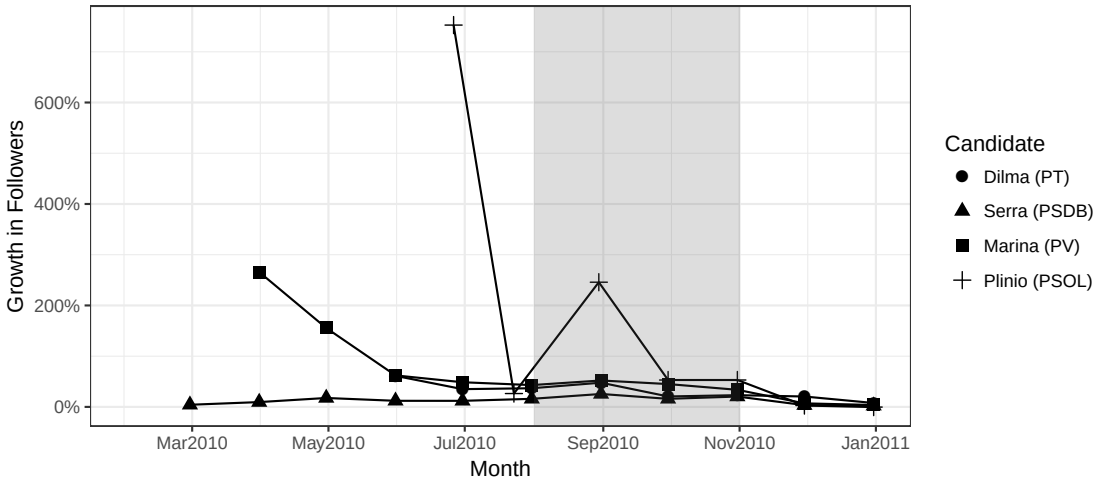
(d) Web search volume for content in politics related to PSOL's candidate (Plinio de Arruda)

Note: Search volume by state of content in politics related to each candidate, as defined by Google itself. The state with the highest search volume is given 100 and the search volume of the other states is a relative measure. These data were obtained from public Google Trends data available online.

Figure 1.2: Candidates' Twitter activity: followers.



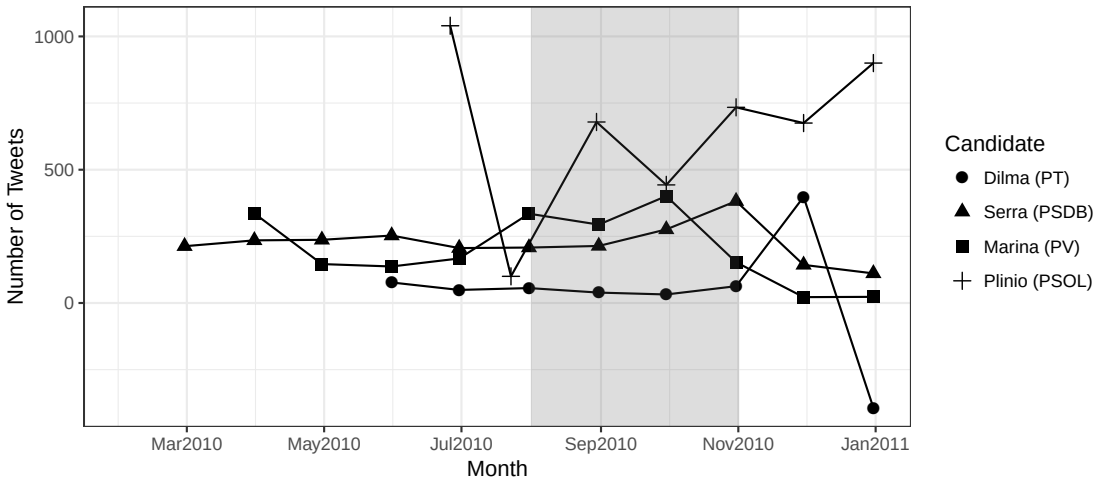
(a) Candidates' number of followers on Twitter over time.



(b) Growth in the number of Candidates' followers on Twitter over time.

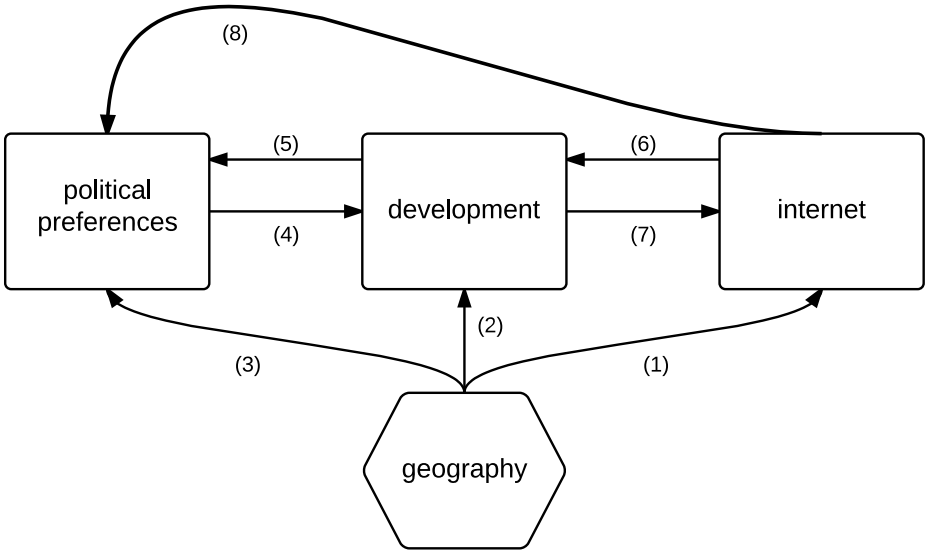
Note: Data obtained from *Twittercounter.com*. The grey area defines the election campaign period. PSDB's candidate created a Twitter account before all the other candidates on May 20, 2009. For the other candidates, data are presented from the date they created an account.

Figure 1.3: Candidates' Twitter activity: tweets.



Note: Data obtained from *Twittercounter.com*. The grey area defines the election campaign period. PSDB's candidate created a Twitter account before all the other candidates on May 20, 2009. For the other candidates, data are presented from the date they created an account.

Figure 1.4: Identification Strategy.



Source: Adapted from Rodrik, Subramanian and Trebbi (2004).

Figure 1.5: Correlation between number of lines (per 10 inhabitants) and number of ISPs.

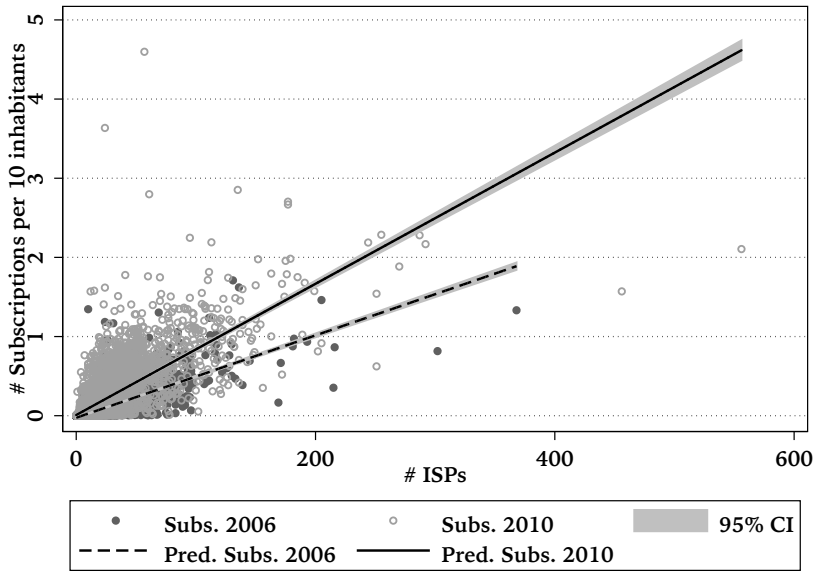
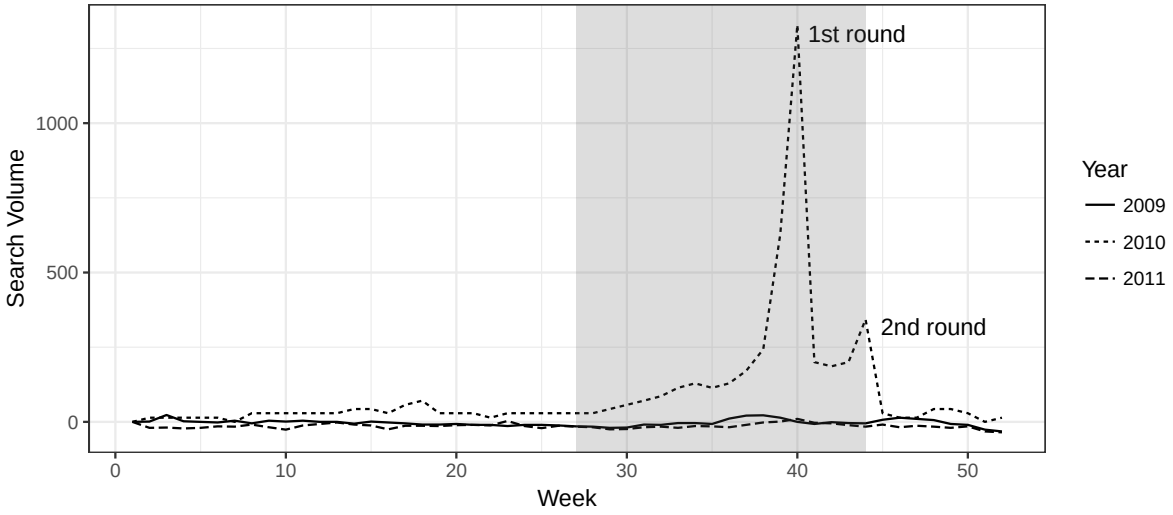


Figure 1.6: Web search volume for terms related to politics on Google Trends.



Note: Web search volume for terms related to politics, as defined by Google itself. The grey area defines the election campaign period. These data were obtained from public Google Trends data available online.

1.7.2 Tables

Table 1.1: Distribution of free broadcast time (in seconds) on TV and Radio in the first round.

Ranking	Time on FTA TV	Time on Radio
1st	638	153
2nd	438	105
3rd	83	19
4th	61	14
5th	55	13
6th	55	13
7th	55	13
8th	55	13
9th	55	13

Note: Data provided by the electoral authority (TSE). Time displayed in seconds.

Table 1.2: Descriptive Statistics for computer ownership and Internet.

	Fraction Total Population	Population
No computer, no internet	0.58	76195260.22
Computer, no internet	0.08	10749440.22
Computer, internet	0.34	45260435.24
Total	1.00	$1.32 \times 10^{+08}$

Note: Data extracted from 2010 Census.

Table 1.3: Correlation between ISPs and Internet Lines.

	(1) Voters Internet	(2) Subs. 2006	(3) Subs. 2010	(4) Δ Subs.
ISPs 2010	0.337***	–	816.581***	–
	[0.027]		[304.392]	
ISPs 2006	–	556.175**	–	–
		[241.789]		
Δ ISPs	–	–	–	991.882***
				[340.851]
Observations	5545	5545	5545	5545
R^2	0.327	0.221	0.303	0.293

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: Standard errors are clustered at the state level. Region fixed effects are also added. The measures of Internet used are respectively (1) the share of voters who reported having Internet access at home in the Census 2010, (2) the number of Internet subscriptions reported by the ISPs in the first term of 2007 to the Anatel, (3) the number of Internet subscriptions reported by the ISPs in the third term of 2010 to the Anatel, and (4) the difference in the number of subscriptions between 2010 and 2007.

Table 1.4: Correlations between measures of Internet.

	Computer	Internet	#ISPs
Computer	1		
Internet	0.501	1	
# ISPs	0.752	0.372	1

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: Standard errors are clustered at the state level. Region fixed effects are also added. The variable Computer is the share of voters who reported have Internet access at home, and Internet is the share of voters who reported having both computer and Internet at home in the Census 2010.

Table 1.5: Overall results of first and second round of the 2010 presidential election.

			1st Round		2nd Round	
Ranking	Candidate	Party	Number of Votes	Percentage	Number of Votes	Percentage
1	Dilma Rousseff	PT	47651434	46.91%	55752529	56.05%
2	José Serra	PSDB	33132283	32.61%	43711388	43.95%
3	Marina Silva	PV	19636359	19.33%	–	–
4	Plínio Sampaio	PSOL	886816	0.87%	–	–
5	Eymael	PSDC	89350	0.09%	–	–
6	José Maria	PSTU	84609	0.08%	–	–
7	Levy Fidélix	PRTB	57960	0.06%	–	–
8	Ivan Pinheiro	PCB	39136	0.04%	–	–
9	Rui Pimenta	PCO	12206	0.01%	–	–
	Valid votes		101590153	91.36%	99463917	93.30%
	Blanks		3479340	3.13%	2452597	2.30%
	Spoilt		6124254	5.51%	4689428	4.40%
	Total		111193747	81.88%	106606214	78.50%
	Abstentions		24610296	18.12%	29197152	21.50%
	Registered Voters		135804433	100.00%	135804433	100.00%

Note: The table displays number and percentage of votes for all candidates in the election in both first and second rounds. It also displays details about turnout, blank and spoilt votes. Data are provided by the Electoral Authority (TSE).

Table 1.6: Descriptive statistics for vote shares.

	Top 50% # ISPs			Bottom 50% # ISPs		
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
Change in Vote Shares 2010-2006						
Δ PT	0.66	7.56	2578	2.97	8.31	2600
Δ PSDB	−7.17	5.81	2578	−7.33	6.31	2600
PV	9.68	5.42	2578	5.85	3.22	2600
Δ PV	6.54	4.26	2578	3.96	2.98	2600
Δ Small	0.61	0.31	2578	0.47	0.35	2600
Instruments						
Distances (Km)	58.67	78.03	2578	97.47	114.49	2600
Log Distances	3.65	0.92	2578	4.20	0.86	2600
Change in Vote Shares 2010-2006 Parliament						
Δ PT	0.83	18.59	2578	7.90	20.63	2600
Δ PSDB	6.52	17.06	2578	2.61	20.04	2600
Δ PV	7.47	6.68	2578	4.37	5.36	2600
Δ Small	−0.01	1.50	2578	0.20	0.82	2600
Change in Vote Shares 2002-1998						
Δ PSDB	15.30	7.14	2547	13.90	7.17	2599
Δ PT	−15.20	7.94	2547	−12.41	9.79	2599
Δ PPS	1.74	4.59	2547	3.15	6.36	2599
Δ Small	0.04	0.07	2547	0.04	0.08	2599

Note: The table shows summary statistics for the groups of municipalities that are above and below the median number of ISPs in 2010. The unit of observation is a municipality-year.

Table 1.7: Descriptive statistics for controls.

	Top 50% # ISPs			Bottom 50% # ISPs		
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
Catholics	0.36	0.07	2578	0.40	0.06	2600
College	0.03	0.02	2578	0.02	0.01	2600
Migrants	0.29	0.09	2578	0.32	0.09	2600
Black	0.06	0.04	2578	0.07	0.05	2600
Urban	0.45	0.03	2578	0.44	0.03	2600
Unemployment	0.09	0.04	2578	0.07	0.04	2600
Inequality	0.52	0.05	2578	0.52	0.06	2600
Median HH Income	6.64	0.39	2578	6.30	0.41	2600
GDPpc 2007	11.61	13.60	2578	6.79	5.74	2600
GDP 2009-2008	1.08	0.13	2578	1.09	0.16	2600
Δ Age 16-17	−0.01	0.00	2578	−0.01	0.00	2600
Δ Age 70+	0.01	0.01	2578	0.01	0.01	2600
Δ Child Benefit	1.13	0.34	2578	1.14	0.33	2600
Δ Female	0.00	0.01	2578	0.00	0.01	2600
Δ Voters	0.07	0.07	2578	0.06	0.09	2600
Log Pop. Density 2006	4.37	1.36	2578	3.18	0.93	2600

Note: The table shows summary statistics for the groups of municipalities that are above and below the median number of ISPs in 2010. The unit of observation is a municipality-year.

Table 1.8: Descriptive statistics for other media.

	Top 50% # ISPs			Bottom 50% # ISPs		
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
TV	0.84	0.08	2578	0.74	0.13	2600
Radio	0.79	0.08	2578	0.73	0.13	2600
I(local newspapers)	0.61	0.49	2578	0.14	0.35	2600
I(local magazines)	0.15	0.36	2578	0.01	0.09	2600
I(Local Radio Station)	0.82	0.38	2578	0.53	0.50	2600
I(Local TV Channel)	0.17	0.38	2578	0.04	0.20	2600
# TV channels received	2.22	0.65	2578	1.78	0.68	2600

Note: The table shows summary statistics for the groups of municipalities that are above and below the median number of ISPs in 2010. The unit of observation is a municipality-year. The variables TV and Radio are the share of voters who reported having a TV or radio set at home in the Census.

Table 1.9: Descriptive statistics for local support.

	Top 50% # ISPs			Bottom 50% # ISPs		
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
I(PT Visit)	0.04	0.65	2578	0.00	0.00	2600
I(PSDB Visit)	0.05	0.74	2578	0.00	0.02	2600
I(PV Visit)	0.04	0.78	2578	0.00	0.00	2600
Δ Partisans PT	0.21	0.09	2578	0.20	0.11	2600
Δ Partisans PSDB	0.18	0.08	2578	0.19	0.10	2600
Δ Partisans PV	0.01	0.02	2578	0.01	0.02	2600
Δ Partisans Small	0.02	0.03	2578	0.01	0.03	2600
I(Mayor PT)	0.33	0.47	2578	0.31	0.46	2600
I(Mayor PSDB)	0.23	0.42	2578	0.23	0.42	2600
I(Mayor PV)	0.02	0.12	2578	0.01	0.11	2600

Note: The table shows summary statistics for the groups of municipalities that are above and below the median number of ISPs in 2010. The unit of observation is a municipality-year.

Table 1.10: OLS with Controls.

	(1) Δ PT	(2) Δ PSDB	(3) Δ PV	(4) Δ Small
Δ ISPs	−0.050***	0.009	0.058***	0.003***
	[0.012]	[0.010]	[0.008]	[0.001]
Controls	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes
Other Media	Yes	Yes	Yes	Yes
Local Support	Yes	Yes	Yes	Yes
Observations	5445	5445	5445	5445
R^2	0.568	0.407	0.695	0.552

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Table 1.11: Reduced Form and First-Stage.

	(1) Δ PT	(2) Δ PSDB	(3) PV	(4) Δ Small	(5) Δ ISPs
Log Distances	0.239**	0.157*	−0.397***	−0.021***	−0.508***
	[0.107]	[0.093]	[0.054]	[0.005]	[0.121]
Constant	−5.884	3.649	−14.141	−3.239***	−232.230***
	[9.888]	[7.116]	[10.053]	[0.648]	[34.640]
Controls	Yes	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes	Yes
Other Media	Yes	Yes	Yes	Yes	Yes
Local Support	Yes	Yes	Yes	Yes	Yes
Observations	5445	5445	5445	5445	5445
R^2	0.567	0.407	0.694	0.552	0.604

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Table 1.12: IV-2SLS using $PV_{t-1} = 0$.

	(1) Δ PT	(2) Δ PSDB	(3) PV	(4) Δ Small
Δ ISPs	−0.470**	−0.309	0.781***	0.042***
	[0.231]	[0.197]	[0.201]	[0.013]
Controls	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes
Other Media	Yes	Yes	Yes	Yes
Local Support	Yes	Yes	Yes	Yes
Observations	5445	5445	5445	5445
R^2	0.484	0.324	.	0.146
$F - Exc.Inst.$	17.620	17.620	17.620	17.620
mean	1.88	−7.25	7.70	0.54

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: In Column (3) the dependent variable is PV's vote shares in 2010. The last line of the table displays the sample mean of the differences in vote shares between 2010 and 2006.

Table 1.13: IV-2SLS using $PV_{t-1} = PPS_{t-1}$.

	(1) Δ PT	(2) Δ PSDB	(3) Δ PV	(4) Δ Small
Δ ISPs	−0.470**	−0.309	0.488***	0.042***
	[0.231]	[0.197]	[0.145]	[0.013]
Controls	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes
Other Media	Yes	Yes	Yes	Yes
Local Support	Yes	Yes	Yes	Yes
Observations	5445	5445	5445	5445
R^2	0.484	0.324	0.211	0.146
$F - Exc.Inst.$	17.620	17.620	17.620	17.620
mean	1.88	−7.25	7.70	0.54

Note: PPS was the party which came in third in the 2006 election. So I take it as a proxy for the share of votes PV would have got if it had fielded a candidate in that election. The last line of the table displays the sample mean of the differences in vote shares between 2010 and 2006. The unit of observation is a municipality-year.

Table 1.14: Period 2002-2006.

	(1) Δ PT	(2) Δ PSDB	(3) PSOL-PSB	(4) Small
Δ ISPs	−0.024	−0.027	0.015*	0.003
	[0.018]	[0.017]	[0.009]	[0.008]
Controls	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes
Other Media	Yes	Yes	Yes	Yes
Local Support	Yes	Yes	Yes	Yes
Observations	5445	5445	5445	5445
R^2	0.794	0.632	0.656	0.672

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: In Column (3) the dependent variable is the vote shares of the third-placed candidate in the 2006 election in levels. As in 2010, this party did not field a candidate in 2002. I do this to guarantee that the results are not being driven by the fact that the third-placed candidate in 2002 is a bad proxy to take differences from. I do the same for small parties in Column (4). The unit of observation is a municipality-year.

Table 1.15: Period 1998-2002.

	(1) Δ PSDB	(2) Δ PT	(3) PPS	(4) Small
Δ ISPs	0.015	−0.018	0.012	0.004
	[0.012]	[0.016]	[0.009]	[0.008]
Controls	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes
Other Media	Yes	Yes	Yes	Yes
Local Support	Yes	Yes	Yes	Yes
Observations	5376	5376	5376	5376
R^2	0.538	0.425	0.506	0.671

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: In Column (3) the dependent variable is the vote shares of the third-placed candidate in the 2002 election in levels. As in 2010, this party did not field a candidate in 1998. I do this to guarantee that the results are not being driven by the fact that the third-placed candidate in 1998 is a bad proxy to take differences from. I do the same for small parties in Column (4). The unit of observation is a municipality-year.

Table 1.16: Turnout and Blank Voting.

	(1) Δ Blanks	(2) Δ Turnout
Δ ISPs	0.001	0.000
	[0.001]	[0.005]
Controls	Yes	Yes
State Dummies	Yes	Yes
Other Media	Yes	Yes
Local Support	Yes	Yes
Observations	5445	5445
R^2	0.249	0.652

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: The turnout share is measured as the number of voters who turned up to voter over the total number of registered voters. The unit of observation is a municipality-year.

Table 1.17: LPM for September's Survey: Internet Use.

	(1) PT	(2) PSDB	(3) PV	(4) Small	(5) Blank
Internet	−3.417***	0.603	1.916***	−0.191	−0.608
	[0.976]	[0.896]	[0.680]	[0.243]	[0.369]
Radio-tv-newspaper	2.377**	−0.665	0.747	−0.032	−0.324
	[1.059]	[0.972]	[0.739]	[0.252]	[0.445]
Word of mouth	3.124***	−0.509	0.866	0.103	−2.188***
	[0.984]	[0.906]	[0.686]	[0.233]	[0.408]
Groups	1.211	−3.790***	−0.713	0.499*	0.430
	[1.024]	[0.936]	[0.688]	[0.258]	[0.372]
Campaign tv-radio	3.863***	1.176	−3.892***	−0.038	−1.465***
	[1.027]	[0.940]	[0.715]	[0.261]	[0.388]
Part. Small	−27.675***	−25.158***	42.795***	19.525***	−3.975*
	[4.861]	[3.307]	[7.520]	[6.211]	[2.326]
Part. PV	−30.540***	−19.907***	61.431***	0.183	−4.256***
	[2.279]	[2.226]	[3.149]	[0.930]	[0.963]
Part. PSDB	−29.029***	45.438***	−5.159***	−0.484	−4.480***
	[1.443]	[1.787]	[1.170]	[0.407]	[0.527]
Part. PT	32.259***	−15.997***	−5.436***	−0.571**	−3.744***
	[0.946]	[0.793]	[0.590]	[0.223]	[0.344]
Pub. sector job	0.251	−2.660	3.803***	0.410	−0.343
	[1.862]	[1.671]	[1.437]	[0.508]	[0.753]
Unemployed	2.794	−3.158	0.413	−0.909*	1.712
	[3.188]	[2.730]	[2.059]	[0.471]	[1.572]
Age	−0.082***	0.086***	−0.060***	−0.006	−0.021**
	[0.027]	[0.025]	[0.017]	[0.007]	[0.010]
Gender	−6.535***	1.996**	1.675***	−0.112	0.101
	[0.867]	[0.783]	[0.578]	[0.219]	[0.369]
University degree	−6.488***	−0.445	7.540***	−0.374	0.638
	[1.553]	[1.460]	[1.389]	[0.395]	[0.764]
High income	−9.907***	6.666***	4.206***	−0.213	2.361**
	[1.726]	[1.742]	[1.510]	[0.431]	[0.955]
Constant	52.201***	23.923***	11.622***	1.852***	8.370***
	[1.997]	[1.805]	[1.305]	[0.516]	[0.891]
Observations	11630	11630	11630	11630	11630
R^2	0.148	0.104	0.091	0.012	0.017

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: These estimations use data from a nationally representative survey (DATAFOLHA, 2010*a*) held in September 2010. The variable Internet is a dummy which takes value one for voters who reported Internet as being their most important source of political information. The unit of observation is a voter, that is, a person that is at least 16 years old.

Table 1.18: LPM for September's Survey: Δ ISPs.

	(1) PT	(2) PSDB	(3) PV	(4) Small	(5) Blank
Δ ISPs	−0.059***	0.029***	0.017**	0.008***	0.023***
	[0.011]	[0.010]	[0.008]	[0.003]	[0.005]
Radio-tv-newspaper	2.048*	−0.505	0.822	−0.067	−0.483
	[1.045]	[0.958]	[0.732]	[0.247]	[0.440]
Word of mouth	2.495**	−0.417	1.249*	0.112	−2.129***
	[0.980]	[0.904]	[0.684]	[0.231]	[0.409]
Groups	0.336	−3.612***	−0.236	0.496*	0.432
	[1.010]	[0.922]	[0.672]	[0.260]	[0.370]
Campaign tv-radio	2.749***	1.384	−3.269***	−0.061	−1.523***
	[1.002]	[0.916]	[0.688]	[0.248]	[0.379]
Controls	Yes	Yes	Yes	Yes	Yes
Infrastruture	Yes	Yes	Yes	Yes	Yes
Observations	11660	11660	11660	11660	11660
R^2	0.152	0.105	0.096	0.012	0.021

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: These estimations use data from a nationally representative survey (DATAFOLHA, 2010*a*) held in September 2010. All the regressions include sociodemographic controls at the municipality level. The unit of observation is a voter, that is, a person that is at least 16 years old.

Table 1.19: LPM for September's Survey: IV-2SLS.

	(1) PT	(2) PSDB	(3) PV	(4) Small	(5) Blank
Δ ISPs	−0.142***	0.063***	0.028**	0.009**	0.040***
	[0.019]	[0.017]	[0.013]	[0.005]	[0.008]
Radio-tv-newspaper	1.944*	−0.462	0.836	−0.065	−0.462
	[1.048]	[0.958]	[0.732]	[0.247]	[0.440]
Word of mouth	2.327**	−0.348	1.271*	0.115	−2.095***
	[0.983]	[0.904]	[0.683]	[0.232]	[0.409]
Groups	0.090	−3.511***	−0.204	0.501*	0.482
	[1.013]	[0.922]	[0.674]	[0.259]	[0.371]
Campaign tv-radio	2.551**	1.466	−3.243***	−0.057	−1.482***
	[1.004]	[0.917]	[0.688]	[0.248]	[0.379]
Socioeconomic Status	Yes	Yes	Yes	Yes	Yes
Infrastructure	Yes	Yes	Yes	Yes	Yes
Observations	11660	11660	11660	11660	11660
R^2	0.148	0.104	0.096	0.012	0.019

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: These estimations use data from a nationally representative survey (DATAFOLHA, 2010*a*) held in September 2010. All the regressions include sociodemographic controls at the municipality level. The unit of observation is a voter, that is, a person that is at least 16 years old.

Table 1.20: LPM for July's Survey: IV-2SLS.

	(1) PT	(2) PSDB	(3) PV	(4) Small	(5) Blank
Δ ISPs	−0.092***	0.108***	0.005	0.002	0.047***
	[0.021]	[0.021]	[0.014]	[0.007]	[0.010]
Radio	2.646**	2.194**	−0.637	0.402	−2.441***
	[1.042]	[1.096]	[0.705]	[0.344]	[0.569]
World of Mouth	1.271	4.267***	−1.980**	0.572	−3.107***
	[1.329]	[1.417]	[0.878]	[0.454]	[0.710]
Newspaper	4.789***	2.232**	0.465	0.123	−3.010***
	[1.067]	[1.127]	[0.749]	[0.352]	[0.607]
TV	7.967***	0.814	−0.976	−0.859**	−3.652***
	[1.141]	[1.238]	[0.822]	[0.404]	[0.717]
Socioeconomic Status	Yes	Yes	Yes	Yes	Yes
Infrastructure	Yes	Yes	Yes	Yes	Yes
Observations	10905	10905	10905	10905	10905
R^2	0.110	0.055	0.041	0.004	0.021

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: These estimations use data from a nationally representative survey (DATAFOLHA, 2010*b*) held in July 2010. All the regressions include sociodemographic controls at the municipality level. The unit of observation is a voter, that is, a person that is at least 16 years old.

1.7.3 Internet Technologies used in Brazil

There are many types of technologies through which households can be connected to the Internet. These technologies offer a wide range of speeds. In Brazil at least ten different types are currently in use. The telecommunications authority of Brazil (Agência Nacional de Telecomunicações – Anatel) makes available termly data on the number of Internet lines each ISP reported to have each year, by municipality. These data are grouped into 11 categories:

(i) xDSL: The most used technology is the DSL (Digital Subscriber Line), that stands for a group of different types of access that make use of the ready built copper wire networking structure of landline telephones. In this group are included the ADSL (Asymmetrical Digital Subscriber Line) technology, HDSL, and VDSL. The Table 1.21 shows the distribution of the usage of the different types of technologies across municipalities. The DSL is the most used one, since the use of an existent structure lower the costs for companies providing the service.

(ii) Satellite: This technology is used in those remote places where the other types of technology are not available, since it does not depend on any built-in structure, just a rooftop antenna to receive the satellite signal. The speed of the Internet provided is on average lower than the one provided by the other types and its quality varies with weather conditions or obstructions. High latency is an issue. Direct-to-Home Services (DHT) satellite broadcasting is also included in this category.

(iii) Cable: We aggregate Cable Modem and HFC (Hybrid fibre-coaxial) technologies under this category. These are normally bundled with TV services, and make use of the cable TV structure. The high bandwidth of the structure translates into high speed Internet. It is the second most used type in Brazil.

(iv) Radio waves: FWA (Fixed Wireless Access), spread spectrum, and MMDS (Multichannel Multipoint Distribution Service) are three different forms of wireless broadband. They are commonly used in areas with low population density, notably rural areas, where building the structure needed to provide the other types of technology may not be economically viable. Some companies also offer these services in urban areas, where cables have not been laid. TV services are also provided using these technologies. The signals are sent via microwave frequencies and received in the subscriber's location by a rooftop antenna. These types normally offer lower speeds and are subject to interruption by weather conditions and physical barriers such as trees or hills.

(v) Fibre-optics: The type used in Brazil is the FTTH (Fiber to the home), where the fiber reaches the boundary of the living space, such as a box on the outside wall of a home. The other technologies in this group, the so called FTTx, are classified according to the reach of the fibre-optic cabling that can vary from the user's house to a street cabinet, and this has quality implications. The capacity of data transmission both of copper wires and fiber optics are both limited by their extensions, with longer extensions meaning lower speeds, but the copper wires are much more limited in this sense in many orders of magnitude.

(vi) PLC: Power-line communication (PLC) is the technology that uses the electric power distribution system to carry data.

(vii) Others: The agency does not specify what are the other types of technologies used in this category.

Table 1.21 displays quantiles of the distribution of the share of each technology across municipalities. As it can be seen, the distributions of xDSL and Cable Modem, those which my instruments are able to predict, dominate all the others.

Table 1.21: Descriptive statistics of the type of access technologies used across municipalities

Technology	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
xDSL	0.00	0.54	0.82	0.71	0.93	1
Cable Modem	0.00	0.11	0.32	0.32	0.47	0.90
Spread Spectrum	0.00	0.03	0.18	0.31	0.56	1
Others	0.00	0.03	0.07	0.16	0.19	1
SAT	0.00	0.01	0.02	0.14	0.08	1
Hybrid	0.00	0.01	0.04	0.10	0.13	1
DTH	0.00	0.00	0.00	0.05	0.00	0.32
MMDS	0.00	0.00	0.00	0.03	0.02	0.16
FWA	0.00	0.00	0.00	0.02	0.01	1
FTTH	0.00	0.00	0.00	0.01	0.01	0.33
PLC	0.00	0.00	0.00	0.00	0.00	0.00

Note: This is extracted from Anatel's data on ISPs' reported number of subscriptions by type of technology. I calculate the share of subscriptions in each municipality for each type of technology and report the quantiles of these shares.

Table 1.22 displays the average share of the number of lines available at municipalities by different bands of speed. The highest average shares are observed in the speed band 2Mbps-34Mbps, which is classified as broadband. With this speed, users are able to watch videos on YouTube, for example. For the lowest speeds in this band, it might not be possible to stream films and TV programs though. It adds to it the fact that actual connection speeds vary. Normally, they are below the one contracted. Admittedly, the speed band 64Kbps–512Kbps gives the users limited access to heavier content. In addition, *coeteris paribus*, a user with a 2Mbps speed would be able to cover four times more content than a user with a 512Kbps speed would. That is, patient people would be able to get the same content, but it would take longer.

On the other hand, according to data from HTTPArchive.com, the average site is now 2.1 MB in size, two times larger than the average site in 2012[22], for example. This is due to the fact that most websites are content rich, have many scripts running

[22]See `http://money.cnn.com/2015/06/16/technology/web-slow-big/`.

in the background and make heavy use of images and interactivity plug-ins. It means that five years ago, website content was more adapted to slower speeds than they are today. These slower speeds would allow for normal browsing – email, blogs, Facebook, Twitter – and websites normally have versions with compressed data that allow for browsing in situations of lower bandwidth.

The bottom line is that the majority of people with Internet access in 2010 would enjoy actual speeds of up to 34Mbps. A smaller number of them, especially those served by the Cable Modem technology, would have actual speeds of up to 2Mbps. At the least, these speeds would allow normal access to social networks in general and news content. The limitations would be manifested only in the access to video content[23].

Table 1.22: Descriptive statistics of the speed of access technologies used across municipalities

Technology	0Kbps–64Kbps	64Kbps–512Kbps	512Kbps–2Mbps	2Mbps–34Mbps	>34Mbps
Cable Modem	0	0.39	0.28	0.47	0.01
xDSL	0	0	0	0.41	0.00
FWA	0	0	0	0.29	0.00
FTTH	0	0.01	0	0.08	0.05
Spread Spectrum	0	0	0	0.06	0.00
Hybrid	0	0	0	0.05	0.01
SAT	0	0	0	0.02	0.09
Others	0	0	0	0.01	0.07
PLC	0	0	0.00	0.00	0
DTH	0.02	0.00	0.07	0	0
MMDS	0	0.12	0.00	0	0

Note: This is extracted from Anatel's data on ISPs' reported number of subscriptions by type of technology and speed. I calculate the share of subscriptions in each municipality for each type of technology and speed, and report the average share for each speed band.

[23]For example, to watch a video on YouTube without pauses it is required a bandwidth of at least 2Mbps.

Chapter 2

Do some electoral systems select better politicians than others? Single- vs dual-ballot elections

2.1 Introduction

Democratic institutions form a complex incentive system that determines which types of citizens become politicians and which politicians get elected. In particular, electoral rules determine how winners are chosen from a set of candidates (ballot) on the basis of voters' rankings of those candidates. Throughout the time, many voting rules have been developed so as to meet different desired properties. Those that allow for more than a ballot in a single election have specific features that, in theory, may provide voters' with more information about candidates and other voters' preferences. This paper examines whether single- and dual-ballot rules, in practice, affect the types of politicians who run for office and win elections.

Dual-ballot elections - those that follow the rule where a second election is held between the two most voted candidates when none of them has obtained a majority in the initial election - are used by many countries in Latin America (for both local and

national elections), in French presidential elections, in the United States gubernatorial primaries, and in Italian local elections. Single-ballot elections are, nonetheless, more ubiquitous around the world. They are normally justified on the basis of the higher financial costs of running two elections instead of just one. The theoretical literature, however, points to higher levels of information disclosure in dual-ballot elections that would benefit voters (Piketty, 2000; Martinelli, 2006), allowing elections to better aggregate preferences. Yet the objective implications of the two systems for both voters' and parties' decisions are not well understood empirically.

On the one hand, dual-ballot elections normally expose voters to a longer period of advertisement and, through first round outcomes, provide valuable information about other voters' preferences and candidates' electoral performance. *Coeteris paribus*, we would think these voters are more informed than those under the single round system. Indeed, the empirical literature has shown that both length of advertisement (Da Silveira and De Mello, 2011) and campaign spending (Levitt, 1994) have a positive effect on election outcomes in general. If a dual-ballot system allows voters to properly screen candidates, parties may respond to that fielding politicians of higher quality, reinforcing the likelihood that voters will choose better candidates.

On the other hand, a longer electoral period and the requirement to vote in two elections may demand more attention and motivation from voters, discouraging participation in the electoral process. It could also be meaningless, when the extra signals produced render those previously held uninformative. Moreover, the signal-to-noise ratio of campaigns can even be reduced in a set up with a larger number of candidates[1] and campaign activities. Aware of this, parties would not be as concerned with the quality of politicians they have and this behavior would attract more and more politicians of inferior quality.

[1]On average, a higher number of candidates is fielded in dual-ballot elections (Fujiwara, 2011).

This ambiguity calls for a careful empirical analysis that, to the best of my knowledge, the literature still lacks. The ethical impossibility of randomizing the assignment of electoral rules to municipalities represents an obvious empirical challenge. Differences in candidates fielded and elected across different municipalities could be incorrectly attributed to the electoral system when, in reality, municipalities that adopt a certain rule have unique features that are correlated with certain types of politicians and policy outcomes. Even though observable factors can be controlled for, there will always remain unobservables that can compromise a causal interpretation of these effects.

By using a quasi-natural experiment that follows from the allocation threshold of single- and dual-ballot elections to Brazilian municipalities, based on electorate size, I can assess the causal effects of the adoption of a dual-ballot system, relative to single-ballot, in a host of relevant outcomes.

First, I verify whether observable candidates' characteristics differ under the two electoral rules. Second, to evaluate unobservable characteristics that are relevant for policy-making, I also compare politicians' performance in office. This can be achieved by analyzing policy outcomes for which the response time of the targeted population is short enough to be observed between elections, such as health care. Following the literature, I also analyze performance in terms of the potential politicians have in attracting discretionary transfers to their constituency.

The main finding of the paper is that female candidates appear less often in the top two and top three positions in the first round of dual-ballot elections. The data supports the interpretation that this difference is not due to gender discrimination but rather associated to underlying political skills, specially experience in mayoral elections. This result is in line with Bhalotra, Clots-Figueras and Iyer (2017) who find an increase in the participation of women candidates from major political parties

in elections that follow a female victory. Most importantly, they also show that this effect is primarily driven by prior candidates contesting again.

Perhaps surprisingly, other observable characteristics - namely age, schooling, and occupational skill level - are not statistically different between the two systems for none of the adopted definitions of effective candidates.[2] In addition, differences in campaign spending and fundraising are also not found.

In general, female candidates absence from dual-ballot elections could be explained through three channels involved in the candidacy process: (i) the woman has to be willing to put herself forward as a candidate; (ii) the party has to offer support by placing her in the ballot; and (iii) voters need to be willing to vote for both men and women. In the case analyzed here, the same parties appear in top positions in both systems but female candidates are fielded at a higher rate in single-ballot elections. An evidence that these parties may not discriminate female candidates *per si*. The same is true for the voter perspective. Because the identification strategy holds municipalities characteristics constant, single- and dual-ballot voters should have similar gender preferences for candidates. Therefore, channels (ii) and (iii) do not offer a plausible explanation for my results. Channel (i) does not seem to be the strongest mechanism behind my findings either. I present evidence[3] that dual-ballot elections are normally more contended and, as such, could repeal the participation of women, who normally avoid competitive environments, specially those of mixed gender. However, female candidates appear at the same rate in lower ranked positions in both systems, revealing that women may actually be willing to participate of dual-ballot elections too, something that downplays this mechanism.

In terms of unobservable characteristics, mayors elected in dual-ballot municipalities are more likely to attract discretionary transfers of capital to their municipalities, but only in election years when they are eligible for re-election. Furthermore, these

[2]Those with a non-zero chance of winning or going to the second round.

[3]De Mello, Firpo and Chamon (2009) also provide evidence in this direction.

municipalities do not present any differences in health outcomes, as measured by the share of non-premature births and the share of mothers with at least one prenatal visit during the mayor's term. These politicians try re-election and are elected at the same rate as in single-ballot, pointing to the interpretation that dual-ballot requires an additional effort from mayors at this stage. If dual-ballot mayors actually performed better than their single-ballot counterparts in general, performance differences should also have been found for term-constrained mayors. Based on this idea, throughout the paper, I provide evidence that supports the view that the political skills required from candidates in order to enter dual-ballot elections are not necessarily associated with higher quality and that female entry is barred because of these very same skills.

Overall, there is suggestive evidence that dual-ballot elections affect entry of effective candidates only through requiring more political experience, but that has no implications for performance in office, except that mayors with re-election interests tend to attract more resources to the municipality in election years – however, these resources are not reverted into better policy outcomes. From the voters' perspective, there is a reduction in turnout in the second round of dual-ballot elections, relative to its first round, followed by reductions in both blank and spoilt votes. This is not explained by closeness in the elections and voter behavior is similar in the comparison of the first rounds of the two types of elections. All together, these results indicate that political experience combined with some loss of voters' interest in the second round tend to attenuate the potential extra demands posed by dual-ballot elections.

This paper has intersections with a number of different strands of the literature. First, it is related to an empirical literature that has tried to assess the effects of electoral rules on politicians' performance in office. Persson, Tabellini and Trebbi (2003) do a cross-country analysis to verify the effects of a host of electoral rules on corruption. They verify that countries that switch from strictly majoritarian to strictly proportional elections present smaller levels of corruption, even though the effect is

small. Gagliarducci, Nannicini and Naticchioni (2011) compare the performance of candidates elected under majoritarian and proportional systems in Italy and show that representatives elected under majority field more bills and are less absent from Congress. This paper compares politicians fielded and elected through two different majoritarian systems instead. Dual-ballot elections not only change the formal way under which votes are translated into a winner, but also consist of an experimental setting where voters and politicians are also exposed to new sets of information. Therefore, has different features to be explored.

Second, there is also a small literature in political science (Fulton, 2014; Roberts, Seawright and Cyr, 2012) that has described how proportional representation electoral systems are associated with greater legislative representation of women relative to majoritarian systems. My findings detect differences in representation within majoritarian systems and in executive positions.

Third, it is also related to the economic literature on gender differences in politics (Brollo and Troiano, 2016; Anzia and Berry, 2011; Bhalotra and Clots-Figueras, 2014). This literature uses samples of elections where women have closely won/lost against male candidates to show that female politicians perform better in office and are less corrupt, on average. I provide new evidence about gender differences at the stage of candidacy for an election and how electoral rules may pose extra barriers to women's political advancement.

Lastly, the closest paper to mine is Nannicini, Bordignon and Tabellini (2016). They show municipalities that choose mayors using dual-ballot systems face less volatility in property taxes. This is due to the fact that in dual-ballot elections centre parties make coalitions with extremist parties[4] - in which case there is an adaption of policies towards the positions of extremist parties - less often than in sin-

[4]This is also related to a theoretical literature on dual-ballot elections that points to gains in efficiency given by the separation between the "communicative" and "decision-making" functions of voting (Martinelli, 2002; Piketty, 2000; Bouton, 2013).

gle round elections, where extremist parties can threaten the victory of centre parties by dividing minority voters. In this case, policies are enacted as a result of strategic considerations induced by the system at the time of the election and are indirectly related to politicians quality. In this paper, I focus on a previous stage: the one in which citizens choose to run for office and the party selects who to support. I use this to infer about the effects of a dual-ballot system on the characteristics of those fielded. I further analyze intermediary policy outcomes, that are plausibly not affected by political ideology, only to try to assess unobserved features that parties may take into account in this selection process.

The paper is organized as follows. In Section 2.2 I describe in detail my identification strategy, the datasets used and the institutional background of the elections considered. Section 2.3 contains the results of all estimations and discusses some of the possible mechanisms leading to the results found. Section 2.4 concludes the paper. All tables and figures are presented in an appendix.

2.2 Institutional Background, Data and Estimation Framework

2.2.1 Data

Election outcomes and candidates' characteristics - Data on mayoral elections, number of registered voters and candidates' characteristics are available from the electoral authority (Tribunal Superior Eleitoral - TSE) for the election years 2000, 2004, 2008, 2012 and 2016. These data, however, need to be matched over the years as to verify previous positions held by the candidates. I developed an algorithm to identify the same candidates over time: (i) for the years where identification numbers for candidates were available, they were used to make these connections (ii) for the elections

where they were not available, I assume a candidate with the same name at the same city at different points in time is the same person. Since this is administrative data, there are some inconsistencies such as misspelled or incomplete names. A partial matching algorithm was used to identify these cases and the results eyeballed so as to verify the valid matches. Candidates' characteristics available in this dataset include date of birth, gender, education, occupation and reported wealth.

Discretionary transfers - Data on discretionary transfers are available from balance sheets reported by the municipalities to the Brazilian National Treasury (Tesouro Nacional). This information is available for all the years up to 2015. Following Brollo and Troiano (2016), I use discretionary transfers of capital as a measure of politician ability to attract resources to the municipality. To get a measure of non-discretionary transfers, I subtract transfers of capital from the total of capital resources the municipality had.

Health Outcomes - Information on all births and number of mother's prenatal visits are provided by the Information System on Live Births (SINASC), available on DataSUS, a system managed by the Ministry of Health. These data are available for all years since 2000 up to 2015. I extract two variables from them: the share of non-premature births and the share of mothers who had at least one prenatal visit during pregnancy.

Campaign spending and donations - This information is also available from TSE, the electoral authority. Candidates are required to report in detail the amount and the sources of donations received and how they were spent. All donations are aggregated into the following categories: those raised in campaign events (*Events*), voters donations (*Voters*), party funding (*Party*), public funding (*Public Funding*) and those for which the source was not identified (*Not Identified*). It is not possible to do the same with spending because candidates report the figures labelled by loose multiple

categories over the years, which makes it difficult to rationalize them into meaningful groups. I therefore explore only the overall total spent by the candidates.

Municipalities' characteristics - The 2000 Census conducted by the national statistics office IBGE (*Instituto Brasileiro de Geografia e Estatística*) is used to obtain detailed characteristics of municipalities at the baseline year. These variables describe voters education, income distribution, demographics and municipalities' infrastructure. For a list of the variables and their definitions see Table 2.2.

2.2.2 Estimation framework

Brazil is a federation with 26 states and a federal district politically organized in 5567 municipalities. Voting is compulsory for all citizens aged 18 years or over, but the vote is voluntary for those aged 16-17 or above 70 years of age. Participation in elections is normally high due to this feature.

Each municipality chooses a new mayor and local legislature every four years. Since 1998 mayors are allowed to run for a second term. The constitution establishes a threshold of 200,000 registered voters to determine whether a municipality holds a dual- (with plurality rule) or a single-ballot in mayoral elections. In dual-ballot elections municipalities above this threshold have a second election with the top two candidates of the first round when none of them has obtained more than 50 percent of the valid votes. This threshold generates a discontinuity that randomly assigns single- and dual-ballot elections to municipalities sufficiently close to it. That is, those municipalities should be similar to each other in both observables and unobservables, with the only difference coming from the discontinuous implementation of distinct electoral systems.[5]

[5]This feature has been used for identification before by Fujiwara (2011) for Brazil, and Nannicini, Bordignon and Tabellini (2016) for Italy.

More formally, let V_{it} be the RDD running variable for municipality i in the election at time t. V_{it} is a variable centered at the threshold and, thus, denotes the distance from the threshold and has positive values for municipalities with dual-ballot and negative values for municipalities with single-ballot. The treatment effect in a close neighborhood of the threshold on outcome Y_{it} is given by:

$$TE = \lim_{v_{it}\downarrow 0} E[y_{it}|v_{it}] - \lim_{v_{it}\uparrow 0} E[y_{it}|v_{it}] \tag{2.1}$$

Under the assumption that the conditional expectation of y_{it} on v_{it} is continuous, the first term on the right-hand side converges to the expected outcome for a dual-ballot municipality, which has as many voters as the single-ballot municipality. Similarly, the single-ballot municipality converges to the expected outcome of a dual-ballot municipality under the same conditions.

The continuity assumption depends on two facts: the non-existence of (i) manipulation in the allocation of single- and dual-ballot rules across municipalities and (ii) other treatments based on the same threshold. As discussed in Fujiwara (2011), this threshold is arbitrary and is not used as a cutoff for the assignment of any other treatment. All the municipalities comply with the rule, and therefore, the regression discontinuity design is sharp.

Following the literature, two different methods are used to estimate average treatment effects (ATE): a p-order polynomial fitted on either side of the threshold and a local polynomial regression.

Let $v_{it} = \#voters_{it} - 200,000$ be the number of registered voters centered around the cutoff in municipality i in time t, then a p-order polynomial can be estimated through:

$$y_{it} = \sum_{k=0}^{p} \beta_k v_{it}^k + \tau_{it} \sum_{k=0}^{p} \gamma_k v_{it}^k + \mu_t + \varepsilon_{it} \tag{2.2}$$

where y_{it} is the outcome of interest in municipality i in time period t, τ_{it} is a dummy that is one when the municipality follows a dual-ballot system, μ_{it} are year fixed effects and standard errors are clustered at the municipality level as the same city can be observed multiple times. γ_0 is the treatment effect, that is, it measures the jump around the cutoff between the two groups of municipalities.

Following Imbens and Lemieux (2008), the local linear regression approach sets $p = 1$ in equation (2.2) above and restricts the sample to municipalities in the optimal interval $v_{it} \in [-b, +b]$ computed using Imbens and Kalyanaraman (2012). Again, γ_0 is the treatment effect at the threshold $v_{it} = 0$.

2.2.3 Validity Tests

The identification strategy is based on the fact that the population threshold used to assign the treatment is arbitrary and, as such, observable and unobservable characteristics of municipalities close enough to this threshold should be statistically similar.

If this assumption holds, there should not be any discontinuities around the threshold for any of the municipalities' characteristics we can observe. Therefore, using a wide array of variables extracted from the Census 2000 and that characterize well a municipality in terms of income, education and infrastructure I assess the validity of this assumption.

These tests are reported in Table 2.1 and can also be confirmed by visual inspection of Figures 2.1, 2.2 and 2.3. All variables, including the share of women in the population, are balanced across the cut-off. Overall, there are no differences in education levels, income, unemployment, as well as sanitation and electricity services.

These results are corroborated by McCrary (2008)'s continuity test. In principle, politicians could manipulate the number of registered voters in a municipality to induce the application of their most preferred electoral rule. McCrary (2008)'s

continuity test verifies whether there is any random sorting of municipalities at the cutoff. The idea of the test is that with sorting the density of the running variable would not be continuous. It tests the null hypothesis of continuity of the running variable density by implementing kernel local linear regressions separately on both sides of the threshold. Figure 2.7 displays the results. The estimated discontinuity is 0.5360 with a standard error of 0.4481 (As it can be seen, there is no evidence of discontinuities in the number of voters around the cutoff).

Taken together, these results suggest that the running variable of the RDD does not show any evidence of manipulation and can be safely used as a local source of exogenous variation in the neighborhood of the threshold that allocates electoral rules to municipalities.

2.2.4 Sample selection and descriptive analysis

Pooling the data of the five electoral cycles produces a sample of over 23,297 elections conducted in a single-ballot system and 259 under dual-ballot.[6] However, because RDD assumptions hold only for those municipalities in the close neighborhood of the threshold, I reduce the sample to those within 75,000 voters from the zero cutoff [7] and calculate the optimal bandwidth using by Imbens and Kalyanaraman (2012). The effective sample size used is composed by 308 elections, with 97 of them following dual-ballot.

In this sample, dual-ballot elections have 0.94 more candidates on average and a maximum number of 13 candidates. Whereas in the full sample, there are 3.54 more candidates on average and the maximum number goes up to 14. In fact, dual-ballot

[6]My data differ from the one used in Fujiwara (2011) by two elections. While he includes electoral cycles from 1996 to 2008, I include electoral cycles from 2000 to 2012. A dataset containing all elections in 1996 is currently not available from TSE.

[7]This was the greatest threshold used in Fujiwara (2011).

has been shown to have a causal effect on the number of candidates fielded in an election (see (Fujiwara, 2011; Nannicini, Bordignon and Tabellini, 2016)).

Since the goal here is to compare candidates' characteristics on both sides of the discontinuity, it is important to investigate what the implications of a higher number of candidates are for the distribution of quality and how a fair comparison between the two systems can be performed.

Vote shares obtained by a candidate reflect many dimensions of quality, such as party affiliation, experience, valence, campaign spending and so on. However, vote shares in dual-ballot elections are on average lower than in single-ballot because (i) more candidates enter the competition and (ii) some citizens vote for their preferred candidate even when she is not likely to win (sincere voting) – as there will be a second round where they can vote making strategic considerations. Hence, the strategic voting component of vote shares is different in each system.

Empirical distributions of aggregate vote shares, broken down by candidates' ranking, are presented in Figure 2.4. If we take the sum of vote shares of the two first-placed candidates in each system, as displayed in Panel (a), the average for single-ballot is 0.76 and for dual-ballot it goes down to 0.72 and has a larger variance, making the difference in means significant. Medians, in turn, are 0.77 and 0.71, respectively, confirming the left skewness of the single-ballot distribution. Both distributions are bimodal. Importantly, dual-ballot presents a longer left tail, illustrating the fact that dual-ballot allows for the entry of candidates that receive very small vote shares and, as we will see later, compare very poorly in terms of observables to other candidates in the race.

Panel (c) shows that we only get an approximation between the aggregate vote shares of the two systems when considering the aggregate of the four first-placed candidates in dual-ballot against the two first-placed in single-ballot.

In the smaller sample of 97 dual-ballot races, there were 43 outright victories and first-round first-placed candidates won in 39 of those with a second round. In the full sample, that was 111 and 108, respectively. In other words, dual-ballot races normally go to a second round but first-round winners often win the election.

Figure 2.6 displays the distribution of the number of elections a municipality has been running dual-ballot within our sample of five electoral races. Because the sample has been reduced to those close to the threshold, we observe only a few cases where the municipalities had a dual-ballot system over all five elections. The bulk of them are newly converted municipalities or have been in the system for two or three races. Therefore, this paper can not say much about the long term effects of the adoption of a dual-ballot system.

We would like to infer what the average quality of groups of candidates fielded under each system is, but there are candidates who enter the election just targeting a non-zero probability of a positive vote share, no matter how small it is. This can be seen in Figure 2.5 which displays the aggregate of vote shares of fourth and lower placed candidates in the two systems. Clearly, the density of zeros or near zero vote shares is higher in dual-ballot elections. Therefore, to make meaningful comparisons between groups it is necessary to find some criteria that determine who the effective candidates are in both types of elections. In the next section, I discuss the problems arising from this sort of classifications and identify the ones to be used here.

2.2.5 Selecting comparable groups of candidates

There is a fundamental problem about comparing the quality of candidates between the two systems. Dual-ballot elections and their lower barriers to entry attract candidates who would not have entered the election if it was not for the dual-ballot rule. These new entrants can be either of high or low quality and make it to the top three candidates, depending on voters' strategies.

If we take the full pool of candidates who enter dual-ballot elections and compare their quality with those in single-ballot, it could be the case that non-effective candidates - those with no chance of getting a positive share of votes if voters had no incentives to vote sincerely - would weight down the average quality, producing a result where dual-ballot candidates would be worse on average, even if some group of top candidates are better than their single-ballot counterparts. The problem is that, *a priori*, we cannot judge by the election results who the effective candidates would be in each of the systems.

When there is a risk of upset victory in a dual-ballot election, a potentially effective candidate could get a relatively low share of votes because voters coordinate to give outright victory to a candidate that is more likely to win the election and avoid the victory of a Condorcet loser in a second round. However, as Bouton and Gratton (2015) show, this is an equilibrium only for dual-ballot elections with a threshold below 50%. Here, all the elections have a threshold of 50% for outright victory. Therefore, we can discard the possibility that dual-ballot elections present effective candidates with low vote shares just due to strategic considerations.

Hence, if the case of upset victory can be ruled out, it is likely that candidates who get very low vote shares in dual-ballot elections are those who enter the race just because the chance of a non-zero vote share is higher. They are likely to be candidates who would perform as badly or worse in single-ballot elections, so we get to observe them only in dual-ballot elections.

Based on the arguments above, assume that election results are a good ex-ante measure of how likely the candidate is to be a serious contestant in an election (ideally we would look at polls held at the beginning of the electoral period, but that is not available for most of the municipalities). We can then have an idea of which candidates are more likely to be in the effective group. Note that by doing this, we do not need the assumption to hold for each candidate's ranking. We just need that the group of

top ranked candidates are the same ex-ante and ex-post. The group of non-effective candidates is then composed by those who get a very small share of votes relative to the other candidates of the system they are competing on.

Based on this argument, we cannot reach a definitive concept of candidate effectiveness for both systems, but we can compare different combinations of best ranked candidates. Throughout the paper I compare the groups of top two placed candidates in both systems and the top two in single-ballot and top three in dual-ballot.

2.3 Results

2.3.1 Are candidates different over observable characteristics?

In this section I evaluate dual-ballot treatment effects over candidates' age, gender, schooling, and occupational skill levels. I will refer to these variables as observable characteristics throughout the text.

Occupational skill levels are used to differentiate those engaged in occupations that require the execution of a more complex range of tasks. These are here classified according to the four skill levels of the International Standard Classification of Occupations 2008 (International Labour Organization, 2012). Candidates belonging to the fourth skill level normally have a university degree and perform complex tasks in a specialized field that demands problem-solving and decision-making in addition to creativity.[8] Examples of professionals in this category are musicians, medical prac-

[8]The other three skill levels are:

1. Skill level 1: Occupations that require basic or no education and involve the performance of simple and manual tasks. Examples: Office cleaners, kitchen assistants.
2. Skill level 2: Occupations that require basic education and involve operating machinery or electronic equipment. Examples: Bus drivers, police officers, hairdressers.
3. Skill level 3: Normally requires a university degree and involve the execution of complex tasks in a specialized field. Examples: Legal secretaries, shop managers, computer support technicians.

titioners, computer systems analysts, etc. A dummy identifying candidates in this skill level, called *High skill*, is used to compare candidates who fall into this category across the two electoral systems.

For all the outcomes and different samples of interest I run a non-parametric model, using Imbens and Kalyanaraman (2012) to select the optimal bandwidth, and a parametric RDD spline polynomial of 3rd and 4th orders – considering the sample of municipalities that are 75,000 voters away from the zero cutoff. Standard errors are clustered at the municipality level to account for any type of serial correlation and year fixed effects control for year-specific characteristics of elections.

I start by showing the results for the sample that includes all candidates running in an election. Panel A in Table 2.3 demonstrates that, overall, there are no statistically significant effects when we go from single- to dual-ballot in any of the characteristics or specifications considered. Similarly, when the set of top ranked candidates in each election is excluded the same results are verified (Panel B in Table 2.3).

As previously discussed, to try to account for the fact that some of the candidates attracted by lower barriers to entry in dual-ballot elections may pull down the overall average quality, I now consider different groups of candidates in each side of the discontinuity. My assumption is that relative low shares of votes in the first round of dual-ballot elections are not given by strategic considerations, but rather reflect voters' preferences. Therefore, my measures of effectiveness are based on the candidates who have a considerable relative large share of the votes in each election.

In Panel A of Table 2.4 I use the subsample of top two candidates in both systems. Remarkably, the probability of a woman appearing in the top two of a dual-ballot election is 10 to 30 percentage points lower, depending on the bandwidth and specification. Out of an already low baseline mean for female candidates in single-ballot of 11 percent. Whereas no statistically significant differences are found with respect to

age, schooling or skill, which implies that the additional women fielded under single-ballot are at least as good as their male counterparts in dual-ballot over these three dimensions.

Since two candidates go to the second round in dual-ballot, three candidates normally contest the elections more closely, while this is true for only two candidates in single-ballot elections. To account for that, I amplify the sample to the top three candidates in dual-ballot against the top two in single-ballot. As displayed in Panel B of Table 2.4, the number of female candidates in dual-ballot races increases but is still lower, on average, when compared to single-ballot, around 10 to 23 percentage points. Again, no differences are verified in the other observable characteristics considered.

As the previous results have shown, there is no discontinuity in the full sample of female candidates fielded in each electoral system, the discontinuity is verified only in the subsamples of effective candidates. That is, women are fielded at the same rate in both systems, but do not get into the elite group as often.

To shed some light on this outcome, I take the subsample containing only female candidates and check for discontinuities in the same characteristics: education, experience and skill. The results are shown in Panel A of Table 2.7. Essentially, there are no detectable discontinuities between characteristics of women fielded in the two systems, except that they seem to be older, on average. I also do the same analysis for the sample containing only males. Panel B in Table 2.7 display the results. Male candidates, in turn, present the same observable characteristics in both systems, on average. This finding aligns with the argument that dual-ballot may demand more experience from politicians, but only for women this requirement is reflected in age. This would be consistent with both parties disproportionally requiring more experience from women than men, or just a reflection of late entry of women in politics.

To sum up, the only significant treatment effect observed thus far is that dual-ballot races have comparatively less female candidates competing in top positions,

despite presenting the same observable characteristics. There seems to be a convergence in all the other observable dimensions in both electoral rules, something that indicates parties' selection process may already be driven by these criteria in general. The question that remains to be answered is whether this result can be attributed to large parties' gender preferences or to some underlying candidates' characteristics that are associated to gender. This question cannot obviously be answered directly, but in the next sections additional results help excluding mechanisms that could be at play.

2.3.2 Do dual-ballot candidates spend more in their political campaigns?

A second round election in Brazil implies additional 28 days of rallies, debates and production of ads for radio and tv. This longer advertisement period should also shift the amount of resources necessary to run a political campaign, as compared to single-ballot. Candidates should, therefore, be able to raise a large amount of funds to make up for these extended costs; a skill that should account for another dimension of candidates' quality that parties may observe.

To test this proposition, I verify whether there is any dual-ballot effect on candidates' campaign funds and spending. The results are displayed in Table 2.6. First, there are no statistically significant differences between the total spent in single-ballot relative to that of dual-ballot elections (*Total spending*). The same is true for the total spent in single-ballot compared to that of a first round in dual-ballot (*Total spending 1st round*). The total spent in the second round of dual-ballot compared with the monthly average spent in single-ballot elections is also statistically similar (*Total spending 1st/2nd round*).

The aggregate of funds raised in both first and second rounds seem to be larger in dual-ballot, although not statistically significant (*Fundraising total*). When broken

down by rounds, there are small and non-significant differences between single-ballot and dual-ballot first round (*Fundraising total 1st round*), but large positive differences between first (monthly average) and second rounds in each system, respectively (*Fundraising total 1st/2nd round*).

In short, there are no differences in the total of resources raised and spent in the two systems, on average. Candidates seem to keep the same pace of spending in the second round and may transfer resources from first to second rounds.

To have a more precise idea of candidates fundraising abilities I also check for discontinuities in the following sources of funding: events, voters, companies, public funding[9], party funding and unknown sources. The results are presented in Table 2.3. They show that either candidates do not focus differently in any donor groups and/or none of the donors seem to take system into account when giving money to campaigns.

All in all, candidates follow a similar path of spending and fundraising in their campaigns despite dual-ballot elections lasting longer. This suggests that candidates have, on average, the same fundraising abilities in the two systems but may require additional political skills to compensate for the supposedly more demanding campaign in dual-ballot. In the next section, I assess whether political characteristics and redistributive politics play a role in balancing these differences.

2.3.3 Are candidates different over political characteristics?

Table 2.8 displays the analysis of discontinuity in key variables characterizing the political process underlying each of the systems: experience as a candidate in mayoral elections, reelection entry, reelection rates and competitiveness.

[9]In Brazil, parties receive public resources to fund political campaigns, proportional to the number of representatives they have in the lower chamber.

The variable experience in mayoral elections counts how many times the politician has been a candidate in a mayoral race at that municipality up to the most recent election. Remarkably, there is a large number of politicians who compete more than twice in the sample. Around 29 percent has competed at least once, 10 percent twice and 3 percent three or four times. [10] The results in the first line of Table 2.8 show that candidates in dual-ballot elections have participated of 0.26 to 0.36 more elections on average when compared to single-ballot elections, which have an average of 0.52 elections. At the same time, both systems do not present any significant differences between a mayor running for reelection (*Incumbent race*) or being reelected (*Reelection win*). A visual representation of these estimates is in Figure 2.8.

This speaks to the absence of women in dual-ballot elections to the extent to that the presence of more experienced male politicians may repeal the chance of entry of new female candidates, given that women do not have had much tradition in politics. It remains to be seen whether this is a result that persists in the long run, in which case part of this effect could be attributed to plain gender discrimination.

Figure 2.9 displays the average share of women fielded in each system over time in the sample of elections used here. When considering all those fielded, independently of ranking, this average is statistically the same in all years, except in 2000, when it is higher. When including only candidates in the top two positions, this average becomes significantly lower for dual-ballot from 2004 to 2012, but displaying a smooth growth in women participation until catching up in 2016. A finding that aids the view that the experience may be a requirement of a more demanding political process and not related to gender per si. As the time passes, more politically experienced women may become available to contest dual-ballot elections in top positions. It is important to highlight, however, how low the participation of women still is in both systems,

[10]Candidates of elections in 2000 are not included in this sample because our data covers all races between 2000 and 2016 and therefore is truncated on the left. however, the possibility of reelection was approved only in 1997 and, therefore, candidates should not be repeatedly competing in the elections not taken into account.

not going over 15% in any period. Something that indicates a slow process of change and that political experience may consist in a strong barrier to entry.

I also investigate whether there are signs of tougher competition in dual-ballot elections that is reflected in vote shares. I first compare the margin of votes of all candidates over the next-placed candidate in the first round of the two systems. As it can be seen in line (2) of Table 2.8, there are small and no significant differences between the two systems. However, when comparing the margin of victory of the first ranked candidate in single-ballot and that of the second round of dual-ballot, I find that this margin is lower for the latter. This is a somewhat surprising result since the higher number of candidates in single-ballot, as compared to the second round of dual-ballot elections, would tend to smooth these margins across candidates. This provides evidence that dual-ballot elections tend to be closer than single-ballot in this context, something that could also explain why women are not seen as often at the top positions in dual-ballot elections.

This interpretation is aligned with experimental evidence that has identified reduction of women's performance in competitive settings, that becomes more severe in mixed gender environments (Gneezy, Niederle and Rustichini, 2003). However, female candidates appear at the same rate in lower positions being fielded by minor parties. This demonstrates that women's entry issue may be more related to parties' endorsement than with their choices of whether to compete or not. The stronger presence of men in these elections may in itself be a repulsing factor for the entry of women, as found in Gagliarducci and Paserman (2012), where female mayors are less likely to finish their mandate when a council is composed solely by male politicians.

In short, the interpretation that receives more support from the data is that there are underlying unobserved political skills associated to entry in dual-ballot elections that not only women seem to lack but also inexperienced male politicians. This

is backed by the fact that large parties are willing to field women in single-ballot elections, despite still not in parity with male candidates.

2.3.4 Are dual-ballot politicians different in terms of performance?

Health outcomes - One dimension over which the performance of politicians has been evaluated is health outcomes. As Fujiwara (2010) and Brollo and Troiano (2016) point out, the population responses to investment in health are fast and relevant enough so they can be evaluated in the short spam of time between different terms, and is one of the few outcomes that are available at the municipality level. Additionally, resources allocated to health in municipalities are tied to population size and income per capita, so it should be homogenous across municipalities in this aspect, even though the mayor has discretionary power in the allocation of these resources. A considerable large share of the population depend on public health care in Brazil (Sistema Unico de Saude – SUS), about 75% of the population according to the Household Surveys [11] data. This number has also been constant over the last decade.

Panel B in Table 2.9 displays treatment effects over two baseline variables characterizing health outcomes: any prenatal visits—the share of mothers who had at least one prenatal visit during pregnancy—and non-premature births, defined as the share of births from pregnancies of 37 weeks or above. Considering the possible existence of political cycles in policy investment so as to improve electoral chances close to elections, I split the four-year term into the sum of the first three years, second and third years and election year – the fourth and last year of the term –, respectively.

The effects of dual-ballot in these variables are numerically small and statistically insignificant. These results are, however, also compatible with both a more effective allocation of health resources and/or a higher share of resources destined to this

[11] Pesquisa Nacional por Amostra de Domicilios – PNAD.

end by male mayors in dual-ballot municipalities. This is true because the presence of female mayors brings up the average health outcomes in single-ballot elections. Therefore, this result does not refute the case for candidates of higher quality in dual-ballot elections.

Discretionary transfers of capital - As in Brollo and Troiano (2016), I use discretionary transfers of capital as a measure of politicians' ability to attract resources to the municipality, since they depend on agreements (convenios) made between the municipality and federal or state governments. The same pattern of non-significance is observed for discretionary transfers of capital in Panel A of Table 2.10. Despite not significant, dual-ballot municipalities still display a large negative effect on the discretionary transfers mayors are able to attract in the first three years of their mandate. Something that is reverted in the election year, when transfers of capital are much larger and significant relative to single-ballot. In terms of economic magnitude, our results show that dual-ballot mayors attract 36 percent more transfers for capital investment than their single-ballot counterparts. Overall, there seems to be a net positive higher value of transfers to dual-ballot municipalities.

Is it that other political incentives affect candidates' performance or candidates elected in dual-ballot elections really have differential characteristics that make them perform better? Term-constrained candidates that still perform better in dual-ballot elections would speak to the hypothesis that the system selects better candidates – namely, career concerned – or at least keeps checks and balances that hold politicians more accountable. To investigate this channel, I exclude term constrained candidates from the sample and run the same specifications. Now, discretionary transfers in election years are large and significant in some of the specifications with a greater sample size. Whereas non-discretionary transfers and overall discretionary transfers remain the same. Since candidates are equally likely to run for reelection and win on both sides of the discontinuity, the higher share of transfers attracted by dual-

ballot may be an extra effort required by the system itself. This supports the idea that dual-ballot elections may require higher political skills that are translated into policies with electoral goals.

2.3.5 Do voters loose interest in the election over time?

The requirement of voting in two elections, rather than in just one, associated with a longer campaign period may discourage voters from participation in the election. To test this hypothesis, I look at whether there are any differences in voter turnout and the share of blank and spoilt votes, which are equivalent to voting for outside options, across the two types of elections. Not surprisingly, a dual-ballot system has no effect on participation relative to that of single-ballot. This is explained by the fact that the first round of the two types of elections are comparable over many campaign dimensions, as we have previously shown.

If there is any type of discouragement it should appear in the second round of dual-ballot elections relative to their own first round. A quick test of difference in means reveals that turnout and the share of spoilt and blank votes are significantly higher in the first round of the election. This result remains even when using a sample of close elections – classified by the difference in vote shares of the top two candidates in the first round – where supposedly voters would have a higher incentive to turnout. This does not obviously rule out all the confounders at play. It could well be that those who vote sincerely for trailing candidates in the first round just give up voting in the second round because they do not have a better option. If that is the case, this is only a different channel through which dual-ballot elections release candidates from the extra demands of a second round, since by knowing that these voters do not turnout, candidates can focus on a smaller set of voters in the second round. All in all, there is suggestive evidence that a second round combines mechanisms that tend to attenuate its own demands. This result helps in explaining why dual-ballot

candidates need more political abilities as opposed to monetary resources or human capital outside politics.

2.3.6 Discussion

The previous results indicate that politics in general attracts the same type of citizens in terms of observable characteristics. That is, candidates are required to have a college degree and/or an occupation that qualifies them to the position. Parties' selection of candidates must therefore be made based on other grounds.

The problem with analyzing unobserved quality of politicians based on performance is differentiating moral hazard aspects from adverse selection. It remains to be seen whether politicians are ex-ante of higher-quality, and voters can observe that only in dual-ballot elections, or there are embedded accountability mechanisms that forces them to perform better. But, in either case, these are features that single-ballot elections seem to lack. The pool of more experienced candidates in dual-ballot, their political campaign strategy and timing of policies with clear electoral goals, support the idea that the screening process in this system, together with the quality of competitors, are more enhanced than in single-ballot.

The results found for women can emerge from a series of mechanisms. First, it could be a result of party discrimination, because parties decide how to allocate candidates. In the sample of elections considered here, major parties are equally present in top positions in both systems. These parties field female candidates more often in single-ballot elections, so the reduction in the proportion of women we observe in dual-ballot top positions does not seem to be consistent with discrimination, unless parties hold gender bias that are specific to dual-ballot elections. The absence of women is more consistent with parties' beliefs that dual-ballot requires more politically experienced candidates, who are able to win a more demanding election with limited campaign spending and strategically acquire and apply resources when in of-

fice targeting political goals. A question that cannot be answered in our setting is whether these requirements are disproportional towards women, because we do not observe the pool of women candidates within each party.

Second, there could be a shorter supply of female politicians in municipalities with dual-ballot elections, even in municipalities with the same proportion of women in the population, on average . That is, women may perceive dual-ballot elections as being more difficult to enter, at least competing through major parties, and may prefer other occupations or alternative public positions Casas-Arce and Saiz (2015), such as local legislatures. However, the similarity in the presence of female candidates representing minor parties is an indication that this is perhaps not a major factor contributing to the results found.

Third, it could also be a result of voters' gender discrimination. My identification strategy accounts for pre-existing gender preferences of the electorate in both sides of the discontinuity. However, voters in dual-ballot elections are subject to a different incentive system that can trigger more discrimination. It might just be more difficult to identify a high-quality female candidate in a dual-ballot setting. This is because people observe the average quality and number of candidates that are fielded in dual-ballot elections over time and they know there is a recurring group of low-quality candidates that make screening more difficult. They just associate women to this group; a sort of pooling equilibrium for women and low-quality male candidates. This would feed back into parties' beliefs that voters tend to prefer male candidates in dual-ballot elections. Nonetheless, it is hard to believe that a major party label would not differentiate a female candidate, even in dual-ballot elections.

For minor parties, the pool of possible candidates is normally composed by underrepresented groups of the population. For example, unionized parties such as PSTU (United Socialist Workers' Party) normally field candidates of working class and participate more often of dual-ballot elections. Moreover, there is evidence in the

literature that women tend to be allocated to positions already known to be difficult to win Sanbonmatsu (2002); Murray (2008); Casas-Arce and Saiz (2015), something that would also explain why minor parties field more women in general. After all, they do not target a victory anyway. It is just that in dual-ballot the probability of a party getting zero votes is lower, making it less prone to shame.

Finally, in terms of the strategic components of dual-ballot elections, threads posed to the victory of mainstream parties by small and/or extremist parties are reduced and, as a result, their bargaining power to influence policy positions of major parties are also reduced (Nannicini, Bordignon and Tabellini, 2016). This factor would work towards the creation of a more permissive and stable environment for policy positions. It would not, however, interfere to a greater extent with the types of politicians elected or the type of performance analyzed here, because the competition is moved towards other grounds.

2.3.7 External validity

The RD empirical strategy is valid for municipalities with a population of around 200,000 voters. That is, municipalities that have just switched to dual-ballot or are close to. It is possible that the results found do not generalize to the rest of Brazilian municipalities.

The longer the time the municipality has been running elections under the dual-ballot system, the more adapted parties can become to the rules, as well as voters, and have a better understanding of the types of candidates each election requires. It is hard to say whether these differences may just fade over time, as part of a natural adaption to the new system. On the other hand, this initial condition could create a persistent gender gap in dual-ballot elections, independently of the characteristics that gave rise to it.

Despite this, the results obtained here speak to the analysis of whether electoral races under rules that supposedly allow for better screening of candidates are worth their costs. It is true that candidates spend, on average, the same financial resources in campaigning, but there are other non-monetary costs such as the time spent in the process, the higher number of candidates to choose from, the enhancement of underrepresentation of women and, more generally, the creation of barriers to entry of newly minted candidates.

My results show that, in practice, the types of candidates elected in each of these systems are virtually the same, with the drawback that dual-ballot mayors may need to make more use of political strategies, such as redistributive politics, to win elections.

In fact, dual-ballot major benefits seem to be related to its strategic components that allow voters to communicate policy preferences. Piketty (2000) defends with theoretical arguments the communicative function of runoff systems in that they allow citizens to vote in extreme candidates in the first round in order to send messages to their most preferred candidate so they can adjust their policies. In contrast, Bouton (2013), theoretically argues that runoff systems with a threshold below 50 percent may lead to the systematic victory of the Condorcet loser in an election with three candidates. Additionally, Nannicini, Bordignon and Tabellini (2016) shows empirically that extremist voters have a lower bargaining power in dual-ballot relative to single-ballot and, as a consequence, have less influence in policies that are implemented by mainstream parties.

2.4 Conclusion

This paper has analyzed whether there are differences in the selection process of candidates, both by parties and voters, when there is a change from single- to

dual-ballot elections under plurality rule. The identification strategy makes use of the quasi-random assignment of these electoral rules to assume that municipalities with population in the neighborhood of the 200,000 voters threshold of assignment are similar in a number of characteristics. This assumption is verified by balance tests for a host of characteristics. McCrary's test confirms the assumption that there is no manipulation in population counts as to induce a municipality to hold elections in a given system.

Candidates' observable characteristics such as occupational skill level, education and age are statistically the same in both systems for different selections of candidates by political strength. Surprisingly, however, there is a gender gap between the two: dual-ballot elections are less likely to have a woman in top positions, as defined by the share of votes obtained in the election.

To investigate whether the gender gap found is due to discrimination or there are other underlying skills that could well be required from both men and women in dual-ballot races, I also look at more specific political characteristics and some nuances of the electoral process. Dual-ballot elections have candidates with a larger experience in mayoral elections, a factor that can be an advantage when competing in parties' primaries but can also repeal other politicians from entry. The low historical presence of women in politics makes experience an important additional barrier to the entry of female candidates.

There is evidence in the literature that women get lower donations in electoral campaigns and, in general, have less resources to spend. The lack of access to financial resources is seen as major contributing factor for not winning an election and could therefore influence parties decisions. I test whether dual-ballot candidates actually raise and spend more money when compared to those in single-ballot. In fact, there are no detectable differences on what is donated and spent across the two systems, even when considering that second round elections would demand more resources.

This finding might impose even more strain on the participation of women, not only through the additional difficulties in gathering donations, but also by having to win an election that has already, on average, less resources per day of campaign.

To further understand these differences in the two systems reflected in the lack of female candidates in dual-ballot, I also look at the performance of winner candidates in office. Following the previous literature Ferraz and Finan (2008); Brollo and Troiano (2016); Gagliarducci and Paserman (2012), I consider a variable that reflects candidates' abilities to attract resources to the municipality - discretionary transfers of capital - and a variable that reflects the effectiveness in the application of non-discretionary resources – prenatal visits and non-premature births. Dual-ballot mayors attract more discretionary transfers of capital only in election years when they can run for reelection. It appears to be a compensatory policy for the comparatively low resources to spend in the campaign. In terms of health outcomes, there are no detectable differences.

All in all, these results are consistent with differentiation of candidates induced by the electoral system over directly unobservable characteristics that are reflected in the lower participation of women in dual-ballot elections.

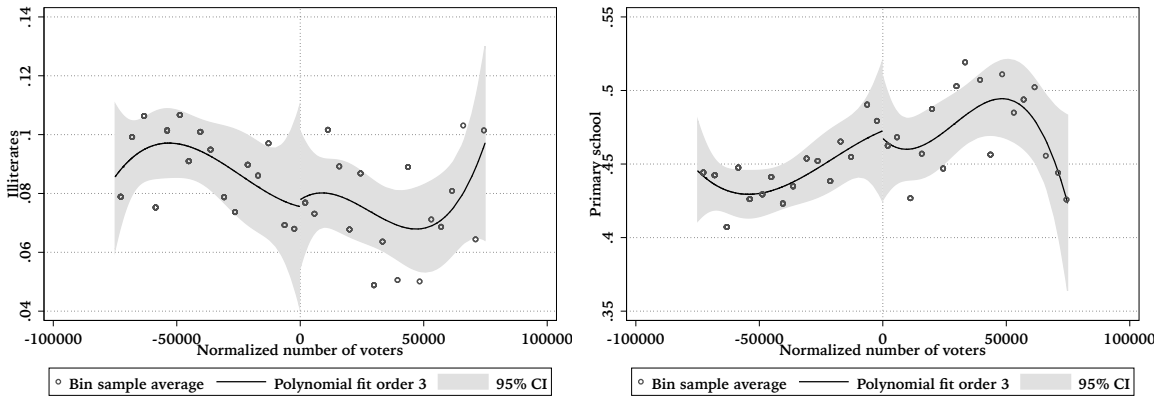
This paper exploited the mechanisms behind the supply side of politicians in dual-ballot elections but has not said anything about how voters respond to the lengthier electoral process and how this feeds back into politicians and party's strategies. Do they tend to participate more or less of the electoral process in the dual-ballot system? Because dual-ballot favors experienced politicians and require more strategic moves - such as attracting more discretionary transfers – it is also important to verify whether these characteristics are also associated with more cases of corruption or are beneficial to the municipalities following the rule.

In terms of public policy, improvements over candidates' characteristics and/or performance induced by simply changing electoral systems might be desirable in any

circumstances where welfare gains would be superior to the costs of setting up the new electoral process. Understanding all these interactions between electoral systems and mechanisms that induce politicians to perform better is an exciting avenue for future research.

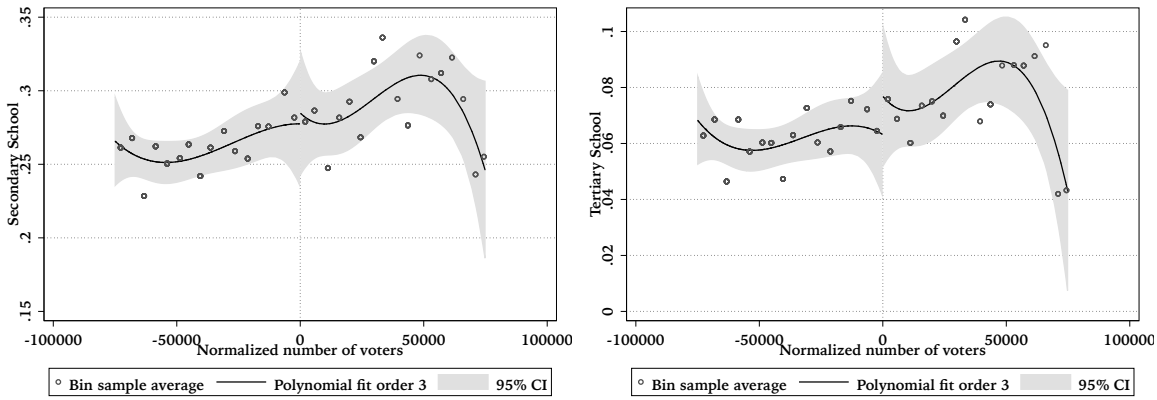
2.5 Appendix

2.5.1 Figures



(a) Share of illiterates (age 18+)

(b) Share of people with basic education (age 18+)

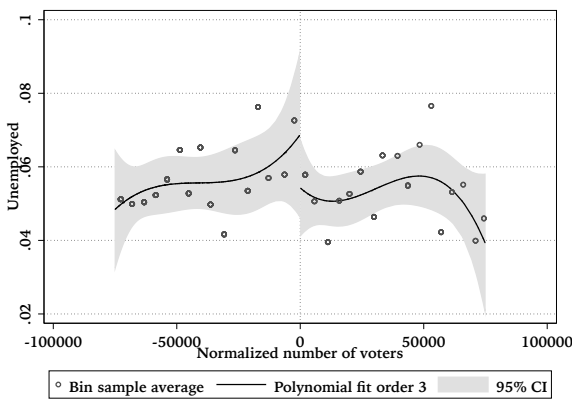


(c) Share of people with secondary education (age 18+)

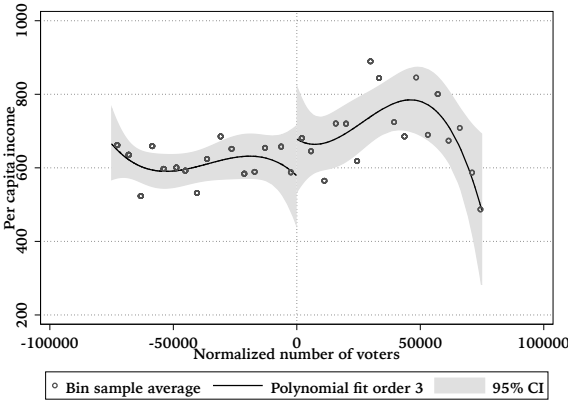
(d) Share of people with tertiary education (age 25+)

Figure 2.1: Balance tests for municipality characteristics: schooling.

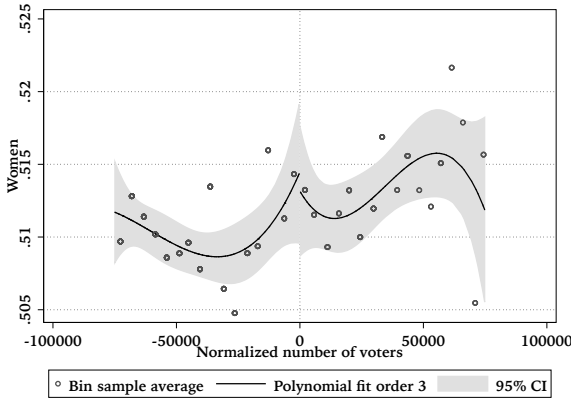
Note: Variables extracted from the 2000 Census. This sample considers all municipalities within a 75,000 distance from the threshold for visualization purposes. Scatter points are averaged over 327 voters, as determined by Imbens and Lemieux (2008).



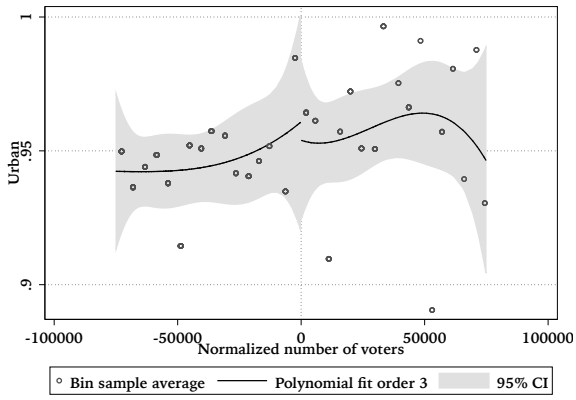
(a) Share unemployed (age 18+).



(b) Per capita income



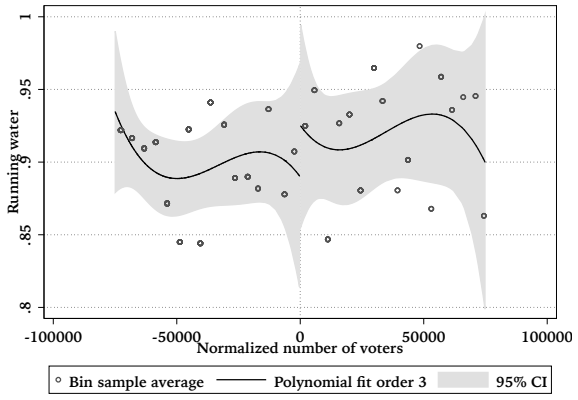
(c) Share of women in the population



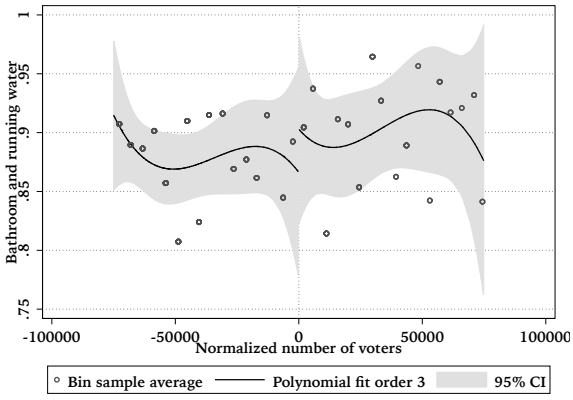
(d) Share of population living in urban area

Figure 2.2: Balance tests for municipality characteristics: income and demographics.

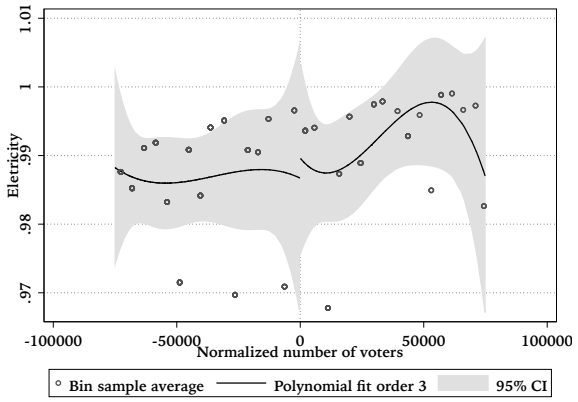
Note: Variables extracted from the 2000 Census. See 2.2 for definitions. This sample considers all municipalities within a 75,000 distance from the threshold for visualization purposes. Scatter points are averaged over 327 voters, as determined by Imbens and Kalyanaraman (2012).



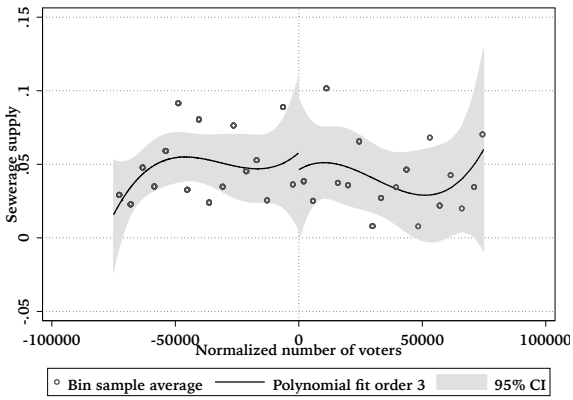
(a) Share of population with running water and bathroom



(b) Share of population with waste collection service



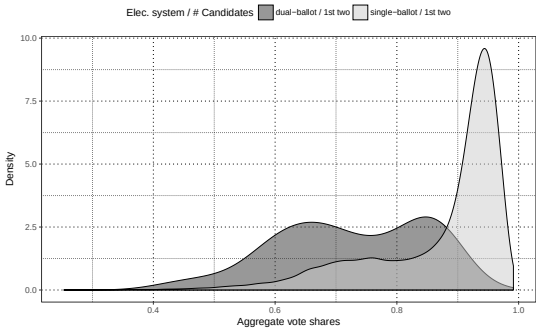
(c) Share of population with eletricity



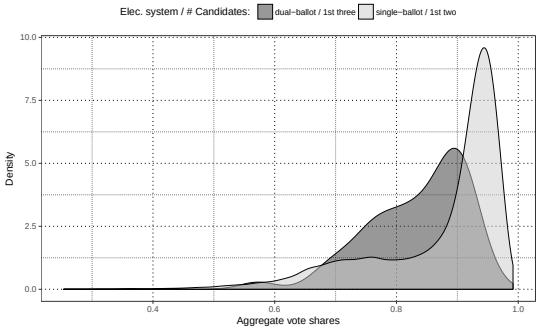
(d) Share of population with running water and sanitation

Figure 2.3: Balance tests for municipality characteristics: infrastructure.

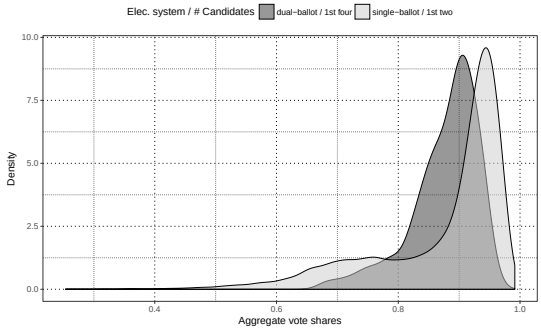
Note: Variables extracted from the 2000 Census. See 2.2 for definitions. This sample considers all municipalities within a 75,000 distance from the threshold for visualization purposes. Scatter points are averaged over 327 voters, as defined by Imbens and Kalyanaraman (2012).



(a) Density of the aggregate vote shares of top two ranked candidates in single- and dual-ballot races.



(b) Density of the aggregate vote shares of top two and three ranked candidates in single- and dual-ballot races, respectively.



(c) Density of the aggregate vote shares of top two and top four ranked candidates in single- and dual-ballot races, respectively.

Figure 2.4: Distribution of vote shares in single- and dual-ballot elections

Note: This sample considers all municipalities within a 75,000 distance from the threshold and all mayoral election held from 2000 to 2016.

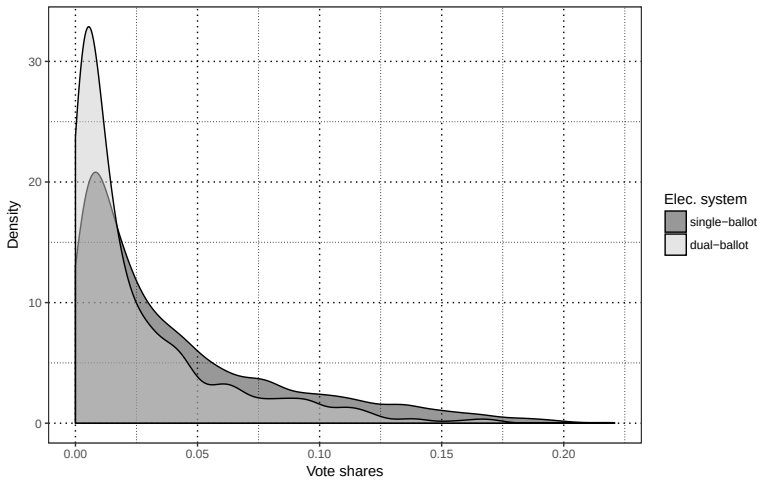


Figure 2.5: Density of aggregate vote shares of fourth or lower placed candidates

Note: The x-axis considers the sum of vote shares of fourth and lower placed candidates. This sample considers all mayoral races between 2000 and 2016 and municipalities within a 75,000 voters distance from the threshold.

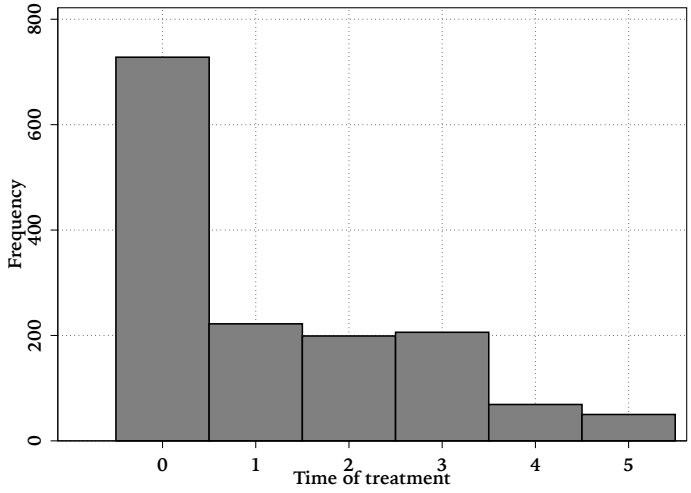


Figure 2.6: Distribution of the number of elections since conversion to dual-ballot.

Note: This sample considers all mayoral races between 2000 and 2016 and municipalities within a 75,000 voters distance from the threshold. Therefore, five means the municipality has run elections under dual-ballot in all races in the sample. Conversely, zero means the municipality has run all elections under single-ballot.

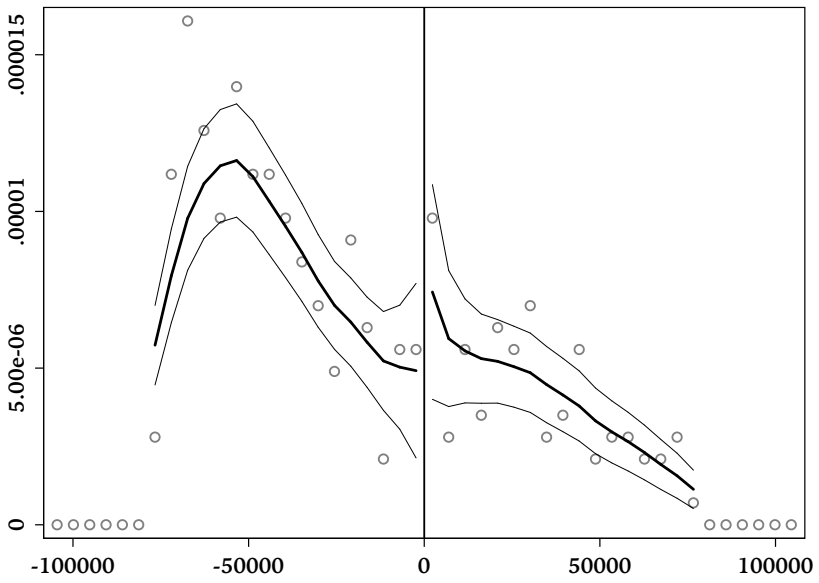
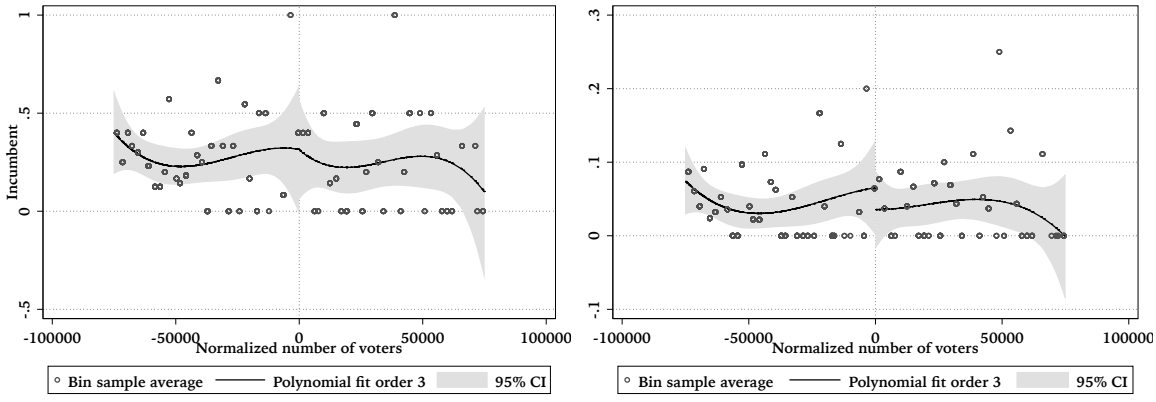


Figure 2.7: McCrary's continuity test for the number of voters in a municipality

Note: Kernel estimation of the density of the number of voters centered around the 200,000 threshold of dual-ballot allocation. Point estimate of discontinuity: .5360 with standard error .4481. Optimal bin-width and bin-size follows McCrary (2008). This sample considers all mayoral races between 2000 and 2016 and municipalities within a 75,000 distance from the threshold for visualization purposes. Test results do not vary with the sample considered.

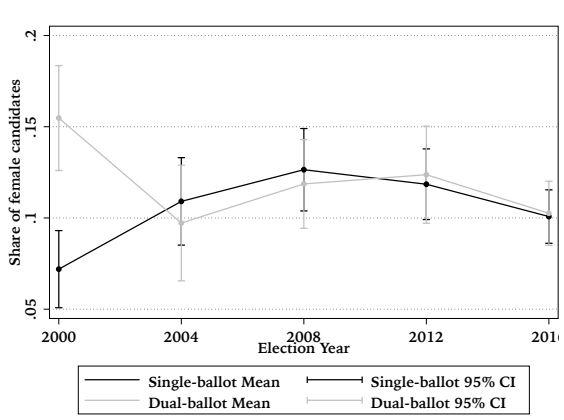


(a) Discontinuity in incumbency.

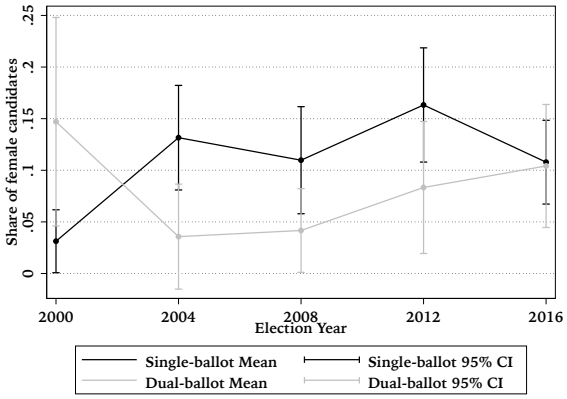
(b) Discontinuity in re-election.

Figure 2.8: Discontinuity in the outcomes of the elections.

Note: This sample considers all mayoral races between 2000 and 2016 and municipalities within a 75,000 voters distance from the threshold.



(a) Share of female candidates in each type of election over time.



(b) Share of female candidates in a top two position in each type of election over time.

Figure 2.9: Share of female candidates in dual-ballot vs Single-ballot.

Note: This sample considers all mayoral races between 2000 and 2016 and municipalities within a 75,000 voters distance from the threshold.

2.5.2 Tables

Table 2.1: Balance tests on municipalities' baseline characteristics.

Panel A: Schooling.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Illiterates	0.0910	15820.94	−0.0034	0.0028	0.0011	0.0089	−0.0028
	[0.0539]	(308)	[0.0081]	[0.0105]	[0.0159]	[0.0178]	[0.0151]
Primary school	0.4401	21925.30	−0.0034	−0.0070	−0.0037	−0.0192	0.0049
	[0.0748]	(308)	[0.0081]	[0.0272]	[0.0333]	[0.0292]	[0.0169]
Secondary School	0.2594	24621.21	−0.0034	0.0057	0.0046	−0.0047	0.0105
	[0.0657]	(308)	[0.0081]	[0.0253]	[0.0305]	[0.0232]	[0.0155]
Tertiary School	0.0614	23889.06	0.0129	0.0131	0.0140	0.0075	0.0184
	[0.0341]	(308)	[0.0086]	[0.0128]	[0.0149]	[0.0125]	[0.0164]
Observations	211		308	308	308	82	40

Panel B: Income and demographics.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Unemployed	0.5421	19184.65	−0.0174**	0.0307	0.0279	0.0327	0.0102
	[0.0548]	(308)	[0.0085]	[0.0196]	[0.0248]	[0.0236]	[0.0121]
Per capita income	611.2953	16780.72	−0.0174**	100.6014	128.3677	67.4590	94.2142**
	[210.4976]	(308)	[0.0085]	[75.3298]	[86.9779]	[85.5042]	[44.7340]
Women	0.5101	34859.90	−0.0174**	−0.0016	−0.0010	−0.0010	0.0018
	[0.0076]	(308)	[0.0085]	[0.0030]	[0.0037]	[0.0025]	[0.0018]
Urban	0.9451	17450.52	−0.0007	−0.0077	−0.0067	−0.0075	−0.0313
	[0.0630]	(308)	[0.0132]	[0.0162]	[0.0226]	[0.0218]	[0.0271]
Observations	211		308	308	308	58	34

Panel C: Sanitation.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Running water	0.9002	25260.67	0.0219	0.0375	0.0254	0.0030	0.0183
	[0.1210]	(308)	[0.0221]	[0.0486]	[0.0739]	[0.0463]	[0.0615]
Bathroom and running water	0.8805	24844.01	0.0219	0.0371	0.0286	0.0062	0.0198
	[0.1377]	(308)	[0.0221]	[0.0541]	[0.0824]	[0.0527]	[0.0243]
Eletricity	0.9868	20292.11	0.0219	0.0021	0.0056	0.0011	0.0006
	[0.0304]	(308)	[0.0221]	[0.0062]	[0.0115]	[0.0098]	[0.0052]
Sewerage supply	0.0467	24352.83	−0.0128	−0.0114	−0.0109	0.0032	−0.0043
	[0.0802]	(308)	[0.0142]	[0.0273]	[0.0478]	[0.0300]	[0.0479]
Observations	211		308	308	308	86	40

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively. See Table 2.2 for definitions.

Table 2.2: Definition of municipalities' baseline characteristics.

Variable	Description
Illiterate	Share of the population aged 18 or above that cannot read or write.
Primary School	Share of the population aged 18 or above with at least primary education.
Secondary School	Share of the population aged 18 or above with at least secondary education.
Tertiary School	Share of the population aged 25 or above with at least a university degree.
Unemployed	Share of the population aged 18 or above who were unemployed.
Per Capita Income	Average per capita for income.
Women	Share of women in the population.
Urban	Share of the population living in urban areas.
Running water	Share of households with water supply.
Bathroom and running water	Share of households with water supply and a bathroom.
Eletricity	Share of households with eletricity.
Sewerage supply	Share of households with sewerage supply.

Note: This table shows the definition of the variables used to measure schooling, income and sanitation at the municipality level. Variables are defined as in the 2000 Census.

Table 2.3: Dual-ballot treatment effects on candidates' observable characteristics.

Panel A: Sample includes all candidates in both single- and dual-ballot races.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Female	0.1079	26822.65	−0.0505	−0.0983	−0.0979	−0.0766	−0.0539
	[0.3104]	(1461)	[0.0354]	[0.0631]	[0.0813]	[0.0506]	[0.0709]
Age	49.8504	24122.24	−0.0505	1.5877	0.4159	1.0060	−0.6161
	[10.6737]	(1461)	[0.0354]	[2.6489]	[3.1055]	[2.3052]	[1.2402]
University Degree	0.7019	31496.82	−0.0505	−0.0822	−0.1389	−0.1436	−0.0146
	[0.4577]	(1461)	[0.0354]	[0.1274]	[0.1456]	[0.0997]	[0.0620]
High Skill	0.6218	19791.07	0.0481	0.0548	−0.0144	0.0414	0.0066
	[0.4852]	(1461)	[0.0426]	[0.1049]	[0.1335]	[0.0997]	[0.1531]
Observations	936		1461	1461	1461	339	181

Panel B: Sample excludes top two and three candidates in single- and dual-ballot races, respectively.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Female	0.1029	33361.95	0.0045	0.0029	0.0545	−0.0020	0.0906
	[0.3041]	(752)	[0.0532]	[0.0893]	[0.1089]	[0.0706]	[0.0885]
Age	48.8971	22200.00	0.0045	2.9944	0.3439	3.0693	0.8956
	[10.7334]	(752)	[0.0532]	[3.8555]	[4.6021]	[3.7320]	[1.8572]
University Degree	0.6388	24132.24	0.0045	−0.1826	−0.1736	−0.1338	−0.0325
	[0.4808]	(752)	[0.0532]	[0.1464]	[0.1812]	[0.1375]	[0.0775]
High Skill	0.5456	20789.34	0.0046	0.0113	−0.0952	−0.0065	−0.2042
	[0.4984]	(752)	[0.0680]	[0.1445]	[0.1990]	[0.1373]	[0.2088]
Observations	515		752	752	752	178	93

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively. See Section 2.3.1 for a definition of high skill.

Table 2.4: Dual-ballot treatment effects on candidates' observable characteristics.

Panel A: Subsample of top two ranked candidates in both single- and dual-ballot races.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Female	0.1140	26380.17	−0.1078*	−0.2199**	−0.3074***	−0.1952**	−0.2577**
	[0.3182]	(615)	[0.0590]	[0.0896]	[0.1093]	[0.0812]	[0.1006]
Age	51.0166	21059.50	−0.1078*	1.3246	0.8900	1.1463	−2.0160
	[10.4957]	(615)	[0.0590]	[2.8500]	[3.3837]	[2.6262]	[1.7522]
University Degree	0.7791	26827.01	−0.1078*	0.0367	−0.1243	−0.0643	0.0350
	[0.4153]	(615)	[0.0590]	[0.1644]	[0.1738]	[0.1501]	[0.0752]
High Skill	0.7150	34166.69	0.0657	0.1306	0.0423	0.0633	0.0800
	[0.4520]	(615)	[0.0794]	[0.1515]	[0.1916]	[0.1098]	[0.1534]
Observations	421		615	615	615	242	114

Panel B: Subsample of top two and top three ranked candidates in single- and dual-ballot races, respectively.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Female	0.1140	27292.10	−0.1055**	−0.1967**	−0.2340**	−0.1638**	−0.1959*
	[0.3182]	(709)	[0.0530]	[0.0935]	[0.1144]	[0.0697]	[0.0987]
Age	51.0166	35036.14	−0.1055**	0.1460	0.2439	−1.0555	−2.3156
	[10.4957]	(709)	[0.0530]	[2.7042]	[3.2674]	[1.9185]	[1.5255]
University Degree	0.7791	26577.84	−0.1055**	0.0188	−0.1601	−0.1022	−0.0161
	[0.4153]	(709)	[0.0530]	[0.1522]	[0.1687]	[0.1265]	[0.0725]
High Skill	0.7150	33480.04	0.0666	0.0707	−0.0080	0.0242	0.0419
	[0.4520]	(709)	[0.0677]	[0.1397]	[0.1686]	[0.0966]	[0.1345]
Observations	421		709	709	709	296	135

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively. See Section 2.3.1 for a definition of high skill.

Table 2.5: Dual-ballot treatment effects on observable candidates' characteristics by gender.

Panel A: All female candidates in both single- and dual-ballot races.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Age	47.8119	30201.10	1.3894	12.6594**	5.7498	8.9580*	1.3894
	[8.7060]	(162)	[2.9905]	[5.8162]	[5.4215]	[4.6990]	[2.9905]
University Degree	0.7822	22937.81	1.3894	0.2446	−0.0815	0.2335	0.2509*
	[0.4148]	(162)	[2.9905]	[0.2141]	[0.2725]	[0.1859]	[0.1420]
High Skill	0.5644	33472.27	0.1179	−0.1956	−0.3894	−0.2429	−0.1409
	[0.4983]	(162)	[0.1387]	[0.2557]	[0.3200]	[0.2013]	[0.3174]
Observations	101		162	162	162	76	30

Panel B: All male candidates in both single- and dual-ballot races.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Age	50.0970	22712.83	−0.7783	0.0091	0.0558	0.6205	−0.7783
	[10.8662]	(1299)	[1.2710]	[2.6314]	[3.0899]	[2.1613]	[1.2710]
University Degree	0.6922	26540.53	−0.7783	−0.1276	−0.1274	−0.1584	−0.0431
	[0.4619]	(1299)	[1.2710]	[0.1366]	[0.1547]	[0.1106]	[0.0630]
High Skill	0.6287	19565.27	0.0439	0.0869	0.0405	0.0670	0.0460
	[0.4834]	(1299)	[0.0467]	[0.0968]	[0.1121]	[0.0892]	[0.1368]
Observations	835		1299	1299	1299	291	159

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively. See Section 2.3.1 for a definition of high skill.

Table 2.6: Dual-ballot treatment effects on campaign donations and spending.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Total spending 1st round	1.5936	24606.31	−0.8073***	−0.0930	−0.1557	−0.1555	−0.4627
	[1.7184]	(499)	[0.2923]	[0.5174]	[0.6049]	[0.4864]	[0.5577]
Total spending 1st/2nd round	0.5312	25575.12	−0.8073***	−0.0006	−0.1466	−0.0073	−0.1727
	[0.5728]	(487)	[0.2923]	[0.2300]	[0.2616]	[0.2094]	[0.1174]
Total spending	1.8251	26482.71	−0.8073***	0.1606	−0.0944	0.1357	−0.6747**
	[1.8846]	(417)	[0.2923]	[0.6049]	[0.7093]	[0.5935]	[0.3042]
Fundraising total 1st round	1.4019	22399.17	−0.8073***	0.0232	0.1211	0.0542	−0.5448**
	[1.6042]	(518)	[0.2923]	[0.4879]	[0.6139]	[0.4765]	[0.2597]
Fundraising total 1st/2nd round	0.4673	27873.58	−0.8073***	0.2707	0.1928	0.2607	−0.0548
	[0.5347]	(518)	[0.2923]	[0.2243]	[0.2817]	[0.1862]	[0.1051]
Fundraising total	1.6721	24737.74	−0.4497	0.3049	0.4446	0.3633	0.1192
	[1.9320]	(518)	[0.3289]	[0.6184]	[0.7567]	[0.5746]	[0.7599]
Observations	358		518	518	518	146	70

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively.

Table 2.7: Dual-ballot treatment effects on sources of campaign funding for top two candidates.

Panel A: Campaign funding (per capita) raised in the first round.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Events	0.0196	22740.40	−0.0179	−0.0111	−0.0460	−0.0245	−0.0085
	[0.1173]	(467)	[0.0172]	[0.0304]	[0.0339]	[0.0228]	[0.0204]
Voters	0.2891	26616.29	−0.0179	−0.0801	−0.0018	−0.0562	−0.0476
	[0.3622]	(467)	[0.0172]	[0.1208]	[0.1804]	[0.1089]	[0.0794]
Companies	0.4894	20056.32	−0.0179	0.0789	0.0971	0.0833	−0.2719**
	[0.8542]	(467)	[0.0172]	[0.2671]	[0.3493]	[0.2732]	[0.1350]
Public funding	0.0076	20816.63	−0.0179	−0.0054	−0.0069	−0.0063	0.0076
	[0.0986]	(467)	[0.0172]	[0.0150]	[0.0170]	[0.0067]	[0.0080]
Party funding	0.2816	19036.88	−0.0179	0.1652	0.2484	0.2651*	−0.1252
	[0.6116]	(467)	[0.0172]	[0.1278]	[0.1775]	[0.1508]	[0.0900]
Unknown	0.0063	7086.51	0.0068	−0.0072	0.0121	0.0002	0.0001
	[0.0989]	(467)	[0.0064]	[0.0099]	[0.0131]	[0.0002]	[0.0003]
Observations	327		467	467	467	41	19

Panel B: Campaign funding (per capita) raised in the election.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Events	0.0215	32912.35	−0.0287	−0.0245	−0.0688	−0.0427	−0.0435
	[0.1231]	(467)	[0.0204]	[0.0399]	[0.0435]	[0.0255]	[0.0355]
Voters	0.3619	32879.41	−0.0287	−0.1231	−0.0441	−0.0430	0.0078
	[0.4526]	(467)	[0.0204]	[0.1346]	[0.1943]	[0.1148]	[0.0901]
Companies	0.5620	21228.35	−0.0287	0.3074	0.5451	0.4938	−0.1573
	[0.9738]	(467)	[0.0204]	[0.3488]	[0.4770]	[0.4054]	[0.1533]
Public funding	0.0078	22013.48	−0.0287	−0.0045	−0.0064	−0.0065	0.0069
	[0.0987]	(467)	[0.0204]	[0.0151]	[0.0171]	[0.0067]	[0.0081]
Party funding	0.3188	23050.91	−0.0287	0.3034*	0.2995	0.1863	−0.1014
	[0.7171]	(467)	[0.0204]	[0.1567]	[0.2203]	[0.1283]	[0.1054]
Unknown	0.0063	7594.75	0.0069	−0.0072	0.0120	0.0002	0.0001
	[0.0989]	(467)	[0.0064]	[0.0099]	[0.0131]	[0.0002]	[0.0003]
Observations	327		467	467	467	43	21

Panel B: Campaign funding (per capita) raised in the first round of single-ballot (monthly average) against the second round of dual-ballot.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Events	0.0065	19681.78	−0.0073	0.0069	−0.0161	−0.0066	0.0010
	[0.0391]	(467)	[0.0059]	[0.0144]	[0.0117]	[0.0067]	[0.0070]
Voters	0.0964	16246.58	−0.0073	−0.0738	−0.0640	−0.0521	−0.0114
	[0.1207]	(467)	[0.0059]	[0.0507]	[0.0666]	[0.0622]	[0.0303]
Companies	0.1631	19823.29	−0.0073	0.2332	0.3336	0.3026	−0.0152
	[0.2847]	(467)	[0.0059]	[0.1807]	[0.2438]	[0.2164]	[0.0652]
Public funding	0.0025	20993.23	−0.0073	−0.0018	−0.0029	−0.0021	0.0024
	[0.0329]	(467)	[0.0059]	[0.0050]	[0.0056]	[0.0022]	[0.0027]
Party funding	0.0939	21581.31	−0.0073	0.1316**	0.0843	0.1159	−0.0094
	[0.2039]	(467)	[0.0059]	[0.0644]	[0.0868]	[0.0689]	[0.0383]
Unknown	0.0021	8251.18	0.0023	−0.0024	0.0039	0.0000	−0.0000
	[0.0330]	(467)	[0.0021]	[0.0033]	[0.0044]	[0.0000]	[0.0001]
Observations	327		467	467	467	45	25
R^2			0.016	0.018	0.022	0.072	0.126

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively.

Table 2.8: Dual-ballot treatment effects on political characteristics.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Experience in Mayoral elections	0.5173	18665.06	−0.1248	0.2458**	0.2632**	0.1927	0.3679**
	[0.8239]	(1251)	[0.0865]	[0.1193]	[0.1205]	[0.1204]	[0.1573]
Margin of votes (1st round)	0.1195	27371.89	−0.0389***	−0.0528*	−0.0001	−0.0232	−0.0242
	[0.1248]	(1163)	[0.0129]	[0.0314]	[0.0345]	[0.0251]	[0.0306]
Margin of votes (1st SB vs 2nd DB)	0.1750	23046.05	−0.1145**	−0.1991***	−0.1210	−0.1145**	−0.1145**
	[0.1534]	(264)	[0.0439]	[0.0744]	[0.0903]	[0.0439]	[0.0439]
Incumbent race	0.3388	31660.89	−0.1145**	0.2007	0.1568	−0.0198	0.0144
	[0.4746]	(267)	[0.0439]	[0.1934]	[0.2151]	[0.1524]	[0.1113]
Reelection Win	0.2240	35966.01	−0.1145**	−0.0424	−0.1744	−0.0583	0.0490
	[0.4181]	(267)	[0.0439]	[0.1875]	[0.2154]	[0.1479]	[0.1100]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively.

Table 2.9: Dual-ballot treatment effects on health outcomes.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Non-premature births	0.3991	20556.50	−0.0035	0.0051	0.0083	0.0045	−0.0014
	[0.2605]	(259)	[0.0039]	[0.0058]	[0.0066]	[0.0054]	[0.0055]
Any prenatal visits	0.6294	20238.87	−0.0035	0.0053	0.0056	0.0071	−0.0013
	[0.4109]	(259)	[0.0039]	[0.0081]	[0.0088]	[0.0088]	[0.0040]
Non-premature births (3 years)	0.4586	21293.10	−0.0035	0.0069	0.0104	0.0054	−0.0045
	[0.3281]	(259)	[0.0039]	[0.0077]	[0.0088]	[0.0072]	[0.0052]
Any prenatal visits (3 years)	0.6963	20278.54	−0.0035	0.0070	0.0085	0.0090	−0.0022
	[0.4410]	(259)	[0.0039]	[0.0092]	[0.0096]	[0.0094]	[0.0043]
Non-premature births (2nd and 3rd years)	0.4598	21402.52	−0.0035	0.0072	0.0095	0.0061	−0.0039
	[0.3296]	(259)	[0.0039]	[0.0072]	[0.0079]	[0.0065]	[0.0047]
Any prenatal visits (2nd and 3rd years)	0.6969	20262.91	−0.0035	0.0058	0.0066	0.0070	−0.0033
	[0.4414]	(259)	[0.0039]	[0.0087]	[0.0093]	[0.0092]	[0.0042]
Non-premature births (election year)	0.4998	23324.97	−0.0035	0.0001	0.0034	0.0013	−0.0010
	[0.0051]	(117)	[0.0039]	[0.0025]	[0.0028]	[0.0020]	[0.0015]
Any prenatal visits (election year)	0.9718	25087.39	0.0024	−0.0017	−0.0076	0.0069	−0.0082
	[0.0246]	(117)	[0.0083]	[0.0145]	[0.0163]	[0.0139]	[0.0103]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively.

Table 2.10: Dual-ballot treatment effects on transfers for capital investment.

Panel A: Transfers for capital investment for the full sample of municipalities.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Discretionary transfers	11.2495	20198.71	0.7054	0.7177	0.4819	1.0945	−0.6584
	[7.4749]	(259)	[0.5963]	[1.2574]	[1.1731]	[1.0130]	[1.1002]
Non-discretionary transfers	9.7850	26258.37	0.7054	1.2183	1.4717	1.3093	−0.2304
	[7.0049]	(259)	[0.5963]	[2.4039]	[2.8766]	[1.9703]	[1.0876]
Discretionary transfers (3 years)	10.9158	24598.71	0.7054	−1.5162	−2.5059	−0.5637	0.2314
	[7.3710]	(259)	[0.5963]	[2.1963]	[2.7931]	[2.1687]	[0.7887]
Non-discretionary transfers (3 years)	9.3919	25862.92	0.7054	1.0124	1.7797	1.1275	−0.2315
	[6.9197]	(259)	[0.5963]	[2.4183]	[2.9401]	[1.9959]	[1.0914]
Discretionary transfers (election year)	14.4140	21211.90	0.7054	0.4458	−0.7116	0.8725	0.9104
	[3.5504]	(117)	[0.5963]	[1.1797]	[1.0613]	[0.8948]	[0.8030]
Non-discretionary transfers (election year)	11.6159	22662.32	−1.0449	−0.5182	0.0560	−0.3493	−2.1525
	[5.0868]	(117)	[1.7881]	[2.7917]	[3.1884]	[2.3069]	[3.1675]

Panel B: Transfers for capital investment in municipalities with first-term mayors.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Discretionary transfers	8.8783	32234.27	0.1901	0.3141	1.1695	0.8175	1.1701
	[8.1251]	(91)	[0.3343]	[0.7328]	[0.7406]	[0.5667]	[1.0578]
Non-discretionary transfers	7.7169	27391.21	0.1901	2.2137	0.4023	0.3054	−0.7342
	[7.2791]	(91)	[0.3343]	[1.7903]	[1.2885]	[1.4465]	[0.8364]
Discretionary transfers (3 years)	8.7493	32678.30	0.1901	0.1336	1.0432	0.5741	0.1660
	[8.0104]	(91)	[0.3343]	[0.7317]	[0.6916]	[0.5475]	[0.3578]
Non-discretionary transfers (3 years)	7.5571	27786.77	0.1901	2.3314	0.9430	0.2077	−1.0827
	[7.1491]	(91)	[0.3343]	[1.7491]	[1.5187]	[1.3130]	[0.8544]
Discretionary transfers (election year)	15.3340	18576.04	0.1901	5.6883***	7.6140***	2.9985	0.6255
	[1.0154]	(28)	[0.3343]	[0.7192]	[1.6457]	[1.6196]	[0.7999]
Non-discretionary transfers (election year)	12.0976	10839.17	2.1985	5.6553	3.1832	6.2464	–
	[3.9729]	(28)	[2.8372]	[5.9781]	[8.8505]	[3.6229]	[.]

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively. Cells with dash indicate there were not enough observations to run a model in that bandwidth.

Table 2.11: Dual-ballot treatment effects on turnout, spoilt and blank votes.

	(1) Mean	(2) BW	(3) OLS	(4) Spline 3rd	(5) Spline 4th	(6) Optimal BW	(7) Half Optimal BW
Blank	0.0294	20742.22	−0.0898	0.0014	−0.0003	−0.0021	−0.0012
	[0.0168]	(233)	[0.0840]	[0.0065]	[0.0085]	[0.0065]	[0.0044]
Spoilt	0.0454	22518.92	−0.0898	0.0034	0.0052	0.0054	−0.0006
	[0.0127]	(233)	[0.0840]	[0.0066]	[0.0079]	[0.0061]	[0.0042]
Turnout	0.8453	33546.13	0.0033	−0.0186	−0.0156	−0.0149	−0.0093
	[0.0371]	(307)	[0.0092]	[0.0128]	[0.0159]	[0.0109]	[0.0171]
Observations	211		307	307	307	119	53
R^2			0.117	0.126	0.135	0.158	0.149

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets. Sample size in parenthesis.

Note: Column (1) displays the corresponding variable mean and standard deviation (in square brackets) in single-ballot elections. Column (2) displays the optimal bandwidth value obtained through Imbens and Kalyanaraman (2012) and the sample size used in the calculation is displayed below in parenthesis. Columns (4) and (5) display parametric estimates considering polynomial splines of 3rd and 4th order, respectively, using a sample of municipalities within a 75,000 voters distance from the zero cutoff. Column (3) displays baseline OLS results using the same sample. Columns (6) and (7) display non-parametric estimates with samples restricted to the optimal bandwidth in (2) and half of it, respectively.

Table 2.12: Differences in turnout, spoilt and blank votes across rounds of dual-ballot elections.

Panel A: Sample of all elections within 75000 voters from the cuttoff.

	Difference	Std. Error	Obs.
Blank	0.0125***	0.0026	273
Spoilt	0.0185***	0.0021	273
Turnout	0.0174***	0.0055	273

Panel B: Sample of close elections within 75000 voters from the cuttoff.

	Difference	Std. Error	Obs.
Blank	0.0109***	0.0027	134
Spoilt	0.0176***	0.0021	134
Turnout	0.0145**	0.0066	134

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Note: Mean test of the differences in the share of votes of the correspondent variable between first and second rounds of dual-ballot elections. Close elections in Panel B are defined as elections where the top candidate received no more than ten percentage points more votes with respect to the second placed candidate in the first round.

Chapter 3

Voting motivations and political corruption

3.1 Introduction

The literature has identified many reasons why citizens turnout to vote. They range from strictly rational, such as the probability of altering election outcomes or the expected returns in social welfare, to strictly behavioral, such as voting to impress others or fulfill a sense of civic duty. It is natural to think that these considerations not only determine voters' turnout but also interact with the choice of candidates. Whether politicians take into account the composition of voter types in the electorate when deciding about rents extraction is still an open question.

According to political agency models, candidates are inherently corrupt and decide about the extraction of rents based on voters' capacity to hold them accountable. Elections are the primary instrument for political accountability. Ousting corrupt politicians from office, nevertheless, will depend on voters knowledge of corruption events as well as how they update their beliefs about the corruption levels. Voters who care less about public policies or who pay extra costs in acquiring information

may have a diminished capacity for accountability. Candidates' knowledge about the size of this group and how it votes may enhance their opportunities for rent extraction without major electoral drawbacks.

In this paper, I use a quasi-naturally generated group of voters with differential political information and motivations for voting to investigate whether this has any consequences to the way they vote and how politicians respond to it. The electoral legislation in Brazil allows internal migrants to remain voting in their municipality of origin if they so wish. Indeed, a large number of them return to vote, as the average gap between the number of resident voters and the number of registered voters in net sender municipalities reveals.

I exploit the variation in registration rates across Brazilian municipalities to analyze whether the size of this group of migrant voters has any implications to politicians accountability. Mayoral corruption is measured using the well known Brazilian random audits program considering three electoral cycles: 2004, 2008 and 2012.

Many identification issues emerge from this setting of double selection: citizens choose to migrate and then where to vote. More than that, there can be a simultaneous causation between corruption, migration and registration rates. For example, local government corruption levels may negatively affect the economy and become a push factor for out-migration. This can either be positively or negatively related to registration rates. On the one hand, voters could choose to remain voting in their municipality of origin to try and improve the quality of politicians elected. On the other hand, these voters could instead get discouraged with politics and request a transfer to vote in their destination. In addition, municipalities with high migration levels have non-observable socioeconomic and political characteristics that are correlated with both the quality of local governments and migration.

The identification of the causal effects of changes in registration rates on corruption is possible by using the total variation of registration rates in destination

municipalities, net of the change in registration rates in the municipality of origin itself, as instruments for changes in registration rates in source municipalities. Because registration rates and migration are intrinsically related, migration by itself could have an effect on corruption; changes in registration rates could also be capturing that. I can separate these effects by adding arguably exogenous measures of in- and out-migration as controls in the equation that estimates the effects of registration rates. For this, I follow the standard approach of the labor/migration literature that uses the proportion of past immigrants of the same origin in a given destination for apportioning contemporaneous migration to the source location, net of its contemporaneous migration itself.

I find that positive changes in registration rates generate a sizable causal increase in the number of corruption events found in a municipality; politicians in these municipalities also perform worse in terms of the attraction of discretionary governmental transfers. Because the bulk of the budget of a municipality is composed by earmarked resources, extracting rents from discretionary transfers would not be a close substitute. At the least, these results imply that politicians in less corrupt municipalities are more concerned with their reputation and try to extract rents from non-monitored resources, which would imply both less cases of corruption found and less cases of non-detectable corruption. The impact of changes in registration rates on corruption is not entirely explained by neither fraud in registration rates nor immigration or return migration.

To try and shed some light on the mechanisms behind these effects, I first identify voters' motivations for keeping their registration in the source municipality. The idea is that when migrant voters decide where to vote they are actually choosing between elections. To evaluate whether the size of an election, and therefore pivotality, plays a role in this decision I exploit the discontinuities in the allocation of the maximum number of councillor seats across municipalities. These thresholds increase the re-

quirement in the number of votes a party needs to elect a councillor and, therefore, the size of an election. Using a fuzzy regression discontinuity design based on these population thresholds, I argue that one strong motivation to remain voting in their city of origin is the size of the election: the likelihood of being pivotal in smaller elections is greater.

Candidates in turn can observe registration rates. There is probably some uncertainty about turnout rates, but historical rates should be a good predictor and voting in Brazil is compulsory. If candidates think that voting motivations for migrant voters have a minor component that corresponds to benefits from public goods and know that they have limited information about their performance in office, conditional on the size of this group of voters, inconsequential extraction of rents can become easier. Because there is positive selection in out-migration, imitating the behavior of resident voters should on average reproduce choices of lower quality, on the assumption that less educated voters are also less sophisticated. Note that this explanation does not require that migrant voters actually return to vote systematically, only that party elites believe that they can return. The total out-migration effects on corruption stemming from the loss of more sophisticated voters (these include both those who return and those who do not) and restricting the pool of citizens that could possibly become better candidates do not seem to be as strong as the one of registration rates. Perhaps because those who return to vote can be persuaded into offsetting the power of remaining voters who care about performance and public goods.

This paper resides in three branches of literature: voting behavior, political corruption and turnout. It relates most closely to the literature on political corruption and electoral accountability. These studies have focused on two mechanisms: electoral institutions that incentivize good performance and press coverage that informs voters about politicians behavior (see e.g. Ferraz and Finan (2011); Snyder and Strömberg (2010). This paper focuses on voting motivation and its consequences for political

corruption. Turnout of migrant voters seems to be consistent with instrumental theories of voting, by which voters condition their decisions on the event of being pivotal (Feddersen and Pesendorfer, 1995). The empirical literature on this is still scarce, with the more relevant studies being the laboratory experiments of Battaglini, Morton and Palfrey (2009) and the structural results of (Kawai and Watanabe, 2013). To a lesser extent, my results are also related to the expressive theories of voting whereby citizens vote because of social pressures (Dellavigna, List and Malmendier, 2015) or because it is a civic duty to do so (Degan and Merlo, 2011).

The paper is structured as follows. Section 3.2 describes the institutional background in Brazilian elections and datasets used. Sections 3.3 and 3.4 focus on implementation issues and empirical strategy. In Section 3.5 I present my first set of results, the effects of registration rates on corruption and performance, along with a discussion of the relationship between migration, registration rates and these outcomes. Section 3.6 tries to understand the mechanisms through which registration rates may affect corruption. Section 3.7 presents some robustness checks and Section 3.8 concludes the paper.

3.2 Institutional Background and Data

This section describes the institutional framework and the data used in the empirical analysis.

3.2.1 Electoral Institutions

Brazil is a federation with 26 states and one federal district divided into 5567 municipalities. Each municipality chooses a new mayor and local legislature every four years. All of them enjoy the same legal status. The Constitution establishes a threshold of 200,000 registered voters to determine whether a municipality holds two-

or one-round mayoral election. Municipalities above the threshold have a second round election with the top two candidates of the first round when none of them has obtained more than 50% of the valid votes. The election for councillors is held simultaneously to that of males and follows the proportional rule to allocate seats. The maximum number of seats available in a municipality is exogenously determined by population thresholds.

All citizens above 18 and below 71 years of age are required to vote and face penalties if they do not and fail to justify their absence. Having a clean electoral record is necessary in order to obtain and regularize key documents, take office as a civil servant and enrol on higher education institutions. To register for voting in a municipality, a voter is requested to present a proof of address and has to have been living there for at least three months. Once the voter is registered, in the case she moves out to another municipality, she has the option to request a registration transfer or remain voting in her municipality of origin. The longer an individual has lived outside her voting municipality the more likely it is that the electoral authority has held a voters' re-registration program; in the case of re-registration, the voter must travel back to the city, attend the local electoral court and present an address proof to be able to keep their electoral register up to date at that municipality.

There have been two such programs at the national level: one in 1986, so as to build an electronic database containing all the electorate information; and one that has gradually been rolled-out since 2008, so as to collect biometric data of all the electorate.

A local electoral authority can, however, run an electorate revision at its own discretion when there is clear evidence of fraud or all of the issues below are detected:

1. the number of voters who request a transfer to a given municipality in a year is 10% higher than in the previous year;

2. the number of registered voters is higher than the number of residents with age 70 or over plus two times the number of residents with ages between 10 and 15.

3. the number of voters in the municipality is higher than 65% of the population estimated by the Brazilian Institute for Geography and Statistics (IBGE) for that year.

These electorate revisions when approved by a judge are not normally run in election years, only in exceptional circumstances. Many legal revision requests have been submitted to electoral authorities over time, specially by opposition parties. All under the allegation of fraud: the existence of "ghost voters". They are basically dead people that still manage to turn up and vote. However, very few of these requests have been approved because not all the conditions above were met, as the thresholds stablished by the law are very high.[1] While it might be true that part of the high registration rates is due to fraud, as we will see, they are also highly correlated with out-migration rates. This indicates that at least part of the voters exceeding the resident population in voting age is composed by those who choose to remain voting in their home municipality.

Another aspect that also facilitates their return is that elections are always held on the first Sunday of October, from 8am to 5pm. For municipalities with second rounds, they are held on the last Sunday of October. The election day is considered a national holiday. This is to guarantee that the majority of voters will not be working on election day. If they do, employers are legally required to allow them time to go and vote. Since voting is mandatory, voters do not necessarily have to cast a vote for a candidate, the options for voting blank or spoiling the vote are also available. For all these reasons, turnout for all elections in Brazil is very high.

[1] Apparently, the criterion used to get to these thresholds was that the number of irregularly registered voters should be high enough to affect the outcome of an election.

3.2.2 Data sources and implementation

Voters Data

The data on the number of registered voters in each municipality is provided by the national electoral authority, Tribunal Superior Eleitoral (TSE). It releases detailed data on the number of voters by age groups in election years. Therefore, the available data is biennial and stretches from 1996 to 2014.

I use IBGE Censuses and Population Counts[2] to construct a database of resident voters by age groups in each municipality. I rely on the annual population estimates provided by DataSus (Departamento de Informática do Sistema Único de Saúde), the statistical division of the Ministry of Health, to calculate the number of resident voters for the remaining years. DataSus takes the population estimates provided by IBGE[3], used to allocate intergovernmental transfers, that are based on population size, across municipalities (see Brollo et al. (2013)), and stratifies them by age. These are the only available estimates of the Brazilian population by municipality and age. I discuss the precision of these estimates in Subsection 3.3.2.

By matching the datasets from the two sources, I can calculate the ratio between the number of registered voters and the number of voting age people who actually live in the municipality. A ratio greater than one indicates that the municipality has more voters than resident voters.

Elections and Candidates Data

Data on mayoral elections outcomes and candidates' characteristics are also available from the electoral authority (TSE) for the election years 1996, 2000, 2004, 2008

[2]IBGE counts the population of all municipalities with up to 150,000 residents 5 years past the Census. I use years 1996 and 2007.

[3]The data and estimates methodology are available at `http://tabnet.datasus.gov.br/cgi/ibge/popdescr.htm`.

and 2012. Here I use the same dataset on candidates characteristics described in Chapter 2.

These characteristics include age, gender, education and occupation.[4] Based on occupational skill levels, I create a dummy variable called *High-Medium Skill*, that is one for groups of occupations that require the execution of more complex tasks and zero otherwise. I use this variable as a proxy for politicians' quality.

Measuring Mayors' Performance

To measure mayors performance I follow Fujiwara (2010) and Brollo and Troiano (2016) and consider the attraction of discretionary transfers of capital.

Data on discretionary transfers are available from Brazilian National Treasury (Tesouro Nacional) for most of the municipalities[5] and years. As in Brollo and Troiano (2016), I consider discretionary transfers of capital as a measure of politicians' ability to attract resources to the municipality, since they depend on agreements (*convênios*) made between the municipality and the federal or state governments. I consider the average transfers received during the first three years of the four-year mayoral term as my variable of interest. This is to net out effects of political cycles in transfers.[6]

Corruption Data

The data on corruption comes from an audit program initiated in 2003 by Controladoria Geral da União (CGU), that randomly selects municipalities for the application of rigorous audits. This data has also been used by Ferraz and Finan (2011) and Brollo et al. (2013). The institution gathers information about irregularities in

[4] As defined by the International Standard Classification of Occupations 2008 (ISCO08) - ILO.

[5] The data is self-reported by the municipalities and is available for 95% of them. I keep in the sample those municipalities that reported transfers for all the years.

[6] All the values are in R$ 2000 prices.

the use of federal transfers and classify them in a scale of corruption according to how serious they are:

(i) *Mild Corruption* These are irregularities more related to the lack of technical knowledge in the management of the resources than with corruption itself. They are mistakes that do not affect the application of the resources in question.

(ii) *Moderate Corruption* These irregularities will affect the execution of the programs funded by the resources involved, but according to the judgement of CGU's auditors do not significantly compromise their outcomes. This category can include the same bad practices of those classified as severe corruption but differ in their intensity.

(iii) *Severe Corruption* According to CGU these are the irregularities that will really compromise the execution of programs funded by the resources involved. Examples of these are favoritism in procurement, purchase of overpriced products, lack of purchase receipts and delay in the execution of infrastructure projects.

The data contain the number of irregularities in each category for the municipalities audited. To relieve concerns about the subjectivity in the classification between moderate and severe corruption, I put together these two types of corruption and call it *corruption*. All municipalities present many cases of mild corruption and this category has a low variability.

I obtained the amount of resources audited in each municipality by scraping the audit reports available at CGU's website[7] and linking them to the database above. All values are in Brazilian Reais considering prices of the baseline year (2000).

[7] `http://www.cgu.gov.br/assuntos/auditoria-e-fiscalizacao/pesquisa-de-relatorios`

Migration Data

Using 1990, 2000 and 2010 censuses, provided by the Brazilian Institute of Geography and Statistics (IBGE), I build a database of all the bilateral flows of migrants across Brazilian municipalities. The Census normally applies a specific questionary about migration to a large sample of the population, so it is representative at the municipality level. The 2010 Census asked how long the individual had lived in their current location, and if the answer was less or equal to 10 years, they further asked the municipality they had lived in before moving. Using the two questions I can approximate the bilateral flows of migrants across municipalities for each year since 2000. The other two censuses only asked individuals for their previous place of residence up to 5 years before the interview.

Defining Local Economies

In the construction of the instruments, to be described in detail later, I have to account for the fact that shocks in a municipality may be diffused among all the components of a local economy. Following the definition of local labor markets used in a recent literature in labor economics (see e.g Dix-Carneiro and Kovak (2015), Costa, Garred and Pessoa (2016) and Dix-Carneiro, Soares and Ulyssea (2016)), I use the geographic units "microregions" constructed by IBGE. These are contiguous municipalities integrated in local economies. Brazil comprehends 558 such units, with an average of 13 municipalities per unit. Economic shocks outside a municipality's "microregion" may be considered plausibly exogenous for that municipality.

Apart from this, I am also interested in excluding from the analysis destination municipalities that are too close to sending municipalities. Because this would imply voters have easy access to their municipality of origin and, thus, to information. In addition, these municipalities may be similar in many respects and, therefore,

destination municipalities characteristics would not be exogenous. Excluding bilateral flows of voters within a local economy will also serve this purpose.

Other Data

From the 2000 Census, I also extract a set of demographic and socioeconomic controls at the municipality level. These variables are the baseline characteristics used in all regressions. They include a measure of income inequality (*Gini Index*), the log of average income (*Log Income*), the rate of unemployment (*Unemployed*), the percentage of public sector workers (*Public Sector Workers*), the share of the population living in a rural area (*Rural*), the percentage of the population with a given education level (*University-Degree, High-School, Illiterates*), and the log of total population (*Log Population*).

3.3 Baseline model, sample selection and measurement error in registration rates

3.3.1 Baseline Model

In this paper, I am primarily interested in estimating the effects of registration rates on the behavior of politicians. Let Y_{it} be the outcome of interest, number of corruption events, for municipality i in time t. I use data on mayoral elections from three electoral cycles, 2004, 2008 and 2012, to estimate the following baseline model:

$$Y_{it} = \alpha + \beta R_{it} + X'_{i,t=2000}\gamma + I_t + \Pi_s + \epsilon_{it},\ i = 1, \ldots, n. \tag{3.1}$$

where R_{it} represents registration rates in municipality i in time t and X includes municipalities characteristics extracted from the 2000 Census (a measure of inequality (Gini Index), the log of average household income, percentage of rural population,

percentage of illiterates, percentage of adults with secondary education, the percentage of civil servants and population size in log units). Because of the variation in state governments and its influence on local politics I include a set of state intercepts (Π_s). To capture elections differences, I also include election-year fixed effects (I_t). Errors are clustered at the municipality level.

The next two subsections discuss how accurate registration rates (R_{it}) are in accounting for migrant voters who remain voting in their municipality of origin and how treated municipalities are selected within those for which corruption data is available.

3.3.2 Registration Rates: Registered Voters vs Resident Voters

I define registration rates as the ratio between the number of registered voters in a municipality and its number of resident voters. In this subsection, I discuss how these two variables have evolved over time and how far they are from each other by different age groups of voters. I also discuss possible sources of measurement error in the number of resident voters.

Figures 3.1 and 3.3 display the average gap between the number of voters registered in a municipality and its number of residents in voting age over a period of nine years, from 1996 to 2012.[8] Out of census years, I use population estimates provided by the Brazilian Ministry of Health (Datasus). I split the sample of municipalities at the 200,000 voters threshold used to allocate single and dual-ballot elections across municipalities, as this will also reflect large cities that receive a high number of migrants.

In Figure 3.1 I include voters with age 16 or above. In census years, the most precise estimates of resident voters, small municipalities (below 200k voters, Figure 3.1(a)) have slightly more resident voters than registered voters. For large municipal-

[8]The electoral authority has biannual releases of voters' data.

ities (above 200k voters, Figure 3.1(b)), this gap is even larger and this is true for all years in the sample.

Figure 3.2 reveals where the noise for small municipalities come from. It includes in the sample only 16-17 year olds. As previously discussed, voting is compulsory in Brazil for citizens between 18 and 70 years of age. But for 16 and 17 year olds voting or registering for voting is voluntary. Therefore, young voters not registering with the electoral authority might account for part of the gap displayed in Figure 3.1.

In Figure 3.3 I exclude these young voters for which voting is voluntary. Note that now the gap between the two measures is inverted for small municipalities only, whereas for large municipalities it has been narrowed. Since 16-17 is an entry threshold, there is also some noise stemming from the disconnect between the date of the Census interviews and birth dates. For example, at the time of the interview a voter may not yet be 16 years old, but by the time of the election she will be and, thus, will have registered with the electoral authority.

There is also some measurement error coming from the fact that voters above 70 years of age are not required to vote as well. It adds to that the fact that this group of people are more sensitive to mobility problems and less likely to vote. Therefore, adding this group to the computation of registration rates would over estimate the number of actual voters in a municipality. Consequently, to increase precision, I include only voters between the ages of 18 and 70 (inclusive) in the registration rates used in all the specifications.

As for the measurement error related to the estimates of the number of residents, its size varies to the extent to which migration and deaths are relatively stable over a 9 years period of time. Since these are population estimates of the Ministry of Health, and it has a centralized system that monitors the number of deaths in a municipality, it is expected that their estimates can account well for this source of variability. In both 1996 and 2007, municipalities with population below 150,000 had

their residents counted. So the data for these years are as precise as in Census years for smaller municipalities.

The level of registration rates is determined in part by the stock of migrants in a municipality which was greatly determined by migratory waves of previous decades. However, the longer an individual has lived outside her voting municipality the more likely it is that she will have asked for a transference and/or a re-registration program has taken place. Therefore, the sample of municipalities with high registration rates should capture recent migration only.

3.3.3 Sample Selection

The sample of municipalities audited by the anti-corruption program, even though obtained through randomization, does not ensure a balance in sample size both across groups and over time. The randomization is based on a single sequence of random assignments through the national lottery and samples (with replacement across years) from the set of all Brazilian municipalities with less than 450,000 inhabitants.[9] Therefore, we have to find some criterion to resample these municipalities appropriately in terms of registration rates.

The sample of audited municipalities made available by CGU comprises 1,158 municipalities. Corruption events are classified as mild, moderate or severe. It includes data from draws realized from 2006 to 2014. An average of 128 municipalities were drawn by year, with a minimum of 24 in 2012 and a maximum of 180 in 2007. Tables 3.1 and 3.2 compare the means of characteristics of municipalities – such as per capita income, education and urbanization – at the top and bottom percentiles of the distribution of registration rates. Except for the Gini coefficient of income, different percentiles are significantly different in means across all dimensions considered.

[9]This includes approximately 92% of all Brazilian municipalities.

Figure 3.4 displays the quantile distribution of registration rates. About 20% of the municipalities are on or above one, that is, have more registered voters than resident voters. However, the closest we get to one the more likely it is that the difference from one is just due to noise. Because the threshold for identifying municipalities with high registration rates is not well defined, we need to analyse in more detail how migration may determine registration rates.

Consider that in time $t = 0$, when there is no movement of voters, voters registration (R_0) is exactly equal to the number of resident voters (N_0), and therefore, $R_0/N_0 = 1$. In time $t = 1$, we allow for both in- and out-migration. The registration rate in time $t = 1$, will now be given by

$$\frac{R_1}{N_1} = \frac{R_0 + InMig_1 - OutMig_1}{N_0 + InMig_1 - OutMig_1} \tag{3.2}$$

However, not everyone who migrates requests a transfer of registration, therefore we have

$$\frac{R_1}{N_1} = \frac{R_0 + (InMig_1 - OutMig_1)}{N_0 + (InMig_1 - OutMig_1)} + \frac{(R_{Out_1} - NR_{In_1})}{N_0 + (InMig_1 - OutMig_1)} \tag{3.3}$$

R_{Out_1} stands for the voters who out-migrate but remain registered in their municipality of origin and R_{In_1} stands for the in-migrants that do not request a transfer to the host municipality. Note that if there are deaths and new citizens entering compulsory voting age, registration rates should still remain equal to one.

The registration rate in $t = 1$ given by equation 3.3 is zero whenever $(R_{Out_1} = NR_{In_1}) \geq 0$. However, this condition is very unlikely. Abstracting from any disturbances stemming from measurement error, whenever $(R_{Out_1} - NR_{In_1}) > 0$, registration rates will be above one in $t = 1$, and below one, otherwise. As the time passes, and migratory flows evolve, registration rates could go either above or below one. In fact, if we start with $R_0/N_0 < 1$ and allow for migratory movements with imperfect registration, there could still be a positive variation in $(R_{Out_1} - NR_{In_1})$ and

we would not observe it by just looking at above one registration rates. The same argument is true when we start with $R_0/N_0 > 1$, and there is a negative variation in $(R_{Out_1} - NR_{In_1})$, such that $R_1/N_1 < 1$. Consequently, a priori, we cannot classify municipalities by just observing registration rates in levels.

However, we can observe the variation in registration rates between elections. And since the imperfect registration rates we are interested in are entirely related to migration[10], we can combine variation in registration rates with variation in migration to elicit municipalities that should be in our sample. If $R_1/N_1 - R_0/N_0 = \Delta RegRate_1 > 0$, even in the case where $R_0/N_0 \neq 1$, registration rates are rising faster than the number of resident voters. If politicians observe this, they can react to it both in terms of policy and campaign efforts.

Based on this argument, the inclusion of a municipality in the sample used here is decided according to the following algorithm:

(i) Take municipalities' full history of registration rates from 1996 to 2012 and calculate differences with respect to the previous election. Keep those municipalities who had at least one positive variation in registration rates.

(ii) Calculate the net out-migration for all these municipalities and retain from the previous sample only those who are above the 25th percentile of net out-migration.

Criterion (i) accounts for the fact that every time there is a positive variation in registration rates, under the assumption of small measurement error, the number of registered voters is growing faster than the number of resident voters. Note that this is independent of the level of registration rates of the municipality. To avoid cases where $(R_{Out_1} - NR_{In_1})$ is negative, which is the case where in-migration would cause the municipality to have more residents than voters, and this would affect politicians

[10]Those related to fraud or measurement error would accumulate confounding factors.

differently, in criterion (ii) I keep only those municipalities that have a reasonably high net out-migration (above the 25th percentile), since out-migration is the only mechanism that would produce the variation in registration rates we are interested in – that is, only sufficiently high out-migration will produce registered voters that are non-residents. Because this excludes municipalities that have variation in registration rates not due to out-migration, it is also useful to reduce measurement error and the possibility that these registration rates are capturing fraudulent registration of voters.

Table 3.3 compares characteristics of this sample of municipalities, referred to as "out-migration sample" throughout the text, against the characteristics of the remaining municipalities in the corruption sample, referred to as "general sample". Municipalities in the out-migration sample present, on average, higher income inequality, lower education levels and income per capita, are more rural and less populated. In addition, a higher percentage of voters in these municipalities work as civil servants. By using this sample, I exploit the variation in intensity of treatment, that is, the variation in registration rates only across municipalities for which net out-migration is above the 25th percentile.

3.4 Identification Strategy

This section formalizes the empirical strategy that allows for the identification of causal effects of changes in registration rates, therefore voting motivations, on corruption and performance.

There are obvious endogeneity problems in registration rates, but the most challenging issue to be taken into consideration is that registration rates are correlated with migration. This correlation induces a simultaneous determination of both registration rates and corruption, because migration can directly affect corruption. I am interested in the effects of out-migration on corruption only through its impact

in the variation of registration rates. Therefore, not only an exogenous variation in registration rates is necessary, but also in migration.

A good IV candidate has to be correlated with registration rates but not directly correlated with migration in addition to the usual exclusion restriction that would solve the problems of simultaneity, measurement error and omitted variables. Because changes in registration rates are a direct consequence of migration, it is not possible to find orthogonal exogenous variation in both of them. The approach adopted here will be to find an IV for registration rates and construct measures of migration flows that can be included as controls in our specification.

The simultaneous determination of registration rates and corruption works through many channels. For example, municipalities with more corrupt politicians would have a higher out-migration rate and these voters displaced by bad governments would hardly be willing to keep their voting duties there. In this case, out-migration and variation in registration rates would be negatively correlated, generating a downward bias in the registration rates coefficient. Conversely, these voters could instead be more interested in politics after having verified the consequences of voting in certain types of politicians and are keen to improve their place of origin. In this case, the coefficient on registration rates would be upward biased.

As previously explained in Section 3.2.2, because Census questions can recover only the movement of people who remain in a destination until the date of the interview, those with short migration spells, length of less than 5 years before the interview, are not captured. This could lead to attenuation bias in the inference of migration effects.

As mentioned earlier, there is also some measurement error in the population estimates used to calculate registration rates, but it is a lesser concern because registration rates will only have an effect on corruption to the extent to which politicians know about it. Politicians know precisely how many voters are there in a munici-

pality, as TSE data is publicly available, and may base their decisions on population estimates as well.

As for the omitted variable bias, one possibility is that corrupt municipalities have an electorate that is more ideological and thus have a higher tolerance to corruption (Boukouras and Koufopoulos, 2008). Politicians respond to that extracting more rents. Ideological out-migrants could probably get a higher value from voting in their home municipalities generating a spurious correlation between registration rates and corruption through ideology. Politicians would extract more rents both with or without higher registration rates.

With these considerations in mind, in the following subsections I develop an instrument for registration rates and plausibly exogenous measures of both in- and out-migration.

3.4.1 Developing an instrument for registration rates

When an internal migrant decides where to vote, she might factor in elections' characteristics, traveling and transference costs. To request a registration transfer, the citizen has to wait three months time after moving and pay a visit to the electoral authority with, among other documents, a proof of address in that municipality. Otherwise, the voter can choose to leave her registration as it is and, in the case she cannot travel to her electoral domicile and vote in election day, she can go to the closest polling place and justify her abstention without any incurring penalties. In this sense, it is probably more costly to request a transfer.

Voters' types may also play a role in this decision. An instrumental voter may be more likely to vote in the host municipality, considering that she will not benefit from any policies implemented in her home municipality, at least directly or contemporaneously. If she does derive benefits from it, differences in the size of the elections, and therefore pivotality, may influence this decision instead. It is important to high-

light that local elections are less ideological than national elections. At the local level voters are more concerned about public goods, property values, taxes, etc.

To try and get rid of this selection process, I focus on exogenous conditions in destination municipalities that might influence voters' decision to request a transfer. Since it is not clear what these conditions could be – for example, peer effects from previous migrants or distance to the local electoral court or distance to the source municipality, etc – I use the decision made by internal migrants who arrive in all the destinations of out-migrants of a given source municipality net of its own possible contribution to it. In other words, I extract the variation in registration rates of municipalities that are destinations of out-migrants of the municipality of interest, but discounting the variation in registration rates stemming from the source municipality itself. To make sure destination municipalities are sufficiently different from source municipalities, I consider only destinations that are outside the source municipalities' microregion (as defined in Section 3.2.2). The idea is that the variation in registration rates of destination municipalities that are independent of the variation in registration rates of a given source municipality will aggregate reasons for not requesting a transfer that are peculiar to those destinations and exogenous to the variation in registration rates of any given source municipality.

More formally, let ΔR_d be the variation in registration rates in destination d with respect to the previous election. Consider the set of all destinations d where the source municipality o has sent voters to, $D_o = \{1, \ldots, N_{od}\}$, with $N_{od} > 3$, and define γ_d as being the proportion of the variation in registration rates in municipality d relative to the total variation in registration rates of all municipalities in D_o.

$$\gamma_d = \Delta R_d / \sum_{D_o} \Delta R_d$$

γ_d is then the factor used to remove the variation in registration rates of the municipality of origin o from each destination d, as in the equation below:

$$\widehat{\Delta R_d} = \Delta R_d - \gamma_d \Delta R_o, \quad d \in D_o \tag{3.4}$$

where $\widehat{\Delta R_d}$ is, therefore, the variation in registration rates of destination d net of the variation in registration rates of the source municipality o. Because the set D_o for each source municipality o is large, a sufficient statistic is needed to summarize these data and be used as an IV. The median seems to be a natural candidate, as it is relatively less skewed towards extremely large or small values and, therefore, may give a better idea of a typical destination municipality for each source.

The median value M_o of all the values $\widehat{\Delta R_d}$, with $d \in D_o$, is the IV for ΔR_o.

That is, $M_o = \widehat{\Delta R_d}_{\left(\frac{N_{od}+1}{2}\right)}$ for an odd number of destinations (N_{od}) and $M_o = \left(\widehat{\Delta R_d}_{\left(\frac{N_{od}}{2}\right)} + \widehat{\Delta R_d}_{\left(\frac{N_{od}}{2}+1\right)}\right) 0.5$, otherwise.

The median destination, in terms of variation in registration rates, net of the variation in registration rates in the origin, should be correlated with variation in registration rates of the origin municipality to the extent to which it proxies for factors in the destination that influence internal migrants in general to request a transfer or not. The necessary exclusion restriction is that variation in registration rates in the median destination municipality, net of the variation stemming from the origin municipality itself, affects corruption in the origin municipality only through its influence on migrants' decision of requesting a transfer.

3.4.2 Developing an exogenous measure of in- and out-migration flows

The thought experiment used for creating an instrument for migration resembles the one in the previous section. We need a variable that captures factors that influence

migration decisions but that is not correlated with contemporaneous economic or political shocks in the source/destination municipality.

A number of empirical studies in the immigration literature uses historical settlement of immigrants as an instrument to estimate the effects of migration on economic outcomes (See e.g. Card, 2001; Dustmann, Hatton and Preston, 2005; Altonji and Card, 2007). The idea is that prospective immigrants choose locations on the basis of the strength of immigrant networks, something that is not correlated with contemporaneous economic shocks, as long as historical migration is sufficiently lagged and cross-sectional permanent differences have been differenced out.

Following the same argument, the instrument developed here uses the 1990s distribution of the stock of internal migrants across municipalities to allocate the new waves of in- and out-migrants to a given municipality. For example, if one-third of the total number of internal migrants in 2000 were living in Sao Paulo and two-thirds in Rio de Janeiro, the instrument allocates one-third of the new migrants of a given year to Sao Paulo and two-thirds to Rio de Janeiro.

Formally, consider the set of all municipalities in the country $I = 1, \ldots, 5565$. In- or out-migration in municipality i will be given by:

$$X_{it} = \frac{X_{i,2000}}{\sum\limits_{I} X_{i,2000}} * \sum_{I} X_{-i,t}, \quad t = 2004, 2008, 2012. \tag{3.5}$$

where $X_{i,2000}$ is the stock of in- or out- migrants in municipality i in 2000, X_{it} is the flow of in- or out- migrants in municipality i in year t considering the date of migration reported by the voter in the census of 2010, and $\sum\limits_{I} X_{-i,t}$ is the aggregate of the flows of in- or out-migrants of all Brazilian municipalities but i. The migration flow in year t considers the aggregate of flows of the previous three years. Migration flows in the election year is excluded to avoid including migrant voters that still have fresh information about the municipality as well as politically motivated migration. The

national flow of internal migration is likely to be exogenous within a single state or municipality. But for large municipalities this assumption may not hold. To deal with this concern, municipality i's contemporaneous flow of migrants is netted out from the total contemporaneous flow in the second term of the equation. The necessary assumption for an exclusion restriction to hold here is that unobserved factors that determine the decision of migrants to leave city i in the 1990s are uncorrelated with contemporaneous political and economic shocks in i.

3.5 Results

This section investigates the effects of registration rates on corruption and implements the IV estimations discussed in Section 3.4.

3.5.1 Effects of registration rates on corruption

This section presents estimates of the effects of registration rates on the number of corruption events. Table 3.1 contains OLS regression results from estimating variants to equation 3.1. The specification in the first column estimates the effects of registration rates on corruption, controlling only for elections intercepts. Columns (2) and (3) extend this specification to include intercepts for the years the municipalities were drawn in the audit program and state intercepts,[11] respectively. Column (4) adds municipal characteristics.

The results suggest that registration rates and their associated effect on politicians choice of rents, on average, had a significant positive effect on the number of corruption events found in a municipality. One percentage point increase in registration rates would imply that 40 more cases of corruption would be found, on average. Audit years and state intercepts reduce this effect by about 12% with a loss in precision.

[11] This sample spams 23 out of 26 states in Brazil.

State intercepts added in Column (3), to account for state-level changes potentially driven by state-specific politics, further reduce this effect to 23.6 corruption cases. This indicates that some of the states that faced the highest variations in registration rates also displayed other characteristics that contributed to the growth of corruption, initially biasing the coefficient.

The inclusion of municipal characteristics in Column (4), however, just absorbs some of the variation in the error term, and does not alter significantly the estimated effect or the estimated precision. In Column (5) the same specification is estimated, but the total value of transfers audited in each municipality is added as a control. This specification addresses concerns about the definition of transfers used here and the so called "political resource course". That is, more transfers leave scope for more corruption. The transfers audited in each municipality are ear marked resources and allocated through exogenous population thresholds. However, as Brollo et al. (2013) shows, some municipalities receive more resources then those compatible with their population size. Therefore, the value of resources audited is potentially endogenous. Adding the value of the resources audited as a control only increases the parameter estimate by 2.4% and slightly improves precision. The number of cases of corruption is really increasing on the amount of resources audited, but it does not affect significantly the coefficient of registration rates on corruption. This indicates that the variation in the amount of resources a municipality receives and that is audited does not follow that of registration rates in our sample, at least not to a great extent. Because of these endogeneity concerns the specification used henceforth is that of Column (4).

3.5.2 The causal effect of registration rates on corruption

Column (1) in Table 3.5 presents the results from the reduced-form specification analyzing the effect of changes in registration rates at destination municipalities on the number of corruption cases found in the source municipality. As we would expect,

the larger the variation in destinations' registration rates the lower the variation in registration rates in the source municipality. Similarly, though more precisely estimated, the first-stage in column (2) displays a direct negative correlation between registration rates in destination and source municipalities.

The first-stage is sufficiently strong, with an F-statistic of around 15. The second stage result shows that an increase in registration rates also lead to an increase in the verified number of corruption cases. On average, 42 extra cases of corruption would be verified for each percentage point increase in registration rates. Moving a municipality from the 90th percentile to the 10th percentile of the distribution of registration rates, leads to a variation of 0.39 percentage point. According to the first-stage results, this would lead to a reduction of 0.34 in median destination registration rates. One percentage point increase in source registration rates means that one percent of voters leave the municipality and none of them request a vote transfer in the destination. To put this number into context, in the sample used to estimate this model, a municipality in the 90th percentile of corruption, had 83 more cases of corruption detected.

3.5.3 Placebo Exercise

To address concerns about the exclusion restriction assumption that the instrument affects corruption only through conditions in destination municipalities that prevent voters from requesting a vote transfer, I run the following placebo exercise.

In Section 3.4, a sample containing only municipalities with out-migration rates above the 25th percentile was chosen to guarantee that changes in registration rates were more likely to stem from migrant voters keeping their home electoral domicile after moving, rather than fraud or measurement error. The remaining sample contains municipalities for which changes in registration rates are supposedly unrelated to migratory movements.

If the change in registration rates for the median destination municipalities proxies for any factor other than those that affect registration rates in the source municipality, when we use this sample to regress registration rates, for the median destination municipalities, on the number of corruption cases, employing the same specification used in the previous section, the coefficient should come out statistically significant. Table 3.6 depicts the results of this exercise. The relationship between these two variables is small, positive and with large standard errors. An indication that registration rates for the median destination municipalities do not capture any other aspect related to the incidence of corruption in an audited municipality. Therefore, omitted variables do not seem to challenge the credibility of the identification strategy used here.

3.5.4 Internal out-migration and corruption

While sections 3.5.1 and 3.5.2 established a connection between positive changes in registration rates and corruption, this section investigates the mechanisms through which this could happen.

As discussed earlier, changes in registration rates could just be capturing the effects of in- and out-migration. Internal migration affects sending and receiving locations through differing mechanisms: by altering the distribution of characteristics of the population such as age, schooling and marital status, the so called compositional effects, and by affecting the capital to labor ratio in the economy, yielding further implications to wages, employment and investment.

The magnitude of the compositional effects depend on the size of the internal migrant population and the degree of substitution and complementarity between movers and residents. If there is positive self-selection in migration and it is sufficiently high, on the assumption of low substitutability, the source area may suffer with a shortage of qualified labor force, for example. Internal migration is therefore a potential source of demographic change and economic growth.

Empirical evidence regards characteristics such as average income (Hicks, 1963; Sjaastad, 1962) and/or regional differences in the return to skills in destination areas (Borjas, Freeman and Katz, 1997; Roy, 1951) as the main attractors for migrants. Therefore, working age people with relatively high levels of schooling and income are normally more likely to migrate to larger metropolitan centres. These variables in turn have been shown to be correlated with voters' participation, abstention and ideological positions.

There is also a literature that demonstrates the population composition by itself actually affects the provision of public goods. The so called intergenerational competition in the allocation of public sector resources. For example, Poterba (1996) shows American states that experience an increase in the elderly population have a correspondent reduction in education spending per child. In Alesina, Baqir and Easterly (1999), the provision of productive public goods by white majorities respond to the variation in the size of minority groups.

To assess the extent to which the effects found through high registration rates are due to labor market changes or population recomposition, I use the flow of in- and out-migrants in the previous three years leading up to an election, as described in Section 3.2.

For a double IV strategy to work here we would need at least one instrumental variable for either registration rates or migration that we could plausibly claim is orthogonal to the other. Otherwise, adding the two IV's would just build up the bias. If we are interested in the effects of registration rates alone, net of the migration effects, the concern about adding migration in the right hand side of the equation is twofold: (i) contemporaneous migration can potentially be related to ex-ante corruption and (ii) it is also indirectly related to corruption via negative economic shocks that are turned into push factors and, perhaps, facilitate corruption. To address this issue

and be able to add migration as a control in the previous specification I construct plausibly exogenous measures of migration as described in Section 3.2.

Table 3.7 displays the estimates of our preferred specification from the previous section including endogenous and plausibly exogenous measures of migration flows. The first four columns present coefficients from an OLS specification, while the last two columns contain the corresponding IV results. In Column (1) endogenous measures of in- and out-migration are added as controls. For parsimony, in Column (2), I include only net-migration, which restricts arrivals and departures to exert equal and opposite effects on corruption. The estimated coefficients for registration rates remain unchanged in both cases. However, the coefficients for in- and out-migration are very close to zero with large standard errors, indicating that attenuation bias may be a concern. The same is true when net out-migration is added. To account for this, in Columns (3) and (4) I add plausibly exogenous migration measures. Besides being more exogenous, they are also more precise to the extent to which a municipality's stock of migrants over a decade is closer to actual migration than yearly flows. Indeed, coefficients on in-, out- and net out-migration become larger (in absolute value) but still with wide confidence intervals that contain zero.

The coefficient in registration rates, however, does not change significantly. While the relationship between in-migration and corruption is positive and large, the coefficient does not meet the conventional standards for statistical significance. Conversely, out-migration has instead a negative and small relationship with corruption, with wide confidence intervals as well. In the last two columns, the two-stages least squares coefficient for registration rates declines by less than 2% both in the more flexible specification with in- and out-migration and the more restrictive with net out-migration.

Considering the results in Column (5), a variation of one percentage point in registration rates, following one percentage point increase in both in- and out-migration

would amount to 57 more cases of corruption found, on average. This would be a stylized case where half of those who out-migrate remain registered in the municipality of origin and half of those who arrive register for voting in the destination municipality.

Taken together these results indicate that changes in registration rates have a significantly large and autonomous effect on corruption when compared with migration. The next section tries to understand why voters would choose to remain voting in their municipality of origin.

3.6 Mechanisms

3.6.1 Why out-migrants would remain voting in their place of origin? A Downsian framework

Consider the traditional Downsian framework where citizens enjoy both instrumental and expressive benefits from voting. The instrumental utility consists of the extra benefit B_i citizen i obtains if her preferred candidate wins. It can be seen as the additional level of public goods the preferred candidate would provide because she is more competent, for example. In contrast, the expressive utility of voting, D_i, is derived from the simple act of voting and is independent of the election outcome. It is the satisfaction of voting itself; some authors have justified it in terms of civic duty.

If citizen i has a perceived likelihood P_i of being pivotal, P_iB_i is the expected level of extra public goods the preferred candidate would be able to provide. But voting is costly. The citizen incurs a cost C_i when she votes regardless of who wins the election. The citizen's rational decision is then voting when:

$$P_iB_i + D_i \geq C_i \tag{3.6}$$

and abstaining otherwise. Therefore, depending on the levels of B_i and D_i voters may turn up to vote or not. Here, I use this framework to help understand migrants' decision of which election to vote. The goal is to identify how the choice of elections may be related to the voters' types.

Assume B_i, D_i and C_i are normally distributed, and C_i has mean zero and variance one.[12] With probability λ a voter believes she will be pivotal. Therefore, $P_i B_i + D_i = V_i$ is also normally distributed with mean $\mu = E(V_i) = \lambda E(B_i) + E(D_i)$.

The framework I use henceforth is similar to the standard Borjas/Roy model of selection in migration. Consider a voter who has moved from source municipality 0 to host municipality 1. The utility of voting in the source municipality is given by

$$v_0 = \mu_0 + \varepsilon_0 - c \tag{3.7}$$

where $\varepsilon_0 \sim N(0, \sigma_0^2)$ and c is an extra cost that can be understood as costs of travel and/or acquiring information about politics in the source municipality.

And in the host municipality,

$$v_1 = \mu_1 + \varepsilon_1 \tag{3.8}$$

where $\varepsilon_1 \sim N(0, \sigma_1^2)$.

A citizen will choose to vote in the host municipality whenever

$$[\mu_1 - (\mu_0 - c)] + (\varepsilon_1 - \varepsilon_0) > 0 \tag{3.9}$$

[12] If we define $B_i = B_i^a - B_i^b$, where a and b are candidates, the voter switches to candidate b every time $B_i \leq 0$. To focus on voters' types and their relationship with the choice of election, I assume $E(C_i) = 0$. That is, citizens will always vote.

Define $\varepsilon = \varepsilon_1 - \varepsilon_0$. The probability that a randomly chosen voter who migrate choose to vote in the host municipality is equal to

$$\begin{aligned} p &= Pr[\varepsilon > (\mu_0 - \mu_1 - c)] && (3.10) \\ &= Pr[\frac{\varepsilon}{\sigma_\varepsilon} > \frac{(\mu_0 - \mu_1 - c)}{\sigma_\varepsilon}] && (3.11) \\ &= 1 - \Phi(z) && (3.12) \end{aligned}$$

where $z = \frac{(\mu_0 - \mu_1 - c)}{\sigma_\varepsilon}$ and $\Phi(.)$ the CDF of a standard normal. Note that z is increasing on the mean benefits obtained in the source municipality and decreasing in the cost of voting in the source municipality and mean benefits of voting in the host municipality; the higher the z the lower the probability of voting in the host municipality. Because we want to focus on selection, we assume average benefits from voting are the same in both locations, that is, $\mu_1 \approx \mu_0$.

The expected utility of voting in the source municipality of a voter who chooses to vote in the host municipality is given by:

$$E(v_0|v_1 > v_0) = \mu_0 + E(\varepsilon_0|\frac{\varepsilon}{\sigma_\varepsilon} > z) \tag{3.13}$$

Similarly, the expected utility of voting in the destination municipality of a voter who chooses to vote in the host municipality is given by:

$$E(v_1|v_1 > v_0) = \mu_0 + E(\varepsilon_1|\frac{\varepsilon}{\sigma_\varepsilon} > z) \tag{3.14}$$

Given the normality of the error term, it can be shown that:

$$E(v_0|v_1 > v_0) = \mu_0 + \frac{\sigma_0\sigma_1}{\sigma_\varepsilon}\Big(\rho - \frac{\sigma_0}{\sigma_1}\Big)\Big(\frac{\phi(z)}{1 - \Phi(z)}\Big) \tag{3.15}$$

$$E(v_1|v_1 > v_0) = \mu_1 + \frac{\sigma_0\sigma_1}{\sigma_\varepsilon}\Big(\frac{\sigma_1}{\sigma_0} - \rho\Big)\Big(\frac{\phi(z)}{1-\Phi(z)}\Big) \tag{3.16}$$

where $\rho = \sigma_{01}/\sigma_0\sigma_1$.

Expressive voting, if perceived as an act of civic duty, may well be very similar across source and destination municipalities. While migrant voters do not benefit directly from policies implemented in the source municipality, they may face a very different perceived probability of being pivotal, depending on the relative number of voters between the two municipalities. If the benefits a voter get from policies implemented in the source municipality are non-zero, pivotality considerations could be a factor weighting up the decision of where to vote. If we assume the average expressive motivation for voting is the same across municipalities, the error terms in the utility equations would be entirely driven by the interaction of pivotality considerations and benefits from public policies.

Having said that, we can derive two relevant different types of voter selection from equations 3.15 and 3.16.

(i) *Positive sorting*: $\sigma_1/\sigma_0 > 1$ and $\rho > \sigma_0/\sigma_1$.

The dispersion of benefits ($\sigma_1/\sigma_0 > 1$) is higher in the host municipality, which means that those who choose to vote there are likely to have a high instrumental motive to vote that is driven by the benefit derived from public policies. Assuming $\lambda_1 < \lambda_0$, this is a case where the benefits from policy compensate the loss in perceived pivotality. The second condition ($\rho > \sigma_0/\sigma_1$) means that the correlation between motivations to vote between source and host municipality are sufficiently high such that voters do not change their types across cities. Voters who get higher instrumental benefits choose to vote in the host municipality. This is the case where there is positive selection of voters into the host municipality.

(ii) *Negative sorting*: $\sigma_0/\sigma_1 > 1$ and $\rho > \sigma_1/\sigma_0$.

The dispersion of benefits ($\sigma_0/\sigma_1 > 1$) is higher in the source municipality, which means that those who choose to vote there are likely to have a high instrumental motive to vote. Assuming $\lambda_1 < \lambda_0$, this is a case where the benefits from perceived pivotality more than compensate the loss in benefits from public goods and the extra costs of voting. The second condition ($\rho > \sigma_0/\sigma_1$) means that the correlation between voting motivations between source and host municipality are sufficiently high such that voters do not change their types across cities. Voters who get higher instrumental benefits choose to vote in the source municipality. In this case, there is a negative selection of voters into the source municipality.

The assumption $\lambda_1 < \lambda_0$ comes from the fact that voters out-migrate more often from small cities to large metropolitan areas. Therefore, the size of mayoral elections between source and destination municipalities may change considerably. The benefit of being more likely to be decisive in an election is only accounted for when there exists at least a small positive benefit stemming from policies. Non-zero benefits from policy in source municipalities could be generated by a natural attachment to home places and/or by relatives that remain there. I test these propositions in the next section.

The effect of election size on migrant voters' choices

In this section, I test the prediction that migrant voters should choose to remain voting in their municipality of origin more often when its election is relatively smaller when compared to that of the host municipality.

To do this, I exploit the discontinuities in the definition of the maximum number of council seats municipalities in Brazil are allowed to have. The Brazilian Constitution uses 24 population brackets to determine the number of seats in each municipality. They start at nine and discontinuously increase by two seats in each subsequent

bracket. These thresholds are displayed in Table 3.8. For example, municipalities with a population below 15,000 are allowed to have a maximum of nine councillors. The largest municipalities, those with population above 8 million, are allowed to have 55 councillors.

The councils themselves decide the number of seats, in theory, based on the budget available to cover wages. Obviously, ruling parties may have different incentives to articulate in favor of more seats or not. It could be either easier or more difficult to get support for their projects. Therefore, these discontinuities allow for a fuzzy design.

The allocation of seats across parties follow the proportional rule: the total number of valid votes in the previous election is divided by the number of seats available to get the number of votes a party need to have gathered to fill a seat. This is the so called electoral quotient. For example, in a municipality that had 6,000 valid votes in the previous election and 9 seats available, the electoral quotient would be approximately 667 votes. If the total number of votes a party receives is equal to 1,334 it gets two seats. The allocation of seats within the party goes to the candidates who obtained most of the votes.

If the width of the thresholds grows faster than the additional number of seats, assuming that the number of voters is proportional to the population, the number of votes a party needs in order to get a seat actually increases in each threshold. This is what happens with the constitutional thresholds, because the number of seats increase in steps of two, but the widths increase following different rules. Therefore, on average, a candidate will need more votes to be elected in a larger municipality.

The idea is that the more votes a candidate needs to be elected, the less likely it is that a voter is decisive (pivotal) in electing that candidate. If a voter moves from a small to a large municipality, which is normally the case, chances are that her actual (and perceived) likelihood of being pivotal in the host municipality is smaller

(Hoffman, Morgan and Raymond, 2013). Moreover, smaller elections may provide a better sense of community, where the social enforcement for voting is greater (Funk, 2010), even from a distance.

I use these discontinuities in the number of seats to compare registration rates of municipalities that are close enough to the thresholds. The constitutional maximum number of seats is used as an instrument for the actual number of seats in a municipality. If out-migrants systematically choose to vote in municipalities with smaller elections, we should observe that municipalities with a higher number of seats experience less variation in registration rates.

Formally, I estimate the following specification:

$$R_i = f(population_i) + \beta C_i + I_t + \Pi_s + \epsilon_{it},\ i = 1, \ldots, n. \tag{3.17}$$

where the maximum number of councillors is used as an instrument for the actual number of councillors, f(.) is a high-order polynomial of the municipality population, I_t time fixed effects, and Π_s state fixed effects. Error terms ϵ_{it} are clustered at the municipality level.

To implement the model, the 24 thresholds are pooled by merging the thresholds together and normalizing population size as the distance from the closest threshold. To keep a municipality in a unique threshold, each interval starts and ends at the midpoints below and above each threshold.

In this framework, for the regression discontinuity design to be valid, we need to assume that conditional regression functions of potential outcomes at the thresholds are continuous, that is, there is no manipulative sorting at the thresholds and there are no other policies implemented over those cutoffs. To my knowledge, there are no other policies that use the same cutoffs. And since the thresholds are determined by the constitution, the only way municipalities could have a higher number of seats

would be by manipulating population estimates. These estimates are provided to the electoral authority by the Brazilian Institute of Geography and Statistics (Instituto Brasileiro de Geografia e Estatistica, IBGE) and the thresholds were established in the constitution of 1989. Unless local politicians were able to bribe IBGE to manipulate population estimates, manipulation would be possible only in Census years by, for example, bringing people in to inflate population counts. This would be, however, very impractical, since IBGE's interviewers visit households.

The continuity assumption also guarantees that the maximum number of seats affects registration rates only through the number of seats the municipality actually has. This is the exclusion restriction of the IV set up. Additionally to this assumption, we also need that the number of seats in a municipality below a threshold is not larger than it would have been had the municipality been allocated to above the threshold. This is the monotonicity assumption. Figure 3.5 displays the jumps in the observed number of seats compared to the constitutional maximum number of seats in that population threshold (indicated by vertical red lines), which suggests that the monotonicity condition is valid.

Table 3.17 estimates the first-stage and the reduced-form regressions. The table reports the estimated coefficients of maximum number of councillors, in a regression where the dependent variable corresponds to each column heading. The column "Full-Sample" is obtained by estimating a single regression on the entire sample. The column "Thresholds = 1000" does the same but considers only municipalities within 1000 inhabitants from a threshold in the first three thresholds. The functional form, $f(population_i)$, is defined either as a forth- (columns (3), (4), (7) and (8)), third- (columns (2) and (6)) or second-order (columns (1) and (5)) polynomial. Columns (4) and (8) also add municipalities characteristics. In all of these cases, the results are very similar. I interpret the results of the most conservative specification.

Columns (1) and (3) report the estimated reduced-form coefficient for the full-sample and the sample within 1000 inhabitants from the threshold, respectively. This coefficient represents the effect of the maximum number of seats on registration rates. Both of them are negative and significant. Columns (2) and (4) report the first-stage estimates for the same two samples, namely the effect of the constitutional maximum number of seats and the actual number of seats in a municipality. The coefficient is positive and highly significant, that is, the higher the threshold of seats the higher the actual number of seats. But the coefficient is smaller than one. This is probably a reflection of the councils' choice of number of seats in accordance with specific interests. Measurement error originating in the population that we use to classify municipalities could also lead to a downward bias. By the estimated coefficients in column (3), when the constitutional maximum number of seats in a municipality increases by two, there is a 0.028 percentage point reduction in registration rates.

Table 3.10 estimates the baseline IV regressions where the maximum allowed number of seats are used as instruments for the actual number of seats observed in a municipality, considering many specifications and the two samples. In the most conservative specification, column (8), an increase of two units in the actual number of seats translates into a 0.12 percentage point reduction in registration rates. In other words, among municipalities where the number of registered voters is greater than the number of resident voters, this ratio is lower for those that face larger elections. This indicates that voters may indeed take into account the size of the election when making a decision of whether to vote in their home municipality.

It is easy to imagine how the number of councillors would affect corruption itself but not registration rates. For example, a larger number of councillors would imply that a mayor would have to negotiate more votes to get her projects approved, depending on the distribution of seats across parties. However, it is not clear how number of seats would directly affect registration rates. One could think that the

higher the number of seats, the more votes a council candidate would need to get elected and municipalities of this type would be more prone to fraud in registration rates. However, if that was the case, we would observe instead a positive effect of the number of councillors on registration rates. Even if there is such effect, our results show that the dominant effect is that at least some voters decide to remain voting in their city of origin based on the size of the election.

3.6.2 Who migrates?

This section investigates the link between the type of selection in migration and its implications to the characteristics of those voters who keep their electoral domicile at their place of origin.

Table 3.11 compares the composition of groups of migrants and non-migrants in terms of education levels, gender, age, the time spent in the host municipality and wages. Column (1) displays the averages for migrants, Column (2) for non-migrants and Column (3) tests the significance of the difference in means between the two groups. The group of non-migrants presents systematically higher means in all education levels except the group of those without formal education. Migrants are younger, include more females and have higher wages. The average time they declared to have lived in the host municipality was 3.7 years.

These statistics could be misleading if those who migrate do so seeking higher levels of education. To address this issue, Table 3.12 considers the same groups and characteristics, but the group of migrants include only those with migration spells of less than three years. The idea is that a degree could not have been completed over this time. For this group, the average spell of migration is 1.4 years. And the same patterns in education levels, age, gender and wages are observed.

In terms of voter behavior, positive selection in migration means that less educated voters remain in the city. If the migrant voters, those who keep their registration in

the origin, are also positively selected, in any case, they would be less informed than before and probably not interested in public goods. Therefore, the municipality looses voters of higher quality both in the out-migration and vote transfer steps.

Even in the case where returning voters can imitate the behavior of relatives or friends living in the source municipality, if we consider that the remaining voters are on average less educated than migrant voters, an imitating behavior would gravitate towards the average behavior of native voters, who are, in theory, worse voters. In addition, even without an imitating behavior, lack of information would imply that it is easier for locals to persuade returning voters in terms of a bad candidate.

To summarise, there is positive selection in those who migrate and, through many mechanisms, the vote of a migrant voter is either not informed or close to that of residents who have lower levels of education and lower wages. And this seems to have implications to the quality of politicians elected and, therefore, the incidence of corruption.

3.6.3 Why do politicians react to it by extracting more rents?

Candidates' quality

This section analyses whether the increase in corruption caused by positive changes in registration rates is due to the entry of candidates of lower quality. I first measure quality in terms of education and occupational skills and then in terms of the performance in attracting discretionary transfers. Table 3.13 depicts the results of linear probability models where highly educated candidates are defined as those with at least secondary education and as high-skilled those with an occupation in a category classified by the International Labour Organization (ILO) as requiring higher skills to be executed.

Column (5) shows that the higher the change in registration rates the lower the probability that the municipality will have candidates ranked in the top three that belong to the group of high-skilled occupations. When education is considered, the 2SLS results are not statistically significant but also display a negative coefficient. The link between skill and corruption may be related to the influence registration rates exert in the entry of candidates.

Columns (3) and (6) investigate whether a municipality is more likely to have an incumbent candidate, given that the candidate is not term constrained, when it displays higher variations in registration rates. Only 279 municipalities in our sample meet this condition; therefore, we may not have enough power to identify an incumbency effect when it is sufficiently small. The coefficient obtained is positive and sizeable, though not statistically significant, indicating that mayors in a first term are more likely to run for reelection in municipalities with higher changes in registration rates. If anything, it is possible that the positive effects of positive changes in registration rates on corruption are driven by migrant voters allowing corrupt mayors to remain competitive when running for reelection.

While the audits investigate the application of earmarked resources, discretionary transfers – transfers from federal and state governments that originate from specific agreements with the local government – are not investigated and their attraction depend entirely on the mayors capacity of setting up these agreements. Table 3.14 contrasts the effects of variation in registration rates on discretionary and non-discretionary transfers (earmarked resources allocated according to population thresholds). Columns (1) and (2) results indicate that variations in registration rates are associated with a reduction in resources but it is only statistically significant for discretionary transfers. The instrumented versions of these models in Columns (3) and (4), respectively, confirm the direction of these effects. A change of one percentage point in registration rates is associated with a reduction of approximately

R$2,823 in discretionary resources, whereas for non-discretionary transfers the coefficient remains negative but still not statistically significant.

Combined with the result that candidates running for election are more likely to be of high skill in municipalities with lower changes in registration rates, this result about performance might indicate that these municipalities have a better pool of candidates and those elected perform better simply because they are of higher quality. Or else voters' motivations could foster accountability both when parties field candidates and politicians manage public resources. Another possible interpretation is that career concerned mayors could try to attract non-auditable resources in order to safely extract rents from them instead. But this channel is less plausible when we compare the total value of resources of the two types. Discretionary transfers are normally of much lower magnitude and would not substitute for rents that can be extracted from non-discretionary transfers; they could be used only as complements when the politician is seeking to extract large rents.

All in all, independently of the specific mechanisms at work, these results show that voter motivation do affect the types of politicians fielded and elected.

3.7 Robustness Checks

3.7.1 Fraud in the registration of voters

While there is evidence that some migrant voters prefer voting in their municipality of origin, the fraudulent registration of voters could also be a source of anomalous variation in the voters to residents ratio. To investigate this possibility, I use data of the recent re-registration program run by the electoral authority undertaken to collect biometric data from the electorate.

Because of its large cost, the program has slowly been rolled out initially to municipalities that have been reported as having some type of irregularity. It started in 2008 with a pilot applied in three municipalities, and expanded in 2010 to include 60 more municipalities. In 2011-2012, 299 more municipalities were added. All together, these municipalities span 24 out of the 26 Brazilian states.

When a re-registration is held in a municipality, both irregular voters and migrant voters may be excluded from the electorate register if they fail to travel back to the municipality and provide the documents needed for it. Therefore, such program will reduce the number of registered voters both by detecting irregularities and by excluding some of the non-resident voters.

First, I look at the list of municipalities where voters have been re-registered and check whether any of them is also in the sample used here. If they are, we require that the new number of registered voters still be superior to the number of residents for all the municipalities that had registration rates above one before the re-registration program. However, only one municipality is common to both samples. When this municipality is excluded from the sample, the results found in previous sections do not change. It is therefore reassuring that the majority of municipalities in the sample were not at the top of the list of those suspected to have fraud in registration rates.

It is, however, informative to analyze how registration rates behaved before and after the program was held. Transition data for municipalities that re-registered their voters during the period considered here are displayed in Tables 3.15 and 3.16. We are interested in evaluating what happened to municipalities who had a registration rate above one before the re-registration. I analyse 2012 and 2010 data separately and use 2008 as the baseline year. In 2012, about 60 percent of the municipalities who were above one in 2008 remained above one after re-registration despite some reduction in registration rates. 88 percent that were below one remained unchanged and 12 percent made the transition to registration rates above one. For those municipalities

who had a positive variation in registration rates in 2012, 32 percent of those below one in 2008 transitioned to above one and 80 percent of those who were above one continued in the same group. These data provide suggestive evidence that even though there might exist some irregularities in the registration of voters, it is not the only determinant of registration rates above one. It is worth mentioning that with biometric data it becomes harder to register ghost voters, for example. So there is no evidence that fraudulent registration has gone up in the year that follows the re-registration program. Additionally, it is hard to believe that the most fraudulent municipalities would have any bargaining power with the federal electoral authority in delaying the program, specially because 'disadvantaged' parties competing in those municipalities would be interested in the re-registration itself.

3.7.2 International migration

A recent literature in economics has linked return migration to transfer of political norms (Batista and Vicente, 2011; Chauvet and Mercier, 2014) and openness to emigration to institutional development(Docquier et al., 2015). If emigration or return migration are large enough to affect registration rates, it could also affect corruption in indirect ways. To address this concern, I look at return migration rates expressed in the Census 2010. This Census asked people whether they were living in another country on July 2005; I use this question to quantify return migration.

Table 3.17 displays descriptive statistics of return migration rates reported in the 2010 Census across municipalities. It also displays a list of municipalities with the highest rates. A total of 0.1 percent of the population in the country and 0.2 percent, on average, across municipalities reported having lived in another country. The highest return rates can be found in southern municipalities that had European colonisation and/or are located in border areas. None of these municipalities are in the sample used here because of the selection criterion of positive net out-migration.

Brazil has historically been an immigration country, with large groups of Italians, Germans and Portugueses being attracted to its southern region to work in the cultivation of coffee during the nineteenth century. Migration from Brazil is a relatively recent phenomenon and basically constituted by first or second generation of European descendants who manage to get an European citizenship based on genealogy. Most of the remaining population faces visa constraints in both Europe and United States. Therefore, return migration levels have been very low throughout the years.

Table 3.18 displays descriptive statistics of immigration rates for the ten years period covered by the 2010 Census across municipalities. It also displays a list of municipalities who had a rate of at least two percent over the period. Panel A displays the data for the period 2001-2005. The average municipality has an emigration rate of 0.01 percent, with a minimum of zero and a maximum of 6.3 percent. The list of municipalities with the highest rates include predominantly border areas. For example, Chui, the municipality with the highest rate, is located in the South of the country and share border with Uruguay. The second municipality in the list, Itaipulandia, also in the South, shares border with Paraguay and had European colonization. In the period of 2006-2010, the average emigration rate goes up to 0.02 percent, again driven by Chui. The second placed municipality, Missal, is also in the South, close to the border with Uruguay and colonized by the Germans. None of these municipalities are included in the sample used in our main results. Thus, neither migration nor return migration seem to be large enough in this period to affect the results obtained here.

3.8 Conclusion

Voters weight differently the various benefits they get from voting. If politicians/parties can identify voters' motivations that are not attached to their perfor-

mance in office they can take advantage of it in order to extract more rents without being punished.

In this paper, I have taken advantage of a quasi-naturally generated group of voters in Brazil that has the option of choosing in which election they would like to vote. The Brazilian electoral legislation allows internal out-migrants to remain voting in their municipality of origin at least until the municipality runs a re-registration program of the electorate. I consider the ratio between the number of registered voters in a municipality and its number of residents as an indicator of how many out-migrants have kept their registration unchanged and are expected to vote in the following election. This allows me to infer about considerations voters make in their choice of election and, more importantly, how politicians after learning about voters' types respond in terms of performance.

I find that voters make considerations about the size of the election, that is, the probability of their vote being decisive in the election. As in Coate, Conlin and Moro (2008), with a relative small number of eligible voters, the equilibrium probability of being pivotal is large enough to motivate individuals with positive costs of voting and offset possible reductions in benefits derived from public goods. When out-migrants choose to vote in the source municipality, they reveal their types to politicians, via registration rates. Candidates' use of this information has then negative externalities to the resident population. The main result of the paper is that there is a link between corruption and voter motivation. On average, for each percentage point increase in registration rates, 42 extra cases of corruption are verified.

I cannot rule out the fact that expressive voting – the consumption value of voting, stemming from civic duty, ethics, or social pressure – also influence corruption in the same context. The concepts embedded in these mechanisms are difficult to measure. What this paper highlights is that candidates' knowledge of voters' types may have negative externalities and that election size itself plays a large role in voters' decisions.

I use this institutional set-up just as a revelation mechanism of voters' types to politicians.

Small elections may have negative externalities for welfare when there is a group of voters whose motivations detached from public goods are known to politicians. The political economy literature has also found that when choosing a candidate voters condition their decisions on the event of being pivotal, and because of it may even vote for less preferred candidates in some cases. These are probably the same voters who make strategic considerations when deciding across elections.

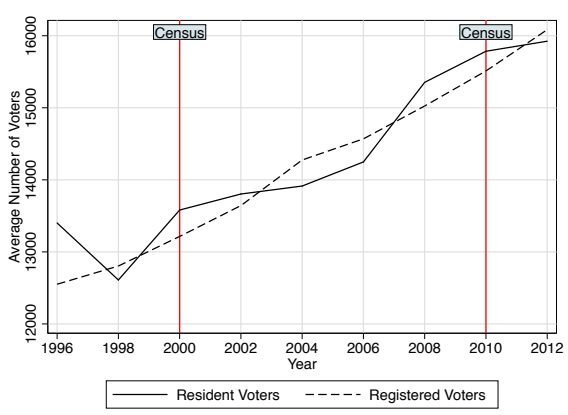
Even though we are in an era where social networks are used to identify voters' types and target advertisement, as mentioned earlier, voters' motivations are difficult to measure and disentangle even at this level of disaggregation. For example, voters may reveal on social networks their policy preferences, but not how they weight these policies in the final vote decision. If they are strategic, knowing about their policy positions is actually uninformative about whom they are voting for. The experiment exploited here provides a quasi-separating mechanism to identify voters' types.

Voters' do not normally choose across elections, but this setting is representative of an experiment where voters would only turnout to vote in an specific election if they believed their vote could make a difference in the election outcome.

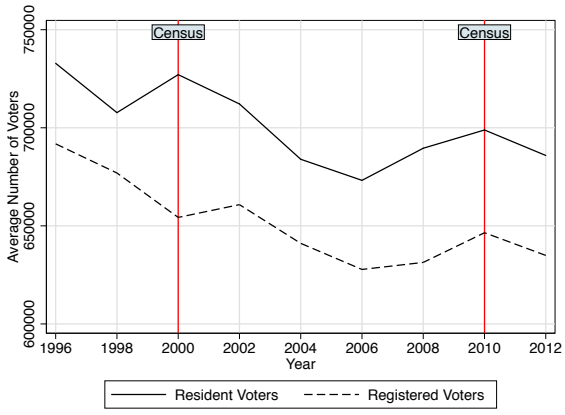
In addition to speaking to theories of political behavior, these results may also have important policy implications. Programs that incentivize turnout, depending on the heterogeneity of their effects across voters of different types, as long as they are known to politicians, may either fortify or abate gains in the local provision of public goods.

3.9 Appendix

3.9.1 Figures

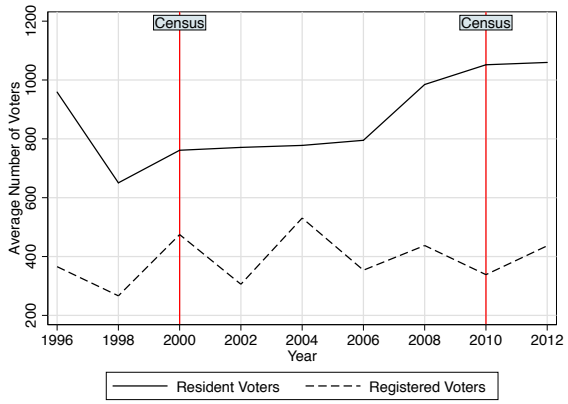


(a) Municipalities with less than 200,000 voters.

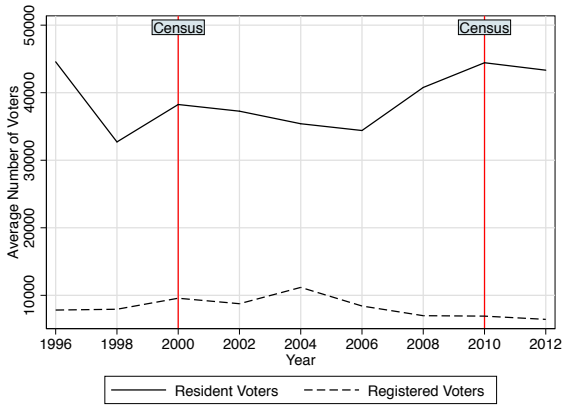


(b) Municipalities with more than 200,000 voters.

Figure 3.1: Average number of resident and registered voters over 16 years of age across municipalities.

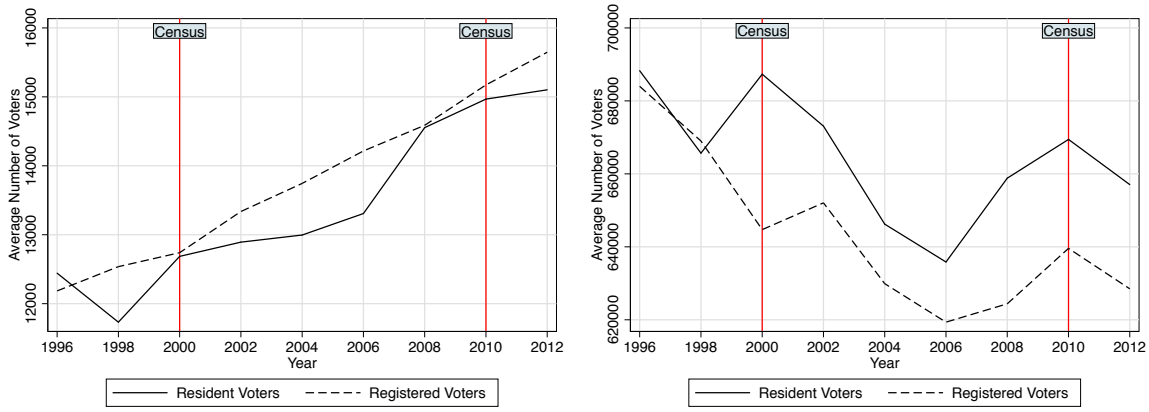


(a) Municipalities with less than 200,000 voters.



(b) Municipalities with more than 200,000 voters.

Figure 3.2: Average number of resident and registered voters with 16-17 years of age across municipalities.



(a) Municipalities with less than 200,000 voters.

(b) Municipalities with more than 200,000 voters.

Figure 3.3: Average number of resident and registered voters over 18 years of age across municipalities.

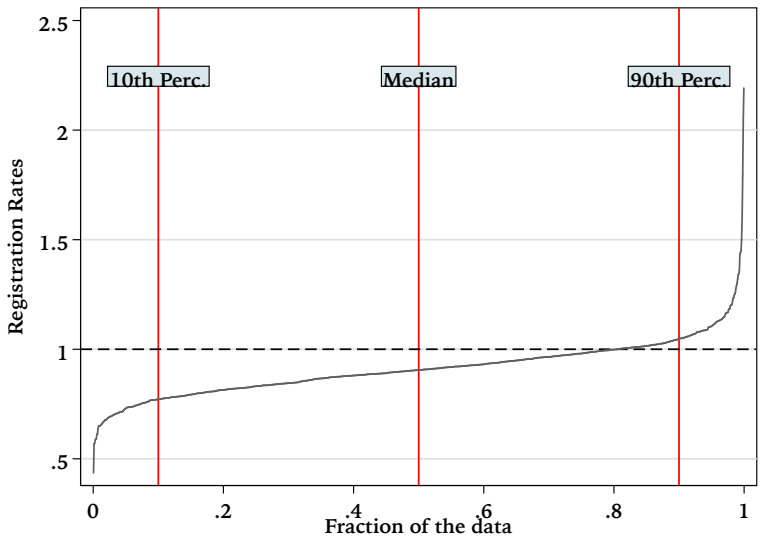


Figure 3.4: Quantile distribution of registration rates in audited municipalities.

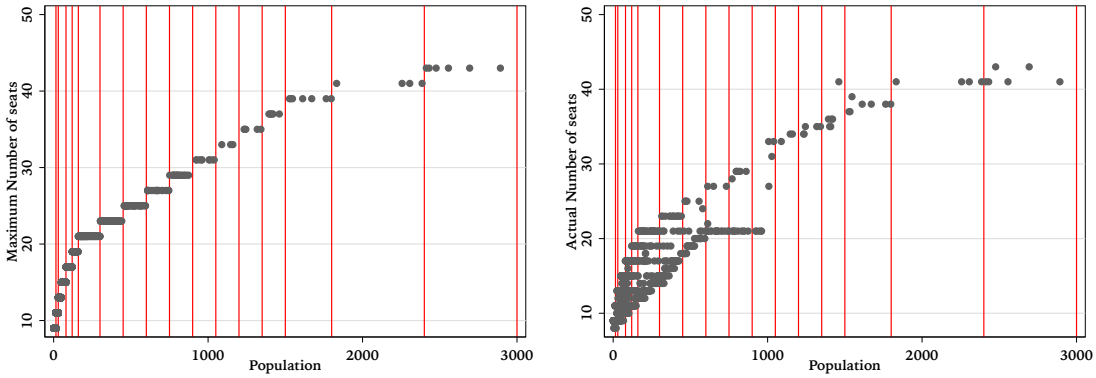


Figure 3.5: Constitutional maximum number of seats compared with actual number of seats by population thresholds (vertical lines).

3.9.2 Tables

Table 3.1: Characteristics of municipalities in the 10th and 90th percentiles of registration rates.

	10th Perc.	90th Perc.	Diff.
Value transfers audited (log)	14.61	15.72	−1.109***
	(0.883)	(1.177)	(0.141)
Gini coefficient	0.557	0.567	−0.00989
	(0.0827)	(0.0721)	(0.0105)
% adults with secondary school	8.794	14.68	−5.890***
	(4.588)	(8.460)	(0.923)
% illiterates	32.17	21.75	10.41***
	(12.44)	(13.01)	(1.721)
% civil servants	8.527	5.684	2.844***
	(5.910)	(3.372)	(0.648)
Per capita income (log)	5.189	5.748	−0.559***
	(0.470)	(0.595)	(0.0725)
% Rural	0.536	0.334	0.202***
	(0.212)	(0.260)	(0.0321)
Population (log)	8.536	10.30	−1.766***
	(0.708)	(1.113)	(0.126)
Observations	108	111	219

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis.

Note: This table compares means of the characteristics of municipalities that have been audited between 2006 and 2014. Columns 1 and 2 display means for municipalities in the 10th and 90th percentiles of registration rates, respectively. Column 3 tests for differences in means between the two groups. The unit of observation is a municipality-year. Municipalities characteristics are extracted from the 2000 Census.

Table 3.2: Characteristics of municipalities below and above median registration rates, without bottom and top 10th percentiles.

	Below Median	Above Median	Diff.
Value transfers audited (log)	14.85	15.01	−0.158
	(1.422)	(1.141)	(0.0877)
Gini coefficient	0.552	0.547	0.00456
	(0.0733)	(0.0620)	(0.00462)
% adults with secondary school	10.60	13.87	−3.265***
	(5.380)	(7.246)	(0.435)
% illiterates	28.15	21.63	6.519***
	(13.11)	(13.25)	(0.898)
% civil servants	7.220	6.173	1.047***
	(4.567)	(3.602)	(0.280)
Per capita income (log)	5.405	5.721	−0.316***
	(0.606)	(0.569)	(0.0400)
% Rural	0.461	0.391	0.0700***
	(0.206)	(0.243)	(0.0153)
Population (log)	9.091	9.611	−0.520***
	(0.953)	(1.060)	(0.0687)
Observations	428	434	862

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis.

Note: This table compares means of the characteristics of municipalities that have been audited between 2006 and 2014. Columns 1 and 2 display means for municipalities below and above the median registration rates, respectively. Column 3 tests for differences in means between the two groups. The unit of observation is a municipality-year. Municipalities characteristics are extracted from the 2000 Census.

Table 3.3: Characteristics of municipalities in the samples selected based on net out-migration.

	Out-Migration Sample	General Sample	Diff.
Value transfers audited (log)	14.98	14.97	−0.0102
	(1.229)	(1.172)	(0.0727)
Gini coefficient	0.560	0.547	−0.0125**
	(0.0618)	(0.0699)	(0.00408)
% adults with secondary school	10.77	13.16	2.385***
	(5.265)	(7.509)	(0.410)
% illiterates	28.91	22.91	−6.000***
	(12.73)	(13.56)	(0.807)
% civil servants	7.562	6.267	−1.295***
	(4.393)	(4.133)	(0.258)
Per capita income (log)	5.369	5.669	0.300***
	(0.520)	(0.591)	(0.0344)
% Rural	0.466	0.400	−0.0658***
	(0.192)	(0.251)	(0.0140)
Population (log)	9.249	9.466	0.217***
	(0.903)	(1.148)	(0.0646)
Observations	437	704	1141

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parenthesis.

Note: This table compares means of the characteristics of municipalities that have been selected on the basis of net out-migration and positive variation in registration rates (out-migration sample), and the other remaining municipalities in the corruption sample, as described in Section 3.3.3. Column 3 tests for differences in means between the two groups. The unit of observation is a municipality-year. Municipalities characteristics are extracted from the 2000 Census.

Table 3.4: Average effects of registration rates on corruption.

Dep.Var.: *#CorruptionCases*	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS
Δ*Reg.Rates*	40.820***	35.750**	23.651**	23.726**	24.308**
	[14.998]	[14.381]	[10.582]	[10.481]	[10.212]
Value transfers audited (log)	–	–	–	–	3.934*
					[2.289]
Elections intercepts	✓	✓	✓	✓	✓
Draws intercepts		✓	✓	✓	✓
State intercepts			✓	✓	✓
Municipal characteristics				✓	✓
Observations	328	328	328	328	328
R^2	0.049	0.131	0.579	0.597	0.604

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: This table reports the effects of registration rates on the number of corruption events. Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. The sample includes municipalities with at least one positive variation in registration rate, with respect to a previous election, and that was above the 25th percentile in net out-migration. State intercepts spam 23 states and the sample includes 19 audit draws of municipalities between 2006-2014. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.5: Registration rates changes and corruption cases.

	Reduced Form	First-Stage	OLS	2SLS
	(1)	(2)	(3)	(4)
Dep. Vars.	$\#CorruptionCases$	$\Delta Reg.Rates$	$\#CorruptionCases$	$\#CorruptionCases$
Median $\Delta^d Reg.Rates$	−37.699*	−0.887***	–	–
	[22.317]	[0.228]		
$\Delta Reg.Rates$	–	–	23.726**	42.489*
			[10.481]	[24.603]
Elections intercepts	✓	✓	✓	✓
Draws intercepts	✓	✓	✓	✓
State intercepts	✓	✓	✓	✓
Municipal characteristics	✓	✓	✓	✓
Observations	328	328	328	328
R^2	0.591	0.460	0.597	0.592
$F - Exc.Inst.$				15.159

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: Columns: (1) Reduced form estimation where the dependent variable is the number of corruption cases found in an audit. The independent variable of interest is the median value of registration rates in all destinations of the municipality's out-migrants; (2) First-stage estimation where the dependent variable is the change in registration rates in a municipality. The independent variable of interest is the median value of registration rates in all destinations of the municipality's out-migrants; (3) These are the same OLS estimates presented in Column (4) of Table 3.4; (4) Two-Stage Least Squares, with Median $\Delta^d Reg.Rates$ used as an instrument for $\Delta Reg.Rates$. Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. The sample includes municipalities with at least one positive variation in registration rate, with respect to a previous election, and that was above the 25th percentile in net out-migration. State intercepts spam 23 states and the sample includes 19 audit draws of municipalities between 2006-2014. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.6: Median destination registration rates and the number of corruption cases - Placebo test.

Dep. Vars.	Reduced Form (1) $\#CorruptionCases$
Median $\Delta^d Reg.Rates$	2.297
	[32.140]
Elections intercepts	✓
Draws intercepts	✓
State intercepts	✓
Municipal characteristics	✓
Observations	521
R^2	0.514

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: The table depicts a reduced form estimation where the dependent variable is the number of corruption cases found in an audit. The independent variable of interest is the median value of registration rates in all destinations of the municipality's out-migrants. Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. The sample includes municipalities with at least one positive variation in registration rate, with respect to a previous election, and that was above the 25th percentile in net out-migration. State intercepts spam 23 states and the sample includes 19 audit draws of municipalities between 2006-2014. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.7: The effects of migration on the number of corruption cases.

	OLS				2SLS	
	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Vars.:	*#CorruptionCases*	*#CorruptionCases*	*#CorruptionCases*	*#CorruptionCases*	*#CorruptionCases*	*#CorruptionCases*
$\Delta Reg.Rates$	23.135**	23.217**	23.025**	23.647**	41.798*	41.486*
	[10.375]	[10.374]	[10.553]	[10.484]	[24.553]	[24.544]
Out-Migration	−0.005	–	–	–	–	–
	[0.003]					
In-Migration	0.008	–	–	–	–	–
	[0.006]					
Net Out-Migration	–	−0.005	–	–	–	–
		[0.003]				
Past Out-Migration	–	–	−5.081	–	−4.737	–
			[4.297]		[4.002]	
Past In-Migration	–	–	21.988	–	19.940	–
			[18.137]		[17.590]	
Past Net Out-Migration	–	–	–	−2.183	–	−2.134
				[3.250]		[2.996]
Elections intercepts	✓	✓	✓	✓	✓	✓
Draws intercepts	✓	✓	✓	✓	✓	✓
State intercepts	✓	✓	✓	✓	✓	✓
Municipal characteristics	✓	✓	✓	✓	✓	✓
Observations	328	328	328	328	328	328
R^2	0.601	0.600	0.599	0.598	0.594	0.593
$F - Exc.Inst.$					15.031	15.166

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: Columns: (1) Estimates equation 3.1 specification by OLS adding in- and out-migration flows as controls; (2) Estimates equation 3.1 specification by OLS adding net out-migration flows as controls; (3) Estimates equation 3.1 specification by OLS adding past in- and out-migration as controls (see details on these measures in Section 3.2); (4) Estimates equation 3.1 specification by OLS adding past net out-migration as controls (see details on these measures in Section 3.2); (5) 2SLS, with Median $\Delta^d Reg.Rates$ used as an instrument for $\Delta Reg.Rates$, and past in- and out- migration added as controls. (6) 2SLS, with Median $\Delta^d Reg.Rates$ used as an instrument for $\Delta Reg.Rates$, and past out migration added as controls. Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. The sample includes municipalities with at least one positive variation in registration rate, with respect to a previous election, and that was above the 25th percentile in net out-migration. State intercepts spam 23 states and the sample includes 19 audit draws of municipalities between 2006-2014. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.8: Council seats thresholds.

Threshold	# of seats	Population Interval
1	9	(0, 15]
2	11	(15, 30]
3	13	(30, 50]
4	15	(50, 80]
5	17	(80, 120]
6	19	(120, 160]
7	21	(160, 300]
8	23	(300, 450]
9	25	(450, 600]
10	27	(600, 750]
11	29	(750, 900]
12	31	(900, 1050]
13	33	(1050, 1200]
14	35	(1200, 1350]
15	37	(1350, 1500]
16	39	(1500, 1800]
17	41	(1800, 2400]
18	43	(2400, 3000]
19	45	(3000, 4000]
20	47	(4000, 5000]
21	49	(5000, 6000]
22	51	(6000, 7000]
23	53	(7000, 8000]
24	55	(8000, ∞)

Note: These are the thresholds that determine the maximum number of councillor seats a municipality can have. They are defined by Article 29 of the Federal Constitution of Brazil. Population is in thousands.

Table 3.9: Registration rates and the number of council candidates.

	Full-Sample		Threshold =1000	
	(1)	(2)	(3)	(4)
Dep. Vars.:	Registration Rates	# Councillors	Registration Rates	# Councillors
Max. # Councillors	−0.041***	0.266***	−0.014**	0.220***
	[0.002]	[0.005]	[0.006]	[0.039]
2^{nd}-order polynomial	✓	✓	✓	✓
3^{rd}-order polynomial	✓	✓	✓	✓
4^{th}-order polynomial	✓	✓	✓	✓
Municipal characteristics	✓	✓	✓	✓
Elections intercepts	✓	✓	✓	✓
Observations	13378	13632	532	540
R^2	0.258	0.337	0.121	0.234

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: The dependent variable is registration rates. Columns (1)-(4) use the full sample of municipalities from 2004 to 2012 to estimate different polynomial specifications; Columns (5)-(8) use the sample of municipalities that are within a 1000 inhabitants from the threshold that the determines the maximum number of councillors. Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.10: Registration rates and the number of council candidates.

	Full-Sample				Threshold =1000			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
# Councillors	−0.159***	−0.153***	−0.147***	−0.153***	−0.075**	−0.075**	−0.075**	−0.062**
	[0.008]	[0.008]	[0.008]	[0.008]	[0.033]	[0.033]	[0.033]	[0.029]
2^{nd}-order polynomial	✓	✓	✓	✓	✓	✓	✓	✓
3^{rd}-order polynomial		✓	✓	✓		✓	✓	✓
4^{th}-order polynomial			✓	✓			✓	✓
Municipal characteristics				✓				✓
Elections intercepts	✓	✓	✓	✓	✓	✓	✓	✓
Observations	13378	13378	13378	13378	532	532	532	532
R^2	.	0.015	0.029	0.132	0.021	0.023	0.023	0.089
$F - Exc.Inst.$	3010.666	2864.696	2800.565	2839.534	30.078	29.894	29.718	31.215

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: The dependent variable is registration rates. Columns (1)-(4) use the full sample of municipalities from 2004 to 2012 to estimate different polynomial specifications; Columns (5)-(8) use the sample of municipalities that are within a 1000 inhabitants from the threshold that the determines the maximum number of councillors. Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.11: Migrants and non-migrants characteristics.

	Movers	***NonMovers***	***Diff. Means***
	(1)	(2)	(3)
Education Levels			
None	0.469	0.546	0.0774***
	(0.115)	(0.139)	(0.00241)
Primary	0.195	0.177	−0.0176***
	(0.0502)	(0.0441)	(0.000894)
Secondary	0.265	0.204	−0.0608***
	(0.0799)	(0.0784)	(0.00150)
Tertiary	0.0709	0.0441	−0.0268***
	(0.0546)	(0.0306)	(0.000846)
Not known	0.00895	0.0286	0.0197***
	(0.0108)	(0.0295)	(0.000595)
Female	0.512	0.474	−0.0384***
	(0.0536)	(0.0285)	(0.000814)
Age	34.60	37.41	2.816***
	(2.823)	(4.423)	(0.0702)
Time Municipality (yrs)	3.691	0	−3.691***
	(0.572)	(0)	(0.00767)
Wage	1002.4	694.2	−308.2***
	(469.6)	(322.0)	(7.626)
Wage (MW)	1.966	1.361	−0.604***
	(0.921)	(0.631)	(0.0150)

Note: The table compares characteristics of migrants and non-migrants using data of the census 2010. The third column display the test for the difference in means between non-movers and movers.

Table 3.12: Migrants and non-migrants characteristics within three years of moving.

	Movers (1)	***NonMovers*** (2)	***Diff. Means*** (3)
Education Levels			
None	0.467	0.546	0.0793***
	(0.127)	(0.139)	(0.00252)
Primary	0.200	0.177	−0.0229***
	(0.0676)	(0.0441)	(0.00108)
Secondary	0.266	0.204	−0.0618***
	(0.0950)	(0.0784)	(0.00165)
Tertiary	0.0739	0.0441	−0.0298***
	(0.0589)	(0.0306)	(0.000907)
Not known	0.0136	0.0286	0.0150***
	(0.0191)	(0.0295)	(0.000734)
Female	0.503	0.474	−0.0286***
	(0.0713)	(0.0285)	(0.00103)
Age	33.53	37.41	3.884***
	(3.331)	(4.423)	(0.0741)
Time Municipality (yrs)	1.399	0	−1.399***
	(0.235)	(0)	(0.00315)
Wage	974.4	694.2	−280.2***
	(475.6)	(322.0)	(7.692)
Wage (MW)	1.911	1.361	−0.549***
	(0.932)	(0.631)	(0.0151)

Note: The table compares characteristics of migrants and non-migrants using data of the census 2010. The sample considered here includes only migrants who had been living in the host city for less than 3 years. The third column display the test for the difference in means between non-movers and movers.

Table 3.13: Registration rates and candidates' quality.

	OLS			2SLS		
	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Vars.:	Highly educated	Highly skilled	Incumbent	Highly educated	Highly skilled	Incumbent
Δ Registration Rate	-0.251^{**}	-0.250^{**}	0.215	-0.244	-0.837^{**}	0.199
	[0.111]	[0.116]	[0.225]	[0.335]	[0.384]	[0.643]
Elections intercepts	✓	✓	✓	✓	✓	✓
State intercepts	✓	✓	✓	✓	✓	✓
Municipal characteristics	✓	✓	✓	✓	✓	✓
Observations	648	648	279	648	648	279
R^2	0.103	0.079	0.122	0.103	0.047	0.122
$F - Exc.Inst.$				13.337	13.337	15.813

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: Columns: (1) and (4) display OLS and 2SLS models, respectively, that estimate the probability of a municipality having a candidate that has at least secondary education; (2) and (5) display OLS and 2SLS models, respectively, that estimate the probability of a municipality having a candidate who occupation is in a category classified as a high skilled occupation as defined in 3.2; (3) and (6) display OLS and 2SLS models, respectively, that estimate the probability of a municipality having an incumbent candidate, given that the candidate is not term constrained; (4) Estimates equation 3.1 specification by OLS adding past net out-migration as controls (see details on these measures in Section 3.2). Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. The sample includes municipalities with at least one positive variation in registration rate, with respect to a previous election, and that was above the 25th percentile in net out-migration. State intercepts spam 23 states. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.14: Registration rates and the attraction of resources.

	OLS		2SLS	
	(1)	(2)	(3)	(4)
Dep. Vars.: Transfers	Discretionary	Non-Discretionary	Discretionary	Non-Discretionary
Δ Registration Rate	−742.309**	−503.564	−2823.717***	−843.468
	[363.892]	[312.754]	[1023.282]	[620.640]
Elections intercepts	✓	✓	✓	✓
Draws intercepts	✓	✓	✓	✓
State intercepts	✓	✓	✓	✓
Municipal characteristics	✓	✓	✓	✓
Observations	317	317	316	316
R^2	0.204	0.181	0.127	0.177
$F - Exc.Inst.$			15.371	15.371

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in brackets.

Note: Columns: (1) and (3) OLS and 2SLS models, respectively, of the changes in registration rates on discretionary transfers using the high registration rates sample (see text for the definition of discretionary transfers); (1) and (3) OLS and 2SLS models, respectively, of the changes in registration rates on non-discretionary transfers using the high registration rates sample (see text for the definition of non-discretionary transfers). Municipal characteristics include population (log), percentage of the population that is literate, percentage of the population with secondary education, percentage of the population that lives in the rural sector, per capita income (log), Gini coefficient for income. The sample includes municipalities with at least one positive variation in registration rate, with respect to a previous election, and that was above the 25th percentile in net out-migration. State intercepts spam 23 states and the sample includes 19 audit draws of municipalities between 2006-2014. Standard errors are clustered at the municipality level. The unit of analysis is an election-year.

Table 3.15: Transitions to registration rates above and below one after re-registration in 2012.

	Δ ***Reg.Rate(2012)*** $<$ ***0***			Δ ***Reg.Rate(2012)*** $\geq$ ***0***		
	Reg.Rate(2012)<1	Reg.Rate(2012)>=1	Total	Reg.Rate(2012)<1	Reg.Rate(2012)>=1	Total
Reg.Rate(2008)<1	87.88	12.12	100.00	68.75	31.25	100.00
Reg.Rate(2008)>=1	39.88	60.12	100.00	20.78	79.22	100.00
Total	47.57	52.43	100.00	29.03	70.97	100.00

Note: The table shows transitions to registration rates above and below one, between 2008 and 2012, for municipalities that re-registered their voters in 2012. The first (second) column shows those who had a negative (non-negative) variation in registration rates after the re-registration program.

Table 3.16: Transitions to registration rates above and below one after re-registration in 2010.

	Δ ***Reg.Rate(2010)*** $<$ ***0***			Δ ***Reg.Rate(2010)*** $\geq$ ***0***		
	Reg.Rate(2010)<1	Reg.Rate(2010)>=1	Total	Reg.Rate(2010)<1	Reg.Rate(2010)>=1	Total
Reg.Rate(2008)<1	100.00	0.00	100.00	0.00	0.00	0.00
Reg.Rate(2008)>=1	77.27	22.73	100.00	0.00	100.00	100.00
Total	80.77	19.23	100.00	0.00	100.00	100.00

Note: The table shows transitions to registration rates above and below one, between 2008 and 2010, for municipalities that re-registered their voters in 2010. The first (second) column shows those who had a negative (non-negative) variation in registration rates after the re-registration program.

Table 3.17: Return migration rates.

Years 2006-2010				
	Rate			
Brazil (Total)	.0014			
	Mean	Std. Dev.	Min	Max
Average Municipality	.0024	.0043	.0001	.0943

Municipalities with the highest rates

Municipality	Rate
Chui	0.0943
Itai	0.0590
Sete Quedas	0.0518
Pato Bragado	0.0384
Santa Helena	0.0375
Missal	0.0362
Diamante D'Oeste	0.0356
Fernandes Tourinho	0.0350
Guaira	0.0312
Mercedes	0.0311
Santa Terezinha de Itaipu	0.0310
Itaipulandia	0.0306
Mundo Novo	0.0301

Note: The table displays descriptive statistics of return migration rates reported in the 2010 Census across municipalities and a list of municipalities with the highest rates.

Table 3.18: Immigration Rates (2001-2010).

Panel A - Years 2001-2005

	Rate			
Brazil (Total)	0.0003			
	Mean	Std. Dev.	Min	Max
Average Municipality	0.0001	0.0011	0	0.0626

Municipalities with the highest rates

Municipality	Rate
Chui	0.0626
Itaipulandia	0.0180
Entre Rios do Oeste	0.0148
Tabatinga	0.0118
Santa Helena	0.0109
Mercedes	0.0101

Panel B - Years 2006-2010

	Rate			
Brazil (Total)	0.0005			
	Mean	Std. Dev.	Min	Max
Average Municipality	0.0002	0.0018	0	0.1154

Municipalities with the highest rates

Municipality	Rate
Chui	0.1154
Missal	0.0184
Tabatinga	0.0167
Santana do Livramento	0.0152
Pato Bragado	0.0135
Santa Helena	0.0130
Sao Miguel do Iguaçu	0.0120
Guaira	0.0118
Campo Alegre	0.0114
Mercedes	0.0113
Bonfim	0.0111
Ponta Pora	0.0104
Pauini	0.0102

Note: The table displays descriptive statistics of immigration rates reported in the 2010 Census across municipalities and a list of municipalities with the highest rates.

Bibliography

Alesina, Alberto, Reza Baqir, and William Easterly. 1999. "Public Goods and Ethnic Divisions." *Quaterly Journal of Economics*, 114(4): 1243–1284.

Altonji, Joseph G., and David Card. 2007. "The effects of immigration on the labor market outcomes of less-skilled natives." In *Immigration, trade, and the labor market.* , ed. John M. Abowd and Richard B. Freeman, 201–234. University of Chicago Press.

Ansolabehere, Stephen, and Shanto Iyengar. 1994*a*. "Does attack advertising demobilize the electorate?" *American Political Science Review*, 88(4): 829–838.

Ansolabehere, Stephen, and Shanto Iyengar. 1994*b*. "Riding the wave and claiming ownership over issues: the joint efffects of advertising and news coverage in campaigns." *Public Opinion Quaterly*, 58: 335–357.

Anzia, Sarah F., and Christopher R. Berry. 2011. "The Jackie (and Jill) Robinson effect: Why do congresswomen outperform congressmen?" *American Journal of Political Science*, 55(3): 478–493.

Batista, Catia, and Pedro C. Vicente. 2011. "Do migrants improve governance at home? Evidence from a voting experiment." *World Bank Economic Review*, 25(1): 77–104.

Battaglini, Marco, Rebecca B. Morton, and Thomas R. Palfrey. 2009. "The Swing Voter's Curse in the Laboratory." *Review of Economic Studies*, 77(1): 61–89.

Benitez, Daniel A., Antonio Estache, D. Mark Kennet, and Christian A. Ruzzier. 2000. "Are cost models useful for telecoms regulators in developing countries?" *World Bank Publications*, 2384.

Bhalotra, Sonia, and Irma Clots-Figueras. 2014. "Health and the Political Agency of Women." *American Economic Journal: Economic Policy*, 6(2): 164–197.

Bhalotra, Sonia, Irma Clots-Figueras, and Lakshmi Iyer. 2017. "Path-Breakers: How Does Women's Political Participation Respond to Electoral Success?" *Economic Journal*.

Borjas, G J, R B Freeman, and L F Katz. 1997. "How much do immigration and trade affect labor market outcomes?" *Brookings papers on economic activity*, 1997(1): 1–90.

Boukouras, Aristotelis, and Kostas Koufopoulos. 2008. "Political Competition, Ideology and Corruption." *SSRN Electronic Journal*, 1–44.

Bouton, Laurent. 2013. "A Theory of Strategic Voting in Runoff Elections." *American Economic Review*, 103(4): 1248–1288.

Bouton, Laurent, and Gabriele Gratton. 2015. "Majority runoff elections: Strategic voting and Duverger's hypothesis." *Theoretical Economics*, 10(2): 283–314.

Brollo, Fernanda, and Ugo Troiano. 2016. "What Happens When a Woman Wins a Close Election? Evidence from Brazil." *Journal of Development Economics*, 122: 28–45.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini. 2013. "The Political Resource Curse." *American Economic Review*, 103(5): 1759–96.

Campante, Filipe R., Ruben Durante, and Francesco Sobbrio. 2013. "Politics 2.0: The multifaceted effect of broadband internet on political participation." *NBER Working Paper Series*.

Campante, Filipe, Ruben Durante, and Francesco Sobbrio. 2014. "Politics 2.0: The Multifaceted Effect of Broadband Internet on Political Participation." *Working Paper*.

Card, D. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Casas-Arce, Pablo, and Albert Saiz. 2015. "Women and Power: Unpopular, Unwilling, or Held Back?" *Journal of Political Economy*, 123(3): 641–669.

Castanheira, Micael. 2003. "Why vote for losers?" *Journal of the European Economic Association*, 1(5): 1207–1238.

Chauvet, Lisa, and Marion Mercier. 2014. "Do return migrants transfer political norms to their origin country? Evidence from Mali." *Journal of Comparative Economics*, 42(3): 630–651.

Coate, Stephen, Michael Conlin, and Andrea Moro. 2008. "The performance of pivotal-voter models in small-scale elections: Evidence from Texas liquor referenda." *Journal of Public Economics*, 92(3-4): 582–596.

Costa, Francisco, Jason Garred, and João Paulo Pessoa. 2016. "Winners and Losers from a Commodities-for-Manufactures Trade Boom." *International Economic Review*.

Czernich, Nina, and Oliver Falck. 2011. "Broadband infrastructure and economic growth." *The Economic Journal*, , (121): 505–532.

Da Silveira, B. S., and J. M. P. De Mello. 2011. "Campaign Advertising and Election Outcomes: Quasi-natural Experiment Evidence from Gubernatorial Elections in Brazil." *The Review of Economic Studies*, 78(2): 590–612.

DATAFOLHA. 2010*a*. "Intenção de Voto para Presidente da República - Avaliação do Presidente Lula." *CESOP-UNICAMP*, September(03288).

DATAFOLHA. 2010*b*. "Intenção de Voto para Presidente da República - Avaliação do Presidente Lula." *CESOP-UNICAMP*, July(03291).

Degan, Arianna, and Antonio Merlo. 2011. "A Structural Model of Turnout and Voting in Multiple Elections." *Journal of the European Economic Association*, 9(2): 209–245.

DellaVigna, Stefano, and E Kaplan. 2007. "The Fox News effect: Media bias and voting." *The Quarterly Journal of Economics*, 122(3): 1187–1234.

DellaVigna, Stefano, and Matthew Gentzkow. 2010. "Persuasion: Empirical Evidence." *Annual Review of Economics*, 2(1): 643–669.

Dellavigna, Stefano, John List, and Ulrike Malmendier. 2015. "Voting to Tell Others." *Working Paper*.

De Mello, J. M. P., Sergio Firpo, and Marcos Chamon. 2009. "Electoral Rules, Political Competition and Fiscal Spending: Regression Discontinuity Evidence From Brazilian Municipalities." *IZA Discussion Papers, No. 4658*.

Dix-Carneiro, Rafael, and Brian K Kovak. 2015. "Trade Reform and Regional Dynamics: Evidence From 25 Years of Brazilian Matched Employer-Employee Data." *NBER Working Paper*, No. 20908.

Dix-Carneiro, Rafael, Rodrigo R Soares, and Gabriel Ulyssea. 2016. "Local Labor Market Conditions and Crime: Evidence from the Brazilian Trade Liberalization." , (9638).

Docquier, Frédéric, Elisabetta Lodigiani, Hillel Rapoport, and Maurice Schiff. 2015. "Emigration and democracy."

Duflo, Esther, Rachel Glennerster, and Michael Kremer. 2007. "Using Randomization In Development Economics Research: a Toolkit." In *Handbook of Development Economics*. Vol. 4, , ed. T. Paul Schultz and John A. Strauss, 3895–3962.

Dustmann, Christian, Tim Hatton, and Ian Preston. 2005. "The labor market effects of immigration." *The Economic Journal*, 115: F297–F299.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2011. "Media and Political Persuasion: Evidence from Russia." *American Economic Review*, 101(7): 3253–3285.

Enli, Gunn Sara, and Eli Skogerbø. 2013. "Personalized Campaigns in Party-Centred Politics." *Information, Communication & Society*, 16(5): 757–774.

Falck, Oliver, Robert Gold, and Stephan Heblich. 2014. "E-lections: Voting Behavior and the Internet." *American Economic Review*, 104(7): 2238–2265.

Feddersen, Timothy J, and Wolfgang Pesendorfer. 1995. "The Swing Voter's Curse." *The American Economic Review*, 86(3): 408–424.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *Quarterly Journal of Economics*, 123(2): 703–745.

Ferraz, Claudio, and Frederico Finan. 2011. "Electoral accountability and corruption: Evidence from the audits of local governments." *American Economic Review*, 101(4): 1274–1311.

Fujiwara, Thomas. 2010. "Voting technology, political responsiveness, and infant health: evidence from Brazil." *Department of Economics, University of British ...*, 83(2): 423–464.

Fujiwara, Thomas. 2011. "A Regression Discontinuity Test of Strategic Voting and Duverger's Law." *Quarterly Journal of Political Science*, 6(3-4): 197–233.

Fulton, Sarah A. 2014. "When Gender Matters: Macro-dynamics and Micro-mechanisms." *Political Behavior*, 36(3): 605–630.

Funk, Patricia. 2010. "Social Incentives and Voter Turnout: Evidence from the Swiss Mail Ballot System." *Journal of the European Economic Association*, 8(5): 1077–1103.

Gagliarducci, Stefano, and M. Daniele Paserman. 2012. "Gender interactions within hierarchies: Evidence from the political arena." *Review of Economic Studies*, 79(3): 1021–1052.

Gagliarducci, Stefano, Tommaso Nannicini, and Paolo Naticchioni. 2011. "Electoral rules and politicians' behavior: A micro test." *American Economic Journal: Economic Policy*, 3(3): 144–174.

Gentzkow, Matthew. 2006. "Television and Voter Turnout." *Quarterly Journal of Economics*, 121(3): 931–972.

Gneezy, Uri, Muriel Niederle, and Aldo Rustichini. 2003. "Performance in competitive environments: gender differences." *Quarterly Journal of Economics*, 118(3): 1049–1074.

Hicks, John. 1963. *Theory of wages.* Springer.

Hoffman, Mitchell, John Morgan, and Collin Raymond. 2013. "One in a Million: A Field Experiment on Belief Formation and Pivotal Voting." *Working Paper*.

Imbens, Guido, and Karthik Kalyanaraman. 2012. "Optimal bandwidth choice for the regression discontinuity estimator." *Review of Economic Studies*, 79(3): 933–959.

Imbens, Guido W., and Thomas Lemieux. 2008. "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*, 142(2): 615–635.

International Labour Organization. 2012. "International Standard Classification of Occupations." I: 433.

Jaber, Ahmed. 2013. "Broadband Internet and Political Behavior: Evidence from the United States." *Working Paper*.

Kawai, Kei, and Yasutora Watanabe. 2013. "Inferring Strategic Voting." *American Economic Review*, 103(2): 624–662.

Kelly, Tim, and Carlo Maria Rossotto. 2012. *Broadband Strategies Handbook.* World Bank Publications.

Knight, Peter T. 2014. *The Internet in Brazil: Origins, Strategy, Development, and Governance.* Bloomington:AuthorHouse.

Levitt, Steven D. 1994. "Using Repeat Challengers to Estimate The Effect of Campaign Spending on Election Outcomes in the US House." *Journal of Political Economy*, 102(4): 777–798.

Lipscomb, Molly, Mushfiq Mobarak, and Tania Barham. 2013. "Development Effects of Electrification: Evidence from the Topographic Placement of Hydropower Plants in Brazil." *American Economic Journal: Applied Economics*, 5(2): 200–231.

Martinelli, César. 2002. "Simple Plurality Versus Plurality Runoff with Privately Informed Voters." *Social Choice and Welfare*, 19(4): 901–919.

Martinelli, César. 2006. "Would rational voters acquire costly information?" *Journal of Economic Theory*, 129(1): 225–251.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698–714.

Miner, Luke. 2015. "The Unintended Consequences of Internet Diffusion: Evidence from Malaysia." *Journal of Public Economics*, 132: 66–78.

Mizukami, Pedro, Jhessica Reia, and Joana Varon. 2013. "Mapping Digital Media: Brazil." Open Society Foundations.

Motta, Marcelo Paiva Da. 2012. "Topologia dos backbones de internet no Brasil." *Sociedade e Natureza*, 21–36.

Murray, R. 2008. "The Power of Sex and Incumbency: A Longitudinal Study of Electoral Performance in France." *Party Politics*, 14(5): 539–554.

Nannicini, Tommaso, Massimo Bordignon, and Guido Tabellini. 2016. "Moderating Political Extremism: Single Round vs Runoff Elections under Plurality Rule." *American Economic Review*, 106(8): 2349–70.

Ostendorp, M, and G Gela. 1997. "Fiber Optic Cables in Overhead Transmission Corridors: A State-of-the-Art Review." Electric Power Research Institute.

Papke, Leslie E., and Jeffrey M. Wooldridge. 1996. "Econometric Methods For Fractional Response Variables with an Application to 401(k) Plan Participation Rates." *Journal of Applied Econometrics*, 11(6): 619–632.

Persson, Torsten, Guido Tabellini, and Francesco Trebbi. 2003. "Electoral Rules and Corruption." *Journal of the European Economic Association*, 1(4): 958–989.

Petrova, Maria, Ananya Sen, and Pinar Yildirim. n.d.. "Social Media and Political Donations: New Technology and Incumbency Advantage in the United States." *Working Paper*.

Piketty, Thomas. 2000. "Voting as Communicating." *The Review of Economic Studies*, 67(1): 169–191.

Poterba, James M. 1996. "Demographic Structure and the Political Economy of Public Education."

Roberts, Andrew, Jason Seawright, and Jennifer Cyr. 2012. "Do Electoral Laws Affect Women's Representation?" *Comparative Political Studies*, 46(12): 1555–1581.

Rodrik, Dani, Arvind Subramanian, and Francesco Trebbi. 2004. "Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development." *Journal of Economic Growth*, 9(2): 131–165.

Roy, A. D. 1951. "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers*, 3(2): 135–146.

Samuels, David J. 2001. "Does Money Matter ? Credible Commitments and Campaign Finance in New Democracies Theory and Evidence from Brazil." *Comparative Politics*, 34(1): 23–42.

Sanbonmatsu, Kira. 2002. "Political parties and the recruitment of women to state legislatures." *Journal of Politics*, 64(3): 791–809.

Sjaastad, Larry A. 1962. "The Costs and Returns of Human Migration." *Journal of Political Economy*, 70(5): 80–93.

Snyder, James M., and David Strömberg. 2010. "Press Coverage and Electoral Accountability." *Journal of Political Economy*, 118(2): 355–408.

Strömberg, David. 2004. "Radio's Impact on Public Spending." *The Quarterly Journal of Economics*, 119(1): 189–221.

Windmeijer, Frank, and João M. C. Santos Silva. 1997. "Endogeneity in Count Data Models: An Application to Demand for Health Care." *Journal of Applied Econometrics*, 12(3): 281–294.