Learn from thy Neighbor: Do Voters Associate Corruption with Political Parties?*

Arieda Muço†

Job Market Paper

Please find the latest version Here

January 6, 2017

Abstract

This paper exploits a randomized audit program to document how information about a corrupt politician causes electoral spillovers on his party. I focus on two types of spillovers: spillovers across types of elections (cross-electoral spillovers) and spillovers across jurisdiction borders (cross-border spillovers). Using detailed data on radio antenna location and coverage, I identify neighboring areas in the same media market. Moreover, via machine learning and text-analysis tools, I take a data-driven approach to create an index that ranks municipalities according to their corruption level. I uncover how information about corruption shapes voting decisions through the structure and geography of the media market. I show that voters hold the party of the incumbent politician accountable in four distinct ways. In municipalities where corruption occurs, voters punish parties in (1) local and (2) national elections. Most importantly, I show that news of a politician's corruption affects his party in neighboring municipalities that share the same media market, and these spillovers affect both (3) local and (4) national elections. Ruling out other potential mechanisms, I show that these findings are consistent with electoral accountability. Finally, a back-of-the-envelope calculation suggests that an audit occurring at the local level causes the party's vote support in the audited neighborhood (audited municipality and its neighbors) to decrease by 10 percentage points.

* I am indebted to Manuel Arellano, Erik Lindqvist, Torsten Persson, Diego Puga, Giancarlo Spagnolo, and David Strömberg for advice, support and encouragement at different stages of this project. I am extremely grateful to Felipe Carozzi for extensive conversations and help. This paper has also benefited from conversations with Samuel Bentolila, Martina Björkman Nyqvist, Guillermo Caruana, Kerem Coşar, Tore Ellingsen, Ruben Enikolopov, Stefano Gagliarducci, Mounir Karadja, Andreas Madestam, Mónica Martínez-Bravo, Anders Olofsgård, Elena Paltseva, Maria Perrotta Berlin, Maria Petrova, and Jakob Svensson. I thank Reinaldo Luz, Fernanda Senra (discussant), Ricardo Lima (research assistant), and Robson Tigre for help with understanding of the Brazilian institutions. The help of the staff at the CGU, especially Giuliana Biaggini, Adriano Souza, Camilla Bezzera, Maria Rezende and Günther Lorenz at FMSCAN/FMLIST, is gratefully acknowledged. I also thank participants in seminars: CEMFI, IIES, SSE, UAB, UCL, Stockholm University, Micro Workshop Stockholm, IFN Stockholm, University of Gothenburg, Max Planck Institute and in conferences: SITE Corruption, Stockholm; Enter Jamboree, Carlos III; IPSWD, Columbia University; ASWEDE, Gothenburg; SUDSWEc, Stockholm University; European Winter Meetings of the Econometric Society, Edinburgh; Spanish Economic Association, Bilbao. All errors are my own.

† Stockholm School of Economics. E-mail: arieda.muco@gmail.com

1 Introduction

Corruption is not a victimless crime. It has been widely argued that corruption is to blame for impediments in growth and economic development (Treisman (2000)). Despite a large literature on corruption, its causes and consequences are not yet fully understood. Institutions that promote accountability, such as elections, are among the main prescriptions for fighting malfeasance. Well-functioning political institutions, however, require an informed electorate. Indeed, if provided with information, voters punish corrupt politicians.[1] It is, however, an open question whether the party of the corrupt politician is also sanctioned.

In this paper, I show that the actions of individual politicians have broad consequences for their party. More precisely, I examine how news of individual malfeasance impacts the party at the local level and how it spills over across jurisdictions. Highlighting important features of information disclosure that have not been previously documented, I find that voters hold the party accountable in four distinct ways. Within municipalities where corruption occurs, voters punish parties in (1) local and (2) national elections. Most importantly, I show that news of a politician's corruption also affects neighboring municipalities that share the same media market, and these spillovers affect both (3) local and (4) national elections. These effects operate through voters' recognition of party names.

To estimate the spillovers from individual corruption, I take advantage of a randomized audit program that was initiated by the Brazilian federal government in 2003. This on-going program aims at fighting corruption in local governments. A public lottery determines which municipalities are audited. Once a municipality is selected, a team of auditors is sent within a week to inspect. After an investigation has been concluded, the auditors provide a detailed report which is made public.

In a country plagued by endemic corruption, audit investigations reveal corruption in 90 percent of the cases. I use the randomized timing of audits with respect to elections to investigate the spillovers of political corruption. My empirical analysis consists of two main parts. First, I investigate the effect of an audit on the electoral outcomes across jurisdictions. Specifically, I examine the impact of these investigations on the audited municipality and its neighbors, both in local and national elections.[2]

Second, I explore how voters respond to reported levels of corruption, given that the

[1] See Ferraz and Finan (2008); Costas-Pérez, Solé-Ollé and Sorribas-Navarro (2012); Bobonis, Cámara Fuertes and Schwabe (2016).

[2] Local and national elections alternate each other ever two years.

intensity of the punishment may depend on the severity of the corruption disclosed. To do so, I use the detailed assessment of the auditors and put together a dataset by scraping more than 2000 PDF reports which span the period from 2003 to 2013. Moreover, via machine learning tools, I take a data-driven approach to construct a corruption index that ranks municipalities according to the amount of disclosed corruption.[3]

To build the corruption index, I create a severe irregularity count that records the number of severe cases of irregularities found. From the audit reports, I also build several other variables that correlate with the amount of corruption uncovered, namely the total amount of irregularities found, the number of pages, the number of lines, and the number of images. I then combine these variables in a principal component analysis (PCA). The main idea behind the PCA is that the *latent* common factor measures the underlying corruption. This methodology has the advantage that once the criteria for classification are set, the index does not rely on subjective decision making.[4]

I start by documenting the effect of an audit in the audited municipality (benchmark case) on vote shares in elections, both for the same level of government where information is disclosed and for higher levels of government (cross-election spillovers). For this purpose, I compare municipalities that are audited before elections with municipalities that are audited after elections.[5] To examine whether higher-level elections are affected by an audit investigation I focus on the parties that compete in local and national elections under the same party label.[6]

Next, I turn to *geographical spillovers* – the effect of the audit on the electoral outcomes in the neighboring municipalities – in elections for the same level of government where information is disclosed (cross-border spillovers), and in elections for higher levels of government (cross-border and cross-electoral spillovers). Following the recent political economy and media literature, which provides evidence of media being the relevant channel of information,

[3]The irregularities fall under the responsibility of the politician in office.

[4]Previous studies such as Ferraz and Finan (2008, 2011); Brollo et al. (2013); Litschig and Zamboni (2011); Timmons and Garfias (2015) in Brazil, Larreguy, Marshall and Snyder Jr (2014) in Mexico, and Bobonis, Cámara Fuertes and Schwabe (2016) in Puerto Rico have relied on hand-coded measures of the audit reports.

[5]Given the randomization of the audit and its timing, these two groups should be homogenous. I check that this is indeed the case by performing some randomization checks, as described in Section 4.1.

[6]There are only two main parties that show up throughout my sample with the same party names, Partido da Social Democracia Brasileira (PSDB) and Partido dos Trabalhadores (PT). These two parties have been the main contestants in the race for president since the democratisation of Brazil. Similar to local elections, I compare municipalities audited before presidential elections with municipalities audited after presidential elections.

I look for an effect in the *informed neighbors* of the audited municipality.[7] Using a unique dataset detailing the location and coverage of the universe of FM antennas in the country, I compute the fraction of the population covered by an antenna in each municipality. *Informed neighbors* are defined as nearby neighbors of the audited municipality which are contiguous (share a vertex or border) and have a positive population fraction covered by the antenna signal emitted from the audited municipality.

The choice of radio rather than other potential media outlets is natural in the Brazilian context. Media penetration in Brazil is dominated by television. However, in most TV broadcasts, state and national channels prevail, making it unlikely for local news to be transmitted by these outlets. Radio does not only have high penetration rates but is also oftentimes locally provided. According to *Latinobarómetro* 40 percent of Brazilians use radio as a source of political information. I do not take newspapers into account since Brazil has one of the lowest rates of newspaper penetration in the world, with an average circulation of only 42 newspaper copies per 1000 inhabitants, (Porto (1985); IBOPE (2012)).

When investigating geographical spillovers, I compare the group of *informed neighbors* that receive information about the audit outcome (in neighboring municipalities) before elections with the group of *informed neighbors* that receive the information after elections. The randomization of the audit makes these two groups comparable.[8]

Several findings emerge from this analysis. An audit reduces the vote support of the incumbent party in local elections in the audited municipality, by 2-3 percentage points. The reelection likelihood of the party remaining in power is also reduced. However, the intensity of corruption does not affect the vote outcome. These findings are consistent with the opposition having hard evidence to effectively attack the incumbent party. Indeed, in the audited municipalities opposition parties are 20 percent more likely to increase their campaign expenditure.

The electoral effects of an audit are not isolated to the audited municipality. They also spillover to neighboring governments. In the informed neighboring municipalities, in local elections, the audit itself does not affect the electoral outcomes but the electoral support for the party is reduced when the level of disclosed corruption increases. These findings are

[7] See for example, Strömberg (2004); DellaVigna and Kaplan (2007); Enikolopov, Petrova and Zhuravskaya (2010); Snyder and Strömberg (2010); DellaVigna et al. (2014); Yanagizawa-Drott (2014); Bursztyn and Cantoni (2015).

[8] I show that these two groups are indeed comparable by performing balancing tests described in the Appendix.

in line with the political science literature, which stresses that the candidate is the most important figure in local elections (Pereira and Mueller (2003); Desposato (2004); Folke, Persson and Rickne (2016)). Voters in neighboring municipalities do not seem to react to an audit investigation because their local politician is not affected. The finding that the level of disclosed corruption matters to voters in neighboring municipalities suggests that the likelihood of voters learning about party characteristics increases with the amount of disclosed corruption.

The electoral effects at higher levels of government are similar in the audited neighborhood (audited municipality and its neighbors). In the neighborhood, the vote support for the party is reduced by 4 percentage points in the audited municipality and by 3 percentage points in the informed neighboring municipalities. These municipalities are also punished more harshly if more corruption is revealed. Taken together, these findings indicate that voters sharing the same information flow behave similarly. The strong electoral effect suggests that these corruption scandals were triggering voters' memories of the party's wrongdoing at the local level. Indeed, anecdotal evidence suggests that the presidential candidates have been using these corruption cases to blame and shame each other in public debates.

Throughout this paper, I maintain the assumption that voters recognize party labels. To test for this, I estimate the impact of an audit investigation on the main party in the coalitions formed for presidential elections. The focus on the main party is necessary because the coalition is recognized as the party of the presidential candidate. If the name of the party is the relevant information for voters, there should be no effect of the audit on the vote share of the coalition. Indeed, this hypothesis is confirmed by the data.

While the electoral effects at higher levels of government and the spillovers across bordering jurisdictions are novel contributions of this paper, the effect of the audit on local electoral outcomes has previously been documented. For instance, using the Brazilian context, Ferraz and Finan (2008) find that incumbent mayors face a reduction of the likelihood of reelection by 7 percentage points in municipalities where at least two violations associated with corruption were reported. Larreguy, Marshall and Snyder Jr (2014) exploit the random timing of the information release in Mexican municipalities to investigate the electoral outcomes of the incumbent political party, given that in Mexico the politician cannot rerun for elections.[9] The authors show that voter sanctioning increases with the increase in the severity of

[9]In Mexico, the audit process is not random, but the timing of the release of the findings relative to elections is random.

disclosed malfeasance. Bobonis, Cámara Fuertes and Schwabe (2016) investigate, in Puerto Rico, how news about the incumbent politician affects his reelection chances. In Puerto Rico the audit is not random, but the timing of municipal audits is predetermined with respect to an election. The authors find that incumbent mayors in municipalities with a negative audit outcome are less likely to run for reelection. It also leads to a significant short-term reduction in municipal corruption but they do not exhibit decreased levels of corruption in subsequent audits.

This paper contributes to several strands of the literature. First, it relates to the broad literature on effective policies designed to reduce corruption (Reinikka and Svensson (2005); Björkman and Svensson (2009); Gonzalez (2016)). Second, it relates to the literature that considers audits to be an effective anti-corruption tool in developing countries (Di Tella and Schargrodsky (2003) and Olken (2007)). Differently from these studies, I show that the effects of audits are magnified due to spillovers. Third, similarly to the growing empirical literature on the electoral effects of access to media (Strömberg (2004); Gentzkow (2006); DellaVigna and Kaplan (2007); Enikolopov, Petrova and Zhuravskaya (2010); Snyder and Strömberg (2010); Adena et al. (2015)), my study confirms the finding that the media structure plays an important role in electoral accountability. In addition and similar to the work by Larreguy, Marshall and Snyder Jr (2014), my findings stress the relevance of local media (rather than national media outlets). Fourth, the paper also contributes to the literature on collective blame, providing an empirical test to the theoretical contribution pioneered by Tirole (1996).

Overall the findings in this paper are also relevant towards understanding why national parties might have incentives to better monitor and improve selection among local candidates. At least in the short run, politicians' behaviour has an impact on the party across government borders and levels, even in a context like Brazil where party labels, in local politics, are considered to be weak.

The rest of this paper is structured as follows. Section 2 discusses the institutional settings. Section 3 describes the data sources and provides summary statistics. Section 4 outlines the empirical analysis and provides results on the effect of an audit in the audited municipality. Section 5 considers geographical spillovers. Section 6 provides additional findings. Section 7 addresses robustness and different placebo tests. Section 8 discusses the findings of the paper and Section 9 concludes the paper.

2 Institutional Setting

2.1 The Audit Program

The randomization of municipal audits in Brazil was initiated in 2003 with the objective of fighting corruption in local governments. The program – which continues to date – is implemented by the Controladoria Geral da União (CGU) and consists of random selection, for audit purposes, of Brazilian local governments. Under examination are the transfers allocated by the federal government to each municipality.

Before the lottery is held, the CGU publishes the list of eligible municipalities jointly with the lottery numbers attributed to each municipality. Once the municipalities to be audited have been chosen, a team of auditors is sent to inspect the use of the transfers, following guidelines from CGU headquarters. In the first rounds of the program, all federal transfers were audited. As the program proceeded, it was decided that only a subset of funds would be inspected. Funds under inspection are also randomly selected and announced jointly with eligible municipalities and lottery numbers attributed. Education and Health are the areas that receive more audits, they are also the areas which have the most cases of mismanagement.[10] State capitals and municipalities with populations above (or below) certain population thresholds are also excluded. If a municipality is drawn, that municipality is re-eligible for selection after a predetermined number of rounds.[11] The rules for exclusion and the funds to be audited are made publicly available the week before the lottery is held. After several days of investigation in the municipality, the auditors compile a report of their findings.[12]

Anecdotal evidence suggests that media covers issues related to the audit outcomes. This implies that citizens learn about the audit outcomes either directly – by observing the auditors in the municipality – or indirectly – from various media sources. For example, in one of the biggest national newspapers, *Folha de S. Paulo*, some of the mayors were mainly mentioned due to the exposure via the audit program.[13] Published on October 1 2004, two days before the first round of municipal elections, the headlines of several articles cite: *"Suspects of fraud*

[10] 25 percent of the overall irregularities in the sample have occurred in the funds related to Education.

[11] In the sample, a municipality is non-eligible for some rounds. The maximum number of rounds for exclusion during my sample period has been 12 rounds. More information on the population cutoffs in Appendix A1.

[12] In the first page of the audit reports the auditors state the days they have spent in the municipality and the average duration is around one week.

[13] Available here: http://acervo.folha.com.br/fsp/2004/10/01/125/.

[will cause an] overthrow of 199 mayors" The title of another article *"40 mayors accused of corruption are running for reelection"* and a third article talking about mayors contesting the CGU investigations where one of the incumbent mayors running for reelections refers to the negative outcome of the audit in his municipality:

> "People knew the opposition would take advantage over this in the elections..."

Another mayor complains about the release of the audit before the municipal elections stating:

> "It is an irresponsibility from the CGU to point out at not proven irregularities in electoral period, so they can be used by the opposition like these [the irregularities] were crimes or diversion of funds.[14] It is a disaster because my opponent is from PT.[15]

Between 2003 and 2013, namely the audit years I use in this analysis, a total of 2053 audits have occurred and more than 400 municipalities have been audited twice and 30 have been audited three times. The timing of the audits is illustrated in Figure 1. Excluding the pilot audits in 2003, where one municipality per state was audited, the number of audited municipalities increased from 50 to 60 municipalities per round after the year 2005.

2.2 Brazilian Municipalities

Brazil is divided into 5,570 municipalities and 27 federal units comprising 26 states and the Federal District, where the national capital, Brasília, is located.[16] Municipal governments correspond to the lowest tier of public administration in the country. They are composed of an executive branch chaired by the mayor (*prefeito*), a vice-mayor, and a multi-member legislative chamber (*câmara municipal*) constituted of council members also known as *vereadores*. Municipalities vary greatly in population, size, wealth and literacy rates.

The Brazilian Constitution guarantees the independence of municipal governments. Local governments are responsible for providing certain goods and services including primary education and healthcare in coordination with the state and the federal authorities. While larger municipalities can generate up to 40 percent of their revenues via taxes, smaller size

[14]The mayor of this municipality, as all mayors, was given the chance to reply to the draft of the report and provide justifications or proofs to overturn the findings. However, for most of the findings, he was not able to provide any credible justification. (Municipality of Laguna in Santa Catarina, audited in lottery 8.)

[15]PT is the presidential party throughout my sample period."

[16]Table A2 provides a more detailed description of the subdivision of the states in the country.

municipalities depend heavily on intergovernmental transfers. Municipalities with a population size below 50,000 inhabitants generate roughly 80 percent of their revenues via federal transfers (e.g Farvacque-Vitkovic and Kopanyi (2014)).

2.3 Elections and Political Parties

National and municipal elections in Brazil are, respectively, held every four years. Thus, there are elections every two years, alternating between national and municipal elections. Elections occur simultaneously throughout the country on the first Sunday of October. Voting is compulsory in Brazil in all types of elections and all Brazilian citizens cast their votes under the same electoral rules.[17] If voters cannot participate in an election, a justification form can be filled out at election centres and post offices. Otherwise, a fine must be paid. Election timing in the sample joint with the lottery rounds are shown in Figure 1.

[Figure 1 Here]

In municipal elections, voters are called to elect the mayor and members of the municipal chamber, *vereadores*. The chamber of local legislators acts as a supervisory authority and the size is directly related to the size of the population in the municipality. The voting system used to elect the municipal government is majority ruling. Exceptions are municipalities with a population size above 200,000 voters, for which the mayoral elections require a second round. Mayors face term limits, they can only get elected for two consecutive terms. If a mayor has already served two consecutive terms, he is eligible for reelection after a term out of office. National legislators and the president are elected in general elections.[18]

Brazil has a multi-party system both at the federal and the local level. More than 30 parties participate in local elections each term; however, only few are relevant at the national level. Since 1994, only two parties have dominated the race for president: *Partido da Social Democracia Brasileira* (PSDB) and *Partido dos Trabalhadores* (PT). Since 1995, the president has been from either of these parties. They are also the only parties that have a presidential candidate throughout the sample period.[19] In presidential elections, the

[17] Participation in the elections is compulsory for literate citizens between 18 and 70 years old. It is non-compulsory for Brazilians aged 16-17 or over 70 or illiterate citizens of any age.

[18] The president is elected through a two-round system. I consider the votes in the first round for municipal and presidential elections.

[19] Fernando Henrique Cardoso (PSDB) was in power from 1995 to 2003. PT has held the presidency since 2003 to 2016. First with Luis Inácio Lula Da Silva (2003-2011) and latter with Lula's handpicked successor Dilma Rousseff (2011-2016).

identities of these two parties are strong and well-recognized. At the local level, the political power is highly fragmented between numerous parties.[20]

PT is a relatively minor player at the local level, holding slightly more than 3 percent of local governments in the 2001 electoral term while PSDB ruled over a sixth of Brazilian municipalities in the same electoral term. Figure 2 shows the distribution of PT and PSDB in the Brazilian territory during the electoral term of 2001.

[Figure 2 Here]

In Brazil, party identities are considered to be weak (Desposato (2004); Pereira and Mueller (2003); Folke, Persson and Rickne (2016)). It is not uncommon for mayors and other candidates in the municipal council to switch parties from one election to the next.[21] For instance, in the municipal elections of the year 2000, 62 percent of the reruning candidates switched party.

Although the political party system revolves around the candidate, there are three main reasons to believe that parties may be recognized as corrupt if their candidates misbehave.[22] First of all, in order to be eligible, all candidates must be supported by a party, making the link between politician and party relevant.[23] Second, when citizens cast their vote, the party of the candidates is quite visible on the ballot. Indeed, the ballot first gives information about the party (and the party number) and then the list of candidates for each party, members of the municipal council and mayor.[24] An illustrative example is shown in Figure 3.

[Figure 3 Here]

Third, candidates may be non-eligible for election either due to term limits or due to other impediments. Parties, on the other hand, have a longer life span than candidates.

[20] The local level does usually have the highest fragmentation. The multitude of parties frequently requires the parties to form coalitions, at any level of the government, as no party has any chance of gaining power alone. The biggest players in local elections are: PTB, PT, PMDB, PSDB, PP and, PFL, respectively standing for Partido Trabalhista Brasileiro, Partido dos Trabalhadores, Partido do Movimiento Democrâtico Brasileiro, Partido da Social Democracia Brasileira, Partido Progressista, Partido da Frente Liberal, now called Democratas.

[21] According to Desposato (2006), the strongest motives for a party change are: access to distributive resources, electoral opportunities, and compatible policy positions.

[22] This has been the case even during recent events in Brazilian politics, where citizens were linking most PT politicians to corruption due to recent political scandals in the country.

[23] They must have been members of the party for at least two years before elections.

[24] The campaign for the candidates is mostly done with party numbers.

2.4 Local Media

As in many other countries in the world, television and radio are the most used media outlets. Television has the largest penetration on a national scale. However, in the year 2000, only 10 percent of the municipalities broadcasted local channels, whereas 34 percent of the municipalities had a local radio station according to Perfil dos Munícipios Brasileiros (2001). According to *Latinobarómetro*, a public-opinion survey, 40 percent of the Brazilian population uses radio as a source of political information. I use information on FM radio coverage to define neighborhoods for audited municipalities.

The use of local radio is motivated by the fact that it is more likely for these outlets to report local news. It is quite rare for local news to be mentioned extensively on state or national TV. The focus on FM radios mainly arises because of reception quality. While Brazil still uses many AM transmitters, which travel farther distances, audio quality is often poor whereas FM broadcasting is capable of a better sound quality.

I do not take into account newspapers as Brazil has one of the lowest levels of newspaper penetration in the world, with only 42 newspaper copies per 1000 inhabitants (Porto (1985); IBOPE (2012)). Community radio stations are also excluded from the analysis. While they certainly constitute another source of local information in the Brazilian municipalities, most stations have low power transmitters with a maximum broadcast range of one kilometre (Boas and Hidalgo (2011)).[25]

3 Data Sources

I collect and combine data from different sources. Three most important sources used are: (i) audit data from which I build a measure of severity of irregularities found by the auditors, (ii) electoral data, and (iii) a measure of antenna coverage. These data are described in detail in this section.

3.1 Audit Reports

The audit reports consist of more than 2000 PDF documents downloaded from the Controladoria Geral da Uniaõ (CGU) - for the period 2003-2013. Each audit starts with a description of the purpose of the auditors in the municipality, the number of days the auditors spent in

[25]Hence, it would not be possible to construct informational neighborhoods, as their signals are unlikely to cross municipal boundaries.

that town and the funds investigated as well as all irregularities found. Although the structure of the reports has been altered from time to time, some features remain similar across lotteries. Within the first five pages, a table, or list, is included, the total amount of funds investigated and the number of programs in each area. From lottery 8 onwards, a summary of irregularities was also introduced in all reports (see Figure 5).[26] For each inspection, there is a detailed description of the findings in the text. Several pictures are also included in the report to further illustrate the wrongdoing.[27]

[Figure 5 Here]

For example, in Figure 6 a simulation of a bidding procedure is documented by the auditors in the municipality of Alvaraes in the state of Amazonas (lottery 2). The call for bids was for cleaning products, while all three bids submitted intended to deliver food.[28] Except from the bid layout being (exactly) the same, the bids also coincide also on the type of the orthographic mistakes. The only slight change the auditors point out is the different use of font.

[Figure 6 Here]

Another fraud in a procurement process is the example in the municipality of Afonso Cunha in Maranhão (lottery 27) where three firms "participated" in the call for bids. The auditors visited the locations of the address registered as the firm domicile and none of them were located as indicated in the bids.

[Figure 7 Here]

In the text, for each irregularity found is also included a statement by the mayor (if provided). There is also a further comment by the auditors in which they explicitly write if they accept the justifications or if they maintain the original statement. Most of the original statements remain unaltered.[29]

[26]The summary in lotteries 34 to 38 is somewhat different because the auditors include a description paragraph instead of the irregularity list.

[27]There is a high variation in the type of irregularity documented such as roads not constructed, abandoned construction sites to poor conditions in the schools and kinder-gardens, non safe storage of food and medicine.

[28]For the validity of the procurement process, the law (Law number 8.666/93) requires the submission of at least three valid bids from different suppliers.

[29]Only a small fraction of the justifications are (partially) accepted.

3.1.1 Text Classification

From the audit reports, it is possible to have municipalities ranked according to the severity of the committed irregularities. Processing more than 2000 reports (a total of 172,768 pages) is challenging for two main reasons: the first, easier to overcome, is that the information in these documents is not stored in a homogeneous format, not only across lotteries but also within the same lottery in the same state. The second is that the amount of training data available – data from which the algorithm should learn – is very limited. When using a machine learning classifier (in supervised learning), it is extremely important to have access to a sufficiently large amount of labeled data for the algorithm to learn so that accurate out-of-sample predictions are made.[30]

Taking advantage of the summary that the auditors provide from the 8-th round of the lottery, I construct a classifier that is able to recognize for each element in the summary if an irregularity found is more severe than others on the same list. To automatise the process (of classifying a total of 104,337 phrases), I created a list of rules to distinguish more severe from less severe cases. Hand-written rules are recognized and extensively used in the machine learning literature in cases where the amount of classified documents available is limited (see Manning, Raghavan and Schütze (2010)). The accuracy of these models can be high.[31]

The classifier works in the following way: given a set of inputs (phrases from a document), call it S, the purpose is to classify each input of the two classes, $C = \{c_1, c_2\}$, indicating whether the irregularity is severe or not. The predicted output would be $c \in C$. To get enough knowledge about the domain of the data with the help of a Brazilian research assistant, I read and classified around 200 reports. The classification was used to test the performance of the classifier and compile the list of rules. Lotteries from 20 - 33 were also used – in the cases where the auditors pointed out the majors irregularities – to build the list of rules for classification.

The list of rules consists of stemmed words, or composition of words, bigrams and other n-grams. All accents are removed from the original documents, and no accents are included in the list of rules. This ensures that if a word is miswritten, i.e missing accents, the algorithm is still able to recognize that word. Upper case letters have also been replaced by lower case

[30] Labeled data is jargon in Machine Learning literature to indicate classified documents by a human. Good practice of the procedure requires that the subset of documents for human classification is selected from a random sample.

[31] Examples are provided in Jacobs and Rau (1990) who report 92 percent precision and a 88.5 percent recall rate while Hayes and Weinstein (1990) report 94 percent recall and 84 percent precision.

letters.

Words like "Fraud", "Collusion", "Fake", which indicate a severe irregularity themselves, are included as simple words. In other cases, I search for combinations of words.[32] Given that in some instances the meaning of the sentence is the same but the wording is different, such as "Firm not located", "Non existing firm", "Ghost Firm" (for an example of a non-located/non-existing firm, see Section 3.1) I ensure these are included as synonyms in the list of rules.[33]

Below, I provide an illustration of the approach taken. Suppose, for example, that the summary written by the auditors for municipality X is the following:

1. Indication of fraud in the procurement process.
2. Resources (from a specific program) used differently than what was foreseen by the program.
3. Payments for non executed services.
4. Non actualisation of pupil's cadastral information.
5. Delay of book delivery in a rural school in the municipality.

Out of five irregularities found by the auditors, three are more severe. For each summary (in each document), I record the total of amount of the most severe irregularities (three in this example).[34] This measure is a count of severe irregularities found in the municipality.[35] The list of rules seems to be quite accurate as it also overlaps with the overview that the auditors have of the most corrupt activities in Brazilian municipalities. An example of this can be found in the final report for the year 2004, where the auditors summarize their annual work.

> Indicators of fraud in the procurement process, availing of ghost firms; non existing or fake bills; overpricing; lack of proof of expenditures; manipulation of

[32] For instance, when a good was supposed to be provided but could not be located by the auditors, I search for "good not loc". The algorithm, in this case, will capture if the auditors have written the phrase in plural form such as "goods not located".

[33] The list of rules has a total of 140 n-grams.

[34] In this simple case, the vocabulary would be: fraud; resource differ, payment non execut.

[35] Classifying the reports as either corrupt or not corrupt is very restrictive. In the random sample that the RA classified, only 7 percent of the municipalities had no instances of corruption.

the procurement process, paralysed construction, non existing object and non reached objectives.[36]

3.1.2 Principal Component

The irregularity count described above may be subject to measurement error because the classifier may fail to recognize a severe irregularity. For example, if the auditors indicate that some goods are not located in the municipality, they write "Bens nao localizados", the regular expression, composed of stemmed words, to classify this phrase would be "ben nao local".[37] This regular expression will also capture a phrase like "beneficiarios nao localizados", indicating that the beneficiaries have not been located in the municipality, which is not a severe irregularity in itself.[38]

To ensure that the irregularity count indeed measures what it should, I create another index, which is the first principal component of a series of variables extracted from the audit reports. The variables are: the number of pages, the number of lines, the number of images included in the PDF, the number of total irregularities – the count of all irregularities summarized for each report.[39] The main idea behind this is that the *latent* common factor will express a dimension of the underlying corruption.

Some preprocessing is needed to use the PCA, i.e the standardisation of the variables. This is done so that variables have comparable scale values and allows us to deal with cases in which variables have different scaling. The first principal component and the severe irregularity count are highly correlated, the correlation coefficient of 0.86. The correlation between both measures is shown in Panel A of Figure 8.

3.1.3 Corruption index

The corruption index, which I use for the main analysis, is the first principal component of all the variables described above (in Section 3.1.2) and the measure of the severe irregularity

[36] Available here: http://www.cgu.gov.br/sobre/auditorias/arquivos/2004-a-2001/relatorio_gestao_cgu_2004.pdf.

[37] It is a sequence of characters that define a search pattern, mainly for use in pattern matching with strings, or string matching.

[38] Usually, the beneficiaries are not found because they have moved to another municipality or because they may have passed away etc. There are some cases when beneficiaries of certain funds were not located and members of the municipal council were making use of these funds. In the latter case, it would make sense to classify this irregularity as more severe.

[39] Also in this case, I discuss in the Appendix how I get the total number of irregularities in the first rounds of the lottery (lottery 3-7).

count (see Section 3.1.1). The distribution of the measure is shown in Panel B of Figure 8.

[Figure 8 Here]

The first principal component of these series is the only eigenvalue with a value larger than 1 and it retains around 80 percent of the explained variation. A visualisation of the variance explained by each eigenvalue and the cumulative variance is shown in Figure 9.

[Figure 9 Here]

The combination all variables that indicate irregularities is done with the main objective to minimize classification error. This measure is dependent on the set of rules I use to construct the irregularity count; thus, it may considered to be partially subjective. However, it has two main advantages: i) it is replicable and, ii) once the rules are set, no subjective judgment is involved during the classification process, thus making it time efficient and less prone to human mistakes.

I test the performance of the corruption index in two different ways.[40] First, I use the sub-sample of municipalities in which the auditors indicate the main problems found in each town. The classifier seems to be scoring correctly among couples of municipalities to which the auditors attribute the same number of irregularities. The reason why the classifier is able to rank among municipalities to which the auditors have scored similarly, for each report, the auditors rank the major irregularities within that same municipality, thus making it non comparable to another municipality.[41] To further clarify this point, consider the case of audited municipalities Andarai in the state of Bahia (lottery 21) and Guatapara in the state of São Paulo (lottery 22). The auditors find 35 irregularities in the first case and 27 irregularities in the second. They point at 3 major irregularities and 3 administrative irregularities in both cases. However, according to human coding, Andarai is more corrupt than Guatapara.[42] The corruption index is able to capture this difference.[43]

[40]These are two tests with external coding. I have also tested my measure with the measure the RA and I created. Indeed, the list of rules was updated until I reached a high correlation with our own hand-coded measure.

[41]For this reason and due to the fact that the coding of the auditors is for a limited sub-sample of the data, I use my measure rather than the coding of the auditors. My sample spans from lottery 3 to 38, making it preferable to have a uniform coding throughout the sample period.

[42]The language used by the auditors is also harsher.

[43]Other examples are municipalities audited in the same lottery (round 21) but in different states such as Gameleira in the state of Pernambuco and Bom Jesus do Araguaia in Mato Grosso where for each municipality, there are 11 administrative irregularities where the first is again more corrupt than the second municipality

I also test my corruption measure versus hand-coded measures in the literature. More on this test can be found in Section 10.4 in the Appendix.

3.2 Electoral Data

Electoral data are obtained from Tribunal Superior Eleitoral (TSE), for municipal elections in 1996, 2000, 2004, 2008 and 2012 and municipal-level outcomes of presidential elections in 1998, 2002, 2006. The vote share is defined by TSE as the number of valid votes scored by the party over the total of valid votes in the municipality. A vote is considered valid if it can be attributed to a candidate (or party). For all elections, the change in the parties' vote share over time can be measured, unless the party participates in one election but not in the other.

I match the political party in the municipality from one time period to the next, in this way I can follow the evolution of the votes of the party over time. A political party might participate in one electoral race without necessarily having participated in the previous. In municipal elections, 43 percent of the parties running in 1996 rerun in the 2000 elections and 46 percent of the parties in the 2000 elections rerun in the 2004 elections. Instead, 59 percent of the incumbent parties run for reelections from 1996 to 2000 and 57 percent show up for reelection from 2000 to 2004.

3.3 Antenna Coverages

Geocoded information on FM antenna locations and signal coverage maps are obtained from FMSCAN.[44] Coverage data are in the form of a map raster or a grid file indicating cells where the signal strength is such that it can be received by a radio. A map of the antenna coverage is shown in Figure 4.

[Figure 4 Here]

I convert the raster file into multiple polygons. Each polygon contains information about a single antenna location and coverage. Next, I intersect each of the files containing antenna

in the pair. Other examples are couples such as: Braga in Rio Grande do Sul (lottery 20) and Meleiro in Santa Catarina audited in lottery 23, Feira Nova do Maranhão in Maranhão and Sao Valerio da Natividade in the state of Tocantins lottery 21.

[44] Information on any country antenna location on their webpage: fmscan.org. I thank Maria Petrova and Ruben Enikolopov for this reference. I cross-check antenna locations with the locations on antennas provided by the National Agency of Telecommunications in Brazil (*Agência Nacional de Telecomunicações*, (ANATEL). Additional data on media presence at the municipal level from 1999 to 2009, are obtained from Perfil dos Munícipios Brasileiros.

information (roughly 6000 antennas) with a file containing information of the population density, for each municipality, at each 20 square kilometres. In this way, I am able to compute the fraction of municipal population and the area that each antenna covers.

To compute the population density at each 20 square kilometres, I retrieve population density estimates of the year 2000 from the Center for International Earth Science Information Network CIESIN-Columbia University (2005). They come in the form of 2.5 arc-minute grid cells. To the gridded population, I overlay a map of Brazil's municipal borders which I download from the IBGE website. In this way, I can identify which points on the grid that are reached by a radio signal and I can compute exactly the population density covered by each FM antenna.[45]

3.4 Other Controls

To obtain information about socio-economic and demographic characteristics, I combine data from two main sources, namely Instituto Brasileiro de Geografia e Estatística (IBGE) and Instituto de Pesquisa Econômica Aplicada (IPEA). The population census carried out in Brazil in 2000 provides several socio-economic and demographic characteristics used as controls in the regressions. Some of these variables are per capita income, income inequality, population density, share of the population that lives in urban areas, and literacy rate. Yearly population and municipal GDP estimates are retrieved from IBGE. For the rest of the variables, I use census variables and intercensal estimates. Spatial variables such as shapefiles, latitude, longitude of the population centroid are also retrieved from IBGE. Using the above, I am able to compute the distance to the nearest state capital, the distance to the federal capital and the vector of contiguous neighbors for each municipality.

3.4.1 Summary Statistics

Table 2 presents some summary statistics of municipal characteristics measured pre-treatment (in year 2000). The municipalities in this sample have a smaller population when compared to the Brazilian average (30,833 inhabitants) also have slightly larger areas (the Brazilian mean municipal area is 1529 km^2). The average vote share for the party winning the municipal elections also winning the first round is 55.22 percent. For municipalities where a second

[45] All geographic computations in this paper are carried out using ArcGIS. I exclude non-local radios and radios for which the coverage goes beyond their contiguous neighbors. The rationale for the latter is that usually antennas covering more municipalities than immediate neighbors are not considered local radios.

round is included, I use the first round data.

[Table 2 Here]

The average number of antennas in the sample is similar to the country average, while the number of contiguous neighbors is smaller, with a country average of roughly six. Other characteristics such as male population, literacy rates, illumination, electricity, Human Development Index (HDI) and income per capita are similar to the country average. The political variables are indicators taking value one if the local government, in the electoral term of 2001, was headed by a mayor from that party and zero otherwise.

4 Benchmark Case

I exploit variation in the timing of audits to document the effect of the audit in the audited municipality across levels of government. For each electoral term, I compare audited municipalities before elections with audited municipalities after elections.[46]

Municipal Elections – To study the effect of the audits on the audited municipalities, I define municipalities audited before municipal elections as the treatment group. The control group consists of municipalities audited within the first eight months of the new administration. Therefore, the control group is composed of municipalities that are: (a) audited during the same electoral term as the municipalities audited before elections, or (b) the handling of funds mainly corresponds to the previous administration. The outcomes of interest in this case are the difference in the shares of votes and a dummy variable taking value one if the incumbent audited party is reelected in municipal elections. I exclude municipalities where the party wins with a vote share above 90 percent. This implies that less than 4 percent of the total sample is removed.[47]

Presidential Elections – When turning to the presidential elections, the question of interest becomes: How does an audit at the local level affect the electoral outcomes of the party in the presidential elections? To answer this question, I focus on local parties which

[46]The same empirical strategy has been used also by Ferraz and Finan (2008); Larreguy, Marshall and Snyder Jr (2014); Bobonis, Cámara Fuertes and Schwabe (2016)

[47]If I include these municipalities, the point estimates remain similar but specifications where the share of votes is the dependent variable are noisier.

have a national counterpart. Therefore, the attention is restricted to the two main players in national politics, namely PT and PSDB. The reason is that the rest of the parties support the presidential candidates of PT and PSDB or enter in coalition with other parties.[48] The treatment group is composed of audited municipalities ruled either by PT or PSDB in the period before presidential elections and the control group is composed of municipalities audited after the presidential elections.[49] The dependent variable is the difference in vote share. In the presidential elections, I only use one cross section – that of the municipalities audited before and after the elections of 2006. The first differences for this cross section are constructed by using the electoral outcomes of the year 2002.

4.1 Test of Balance

Before proceeding with the main analysis, I perform several randomization checks despite the randomization carried out by the CGU. I check if treatment and control groups differ significantly on observable characteristics by mean comparison for both samples. In Table 1, the groups show no significant difference on most characteristics. However, there are two main differences that require particular attention: the corruption index and the PT dummy.

The first emerges due to the difference in constructing the corruption index which concerns rounds 3-7 (see Section 10.3 in the Appendix). However, there are no significant differences when I exclude those lotteries. In order to maximise the sample size, I include these lotteries in the main sample for the analysis. In the Appendix, I also show that the main results in the paper remain unaltered, excluding these lottery rounds.

Another non-balanced characteristic is the indicator variable PT. The test for equality of means for these coefficients is rejected at conventional significance levels (see the corresponding column in Table 1), indicating that mayors from PT are more likely to be audited before municipal elections and less likely to be audited before presidential elections. This would be a source of concern if the audit was aimed to harm PT's municipal elections and favour them during presidential elections. However, this scenario seems unlikely as it would make more sense for the party that has more control over the CGU throughout the sample period, to consistently refrain from harming itself before elections and attempting to harm its opponents in all types of elections.

[48] In the 2002 and 2006 elections, respectively, 6 and 7 coalitions took place in municipal elections.

[49] The treatment group is the set of municipalities audited in the period April 2003-July 2006 and the control group consists of all audited municipalities in the period from May 2007 to the end of the year 2011. These results are robust to varying the time span of the control group.

To address these potential concerns, I downloaded all publications regarding the random audits in the *Official Gazette*. This information is also available on the CGU website. In this way, we have the list of eligible municipalities and the numbers allocated to each municipality. As shown in Figure 10, the allocation of numbers, from 0 to 9, across lotteries from 3-31 mimics a uniform distribution. Lottery 7 is missing because the set of numbers attributed is missing in the corresponding publication.[50] This is a useful test in case one might be worried that numbers are assigned in such a way that it would increase the chances of selecting some municipalities and reducing the chances of selecting others.

[Figure 10 Here]

Moreover, I link these data to the audit data and for each lottery, I am able to compute the fraction of selected numbers per lottery over the total of times that a specific number is allocated to a municipality. These fractions are depicted in Figure 11. Even in this case, the distributions seem to be uniform across lotteries. This piece of evidence may be useful to attenuate possible concerns about a potential randomization failure.

Third, I run a linear probability model to estimate whether a municipality headed by PT is more or less likely to be audited. I also restrict the sample to municipalities where mayors are either from PT or PSDB. Results seem to support the fairness of the lottery. This procedure is discussed in detail in the Appendix.

[Figure 11 Here]

Fourth, the literature has established (and treats) these audits as random.[51]. In light of the above evidence, I have no reason to believe that the randomization was manipulated and used for political purposes. Anyhow, to deal with imbalances on the party dummies, I include them in all regressions throughout the paper. For the rest of the characteristics, I present specifications with and without controls. Results are similar across specifications.

4.2 Empirical Strategy: Benchmark Case

Given that more than 2000 audits have occurred throughout the sample period, I am able to use both cross section and time variation in my analysis. The estimating equation is expressed in first differences and can be written as:

[50] As in all publications, the number of municipalities per state and those that are non-eligible because of a previous type of audit (random or not) are clearly stated.

[51] Examples regarding this consensus in the literature can be found in Ferraz and Finan (2008, 2011); Brollo et al. (2013); Litschig and Zamboni (2011); Silva (2010)

$$\Delta E_{pit} = \alpha + \beta_0 \text{Audit}_{pit-s} + \nu_s + X_i^T \delta + \eta_p + \tau_t + \varepsilon_{pit} \quad (1)$$

where ΔE_{pi} is the change in the electoral outcome of interest for party p in municipality i in elections, local or national. Audit_{pit-s} is also an indicator variable. It takes value one if the municipality i with mayor from party p was audited some time s before elections and value zero if audited after elections. Note that the model is estimated in first differences, and the audit dummy can also be written as $\Delta\text{Audit}_{pit-s}$ given that at time $t-1$ the municipality had no audit. Estimating the model in first differences allows me to take into account time-invariant characteristics in the municipality. ν_s is a set of state dummies needed due to the fact that the lottery randomization is stratified at the state level. X_i is a set of controls as predetermined to the audit, measured in the year 2000. η_p are the party fixed effects, τ_t are time fixed effects and ε_{pit} is the idiosyncratic error term. The identifying assumption requires that once it is conditioned for state fixed and time effects, Audit_{pit-s} is uncorrelated to the error term. Conditioning for time effects is necessary due to multiple time periods pooled together.

β_0 is the causal effect of the audit on the electoral outcomes of municipality i. We expect the audit to have no effect, in case voters have correct priors about the party. If individuals interpret the fact that the auditors show up in the municipality as a signal of a wrongdoing occurring, a negative effect of the audit on electoral outcomes of the party is expected.

A point worth stressing is that the unit of observation in my analysis is the party in the municipality. I follow the evolution of the electoral outcome of the incumbent party over time. A further note regards the feature of the data: In municipal elections, given that few parties show up for more than two consecutive terms, I cannot fully exploit the longitudinal feature of the data. The same applies to municipal elections in case of geographical spillovers. The design is simpler for presidential elections (audited municipalities and neighboring municipalities). In presidential elections, both parties show up in both periods, thus allowing me to take first differences.

In municipal elections, the focus is on incumbent political parties which rerun in the upcoming elections. To ensure that parties audited before elections do not differentially exit, I estimate a linear probability model of the exit dummy on the treatment dummy.[52] Results shown in Table A7 in the Appendix, suggest that there are no differences in exits of incumbent

[52] Party exits, although they might be a consequence of the policy, would cause selection in the sample.

parties in municipalities.[53] This exercise is not necessary for presidential elections, because both parties are always present throughout the duration of the sample period.

To understand the role of corruption of the politician on the party's electoral outcome, I interact the main regressor with the corruption index, explained in Section 3.1.3. The equation of interest becomes:

$$\Delta E_{pit} = \alpha + \beta_0 \text{Audit}_{pit-j} + \beta_1 \text{I}_{cit-1} + \beta_2 \text{Audit}_{pit-j} \times \text{I}_{cit-1} + \nu_s + X_i^T \delta + \tau_t + \eta_p + \varepsilon_{pit} \quad (2)$$

where I_{cit-1} is the corruption index in municipality i committed by the incumbent mayor, from party p and his administration. The parameter of interest is β_2 which measures the effect of the audit conditional on the levels of irregularities found in the municipality. The standard errors are clustered at the same level where the treatment occurs, namely at the municipality level.

4.3 Results: Benchmark Case

In panel A of Table 3, results from Equation 1 are shown. Columns (1) and (2) show a linear probability model where I regress the dummy for reelection of the party on the audit dummy in municipality i. In columns (3) and (4) and columns (5) and (6), the outcome variable is the change in vote shares of the party in municipal and presidential elections.

[Table 3 Here]

It appears that if the incumbent administration, headed by the mayor, is audited before elections, the mayor's party is less likely to be reelected as the ruling party in that municipality. The support for the party, expressed as the share of valid votes in the municipality, is also reduced. The coefficient remains similar in size and magnitude with he inclusion of different controls (shown in even numbered columns). Parties associated with an incumbent politician audited before elections are 9 percent less likely to be reelected as the ruling party and they obtain 2 percentage points lower vote shares than parties associated with politicians audited after elections. The elections under consideration are close elections.[54]

[53] I estimate a linear probability model on the likelihood of exit of the political party from the municipality. The coefficient in this case is not significant suggesting the lack of selection of parties into elections. Parties audited before elections are not more likely to exit from the electoral race. The coefficient goes in the opposite direction and is not statistically significant.

[54] Indeed, 70 percent of the sample are composed of parties that, in the first round, received a vote share between 40 and 60 percent of the valid votes in the municipality.

In presidential elections, point estimates suggest different magnitudes when compared to municipal elections. In municipalities where the audited mayor – before elections – is either from PSDB or PT, the vote share obtained by this party in the presidential elections is lower than in similar municipalities audited after elections. The size of the coefficients lies around 4 percentage points. The effect is both statistically and economically significant.[55]

These result imply that audits in local governments do not only have consequences for the electoral outcomes of the incumbent party in local elections but also on higher levels of government, such as the presidential elections.

In Panel B of the same table, results from Equation 2 are shown. Columns (1)-(4), correspond to outcomes in municipal elections. In columns (1) and (3), I interact the policy dummy with the corruption index. In columns (2) and (4), I present a more flexible functional form where the policy dummy is interacted with terciles of the corruption index, allowing the slope to vary for each tercile. The interaction terms are not significant in any of the specifications, indicating that the effect of the pre-electoral audit in municipal elections does not depend on corruption levels. The F-stat on the joint significance of the main effects and the interaction term, except from column (4), indicates that interaction terms are jointly significant.

Coefficient estimates for presidential elections are shown in columns (5) and (6). In column (5), the corruption index enters the equation linearly. The coefficient on the interaction term in column (5) is statistically significant. The F-stat on the interaction terms fails to reject the null. In column (6), the treatment dummy is interacted with terciles of the corruption index. The coefficient on the third tercile indicates that moving from the lower tercile, omitted category, to the upper tercile, the votes of the incumbent party are reduced by 9 percentage points. The reduction in votes seems to be statistically and economically significant.

Taken together, these results indicate that: i) electoral accountability does also operate through party identities, and ii) that this effect spills over to different levels of government. The latter has important implications for the role of parties in selecting and monitoring their candidates. Apparently, individual malfeasance has strong repercussions on the party.

The findings in municipal elections may seem puzzling when compared to those by Ferraz and Finan (2008). In their case, the exercise leads to negative but insignificant coefficients

[55]I have attempted to estimate the presidential outcomes on the sub-sample of municipalities (with mayors from PT and PSDB). I do not have much power for this specification. However, the coefficient remains similar in magnitude. Moreover, I add controls for the timing of municipal elections. Results remain unaltered.

(in most specifications) for the audit dummy on electoral outcomes. They find electoral punishment of the candidate, depending on the disclosed corruption levels. It should be kept in mind that there are some key differences between their work and mine: First, they focus on the incumbent candidate, the most important figure in Brazilian politics. I instead focus on the party, less important at the political stage (60 percent of the candidates in the elections of 1996 showed up in those of the year 2000 with a different party label). Second, they focus on all eligible candidates, not only those rerunning for elections, imputing a zero to the reelection dummy (the main regressor) if the candidate does not show up for reelection. I follow the parties that rerun for election, either with the same politician or with another candidate in case the politician does not rerun or is non-eligible. Third, they only use one cross section of municipalities whereas I extend the sample for three different cross-sections. As a robustness check, I also show results using only one electoral term, meaning the same time period where they base their analysis. In doing so, only 1/2 of their sample coincides with mine.

5 Geographical Spillovers

To understand whether and how the electoral effects of audits spill over on neighboring municipalities, a definition of a relevant neighborhood is required.[56] The media literature (DellaVigna et al. (2014); Enikolopov, Petrova and Zhuravskaya (2010); Yanagizawa-Drott (2014); Bursztyn and Cantoni (2015)) suggests that radio and TV play an important role in information transmission to citizens; hence, I use information on FM transmitters to define the treated areas, namely informed neighbors.

I define as informed neighbor a municipality that shares a border or a vertex with the audited municipality and receives antenna coverage from the latter. In this way, I am able to compare informed neighbors of an audited municipality before elections with informed neighbors of an audited municipality after elections. This empirical strategy is similar in spirit to the one used in the benchmark case.[57] Given the randomization in audits, both groups should be homogeneous.

In the case of local elections, I study the electoral effect of an audit on the audited party

[56] Several reasonable candidates come to mind, for example: municipalities within a certain distance of the audited town, municipalities within the same region, trading partners, municipalities sharing common administrations such as courts, etc.

[57] I exclude neighbors which were audited themselves because I want to understand the impact of the audit on the neighbors if they were not audited.

in neighboring municipalities. The sample consists of informed neighbors that at the time of the audit have had a positive vote share of the party. Note that in this case, the party is not necessarily the party of the incumbent politician. When turning to the effects on the presidential elections, the sample is restricted to all informed neighbors of an audited municipality – with a head of government from either PT or PSDB. The reason is that in presidential elections, both major parties have a positive vote share throughout the country.

In order to avoid that the same municipality shows up in a treatment and a control group simultaneously – in municipal and presidential elections – I proceed in the following way: if a municipality is the neighbor of two different audited municipalities, I allocate the municipality to either treatment or control as a function of the antenna coverage the municipality receives from the audited neighbor. The municipality is assigned to the treatment group if it receives more coverage from a town audited before elections and in the control group if it receives more coverage from a town audited after elections. If a municipality receives equal coverage by two different towns, one before elections and the other after, the municipality is allocated to the treatment group as it is supposed to learn about the audit before elections.

To ensure that the groups are indeed comparable, I perform randomization checks discussed in Section 10.5 in the Appendix.

5.1 Empirical Strategy: Geographical Spillovers

Similar to the empirical strategy used in the Benchmark Case (see Section 4.2), the estimating equation is the following:

$$\Delta E_{pjt} = \alpha_1 + \gamma_0 \text{Audit}_{pit-s} + \nu_s + X_j^T \delta + \tau_t + \eta_p + \varepsilon_{pjt} \tag{3}$$

In this case, i is the audited municipality and j corresponds to the informed neighbor. E_{pjt} is the difference in the vote share in the municipality j, Audit_{pit-j} is an indicator taking value one in case there was an audit in a nearby municipality i at time s previous to elections and value zero in case the audit in town i occurred after elections.

All remaining variable definitions are as in Equation 1. γ_0 captures the average impact of the audit in a nearby town i, with incumbent politician from party p, on the votes of that party in the closest upcoming elections in the informed town j.

When testing for the differential effect of the audit – conditioning on levels of corruption in the audited municipality – the above equation becomes:

$$\Delta E_{pjt} = \alpha_1 + \gamma_0 \text{Audit}_{pit-s} + \gamma_1 \text{I}_{cit-1} + \gamma_2 \text{Audit}_{pit-s} \times \text{I}_{cit-1} + \nu_s + X_j^T \delta + \tau_t + \eta_p + \varepsilon_{pj} \quad (4)$$

where I_{cit-1} is the level of irregularities committed by the municipal administration in the nearby municipality. The interest lies in the interaction term, γ_2 which captures the effect of the audit in the nearby municipality conditioning on the corruption level found in that municipality.

When estimating Equations (3) and (4), standard errors are clustered at the Mesoregion level. Mesoregion is a subdivision of the Brazilian states, by IBGE, which clusters together municipalities by geographic proximity and common characteristics. Brazilian states are subdivided into a finer level, also constructed for statistical purposes, namely Microregions.[58] Although Micro and Mesoregions were created for statistical purposes and do not constitute an administrative area, they are frequently mentioned in the audit reports.[59]

To check on the robustness of the inference, I also cluster standard errors at the Microregion level and also correct for spatial dependence. A further check is carried allowing for spatial dependence among neighboring municipalities 100 kilometres apart, following Conley (1999) with a linear decaying in the spatial weighting (a la Bartlett).[60]

5.2 Results: Geographical Spillovers

Results for the geographical spillovers are presented in Table 4. Panel A presents results from Equation (3). Columns (1) and (2) contain the electoral outcomes for municipal elections whereas (3) and (4) contain those for the presidential elections.

[Table 4 Here]

In local elections, there is no effect of the audit in the neighboring town on the votes of that party on the informed neighbors. The coefficient γ_0 is not statistically significant, with and without inclusion of controls. In presidential elections, the coefficient is statistically significant and negative. If there is an audit in a nearby town, the party loses votes in

[58] See Table A2 for more on the subdivision of the Brazilian states.

[59] Some municipalities within the same Mesoregion have joint federal programs. Auditors also mention municipalities in the micro or Mesoregion to compare prices among municipalities or to mention that some beneficiaries of different funds, such as Bolsa Escola, have moved to municipalities within the same e.g. Micro/Mesoregion.

[60] I use the version of the STATA command by Hsiang (2010). Results available upon request.

presidential elections. The behaviour of the voters in neighboring municipalities is similar to the behaviour of the voters in the audited municipality itself. However, there is a small difference with the coefficient estimates in the audited municipality. The coefficient on β_1 is larger than γ_1, the corresponding estimate in the case of geographical spillovers. A t-statistics on the equality rejects the null hypothesis that these coefficients have the same magnitude.

In Panel B, variations from Equation (4) are shown. The coefficients of interest are the interaction terms, either on the linear specification columns (1) and (3) or with a more flexible functional form, columns (2) and (4). For municipal elections, in both types of specification, the corruption index enters the equation with a negative sign. Moving from the first to the second tercile of the distribution of the corruption index decreases the votes of the party by roughly 8.3 percentage points while moving from the first to the third by 9 percentage points. In the presidential sample, moving from the first to the third tercile, the vote share is reduced by 9.8 percentage points. The coefficient of interest, in these separate regressions, is negative and statistically significant, indicating that for higher levels of the corruption index the political party is punished harsher in nearby municipalities.

In municipal elections, the findings suggest that voters do not change behaviour once an investigation in the nearby town occurs. Indeed, in nearby municipalities the incumbent mayor is not audited. The votes under considerations are those of any party which has a neighbor whose politician was audited before elections. However, the coefficient on the interaction term indicates that voters do learn about party characteristics as they became more salient.

Turning to cross-border and cross-electoral spillovers, a similar argument holds. Here again the results confirm the similar behaviour of citizens in neighboring municipalities with those in the audited municipality itself, highlighting that when voting for higher levels of government, areas sharing the same information network behave similarly.

From the sample size, it seems that on average, for each audited town, two neighboring municipalities are affected. A back-of-the-envelope calculation suggests that in the audited neighborhood (audited municipality and the informed neighbors), the vote share of the audited party goes down by 10 percentage points. This is a sizeable effect both politically and economically.

6 Are party labels recognized?

To understand whether party names are indeed recognized, I take advantage of coalition formation in national elections. I test if an audited party in municipal elections harms the electoral prospect of the main party in the coalition. If voters recognize party labels, no effect on the coalition is expected. If the party is linked to candidates with a high enough amount of disclosed corruption, it might as well be the case that voters vote less for the coalition.

To answer this question, I use as the definition coalition parties that have been in the same coalition with either PT or PSDB in the previous presidential term.[61] The estimating equation is similar to the above, the only difference is the main regressor.

$$\Delta E_{pit} = \alpha_0 + \beta_0 \text{Audit}_{rit} + \nu_s + X_i^T \delta + \eta_p + \varepsilon_{pit} \tag{5}$$

where Audit_{rit} is an indicator taking value one if party r in coalition with party p is audited before elections and zero if audited after elections. Results are shown in Table 5. Columns (1)-(3) present results for the baseline regression with progressive addition of controls. In columns (4)-(6) the policy dummy is again interacted with the irregularity index. In column (7) and (6) I interact the policy dummy with the quadratic term of the irregularity index.

[Table 5 Here]

Notice that columns (1) and (2) indicate that the audit of one of the parties on the coalition seem to have no effect on the votes of the main party in the coalition. This also seems to be the case when I interact the investigation dummy with the corruption index or with different terciles of the corruption index. These findings suggest that indeed party names seem to convey the relevant information to the voters.[62]

6.1 Heterogeneous treatment

In the case voters truly react to news about corruption, it is important to understand if voters change behaviour relative to the timing of information provision. To answer this question, I

[61] In 2002, parties that were in coalition with PT are PC do B, PCB, PL and PMN. Instead, the only party in coalition with PSDB was PMDB. In this exercise, I do not consider areas where the mayor was from PCB as it was part of another coalition in the 2006 elections. These parties are more or less the main coalition parties throughout the sample period under consideration.

[62] Results, not shown, are very similar in the neighboring municipalities.

allow for the treatment to be heterogeneous with regards to the timing of the audit before presidential elections. The estimating equation is stated below:

$$\Delta E_{pit} = \alpha + \beta_0 \text{Audit}_{piy-1} + \beta_1 \text{Audit}_{piy-2} + \beta_2 \text{Audit}_{piy-3} + \nu_s + X_i^T \delta + \eta_p + \varepsilon_{pit} \quad (6)$$

Audit_{piy-1} is an indicator variable taking value one in case the investigation occurred one year before presidential elections and zero otherwise. Audit_{piy-2} takes value one in case the audit happened two years before elections and Audit_{piy-3} three years before elections. The remaining variables are defined as in Equation 1. Even in this case, in order to understand if the electoral punishment varies with the level of irregularities found in the municipality, I interact all pre-electoral audit dummies with the corruption index.

Results with different variations of Equation 6 are in shown in Table 6.

[Table 6 Here]

The F-test of joint significance rejects the null of all coefficients being zero in each specification. These findings indicate that regardless of the timing of the audit, voters use audit information when deciding whom to vote for in national elections. The same pattern seems to be confirmed in case the policy dummies are interacted with the corruption index.

7 Placebos and Robusteness

7.1 Robusteness

To make sure that my estimates are robust to different sample selections, I run different robustness tests among different samples that I use for the main analysis.

7.1.1 Choice of the control group

Municipal Elections – As mentioned above, in municipal elections in the audited municipality, the treatment group is composed of municipalities audited before elections and the outcome of the audit is also announced before elections. The control group is composed of municipalities audited after elections, either during the same mayoral term, or those audited within the first months of the subsequent administration. The latter is justified due to the fact that most of the irregularities belong to the previous administration. However, it may also be the case that some of the municipalities in the control group are misclassified. To deal with this concern,

I restrict the sample to municipalities audited just before the municipal elections of the year 2004 and just after the same election during the same electoral term. The reason why I only use the electoral term of 2004 is because it is the only electoral term in which municipalities are audited before and after elections during the same municipal administration. In the rest of the electoral term, an audit after election coincides with a new administration.

Results are provided in the top panel of Table 7. Coefficients and standard errors are a bit larger for the difference in votes. The coefficients on the reelection dummy are also slightly larger. These differences when compared to coefficients shown in the main specification are due to the differences in the samples used.

[Table 7 Here]

Presidential Elections – To test the robustness of the estimates in the presidential sample, I consider an alternative strategy which is selecting treatment and control groups from the subset of municipalities audited before elections. I compare municipalities where the mayor is either from PT or PSDB with municipalities in which the mayor is from any other party.[63] Results are shown in the bottom panel of Table 7. Apparently, there are not any major differences between the point estimates in these samples and those that I use in the main specifications. The coefficient on β_1 is similar across those presented in panel B of Table 7.

7.1.2 Excluding the first lottery rounds

Given that the auditors did not provide a list of irregularities found in lotteries from 3-7, I exclude these lotteries in order to test the robustness of the estimates. This further robustness test demonstrates that there are no significant differences among the coefficient estimates in the full sample and those in the samples where lotteries 3-7 are excluded. Tables are shown in Tables 8 and Table 9, respectively, for the effect of the audit in the audited municipality and its neighbors.

[Table 8 Here]

[Table 9 Here]

[63] In order to be able to assign parties to outcome variables, I use a random random number generator which attributes a label to each municipality audited before elections, either PT or PSDB.

7.2 Placebos

To ensure that the empirical strategy is valid, I employ two different placebo strategies. The first consists of using a placebo outcome variable. In the second I create a placebo treatment. The first consists of estimating Equations (1) and (3), I use the differences in the previous vote shares as the outcome variable. Namely, the differences between vote shares in the years of 2000 and 1996, with respect to municipal election, and those between 2002 and 1998 in presidential elections. Note that these elections took place before the audit program was initiated so it is expected that the treatment has no effect on previous vote outcomes. Results for the benchmark case are shown in Table 10 and those for the geographical spillovers are shown in Table 11. The coefficient of interest is not statistically significant; thus it is not possible to reject the null hypothesis of no effect of the audit on previous votes.

[Table 10 Here]

[Table 11 Here]

A difference between these samples used in the placebo and those used in the main analysis is due to the fact that some municipalities did not yet exist in 1996 or 1998.[64] Another difference in the samples, which only refers to the municipal sample, arises from the fact that the incumbent party in 1996 has not participated in the municipal elections of 2000, not making it possible to generate the dependent variable, namely the difference in vote shares.

The second placebo consists in randomly allocating treatment across municipalities. This is done via simulation of a uniform distribution. I (randomly) assign municipalities to the treatment group for which the value of the random variable is above the mean and to control the control group municipalities for which the variable assumes a value below the mean. These findings indicate that the null hypothesis is not rejected more often than in 5 percent of the cases. Figure 13 depicts the distribution of the coefficient from the placebo treatment and the vertical line indicates the true coefficient estimate.

[Figure 13 Here]

The random allocation of the placebo treatment has no effect on the vote share in the municipality. Both placebo tests confirm the validity of the empirical strategy. I continue to interpret my findings as valid.

[64] New municipalities were born during the 1990's from the split of several municipalities.

8 Mechanism

8.1 The Audit Investigation

The evidence provided in this paper supports the idea that voters react to new information in neighboring governments, and at higher levels of government. These are novel findings. Previous analyses had restricted the attention to the local elections within the audited municipality. Moreover, these results show how party labels become important even in a setting where the link between the candidate and the party is considered to be weak.

It is important to note that the coefficient estimates point (except in the case of cross-border spillovers) at a negative impact of the audit on the electoral outcomes. The negative electoral effect of audits may seem puzzling at first; the information disclosed regarding municipal practices is not necessarily negative. However, it is worth bearing in mind that the content of the audit reports shows that corruption and other irregularities are widespread in Brazilian municipalities. In the sample, roughly 90 percent of municipalities have committed a major irregularity. This is in line with Ferraz and Finan (2008), Brollo et al. (2013) and Litschig and Zamboni (2011).

The negative effect of audits may arise for several reasons. Some of them are discussed shortly here below:

Media Bias: Radios may be biased towards only transmitting the negative news disclosed in the audits and forgo positive news. Evidence in this direction for Brazil can be found in Aldé, Mendes and Figueiredo (2008) which shows that most of the information on candidates reported by the press in 2006 presidential elections was negative (for both PT and PSDB).

Cognitive Bias: Individuals might only recall negative information and forget positive news, such as *negativity bias* (surveyed in Skowronski and Carlston (1989)). This behaviour has been shown to extend into belief formation about candidates in Lau (1982).

Negative Campaigning: Opponents might use negative audits in a more effective way as compared to the use of positive audit outcomes by the audited party itself.[65] In any

[65] This relates to the widespread use of negative campaigning in Brazilian national elections, which has been documented in Sampaio (2013) for the 2010 presidential dispute and Steibel (2005) for previous presidential

case, for voters to learn about the audit contents, the information needs to be disclosed either by the media or in popular discussion. Anecdotal evidence suggests that voters indeed learn about the corruption scandals unveiled from the randomized audit program. As mentioned in Section 2, these audits have been used actively in the campaign prior to local elections. There is also anecdotal evidence indicating that local audit outcomes were used for campaign purposes in presidential elections. For example, in a BBC report of October 2006, some of the voters' views indicate they do indeed learn about the corruption scandals.[66]

If one or several of the above mechanisms are in place, they could lead to a negative average effect of audits on electoral outcomes, as negative information would be used disproportionately in the ballot box. Without data on media content or cognitive bias, I cannot test for either of the first two arguments. However, I provide some evidence in line with the argument that the opposition parties might have been using these audits for campaign purposes.

For this purpose, I use data on campaign expenditures of parties and candidates for the electoral term of 2004, 2008 and 2012. I regress the logarithm of the total expenditure of each party on the audit dummy. In Panel A of Table 12, columns (1)-(4) show the spending pattern of the opposition parties in the audited municipalities. The sign and magnitude of this coefficient indicate that indeed, in local elections in audited areas, the opposition parties increase their campaign expenditures. The increase is roughly 20 percent more than in areas where information about the audit outcome is revealed after elections.

Panel B, columns (5)-(8) show the spending pattern of the opposition parties that participated in the electoral race. Apparently, these parties are decreasing spending. These results seem to point at the fact that parties are now diverting resources to safer areas, where the winning likelihood is higher.

Regarding the presidential elections, I cannot proceed in the same way for the main reason that party spending is not detailed at the municipality level. However, in presidential elections, the main findings of this paper go in line with the possibility that the voters are reminded about the party's behaviour at the local level. It is likely that the big scandals at the national level, where the main parties were blaming and shaming one another about

races.

[66] http://kochhars.com/2/hi/talking_point/6091768.stm and television debate http://eleicoes.uol.com.br/2006/campanha/ultnot/2006/10/08/ult3750u1308.jhtm

corruptness, served as a reinforcing mechanism to voters' memory about the corruption of the parties at the local level. After all, they were used by the presidential candidates to attack each other during the presidential (televised) debate.[67]

8.2 Alternative Mechanism

Vote-Buying: The findings in this paper are also consistent with an alternative interpretation. The effect on vote reduction may not occur because voters punish corruption but for the reason that now the incumbent party is no longer able to buy votes. After all, vote-buying is present at any level of government in Brazil.[68] This scenario is difficult to test without data on actual vote-buying.

However, I am providing two pieces of evidence that may rule out this alternative story. The first is related to municipal elections. In this case, I refer to the findings from the campaign expenditures. As it can be seen from Panel A in Table 12, there is no significant decrease in the amount spent before elections by the incumbent parties. This evidence is in line with the hypothesis that the incumbent parties, in the audited municipalities, are not more likely to be financially constrained after an audit. Potentially, their vote-buying ability remains unaltered. However, in the event that the incumbent party is not able to do "effective" vote-buying, because now voters are more willing to sell their vote to a less corrupt party, this may still be considered as a form of electoral punishment. A more careful investigation of the vote-buying mechanisms at the local level, with better data, is left to future research.

[Figure 12 Here]

At the presidential level, the additional evidence is brought using data from the *Latinobarómetro* surveys. In the wave of the year 2006, *Latinobarómetro* included a question related to vote-buying. The question is: *Have you known of someone in the last elections who was pressured or received something to change his vote in a certain way?*

[67] see http://news.bbc.co.uk/2/hi/6095820.stm and http://www.washingtonpost.com/wp-dyn/content/custom/2006/09/28/CU2006092800771.html

[68] The most clamorous vote-buying scandal during the period of my study and the history of the country is the vote-buying scandal that broke out on June 6, 2005 when Brazilian Congressional Deputy Roberto Jefferson told the Brazilian newspaper *Folha de S.Paulo* that the ruling Partido dos Trabalhadores (PT) had paid a number of Congressional deputies monthly salaries (around USD 12,000 at the time) to vote for the legislations that were favoured by the ruling party.

In Table 13, I provide evidence that individuals in areas audited before presidential elections, and its neighbors, provide a similar answer to individuals in the control group. This piece of evidence does not rule out the possibility completely; however, the results go in the expected direction.

[Figure 13 Here]

Parallel Auditing: A second concern, which only regards neighboring municipalities, is the following: It may be the case that voters in neighboring municipalities are not responding to information about the party but they are responding to the disclosed corruption level in their municipality. Indeed, a potential worry might be that in case the auditors find enough corruption in a municipality, the CGU may start parallel investigations in these municipalities. This does not seem to be a potential explanation for the following reasons: First, the vote shares under consideration in the neighboring municipalities in local elections are not those of the incumbent mayor. Second, I use data from the publication of the lottery rules. In these documents, the CGU also states the municipalities that are not to be audited because they have been subject to (non random) audits. The amount of these municipalities is small, and they are thus unable to influence the share of votes (in the subsample where the incumbent party is of the same party as in the audited municipality). In light of the above evidence, I continue to interpret my findings as evidence of electoral punishment.

9 Conclusions

This paper shows the extent to which information about corruption of local politicians affects the electoral outcomes of the party, not only in the same government to which the information corresponds, but it also spills over across jurisdictions and at higher levels of elections. The findings are consistent with the hypothesis that information regarding a party member at one level of government may be used by voters to decide for whom to vote across levels of elections and borders of government. Within the municipality where the wrongdoing occurs, the party of the corrupt incumbent experiences a vote reduction and a decline in the likelihood of reelection. Voters punish the party also in national elections in the audited neighborhood (audited municipality and its neighbors), suggesting that areas sharing the same media market behave similarly. In neighboring municipalities, in local elections, voters punish the party only if the disclosed corruption is high enough. These results have two

main implications. First, they show that tools meant to control corruption can have positive accountability spillovers. Second, they suggest that political parties might have incentives to better select and monitor their local members.

References

Adena, Maja, Ruben Enikolopov, Maria Petrova, Veronica Santarosa, and Ekaterina Zhuravskaya. 2015. "Radio and the Rise of The Nazis in Prewar Germany." *The Quarterly Journal of Economics*, 130(4): 1885–1939.

Aldé, Alessandra, Gabriel Mendes, and Marcus Figueiredo. 2008. "Tomando partido: imprensa e eleições presidenciais em 2006." *Política & Sociedade*, 6(10): 153–172.

Björkman, Martina, and Jakob Svensson. 2009. "Power to the People: Evidence from a Randomized Field Experiment on Community-Based Monitoring in Uganda." *The Quarterly Journal of Economics*, 124(2): 735–769.

Boas, Taylor C, and F Daniel Hidalgo. 2011. "Controlling the airwaves: Incumbency advantage and community radio in Brazil." *American Journal of Political Science*, 55(4): 869–885.

Bobonis, Gustavo J, Luis R Cámara Fuertes, and Rainer Schwabe. 2016. "Monitoring Corruptible Politicians." *American Economic Review*, 106(8): 2371–2405.

Brollo, Fernanda, Tommaso Nannicini, Roberto Perotti, and Guido Tabellini. 2013. "The Political Resource Curse." *American Economic Review*, 103(5): 1759–1796.

Bursztyn, Leonardo, and Davide Cantoni. 2015. "A tear in the iron curtain: The impact of western television on consumption behavior." *Review of Economics and Statistics*.

Center for International Earth Science Information Network CIESIN-Columbia University, Center. 2005. "Gridded Population of the World, Version 3 (GPWv3)."

Conley, Timothy G. 1999. "GMM estimation with cross sectional dependence." *Journal of econometrics*, 92(1): 1–45.

Costas-Pérez, Elena, Albert Solé-Ollé, and Pilar Sorribas-Navarro. 2012. "Corruption scandals, voter information, and accountability." *European Journal of Political Economy*, 28(4): 469–484.

DellaVigna, Stefano, and Ethan Daniel Kaplan. 2007. "The Fox News Effect: Media Bias and Voting." *The Quarterly Journal of Economics*, 122(3): 1187–1234.

DellaVigna, Stefano, Ruben Enikolopov, Vera Mironova, Maria Petrova, and Ekaterina Zhuravskaya. 2014. "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia." *American Economic Journal: Applied Economics*, 6(3): 103–132.

Desposato, Scott W. 2004. "The impact of federalism on national party cohesion in Brazil." *Legislative Studies Quarterly*, 29(2): 259–285.

Desposato, Scott W. 2006. "Parties for rent? Ambition, ideology, and party switching in Brazil's chamber of deputies." *American Journal of Political Science*, 50(1): 62–80.

Di Tella, Rafael, and Ernesto Schargrodsky. 2003. "The role of wages and auditing during a crackdown on corruption in the city of buenos aires*." *Journal of law and economics*, 46(1): 269–292.

Enikolopov, Ruben, Maria Petrova, and Ekaterina Zhuravskaya. 2010. "Media and political persuasion: Evidence from Russia." *American Economic Review*, 101(7): 3253–85.

Farvacque-Vitkovic, Catherine D, and Mihaly Kopanyi. 2014. *Municipal finances: A handbook for local governments.* World Bank Publications.

Ferraz, Claudio, and Frederico Finan. 2008. "Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes." *The 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.

Folke, Olle, Torsten Persson, and Johanna Rickne. 2016. "The Primary Effect: Preference Votes and Political Promotions." *American Political Science Review*.

Gentzkow, Matthew. 2006. "Television and voter turnout." *The Quarterly Journal of Economics*, 931–972.

Gonzalez, Robert M. 2016. "Social Monitoring and Corruption in Developing Countries."

Hayes, Philip J, and Steven P Weinstein. 1990. "CONSTRUE/TIS: A System for Content-Based Indexing of a Database of News Stories." Vol. 90, 49–64.

Hsiang, Solomon M. 2010. "Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America." *Proceedings of the National Academy of sciences*, 107(35): 15367–15372.

IBOPE. 2012. "Media Book 2012: Hábitos de Mídia e Investimento Publicitário em 2011." *IBOPE*.

Jacobs, Paul S., and Lisa F Rau. 1990. "SCISOR: Extracting information from on-line news." *Communications of the ACM*, 33(11): 88–97.

Larreguy, Horacio A, John Marshall, and James M Snyder Jr. 2014. "Revealing Malfeasance: How Local Media Facilitates Electoral Sanctioning of Mayors in Mexico." National Bureau of Economic Research.

Lau, Richard R. 1982. "Negativity in political perception." *Political Behavior*, 4(4): 353–377.

Litschig, Stephan, and Yves Zamboni. 2011. "Audit risk and rent extraction: evidence from a randomized evaluation in Brazil." *WP*.

Manning, Christopher D., Prabhakar Raghavan, and Hinrich Schütze. 2010. *Introduction to Information Retrieval, Cambridge University Press. 2008. ISBN-13 978-0-521-86571-5.* Cambridge Univ Press.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy*, 115(2).

Pereira, Carlos, and Bernardo Mueller. 2003. *Partidos fracos na arena eleitoral e partidos fortes na arena legislativa: a conexão eleitoral no Brasil.* SciELO Brasil.

Porto, Mauro. 1985. *Mass media and politics in democratic Brazil.*

Reinikka, Ritva, and Jakob Svensson. 2005. "Fighting corruption to improve schooling: Evidence from a newspaper campaign in Uganda." *Journal of the European Economic Association*, 3(2-3): 259–267.

Sampaio, Thiago. 2013. "Ataques e contra-ataques: campanhas negativas na disputa presidencial de 2010." *mimeo*.

Silva, Pedro. 2010. "Learning to fear the Inspector-General: measuring spillovers from anti-corrupt policies."

Skowronski, John J, and Donal E Carlston. 1989. "Negativity and extremity biases in impression formation: A review of explanations." *Psychological bulletin*, 105(1): 131.

Snyder, James M, and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy*, 118(2): 355–408.

Steibel, Fabro B. 2005. "Campanha negativa: construindo o objeto de estudo." *Contemporánea*.

Strömberg, David. 2004. "Radio's impact on public spending." *The Quarterly Journal of Economics*, 189–221.

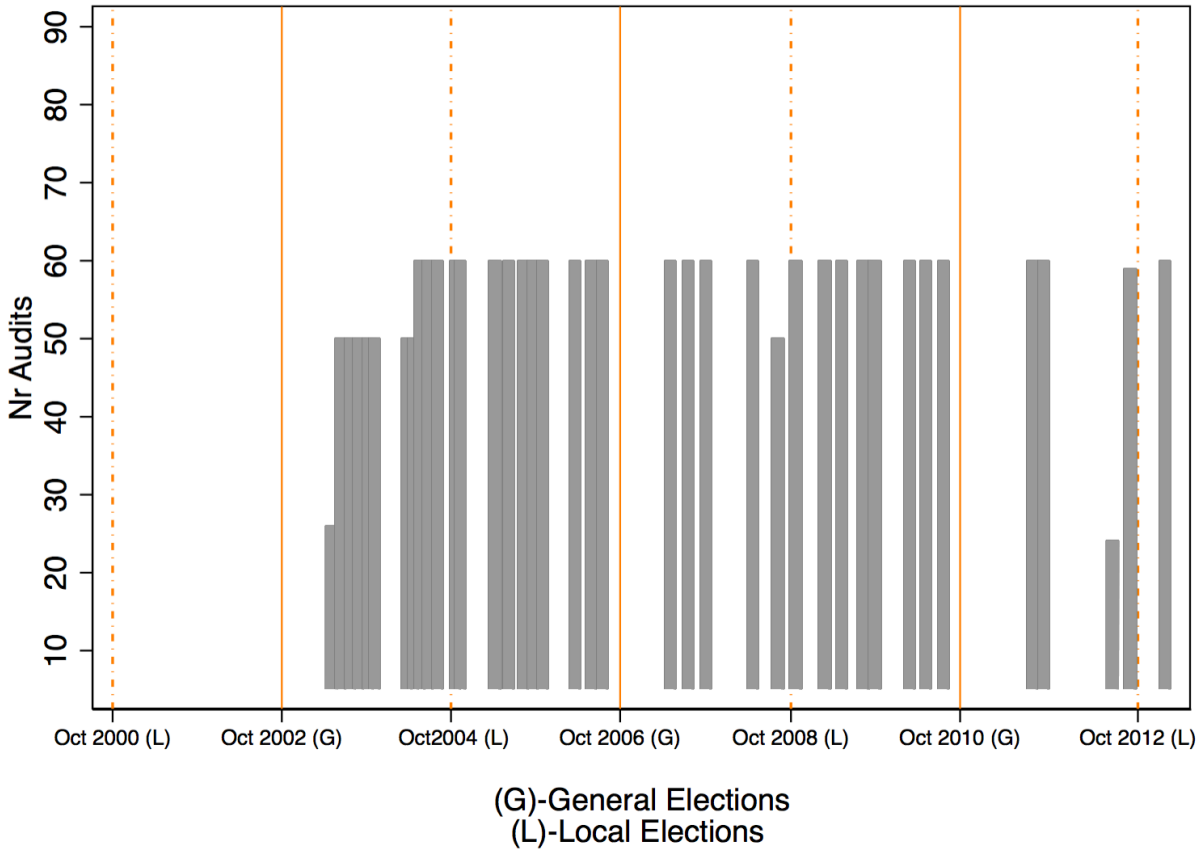
Timmons, Jeffrey F, and Francisco Garfias. 2015. "Revealed Corruption, Taxation, and Fiscal Accountability: Evidence from Brazil." *World Development*, 70: 13–27.

Tirole, Jean. 1996. "A theory of collective reputations (with applications to the persistence of corruption and to firm quality)." *The Review of Economic Studies*, 63(1): 1–22.

Treisman, Daniel. 2000. "The causes of corruption: a cross-national study." *Journal of public economics*, 76(3): 399–457.

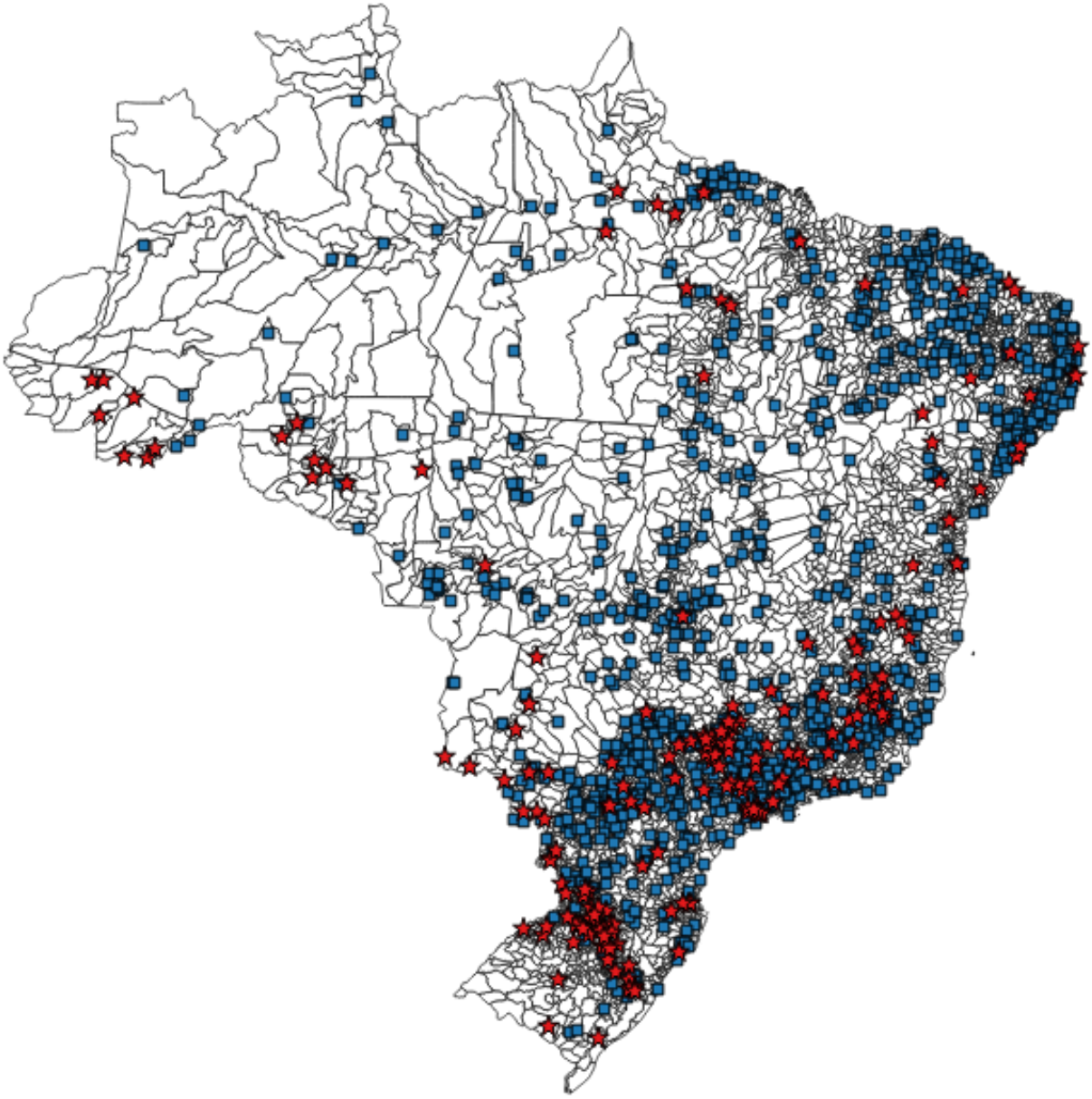
Yanagizawa-Drott, David. 2014. "Propaganda and conflict: Evidence from the Rwandan genocide." *The Quarterly Journal of Economics*, 129(4): 1947–1994.

Figure 1: Audit and Election Timeline



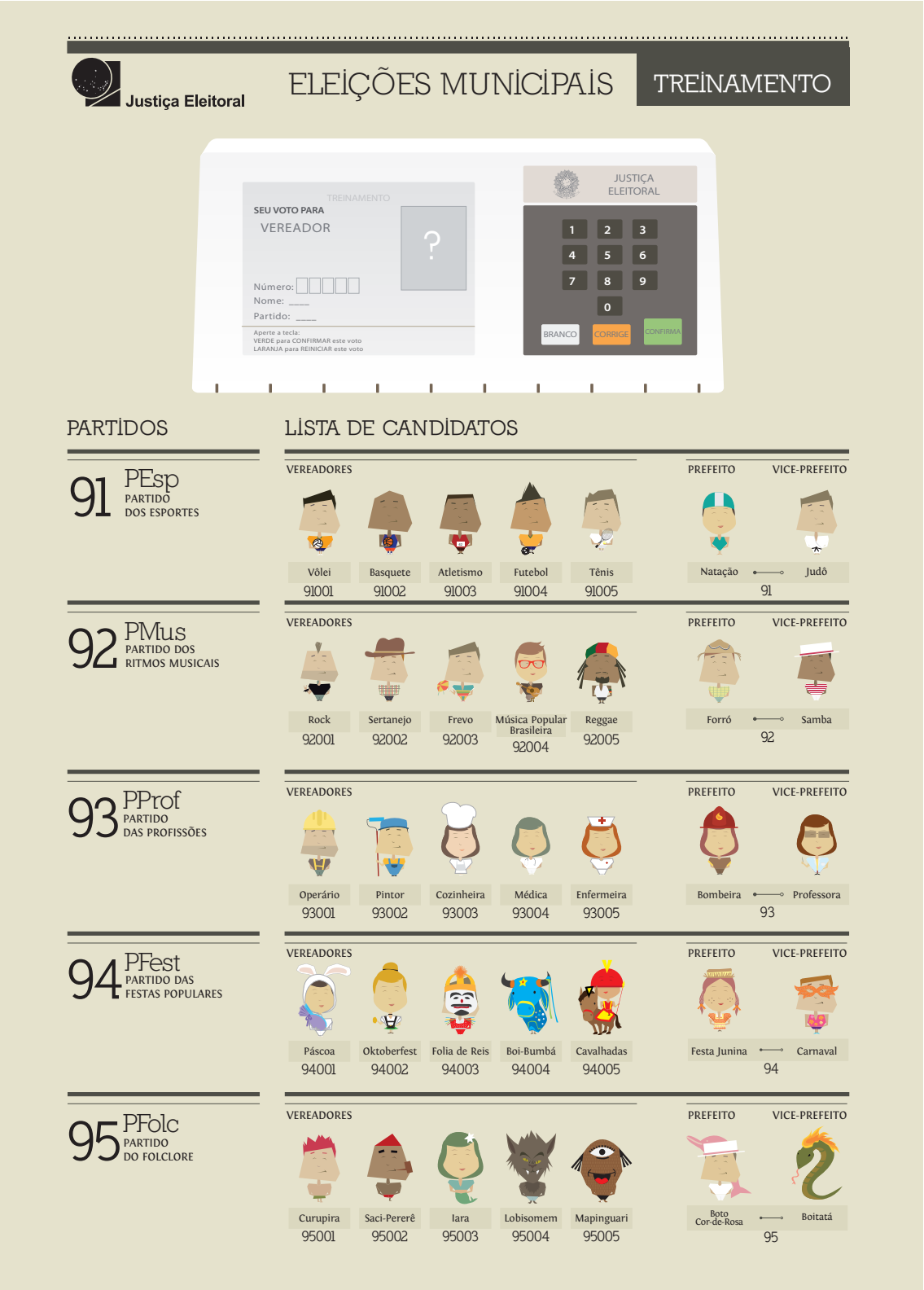
Notes: In the horizontal axes lottery rounds and the corresponding elections are depicted. The vertical axes show the number of municipalities drawn per round.

Figure 2: Distribution of PT and PSDB



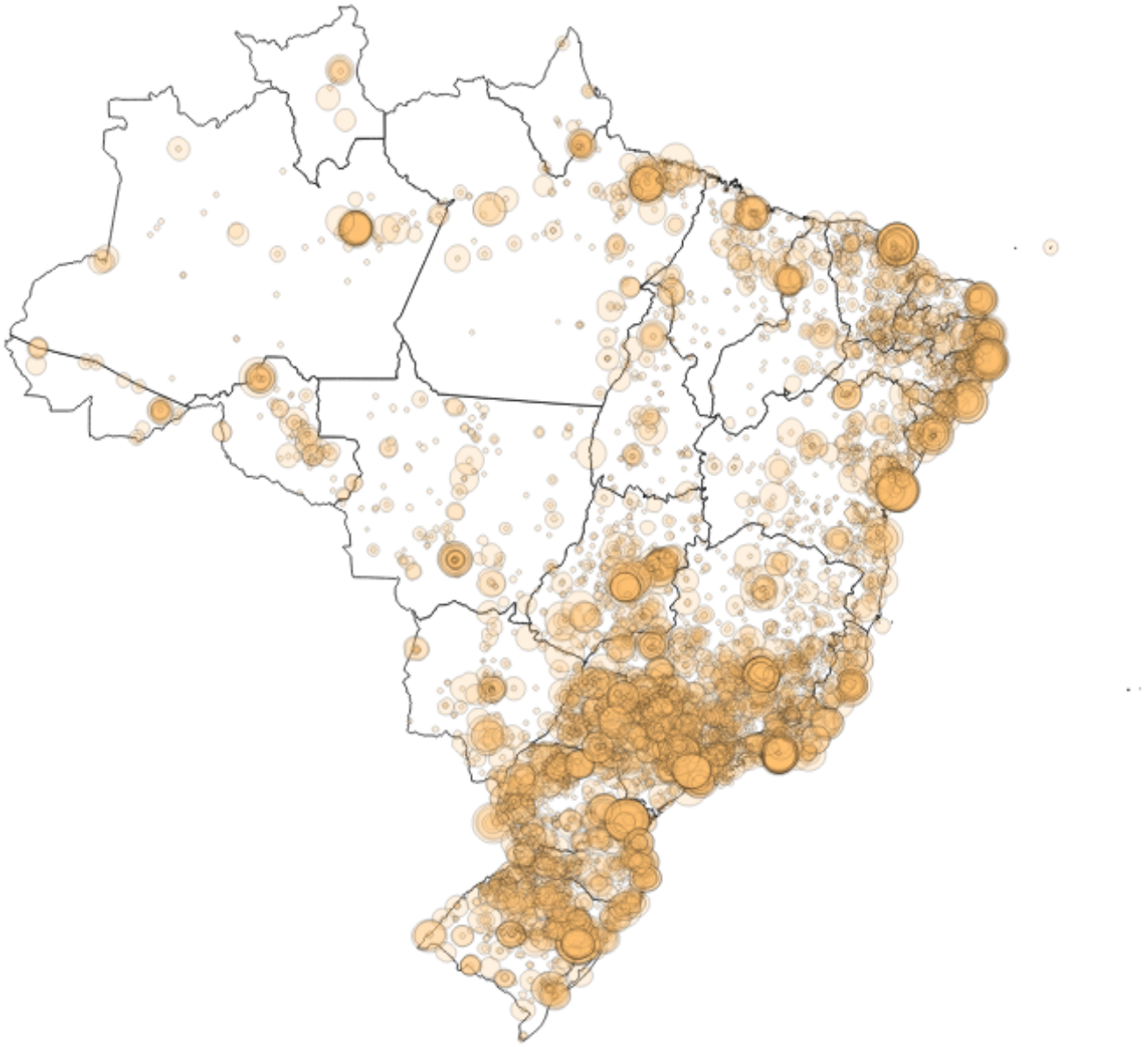
Notes: Municipalities that had a mayor either from PT (in Red) or PSDB (in Blue) during the 2001 electoral term.

Figure 3: Parties and Candidates



Notes: Example made by the TSE. It depicts the voting machine and the list of candidates and parties. Voters are required to enter the numbers associated with the pictures. The first vote is for the members of the municipal chamber and second vote is for the mayor.

Figure 4: Radio Coverage



Notes: FM antenna coverages in the Brazilian territory.

Table 1: Balance Test (Audited Municipalities)

	Municipal Elections			Presidential Elections		
	Treated	Control	*t-stat*	Treated	Control	*t-stat*
Corrupt. Index	-0.02	0.17	-1.62	0.09	-0.10	0.98
Corrupt. Index (Reduced Sample)	0.04	0.17	-1.04	0.28	-0.10	1.59
Pop. (2000)	24452.06	27508.28	-1.01	24026.82	32007.67	-1.49
Literacy rate	78.22	78.14	0.09	76.82	78.62	-1.45
Frac. Urban	0.60	0.60	-0.07	0.59	0.60	-0.51
Frac. Men	0.51	0.51	1.19	0.51	0.51	0.86
Frac. Illumination	0.22	0.23	-0.39	0.21	0.22	-1.21
Electricity TV	75.21	75.50	-0.22	71.00	73.31	-1.04
HDI	0.70	0.70	0.20	0.69	0.70	-1.29
Income per capita	171.22	168.84	0.38	160.26	170.15	-1.08
Contiguous Nbd	5.82	6.11	-2.23	6.07	5.89	0.86
Nr Antennas	2.71	2.80	-0.52	2.87	2.63	0.92
Latitude	-15.65	-15.93	0.48	-14.22	-15.31	1.27
Longitude	-45.92	-45.86	-0.14	-47.02	-46.44	-0.86
Max Distance	83.76	82.32	0.31	101.36	93.57	0.79
Min Distance	16.17	17.07	-1.08	19.05	17.76	0.67
Dist. State capital	232.43	249.25	-1.56	262.96	242.16	1.22
Dist. Federal capital	1160.46	1111.18	1.64	1115.39	1074.70	0.83
Surface	1279.32	1793.75	-1.08	2269.20	2001.69	0.37
Electorate	15784.11	17611.44	-0.98	15853.48	21267.95	-1.56
Residual Vote	0.07	0.07	-0.90	0.14	0.14	-0.95
Munic. Council	9.42	9.38	0.52	11.02	11.52	-1.64
PT	0.11	0.07	2.32	0.19	0.38	-4.37
PSDB	0.16	0.17	-0.23			
PTB	0.05	0.05	0.33			
PMDB	0.23	0.26	-0.98			
PP	0.09	0.11	-1.03			
PFL	0.14	0.16	-0.72			
Obs	594	379		181	206	

Notes: Treated is the group of municipalities audited before elections. Control is the group of municipalities audited after elections. Population and Literacy rate population and literacy rate in the municipality. Frac. Urban, Men and Illumination are, respectively, the fraction of urban population, men and household reporting to have sources of illumination. Electricity TV is the rate of households reporting to have current electricity and TV in the household. HDI stands for Human Development Index, measured at the municipality. Income per capita is measured in $Reais. All the above are measured in the year 2000. Contiguous Nbd measures the number of contiguous municipalities, sharing a border or a vertex with a nearby municipality. Nr of Antennas measures the number of FM antennas placed in the municipality. Latitude and Longitude are coordinates of the population centroid in the municipality. Max and Min Distance are, respectively, the farthest and the shortest distance from the population centroid of the audited municipality to the population centroid of the contiguous neighbors. Dist. to State and Federal capital measure the distances of municipalities from the state capital and Brasília, the federal capital. Surface is the municipal area measured in square km. Electorate is the total electorate in the municipality, Residual votes is the fraction of invalid votes, both measured in the electoral term previous to the audit, namely the year 2000 for municipal elections and 2002 for presidential elections. Munic. Council is the size of the municipal council measured as in 2000 electoral term. PSDB, PT, PTB, PMDB, PP, PFL are indicators taking the value of one if its the party of the audited mayor.

Table 2: Descriptive Statistics

VARIABLES	(1) mean	(2) sd	(3) min	(4) max
Literacy rate	77.86	12.51	39.34	98.13
Pop. (2000)	25,414.61	46,006.36	795.00	461,534.00
Frac. Urban	0.59	0.23	0.04	1.00
Frac. Men	0.51	0.01	0.47	0.58
Frac. Illumination	0.22	0.06	0.03	0.32
Electricity TV	74.34	20.30	6.24	99.35
HDI	0.70	0.08	0.47	0.92
Income per capita	167.04	94.58	34.53	834.00
Contiguous Nbd	5.96	1.99	0.00	14.00
Nr Antennas	2.71	2.64	0.00	14.00
Dist. State capital	241.45	162.25	0.00	1,154.59
Dist. Federal capital	1,132.96	456.29	33.08	2,695.80
Surface	1,741.08	7,189.62	15.36	159,695.94
Electorate	16,340.64	28,722.38	979.00	313,221.00
Residual Votes	0.07	0.03	0.01	0.44
Munic. Council	11.01	2.87	9.00	21.00
Number Parties	2.79	1.04	1.00	9.00
Vote Share	55.22	11.67	23.89	100.00
PSDB	0.21	0.41	0.00	1.00
PT	0.04	0.20	0.00	1.00
PTB	0.06	0.24	0.00	1.00
PMDB	0.23	0.42	0.00	1.00
PP	0.12	0.32	0.00	1.00
PFL	0.17	0.37	0.00	1.00
Obs	1089	1089	1089	1089

Notes: Summary statistics for all municipalities under analysis. Literacy rate is the fraction of literate population, Frac. Urban, Men and Illumination TV are, respectively, the fraction of urban population, men in the population and households that report both having electricity and TV in the household. HDI is the human development index, income per capita is that of the household in the municipality. Contiguous Nbd is the number of contiguous neighbors for each municipality, Nr of Antennas is the average number of antennas in the municipality. Distance to State and Federal Capital are measured in kilometres. Electorate is the number of people eligible to vote. Residual Vote is the fraction of invalid votes over total of valid votes. Council, Number of Parties and Vote Share are, respectively, the size of the municipal council, the number of parties in the municipality and the vote share of the winning party in year 2000. The political variables are indicators taking the value of one if the local government, in the electoral term of 2000 (pre-audit) was headed by a mayor from that party and zero otherwise.

Figure 5: Summary Example

Ministério da Agricultura, Pecuária e Abastecimento:

1.1. Inobservância de preceito da Lei nº 9.452/97 pela Prefeitura Municipal;
1.2. Desorganização formal de documentação referente a processo licitatório, em descumprimento ao disposto no artigo 38 da Lei nº 8.666/93;
1.3. Realização de licitação na modalidade Convite quando caberia Tomada de Preços; e, simulação de realização de processo licitatório;
1.4. Guarda inadequada de trator adquirido pelo Programa PRODESA.

Ministério das Cidades:

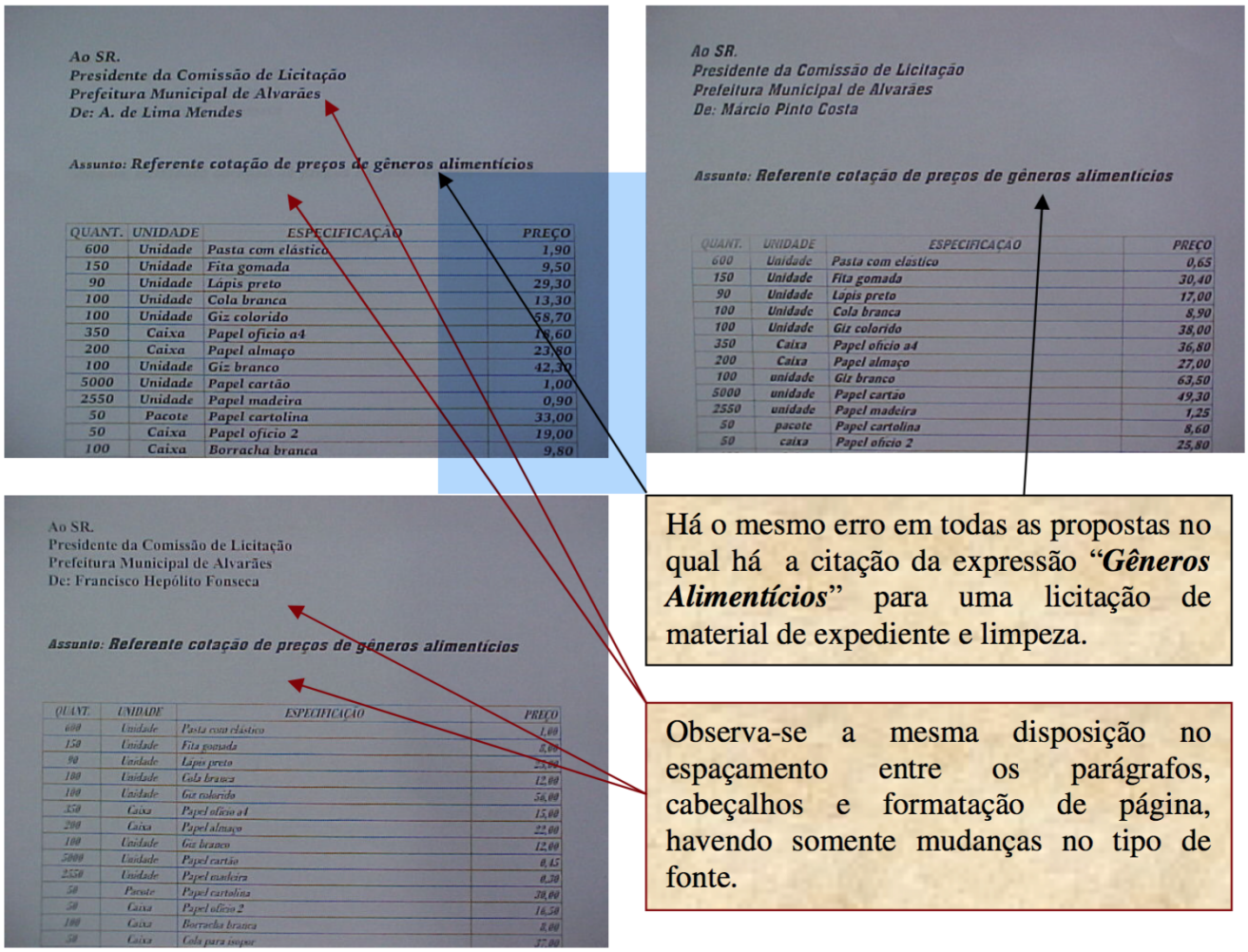
1.1. Inobservância às exigências do artigo 38 da Lei nº 8.666/93 quanto à formalização de processo;
1.2. Inobservância de preceito da Lei nº 9.452/97 pela Prefeitura Municipal.

Ministério das Comunicações:

1.1. Inexistência de postos de atendimento pessoal a usuários do serviço telefônico fixo comutado – STFC;
2.1. Restrição do acesso à Internet.

Notes: Example of a summary in the municipality Curiuva in the state of Paraná in lottery 11. For each ministry, the irregularities found are listed. Lei nº 9.452/97 and Lei nº 8.666/93 indicate, respectively, the law on Public Procurement and that on the announcement of the received funds (from the federal authorities).

Figure 6: Simulation of Procurement Process

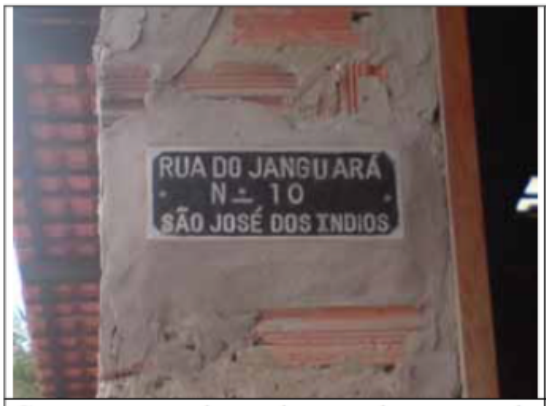


Notes: An example of a simulation of a procurement process. The call for bids was for cleaning products. All bids submitted intended to supply food. Not only is the layout of the bids the same, but they also coincide also on the type of the orthographic mistakes. The only slight change the auditors point out is the font used across bids.

Figure 7: Simulation/Fraud in Procurement Process

Estrada que dá acesso ao endereço informado como sendo da licitante de CNPJ nº 07.981.189/0001-90.

Fim da estrada onde deveria estar instalada a sede da licitante de CNPJ nº 07.981.189/0001-90.



Comprovante de endereço da rua onde deveria funcionar a empresa de CNPJ nº 07.981.189/0001-90.

Placa indicativa do bairro onde deveria funcionar a empresa de CNPJ nº 07.981.189/0001-90.

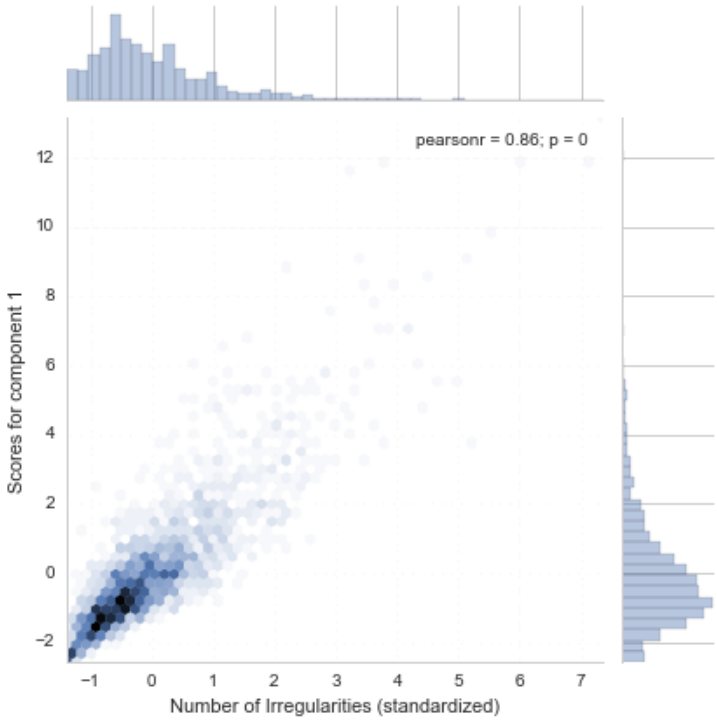
Placa indicativa do bairro onde deveria funcionar a empresa de CNPJ nº 06.313.322/0001-77.

Rua onde deveria funcionar a empresa de CNPJ nº 06.313.322/0001-77.

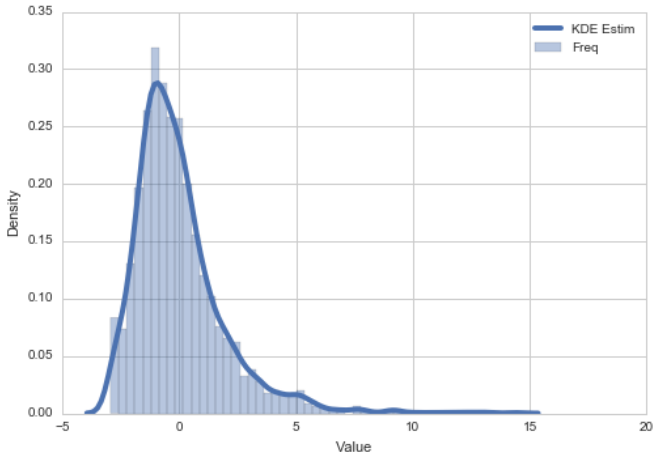
Notes: A second example of fraud in a procurement process. Three firms "participated" in the call for bids. The winner of the contract was declared to be the firm with the unique identification number (CNPJ) 07.981.189/0001-90. However, none of the firms were located in the domiciles indicated in the documents.

Figure 8: Corruption Measures

Panel A

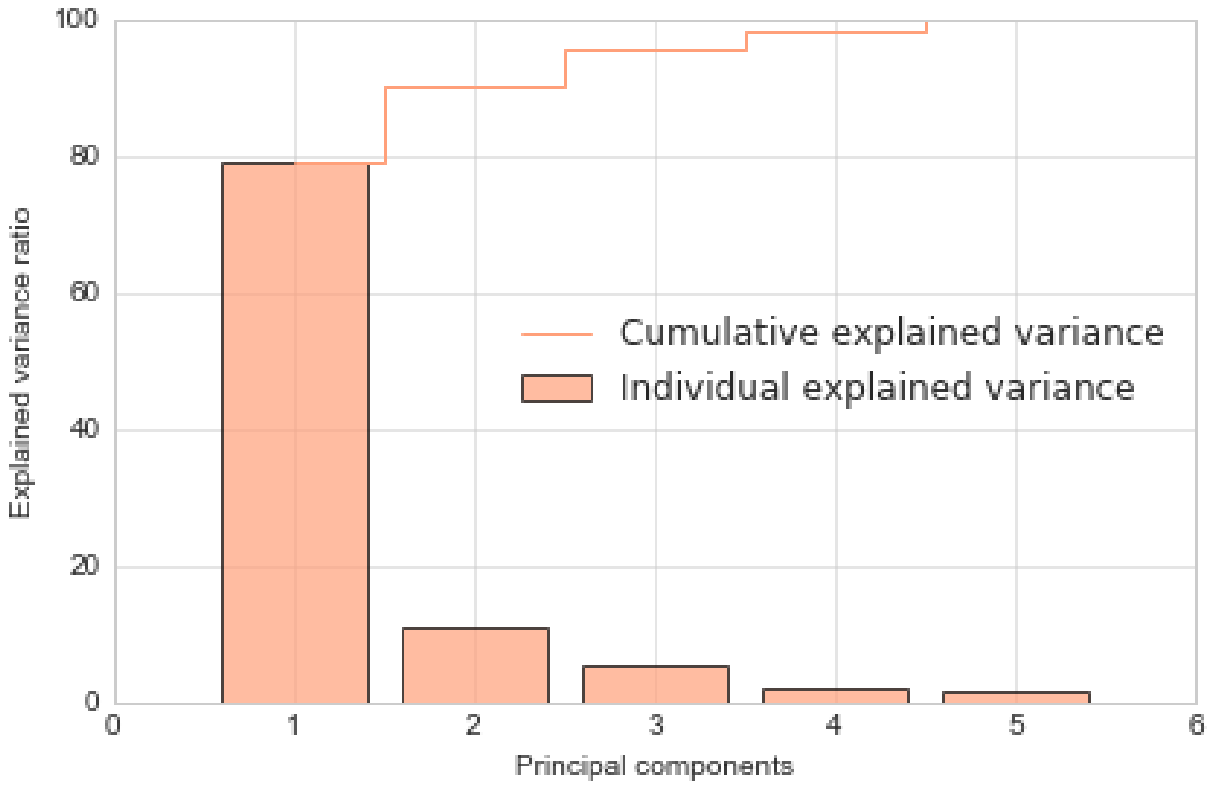


Panel B



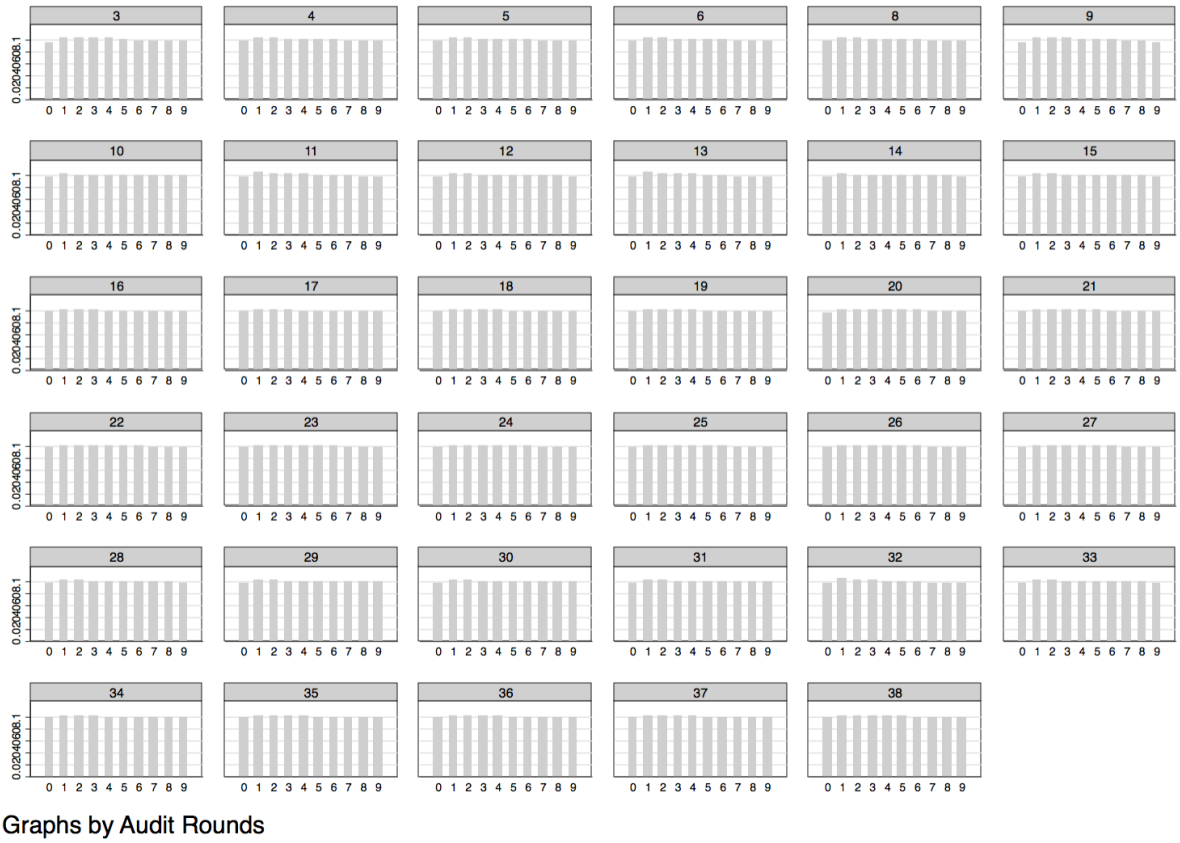
Notes: Panel A depicts the correlation of the count of major irregularities and the first principal component of the following variables: number of pages, number of lines, number of images and total amount of irregularities found in the municipality. Panel B depicts the distribution of the first principal component of all measures jointly.

Figure 9: Explained Variation



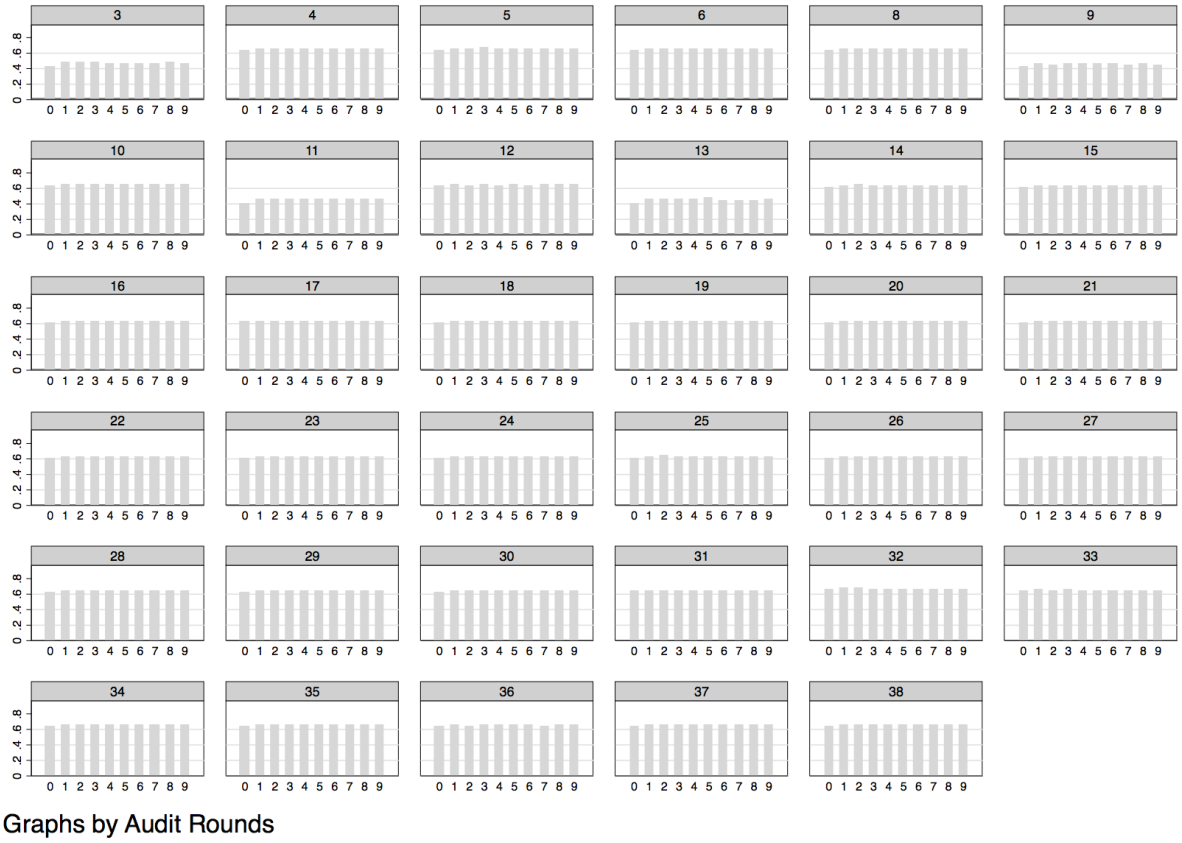
Notes: Explained variance principal component. The first principal component retains 80 percent of the explained variation.

Figure 10: Numbers Attributed per Lottery



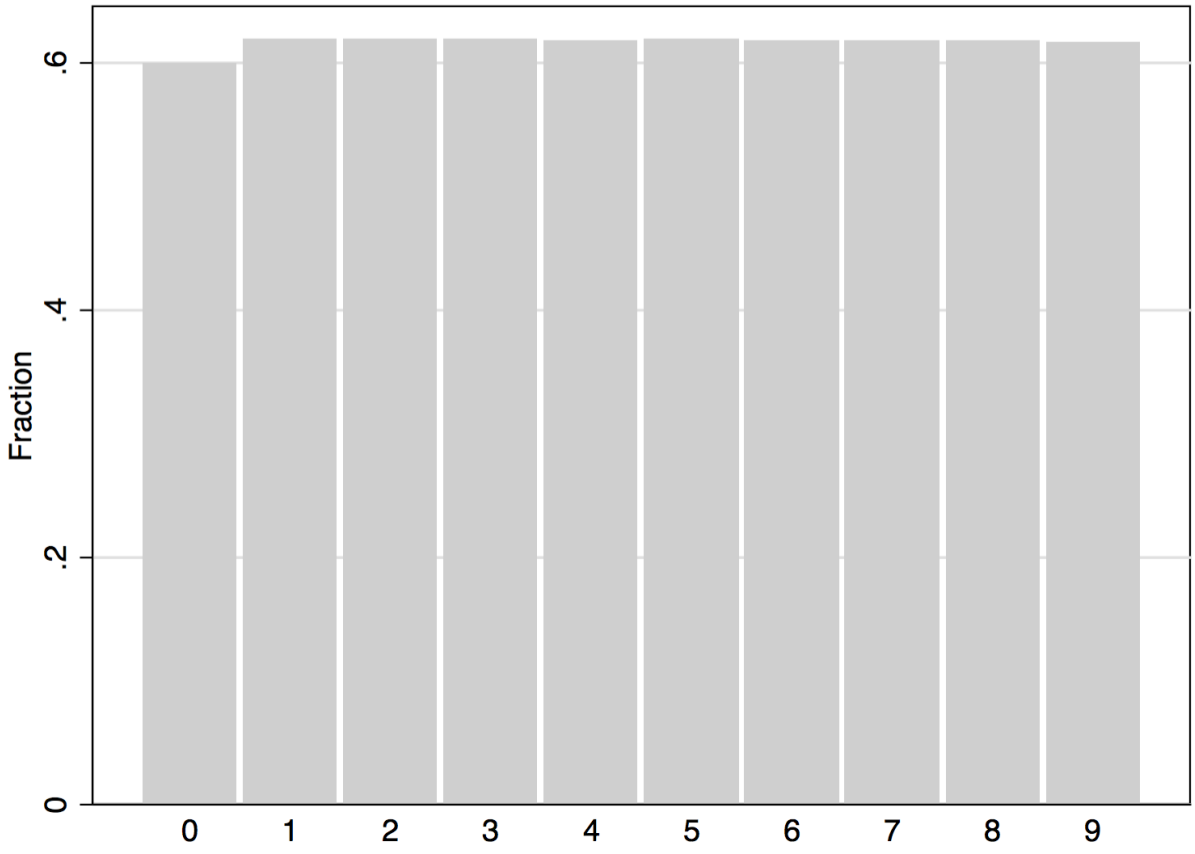
Notes: Each graph plots the numbers attributed to municipalities for a specific lottery. The distribution of numbers mimics a uniform distribution consistent with a fair lottery.

Figure 11: Numbers Selected per Lottery



Notes: Each graph plots the fraction of numbers selected over the total that number was attributed in that lottery. The distribution of numbers within each lottery mimics a uniform distribution consistent with a fair lottery.

Figure 12: Attributed and Selected Lottery Numbers



Notes: Fraction recording the number of times that the number is selected over the number of times that the number is allocated for lotteries 3-6 and 7-31.

Table 3: Direct Effect of Audit on Electoral Outcomes in the Audited Municipality

	Local Elections				Presidential Elections	
	Reelection		Δ VS		Δ VS	
PanelA	(1)	(2)	(3)	(4)	(5)	(6)
Audit	-0.094***	-0.099***	-1.952*	-2.309**	-4.135**	-4.731***
	(0.035)	(0.036)	(1.060)	(1.099)	(1.698)	(1.667)
R-squared	0.066	0.116	0.091	0.145	0.290	0.334
State FE	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓
PanelB						
Audit	-0.111***	-0.121**	-2.609**	-1.331	-4.183***	0.006
	(0.036)	(0.058)	(1.111)	(1.979)	(1.562)	(2.420)
Audit * Corrupt. Index	0.013		0.249		-1.662**	
	(0.019)		(0.536)		(0.731)	
Audit * 2^o Tercile		0.054		1.949		-3.364
		(0.081)		(2.577)		(3.592)
Audit * 3^o Tercile		0.055		-0.225		-8.807**
		(0.082)		(2.587)		(3.488)
Corrupt. Index	-0.041**		-1.012**		0.391	
	(0.018)		(0.483)		(0.609)	
2^0 Tercile		-0.108*		-4.267**		0.216
		(0.064)		(2.108)		(2.445)
3^0 Tercile		-0.146**		-3.224		1.352
		(0.068)		(2.196)		(2.877)
Observations	973	973	973	973	387	387
R-squared	0.123	0.067	0.150	0.070	0.480	0.482
State FE	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓
Controls	✓	✓	✓	✓	✓	✓
Mean Corrupt. Index	0.0556	0.0556	0.0556	0.0556	-0.0121	-0.0121
F-test joint significance	5.120	2.369	3.383	0.480	4.365	4.475

Notes: State fixed effects and population measures in the year 2000 are included in all regressions. Municipal elections, columns (1) - (4), also include term fixed effects. Controls are: Income per capita, fraction of urban population, fraction of male population and literacy rate measured in 2000 and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Panel B in columns (1), (3) and (5) show an interaction with the corruption level in linear form. Columns (2), (4) and (6) allow for a more flexible form in which terciles of the corruption index are interacted with all regressors. Standard errors clustered at the municipality level in parentheses. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 4: Electoral Outcomes in neighboring Municipalities

	Local Elections		Presidential Elections	
	Δ VS		Δ VS	
PanelA	(1)	(2)	(3)	(4)
Audit	0.313	-0.135	-3.308*	-3.491*
	(1.679)	(1.669)	(1.858)	(1.872)
R-squared	0.087	0.177	0.322	0.345
State FE	✓	✓	✓	✓
Party FE	✓	✓	✓	✓
Controls		✓		✓
PanelB				
Audit	-0.645	5.195	-3.134*	2.120
	(1.663)	(3.273)	(1.708)	(2.175)
Audit * Corrupt. Index	-2.130**		-2.306***	
	(0.917)		(0.783)	
Audit * 2^0 Tercile		-8.297*		-4.858
		(4.849)		(3.844)
Audit * 3^0 Tercile		-9.084**		-9.847***
		(4.351)		(3.218)
Corrupt. Index	0.913		0.196	
	(0.664)		(0.862)	
2^0 Tercile		4.973		1.402
		(4.217)		(2.777)
3^0 Tercile		4.814		-1.432
		(3.393)		(3.080)
Observations	465	465	871	871
R-squared	0.139	0.140	0.490	0.499
State FE	✓	✓	✓	✓
Party FE	✓	✓	✓	✓
Controls	✓	✓	✓	✓
Mean Corrupt. Index	-0.00490	-0.00490	-0.00515	-0.00515
F-test joint significance	1.950	1.667	4.468	3.553

Notes: State fixed effects and population measure in the year 2000 are included in all regressions. The municipal elections column (1) - (2) also include term fixed effects. Panel B, columns (1) and (3) show an interaction with the corruption level in linear form. Columns (2) and (4) allow for a more flexible form in which terciles of the corruption index are interacted with all regressors. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the mesoregion area. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 5: Parties in coalition

Dependent Variable: Change in the Vote Share of the main coalition party							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Audit	0.334	0.345	0.298	0.272	0.425	0.339	2.255
	(1.472)	(1.303)	(1.319)	(1.379)	(1.330)	(1.338)	(2.243)
Audit * Corrupt. Index				-0.926	-0.767	-1.073	
				(0.828)	(0.778)	(0.763)	
Audit * 2^0 Tercile							-1.145
							(3.014)
Audit * 3^0 Tercile							-4.436
							(3.536)
Corrupt. Index				-0.162	-0.006	0.092	
				(0.442)	(0.404)	(0.430)	
2^0 Tercile							-2.157
							(2.394)
3^0 Tercile							-0.519
							(2.598)
Observations	498	498	498	489	489	489	489
R-squared	0.339	0.484	0.502	0.440	0.481	0.502	0.460
State FE	✓	✓	✓	✓	✓	✓	✓
Party FE		✓	✓		✓	✓	✓
Controls			✓			✓	✓

Notes: The effect of the audit on the electoral outcomes of the main party in the coalition. State fixed effects, party fixed effect, and population measured in the year 2000, are included as controls in all regressions. Controls include income per capita and population, the fraction of the urban population, the fraction of the male population and the literacy rate, surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the municipality level in parenthesis. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 6: Heterogeneous Treatment (Presidential Elections)

Dependent Variable: Change in the Vote Share				
PanelA	(1)	(2)	(3)	(4)
1 year Before Elections	-4.970**	-5.534**	-4.387*	-4.890**
	(2.246)	(2.312)	(2.302)	(2.395)
2 year Before Elections	-3.743*	-3.675*	-3.725*	-3.581*
	(1.909)	(1.919)	(1.920)	(1.938)
3 year Before Elections	-3.084	-3.386	-4.066*	-4.391*
	(2.095)	(2.131)	(2.360)	(2.408)
1 year Before Elections * Corrupt. Index			-1.873**	-1.737**
			(0.864)	(0.845)
2 year Before Elections * Corrupt. Index			-1.209	-1.225
			(1.022)	(1.047)
3 year Before Elections * Corrupt. Index			-2.997	-3.166
			(2.034)	(2.114)
Observations	387	387	387	387
R-squared	0.458	0.473	0.466	0.481
State FE	✓	✓	✓	✓
Party FE	✓	✓	✓	✓
Controls		✓		✓
PanelB				
1 year Before Elections	-3.124	-3.026	-2.956	-2.714
	(2.661)	(2.868)	(2.611)	(2.852)
2 year Before Elections	-3.298	-3.823	-4.098	-4.350
	(2.563)	(2.820)	(2.474)	(2.658)
3 year Before Elections	-1.190	-1.070	-2.391	-2.192
	(2.378)	(2.555)	(2.775)	(2.994)
1 year Before Elections * Corrupt. Index			-1.932**	-2.081**
			(0.844)	(0.856)
2 year Before Elections * Corrupt. Index			-2.851**	-3.040**
			(1.252)	(1.433)
3 year Before Elections * Corrupt. Index			-1.751	-1.983
			(2.095)	(2.246)
Observations	871	871	871	871
R-squared	0.459	0.476	0.476	0.491
State FE	✓	✓	✓	✓
Party FE	✓	✓	✓	✓
Controls		✓		✓

Notes: Panel A and B show results from the audited municipality and neighbors, respectively. State fixed effects and population measure in the year 2000 are included in all regressions. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the municipality level for the audited municipality and at the mesoregion area in the geographical spillovers.*** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 7: Robustness (Change of the control group)

Outcome	Party Reelected			Changes in Vote Share		
	(1)	(2)	(3)	(4)	(5)	(6)
Audit	-0.128*	-0.130*	-0.148*	-4.779*	-4.638*	-5.296**
	(0.076)	(0.074)	(0.077)	(2.525)	(2.499)	(2.571)
Audit * Corrupt. Index			0.002			-0.235
			(0.046)			(1.538)
Corrupt. Index			-0.032			-1.014
			(0.041)			(1.367)
Observations	302	302	302	302	302	302
R-squared	0.140	0.220	0.224	0.161	0.230	0.235
State FE	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓
Controls		✓	✓		✓	✓

Outcome: Changes in Party's Vote Share						
	(1)	(2)	(3)	(4)	(5)	(6)
Audit	-3.567*	-4.095*	-4.104*	-3.387	-3.955*	-3.956*
	(2.151)	(2.132)	(2.188)	(2.077)	(2.083)	(2.133)
Audit * Corrupt. Index				-3.398***	-3.384***	-3.300***
				(1.040)	(1.038)	(1.106)
Corrupt. Index				0.224	0.148	0.082
				(0.623)	(0.631)	(0.656)
Observations	558	558	558	558	558	558
R-squared	0.122	0.123	0.135	0.137	0.139	0.150
State FE	✓	✓	✓	✓	✓	✓
Party FE		✓	✓		✓	✓
Controls			✓			✓

Notes: Municipal election are shown in the top panel. The bottom panel shows results from presidential elections. State fixed effects and population measure in the year 2000 are included in all regressions. In municipal elections, term fixed effects are also included. The dependent variable in the top panel is, respectively, the change in the reelection status in columns (1) - (3) and the difference in vote shares in columns (4) - (6). State fixed effects and population, measured in year 2000, are included as controls. Results for presidential elections are presented in the bottom panel, where the treatment group is a municipality audited in one of the two main political parties and the control group is composed of municipalities audited in any other party. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the municipality level, in parentheses.*** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 8: Excluding Lotteries 3-7 (Audited Municipality)

	Local Elections				Presidential Elections	
	Reelection		Δ VS		Δ VS	
PanelA	(1)	(2)	(3)	(4)	(5)	(6)
Audit	-0.097**	-0.106***	-2.097*	-2.644**	-4.743**	-5.380***
	(0.040)	(0.041)	(1.249)	(1.292)	(1.934)	(1.918)
R-squared	0.088	0.147	0.098	0.154	0.283	0.329
State FE	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓
PanelB						
Audit	-0.112***	-0.117***	-2.165*	-2.875**	-4.125**	-4.348**
	(0.039)	(0.041)	(1.245)	(1.298)	(1.788)	(1.809)
Audit * Corrupt. Index	0.016	0.015	0.555	0.418	-1.706**	-1.593**
	(0.019)	(0.020)	(0.542)	(0.539)	(0.796)	(0.788)
Corrupt. Index	-0.047***	-0.043**	-1.077**	-0.975**	0.132	0.027
	(0.018)	(0.018)	(0.500)	(0.493)	(0.594)	(0.637)
Observations	848	848	848	848	332	332
R-squared	0.082	0.155	0.063	0.159	0.458	0.476
State FE	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓
Mean Corrupt. Index	0.0965	0.0965	0.0965	0.0965	0.0424	0.0424
F-test joint significance	5.436	4.727	2.554	2.871	4.217	4.268

Notes: Municipal elections are shown in the top panel. They also include state fixed effects and population as of the year 2000. The bottom panel shows results from presidential elections. In the municipal elections, columns (1)-(4), term fixed effects are included. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the municipality level in parentheses.*** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 9: Excluding Lotteries 3-7 (Geographical Spillovers)

	Local Elections		Presidential Elections	
	Δ VS		Δ VS	
PanelA	(1)	(2)	(3)	(4)
Audit	-0.582	-0.315	-4.653**	-5.044**
	(1.805)	(1.847)	(2.225)	(2.183)
R-squared	0.076	0.213	0.310	0.343
State FE	✓	✓	✓	✓
Party FE	✓	✓	✓	✓
Controls		✓		✓
PanelB				
Audit	-0.879	-0.729	-4.082**	-4.118**
	(1.750)	(1.840)	(1.968)	(2.003)
Audit * Corrupt. Index	-2.105**	-2.120**	-2.417***	-2.582***
	(0.863)	(0.847)	(0.833)	(0.856)
Corrupt. Index	1.487**	1.205	-0.313	-0.178
	(0.679)	(0.737)	(0.987)	(1.132)
Observations	383	383	746	746
R-squared	0.089	0.225	0.474	0.495
State FE	✓	✓	✓	✓
Party FE	✓	✓	✓	✓
Controls		✓		✓
Mean Corrupt. Index	0.0112	0.0112	0.0346	0.0346
F-test joint significance	2.455	2.183	5.476	5.689

Notes: Municipal election are shown in the top panel. The bottom panel shows results from presidential elections. State fixed effects are included and population as of the year 2000. In the municipal elections, columns (1)-(4), term fixed effects are included. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the mesoregion area. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 10: Placebo Audited Municipality

	Local Elections				Presidential Elections	
	Reelection		Δ VS		Δ VS	
PanelA	(1)	(2)	(3)	(4)	(5)	(6)
Audit	-0.025	-0.014	-0.362	-0.221	0.538	0.713
	(0.047)	(0.049)	(1.658)	(1.720)	(1.368)	(1.445)
Observations	540	540	540	540	325	325
R-squared	0.064	0.130	0.131	0.219	0.812	0.818
State FE	✓	✓	✓	✓	✓	✓
Party FE		✓		✓		✓
Controls		✓		✓		✓

Notes: All regressions include state fixed effects and population as of the year 2000. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the municipality level in parenthesis. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

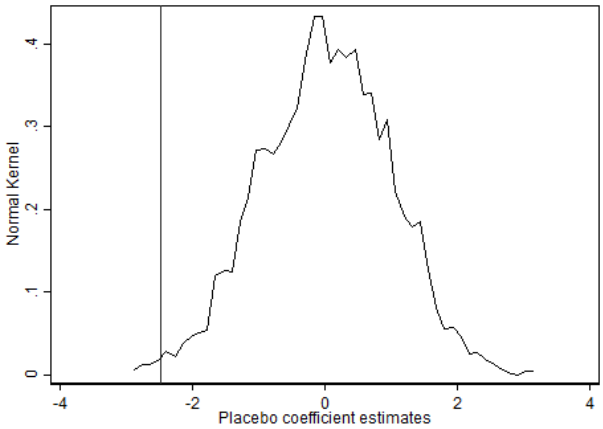
Table 11: Placebos Geographical Spillovers

	Local Elections		Presidential Elections	
	Δ VS		Δ VS	
	(1)	(2)	(3)	(4)
Audit	-0.171	0.168	-1.028	-1.353
	(2.372)	(2.678)	(1.246)	(1.185)
Observations	255	255	828	828
R-squared	0.185	0.232	0.817	0.823
State FE	✓	✓	✓	✓
Party FE		✓		✓
Controls		✓		✓

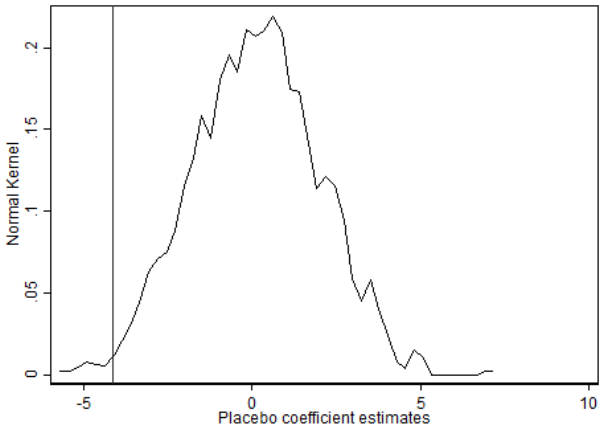
Notes: In the municipal regressions, columns (1) and (2), term fixed effects are included. All regressions include state fixed effects and population as of the year 2000. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors clustered at the municipality level and at the mesoregion level respectively for audited municipality and neighboring municipalities in parenthesis. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Figure 13: Coefficient Distribution for Placebo Test

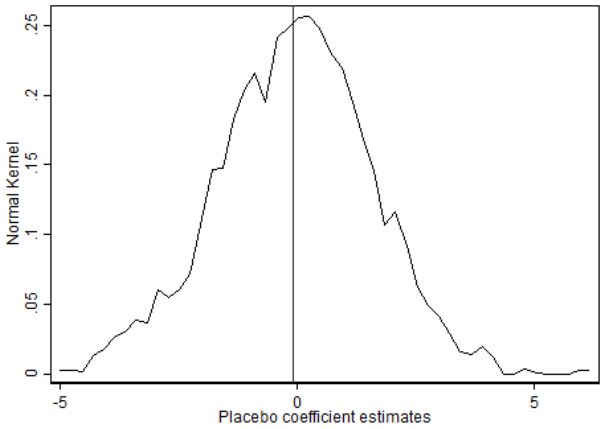
Panel A: Audited Municipal



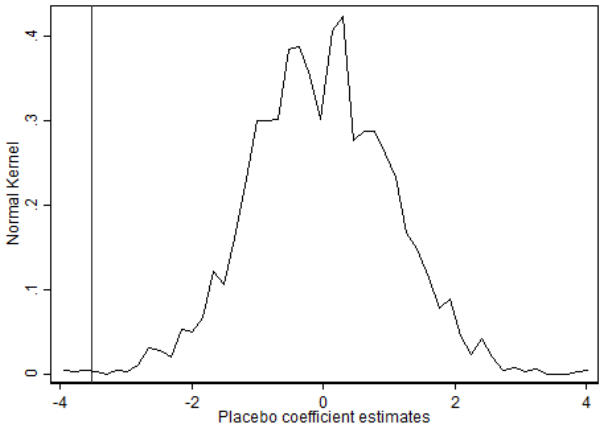
Panel B: Audited Presidential



Panel C: neighbors Municipal



Panel C: neighbors Presidential



Density of predicted β_0 from the placebo regressions. Each vertical line represents the estimated value of the true coefficient.

Table 12: Campaign Expenditure

	Opposition Parties				Incumbent Parties			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A								
Audit	0.229**	0.242**	0.208*	-0.174	-0.137	-0.179	-0.232*	-0.347
	(0.114)	(0.116)	(0.116)	(0.213)	(0.127)	(0.130)	(0.128)	(0.253)
Audit * Corrupt. Index			-0.010				0.029	
			(0.059)				(0.066)	
Corrupt. Index			-0.041				-0.027	
			(0.051)				(0.056)	
Audit * 2^0 Tercile				0.844***				0.237
				(0.281)				(0.313)
Audit * 3^0 Tercile				0.272				0.114
				(0.285)				(0.309)
2^0 Tercile				-0.313				-0.100
				(0.225)				(0.248)
3^0 Tercile				-0.213				-0.099
				(0.230)				(0.264)
Observations	1,400	1,400	1,400	1,400	817	817	817	817
R-squared	0.310	0.344	0.363	0.368	0.418	0.438	0.477	0.478
Panel B								
	Opposition Parties				Party of the Audited Neighbour			
Audit	-0.393*	-0.424*	-0.288	-0.475	-0.346	-0.376	-0.303	-0.312
	(0.221)	(0.229)	(0.219)	(0.329)	(0.267)	(0.276)	(0.289)	(0.457)
Audit * Corrupt. Index			0.087				0.009	
			(0.089)				(0.106)	
Corrupt. Index			-0.081				0.045	
			(0.062)				(0.087)	
Audit * 2^0 Tercile				0.496				0.106
				(0.497)				(0.551)
Audit * 3^0 Tercile				0.043				-0.208
				(0.457)				(0.521)
2^0 Tercile				-0.759*				-0.265
				(0.387)				(0.550)
3^0 Tercile				-0.190				0.190
				(0.400)				(0.544)
Observations	766	766	766	766	412	412	412	412
R-squared	0.366	0.402	0.455	0.459	0.370	0.387	0.439	0.441
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓	✓	✓
Term	✓	✓	✓	✓	✓	✓	✓	✓
Party x Term FE		✓	✓	✓		✓	✓	✓
Controls			✓	✓			✓	✓

Notes: Panels A and B show results for the audited and neighboring municipalities, respectively. Column (1) - (4) depict results for the opposition parties, columns (5) - (8) show results regarding incumbent parties. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors are clustered at the municipality level in the case of the audited municipality and at the mesoregion area in the case of geographical spillovers. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

Table 13: Vote-buying

	Audited Municipality			Informed Neighbour		
	(1)	(2)	(3)	(4)	(5)	(6)
Audit	0.014	-0.060	-0.087	0.087	-0.025	-0.025
	(0.076)	(0.155)	(0.155)	(0.060)	(0.126)	(0.127)
Observations	125	125	125	405	405	405
R-squared	0.000	0.122	0.173	0.005	0.062	0.100
State FE		✓	✓		✓	✓
Population		✓	✓		✓	✓
Controls			✓			✓

Notes: The dependent variable is an indicator taking value 1 if the individual replies yes to the question: "Have you known of someone in the last elections who was pressured or received something to change his vote in a certain way?". Survey weights are included in all regressions. In the last column, individual controls such as age, religion, sex, use of email, home ownership, TV ownership, expenditure in health, family income, and health status are included. Population is the municipal population. Robust standard errors in parenthesis. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

10 Appendix

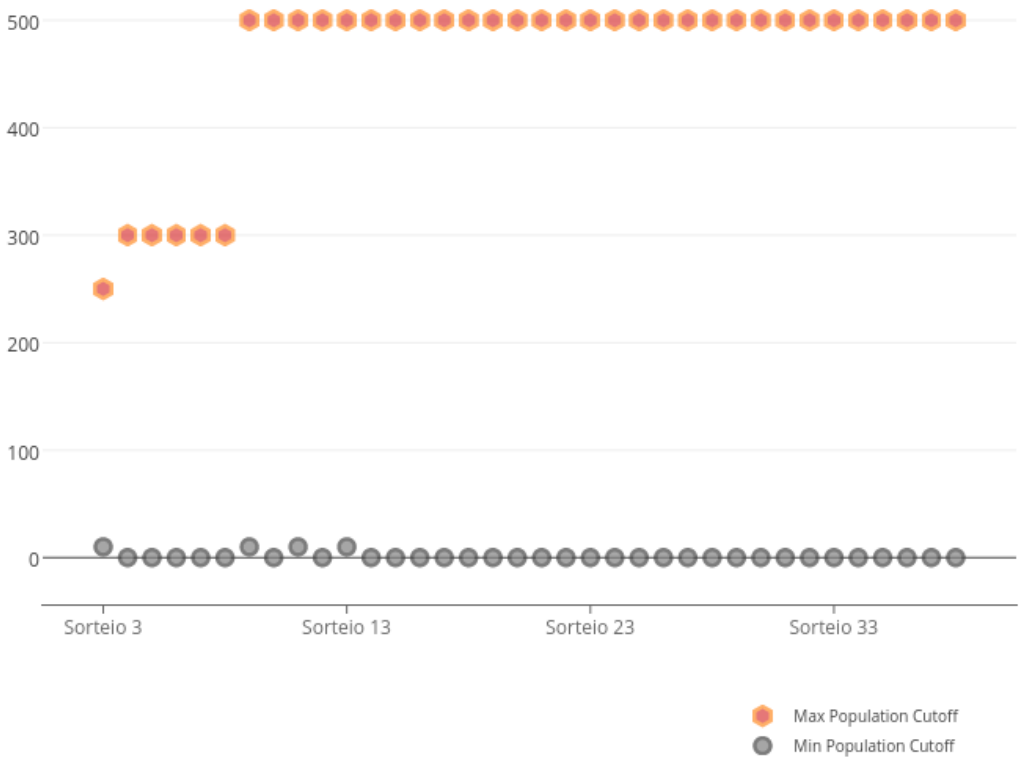
10.1 More on the audit program

Although the rules for exclusion of municipalities in the lottery change over time, 85 percent of the municipalities have been eligible for selection during my sample period. In Table A1, for each lottery round, the date the lottery took place, the date of publication of lottery rules, the date of publication of audit outcomes and replacement rules are shown. In Figure 14, the population cutoffs for each lottery are plotted. The population estimates are based on the IBGE yearly projections.[69]

Another difference between the lottery rounds is the stratification. In earlier rounds, up to lottery 8, the stratification was done at the state level. In later rounds, states with a smaller number of municipalities are bundled together (see Table A2). For example in lottery 9, states Acre, Amapá, Amazonas, Rondônia, and Roraima were bundled together and one municipality was to be selected. In lottery 10 Acre, Amapá, Rondônia, and Roraima were bundled together and one municipality was to be selected from the state of Amazonas separately.

[69]Except from lottery 32, which had two stages, the rest of the lotteries only had one stage. Each stage implied two overlapping population brackets. Within each population bracket, 30 municipalities were audited.

Figure 14: Population Cutoffs for Each Lottery



Notes: Population brackets for each lottery expressed in 1000 of inhabitants. In grey, minimum population cutoff and in orange maximum population cutoff. The population brackets, except the first rounds, have been pretty constant over time.

Table A1: Population Cutoffs

Lottery	Draw Date	Date Official	Date Pub.	Replacement
3	18/06/2003	11/06/2003	03/09/2003	past 3 lotteries
4	30/07/2003	24/07/2003	17/10/2003	past 3 lotteries
5	03/09/2003	28/08/2003	03/09/2003	past 3 lotteries
6	15/10/2003	09/10/2003	12/02/2004	past 3 lotteries
7	12/11/2003	10/11/2003	27/04/2004	past 6 lotteries
8	30/03/2004	24/03/2004	11/08/2004	past 6 lotteries
9	29/04/2004	26/04/2004	27/09/2004	past 6 lotteries
10	26/05/2004	17/05/2004	27/09/2004	past 6 lotteries
11	30/06/2004	08/06/2004	12/01/2005	past 6 lotteries
12	11/08/2004	26/07/2004	22/02/2005	past 6 lotteries
13	27/10/2004	23/09/2004	15/04/2005	past 6 lotteries
14	17/11/2004	16/11/2004	24/06/2005	past 12 lotteries
15	14/04/2005	12/04/2005	30/09/2005	past 12 lotteries
16	09/06/2005	03/06/2005	22/11/2005	past 12 lotteries
17	16/08/2005	08/08/2005	22/02/2006	past 12 lotteries
18	27/09/2005	14/09/2005	30/05/2006	past 12 lotteries
19	07/11/2005	27/10/2005	22/08/2006	past 12 lotteries
20	23/03/2006	15/03/2006	29/12/2006	past 12 lotteries
21	02/06/2006	24/05/2006	06/03/2007	past 12 lotteries
22	19/07/2006	12/07/2006	09/07/2007	past 12 lotteries
23	09/05/2007	24/04/2007	25/01/2008	past 12 lotteries
24	24/07/2007	18/07/2007	01/04/2008	past 12 lotteries
25	09/10/2007	1/10/2007	06/06/2008	past 12 lotteries
26	30/04/2008	17/04/2008	12/12/2008	past 6 lotteries
27	29/10/2008	22/10/2008	27/04/2009	past 12 lotteries
28	12/05/2009	07/05/2009	11/11/2009	past 3 lotteries
29	17/08/2009	11/08/2009	19/03/2010	past 3 lotteries
30	05/10/2009	28/09/2009	09/07/2010	past 3 lotteries
31	01/03/2010	23/02/2010	06/10/2010	past 3 lotteries
32	10/05/2010	30/04/2010	21/01/2011	past 3 lotteries
33	26/07/2010	20/07/2010	20/01/2011	past 3 lotteries
34	15/08/2011	05/08/2011	07/03/2012	past 3 lotteries
35	03/10/2011	27/09/2011	24/04/2012	past 3 lotteries
36	23/07/2012	11/07/2012	17/01/2013	past 3 lotteries
37	08/10/2012	04/10/2012	04/04/2013	past 3 lotteries
38	04/03/2013	18/02/2013	18/09/2013	past 3 lotteries

Notes: Lottery stands for the round of the lottery. Draw Date is the date in which the lottery took place for that specific round. Date Official is the date the rules of the lottery were published. Date Pub. is the date in which the audit reports were published. Replacement indicates which lotteries were eligible for selection among those in the published population brackets.

Table A2: Division of Brazilian States

State	Municipalities	Microregion	Mesoregion
Acre	22	2	5
Alagoas	102	3	13
Amapá	16	2	4
Amazonas	62	4	13
Bahia	417	7	32
Ceará	184	7	33
Espírito Santo	78	4	13
Goiás	246	5	18
Maranhão	217	5	21
Mato Grosso	141	5	22
Mato Grosso do Sul	78	4	11
Minas Gerais	853	12	66
Pará	143	6	22
Paraíba	223	4	23
Paraná	399	10	39
Pernambuco	185	5	19
Piauí	224	4	15
Rio De Janeiro	92	6	18
Rio Grande do Norte	167	4	19
Rio Grande do Sul	496	7	35
Rondônia	52	2	8
Roraima	15	2	4
Santa Catarina	293	6	20
São Paulo	645	15	63
Sergipe	75	3	13
Tocantins	139	2	8

Notes: Subdivision of states into municipalities, microregions and, mesoregions. Distrito Federal (Federal Districtis) not included. According to the Brazilian Constitution the Federal District is not separable into municipalities.

10.2 Selection Probability

To further address the large t-statistics for the PT dummy in the presidential sample, I use all documentation on the published lotteries in the *Official Gazette* and available on the CGU website. To investigate if a mayor from PT is more or less likely to be selected for an audit, I run a linear probability model of the audit probability on a dummy taking value one if the municipality is headed by a PT mayor. In Panel A of Table A3 the full sample of eligible municipalities for each lottery is included. In Panel B, I restrict the sample to the municipalities that have a mayor either from PT or PSDB. The subsample of the municipalities actually audited constitutes the presidential sample.

[Table A3 Here]

Findings suggest that having a mayor from PT does not differentially affect the likelihood of selection. This is also the case when lottery fixed effects are added. In column (2), results from within lottery variation are shown. In column (3), I add the probability of selection of each municipality in the lottery. The latter is computed as the number of municipalities to be selected over the total municipalities eligible in the state or group of states. In column (4), both lottery fixed effects and the probability of selection are included in the regression. In column (5), I also include group fixed effects. In column (6) I include state fixed effects instead of group fixed effects. The coefficient on the PT changes only marginally indicating that PT does not receive any preferential treatment.

Table A3: Likelihood of an Audit

PanelA	(1)	(2)	(3)	(4)	(5)	(6)
PT	-0.0005	-0.0005	-0.0008	-0.0007	-0.0005	-0.0005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Selection Prob			2.8693***	2.9183***	5.9840***	1.5288***
			(0.519)	(0.524)	(1.938)	(0.589)
Observations	156,551	156,551	156,551	156,551	156,551	156,551
R-squared	0.000	0.001	0.001	0.001	0.003	0.002
Lottery FE		✓		✓	✓	✓
Probability			✓	✓	✓	✓
State FE						✓
Group FE					✓	
PanelB						
PT	0.0006	0.0010	0.0002	0.0006	0.0001	0.0003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Selection Prob			3.0889***	3.1315***	1.2639	1.0018
			(1.072)	(1.076)	(5.218)	(1.356)
Observations	35,618	35,618	35,618	35,618	35,618	35,618
R-squared	0.000	0.001	0.001	0.002	0.004	0.004
Lottery FE		✓		✓	✓	✓
Probability			✓	✓	✓	✓
State FE						✓
Group FE					✓	

Notes: Likelihood of audit on party dummy. Selection Prob is the selection probability of the municipality in each lottery. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

10.3 Scraping and Coding Lotteries 3-7

Lotteries 3-7 are problematic to scrape and code because the information in these PDF is not stored in a homogeneous format. This phenomenon does not only affect reports across time but also those within the same state (and within lottery) where the same auditors are investigating and compiling the report. The auditing procedure, however, seems to be similar over time and across states.

In order to maximise the sample size, I include the total amount of irregularities found by the auditors and create a count of severe irregularities. Provided that these documents do not have a summary indicating the irregularities found by the auditors, I try to locate the paragraph where the irregularity is described.[70]

In most of the cases, the paragraph can be located by searching words like "Constatação da Fiscalização" (Audit Findings). In some cases, a phrase summarising the paragraph is included before the irregularity description. Figure 15 provides an example. Usually the summary phrase is followed by a word such as "Fato" (fact) and "Evidencia" (Evidence – brought by the auditors).

[Figure 15 Here]

I am able to systematically count the number of times a line starts with "Constatação da Fiscalização" and how many times a line starts with "Fato". Having two different ways of locating the paragraph where the irregularity is included allows me to reduce the measurement error in the variable. Both approaches point to a similar result, allowing me to get the total amount of irregularities encountered by the auditors in the municipality.[71]

The second step is to get the irregularity count from these reports. For this purpose, I start by extracting the first three sentences in the paragraph where the auditors describe the irregularity. Next, I again follow two different approaches. First, I classify the text with the same method in Section 3.1.1. Second, I use a Naive Bayes classification technique.

The classifier is based on Bayes' Theorem with the assumption of independence among

[70] This is done again, by regular expressions.

[71] This variable may be prone to measurement error especially belonging to lotteries 3 and 4. The reason being that in earlier lotteries the auditors were not only auditing every fund transferred, they were also writing about all findings, both positive and negative. From lottery 5, auditors only point out irregularities found. Namely, they only point out only irregularities, or negative facts, which, depending on the situation, may or not be improved by the municipal management.

predictors.[72] Regardless of the assumption, the classifier has a very accurate performance. I am able to use the Naive Bayes because the training sample, automatised classified phrases from the reports 8-38, is large enough and the out of sample prediction is similar in magnitude. When working with the Naive Bayes classifier, the text pre-processing is different.[73] First, stopwords need to be removed from the text. Stopwords are common words in a language. In English, the equivalent would be: "of", "to" , "on". I do so mainly using the packages in Natural Language Toolkit (NLTK) for Portuguese. The library works well in English but it has some limitation for other languages such as Portughuese. For this reason, I include another list of stopwords, which I create myself, from the most commonly occurring words among the PDF reposts. The classifier, as expected, provides results similar to the first methodology. The performance of the Naive Bayes is in Table A4. The test on the classifier indicates that it is a good one; both in-sample and out-of-sample forecasts have a good precision, recall, and f1-score.

[Table A4 Here]

[72] A Naive Bayes classifier assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature.

[73] Some steps are similar such as accent removal and conversion of upper case letters to lower case letters.

Table A4: Naive Bayes Classifier

In sample forecast				
	precision	recall	f1-score	support
0.0	0.94	0.88	0.91	52936
1.0	0.82	0.91	0.86	32129
avg / total	0.90	0.89	0.89	85065
Out of sample sample forecast				
0.0	0.87	0.94	0.90	17269
1.0	0.90	0.81	0.85	12504
avg / total	0.88	0.88	0.88	29773

Notes: In-sample forecast refers to the performance of the classifier in the training sample. Out-of-sample forecast refers to the performance of the classifier in the test sample. In sample forecast refers to the performance of the classifier in the training sample. Out of sample forecast refers to the performance of the classifier in the test sample. Precision (or positive predicted value) and recall (or sensitivity) are, respectively. the fraction of retrieved instances that are relevant and the fraction of relevant instances that are retrieved. Both precision and recall are based on an understanding and measure of relevance. F1-score (also F-score or F-measure) is a measure of a test's accuracy. It considers both the precision p and the recall r of the test to compute the score. The f1-score can be interpreted as a weighted average of the precision and recall, where an f1-score reaches its best value at 1 and its worst value at 0. Support stands for the number of samples that support the classification. In this particular case, support counts the number of phrases that support the classification.

Figure 15: (Slight) Differences in audit reports

Panel A: Example of an Audit Report with a summary phrase before the full description of the irregularity.

1.1)Constatação da Fiscalização:

Deficiência no pagamento automático da Bolsa Criança Cidadã, realizado via cartão magnético.

Fato

Encontrou-se o programa em questão em bom andamento no Município no que tange a sua gestão. O Conselho Municipal de Assistência Social tem uma participação ativa tornando, desta forma, as ações sociais de fácil acesso à comunidade. Quanto à Comissão Municipal de Erradicação do Trabalho Infantil – CMETI, esta foi regularmente criada e encontra-se em funcionamento.

Panel B: Example of an Audit Report without a summary phrase before the full description of the irregularity.

1.1)Constatação da Fiscalização:
Fato:
Os veículos utilizados pelo Município nas atividades de endemias são insuficientes em número para atender as necessidades do programa e não são adequados pelos seguintes fatos: são veículos de capacidade reduzida para conduzir toda a equipe, para fazer o transporte do material a ser utilizado e por não chegarem a áreas de difícil acesso pelo fato de não ter tração.

Evidência:
O fato foi evidenciado por meio de inspeção "in loco" dos veículos e entrevista junto ao pessoal envolvido no programa.

10.4 Correlation with other measures in the literature

I perform a second test on the corruption measure using two independently coded measures of corruption. The first one is the coding by Ferraz and Finan (2011), henceforth FF (available on the AER webpage) and the data by Garfias and Timmons (2015), GT henceforth.[74] The second measure was coded similar in spirit to the measure by FF. I match these data and use the subsample where the coders mainly agree. This is the subsample where the coders attribute to the same municipality the same number of corrupt activities or they disagree by 1 corrupt activity. The correlation among both measures in this subsample is of 0.99. To test the performance of my classifier, I regress it on the corruption measure by FF and GT conditioning for state fixed effects. The rationale for including state fixed effects is due to the fact that my measure depends heavily on the language used by the auditors which varies across states (in different states there are different teams of auditors). I report the partial correlation coefficient and the R-squared of the regression. I also present a binned scatterplot of corruption index and the measure of corrupt activities in FF and the overlaid regression line.[75]

[74] I thank Claudio Ferraz, Frederico Finan, Francisco Garfias and, Jeffrey Timmons for sharing their data with me.

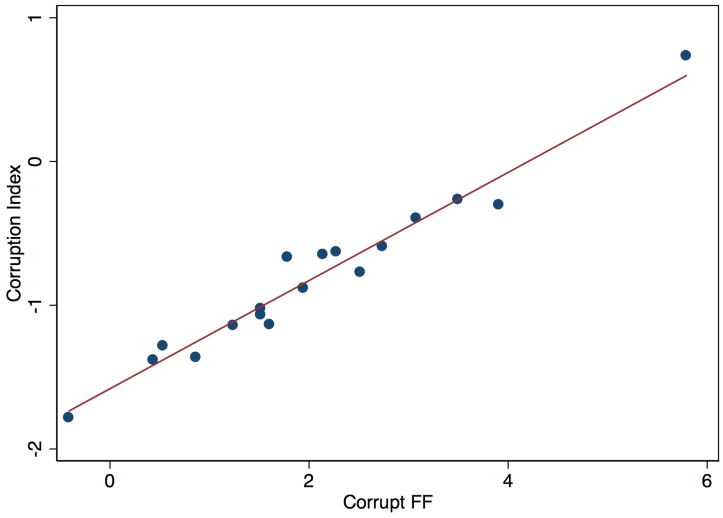
[75] These graphs were produced using binscatter, a user-written Stata command written by Michael Stepner, with input from Jessica Laird and Laszlo Sandor.

Table A5: Correlation with FF and GT

	(1) Corrupt. Index	(2) Corrupt. Index	(3) Sd Major Irreg	(4) Sd Major Irreg
Corrupt FF	0.376***		0.274***	
	(0.041)		(0.025)	
Corrupt GT		0.352***		0.260***
		(0.045)		(0.033)
Observations	99	99	99	99
R-squared	0.751	0.725	0.730	0.704

Notes: Partial correlation of the corruption index, in columns (1) and (2), and the standardised measure of irregularity count, in columns (3) and (4), with the measures of corrupt activities as in Ferraz and Finan (2011) and Garfias and Timmons (2015).

Figure 16: Correlation with FF and GT



Notes: Binned Scatterplot of the Corruption Index and Corruption Activities (as in FF and GT).

10.5 Exogeneity Check (Geographical Spillovers)

Given the randomization in audits, both groups should be homogeneous. There should be no major differences between the treatment and the control group. To assess this, I test the validity of the exogeneity assumption by using predetermined municipal characteristics. I run the following specification:

$$\text{Audit}_{pit-s} = \lambda_0 + \lambda_1 x_j + \eta_c + pop_{2000} + \varepsilon_{pjt} \tag{7}$$

In this case, i is the audited municipality and j corresponds to the informed neighbor. Audit_{pit-j} is an indicator taking value one in case there was an audit in a nearby municipality i at time s previous to elections and value zero in case the audit in town i occurred after elections. x_j is a characteristic of municipality j, measured in the year 2000. η_c state fixed effects and pop_{2000} are included in each pairwise regression given that they are determinants of the eligibility of the municipality. λ_1 is expected to be zero in each regression. This is indeed the case for most of the municipal characteristics; however, there are some differences in the presidential sample. These differences vanish if instead of controlling for state fixed effects and population, I control for the selection probability of the municipality.

If I follow the latter strategy, it requires for the sample to be restricted to the municipalities that are eligible for selection in the same lottery as their (audited) neighbor. This reduces the samples by hundreds of observations (municipal and presidential), thus the samples will not coincide with those used for the main analysis.[76] This test provides further credibility to the identification strategy. I present results both with and without the vector of municipal characteristics as controls. Similarly to the specifications used for the geographical spillovers, I cluster the standard errors at the mesoregion area. These results are shown in Table A6.

[Table A6 Here]

[76] Other possibilities would be to control for the probability of selection in the same year or the same electoral term as the audited neighbor. In these cases, an average of the selection probability must be attributed given that the municipality is eligible for selection various times within the same year. These results are available upon request.

Table A6: Exogeneity Checks neighbors

	Municipal	Presidential
Corrupt. Index	-0.0272	-0.0168
	(0.0321)	(0.0247)
Literacy Rate	-0.0030	0.0092
	(0.0039)	(0.0055)
Frac. Urban	0.2288	0.1540
	(0.2599)	(0.1767)
GDP per Capita	2.1871	-2.0610
	(2.7150)	(3.0867)
Income per Capita	1.2777	1.6207
	(1.4676)	(0.9219)*
Electricity TV	-0.0016	0.0040
	(0.0019)	(0.0027)
HDI	-0.9477	1.0830
	(1.1132)	(0.7097)
Income per capita	-0.0005	0.0005
	(0.0007)	(0.0005)
Coverage	-0.0491	0.1525
	(0.0743)	(0.0671)**
Latitude	-0.0163	-0.0145
	(0.0329)	(0.0187)
Longitude	-0.0604	-0.0268
	(0.0596)	(0.0171)
Dist. State capital	0.0003	0.0003
	(0.0004)	(0.0002)
Dist. Federal capital	0.0002	0.0003
	(0.0003)	(0.0002)
Surface	0.0000	0.0003
	(0.0000)	(0.0002)
Electorate	-0.0000	-0.0000
	(0.0000)	(0.0000)
Residual Vote	-0.9491	-1.9190
	(1.3666)	(0.9970)*
Munic. Council	0.0030	0.0016
	(0.0092)	(0.0050)
PT	0.0419	-0.1792
	(0.0567)	(0.0889)**
PSDB	0.0222	
	(0.1106)	
PMDB	-0.0775	
	(0.0761)	
PP	-0.0878	
	(0.1177)	
PFL	0.0701	
	(0.0812)	

Notes: Coefficients and standard errors reported in this table result from a regression in which the dependent variable is an indicator taking value one in case the neighbor was audited before elections and zero after elections. The regressor is defined by each row. In the regression are also included population and state fixed effects as controls. Standard errors are clustered at the mesoregion area. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.

10.6 Exit of parties

To investigate for potential differential exit of parties, I use a linear probability model where I regress the exit dummy on the audit dummy and other controls. The exit dummy takes value one in case the party does not show up in upcoming elections (and zero otherwise). The results are shown in Table A7.

Columns (1) - (4) present results for the audited municipality whereas those from (5) - (8) are related to exit of the audited party in the case of the geographical spillovers. The interaction of the treatment with the corruption index is shown in columns (3), (4) for audited municipalities and (7), (8) for neighboring municipalities. Across columns, the coefficient on the audit dummy is not statistically different from zero and this also applies to the interaction term. These results indicate that there is no differential exit of parties.

Table A7: Party Exits

	Audited Municipality				Geographical Spillovers			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Audit	-0.038	-0.030	-0.037	-0.031	-0.016	-0.014	-0.016	-0.016
	(0.024)	(0.024)	(0.024)	(0.024)	(0.045)	(0.046)	(0.046)	(0.047)
Audit * Corrupt. Index			-0.006	-0.003		-0.010		-0.010
			(0.012)	(0.013)		(0.018)		(0.019)
Corrupt. Index			0.008	0.002		0.010		0.008
			(0.010)	(0.011)		(0.017)		(0.017)
Observations	1,834	1,834	1,828	1,828	898	898	898	898
R-squared	0.140	0.175	0.141	0.175	0.142	0.143	0.165	0.165
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Party FE	✓	✓	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓		✓

Notes: All regressions include state fixed effects, time fixed effects and also population as of the year 2000. Columns (1) - (4) show results in the audited municipality. Columns (5) - (8) show results from municipal elections in neighboring municipalities. Controls include: income per capita, fraction of the urban population, fraction of the male population, and the literacy rate, and surface area, latitude and longitude, distance to state capital, Human Development Index, and electricity in the household. Standard errors are clustered at the municipality level in case of the audited municipality and at the mesoregion area in the case of geographical spillovers. *** Significant at the 1 percent level, ** Significant at the 5 percent level, * Significant at the 10 percent level.