Information About Corruption and Politicians' Proposals*

Guillermo Lezama†

[CLICK HERE FOR THE LATEST VERSION]

October 8, 2023

Abstract

This paper explores the impact of information about corruption on politicians' proposals. Using text analysis on 13,344 manifestos from the 2012 mayoral elections in Brazil, this study examines how revealing corruption through audits of public funds influenced discussions on the policy areas under scrutiny. The results indicate that the disclosure of these irregularities led to an increased discussion of the policy areas covered by the audit by opposition parties in high-corruption cities. However, incumbents in high-corruption cities respond to this disclosure, decreasing the discussion of the policy areas covered by the audit. In high-corruption municipalities, disclosing irregularities made incumbents to employ more populist language. In municipalities with high corruption, disclosing irregularities caused incumbents to use more populist language. Meanwhile, opposition candidates in low-corruption municipalities adopted less extreme and populist language than their counterparts in non-audited, low-corruption municipalities. These findings contribute to our understanding of how information on corruption in public expenditures influences politicians' agendas and the ideological framework of their proposals.

JEL codes: D72, P16, H7

*Please do not cite or circulate without author's permission.

†Department of Economics, University of Pittsburgh. guillermo.lezama@pitt.edu

1 Introduction

In electoral campaigns, politicians have the opportunity to address pressing issues and offer voters information and policy proposals that can shape the future of their communities. However, this period also presents a unique occasion for politicians to promote populist and extreme narratives, often divorced from facts. Misinformation, populist rhetoric, and hateful communication are increasingly prevalent in electoral campaigns, with significant real-life consequences. Consequently, there is a growing interest in understanding how politicians select and communicate their messages in response to these challenges. An unexplored factor that could significantly influence the politicians' agenda and the political climate is the presence of corruption scandals. Examining the role of exposing corruption and other governmental information in electoral campaigns becomes crucial in understanding the interplay between institutional oversight, electoral accountability, and the increasing trend of populist rhetoric.

While the benefits of revealing irregularities in the use of public funds, such as aiding in the removal of corrupt politicians and reducing future corruption, are well-established (e.g., Avis, Ferraz, & Finan, 2018; Ferraz & Finan, 2008), relatively less attention has been paid to the impact of unveiling these irregularities on the electoral campaign and political discourse. In other words, our understanding of how government audits specifically affect politicians' agendas and the utilization of populist rhetoric remains limited. On the one hand, government audits can impact politicians' agendas by revealing problems in the government that need to be addressed. The availability of information about government activities has been found to correlate with the distribution of policy issues discussed in politicians' manifestos (Abou-Chadi, Green-Pedersen, & Mortensen, 2020; Seeberg, 2022; Williams, Seki, & Whitten, 2016). On the other hand, corruption is often associated with the use of populist rhetoric (Berman, 2021; Mudde & Rovira Kaltwasser, 2018). Consequently, these scandals can also influence the content of politicians' communication and the extent to which populist rhetoric is employed.

This paper examines the impact of exposing information about corruption cases on the public sector on the proposals put forth by politicians during an election. Specifically, this study addresses two key questions. First, it explores whether politicians adjust their agendas in response to reports that uncover irregularities. I explore whether these changes indicate responsiveness to such scandals by addressing the policy areas involved in the corruption scandals and using vocabulary used in the audit reports. Second, the paper investigates whether these scandals can contribute to the rise of extremism and populism. Previous research has demonstrated that reputational shocks can impact ideological positioning (e.g., Bernhardt, Buisseret, & Hidir, 2020; Buisseret & Van Weelden, 2022; Groseclose & Milyo, 2005; Serra, 2010). Revealing corruption reveals information about the incumbent's type that can be viewed as a negative valence or reputation shock. Therefore, this study examines how these shocks influence the choice of policy dimensions and assess their potential role in fueling ideological extremism (moving away from the center on the left-right axis) and populism (using anti-elite rhetoric (Gennaro, Lecce, & Morelli, 2021)).

Municipalities in Brazil provide an ideal context for analyzing these issues. In 2003, Brazil implemented an anti-corruption program to enhance government transparency and reduce corruption, which yielded successful results. The assignment to be audited or not was determined by a lottery system. These audits generated a flow of data, disseminating information about public affairs to politicians, including incumbents, candidates, and the electorate. As a result, they caused reputation shocks, either detrimental when uncovering high levels of corruption or beneficial when revealing low levels of corruption. The identification strategy in this study comes from Ferraz and Finan (2008) and relies on utilizing the random variation that arises when municipalities undergo audits. I exploit the timing variation of the audits, whether conducted before or after the 2012 election, to compare municipalities with comparable corruption levels but differing in their awareness of the audit report prior to the election. Additionally, candidates for the mayor position in Brazil are required to submit a manifesto before the election, which serves as a valuable source for understanding politi-

cians' priorities and communication strategies (Cagé, Le Pennec, & Mougin, 2022; Catalinac, 2018; Le Pennec, 2022; Williams et al., 2016). Overall, I analyze whether audit information influenced the formulation of the manifesto.

The results of this paper clearly show that audits and the associated information about irregularities influence the content of political manifestos. When candidates are informed about an audit report before an election, they are more likely to use terms from that report in their manifestos than if they learn of it afterward. However, incumbents tend to avoid or reduce discussion on areas where the audit report indicates high corruption, while challengers tend to emphasize those very areas. This pattern indicates that while audit reports guide candidates' proposals, candidates also tactically choose topics that they believe will strengthen their position in campaigns, in line with existing theoretical literature (Riker, 1996). Next, it is demonstrated that an audit in high-corruption cities leads to the incumbent adopting a more populist approach. The frequency of words related to populism is analyzed in candidates' proposals using specific dictionaries (Gennaro et al., 2021; Mendes, 2021). A supervised method is employed to assess the level of extremism and partisanship in manifestos, following the approach of Le Pennec (2022), where labels are assigned based on the known political orientation of Brazilian political parties. In cities with low levels of corruption, the disclosure of the audit report, which represents a positive reputation shock for the incumbent mayor, leads the challenger to adopt a manifesto that utilizes less of a populist rhetoric and less ideologically extreme vocabulary.

The findings of this paper demonstrate that politicians respond to corruption scandals by modifying their agendas, vocabulary, and rhetoric. These effects reveal that revealing corruption can influence electoral accountability and politicians' priorities in their communication. However, the study also shows that politicians are strategic in their responses, addressing topics they perceive as an advantage. Additionally, it is revealed that increased transparency can have unintended consequences, as incumbents tend to increase their use of populist language and exhibit higher levels of extremism after an audit that unveils a

significant number of corruption cases. In summary, audits concerning the use of public funds shape the communication strategies of politicians by influencing the topics addressed and the ideological positioning reflected in their proposals.

Contributions. This paper speaks to several strands of the literature. First, it relates to previous studies demonstrating that politicians react to disclosing information about them in terms of effort and electoral choices (Cavalcanti, Daniele, & Galletta, 2018; Poblete-Cazenave, 2021; Snyder Jr & Strömberg, 2010, e.g.). This paper explores how reputation shocks from new information affect candidates' agendas and the political positioning of their proposals.

Second, this study contributes to the growing body of research on the factors driving the rise of populism and extremism. To the best of my knowledge, the effect of changes in the informational environment on populism has not been previously studied. Additionally, this study shows how the supply of populism differs in a high-corruption from a low-corruption setting. Berman (2021) reviews the literature about the causes of the rise of populism and mentions the role of corruption in it. This paper explores voluntarist theories that suggest populism is a deliberate choice made by politicians and parties to attract more votes (e.g., Gennaro et al., 2021). Regarding extremism, this paper contributes to the theoretical literature that analyzes how politicians respond with their policy positions and political agenda to changes in reputation by conducting empirical tests (e.g. Bernhardt et al., 2020; Buisseret & Van Weelden, 2022; Dragu & Fan, 2016; Serra, 2010).

Third, this paper expands on the existing literature that employs text analysis on speeches or political manifestos to address different questions in political science (e.g. Cagé et al., 2022; Catalinac, 2018; Crabtree, Golder, Gschwend, & Indriason, 2020; Gennaro et al., 2021; Le Pennec, 2022). It investigates a new factor impacting communication strategies - information about government actions, specifically audits. The study finds that audits can influence the topics discussed in subsequent elections and affect the usage of populist and extremist content, particularly in the case of a negative reputation shock.

Finally, this paper is similar to previous research examining the effects of disclosing information about government actions, specifically audits. These studies have shown that audits can have an impact on corruption, economic activity, and hiring practices Avis et al. (2018); Colonnelli and Prem (2020); Ferraz and Finan (2008, 2011); Gonzales (2021); Lauletta, Rossi, and Ruzzier (2020). For instance, Amorim (2022) found that transfers for the health sector from the federal government decreased in municipalities with a high number of irregularities after being audited. This study builds on this existing body of work by examining audits' impact on politicians, incumbents, and challengers. Specifically, it looks at the effects on local political leaders' political positioning and agenda, which could have medium- and long-term effects.

The paper is structured as follows: Section 2 introduces the institutional context and data. Section 3 discusses the potential mechanisms in the interplay between the audits and the manifestos' content. Section 4 outlines the empirical approach. Section 5.1 empirically evaluates the impact of audits on politicians' agendas. Section 6 empirically explores the effect of audits on politicians' rhetoric. The papers' conclusions are in section 7.

2 Background and Data

2.1 Brazilian Anti-Corruption Audit Program

The Brazilian Federal Government initiated the random audit program in 2003, which was later transformed into a new program with different characteristics in 2015. The program aimed to monitor local municipalities' use of public funds through the State Comptroller (Controladoria Geral da Uniao; CGU) by auditing municipalities chosen through a public lottery. All municipalities with a population under 500,000 were eligible to participate. Once selected, 10-15 CGU auditors would visit the municipality for 1-2 weeks to gather information and documentation on their use of funds over the past 3-4 years. The auditors then compiled a report that was sent to the city council and prosecutors and published it on the CGU's

website.

Over 40 editions, 1955 municipalities were audited in 2180 inspections. The probability of being audited varied for mayors depending on the state and the audit period (Avis et al., 2018). The number of lotteries per year and municipalities audited on each lottery changed over time. From lotteries 28 to 33, smaller municipalities were audited in all sectors. In contrast, larger ones were audited in specified sectors (see Table A.1 to see the areas that were audited according to the municipality's population). After lottery 34, all municipalities were audited in designated sectors at the time of the lottery.[1]

Researchers have used the data collected from the random audit program in Brazil to study its impact on corruption and other variables. The findings showed that the audits influenced election outcomes (Ferraz & Finan, 2008) and helped reduce corruption levels over time (Avis et al., 2018). Other studies, such as those by Ferraz and Finan (2011), Brollo and Troiano (2016), Colonnelli and Prem (2020), have also used the data to investigate the effects of the audits.

In my analysis, I use the CGU's audit data to evaluate the impact of unveiling corruption about the use of public funds on candidates' proposals. I focus on municipalities audited before the 2012 election (from lotteries 28 to 35). To measure the effect of audits, I use CGU data that shows the level of corruption found in each audit. I classify acts of moderate or severe corruption as corruption cases, following the methodology of Avis et al. (2018), grouping both kinds of irregularities into one category. For each audit, I also have data on the number of irregularities found in each sector (Table A.2). This allows me to examine the impact of audits on election results while considering the number of corruption cases. Finally, I only include municipalities with a population of less than 500,000 and exclude capital cities in my sample to ensure that the audited and non-audited municipalities are similar.

[1]Avis et al. (2018) and Ferraz and Finan (2008) make a complete description of this program.

2.2 2012 Municipal Elections and Municipality Characteristics

Brazil is comprised of 5,568 municipalities that are responsible for providing essential services, such as water, sanitation, health, and education, among others. Characteristics data for these municipalities were obtained from the 2011 Pesquisa de Informações Básicas Municipais (MUNIC) survey conducted by the Statistics and Geography Institute (IBGE). The mayors of these municipalities are elected every four years in October elections, which also include elections for vice-mayors and city councilors. In cities with a population over 200,000, a second round of elections is only held if no candidate receives more than 50% of the valid votes in the first round.

Table 1: Descriptives: Issues and Ideological Content

	Rounds 28 - 35 (2009-2012)		Rounds 36 - 38 (2012-2013)		Other	
	Mean	SD	Mean	SD	Mean	SD
Age	48.70	10.53	48.13	10.35	48.58	13.66
% College Studies	0.55	0.50	0.58	0.49	0.56	0.50
% Women	0.14	0.35	0.13	0.34	0.13	0.34
% Same party as the President	0.10	0.30	0.11	0.32	0.12	0.32
% Same party as the Governor	0.15	0.35	0.13	0.33	0.15	0.36
Running for re-election	0.18	0.38	0.19	0.39	0.18	0.38
Number of Candidates	3.38	1.50	3.39	1.45	3.29	1.46
Observations	1129		423		11548	

Notes: This table shows means and standard deviations for candidates' characteristics according to whether the municipalities where they are running were audited or not. Only candidates for which a manifesto was retrieved are considered. Data from TSE.

The data for the 2012 municipal elections in Brazil was obtained from the Superior Electoral Court (TSE). This data also provided information about the candidates' characteristics and showed that the candidates in both audited and non-audited municipalities were similar. In these elections, it is common for more than two parties to compete, and national parties often form coalitions to support a mayoral candidate. It is important to note that mayors in Brazil can only be re-elected once for consecutive terms. Table 1 shows the means of candidates' characteristics, which overall are similar in both audited and non-audited municipalities. 31 parties have candidates. In the 2012 municipal elections, three parties had more

than 10% of the total number of candidates: the Brazilian Social-Democratic Party (PSDB) with 15%, the Workers' Party (PT) with 12.3%, and the Brazilian Democratic Movement (MDB) with 10.8%.

Table 2 and Table 3 show means and standard deviations for municipal characteristics. Due to the empirical design, there should be a balance between the municipalities audited before the election and the control groups (municipalities not audited during 2009-2012). The results indicate that the municipalities do not differ significantly based on these variables between the group of those audited between 2009 and 2012 and those not audited in that period (Table 2), and between the group of those audited between 2009 and 2012 and those audited in after the 2012 election before 39th round (Table 3).

Table 2: Mean Comparisons between Audited and Nonaudited Municipalities

	Control	Treatment	Difference
GDP pc	12886.52	10805.9	-819.3
	[14487.15]	[9571.88]	[569.971]
Share Illiterate (%)	85.3472	83.51522	-0.0758
	[8.86]	[9.33]	[0.258]
Share Urban	0.6374198	0.6262824	0.00446
	[0.22]	[0.21]	[0.007]
Share Secondary Education and above	0.2156972	0.2083799	0.000671
	[0.08]	[0.08]	[0.003]
Share of Bureaucrats with Superior Education	0.3069009	0.2967605	-0.00219
	[0.11]	[0.11]	[0.004]
HDI	0.6598012	0.6443488	-0.00108
	[0.07]	[0.07]	[0.002]
AM radio	0.2092931	0.1987315	-0.00137
	[0.41]	[0.4]	[0.025]
Gini	0.5013802	0.5095829	-0.0014
	[0.07]	[0.06]	[0.002]
Population (logs)	9.377024	9.470213	0.00526
	[1.09]	[1.1]	[0.032]
Audited Previously	0.2499018	0.2635983	-0.0113
	[0.43]	[0.44]	[0.021]
Observations	5090	478	

Notes: Estimates are means and standard deviations (in brackets) of various municipal characteristics by places that have been audited in the period 2009-2012 (Treatment) and places that have not been audited in that period (control). The difference and corresponding standard error (in brackets) are computed on the basis of a regression that controls for state.

Table 3: Mean Comparisons between Audited and Nonaudited Municipalities

	Control	Treatment	Difference
GDP pc	11700.37	10805.9	-483.2
	[14514.56]	[9571.88]	[1320.197]
Share Illiterate (%)	84.28	83.52	-0.16
	[8.84]	[9.33]	[0.546]
Share Urban	0.64	0.63	0
	[0.22]	[0.21]	[0.014]
Share Secondary Education and above	0.22	0.21	0
	[0.08]	[0.08]	[0.005]
Share of Bureaucrats with Superior Education	0.31	0.3	-0.01
	[0.11]	[0.11]	[0.009]
HDI	0.65	0.64	0
	[0.07]	[0.07]	[0.004]
AM radio	0.2	0.2	0
	[0.4]	[0.4]	[0.041]
Gini	0.5	0.51	0
	[0.06]	[0.06]	[0.004]
Population (logs)	9.41	9.47	0.07
	[1.11]	[1.1]	[0.053]
Audited Previously	0.24	0.26	0.03
	[0.43]	[0.44]	[0.043]
Observations	165	478	

Notes: Estimates are means and standard deviations (in brackets) of various municipal characteristics by places that have been audited in round 28 to 35 (2009-2012) (Treatment) and places that have been audited in rounds 36 to 38 (control). The difference and corresponding standard error (in brackets) are computed on the basis of a regression that controls for state.

2.3 Party Manifestos

According to the electoral law, candidates running for mayor must submit their manifestos before the election. This has been a requirement since 2009 for all mayoral, gubernatorial, and presidential candidates. Consequently, the manifestos from the 2012, 2016, and 2020 municipal elections can be found on the Electoral Authority (Tribunal Superior Eleitoral - TSE) website, various news outlets, and each candidate's personal websites.

The dataset for this study was constructed by obtaining the manifesto documents in PDF format from the TSE website. For the 2012 election, 16,173 documents were uploaded, of which 13,724 were retrieved. After preprocessing, 13,344 texts from 5,140 municipalities were available (the preprocessing steps and exclusion criteria are explained in the appendix). Each manifesto was linked to the candidate data using a unique identifier obtained during data scraping. The average number of words per manifesto was 2150 (Table 4).

2.3.1 Overlap Between Audit Reports and Manifestos

In order to measure the extent to which the vocabulary of the audit report is reflected in the party manifesto, an analysis of the overlap between both texts is conducted. To accomplish this, a list of all words present in a party's proposal (excluding stop words) is generated, and the same process is repeated for each manifesto. For each candidate, the proportion of words in the audit report that also appear in their manifesto is computed, thereby serving as the measure of overlap. For each manifesto j, in a municipality with an audit-report i, this measure is computed as:

$$\frac{\sum_{w \in d} \mathbb{1}[w \in d \bigcap j]}{m_d} \tag{1}$$

where w is each word, and m_j is the total number of words on document d.

This measure serves as an initial quantitative approach to assess the degree to which politicians incorporate the findings of an audit report into their electoral campaigns. It offers an objective and transparent measure of content utilization, contributing to the evaluation of the alignment between campaign proposals and the outcomes of the audit.

2.3.2 Proposals' Topics

The manifestos were organized into topics using headings. The entire corpus consisted of 3,445,957 lines (on average, each line contained 7.6 words). Thus, to determine which topic each line discussed, a Multinomial Naive Bayes classifier was used to assign probabilities to each line of text into ten topics: six specific (Bureaucracy, Social, Health, Urban, Economics, and Crime), two general (Titles and Introduction/Other), and one residual (Unrecognizable words). Examples of headings for each topic can be found in Table A.3.

A random sample of 100 manifestos (1%) was manually classified into topics. I follow some preprocessing steps to train and fine-tune the model and then apply it to the entire corpus, as described in Appendix B. The final policy issues' distribution is presented in

Table A.4.

Classifying each line regarding the policy issue they discussed involved creating a vector of word frequencies for each line L, with frequency 0 if the word was not in the line. The algorithm used a Multinomial Naive Bayes classifier to calculate $P(w|C_k)$ for each word w and topic C_k [2]. This machine learning algorithm, commonly used for text classification, uses the Bayes theorem and the independence assumption (naive assumption) between features. Each line L is considered a bag of words, with each word as a feature. The algorithm calculates the likelihood of L given each topic C_k based on the presence of certain words. Let

$$P(L|C_k) = \frac{\text{prior} \times \text{likelihood}}{\text{evidence}} = \frac{(\sum_{i=1}^{n} w_i)! \times \prod_{i=1}^{n} p(w_i|C_k)}{\prod_{i=1}^{n} w_i!} \tag{2}$$

I used an 80-20 train-test split. I calculated two predictions: a soft prediction and a hard prediction. The soft prediction calculates the probability of each line being classified into each topic. After computing $P(L|C_K)$ for each line L, I calculated π_K for each document d by multiplying these probabilities by the number of words in the line and dividing it by the total number of words in the document. d_k measures the predicted share of the document that discusses topic k.

$$d_k = \sum_{L \in d} P(L|C_K) \times \frac{\mid w \in L \mid}{\mid w \in d \mid} \tag{3}$$

where $\mid . \mid$ indicates the cardinality of the set.

I assigned each line to the topic with the highest probability for the hard prediction. I created a binary variable for each topic to determine whether the line belongs to topic k. Then, I multiply it by the share of words in document d corresponding to that line.

$$\omega_k = \sum_{L \in d} \mathbb{1}[k \in \arg\max_k P(L|C_K)] \times \frac{\mid w \in L \mid}{\mid w \in d \mid} \tag{4}$$

This method provides an accuracy of 62% when looking at the test set's classification. As

[2] Long short-term memory networks (a deep learning algorithm), and K-neighbors were also used. With different parameters, these algorithms provided lower accuracy rates and higher log-loss.

a benchmark, if topics were chosen randomly (using the distribution on the sample), the accuracy would be 16.9%.

Table A.4 shows the distribution of the share of the document dedicated to each topic.

2.3.3 Partisanship and Extremeness

Extremeness

To measure to what degree a candidate's manifesto is similar to the other manifestos of that candidate's party, I follow Le Pennec (2022) and compute partisan scores for each document. This paper is built on the *Wordscores* method (Laver, Benoit, & Garry, 2003). To do that, I first labeled each party on whether they are left-wing, right-wing, or center as explained in subsection D.1. The approach uses the word counts in each document.[3] I computed the frequencies p_w^R and p_w^L that represent how frequent a word w is in all the manifestos in the left or right

$$p_w^i = \frac{\sum_{j \in i} c_{wj}}{\sum_{j \in i} m_j}$$

where c_{wj} is the counts of word w in document j, and m_j is the total number of words of document j.

Using these frequencies, I can compute the right-wing score of each word w

$$s_w = \frac{p_w^R}{p_w^R + p_w^L} - \frac{p_w^L}{p_w^R + p_w^L} \tag{5}$$

A word only used by right-wing parties will receive a score of 1, while a word only used by left parties would get a score of -1. In Table A.5, we can see the words that received the highest scores for the right-wing and left-wing parties. The ones on the left show language usually related to parties with a socialist ideology (socialism, capitalist, deliberative, dominant, etc.), plus specific words such as petistas (member of the PT). On the right-wing side, there are several words related to party names, such as democrats, and Christian.

[3] In computing manifestos' score, I applied the same steps as described in appendix Appendix B. I also excluded any word present in more than 95% of the manifestos or in less than 0.5% of the manifestos.

A manifesto j score is then calculated by:

$$S_j = \frac{\sum_w p_{wj} \times s_w}{S_R} \tag{6}$$

where S_R is the score of the aggregation of all the manifestos of the right-wing group.[4]

S_j is the positioning of a party in the left-right axis. I compute $\mid S_j \mid$ as a measure of Extremeness.

Partisanship

To measure Partisanship, instead of defining three labels (left, right and center), I define a label for each party. For party t, I compute

$$s_w = \frac{p_w^t}{p_w^t + p_w^{-t}} - \frac{p_w^{-t}}{p_w^t + p_w^{-t}} \tag{7}$$

where p_w^t and p_w^{-t} are the frequencies of a word in the set of manifestos of party t and the set of manifestos that are not t, respectively. s_w^t takes the value of 1 for party t, if it is a word used exclusively by party t, while it takes the value of -1 if used exclusively by all the parties except t. Let j be a manifesto of party T, then,

$$S_{j \in T} = \frac{\sum_w p_{wj} \times s_w^T}{S^T} \tag{8}$$

Finally, to compute partisanship, I removed all the parties that had less than 1.5% of the candidates. This was done to not create much distortion on the word scores.

2.3.4 Populism

I used a dictionary to compute the share of each document with populist content following Gennaro et al. (2021)[5] and Mendes (2021). In the former case, I translated the dictionary

[4]This is done to preserve the distance between the reference texts (Martin & Vanberg, 2008)

[5]Gennaro et al. (2021) discuss how their dictionary represents well the people-vs-elite rhetoric, which is a distinctive feature of populism and the dimension it usually used in the empirical literature to measure populism.

into Portuguese. The final dictionary is available in Appendix C.

To measure the presence of populism, I calculated the tf-idf matrix for each document, which considers the frequency of words and reduces the weight of words that frequently appear in other documents. Then I summed the values of words present in the populism dictionary.

2.4 Descriptives

The average number of words per manifesto is 2150, with a median of 1446, as shown in Table 4. Social policy issues are the most frequently discussed in these proposals, but this is expected given the wide range of topics covered. It is important to note that the scores for extremism, partisanship, and populism lack direct interpretation. Regarding the L-R score (the one that generates the extremeness measure), the median document is close to the score of the reference document on the right. The score of 0 falls at the 36th percentile.

Table 4: Descriptives: Issues and Ideological Content

	Mean	**p50**	**p90**	**se**	**N**
Bureaucracy	0.08	0.07	0.13	0.04	13344
Social	0.36	0.36	0.47	0.09	13344
Health	0.10	0.10	0.16	0.05	13344
Urban	0.12	0.11	0.19	0.06	13344
Economic	0.15	0.15	0.23	0.06	13344
Crime	0.02	0.02	0.04	0.02	13344
Overlap	0.11	0.10	0.19	0.07	1554
Populism	0.05	0.04	0.12	0.06	13706
Extremeness	2.47	2.14	4.77	2.18	13322
Count	2150	1446	4472	2578	13344

Notes: This table shows means, median, percentile 90th, standard deviations and the number of observations for the share dedicated to each topic in manifestos as described in Equation 3, the scores for extremeness, partisanship, and position in the L-R score, the sum of the terms associated to populism in the tf-idf matrix for each document, and word count.

Figure 1 shows the distribution of the left-right score for the group of parties in the left, right, and center of the ideological spectrum. In the appendix, Figure D.2, and Figure D.3 show the distribution for the parties with more candidates on the left, and right, respectively. Figure A.1 shows that variation exists across municipalities and over time (2012 and 2020)

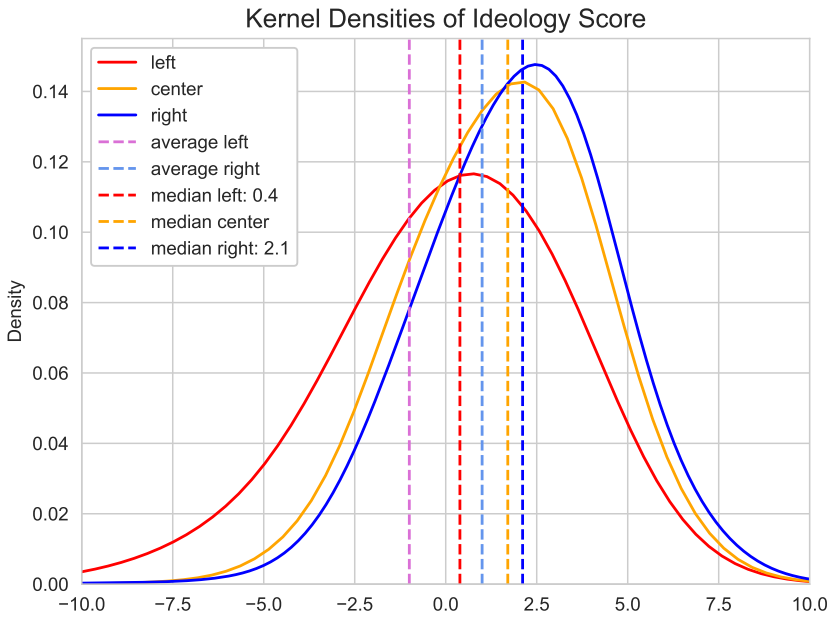


Figure 1: Ideological Scores' Densities for the Left, Center and Right

in the share of how many parties use any populist vocabulary in their manifestos.

3 Information about Corruption and Manifestos

This section aims to provide a brief explanation of why there is an expected effect of the availability of information about irregularities in the government on the proposals outlined in political candidate manifestos. Past audits serve as a proxy for increased knowledge about government actions and irregularities in using public funds in local governments. The effect of the audits could be synthesized through two main channels.

First, the increasing amount of information could have an effect in and of itself (an information shock) by informing the public about government issues (e.g., how much the local government spends on hospital wages). Politicians have responded to audits by changing their practices in government through political selection and a disciplining effect (Avis et al., 2018; Gonzales, 2021; Lauletta et al., 2020). This paper analyzes whether the audit information could affect how candidates frame their proposals and communicate with voters.

Second, in a municipality with a high (low) corruption level, the audit could lead to a negative (positive) reputation shock for the mayor. I look at these effects on the agenda (how much of each policy issue is discussed), and in the ideological content of the language employed. Studies have investigated the impact of an audit's reputational shock on election outcomes (Cavalcanti et al., 2018; Ferraz & Finan, 2008; Poblete-Cazenave, 2021).

3.1 Information about Corruption on Issue Selection

Each manifesto line addresses a specific policy issue, determining the candidate's emphasis on each topic. The distribution of issues in a manifesto likely aims to increase the relevance of those topics in the campaign and influence voters when determining their preferences. There are at least two ways that information about government actions can impact a candidate's agenda.

First, from an electoral accountability perspective, auditing can raise awareness of bureaucratic, administrative, and governance issues among voters, which could prompt candidates to address these issues in their campaigns if the results of the audit were negative. This could also lead to changes in what voters and politicians consider as optimal policies in each area, causing politicians to describe their proposals in greater detail for these issues (Abou-Chadi et al., 2020; Williams et al., 2016). Both incumbent and challenger may address the issue (Seeberg, 2022). This could happen even if there is not a negative reputation shock, and it only provides new information that affects their policy views and voters' demands (e.g. Gagliarducci, Paserman, & Patacchini, 2019). Candidates may also want to reproduce this information in their campaigns to raise awareness among voters. This highlights the role that auditing plays in shaping the policy debate and determining the issues that are discussed during election cycles. This means that audits that find a high number of irregularities in specific areas could lead to candidates discussing these topics more.

Second, the literature on issue selection also addresses the relationship between reputation, electoral advantage, and issue selection. Following Riker (1996), a candidate will appeal

to a specific issue only if they dominate the other candidates in terms of persuasion. A reputational shock resulting from the audit could be seen as affecting the perceived advantage on that policy area. However, Aragonès, Castanheira, and Giani (2015) and Dragu and Fan (2016) show that in specific contexts, parties with a disadvantage in a topic could choose to discuss it more.

In conclusion, the availability of information about government activities, as indicated by audits, can have a significant impact on the distribution of policy issues discussed in political candidate manifestos. The information provided by audits can serve as a source of new policies for candidates and highlight problems that need attention. This may lead to an increase in the attention given to certain topics by all candidates. However, reputational shocks can also influence the extent to which a candidate addresses certain issues in their manifesto. In a setting where corruption cases shape reputation, a positive (negative) reputation shock can result in an increase (decrease) in the attention given to a specific issue by the incumbent candidate, while the opposite is true for their challenger.

3.2 Information about Corruption on Partisanship and Extremeness

Manifestos are placed on the left-right political axis based on their national party affiliation, which allows for the calculation of two measures: partisanship and extremeness. There is a body of literature that examines how changes in reputation can affect policy positioning. An increase or decrease in reputation can lead to a shift toward the center or towards extremism.

Starting from a similar reputation level, an increase in reputation can result from a positive shock for an incumbent or a negative shock for a challenger. If a candidate experiences an increase in reputation, they are likely to adopt a policy closer to their preferred stance, while their opponent moves towards the center (e.g. Serra, 2010). However, after the reputation has increased, the candidate may choose to emphasize this valence advantage to win (Groseclose & Milyo, 2005). As a result, they may converge on their policy position.

This paper is relevant to the context of Bernhardt et al. (2020), as the candidates for the mayor position are running simultaneously with the city council candidates. Even if the mayoral candidates lose, they aim to retain as many votes as possible. If the popularity advantage is small, they adopt a policy the median voter prefers. For moderate advantages, the disadvantaged candidate adopts a policy closer to its core supporters to retain as many seats as possible, while the advantaged candidate does not move towards the other candidate unless the popularity advantage is substantial.

It is important to note that candidates and parties do not always propose their ideal policies and often move towards the center, which can result in a bias towards the center in national party positioning. Therefore, any movement after an audit will also affect partisanship.

In conclusion, the policy choices made by candidates, as measured by their manifestos, may be influenced by reputational shocks, but the direction of this influence can vary. Nevertheless, the results of audits are expected to impact these choices.

3.3 Information about Corruption on Populism

Several papers have reviewed the determinants of populism (Berman, 2021; Guriev & Papaioannou, 2022). They show that different mechanisms can explain the rise of populism.

For example, corruption is often a topic that is associated with populist rhetoric (Berman, 2021). This paper, populism is measured using a dictionary that focuses on how populist politicians use the rhetoric of "us vs. the elite." Audits can make corruption and transparency issues more prominent, leading politicians to respond strategically by incorporating populist rhetoric. This is because we know that politicians supply populist rhetoric based on the demands of their audience (Gennaro et al., 2021). As a result, increasing public awareness of corruption in government may result in an increase in the strategic use of populist rhetoric, particularly by the opposition. However, even the incumbent mayor could increase the usage of this language if the irregularities disclosed are not their responsibility.

Another effect could be through electoral competitiveness. Studies show that negative reputational shocks on the incumbent can lead to increased electoral competitiveness Poblete-Cazenave (2021). In this context, using populist language could serve as a way to differentiate.

Finally, releasing more accurate information to the public may also decrease the use of populist language if it helps prevent the spread of false news (Guriev & Papaioannou, 2022). This is particularly relevant in a municipality with low levels of corruption, where the audit could make it difficult for opposition candidates to accuse the incumbent of being part of the "corrupt elite" (Guriev & Papaioannou, 2022).

4 Empirical Analysis

4.1 Effect of Information about Corruption on Manifestos

This paper investigates the causal impact of information about irregularities in the use of public funds on political discourse and campaign proposals of politicians. To account for the possible differential effect of the audit results, it is important to consider the corruption level found in the audits when analyzing the impact of the audit on political communication. This allows for a more nuanced examination of the relationship between government audits and political discourse.

Following Ferraz and Finan (2008), I exploit the timing of the audit to look at these differential effects. Some of the audited municipalities were drawn close to the election date (October 2012). Thus, the audit measured the number of irregularities in that municipality for the mayor who was in office between 2008 and 2012. Still, the results of it were not available before the election. Thus, I use this set of municipalities as a control group for those municipalities that were audited before the election and for which the audit results were also disclosed before the election or audited closely after the election. The treatment group consists of municipalities drawn to be audited between the 28th and the 35th lottery.

The control group is those audited between the 36th and the 38th lottery. This strategy helps estimate the effect of the audit (disclosing information about the government's actions to the public) conditioning on the level of corruption found.

The claim about estimating the causal effect of an audit comes from the fact that municipalities were randomly drawn into being audited just before or after the election. Before showing the model to estimate, Table 3 shows observables are balanced across both groups as expected. There are not many differences between both groups. I cannot reject the null when I test for joint significance (F-test= 1.05; p-val= 0.4306).

I estimate the following model for incumbents and challengers

$$Outcome_{imst} = \alpha + \beta_0 Disclosure_{mst} + \beta_1 Disclosure_{mst} \times High - Corruption + \tag{9}$$

$$+\beta_2 High - Corruption + \gamma Controls_{imst} + \nu_s + \varepsilon_{imst}$$

where Y_{imst} is the outcome variable for candidate i in municipality m in state s at time t. $Disclosure_{mst}$ is a binary variable that represents if a municipality was audited and the result was disclosed before the election, $High - Corruption$ is a binary variable that represents whether the audit gave a number of acts of corruption cases higher than the median. The vector $Controls_{imst}$ consists of a set of municipal: the GDP per capita (logs), the share of people who are not illiterate, the share of people living in an urban area, if there is an AM radio in the city, the Gini index, population dummies [6], dummies for the number of candidates competing for the mayor position in that municipality, whether the municipality was already audited before the 28th lottery and the total count of words in the document (logs; except when the dependent variable is the word-count). ν_s represents state fixed-effects. For all estimations, I clusterize at the state level.

The dependent variables considered are the share dedicated to each topic, a variable that

[6]The categories are less than 20,000, between 20,000 and 50,000, between 50,000 and 100,000 and larger than 100,000. The thresholds were chosen to be the same that CGU chose in lotteries to assign the topics that the audits would cover

represents the usage of populist words and an index to represent how extreme the ideological score was. We can consider β_1 as the average causal impact of the audit conditional on a low number of irregularities, while $\beta_1 + \beta_2$ measures the average causal impact of the audit conditional on a high number of irregularities.

5 Impact of the information about irregularities on the content of the manifesto

In this section, I investigate whether political candidates adjust their campaign proposals based on the revelation of corruption information in their respective municipalities.

5.1 Incorporation of Corruption Information into Proposals

When corruption cases come to light, they highlight issues that municipalities must address. It is reasonable to assume that politicians would use this information in their campaigns. This subsection assesses whether candidates incorporate terms or topics from the corruption reports released before an election in their campaign pledges. More specifically, I determine if candidates use words from the corruption reports in places where the report was released before the election, compared to places where it was not.

Table 5: Does the Disclosure of the Audit Report inform Manifestos?

Outcome: Overlap Between the Audit-Report and the Manifesto (% of Words on the Audit-Report)

	(1)	(2)	(3)	(4)
Disclosure	0.010***	0.014**	0.009**	0.012*
	(0.003)	(0.005)	(0.004)	(0.007)
High-Corruption x Disclosure			0.006	0.005
			(0.006)	(0.011)
High-Corruption			-0.021***	-0.018
			(0.005)	(0.013)
Candidate	Challengers	Incumbents	Challengers	Incumbents
Observations	459	248	459	248
R-squared	0.850	0.841	0.856	0.846
Mean of DV	0.110	0.108	0.110	0.108
$\beta_0 + \beta_1$			0.0155	0.0172
p-value			0.000605	0.0270
Controls	Yes	Yes	Yes	Yes

Notes: Estimates are derived from Equation 9 for columns 3 and 4 using separate regressions for incumbents and challengers. In columns 1 and 2, no interaction with $High - Corruption$ is included. Dependent Variable is a count of all the words in the audit report that are also in the manifesto, divided by the total number of words in the audit report. Only municipalities audited after 2008 are considered (rounds 28-38). Disclosure is a binary variable indicating whether a municipality was audited between 2009 and 2012 and whether the report was made public before the election (rounds 28-35). High-Corruption is a binary variable defined based on the total number of irregularities found during the audit of the municipality, compared to the median number of irregularities identified. Incumbents refers to candidates running for re-election, while challengers are opposition candidates running in a municipality where an incumbent is also contesting. I exclude candidates with manifestos containing fewer than 100 words. All regressions include state fixed-effects. Control variables at the municipality level encompass GDP per capita (in logs), share of illiteracy, share of urban population, Gini index, and indicator variables for populations below 20,000, between 20,000 and 50,000, between 50,000 and 100,000, and above 100,000. There are also binary variables representing whether the candidate belongs to the President's party, the Governor's party, whether the municipality was audited before 2009, and whether there is an AM radio station in the municipality. Clusters are defined at the state level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 5 provides the outcomes derived from Equation 9. The main focus is on the frequency with which words from the municipal corruption report appear in a candidate's campaign promises. In Table 5, Columns 1 and 2 illustrate the impact of releasing a corruption report before an election on both new candidates and incumbents. Columns 3 and 4 focuses on analyzing the effect in municipalities with high and low numbers of reported irregularities. Additionally, considering that the length of the corruption report or the length of the campaign proposals might influence outcomes, the regressions include controls for the number of service orders and the length of the proposals (both in logs). Results from columns 1 and 2 indicate that candidates tend to incorporate more of its content into their campaign promises when a corruption report is made public before an election. This observation holds

true for both new entrants and incumbent politicians. Specifically, there is a 9% increase for new candidates and a 13% increase for incumbents in the use of terms from these reports. Moreover, columns 3 and 4 confirm that this pattern persists irrespective of the number of reported issues in the municipality.

In conclusion, the findings suggest that the timely release of corruption reports influences how candidates craft their campaign messages. If these reports did not offer any fresh insights to the candidates or if the candidates chose to overlook them, there would likely be no notable difference in their usage of report content. However, a discernible difference is observed, implying that candidates view these reports as significant and believe that their constituents do as well.

5.2 Topics in Candidates' Agenda

The evidence thus far suggests candidates integrate corruption-related information into their discourse. However, the exposure of such corruption cases might also lead candidates to adapt their focus on various policy areas where irregularities were observed. If certain irregularities make a topic more pressing, candidates might emphasize it further. However, incumbents might also opt to steer clear of those subjects, especially if the highlighted irregularities portray them in an unfavorable light in terms of policy effectiveness.

Table 6: Do irregularities on a topic influence the extent to which it is discussed?

Outcome: Share of Topic on Manifestos (% of Words on the Manifesto)

	(1)	(2)	(3)	(4)	(5)	(6)
	Social	Health	Economy	Social	Health	Economy
Disclosure	0.007	0.012*	0.011	-0.014	-0.009	-0.002
	(0.016)	(0.006)	(0.008)	(0.025)	(0.011)	(0.015)
High-Corruption x Disclosure	0.012	0.011	0.012	-0.021	-0.012	-0.053***
	(0.019)	(0.014)	(0.013)	(0.029)	(0.014)	(0.013)
High-Corruption	-0.018	-0.001	0.003	0.024	0.014	0.052***
	(0.016)	(0.012)	(0.012)	(0.028)	(0.011)	(0.011)
Candidate	Challengers	Challengers	Challengers	Incumbents	Incumbents	Incumbents
Observations	474	474	474	257	257	257
R-squared	0.14	0.21	0.14	0.16	0.22	0.21
$\beta_0 + \beta_1$	0.02	0.02	0.02	-0.04	-0.02	-0.06
pval	0.18	0.04	0.04	0.04	0.12	0.00
Mean of DV	0.36	0.11	0.15	0.38	0.11	0.15
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Estimates are derived from Equation 9. Dependent Variable is the share of each manifesto dedicated to each topic (measured in number of words). Only municipalities audited after 2008 are considered (rounds 28-38). Disclosure is a binary variable indicating whether a municipality was audited between 2009 and 2012 and whether the report was made public before the election (rounds 28-35). High-Corruption is a binary variable defined based on the total number of irregularities found during the audit of the municipality, compared to the median number of irregularities identified. Incumbents refers to candidates running for re-election, while challengers are opposition candidates running in a municipality where an incumbent is also contesting. I exclude candidates with manifestos containing fewer than 100 words. All regressions include state fixed-effects. Control variables at the municipality level encompass GDP per capita (in logs), share of illiteracy, share of urban population, Gini index, and indicator variables for populations below 20,000, between 20,000 and 50,000, between 50,000 and 100,000, and above 100,000. There are also binary variables representing whether the candidate belongs to the President's party, the Governor's party, whether the municipality was audited before 2009, and whether there is an AM radio station in the municipality. Clusters are defined at the state level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 6 shows the results estimated from Equation 9. It assesses the influence of audit report disclosures on the frequency of manifesto words related to health policies (columns 1 and 4), education policies (columns 2 and 5), and economic policies (columns 3 and 6). Columns 1-3 gauge the effects of these disclosures on challengers, whereas columns 4-6 focus on incumbents. Each evaluation incorporates an interaction with a binary variable, indicating if the proportion of irregularities associated with a given policy topic exceeds or falls short of the median across municipalities. This facilitates understanding the impact of report disclosures on manifesto content related to a specific policy area, particularly in cities with a notable concentration of related irregularities. Columns 1-3 reveal that challengers in cities with pronounced irregularities in a particular policy area tend to allocate more

manifesto space to that topic. The significance of this trend is evident in columns 2 and 3, with a notable increase of 22% in discussions related to health policy and 16% for economic matters.

Conversely, columns 4-6 show that incumbents in cities with substantial irregularities in certain areas are less inclined to address those subjects. This avoidance is especially pronounced for social and economic policies, with reductions of 9% and 37%, respectively, in the mean share of words allocated to these topics.

While all candidates draw upon audit report data, disclosure effects manifest differently in their broader discourse. Opposition figures delve deeper into areas with highlighted irregularities, capitalizing on potential electoral gains, whereas incumbents exhibit reticence, likely to sidestep perceived electoral vulnerabilities. This behavior underscores the significance of reputation in guiding topic selection during campaigns.

This behavioral trend aligns with predictions made by Riker (1996). A significant concentration of irregularities in a specific policy realm can potentially tarnish perceptions of an incumbent's administrative efficacy, prompting them to downplay the subject. Conversely, challengers perceive an opportunity to gain comparative ground and thus intensify their focus on these areas.

5.3 The Average Effects of an Audit

An alternative approach to understanding the implications of disclosed audit reports is by comparing the effects of audit disclosures, which scrutinize spending in a specific policy domain to those municipalities where spending in that domain was not evaluated. The outcome of an audit could reflect unfavorably on an incumbent or not. As such, predicting the precise impact of an audit on the incumbent becomes challenging because the reputational and informational effects could move in opposite directions, potentially canceling each other out. In contrast, the effect on challengers is more discernible. Both informational and reputational factors align for challengers, suggesting that an audit focused on a specific

topic should amplify their discussions on that matter.

This exploration capitalizes on the fact that all small municipalities underwent comprehensive expenditure audits. In contrast, in many of the lotteries, municipalities with populations exceeding 50,000 were seldom audited on health expenditure. Observing changes in health policy discussions in municipalities that didn't undergo health expenditure audits would be concerning. Such a pattern would suggest that other external factors, not the audit report's content, influence the political agenda.

Table 7: Do audits on Health Expenditures influence the extent to which Health Policies are discussed?

Outcome: Share of Health Policies on Manifestos (% of Words on the Manifesto)

	(1)	(2)	(3)	(4)
Audited	0.009***	0.002	0.001	0.011
	(0.003)	(0.007)	(0.004)	(0.017)
Candidate	Challengers	Challengers	Incumbents	Incumbents
Population	Below 50K	Above 100K	Below 50K	Above 100K
Observations	3,283	325	2,053	110
R-squared	0.063	0.140	0.053	0.413
Mean of DV	0.106	0.0875	0.111	0.0900

Notes: Estimates are derived from Equation 9 without interacting with High-Corruption. Dependent Variable is the share of each manifesto dedicated to Health Policies (measured in number of words). Audited is a binary variable indicating whether a municipality was audited between 2009 and 2012 and whether the report was made public before the election (rounds 28-35). All municipalities are considered to estimate these models. Columns 1 and 3 only look at municipalities with a population below 50,000. Columns 2 and 4 show results for municipalities with a population above 100, excluding those audited on rounds 30, 33 and 35. Incumbents refers to candidates running for re-election, while challengers are opposition candidates running in a municipality where an incumbent is also contesting. I exclude candidates with manifestos containing fewer than 100 words. All regressions include state fixed-effects. Control variables at the municipality level encompass GDP per capita (in logs), share of illiteracy, share of urban population, Gini index, and indicator variables for populations below 20,000, between 20,000 and 50,000, between 50,000 and 100,000, and above 100,000. There are also binary variables representing whether the candidate belongs to the President's party, the Governor's party, whether the municipality was audited before 2009, and whether there is an AM radio station in the municipality. Clusters are defined at the state level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 7 presents the audit effects for both challengers and incumbents on the depth of their health discourse. The control group comprises municipalities that escaped any audits. Columns 1 and 2 display the effects on challengers, while columns 3 and 4 focus on incumbents. Columns 1 and 3 consider the sample of municipalities audited on health expenditures, whereas columns 2 and 4 consider those not audited on this front.

Column 1 elucidates that health expenditure audits significantly amplify health policy discussions for challengers. Conversely, Column 2 reveals an absence of this effect in municipalities spared from health expenditure audits – an expected outcome given the lack of new information or perceived electoral advantages on this subject. For incumbents, neither set of municipalities yields significant outcomes (columns 3 and 4).

The observed divergence from earlier findings - where challengers intensify discussions on irregularity-heavy topics while incumbents retreat - may stem from the dual nature of an audit as both a reputational and informational signal. On average, the opposition leverages audit reports to bolster discussions on audited topics. Given insights from subsection 5.1, we know the opposition harnesses audit details, accentuating discussions when irregularities surface. Thus, it's logical for a topic-centric audit to spur opposition discussions on that topic. For incumbents, however, the informational and reputational dimensions of the audit may pull in different directions. This might elucidate why a mere audit, devoid of a high irregularity count (as documented in subsection 5.2), fails to shift the discourse.

In summary, these findings resonate with prior evidence, underscoring that the specific content and focus of audit reports influence the weight politicians assign to various topics in their discourse.

5.4 Discussion

The results presented offer compelling evidence that government audits notably influence the content of political proposals during electoral campaigns. It becomes evident from the findings that the disclosure of audit reports informs and guides the crafting of candidates' proposals. Both incumbents and challengers appear to incorporate language from these reports into their campaign narratives, largely setting aside the reputational implications these audits might have for the sitting mayor. This attests to the presence and influence of the informational channel.

Nevertheless, reputation plays a pivotal role when candidates decide which broader topics

to delve into. Audit disclosures, particularly when released pre-election, deter incumbents from extensively discussing areas riddled with irregularities. Conversely, Challengers seize upon these issues, amplifying discussions around them. This behavior resonates with the theoretical perspective suggesting candidates prioritize issues where they perceive a reputational edge (Riker, 1996; Seeberg, 2022). Interestingly, the data suggests that while specific topic audits reshape the discourse of challengers, incumbents' discussions remain relatively unaffected.

In closing, these insights show how corruption-related information steers political communication and discourse surrounding distinct issues. The findings underscore the informational utility of audit revelations for proposal formulation and the strategic considerations candidates weigh, pivoting their narratives based on perceived strengths and vulnerabilities in different policy spheres.

6 Impact of audits on the use of ideological rhetoric

Unveiling corruption may induce significant shifts in candidates' ideological stances. Specifically, such disclosures might prompt candidates to adopt more extreme, partisan, or populist positions in their proposals. Scores are assigned to each manifesto to measure extremism and partisanship, as detailed in subsubsection 2.3.3. Populism is measured using a methodology outlined in subsubsection 2.3.4.

Table 8: Do irregularities on a topic influence the ideological framework?

Outcome: Score for each dimension on Manifestos

Panel A: Challengers						
	(1)	(2)	(3)	(4)	(5)	(6)
	Populism	Extremeness	Partisanship	Populism	Extremeness	Partisanship
Disclosure	0.002	-0.183*	-0.055	-0.301**	-0.661**	-0.254
	(0.046)	(0.096)	(0.087)	(0.138)	(0.244)	(0.266)
High-Corruption x Disclosure				0.387**	0.348	0.194
				(0.154)	(0.465)	(0.611)
High-Corruption				-0.279**	-0.026	0.105
				(0.110)	(0.368)	(0.630)
Observations	3,948	3,946	3,489	474	474	422
R-squared	0.111	0.058	0.025	0.134	0.154	0.126
Mean of DV	0.0800	2.248	-0.0933	0.130	2.185	-0.191
$\beta_0 + \beta_1$				0.0861	-0.313	-0.0599
p-value				0.569	0.546	0.894
Panel B: Incumbents						
	(1)	(2)	(3)	(4)	(5)	(6)
	Populism	Extremeness	Partisanship	Populism	Extremeness	Partisanship
Disclosure	-0.023	0.212	0.073	0.063	0.914*	0.781
	(0.058)	(0.250)	(0.214)	(0.096)	(0.520)	(0.657)
High-Corruption x Disclosure				0.273	-1.466*	-1.881*
				(0.170)	(0.721)	(1.048)
High-Corruption				-0.322	1.233*	1.067
				(0.206)	(0.628)	(1.092)
Observations	2,298	2,261	2,181	257	254	238
R-squared	0.15	0.05	0.03	0.28	0.16	0.15
Mean of DV	-0.18	2.13	-0.04	-0.18	2.24	-0.06
$\beta_0 + \beta_1$				0.34	-0.55	-1.10
p-value				0.03	0.43	0.30

Notes: Estimates are derived from Equation 9. Dependent Variable in columns 1 and 4 is the share of each manifesto dedicated to words associated with populism (weighted by tf-idf). Dependent Variable in columns 2 and 5 are Ideological Extremeness Score and in 3 and 6 are Partisanship scores. For columns 1-3 all municipalities are considered. For columns 3-6 only municipalities audited after 2008 are considered (rounds 28-38). Disclosure is a binary variable indicating whether a municipality was audited between 2009 and 2012 and whether the report was made public before the election (rounds 28-35). High-Corruption is a binary variable defined based on the total number of irregularities found during the audit of the municipality, compared to the median number of irregularities identified. Incumbents refers to candidates running for re-election, while challengers are opposition candidates running in a municipality where an incumbent is also contesting. I exclude candidates with manifestos containing fewer than 100 words. All regressions include state fixed-effects. Control variables at the municipality level encompass GDP per capita (in logs), share of illiteracy, share of urban population, Gini index, and indicator variables for populations below 20,000, between 20,000 and 50,000, between 50,000 and 100,000, and above 100,000. There are also binary variables representing whether the candidate belongs to the President's party, the Governor's party, whether the municipality was audited before 2009, and whether there is an AM radio station in the municipality. Clusters are defined at the state level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

Table 8 shows the relationship between corruption disclosure and these ideological indi-

cators. Columns 1-3 of Table 8 present the average effects of audit disclosure on challengers and incumbents, comparing them to non-audited municipalities. Columns 4-6 introduce an interaction with a binary variable, indicating whether the number of irregularities is above or below the median. This distinction helps in discerning potential divergent effects in municipalities characterized by significant public sector irregularities.

Upon examining columns 1-3 (Panels A and B), we find a modest influence of the audit on ideological orientations. Notably, there's an observable moderation in the rhetoric of opposition candidates following the audit. However, a clearer picture emerges when considering the effects of audit reports conditional upon corruption levels.

Firstly, in Panel A (columns 4 and 5), opposition candidates in municipalities with fewer instances of corruption seem to moderate their positions, especially distancing from populist rhetoric. In terms of populism, this change represents a decrease of 0.3 standard deviations relative to all candidates. Moreover, there's a decline in ideological extremeness by 0.67 points, representing a 30% drop from the average. This suggests that the subtle effect spotted in column 2 primarily stems from municipalities with a lower count of irregularities.

Secondly, as shown in Panel B (column 4), incumbents in areas with above-median corruption counts tend to adopt a more populist discourse post-audit. This trend suggests that incumbents might aim to deflect attention from salient irregularities. In contrast, no significant changes appear in the partisanship dimension (columns 3 and 6).

The presented analysis underscores the varying impacts of audits on the ideological content of campaigns for incumbents and challengers. Notably, the effects are nuanced, contingent on whether a candidate is contesting in a municipality characterized by low or high corruption levels. These observed effects persist even when employing alternate thresholds to categorize municipalities by their irregularity count (e.g., the 75th percentile instead of the median).

6.1 Discussion of the Results

This section shows that incumbents and challengers react differently in their campaigns in terms of their ideological framework to the disclosure of corruption cases. For incumbents in municipalities marked by substantial irregularities, there appears to be a discernible shift toward populist rhetoric. This might serve as a strategic effort to galvanize their base of support. An exploratory analysis in the appendix (Table A.8 examines if this shift can be attributed to increased usage of terms like 'corruption,' 'corrupt,' or 'transparency.' However, the findings indicate that the shift does not hinge on these specific terms but likely emerges from a broader populist narrative.

Interestingly, challengers do not respond in high-corruption municipalities to the disclosure of corruption. However, when information does not directly harm the incumbent (low-corruption municipalities), challengers often moderate their stance and avoid strong populist rhetoric. This suggests that appealing to the median voter is crucial in less competitive elections.

A potential interpretation for this moderation in rhetoric, especially among challengers, is a strategic shift to emphasize local concerns over broader ideological or populist themes. The differentiation between localized promises (e.g., "build a school") versus national or aspirational goals (e.g., "enhance education and combat inequality") is challenging without specialized algorithmic analysis. If reference to corruption cases were solely driving this shift, one would expect incumbents to exhibit a similar pattern. As such, it's vital to consider this potential localized focus when interpreting the observed moderation in challengers' rhetoric. Ultimately, this trend, signaling increased civility among opposition candidates post-audit, may well augur a more constructive political atmosphere for voters.

In sum, these findings emphasize that revelations of corruption shape the political narrative and influence candidates' strategies in response to reputational shocks during campaigns.

7 Conclusions

In this comprehensive analysis of 13,344 manifestos from Brazil's 2012 municipal election, I sought to elucidate the consequences of an auditing program on the topics and ideological framework used within candidates' manifestos. The findings show that audits affect political communication and electoral narratives. Yet, these effects are not uniform — they are heterogeneous by the level of corruption and whether a candidate is an incumbent or a challenger.

Candidates, without a doubt, assimilate and respond to audit findings. However, their emphasis on specific policy areas depends on the prevalence of irregularities in that area. Challengers focus on areas with irregularities, while incumbents strategically avoid them. This behavior aligns with their strategic positioning based on perceived strengths and weaknesses (Riker, 1996).

Audit disclosures often lead incumbent candidates in corruption-heavy municipalities to adopt a distinct populist tone. This contrasts sharply with the moderate rhetoric of challengers in low-corruption areas. Surprisingly, this study suggests populism is not limited to newcomers, as the literature suggests (Berman, 2021; Guriev & Papaioannou, 2022), but is also embraced by incumbents, especially those seeking re-election. In the context of Latin America's political landscape, this warrants further detailed investigation.

Previous studies have explored electoral outcomes and corruption trends following corruption revelations (Avis et al., 2018; Ferraz & Finan, 2008). This research extends that discussion, presenting audits as instruments that reshape political responsiveness and drive nuanced ideological shifts, sometimes steering narratives toward populism.

In conclusion, future research should explore how the quality of information about corruption scandals affects candidate responsiveness and political ideologies. This paper examined the impact of exposing corruption using federal government data. It raises the question: Would there be different effects if allegations came from the media? Moreover, Ferraz and Finan (2008) observed effects on reelection outcomes. It prompts further inquiry: Are these

changes in election results due to electoral incentives influencing ideology shifts or topic focus? Essentially, can changes in ideologies counter potential reputational damages?

References

Abou-Chadi, T., Green-Pedersen, C., & Mortensen, P. B. (2020). Parties' policy adjustments in response to changes in issue saliency. *West European Politics*, *43*(4), 749–771.

Amorim, G. (2022). *Monitoring transfers to public health: Evidence from randomized audits in brazil* (Tech. Rep.).

Aragonès, E., Castanheira, M., & Giani, M. (2015). Electoral competition through issue selection. *American journal of political science*, *59*(1), 71–90.

Avis, E., Ferraz, C., & Finan, F. (2018). Do government audits reduce corruption? estimating the impacts of exposing corrupt politicians. *Journal of Political Economy*, *126*(5), 1912–1964.

Berman, S. (2021). The causes of populism in the west. *Annual Review of Political Science*, *24*, 71–88.

Bernhardt, D., Buisseret, P., & Hidir, S. (2020). The race to the base. *American Economic Review*, *110*(3), 922–42.

Bolognesi, B., Ribeiro, E., & Codato, A. (2022). A new ideological classification of brazilian political parties. *Dados*, *66*.

Brollo, F., & Troiano, U. (2016). What happens when a woman wins an election? evidence from close races in brazil. *Journal of Development Economics*, *122*, 28–45.

Buisseret, P., & Van Weelden, R. (2022). Polarization, valence, and policy competition. *American Economic Review: Insights*, *4*(3), 341–352.

Cagé, J., Le Pennec, C., & Mougin, E. (2022). Money and ideology: Evidence from candidate manifestos.

Catalinac, A. (2018). Positioning under alternative electoral systems: evidence from japanese candidate election manifestos. *American Political Science Review*, *112*(1), 31–48.

Cavalcanti, F., Daniele, G., & Galletta, S. (2018). Popularity shocks and political selection. *Journal of Public Economics*, *165*, 201–216.

Colonnelli, E., & Prem, M. (2020). Corruption and firms. *Available at SSRN 2931602*.

Crabtree, C., Golder, M., Gschwend, T., & Indriason, I. H. (2020). It is not only what you say, it is also how you say it: The strategic use of campaign sentiment. *The Journal of Politics*, *82*(3), 1044–1060.

Dragu, T., & Fan, X. (2016). An agenda-setting theory of electoral competition. *The Journal of Politics*, *78*(4), 1170–1183.

Ferraz, C., & Finan, F. (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, C., & Finan, F. (2011). Electoral accountability and corruption: Evidence from the audits of local governments. *American Economic Review*, *101*(4), 1274–1311.

Gagliarducci, S., Paserman, M. D., & Patacchini, E. (2019). *Hurricanes, climate change policies and electoral accountability* (Tech. Rep.). National Bureau of Economic Research.

Gennaro, G., Lecce, G., & Morelli, M. (2021). Mobilization and the strategy of populism theory and evidence from the united states.

Gonzales, M. (2021). Politics never end: Public employment e ects of increased transparency.

Groseclose, T., & Milyo, J. (2005). A measure of media bias. *The Quarterly Journal of Economics*, *120*(4), 1191–1237.

Guriev, S., & Papaioannou, E. (2022). The political economy of populism. *Journal of Economic Literature*, *60*(3), 753–832.

Lauletta, M., Rossi, M., & Ruzzier, C. (2020). *Audits and the quality of government* (Tech. Rep.). (Working Paper N 141, Universidad de San Andrés. Available at `https://webacademicos.udesa.edu.ar/pub/econ/doc141.pdf`)

Laver, M., Benoit, K., & Garry, J. (2003). Extracting policy positions from political texts using words as data. *American political science review*, *97*(2), 311–331.

Le Pennec, C. (2022). *Strategic campaign communication: Evidence from 30,000 candidate manifestos* (Tech. Rep.).

Martin, L. W., & Vanberg, G. (2008). A robust transformation procedure for interpreting political text. *Political Analysis*, *16*(1), 93–100.

Mendes, M. S. (2021). 'enough'of what? an analysis of chega's populist radical right agenda. *South European Society and Politics*, *26*(3), 329–353.

Mudde, C., & Rovira Kaltwasser, C. (2018). Studying populism in comparative perspective: Reflections on the contemporary and future research agenda. *Comparative political studies*, *51*(13), 1667–1693.

Poblete-Cazenave, R. (2021). Reputation shocks and strategic responses in electoral campaigns. *Available at SSRN 3786253*.

Riker, W. H. (1996). *The strategy of rhetoric: Campaigning for the american constitution*. Yale University Press.

Seeberg, H. B. (2022). First avoidance, then engagement: Political parties' issue competition in the electoral cycle. *Party Politics*, *28*(2), 284–293.

Serra, G. (2010). Polarization of what? a model of elections with endogenous valence. *The Journal of Politics*, *72*(2), 426–437.

Snyder Jr, J. M., & Strömberg, D. (2010). Press coverage and political accountability. *Journal of political Economy*, *118*(2), 355–408.

Tarouco, G. d. S., & Madeira, R. M. (2015). Os partidos brasileiros segundo seus estudiosos: análise de um expert survey. *Civitas-Revista de Ciências Sociais*, *15*, e24–e39.

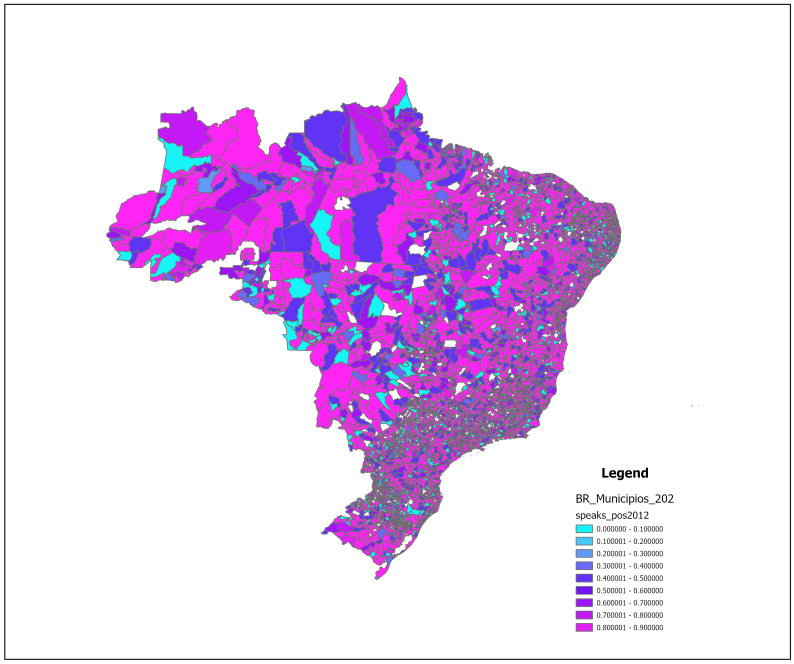
Williams, L. K., Seki, K., & Whitten, G. D. (2016). You've got some explaining to do the influence of economic conditions and spatial competition on party strategy. *Political Science Research and Methods*, *4*(1), 47–63.

Contents

A Figures and Tables

A.1 Figures

2012



2020

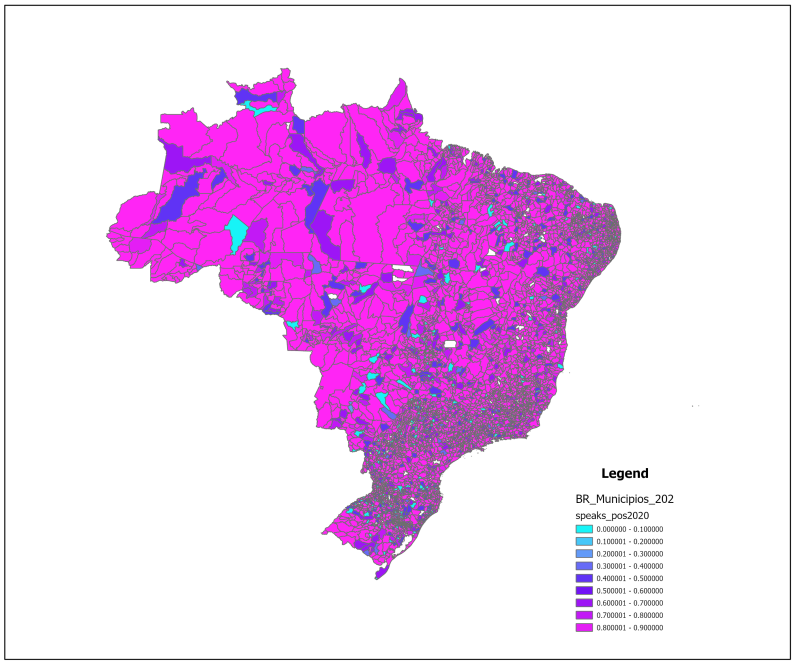


Figure A.1: Percentage of local party manifestos in each municipality that include a populist word

A.2 Tables

Table A.1: Topics covered by the audits in each lottery

Topic by CGU	**Topic**	**Population Ranges**(thousands)	**Lotteries**								
			28	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
Agriculture	Econ	20<x<50		X			X				
		50<x<100		X			X				
		x>100		X			X				
Commerce	Econ	20<x<50		X			X				
		50<x<100		X			X				
		x>100		X			X				
Crime	Crime	20<x<50			X	X		X			
		50<x<100			X	X		X			
		x>100			X	X		X			
Culture	Social	20<x<50		X			X				
		50<x<100		X			X				
		x>100		X			X				
Education	Social	20<x<50	X	X	X	X	X	X	X	X	X
		50<x<100	X	X	X	X	X	X	X		X
		x>100	X			X			X		X
Health	Health	20<x<50	X	X	X	X	X	X	X	X	X
		50<x<100	X	X	X	X	X	X		X	
		x>100			X			X		X	
Housing	Urban	20<x<50	X								
		50<x<100	X								
		x>100	X								
Industry	Econ	20<x<50			X	X		X			
		50<x<100			X	X		X			
		x>100			X	X		X			
Sanitation	Urban	20<x<50	X								
		50<x<100	X								
		x>100	X								
Science and Technology	Social	20<x<50			X	X		X			
		50<x<100			X	X		X			
		x>100			X	X		X			
Services	Bureau/Urban	20<x<50		X			X				
		50<x<100		X			X				
		x>100		X			X				
Social Assistance	Social	20<x<50	X	X	X	X	X	X			
		50<x<100	X	X	X	X	X	X			
		x>100		X			X				
Social Development	Social	20<x<50							X	X	X
		50<x<100							X	X	X
		x>100							X	X	X
Urban Planning	Urban	20<x<50	X								
		50<x<100	X								
		x>100	X								

Table A.2: Number of Irregularities per Policy Area

	N	Percentage
Bureaucracy	58	0.2%
Crime	138	0.4%
Economics	1,494	4.1%
Health	10,207	28.2%
Social	21,644	59.8%
Urban	1,818	5.0%
Other	811	2.2%

Percentages refer to the percentage of total cases in lotteries between the 28th and the 38th
Source: TSE

Table A.3: Examples of Headings for Each Topic

Introduction	**Administrative / Bureaucracy**	**Social**	**Health**	**Urban**	**Economics**	**Crime**	**Other**
Concluding Remarks	Administration	Childhood and Elders	Health	Cleaning	Agriculture	Crime	Communication
General Comments	Administration and Participation	Culture		Cleaning and Environment	Agriculture and Environment		Events Center
Introduction	Government	Culture and Sport		Housing	Development		Funerals
	Management	Culture and Sports		Infrastructure	Economic Development		Religion
	Participation	Culture and Tourism		Infrastructure and Services	Economic Development and Sustainable Development		
	Public Finance	Culture, sport and tourism		Infrastructure and transport	Employment		
	Public Servers	Disability		Street lights	Employment and Income		
		Education		Natural Disasters	Environment		
		Education and Culture		Sanitation	Environment and Agriculture		
		Education and Sports		Sanitation and Environment	Industry and Commerce		
		Education, Culture and Sports		Streets	Production		
		Elders		Transit	Rural		
		Gender		Transport	Solidarity Economy		
		Social		Urban	Sustainable Development		
		Social Assistance		Urban and housing			
		Social Development		Urban Development			
		Social Policies		Water			
		Social Policy					
		Sports					
		Sports and Culture					
		Sports and Tourism					
		Sports and Youth					
		Sports, culture, youth and tourism					
		Tourism					
		Tourism, Culture and Sports					
		Youth					
		Youth, Gender and Elders					

Table A.4: Distribution of Topics in the Corpus

Outcome: Share of Topic on Manifestos (% of Words on the Manifesto)

	Sample	Soft Prediction
Titles	2.9	2.1
Introduction & Other topics	14.9	15.8
Administrative / Bureaucracy	8.3	8.7
Social	31.3	35.4
Health	9.5	8.7
Urban	10.6	9.6
Economics	13.9	10.7
Crime	2.5	2.2
Unrecognizable characters	6.2	1.0

Notes: Column 1 represents the topic distribution in the sample of 100 manifestos that was used to tune the model. Column 2 represents the topic distribution using the soft measure described in section 2.

Table A.5: Words with the lowest (left-wing) and highest (right-wing) scores

Leftist Words		Right-wing Words	
cidasc	[a]	democratas	democrats
petista	[b]	cristao	Christian (masc)
capitalista	capitalist	crista	Christian (fem)
petistas	[b]	democrata	democrat
socialismo	socialism	republicano	republican
inverter	reverse	farei	will do
deliberativos	deliberative	indeb	[c]
desiguais	unequal	renova	renew
socialistas	socialists	equoterapia	equine therapy
dominante	dominant	grafias	spellings

Notes:

[a] *cidasc:* Companhia Integrada de Desenvolvimento Agrícola do Estado de Santa Catarina (Cidasc). This is an agropecuarian policy in Santa Catarina.

[b] *petista:* Member of the Workers Party (PT)

[c] *indeb:* Basic Education Development Index (Indeb)

Table A.6: Mean Comparisons between Audited and Nonaudited Municipalities (High Corruption)

	Control	Treatment	Difference
GDP pc	12886.52	10805.9	-819.3
	[14487.15]	[9571.88]	[569.971]
Share Illiterate (%)	85.3472	83.51522	-0.0758
	[8.86]	[9.33]	[0.258]
Share Urban	0.6374198	0.6262824	0.00446
	[0.22]	[0.21]	[0.007]
Share Secondary Education and above	0.2156972	0.2083799	0.000671
	[0.08]	[0.08]	[0.003]
Share of Bureaucrats with Superior Education	0.3069009	0.2967605	-0.00219
	[0.11]	[0.11]	[0.004]
HDI	0.6598012	0.6443488	-0.00108
	[0.07]	[0.07]	[0.002]
AM radio	0.2092931	0.1987315	-0.00137
	[0.41]	[0.4]	[0.025]
Gini	0.5013802	0.5095829	-0.0014
	[0.07]	[0.06]	[0.002]
Population (logs)	9.377024	9.470213	0.00526
	[1.09]	[1.1]	[0.032]
Audited Previously	0.2499018	0.2635983	-0.0113
	[0.43]	[0.44]	[0.021]
Observations	5090	478	

Notes: Estimates are means and standard deviations (in brackets) of various municipal characteristics by places that have been audited in round 28 to 35 (2009-2012) and places that have been audited in rounds 36 to 38 with a high level of corruption (defined by the median of irregularities found). The difference and corresponding standard error (in brackets) are computed on the basis of a regression that controls for state.

Table A.7: Mean Comparisons between Audited and Nonaudited Municipalities (Low Corruption)

	Control	Treatment	Difference
GDP pc	17136.57	13572.12	-1596.1
	[24649.26]	[12166.64]	[3520.809]
Share Illiterate (%)	87	87.39	1.371*
	[8.65]	[8.02]	[0.731]
Share Urban	0.64	0.66	0.0245
	[0.25]	[0.21]	[0.034]
Share Secondary Education and above	0.22	0.22	0.0141
	[0.08]	[0.09]	[0.012]
Share of Bureaucrats with Superior Education	0.32	0.3	-0.0118
	[0.14]	[0.1]	[0.026]
HDI	0.68	0.67	0.00408
	[0.06]	[0.07]	[0.007]
AM radio	0.25	0.19	-0.00734
	[0.44]	[0.4]	[0.092]
Gini	0.49	0.5	0.00901
	[0.07]	[0.06]	[0.011]
Population (logs)	9.4	9.33	0.101
	[1.28]	[1.17]	[0.197]
Audited Previously	0.2	0.27	0.131*
	[0.41]	[0.45]	[0.074]
Observations	44	108	

Notes: Estimates are means and standard deviations (in brackets) of various municipal characteristics by places that have been audited in round 28 to 35 (2009-2012) and places that have been audited in rounds 36 to 38 with a low level of corruption (defined by the median of irregularities found). The difference and corresponding standard error (in brackets) are computed on the basis of a regression that controls for state.

Table A.8: Do irregularities on a topic influence the frequency of specific words?

Outcome: Frequency of words on Manifestos (% of Words on the Manifesto)

	Panel A: Challengers			
	(1)	(2)	(3)	(4)
	Corruption	Transparency	Corruption	Transparency
Disclosure	-0.000	-0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High-Corruption x Disclosure			0.000	0.000
			(0.000)	(0.000)
High-Corruption			0.000	-0.001*
			(0.000)	(0.000)
Observations	3,948	3,948	474	474
R-squared	0.027	0.069	0.130	0.107
Mean of DV	5.46e-05	0.00200	6.14e-05	0.00181
$\beta_0 + \beta_1$			-4.81e-05	0.000499
p-value			0.605	0.261
	Panel B: Incumbents			
	(1)	(2)	(3)	(4)
	Corruption	Transparency	Corruption	Transparency
Disclosure	-0.000*	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
High-Corruption x Disclosure			-0.000	0.000
			(0.000)	(0.000)
High-Corruption			0.000	-0.000
			(0.000)	(0.000)
Observations	2,298	2,298	257	257
R-squared	0.010	0.056	0.093	0.254
Mean of DV	2.03e-05	0.00185	1.01e-05	0.00173
$\beta_0 + \beta_1$			-3.13e-05	0.000276
p-value			0.556	0.419

Notes: Estimates are derived from Equation 9. Dependent Variable in columns 1 and 3 is the share of each manifesto dedicated that start with 'corrup'. Dependent Variable in columns 2 and 4 is the share of each manifesto dedicated that start with 'transp'. For columns 1-2 all municipalities are considered. For columns 3-4 only municipalities audited after 2008 are considered (rounds 28-38). Disclosure is a binary variable indicating whether a municipality was audited between 2009 and 2012 and whether the report was made public before the election (rounds 28-35). High-Corruption is a binary variable defined based on the total number of irregularities found during the audit of the municipality, compared to the median number of irregularities identified. Incumbents refers to candidates running for re-election, while challengers are opposition candidates running in a municipality where an incumbent is also contesting. I exclude candidates with manifestos containing fewer than 100 words. All regressions include state fixed-effects. Control variables at the municipality level encompass GDP per capita (in logs), share of illiteracy, share of urban population, Gini index, and indicator variables for populations below 20,000, between 20,000 and 50,000, between 50,000 and 100,000, and above 100,000. There are also binary variables representing whether the candidate belongs to the President's party, the Governor's party, whether the municipality was audited before 2009, and whether there is an AM radio station in the municipality. Clusters are defined at the state level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

B Preprocessing Manifestos

The preprocessing steps were:

- Converted each pdf files into a json format.
- Any empty json files was removed.
- For the topic analysis, the next step was transforming each entry in the json file into a line in the pdf.
- Stop words were removed from each line. The stop words used for the topic analysis were those included in nltk for the Portuguese language.
  For the analysis where I looked at the frequency of the populist words, I also excluded the parties names, the state names, and the names of each candidates. This was done to reduce the total count of words and imbalances that could be generated by the use of these removed terms.
- The tokenization process was then performed and punctuation signs were removed. In all cases 1-word n-grams where used.

C Populist dictionary

The dictionary that was used is a translation of the one described in the appendix of (Gennaro et al., 2021). I translated that dictionary into Portuguese. The dictionary they used is one of stemmed words. Thus, I had to look at all the possible words that have a similar stem.

- 'casta', 'classe', 'classes', 'elite', 'elites', 'elitista', 'elitistas', 'elitismo', 'elitização', 'elitizado', 'establishment','estabelecimento', 'estabelecimentos', 'corrup', 'corrupta', 'corrupto', 'corruptor', 'corruptos', 'corrupção', 'corrupções', 'corruptas', 'corruptores', 'corrupça' ,'corrompe', 'regime', 'regimentais', 'regimento', 'regimentos', 'regimes', 'propaganda', 'propagandas', 'proeminente', 'proeminentes', 'arrogância','arrogante', 'trair', 'traição', 'trais', 'promessa', 'promessas', 'promessasmas', 'vergonha', 'vergonhosa', 'vergonhoso', 'vergonhosos', 'vergonhosamente', 'vergonhosas', 'desavergonhado', 'descarado', 'descarada', 'absurdamente', 'absurdas', 'absurdo', 'absurdos', 'absurda', 'disparatado', 'referendo', 'referendum', 'referenda', 'referendada', 'referendado', 'referendadas', 'referendados', 'referendos' 'povo', 'gente', 'povos', 'tradição', 'tradicioalista', 'tradicionais', 'tradicional', 'tradicionalismo', 'tradicionalista', 'tradicionalistas', 'tradicionalmente', 'tradições', 'tradiçoes', 'direta', 'politicos', 'estadista', 'governar', 'antidemocrata', 'engano', 'fraude', 'dolo'.

I also incorporated some words that were available in the dictionary described by Mendes (2021). This is a dictionary in portuguese to identify populism speeches:

- 'voz', 'verdade', 'verdadeira', 'verdadeiramente','verdadeiras', 'verdadeiro', 'verdadeiros', 'verdades', 'verdadeiraos','mentira', 'mentiras', 'oligarquia', 'oligarquias', 'clientelismo'.

D Party Classification

D.1 Parties

I follow the classification made by Tarouco and Madeira (2015). They surveyed Brazilian experts to get the ideological positioning of the parties. They do not classify parties as left, right or center. I decided to group all parties between 1 and 4 as left-wing, and all parties between 5 and 7 as right. The remaning are classified as center parties.

Left	Center	Right
PCO	MDB	PTB
PSTU	Avante	SDD
PSOL	PMN	Podemos
PCB	PHS	PRTB
PCdoB	PSDB	PRB
PT	PSD*	PTC
PSB		PRP
PDT		PR
PV		PSL
PPS		PSC
UP*		DC
PPL*		Progre
		DEM

Table D.9: Parties in the 2012 Election and how they are labeled to measure extremeness

[a] Notes: This table shows distribution of parties in the left, center and right-wing groups following Tarouco and Madeira (2015). They are ordered from left (above) to right (below). * Parties classified using Bolognesi, Ribeiro, and Codato (2022).

D.2 Ideological Scores' Distributions

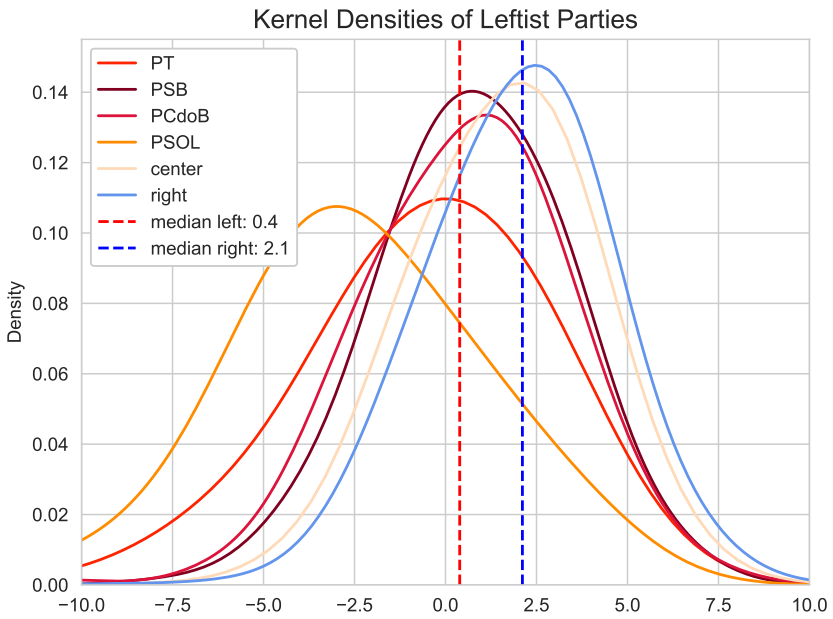


Figure D.2: Ideological Scores' Densities for selected Left-wing parties

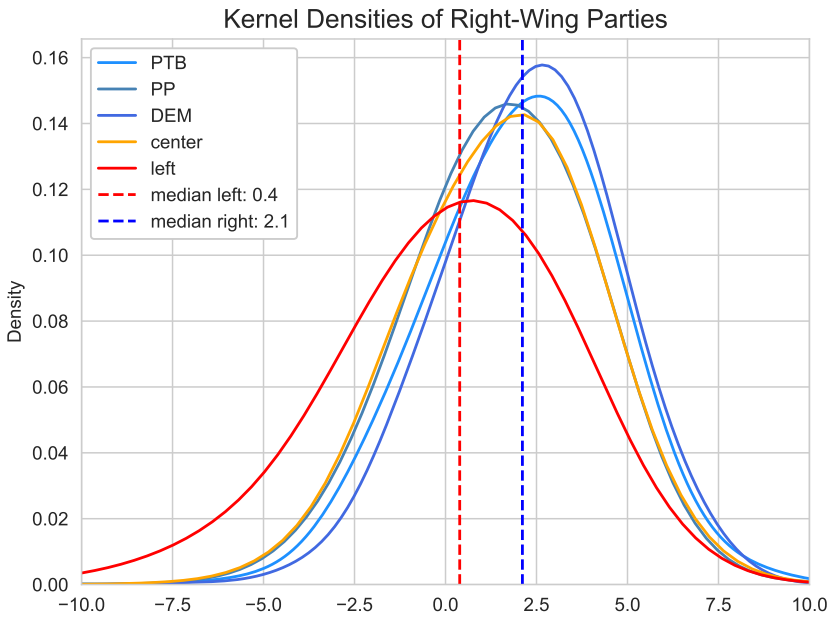


Figure D.3: Ideological Scores' Densities for selected Right-wing parties