Pontifícia Universidade Católica
do Rio de Janeiro

Nathalia Machado Sales

Essays on Public Procurement and Political Economy

Tese de Doutorado

Thesis presented to the Programa de Pós–graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Economia.

Advisor : Prof. Juliano Assunção
Co-advisor: Prof. Ricardo Dahis

Rio de Janeiro
September 2024

Pontifícia Universidade Católica
do Rio de Janeiro

Nathalia Machado Sales

Essays on Public Procurement and Political Economy

Thesis presented to the Programa de Pós–graduação em Economia of PUC-Rio in partial fulfillment of the requirements for the degree of Doutor em Economia. Approved by the Examination Committee.

Prof. Juliano Assunção
Advisor
Department of Economics – PUC-Rio

Prof. Ricardo Dahis
Co-advisor
Department of Economics – Monash University

Prof. Claudio Ferraz
Department of Economics – PUC-Rio
Department of Economics – University of British Columbia

Prof. Juan Rios
Department of Economics – PUC-Rio

Prof. Joana Naritomi
Department of Economics – London School of Economics and Political Science

Prof. Bernardo Ricca
Department of Economics – Insper

Rio de Janeiro, September 26th, 2024

Nathalia Machado Sales

MSc. in Economics (UFRJ, 2020), BA. in Economics (UFRJ, 2018). During the doctorate programme, was visiting PhD student at the Columbia University Graduate School of Economics (2023/2024).

Bibliographic data

Sales, Nathalia Machado

Essays on Public Procurement and Political Economy / Nathalia Machado Sales; advisor: Juliano Assunção; co-advisor: Ricardo Dahis. – Rio de janeiro: PUC-Rio, Departamento de Economia, 2024.

v., 136 f: il. color. ; 30 cm

Tese (doutorado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia.

Inclui bibliografia

1. Economia – Teses. 2. Economia do Desenvolvimento – Teses. 3. Economia Política – Teses. 4. Economia do Setor Público – Teses. 5. Compras Públicas;. 6. Gestão Fiscal;. 7. Auditorias de Corrupção;. 8. Incentivos à Reeleição;. 9. Tratamento Preferencial. I. Assunção, Juliano. II. Dahis, Ricardo. III. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. IV. Título.

CDD: 620.11

To my parents, Andrea and Jorge.

Acknowledgments

Pursuing a PhD requires a long period of dedication. The years go by quickly and at the same time slowly. First come the courses, exams, and assignments. Then, the research. Research that is not linear and, perhaps because of that, is such a big challenge. Resilience, that is the word. Not knowing exactly where you will end up and still continuing, little by little, every day.

To reach the end of this process, many things are necessary. The most fundamental, without a doubt: the support of people. First and foremost, I want to express my immense gratitude to my parents, Andrea and Jorge, who have always provided me with the support and encouragement necessary to keep moving forward. I would also like to thank Hebert, my life partner, who, with all his patience, has listened to me in difficult times, understood my absences when necessary, and always encourages me to be better.

I also thank my advisor, Juliano, who agreed to join this process midway, always providing excellent feedback. Likewise, my co-advisor, Ricardo, who began this process as an advisor, became a co-author, and who has always been present and available even after moving to the other side of the world.

Additionally, I want to thank Ricardo, Thiago, Bernardo, and Lucas for the exceptional team we formed on the MiDES project. Over nearly two years of dedicated effort, we worked together to develop this dataset, a portion of which is explored in the first chapter.

I am also thankful to Thiago for his partnership on the second chapter of this thesis and for introducing me to this discussion during my time as a research assistant at the World Bank. Additionally, I would like to thank Martin, my co-author on the third chapter, for his effort to make it work through regular meetings and great discussions.

I also extend my thanks to the colleagues and professors from PUC for their valuable feedback during workshop presentations. Moreover, I want to thank Bianca for her assistance with administrative issues throughout this process.

Finally, I thank professors Michael Best and Bernard Salanié for their reception during my visiting period at Columbia University and for their valuable comments that enriched this project.

* This research was financed by: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

Abstract

Sales, Nathalia Machado; Assunção, Juliano (Advisor); Dahis, Ricardo (Co-Advisor). **Essays on Public Procurement and Political Economy**. Rio de Janeiro, 2024. 136p. Tese de doutorado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This dissertation comprises three chapters, with two dedicated to public procurement and one to political economy. The first chapter examines Brazilian municipal procurement during the COVID-19 pandemic, where legislation allowed for greater discretion exceeding usual thresholds. We document a rapid increase in discretion, particularly for emergency-related goods, and in municipalities with worse fiscal management. Interestingly, no evidence suggests that this led to the selection of suppliers associated with favoritism or corruption, nor to higher prices for a set of crisis-related products. The increased flexibility enabled public buyers to access more suppliers outside municipal boundaries. We discuss whether this mitigated mortality in financially constrained municipalities. The second chapter analyzes a policy that requires public procurement entities in Brazil to set-aside part of their purchases exclusively for SMEs. Our key finding is that the use of set-asides for SMEs reduces competition in auctions. We further investigate the impact of reduced competition on prices and discuss compliance with the policy. Finally, in the third chapter, we extend previous manual attempts at classifying corruption audit reports with a Large Language Model (LLM) to encode reports from Brazilian municipalities. We then apply our extended corruption data to reassess the impact of reelection incentives on corruption. We find some evidence that reelection incentives reduce corruption, corroborating existing findings in the literature. However, the effect sizes are smaller and the effects are only statistically significant for one of the three outcome variables. We introduce alternative explanations to the empirical findings.

Keywords

Public Procurement; Fiscal Management; Corruption Audits; Reelection Incentives; Set-Asides

Resumo

Sales, Nathalia Machado; Assunção, Juliano; Dahis, Ricardo. **Ensaios em Compras Públicas e Economia Política**. Rio de Janeiro, 2024. 136p. Tese de Doutorado – Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Esta tese é composta por três capítulos: dois dedicados a compras públicas e um à economia política. O primeiro capítulo examina as aquisições realizadas por municípios brasileiros durante a pandemia de COVID-19, quando a legislação flexibilizou os limites para dispensa de licitação. Os dados mostram um aumento significativo no uso de dispensas, especialmente para bens relacionados ao combate da pandemia e em municípios com pior gestão fiscal. Curiosamente, não há evidências de que o aumento nas dispensas tenha levado à seleção de fornecedores associados ao favoritismo ou à corrupção, nem a preços mais altos para produtos relacionados à crise. A maior flexibilidade permitiu que as unidades compradoras acessassem mais fornecedores fora dos limites municipais. Discutimos se essa medida atenuou os impactos na mortalidade em municípios com restrições financeiras. O segundo capítulo analisa uma política brasileira que estabelece a reserva de lotes em leilões exclusivamente para a competição de pequenas e médias empresas (PMEs). A principal conclusão é que o uso de lotes exclusivos para PMEs reduz a concorrência nos leilões. Investigamos também o impacto da política sobre os preços e discutimos a adesão à política. Por fim, no terceiro capítulo, ampliamos tentativas manuais anteriores de classificar relatórios de auditoria de corrupção, utilizando um Modelo de Linguagem de Grande Escala (LLM) para codificar relatórios de municípios brasileiros. Em seguida, utilizamos nossos dados para reavaliar o impacto dos incentivos à reeleição sobre a corrupção. Encontramos algumas evidências de que os incentivos à reeleição reduzem a corrupção, corroborando achados anteriores. No entanto, os efeitos são menores e estatisticamente significativos apenas para uma das três variáveis de corrupção analisadas. Apresentamos explicações alternativas para esses resultados.

Palavras-chave

Compras Públicas; Gestão Fiscal; Auditorias de Corrupção; Incentivos à Reeleição; Tratamento Preferencial

Table of contents

List of figures

List of tables

"Nothing is easier than spending the public money. It does not appear to belong to anybody. The temptation is overwhelming to bestow it on somebody."

Calvin Coolidge, .

1
Public Procurement Under an Emergency: Assessing the Trade-off Between Rules and Discretion

Abstract

Does more discretion lead to worse procurement outcomes? I address this question in the context of Brazilian municipal procurement during the COVID-19 pandemic, where legislation allowed for greater discretion in crisis-related purchases, exceeding usual thresholds. Using tender-level data, I exploit the timing of the shock and the spatial heterogeneity across municipalities provided by a fiscal management index. Following the pandemic's onset, I observe a general rise in the use of tender waivers, particularly for emergency-related goods. Moreover, municipalities with bad management began using a higher share of tender waivers relative to total tenders, compared to those with good management. However, I find no evidence that supplier selection deteriorated through the hiring of firms associated with favoritism or corruption, and there is no statistically significant increase in prices for a set of crisis-related products. Evidence suggests that the greater flexibility allowed public buyers to access more suppliers outside municipal boundaries, yet it was insufficient to mitigate the adverse mortality effects in financially constrained municipalities.

1.1
Introduction

Government purchases constitutes a substantial share of the economic activity. Between 2002 and 2019, the Brazilian government awarded approximately 12% of its GDP procuring goods and services from the private sector (Thorstensen and Giesteira, 2021a) - a percentage in line with the recent global average (Bosio et al., 2022). Most of this expenditure was concentrated within the federal government (7.5%), but municipalities play an important role in this process - municipal procurement was equivalent to 3% of GDP in the same period, or about 25%-30% of total purchases (Thorstensen and Giesteira, 2021a).

Despite the economic significance of government procurement, this topic remains relatively understudied in Brazil. This is particularly evident in more decentralized contexts. The lack of studies on municipal purchases is not coincidental and can be largely explained by the challenge of accessing municipal data, which varies in quality and availability depending on each State Audit Court (TCE). Given that, I contribute to fill this gap by exploring purchases made at the municipal level using a newly created dataset comprising harmonized microdata for more than 2,100 Brazilian municipalities. While my analysis focuses on a specific context, this study represents a first attempt to open the black box of municipal procurement practices.

In this paper, I examine how the relaxation of procurement regulations during the COVID-19 pandemic affected the purchasing behavior of public buyers. During emergencies, public procurement rules are often made more flexible to support governments in increasing spending and reducing damage. Around the world, flexibility has taken various forms, including increased use of negotiated contracts and direct contracting, more flexible pricing strategies, more frequent renegotiation, and expedited timelines (Bandiera et al., 2021b). In Brazil, the federal government began to relax the procurement rules during the initial weeks of the global health crisis. At first, legislation changed to significantly expand the possibilities for bypass tenders and exercise discretion in acquiring goods and services related to the pandemic needs. After a while, flexibility also increased for other types of goods, but at a much smaller magnitude.

In this context, I address several key questions. I start by analyzing the first order effect - the impact of the regulatory changes on the utilization of tender waiver. Second, I investigate if the increase in discretion primarily affected goods outlined in legislation - those related to emergency response - or if there has been policy leakage. Additionally, I explore whether this has led to higher prices for relevant products and if the increased use of exemptions has affected supplier selection, such as favoring ineligible, local, or politically connected firms.

To answer these questions, I exploit data of more than 280,000 purchases made by Brazilian municipalities' between January 2019 and December 2020. One empirical challenge in this context is the lack of a control group since all the municipalities were affected by both the virus and the law.[1.1] Therefore, my identification strategy leverages on weekly longitudinal variation across municipalities, exploiting the timing of the first detected case of COVID-19 in the country, as well as the cross-sectional variation provided by the

[1.1] Although municipalities were not affected at the same time, I chose not to use a staggered Difference-in-Differences approach because the timing of the emergency was not random. There was a clear pattern of which municipality was infected at first. As illustrated in Figure 1.A.6, the municipalities that reported the first cases were those in the first quartile of population distribution.

baseline fiscal capacity of each municipality. This setting captures the average differential change in procurement outcomes, before and after the pandemic outbreak, in municipalities with high baseline fiscal management capacity relative to municipalities with low fiscal management capacity.

In the sections below, I first document that there was a rapid increase in discretion following the pandemic outbreak. The average proportion of purchases carried out through tender waiver increased from 24% to 39%. As expected, the rise in discretion was much greater for emergency-related goods, jumping by 34 percentage points and reaching an average of 57% in the post-pandemic period. In terms of fiscal management, I find that both good and bad management municipalities experienced an increase in the use of tender waiver. However, the increase was more pronounced for the latter group, whether considering all products (+2.9 p.p) or restricting the sample to crisis-related goods (+4.8 p.p). Analysis of emergency goods further reveals that this effect was predominantly driven by purchases of pharmaceutical drugs, hospital materials, and personal protective equipment.

To evaluate whether increased autonomy led public buyers to purchase cheaper or more expensive products, I analyze tenders with detailed item-level data, focusing on a sample of standardized healthcare-related products. Despite the homogeneous nature of these products, I show that some exhibit considerable price variation. Contrary to expectations, however, only a few show differences before and after the pandemic. Furthermore, I can not reject the hypothesis that there is no differential effect on prices between competitive and non-competitive tenders. Although the coefficients are positive, indicating higher prices under discretion, they are not statistically significant.

When looking at supplier selection, I find no evidence that discretion increases the likelihood of hiring firms more susceptible to corruption. The estimates show that tender waivers do not have a significant impact on the chances of contracting firms deemed ineligible for procurement. This is an interesting finding, given that one of the precedents established by the legislation was the possibility of hiring debarred suppliers in the case of being the only available option. Additionally, my results indicate that tender waivers do not raise the likelihood of contracting firms judged ineligible following contract award, which could suggest that these purchases were more prone to irregular practices.

Another concern in public procurement is the existence of favoritism. Favoritism is an issue when there is a tendency to prioritize politically connected firms or local firms without adequate justification or transparent evaluation criteria, often for reasons unrelated to quality or efficiency. One common argument is that decision-makers may bypass competitive processes to favor specific suppliers. Still, I find no significant effect indicating that discretion raises the likelihood of contracting politically connected firms. Moreover, purchases made through tender waivers are 13 percentage points less likely to be awarded to local firms compared to competitive tenders.

Finally, I examine how municipalities' management status and the use of tender waivers during the pandemic affected mortality outcomes, measured by excess mortality per 100,000 inhabitants and its percentage variation (the p-score). I show that municipalities with worse fiscal management experienced, on average, an increase of 3 deaths per 100,000 inhabitants when compared to those with good management, in line with previous findings from Barros Barbosa et al. (2022). Moreover, the interaction between prior fiscal capacity and the utilization of tender waivers during the pandemic suggests that increased discretion may have been ineffective in alleviating the adverse effects of fiscal mismanagement.

This paper is mostly related to the literature that studies the effects of discretion on public procurement. How much public procurement should be regulated and to what extent buyer discretion should be allowed has been one of the longstanding debates in the literature. The optimal approach varies depending on institutional context (Bosio et al., 2022) and whether agents are aligned with public service goals (Kelman, 1990; Bandiera et al., 2009; Carril et al., 2021).[1.2]

In general, the existing literature presents both positives and negatives effects generated by less stricter rules in a variety of countries. While discretion can lead to inefficiencies, favoritism, and corruption, it can also lead to improved contract performance. Decarolis et al. (2020) and Coviello et al. (2018) shed light on both aspects of discretion in the Italian context. The former shows that purchases made under discretion increases the risk of corruption as contracts are more likely to be awarded to firms under investigation, and results in higher awarded prices. However, it also reduces cost overruns and delays. The later, document that discretion increases the probability that the same firm wins repeatedly, yet this does not deteriorate procurement outcomes and, in fact, reduces service execution time.

Some studies emphasize the costs of discretion outweighing its benefits.

[1.2]When agents are aligned, rules can produce worse results than discretion. In that case, having more flexibility to adapt to various situations would be preferable (Carril et al., 2021). When agents are misaligned they might behave sub-optimally, either because they are corrupt, allowing higher prices to extract rent, or because they are lazy and act inefficiently thus leading to higher prices. In the second case, strict rules can backfire. Imposing a higher cost on lazy agents may increase inefficiency and lead to worse outcomes (Bandiera et al., 2009).

Palguta and Pertold (2017a) argue that discretion leads to increased rent-seeking by anonymously owned firms in the Czech Republic. Similarly, Szucs (2023) documents that discretion increases the selection of less productive contractors and politically connected firms, without improving ex-post contract execution in Hungary. In contrast, other studies emphasize the advantages of discretion over its costs. Bandiera et al. (2021a) presents findings from an experiment in Pakistan, in which giving more autonomy to officials reduces prices by 9% without reducing the quality of goods. Additionally, Fazio (2022), finds that discretionary contracts in Brazil may have higher prices but often result in better-quality purchases, suggesting potential efficiency gains.

Concerning the COVID-19 pandemic, there is still little evidence. Drawing on data from Colombia, Gallego et al. (2020) observe that municipalities with a higher risk of corruption in the baseline tend to respond using a larger proportion of non-competitive contracts during the pandemic, especially for crisis-related purchases. In these locations, contracts are more prone to experiencing cost overruns, being awarded to campaign donors, and exhibiting budget and time extensions. I contribute to this literature by providing evidence from Brazil, one of the countries hardest hit by the pandemic. In the Brazilian case, the use of non-competitive methods increased significantly, yet there is no evidence that these contracts were more likely to be awarded to high-risk suppliers. Moreover, prices for products purchased through discretion are not statistically significantly higher compared to those bought under competition.

More specifically, this paper contributes to the ongoing debate regarding the balance between regulation and discretion in emergency situations (Bandiera et al., 2021b). Emergency procurement has historically been related to misconduct, and while discretion may worsen these issues, it also allows for quicker delivery of supplies. It is crucial to balance the benefits of transparency and competition with the urgency of saving lives. Regarding this point, I show that municipalities with lower initial fiscal capacity were more severely affected by the pandemic, experiencing worse mortality outcomes, while the legislative flexibility failed to reduce the gap with good management municipalities.

Finally, this paper relates to a growing literature discussing how episodes of crisis and catastrophic events, such as wars (Querubin et al., 2011), natural disasters (Leeson and Sobel, 2008) and epidemics (Khemani, 2020) increases rent-seeking and imposes governance challenges. The most common prediction is that opportunistic behaviors become more pronounced when governments have to quickly spend large amounts of resources, especially through exceptional regimes. In Brazil, potential misallocation in the procurement for health supplies has come under scrutiny and has gained high levels of public attention uncovering a lot of suspicious cases.[1.3] Despite the anecdotal evidence, there is no indication of rent-seeking in my setting. In fact, the overall percentage of firms punished after awarding an emergency-related contract decreased, and the share of politically connected suppliers remained stable during the pandemic.

The remainder of this paper is divided as follows. Section 1.2 provides an overview of the Brazilian public procurement system and discusses the institutional context of legislative changes during the pandemic. Section 1.3 describes the data sources and the main sample used for analysis. Section 1.4 details the identification strategy and presents findings on the impact of fiscal management and discretion on procurement outcomes. Section 1.5 focuses on health outcomes, detailing the identification strategy and presenting related results. Finally, in Section 1.6, I discuss the conclusions drawn from the study.

1.2 Institutional Background

1.2.1 General Framework

Brazil has a comprehensive legislation that establishes general rules for public tenders and contracts across all levels of government - federal, state, and municipal. For almost twenty years public procurement was governed by Law No. 8,666/1993. During the period covered in this paper, this was the prevailing legislation. In 2021, a new law was enacted, introducing a new legal framework for public procurement. The idea was to update the former legislation towards making the purchase process faster and more efficient. Both laws remained in effect until the end of 2023, with only the new law staying in force from 2024 onward.

At the municipal level, procurement oversight is conducted by the State Audit Courts (TCEs, henceforth). These independent bodies are established in each of Brazil's twenty-six states plus the Federal District.[1.4] The primary responsibility of TCEs is to ensure transparency, accountability, and legality in the management of public funds and resources. In addition to supervising

[1.3]This BBC report shows that in June 2020 there were already several corruption investigations in seven states, totaling more than 1 billion BRL.

[1.4]In the states of Rio de Janeiro and São Paulo, the TCEs oversee all municipalities apart from the capitals, which have their own audit court. Three states – Bahia, Goiás, and Pará – have two audit courts: one that oversees the state government and one that oversees all municipal governments within the state.

public tenders, these courts are tasked with supervising the fiscal policies of states and municipalities, including taxation, spending, and budget execution.

Law 8,666/1993 provided for seven different types of procurement methods. As a rule, governments are expected to run competitive tenders. The specific method used, such as reverse auctions, invitations to tender or framework agreements, depends on the nature of the object to be acquired and its estimated value.[1.5] In exceptional circumstances or when purchase values are small, officials can waive tenders and directly contract with suppliers.

According to Dahis et al. (2023), approximately 30%-40% of municipal purchases are made through non-competitive methods, yet they account for less than 20% of the total purchase value. Non-competitive tenders encompass two categories: non-requirement tenders (*inexigibilidade*) and tender waivers (*dispensa*). The former is applicable when competition is impractical, such as when there is only one available supplier or for the acquisition of unique goods or specialized professional services. The latter occurs when competition is feasible, but the government chooses not to carry out the tender process. This is possible for small-value purchases - under a specific threshold established by procurement law - or in emergency situations. Under the old legislation, waivers were allowed for purchases below R$ 17,600 for general goods and services and R$ 33,000 for engineering services.

Among competitive methods, reverse auctions stand out as the most common. Since 2005, auctions have been mostly done electronically, leading to a significant reduction in procurement participation costs and increase in transparency. In this procedure, tender notice is freely available on the internet and participants can see others' bids, but do not have access to the identity of their competitors.

1.2.2 Changes in Procurement Regulations

At the onset of the pandemic, a set of measures concerning public procurement started to be implemented by the government. In early February 2020, before any official cases were confirmed in Brazil, the government anticipated and enacted Law 13,979/2020. This law provided general regulations to address the public health emergency, such as social distancing and quarantine, but also extended the possibility for direct contracting, adding a new scenario for exceptions to those already outlined in standard legislation.

The law set a great precedent for the acquisition of **health** goods and services necessary for combating the pandemic. For such acquisitions, tender

[1.5] See Table 1.A.4 for more details.

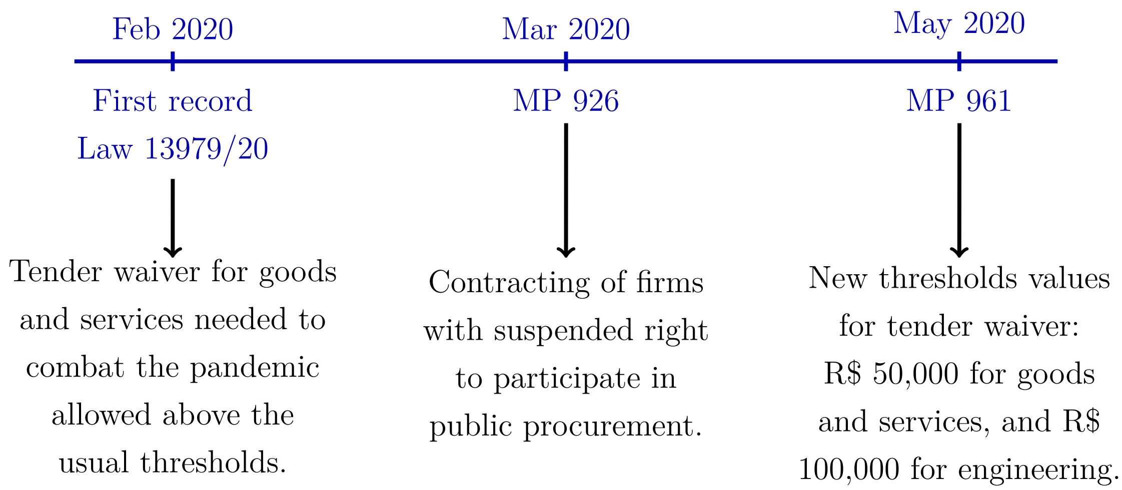
waivers were permitted beyond the usual limits established by the general rule. This new form of contracting was settled to be temporary and to last until the end of the public health emergency situation.[1.6] It is worth mentioning that Law 13,979 was based on the premise of an existing state of calamity. In other words, there was no need to declare or prove a state of calamity for its application. However, similar to Law 8,666, procedural justification was still required for its utilization.

In March 2020, Law 13,979 was amended by *Medida Provisória* 926 (MP 926), bringing some important changes. First, it clarified that tender waivers, applicable to amounts exceeding specified thresholds, could extend to the acquisition of **all** goods and services necessary to address the emergency, including those related to engineering. The law also established a precedent allowing the hiring of suppliers previously declared ineligible and debarred from procurement participation in cases where they were the only available supplier. Additionally, if there were difficulties in finding suppliers, the authority had the option to waive the requirement for tax and labor documentation, provided that a specific justification was given. Finally, for auctions intended to procure goods and services for emergency purposes, the law stipulated that the deadlines for procedures, whether electronic or in-person, should be halved.

All the measures mentioned above were intended to address the significant demand directly generated by the pandemic without changing the framework for other public contracts. To mitigate the effects of the economic crisis that accompanied the public health emergency, a second *Medida Provisória* (MP 961) was passed in May 2020. This MP raised the overall thresholds for which tender waivers were allowed during the state of public calamity. For general goods and services the threshold increased from R$ 17,600 to R$ 50,000, while for procurement related to engineering, it rose from R$ 33,000 to R$ 100,000. The timeline containing the changes in procurement regulation is depicted in Figure 1.1.

[1.6] In other emergency contracts not related to addressing the pandemic needs, Law 8,666/1993 continued to apply.

Figure 1.1: Timeline of Legislation Changes



1.3 Data

1.3.1 Public Procurement

All public procurement variables comes from *Microdados de Despesas de Entes Subnacionais (MiDES)*, a dataset that harmonize microdata on public procurement and budget execution collected from State Audit Courts.[1.7] The data contains information of all public purchases made by more than 2,100 municipalities, including competitive and non-competitive methods, spanning different coverage periods. In this paper, I limit the sample to the years 2019 and 2020, focusing specifically on product purchases and excluding those related to services or construction. Moreover, I use five of the six states available, which are Ceará (CE), Pernambuco (PE), Minas Gerais (MG), Rio Grande do Sul (RS) and Paraná (PR).[1.8] With these data, I can identify the notice date, purchase method, tender description, total value, and suppliers for each tender.

Using textual analysis on tender description, I identify and manually classify purchases directly related to the pandemic - emergency goods henceforth. These goods includes personal protective equipment (PPE), pharmaceutical drugs, medical equipment, hospital materials[1.9], antigen tests, cleaning mate-

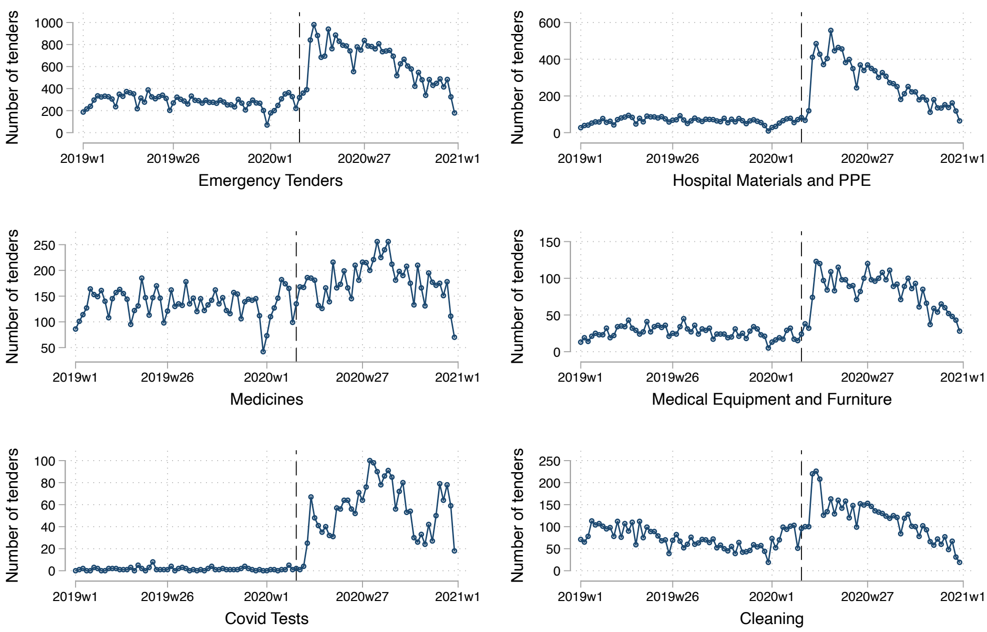
[1.7] More detailed information can be found in Dahis et al. (2023). The dataset is available on Base dos Dados.

[1.8] The state of PB was excluded due to a lack of information on tender notice or homologation date, turning it unsuitable for constructing a weekly panel.

[1.9] Besides PPE, hospital materials may include thermometers, nasogastric tubes, catheters, and etc.

rials, and others related to the health infrastructure, such as ambulances and hospital beds. The first panel of Figure 1.2 displays the increase in the number of tenders for emergency goods after the first COVID-19 case. The same can be observed in the other panels, where emergency goods are broken into different categories. While hospital materials and PPE peak in the first few weeks, the purchase of medicines and COVID tests peaks much later.

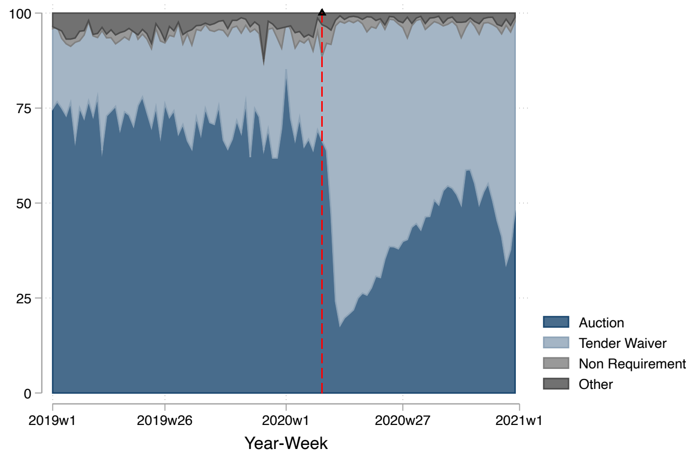
Figure 1.2: Number of Tenders Classified as Emergency-Related (2019-2020)



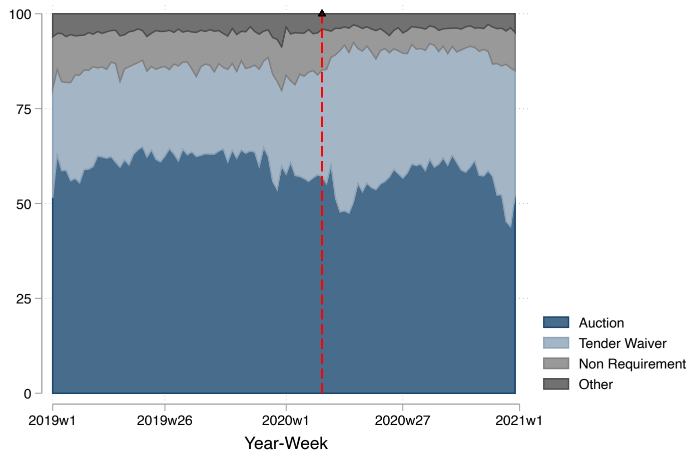
Note: This figure shows the weekly count of tenders classified as emergency-related from January 2019 to December 2020, considering all procurement methods. The dashed line marks week 9 of 2020, which corresponds to the week of the first recorded COVID-19 case in Brazil. The classification of these tenders was achieved through text analysis of the tender descriptions (Refer to Section 1.3.1 for further details). Quadrants 2 through 6 represent different categories of goods within the emergency definition. The first quadrant aggregates the counts from quadrants 2 to 6.

Figure 1.3 depicts the usage of competitive and non-competitive methods over time for emergency (Panel A) and non-emergency (Panel B) goods. The later category includes all the products not classified within the emergency definitions mentioned above. In the case of emergency goods, there was a substantial reduction in the use of auctions, accompanied by an increase in tender waiver. Non-requirement and other methods remained almost stable. Conversely, for non-emergency goods, this movement was much more subtle, reflecting the changes in the exemption thresholds for general purchases promoted by MP 961. This underscores an interesting point: there appears to have been no policy leakage, as the legislation was designed to increase flexibility in the procurement of goods necessary to combat the pandemic in the first place.

Figure 1.3: Use of Purchase Methods Over Time



A: Emergency Tenders



B: Non-emergency Tenders

Note: This figures illustrates the composition of procurement methods from January 2019 to December 2020. The category "Other" encompasses all methods not included in Auctions, Waiver or Non-Requirement (See Table 1.A.4 for a detailed description of procurement methods). The red dashed line marks week 9 of 2020, which corresponds to the week of the first recorded COVID-19 case in Brazil. Emergency tenders include products from categories such as hospital materials and PPE, medicines, medical equipment and facilities, COVID tests, or cleaning materials (See Section 1.3.1 for details). Non-emergency tenders refer to those not classified as emergency-related.

For all states except Pernambuco, I also have detailed information on items purchased in each tender. Along with product descriptions, the data includes quoted price, unit price, quantity, and their respective units of measure. A limitation of this data is the absence of standardization for items, as there is no coding or catalog system in place. As the description field is open, the same product might have different spellings, imposing a series of challenges to classify them. The process of refining raw item data to reach the tractable sample used in this paper can be summarized in a few steps.

I start by narrowing the data to the tenders classified as emergency related. This is possible because both the aggregated tender-level data and the item-level data share a common ID, allowing them to be connected. The result is a dataset with more than 1.7 million of items purchased from 2019 to 2020. After removing stopwords, special characters, and supply units from the item descriptions, I generate a frequency list using the first two words of each remaining description.[1.10] From the 50 most common products, I select those recognized for their homogeneity.[1.11] These include 13 pharmaceutical drugs and 13 materials, some of which are standardized by regulation, such as N95/PFF2 masks and ethyl alcohol, and others with minimal product differentiation, like detergent and garbage bags. Additionally, I select 3 medicines that gained popularity during the pandemic due to their perceived potential for treating the disease, such as Chloroquine, Azithromycin, and Ivermectin (known as COVID-19 kit). The list containing all selected products is available in Table 2.A.6.

Following the approach adopted by Fazio (2022), I define a product as the combination of the item itself plus its unit measure. This implies that products are considered the same only if they share the same unit measure. Therefore, for each product, I select the most common and standardized unit, discarding the remaining observations. For example, all pharmaceutical drugs in my selected sample are bidden in Pills. Pills are selected for their frequent occurrence among the most common units and their higher level of standardization, as they are typically sold in single units. In contrast, other forms such as liquid oral and injectable formulations are often sold in ampules or syringes of different sizes (Fiuza et al., 2023).

Finally, for all selected items, I remove the top and bottom 1% of unit prices, resulting in a dataset with 63,140 observations. Among these, 6,946 were lacking unit price information, leading to a final dataset of 56,194 items with available prices. It is noteworthy that 68.2% of these items were purchased through auctions, compared to just 8.8% acquired via tender waivers. Even though the sample comprises only 29 products, it represents a significant monetary value, totaling more than 1 billion BRL in the analyzed period.

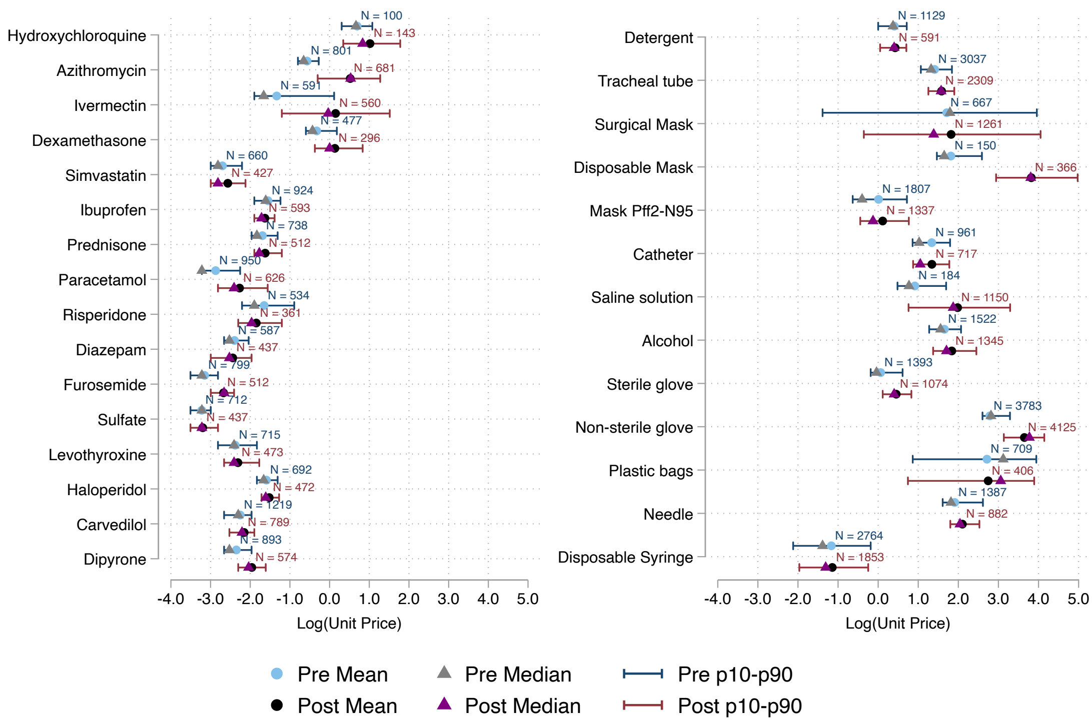
Figure 1.4 shows that despite the homogeneous nature of the products, some of them exhibit significantly price variation. This suggests that different

[1.10] Despite extensive efforts to standardize descriptions, the cleaning process results in 120,000 unique product entries based on the first two words. Filtering for those mentioned 10 or more times reduces the list to 15,000 unique products. The most frequently mentioned are garbage bags (12,556 occurrences), suture thread (7,182 occurrences), and urethral probe (7,178 occurrences).

[1.11] This follows the usual approach found in the literature (Szerman, 2012; Bandiera et al., 2021a; Fazio, 2022; Fiuza et al., 2023).

bureaucrats pay different amounts for some identical or very similar products.[1.12] Second, contrary to expectations, I observe little change before and after the pandemic. For most common medicines, price variation is small and shows no substantial differences across the pandemic period. However, for medicines in the COVID-19 kit, such as Hydroxychloroquine, Azithromycin, and Ivermectin, there was an increase in price dispersion, accompanied by a rise in the average price. In the acquisition of materials, the greatest differences are observed for disposable masks, saline solution and non-sterile gloves. Some products like surgical masks and plastic bags exhibit substantial price variation even before the crisis.

Figure 1.4: Summary Statistics of Selected Products



Note: The figure presents summary statistics (mean, median, 10th percentile, and 90th percentile) for the log-transformed unit prices of all products listed in Table 2.A.6 before and after the COVID-19 pandemic. Prices below the 1th percentile and above the 99th percentile were excluded from the analysis.

1.3.2 Fiscal Management Index

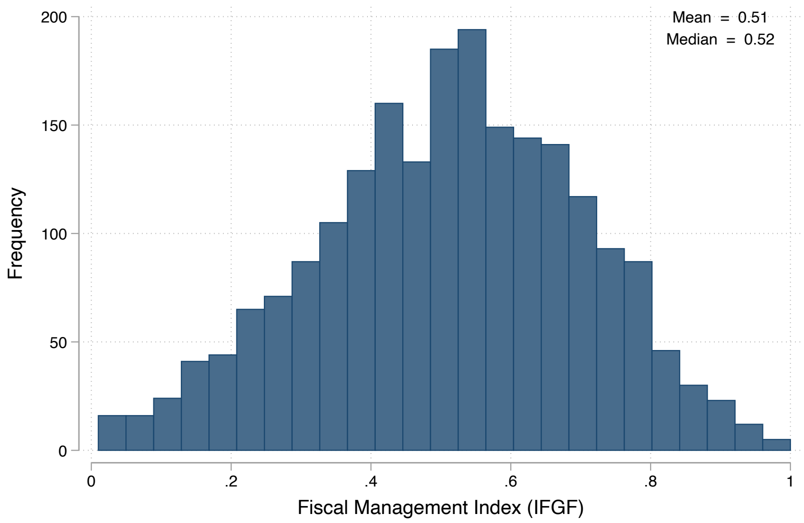
The fiscal management index (*Índice de Gestão Fiscal*) was obtained from Firjan. The index evaluates the financial performance of Brazilian municipalities annually based on data from the *Sistema de Informações Contábeis e Fiscais do Setor Público Brasileiro* (Siconfi). It is composed of four equally

[1.12] The observed price dispersion for homogeneous goods is not uncommon in the public sector and have been documented in previous studies (Bandiera et al., 2009, 2021a).

weighted indicators: Autonomy, Payroll, Liquidity, and Investments. Autonomy measures a municipality's capacity to manage its administrative structure using its own revenue, while Payroll measures the portion of revenue allocated to personnel expenditures. Liquidity indicates whether the municipality has sufficient cash flow to cover remaining balance, and Investment assesses the proportion of revenue invested. Ranging from 0 to 1, a higher score denotes better municipal management. The index distribution is illustrated in Figure 1.5.

From this metric, I construct the bad management indicator variable employed in the regressions. Municipalities falling below the 50th percentile of the management index are assigned a value of one, while those above the 50th percentile are assigned a value of zero. As shown in Figure 1.5, the median value of the fiscal management index is 0.52.

Figure 1.5: Management Index Distribution



Note: This figure displays the distribution of the Fiscal Management Index (IFGF) obtained from FIRJAN for municipalities in 2019. Higher scores indicate better fiscal management performance.

A natural question is whether this index correlates with other municipal characteristics. To investigate this, I perform a regression of the index on several variables: the logarithm of GDP per capita, population size, proportion of revenue allocated to healthcare, total healthcare expenditure, latitude, longitude, municipal HDI, and the percentage of skilled workers in the municipality's direct administration. The residuals of this regression are displayed in Figure 1.A.5. While these variables account for one-third of the index variation, a substantial amount remains unexplained.

Table 1.1 presents the main descriptive statistics of municipalities' characteristics in the baseline. As expected, municipalities with bad management exhibit lower GDP per capita and smaller population size. They also allocate a smaller portion of their revenue to healthcare and have fewer hospital beds per 100,000 inhabitants. In terms of bureaucracy quality, these municipalities have 37% high-skilled employees working in direct administration, which is 5 percentage points lower than municipalities with good fiscal management.

Table 1.1: Baseline Descriptive Statistics at the Municipality Level

	All	Bad Mgmt	Good Mgmt
GDP per capita	23,533.79	15,915.00	31,425.30
	(23,312.74)	(11,909.85)	(28,931.10)
Population	29,588.83	18,478.13	41,098.57
	(114,275.86)	(29,361.48)	(159,469.23)
Total committed value in healthcare per capita (log)	6.76	6.67	6.86
	(0.40)	(0.39)	(0.38)
Share of municipal revenue applied to health	22.32	22.01	22.66
	(4.52)	(4.51)	(4.46)
Hospital beds per 100,00 inhabitants	124.10	109.92	138.59
	(158.81)	(146.55)	(169.40)
Share of high skill employees	0.39	0.37	0.42
	(0.11)	(0.10)	(0.11)
Management Index	0.51	0.36	0.67
	(0.19)	(0.12)	(0.10)
Autonomy Index	0.42	0.19	0.65
	(0.37)	(0.26)	(0.33)
Payroll Index	0.59	0.44	0.74
	(0.29)	(0.26)	(0.22)
Liquidity Index	0.59	0.47	0.72
	(0.30)	(0.31)	(0.21)
Investment Index	0.46	0.35	0.57
	(0.26)	(0.20)	(0.27)
Observations	2,118	1,076	1,041

Note: This table reports baseline descriptive statistics at the municipality level. The Fiscal Management Index (IFGF) data is from 2019. Municipalities are categorized as 'Bad Management' if they fall below the median of the index, and 'Good Management' if they are above the median. Additional information includes GDP, population, the share of municipal revenue allocated to health, total committed healthcare expenditure, and the number of hospital beds, all from 2019. The share of high-skilled employees in public administration is derived from the 2018 Survey of Basic Municipal Information (MUNIC). For further details on the data, see Section 1.3.

1.3.3 Other Datasets

This paper also uses a couple of other datasets. I gather firms' registry data from the *Cadastro Nacional da Pessoa Jurídica*, a database managed by the *Receita Federal*, which contains information on firms' location, legal structure, and owners for the universe of Brazilian firms.[1.13] I merge firms'

[1.13]The *Cadastro Nacional da Pessoa Jurídica* is available on Base dos Dados.

registry data with procurement data using suppliers' 14-digit tax identifiers (CNPJ). Using the firms' location, I construct a measure of local suppliers. A supplier is considered local if any of the firms' branch is situated in the same municipality as the buyer unit.[1.14]

To construct the measures of political connection, I merge procurement and firms' registry data with electoral data from *Tribunal Superior Eleitoral* (TSE). Considering the four-year political term of Brazilian mayors, those in office in 2020 were in their final year and were elected in 2016. Then, I collect data from the 2016 municipal election, including mayoral election results, party coalitions, campaign expenses, and donations. A firm is considered politically connected if it satisfies at least one of the following criteria:

1. The mayor was the owner of the supplier firm.
2. Members of the mayor's party or the mayor's coalition parties were owners of the supplier firm.
3. The firm donated directly to the mayor or to the mayor's coalition in the 2016 election (CNPJ donation, prohibited from 2016).
4. Partners of the firm donated directly to mayors or to the mayor's coalition in the 2016 election (CPF donation).
5. The firm was hired for the mayor's election campaign.

Additionally, I collect data from the *Cadastro de Empresas Inidôneas e Suspensas* (henceforth, CEIS), public available on the Transparency Portal. This dataset contains records of sanctions imposed on both establishments and individuals involved in irregular activities, which may include misconduct in tender procedures, fiscal fraud, and contract fraud with the public administration. Once the establishment is listed and officially debarred, it is no longer allowed to have contracts with the government until the sanctions expire. By using the start and end dates of the sanctions, I create two variables: firms punished before being contracted, potentially indicating worse selection, and firms punished after being awarded a contract, possibly indicating irregularities during the contract period.

Other independent variables used in the regressions were obtained from various sources. GDP per capita, population, HDI, and the percentage of skilled workers in the municipality's direct administration are all sourced from

[1.14] Some of the suppliers have a 11-digit tax identifier (CPF) and then can not be identified. As there is a high probability that these suppliers are local, we are underestimating the degree of the same municipality suppliers.

IBGE.[1.15] The total amount committed to healthcare comes from MiDES. Healthcare variables, such as the proportion of municipal revenue applied to health, hospital beds rate, COVID-19 cases, and deaths come from DATA-SUS.[1.16]

The main descriptive statistics are presented in Table 1.2. On average, 24% of purchases were made through tender waiver in the pre-pandemic period. This number increased to 39% after the pandemic. Upon examining the groups categorized by good and bad management, it is possible to see that both exhibit similar percentages in both periods. However, municipalities with worse management exhibited slightly higher growth in the utilization of tender waivers, with an increase of 16 percentage points compared to 14 percentage points for those with better management. In the sub-sample of emergency goods, which increased the participation from 11% to 23% of total tenders, the growth of tender waiver was even higher, going from 23% to reaching 57%. In this case, bad and good management municipalities show an increase of 40 and 30 percentage points, respectively.

Table 1.2 also presents some characteristics of the winning firms for the sample of emergency tenders. On average, the percentage of tenders awarded to a punished (ineligible) supplier is virtually zero. This remains consistent both before and after the pandemic, across both groups. Therefore, this data fails to capture any increase in the hiring of punished firms in response to MP 926. In addition, the percentage of firms debarred from participating in procurement after being awarded a contract decreases from 17% to 12%. Surprisingly, both groups experience a similar decline.

Equally interesting is the low average of tenders involving politically connected firms across both periods. Only 1% of emergency tenders were awarded to politically connected firms, both before or after the pandemic. A slight decrease of 1 percentage point is observed for municipalities with bad fiscal management.

In terms of selecting local suppliers, there is a considerable decline in the proportion of tenders awarded to them, decreasing from 26% to 20%. While municipalities with bad management contract fewer local firms, potentially reflecting their smaller size and limit availability of local suppliers, both groups experience a reduction during the pandemic.

[1.15]The information on skilled and unskilled workers in the municipality's direct administration is provided by the *Pesquisa de Informações Básicas Municipais* (MUNIC) from 2018.

[1.16]The share of municipal revenue applied to health and the information on hospital beds were obtained from IEPS Data website.

Table 1.2: Descriptive Statistics at the Tender Level

	All		Bad Mgmt		Good Mgmt	
	Pre	Post	Pre	Post	Pre	Post
Panel A: All Tenders, Products						
Share of Tender Waiver	0.24	0.39	0.22	0.38	0.25	0.39
	(0.43)	(0.49)	(0.41)	(0.49)	(0.43)	(0.49)
Share of Reverse Auction	0.62	0.53	0.67	0.55	0.60	0.51
	(0.49)	(0.50)	(0.47)	(0.50)	(0.49)	(0.50)
Share of Tenders Related to Emergency	0.11	0.23	0.12	0.24	0.11	0.23
	(0.32)	(0.42)	(0.32)	(0.43)	(0.31)	(0.42)
Avg Tender Waiver Value (log)	8.54	8.97	8.65	9.14	8.49	8.88
	(1.84)	(1.76)	(1.86)	(1.71)	(1.83)	(1.78)
Share of Tenders With Supplier Information	0.75	0.77	0.71	0.75	0.77	0.78
	(0.43)	(0.42)	(0.45)	(0.43)	(0.42)	(0.41)
# Tenders	161,228	124,423	53,372	40,068	107,856	84,355
Panel B: Emergency Tenders, Products						
Share of Tender Waiver	0.23	0.57	0.14	0.54	0.28	0.58
	(0.42)	(0.50)	(0.35)	(0.50)	(0.45)	(0.49)
Share of Reverse Auction	0.71	0.40	0.81	0.43	0.66	0.38
	(0.45)	(0.49)	(0.39)	(0.50)	(0.47)	(0.48)
Avg Tender Waiver Value (log)	8.44	9.27	8.65	9.38	8.38	9.21
	(1.73)	(1.73)	(1.56)	(1.63)	(1.77)	(1.77)
Share of Tenders With Supplier Information	0.84	0.83	0.80	0.82	0.86	0.84
	(0.37)	(0.38)	(0.40)	(0.39)	(0.35)	(0.37)
Share of Punished Firms (before)	0.00	0.00	0.00	0.00	0.00	0.00
	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)	(0.03)
Share of Punished Firms (after)	0.17	0.12	0.17	0.12	0.16	0.12
	(0.37)	(0.33)	(0.38)	(0.32)	(0.37)	(0.33)
Share of Politically Connected Firms	0.01	0.01	0.02	0.01	0.01	0.01
	(0.11)	(0.09)	(0.12)	(0.10)	(0.10)	(0.08)
Share of Local Firms	0.26	0.20	0.20	0.13	0.28	0.23
	(0.44)	(0.40)	(0.40)	(0.34)	(0.45)	(0.42)
# Tenders	18,041	28,885	6,250	9,747	11,791	19,138

Note: This table reports descriptive statistics at the tender level. Municipalities are categorized as "Bad Mgmt" if they fall below the median of the 2019 Fiscal Management Index (IFGF), and "Good Mgmt" if they are above the median. To calculate the average value of tender waiver contracts, values below the 5th percentile and above the 95th percentile were excluded. The first four variables include all tenders, while variables below 'Emergency' are restricted to tenders classified as emergency-related (See Section 1.3.1 for details). In both panels, only goods (products) are included. Panel B includes only reverse auction and tender waiver methods. Standard deviations appear in parenthesis.

1.4 The Impact of Fiscal Management and Tender Waivers on Procurement Outcomes

1.4.1 Empirical Strategy

Assuming that the pandemic caused an exogenous shock to public procurement at the local level, I employ a modified version of the difference-in-differences method to analyze the impact of this shock on the utilization of discretionary contracts and supplier selection. The traditional difference-in-differences method compares changes in outcomes over time between a population experiencing the shock (the treatment group) and one that does not (the comparison group). In this scenario, nearly all municipalities were affected by the pandemic, making it challenging to identify pure treatment and control groups. Nevertheless, we can expect that some places increased their use of discretionary methods more than others. This difference could occur simply because some areas were more severely affected, requesting more urgency in purchases, or due to opportunistic behavior.

To explore possible differential effects, I use the fiscal management index presented in Section 1.3.2. The index is positively correlated with GDP, thus partially indicating that municipalities are low or high income, but can also reflect state capacity and bureaucratic quality. For instance, we may examine if municipalities with lower levels of fiscal management were more inclined to use discretionary contracts during the emergency period. Additionally, given that purchases were made through discretionary procedures, we can explore whether these municipalities exhibited a higher propensity for opportunistic behavior when procuring goods.

Therefore, my identification strategy leverages on weekly longitudinal variation across 2,108 municipalities, exploiting the timing of the first detected case of COVID-19 in the country, as well as the cross-sectional variation provided by the baseline fiscal capacity of each municipality. Employing this variation of difference-in-differences— one difference with spatial heterogeneity— I compare procurement outcomes, before and after the outbreak of the pandemic, in municipalities with low baseline fiscal management relative to municipalities with high baseline fiscal management. Based on this framework, I estimate the following equation:

$$y_{i,m,t} = \alpha_m + \lambda_t + \beta(\text{Post Outbreak}_t \times \text{Bad Mgmt}_m) + X_{m,t} + \epsilon_{i,m,t} \quad (1.1)$$

Where $y_{i,m,t}$ denotes the outcomes for tender i in municipality m during week t. The outcomes encompass an indicator for tender waiver, as well as a set of supplier-related characteristics, including an indicator for punished, politically connected and local firms. The term Post Outbreak_t is an indicator variable that takes the value of one after the first COVID-19 record, remaining constant across municipalities. The variable Bad Mgmt_m denotes an indicator variable in the main specification. Municipalities falling below the 50th percentile of the management index in 2019 are assigned a value of one, while those above are assigned a value of zero. As a robustness check, I also provide the same estimates weighted by Inverse Probability Weighting in the appendix, after matching municipalities by population size.

When employing the indicator for tender waiver as the dependent variable, I estimate the regression using three different samples: one containing all goods, and the other two consisting only of emergency and non-emergency goods, respectively. In the case of suppliers' outcomes, the analysis is restricted to emergency goods. The identifying assumption guiding this empirical strategy is the same as in the traditional difference-in-differences: in the absence of a pandemic outbreak, the outcomes in municipalities with high and low baseline management index would have followed similar trajectories in outcomes.

All regressions include year-week fixed effects (λ_t) and municipality fixed effects (α_m) to control for changes over time that are common to all municipalities and for unobserved and time-invariant characteristics of each municipality, respectively. Furthermore, all regressions include controls for municipal characteristics ($X_{m,t}$), such as GDP per capita, hospital bed density, total healthcare expenditure, the proportion of municipal revenue allocated to healthcare, and COVID-19 incidence, measured as new cases per 100,000 inhabitants.

In order to assess the impact of the pandemic shock on prices, I estimate the Equation 2.1 using item-level data. In that case, the subscript i represents the item rather than the tender and the dependent variable is the is the natural logarithm of the unit price. This regression is estimated using the sample of goods outlined in Table 2.A.6. In addition to incorporating year-quarter and municipality fixed effects, and controlling for municipal characteristics, this specification also includes the interaction between product category and year-quarter to capture variations in trends across different product types.

While this exercise provides valuable insights, the specification from Equation 2.1 does not explicitly address whether the use of discretion directly results in better or worse outcomes. To analyze the differential effect between municipalities with bad and good management, given a procurement awarded by tender waiver, I estimate the following triple differences equation:

$$\begin{aligned} y_{i,m,t} = \alpha_m + \lambda_t &+ \beta_1(\text{Post Outbreak}_t \times \text{Bad Mgmt}_m) \\ &+ \beta_2(\text{Post Outbreak}_t \times \text{Bad Mgmt}_m \times \text{TW}_i) \\ &+ \beta_3(\text{Post Outbreak}_t \times \text{TW}_i) \\ &+ \beta_4 \text{TW}_i + X_{m,t} + \epsilon_{i,m,t} \end{aligned} \tag{1.2}$$

Where $y_{i,m,t}$ represents the same outcomes as Equation 2.1, excluding the tender waiver indicator (TW_i), which now is an independent variable. The parameter of interest is β_2. The first difference is basically the difference in outcomes between bad and good management municipalities before and after the pandemic when using tender waiver. The second difference denotes the difference between both groups, before and after the pandemic, when using competitive methods. By subtracting both terms, I get the effect of discretion.

$$\begin{aligned} \beta_2 = [(&E[Y|\text{Bad Mgmt} = 1, \text{TW} = 1, \text{Post} = 1] - E[Y|\text{Bad Mgmt} = 1, \text{TW} = 1, \text{Post} = 0]) \\ &- (E[Y|\text{Bad Mgmt} = 0, \text{TW} = 1, \text{Post} = 1] - E[Y|\text{Bad Mgmt} = 0, \text{TW} = 1, \text{Post} = 0])] \\ &- [(E[Y|\text{Bad Mgmt} = 1, \text{TW} = 0, \text{Post} = 1] - E[Y|\text{Bad Mgmt} = 1, \text{TW} = 0, \text{Post} = 0]) \\ &- (E[Y|\text{Bad Mgmt} = 0, \text{TW} = 0, \text{Post} = 1] - E[Y|\text{Bad Mgmt} = 0, \text{TW} = 0, \text{Post} = 0])] \end{aligned}$$

1.4.2 Results

Section 1.3 provided an overview of the use of discretion following the pandemic outbreak, revealing an increase in the use of tender waivers by municipalities, especially in emergency-related purchases. In this section, I examine whether there is a differential effect between municipalities with bad and good fiscal management and investigate how the use of tender waivers affects supplier selection and unit prices of crisis-related goods.

Table 1.3 presents the primary findings using various sample specifications. Column 1 covers all products, while Column 2 and 3 provides estimates using only non-emergency and emergency products, respectively. Additionally, Columns 4-7 represent subcategories of Column 3. Considering the most comprehensive sample, I find that the increase in tender waiver adoption after the pandemic was, on average, 2.9 percentage points higher in municipalities exhibiting worse fiscal management. The effect was particularly pronounced for emergency goods, with a positive variation of 4.8 percentage points. Therefore, the stronger the shock generated by the legislation, the more bad management municipalities increase the use of tender waivers compared to other municipalities. In the appendix, I provide a robustness check where municipalities are matched by population size. Table 1.A.1 summarizes the results of this robustness check, showing coefficients that follow the same direction but are smaller in magnitude.

Regarding the type of emergency goods, Table 1.3 shows that the impact appears to be primarily driven by the procurement of hospital materials, and medicines. Hospital materials comprise a range of items, including catheters, syringes, thermometers, oximeters, and personal protective equipment (PPE).

Table 1.3: Impact of Municipalities' Fiscal Management on Tender Waivers

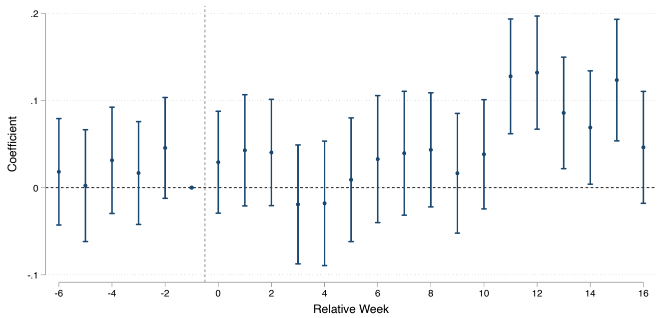
	TW All (1)	TW Non-Emergency (2)	TW Emergency (3)	TW Hosp Material and PPE (4)	TW Medicines (5)	TW Medical Equip (6)	TW Cleaning (7)
Post x Bad Mgmt	0.0296***	0.0243***	0.0480***	0.0423*	0.0427**	0.0301	-0.0116
	(0.0090)	(0.0082)	(0.0158)	(0.0243)	(0.0170)	(0.0387)	(0.0209)
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipalities FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.2424	0.2440	0.2304	0.1438	0.3159	0.2734	0.1337
r2	0.1932	0.1876	0.3741	0.5062	0.4925	0.4786	0.4557
N	282,989	236,581	46,390	15,967	16,647	4,993	9,634

Note: This table reports the effects of municipalities' fiscal management on the likelihood of using tender waivers. Estimates are derived from Equation 2.1. The variable "Bad Mgmt" is an indicator for municipalities below the median of the management index. Column 1 includes all tenders, Column 2 includes only those not classified as emergency-related (Column 3). The remaining Columns (4-7) are subcategories of Column 3. Controls include GDP per capita, population, hospital beds rate, health commitment value, share of municipalities' revenue allocated to health and COVID-19 incidence, measured as new cases per 100,000 inhabitants. Standard errors are clustered at municipality level. *** p<0.01, ** p<0.05, * p<0.1.

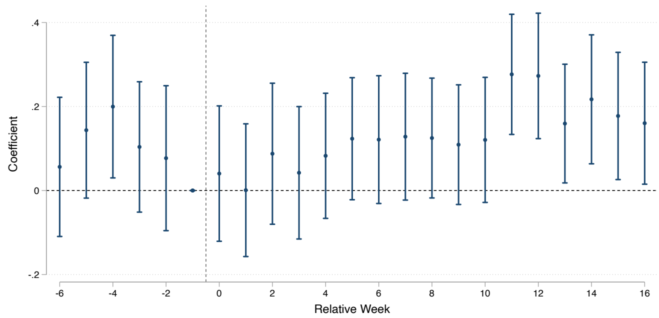
Figure 1.6 illustrates a dynamic estimation starting six weeks prior the first reported COVID-19 case. Time zero is set at week 9 of 2020. The figure reinforces the more pronounced use of tender waivers among bad management municipalities, especially after 10 weeks. In addition, it supports the main identifying assumption of the empirical strategy adopted in this study — both groups appear to have similar trends before the outbreak. Further evidence supporting the parallel trends assumption is provided in Figure 1.A.2, which plots the average share of waived tenders per week for both groups. Despite significant weekly variation, both groups exhibit similar trends before the pandemic, with bad management municipalities consistently using a small proportion of tender waivers, especially for emergency-related purchases.

The noted use of legislative flexibility by municipalities with bad management may generate two non-trivial effects. First, less regulation may encourage greater opportunistic behavior, and municipalities with worse management, if also interpreted as those with weaker state capacity, would be more likely to engage in this type of behavior. Second, the flexibilization could actually lead to more efficient spending practices in these municipalities by lowering the regulatory burden, potentially helping those with the greatest financial constraints.

Figure 1.6: Impact of Municipalities' Fiscal Management on Tender Waivers



A: All Tenders



B: Emergency

Note: This figure reports the weekly effects of municipalities' fiscal management on the likelihood of using tender waivers. Estimates are derived from Equation 2.1. The variable "Bad Mgmt" is an indicator for municipalities below the median of the management index. Panel A reports the coefficients for all products, while Panel B reports the coefficients for goods classified as emergency-related. Controls include GDP per capita, population, hospital beds rate, health commitment value, share of municipalities' revenue allocated to health and COVID-19 incidence, measured as new cases per 100,000 inhabitants. The omitted category is the week preceding the first COVID-19 confirmed case. Standard errors are clustered at the municipality level. The confidence of interval is at the 95% level.

Following, I investigate whether the greater flexibility affected the selection of suppliers when procuring emergency goods. Variables associated with firms' characteristics may signal potential misconduct in the case of hiring punished firms, or favoritism in the case of hiring politically connected and local firms. On the other hand, buying from politically connected and local firms could also reflect the procurement entity's efforts to hire suppliers with known reputation and quality (Fazio, 2022). Table 1.4 presents the results related to supplier selection.

The odd-numbered columns in Table 1.4 show no difference on the allocation of tenders to punished, politically connected, or local firms between municipalities with different degrees of fiscal management. In other words, despite the greater increase in tender waiver utilization among bad management municipalities, there is no indication of misconduct or favoritism when looking to the characteristics of the firms being selected. Actually, descriptive statistics in Table 1.2 indicate a general decline in the proportion of firms sanctioned after winning tenders, with similar trends across both groups.

How about the direct effect of buying under discretion? The even-numbered columns in Table 1.4 provide the estimates using the triple interaction term Post $\times$ Bad Mmgt $\times$ TW. This approach allows for examining whether procurements made via tender waivers in municipalities with bad management yield different outcomes compared to those conducted through competitive methods or in municipalities with better management practices, thereby assessing potential heterogeneity in treatment effects.[1.17]

The analysis reveals no significant difference between tender waivers and competitive tenders on three supplier characteristics: firms sanctioned before or after being awarded a tender, and politically connected firms. Although the coefficient for firms punished ex-post is negative and larger for competitive tenders, suggesting a stronger reduction in this group, the analysis lacks sufficient statistical power to confirm this finding.

Moreover, I find that exercising discretion in procurement during the pandemic decreased the likelihood of hiring local suppliers by 13 percentage points. The greater flexibility provided by tender waivers allowed public buyers to access more suppliers outside municipal boundaries in a context marked by supply chain disruptions. Figure 1.7 plots the coefficients over time, showing that the effects are concentrated two to three months after the outbreak.

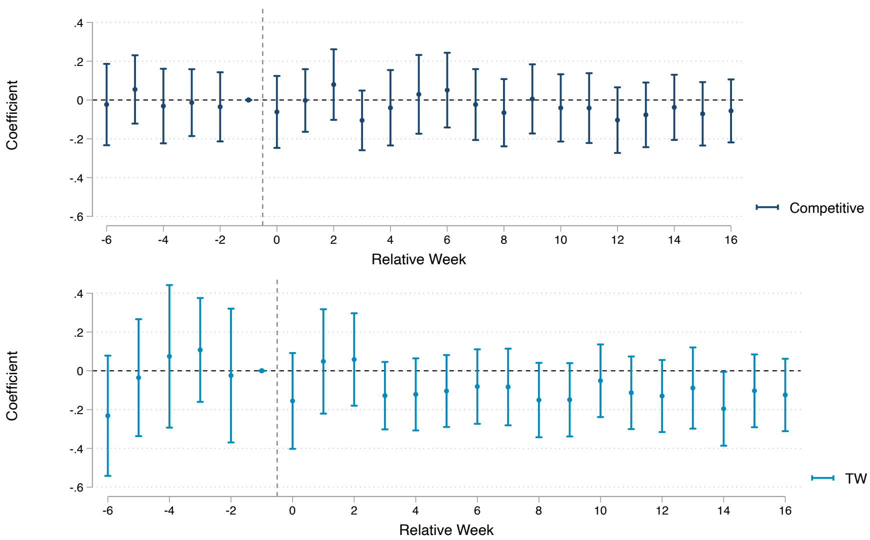
[1.17] See Olden and Møen (2022) for a broad discussion of the triple difference estimator.

Table 1.4: Impact of Management and Tender Waiver on Suppliers' Selection

	Political Connection		Punished (before)		Punished (after)		Local Firm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post x Bad Mgmt	-0.0019	-0.0018	0.0001	-0.0001	-0.0101	-0.0186	-0.0129	0.0125
	(0.0030)	(0.0032)	(0.0007)	(0.0012)	(0.0094)	(0.0132)	(0.0114)	(0.0112)
Post x Bad Mgmt x TW		-0.0033		-0.0006		0.0168		-0.1307***
		(0.0066)		(0.0013)		(0.0191)		(0.0371)
Post x TW		0.0022		-0.0004		-0.0287**		0.0020
		(0.0029)		(0.0011)		(0.0134)		(0.0231)
Bad Mgmt x TW		0.0053		0.0017		0.0233		0.1114***
		(0.0066)		(0.0012)		(0.0166)		(0.0429)
TW		-0.0036		-0.0014*		-0.1958***		0.0226
		(0.0028)		(0.0008)		(0.0099)		(0.0310)
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipalities FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.0123	0.0123	0.0007	0.0007	0.1673	0.1673	0.2569	0.2569
r2	0.1633	0.1634	0.0833	0.0836	0.1520	0.2054	0.2303	0.2320
N	37,795	37,795	37,795	37,795	37,795	37,795	37,795	37,795
Post x Bad Mgmt for TW = 1		-0.0051		-0.0007		-0.0018		-0.1182***
		(0.0063)		(0.0006)		(0.0123)		(0.0364)
Post x Bad Mgmt for TW = 0		-0.0018		-0.0001		-0.0186		0.0125
		(0.0032)		(0.0012)		(0.0132)		(0.0112)

Note: The data used in this table includes only non-competitive purchases made through tender waiver. Controls include GDP per capita, hospital beds rate, health commitment value, share of municipalities' revenue spent on health and the incidence of COVID-19 cases, measured by new cases per 100,000 inhabitants. Standard errors are clustered at municipality level.

Figure 1.7: Impact of Tender Waivers on Local Suppliers



Note: This figure presents the coefficients from a dynamic specification of Equation 2.2. In this sample, only goods classified as emergency-related are considered. Controls include GDP per capita, hospital beds rate, health commitment value, share of municipalities' revenue spent on health and the incidence of COVID-19 cases, measured by monthly new cases per 100,000 inhabitants. Standard errors are clustered at municipality level. The confidence of interval is at the 95%.

Finally, I investigate whether the increased procurement flexibility resulted in higher prices. While higher prices paid by the government could indicate overpricing or corruption, it could also indicate the procurement of higher-quality products (Fazio, 2022). Yet, considering the context of the public health crisis, it is unlikely that any price increase emerges from the acquisition of higher-quality products.

Yemeke et al. (2023) presents qualitative evidence from Zimbabwe, arguing that besides increasing costs and delays, the shortages in COVID-19-related products led to increased risks to the quality of medical products. According to the authors, there was an increased usage of the informal market, with unregistered medical products being sold with less oversight by the regulator and an influx of non-traditional suppliers. In addition, there was a significant increase in medical products procured under special provisions that waived registration requirements. This aligns with certain measures allowed by the Brazilian government during the pandemic, such as the procurement of second-hand equipment and the exemption from preliminary studies for the purchase of essential goods and services.[1.18]

To evaluate the potential impact on prices, I conduct the same regressions using the logarithm of the unit price as the dependent variable, focusing on the product sample detailed in Table 2.A.6. Column 1 of Table 1.5 shows a positive coefficient, indicating a 2% increase in prices for bad management municipalities compared to those with good management. Column 2 also presents a positive coefficient when interacting with the tender waiver indicator, suggesting a 8,6% price increase for non-competitive tenders. However, neither coefficients are statistically significant. For this reason, I can not reject the hypothesis of no differential effect on prices between municipalities with different management status or between competitive and non-competitive tenders.

[1.18] MP 926/2020 allowed the acquisition of second-hand equipment provided that the supplier assumed responsibility for its condition and operation. Additionally, the legislation waived the requirement for a preliminary study, only requiring a basic or simplified term. The preliminary technical study, also referred to as ETP, is a document outlining the contracting needs of the public entity and proposing potential solutions.

Table 1.5: Impact of Management and Tender Waivers on Price

	Ln(Unit Price)	
	(1)	(2)
Post x Bad Mgmt	0.0203	0.0029
	(0.0211)	(0.0214)
Post x Bad Mgmt x TW		0.1137
		(0.1423)
Post x TW		0.0996
		(0.0891)
Bad Mgmt x TW		-0.0486
		(0.1389)
TW		0.0871
		(0.0863)
Year-Quarter FE	Yes	Yes
Municipalities FE	Yes	Yes
Controls	Yes	Yes
Product Category * Year-Quarter	Yes	Yes
Mean	-0.0329	-0.0329
r2	0.9086	0.9090
N	55,360	55,360
Post x Bad Mgmt for TW = 1		0.1165
		(0.1399)
Post x Bad Mgmt for TW = 0		0.0029
		(0.0214)

Note: This table reports the effects of municipalities' fiscal management and the use of tender waivers on the prices of selected products (see Table 2.A.6). The first column presents the estimates from Equation 2.1, while the second column presents the estimates from Equation 2.2. The variable "Bad Mgmt" is an indicator for municipalities below the median of the management index. Controls include GDP per capita, population, hospital beds rate, health commitment value, share of municipalities' revenue allocated to health and COVID-19 incidence, measured as new cases per 100,000 inhabitants. Product category refers to the categories presented in Table 2.A.6. Standard errors are clustered at municipality level. *** p<0.01, ** p<0.05, * p<0.1.

So far, my findings indicate that emergency-related contracts awarded through tender waivers during the pandemic do not present higher unit prices when compared to competitive tenders, nor do they exhibit a higher probability of hiring politically connected or debarred firms. Moreover, these tenders also point to a greater decline in the hiring of local firms when compared to competitive methods.

Another interesting question is whether the increase in discretion translates into better provision of public health services. Despite increasing the risks of inefficiency and corruption, such procurement practices may also facilitate faster responses, thus improving health outcomes. The next section discuss these points.

1.5 The Impact of Fiscal Management and Tender Waivers on Mortality

1.5.1 Empirical Strategy

In this Section, I investigate whether a worse fiscal situation in the baseline contributed to a greater impact of the pandemic on mortality and, if so, whether the flexibility allowed by procurement legislation helped mitigate this effect. The first estimation is similar to the one outlined in Equation 2.1, but now at the municipal level.

$$y_{m,t} = \alpha_m + \lambda_t + \beta(\text{Post Outbreak}_t \times \text{Bad Mgmt}_m) + X_{m,t} + \epsilon_{m,t} \tag{1.3}$$

To assess the impact of tender waivers on mortality I estimate the following equation, similar to Equation 2.2:

$$\begin{aligned} y_{m,t} = \alpha_m + \lambda_t &+ \beta_1(\text{Post Outbreak}_t \times \text{Bad Mgmt}_m) \\ &+ \beta_2(\text{Post Outbreak}_t \times \text{Bad Mgmt}_m \times \text{Share TW}_{m,t}) \\ &+ \beta_3(\text{Post Outbreak}_t \times \text{Share TW}_{m,t}) \\ &+ \beta_4\text{Share TW}_{m,t} + X_{m,t} + \epsilon_{m,t} \end{aligned} \tag{1.4}$$

Where $y_{m,t}$ represents mortality measures in municipality m during month t. The parameters of interest are β in Equation 1.3 and β_2 in Equation 1.4. The term Post Outbreak$_t$ is an indicator variable that takes the value of one from April 2020 across all municipalities. Although the first COVID-19 cases were reported in week 9 of 2020, the municipalities included in this study only began reporting cases in week 13, the last week of March (see Figure 1.A.6). Therefore, to examine health-related outcomes, April is chosen as the starting point for the post-outbreak period.

As before, the variable Bad Mgmt$_m$ denotes an indicator variable. Municipalities falling below the 50th percentile of the management index in 2019 are assigned a value of one, while those above the 50th percentile are assigned a value of zero. The variable Share TW$_{m,t}$ denotes the percentage of waived tenders over all tenders carried out in each municipality and month.[1.19] Additionally, I include year-month fixed effects (λ_t), municipality fixed effects (α_m), and controls at the municipality-level ($X_{m,t}$), such as GDP per capita, total

[1.19] It is worth noting that the distribution of this variable is highly skewed towards zero (Figure 1.A.8), which can pose challenges for its interpretation.

healthcare expenditure, hospital bed rate and the share of revenue allocated to health.

The first mortality outcome is excess mortality. Originally, excess mortality is defined as the exogenous increase in total deaths in a municipality in a given period relative to the historical average for that location. This measure captures the overall effect of the pandemic, including the deaths directly caused by COVID-19 infections as well as the deaths indirectly resulting from the crisis. The raw number of excess mortality offers a good sense of scale, but it is not easily comparable across municipalities with large differences in population size. For this reason, I compute excess mortality per 100,000 inhabitants. Then, $\text{Excess Mortality}_{m,t}$ represents the number of deaths per 100,000 inhabitants in municipality m during month t that exceed the recent historical average measured between 2015 and 2019. Formally:

$$\text{Excess Mortality}_{m,t} = (\text{Mortality Rate}_{m,t} - \overline{\text{Mortality Rate}}_{m,t}) \tag{1.5}$$

Moreover, to improve the interpretation of this measure and comparisons across municipalities, I also present the excess mortality as the percentage difference between the reported and historical mortality rates. This metric is called the P-score and is defined as:

$$\text{P-score}_{m,t} = \frac{(\text{Mortality Rate}_{m,t} - \overline{\text{Mortality Rate}}_{m,t})}{(\overline{\text{Mortality Rate}}_{m,t})} \times 100 \tag{1.6}$$

For instance, if a municipality had a P-score of 10% in a given month in 2020, that would mean the mortality rate per 100,000 inhabitants for that month was 10% higher than the average for that month in the previous five years.

1.5.2 Results

The pandemic of COVID-19 posed significant challenges for fiscal policy in Brazil. At the local level, these challenges were exacerbated as states and municipalities faced limitations in quickly increasing their fiscal resources. Unlike the federal government, subnational entities lack mechanisms to mitigate economic impacts through increased debt. Moreover, changes in tax laws involve deadlines that are longer than the urgent response times needed for a pandemic. Consequently, the ability of states and municipalities to effectively respond depends on their existing fiscal capacity to deliver public services and on federal transfers (Barros Barbosa et al., 2022).

The first hypothesis tested in this section is whether municipalities with lower fiscal capacities prior to the pandemic faced more severe mortality impacts in the beginning of the pandemic. This could be attributed to challenges in expanding hospital capacity, implementing effective restrictions on virus transmission, or procuring personal protective equipment and other essential inputs.

Column 1 of Table 1.6 indicates that municipalities with bad fiscal management in the baseline experienced, on average, an increase of 3 deaths per 100,000 inhabitants when compared to municipalities with good management. Additionally, Column 3 shows that the P-score for municipalities with bad management was, on average, 5.6 percentage points higher. This effect goes in the same direction as the one pointed by Barros Barbosa et al. (2022), which uses an indicator of Payment Capacity (CAPAG) from the National Treasury to assess the impact of fiscal conditions on mortality outcomes.[1.20]

The difference in mortality is further illustrated in Figure 1.8, particularly noticeable from May to June 2020. These findings also corroborate those of Barros Barbosa et al. (2022), highlighting that the impact was mitigated following the implementation of Complementary Law 173/2020 in June 2020. The law was enacted in April 2020 to allocate federal resources to municipalities, but the actual distribution of funds only started in June 2020.

In explaining why municipalities with better fiscal capacity exhibit lower mortality rates, Barros Barbosa et al. (2022) discuss that the main mechanism was the expansion and improvement of hospital care capacity, rather than local social distancing measures. The authors show that these municipalities have expanded the number of hospital beds and increased the presence of doctors and nurses compared to those with worse fiscal conditions. This suggests that municipalities with greater fiscal capacity were better equipped to handle the pandemic by enhancing their healthcare infrastructure. In contrast, municipalities with weaker fiscal conditions faced challenges in scaling up their healthcare services, potentially exacerbating mortality rates during that period.

Given these findings, the second hypothesis tested in this section is if the flexibility allowed by procurement legislation helped attenuate the negative effects caused by the pandemic, especially for financially constrained municipalities. For instance, the increased use of tender waivers may have enabled municipalities to quickly procure essential medical supplies and services, thereby mitigating some of the negative impacts of fiscal constraints on public health

[1.20] The CAPAG is composed of three indicators that measure municipal debt, current savings, and liquidity. Each municipality is assigned a score ranging from A to D, with those receiving an A having the best fiscal conditions.

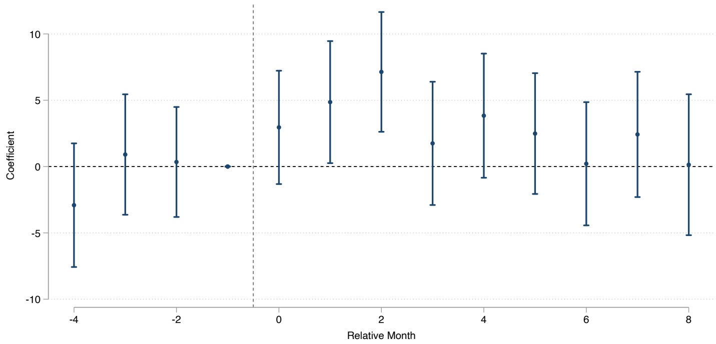
outcomes. Columns 2 and 4 of Table 1.6 presents the results of the estimation, showing that the coefficients for the triple interaction are not statistically different from zero. This indicates that an increase in the use of tender waivers did not attenuate the difference in mortality rates between municipalities with good and bad fiscal management.

Table 1.6: Impact of Management and Tender Waivers on Mortality

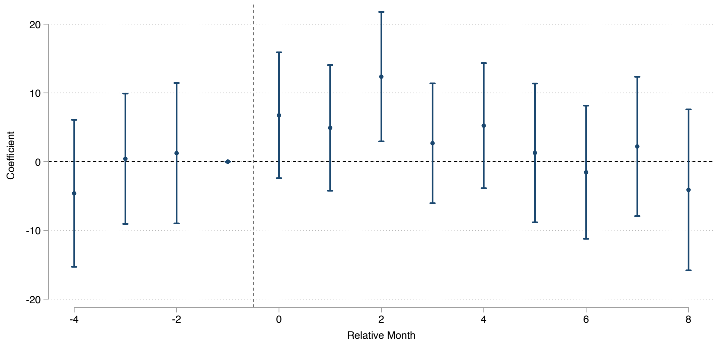
	Excess Mortality		P-Score Mortality	
	(1)	(2)	(3)	(4)
Post x Bad Mgmt	3.0569***	2.9631***	5.6107***	7.0521***
	(0.7334)	(0.9851)	(1.5719)	(2.1177)
Post x Bad Mgmt x Share TW		0.0095		-0.0480
		(0.0249)		(0.0580)
Post x Share TW		0.0090		0.0276
		(0.0187)		(0.0409)
Bad Mgmt x Share TW		-0.0187		0.0034
		(0.0195)		(0.0505)
Share TW		0.0181		0.0400
		(0.0143)		(0.0324)
Year-Month FE	Yes	Yes	Yes	Yes
Municipalities FE	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
Mean	3.8753	3.8753	12.9547	12.9547
r2	0.0699	0.0700	0.0735	0.0737
N	44,450	44,450	44,438	44,438

Note: This table reports the effects of municipalities' fiscal management on mortality, and the effects of tender waivers on mortality. Estimates are derived from Equation 1.3 and 1.4. The variable "Bad Mgmt" is an indicator for municipalities below the median of the management index. The variable "Share TW" denotes the percentage of waived tenders over all tenders in each municipality and month. Controls include GDP per capita, population, hospital beds rate, health commitment value, and share of municipalities' revenue allocated to health. Standard errors are clustered at municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Figure 1.8: Impact of Fiscal Management on Mortality



A: Excess Mortality



B: P-score Mortality

Note: This figure presents the coefficients from a dynamic specification of Equation 1.3. Controls include GDP per capita, hospital beds rate, health commitment value and the share of municipalities' revenue spent on health. Standard errors are clustered at municipality level. The confidence of interval is at the 95%.

1.6 Conclusion

During emergencies, there is a tough choice between making procurement faster, thereby amplifying the susceptibility to corrupt practices and inefficiencies, and the urgent need to save lives. Finding the right balance is a challenge.

To face the pandemics needs, the Brazilian government increased flexibility in procurement legislation, extending the possibility for direct contracting and adding a new scenario for exceptions to those already outlined in standard legislation. One important finding is that despite the increased flexibility, there were no significant leaks in policy. In other words, the huge increase in discretion was mainly driven by emergency-related goods, as originally set by the legislation.

The second finding is that both good and bad management municipalities increased their use of tender waivers, but the later experienced a more significant rise, especially in emergency-related purchases. On one hand, this could raise concerns about opportunistic behavior. On the other hand, less regulation could actually lead to more efficient spending practices, helping municipalities with worse management to alleviate financial constrains.

As shown earlier, the results of this study suggest that more discretion did not lead to worse supplier selection. The percentage of politically connected suppliers remained stable, while the share of firms listed as punished after being selected decreased. Overall, neither bad management municipalities presented higher likelihood of hiring those firms, nor did the purchases made through tender waivers.

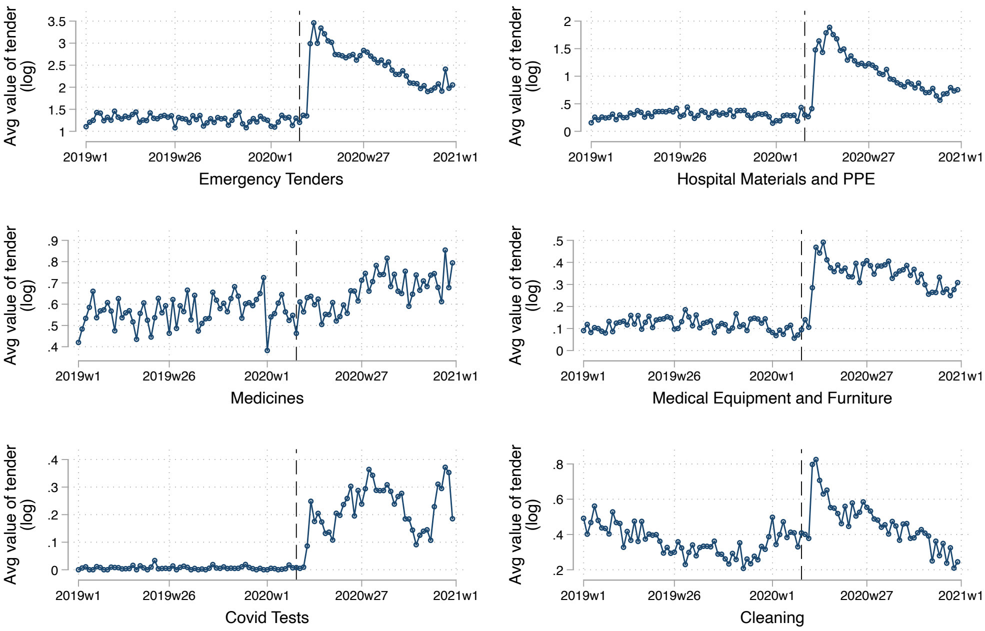
Regarding economic efficiency, the data indicate that prices for most of the selected products did not change significantly, except for certain items like COVID-19 kit medicines and disposable masks. In addition, the results show that buying through tender waivers did not result in higher prices for a set of emergency goods. Although the estimated coefficient is relatively high and suggests a possible increase, it lacks statistical significance.

A potential limitation of the price analysis is the low percentage of tender waivers in the reduced sample (8.8%), compared to 23% for all emergency-related goods. This results in a low number of observations of non-competitive tenders, which can make it difficult effectively applying a triple interaction model. Future research could apply advanced machine learning techniques to enhance product classification based on item descriptions, which could potentially increase the number and the quality of observations.

Finally, this paper highlights the critical importance of financial stability in effectively responding to health crises. Municipalities with bad management prior to the pandemic experienced higher mortality early on, likely due to difficulties in expanding their healthcare services. My analysis reveals that using tender waivers increased the likelihood of selecting suppliers from outside municipal boundaries, which may have contributed to alleviate local supply shortages. However, the increased use of discretion appeared insufficient to close the gap in excess mortality between good and bad managed municipalities. An interesting avenue for future research is to understand the determinants of selecting non-local suppliers and the geographical factors influencing supply chain dynamics during emergencies.

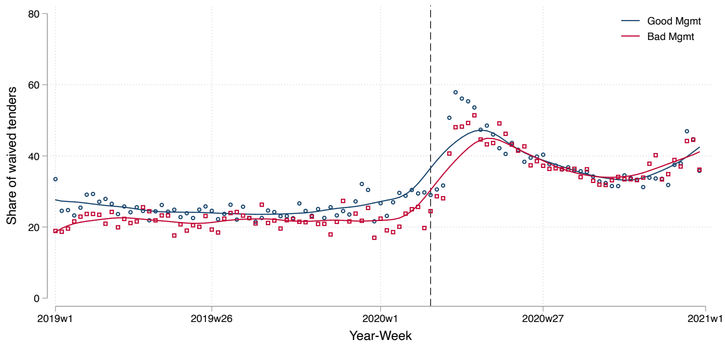
1.A Appendix

Figure 1.A.1: Average Value of Tenders Classified as Emergency-Related (2019-2020)

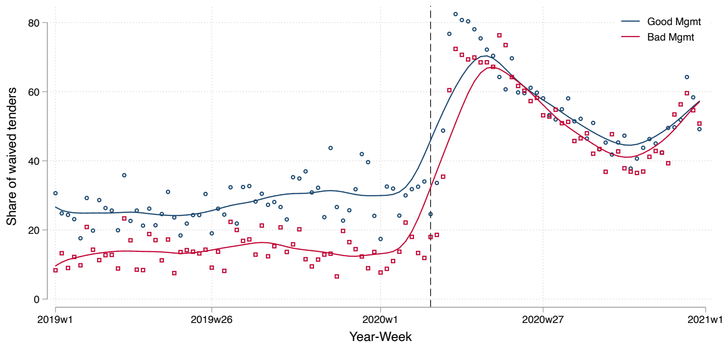


Note: This figure shows the average value of tenders classified as emergency-related per week from January 2019 to December 2020, considering all procurement methods. The dashed line marks week 9 of 2020, which corresponds to the week of the first recorded COVID-19 case in Brazil. The classification of these tenders was achieved through text analysis of the tender descriptions (Refer to Section 1.3.1 for further details). Quadrants 2 through 6 represent different categories of goods within the emergency definition. The first quadrant aggregates the counts from quadrants 2 to 6.

Figure 1.A.2: Use of Tender Waivers by Municipalities' Management (2019-2020)



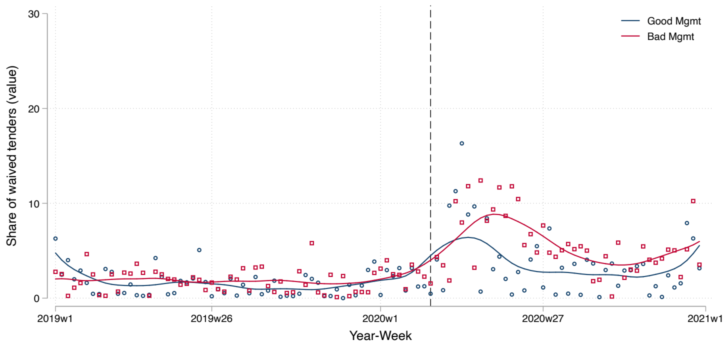
A: All Tenders



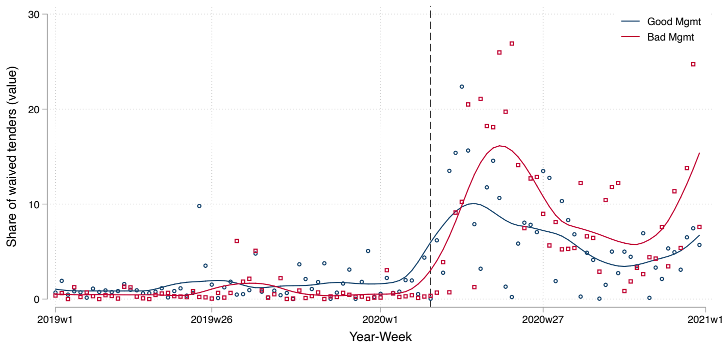
B: Emergency Tenders

Note: This figure shows the average percentage of purchases made through tender waivers from the first week of 2019 until the last week of 2020, by management status. Municipalities are classified as "Bad Mgmt" if they fall below the median of the Fiscal Management Index (IFGF), and "Good Mgmt" if they are above the median. The distribution of the share of tender waivers is smoothed using locally weighted regression (lowess) with a bandwidth equal to 20% of the data. The dashed line marks week 9 of 2020, which corresponds to the week of the first recorded COVID-19 case in Brazil.

Figure 1.A.3: Proportion of Funds Allocated to Tender Waivers by Municipal Management (2019-2020)



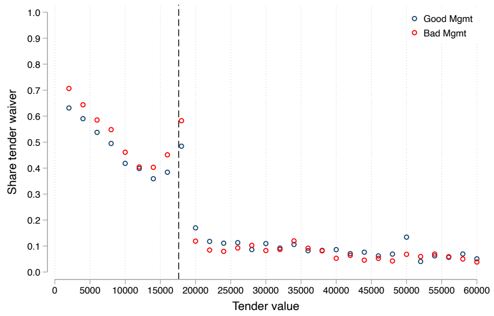
A: All Tenders



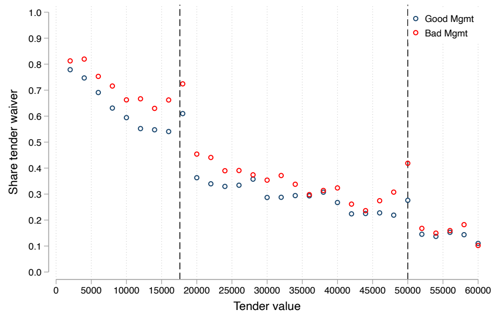
B: Emergency Tenders

Note: This figure shows the average percentage of funds allocated to tender waivers relative to the total expenditure on tenders - considering only goods - from the first week of 2019 to the last week of 2020, by management status. Municipalities are classified as “Bad Mgmt” if they fall below the median of the Fiscal Management Index (IFGF), and “Good Mgmt” if they are above the median. The distribution of the share of tender waivers is smoothed using locally weighted regression (lowess) with a bandwidth equal to 20% of the data. The dashed line marks week 9 of 2020, which corresponds to the week of the first recorded COVID-19 case in Brazil.

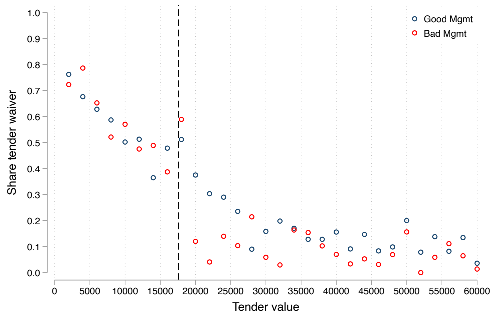
Figure 1.A.4: Share of Waived Tenders by Municipalities' Management



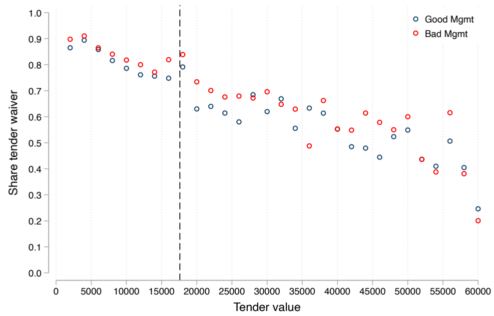
A: All Tenders, before COVID-19



B: All Tenders, after COVID-19



C: Emergency Tenders, before COVID-19

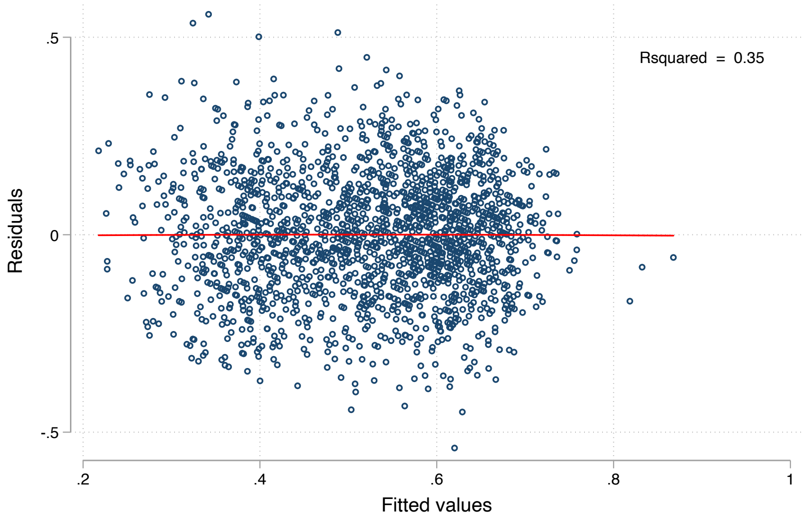


D: Emergency Tenders, after COVID-19

Note: This figure presents the proportion of purchases made through tender waivers within each R$2000 bin. Panels A and B display the share of tender waivers for all tenders, while Panels C and D restrict the data to emergency tenders. In both cases, only goods (products) are included. All figures indicate the threshold of R$ 17,600 established by Law No. 8,666/1993, which was waived for emergency-related purchases by Law 13,979/2020. Additionally, Figure B displays the threshold of R$ 50,000 set by MP 961 for all goods and services other than emergency-related ones. The red dots represent municipalities with bad fiscal management (below the median of the IFGF index), while the blue dots indicate municipalities with good fiscal management (above the median of the IFGF index).

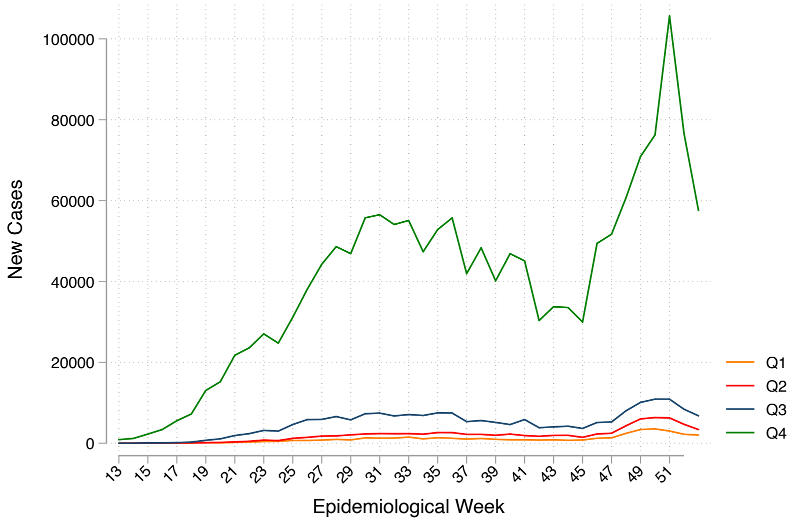
.

Figure 1.A.5: Residuals of IFGF Regression on Municipal Characteristics



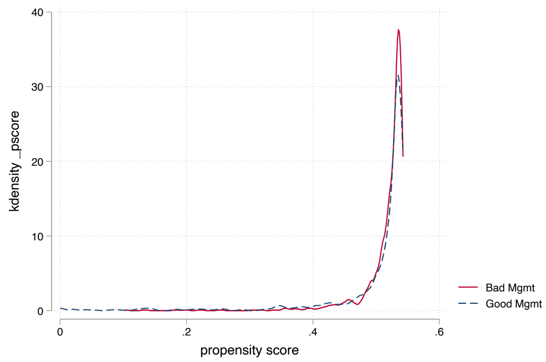
Note: This plot shows the residuals from a regression analysis of the fiscal management index (IFGF) against several predictor variables: GDP per capita, population size, health commitment value, share of municipalities' revenue allocated to health, latitude, longitude, municipal human development index, and the percentage of high-skilled employees in public administration.

Figure 1.A.6: Weekly New COVID-19 Cases by Population Quartiles

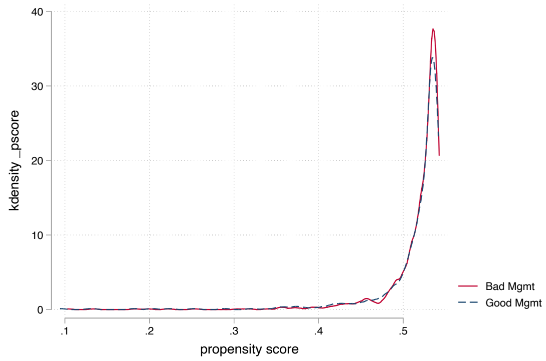


Note: This Figure presents the number of new cases of COVID-19 per week, by population quartiles, across 2020. Data are restricted to the states of Ceará, Minas Gerais, Paraná, Pernambuco, and Rio Grande do Sul. While the first COVID-19 case in Brazil was reported in week 9 of 2020, the first case within the sample of states examined in this paper occurred in week 13. The average population of each quartile in this sample is as follows: 3210, 6897, 15165 and 93158 inhabitants, respectively.

Figure 1.A.7: Propensity Score Matching



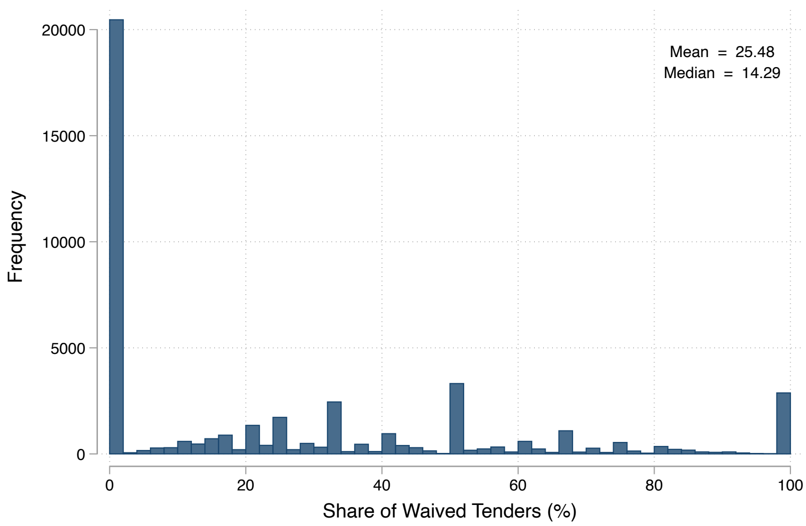
A: Before Matching



B: After Matching

Note: This figure presents the propensity score for municipalities classified as Bad Management and Good Management before and after matching by population size.

Figure 1.A.8: Distribution of the Share of Waived Tenders



Note: This figure displays the distribution of the share of waived tenders as a proportion of total tenders by municipality and year-month.

Table 1.A.1: Impact of Municipalities' Fiscal Management on Tender Waivers - Inverse Probability Weighting

	TW All (1)	TW Non-Emergency (2)	TW Emergency (3)	TW Hosp Material and PPE (4)	TW Medicines (5)	TW Medical Equip (6)	TW Cleaning (7)
Post x Bad Mgmt	0.0255***	0.0218***	0.0379**	0.0273	0.0424**	0.0470	-0.0176
	(0.0085)	(0.0079)	(0.0166)	(0.0261)	(0.0181)	(0.0376)	(0.0214)
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipalities FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.2424	0.2440	0.2304	0.1438	0.3159	0.2734	0.1337
r2	0.2009	0.1956	0.3811	0.5105	0.5004	0.4912	0.4668
N	282,989	236,581	46,390	15,967	16,647	4,993	9,634

Note: This table reports the effects of municipalities' fiscal management on the likelihood of using tender waivers. The variable "Bad Mgmt" is an indicator for municipalities below the median of the management index. Column 1 includes all tenders, Column 2 includes only those not classified as Emergency-related (Column 3). The remaining Columns (4-7) are subcategories of Column 3. The regressions are weighted by inverse probability weighting (IPW) to account for population size variations. Controls include GDP per capita, population, hospital beds rate, health commitment value, share of municipalities' revenue allocated to health and COVID-19 incidence, measured as new cases per 100,000 inhabitants. Standard errors are clustered at municipality level. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.A.2: Impact of Management and Tender Waiver on Suppliers' Selection - Inverse Probability Weighting

	Political Connection		Punished (before)		Punished (after)		Local Firm	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Post x Bad Mgmt	-0.0000	-0.0015	-0.0001	-0.0005	-0.0144	-0.0239*	-0.0087	0.0151
	(0.0031)	(0.0032)	(0.0007)	(0.0011)	(0.0107)	(0.0131)	(0.0128)	(0.0127)
Post x Bad Mgmt x TW		-0.0007		-0.0004		0.0175		-0.1232***
		(0.0072)		(0.0013)		(0.0194)		(0.0362)
Post x TW		0.0028		-0.0007		-0.0349***		-0.0119
		(0.0032)		(0.0011)		(0.0111)		(0.0214)
Bad Mgmt x TW		0.0064		0.0018		0.0177		0.1011***
		(0.0065)		(0.0011)		(0.0175)		(0.0391)
TW		-0.0039		-0.0014**		-0.1907***		0.0410
		(0.0031)		(0.0007)		(0.0089)		(0.0250)
Year-Week FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipalities FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.0123	0.0123	0.0007	0.0007	0.1673	0.1673	0.2569	0.2569
r2	0.1678	0.1679	0.0847	0.0852	0.1538	0.2065	0.2314	0.2337
N	37,795	37,795	37,795	37,795	37,795	37,795	37,795	37,795
Post x Bad Mgmt for TW = 1		-0.0022		-0.0008		-0.0064		-0.1081***
		(0.0065)		(0.0006)		(0.0137)		(0.0351)
Post x Bad Mgmt for TW = 0		-0.0015		-0.0005		-0.0239*		0.0151
		(0.0032)		(0.0011)		(0.0131)		(0.0127)

Note: The data used in this table includes only non-competitive purchases made through tender waiver. The regressions are weighted by inverse probability weighting (IPW) to account for population size variations. Controls include GDP per capita, hospital beds rate, health commitment value, share of municipalities' revenue spent on health and the incidence of COVID-19 cases, measured by new cases per 100,000 inhabitants. Standard errors are clustered at municipality level.

Table 1.A.3: List of Selected Materials

Product	**Unit**	**Expenditure (millions R$)**
Azithromycin (500 mg)	Pill	87.59
Carvedilol (12,5 mg)	Pill	57.66
Catheter	Unit	78
Dexamethasone (4mg)	Pill	10.68
Detergent	500 ml	6.13
Diazepam	Pill	26.94
Dipyrone (500mg)	Pill	45.87
Disposable mask	Unit	17.73
Disposable needle	Unit	10.44
Disposable syringes	Unit	45.18
Ethyl alcohol	1 liter	21.34
Furosemide (40mg)	Pill	20.13
Gloves for non-surgical procedures	Box 100 units	247.89
Gloves for surgical procedures	Pair	9.5
Haloperidol (5mg)	Pill	35.15
Hydroxychloroquine (400mg)	Pill	7.89
Ibuprofen (600mg)	Pill	138.8
Ivermectin (6mg)	Pill	20.61
Levothyroxine (25mcg)	Pill	22.06
Masks (n95, pff2)	Unit	17.95
Paracetamol (500mg)	Pill	39.44
Plastic bags (100 liter)	Package 100 units	16.08
Prednisone (20mg)	Pill	10.36
Risperidone (2mg)	Pill	5.83
Saline solution (0,9%)	500 ml	21.95
Simvastatin (20mg)	Pill	35.56
Sulfate (40mg)	Pill	13.92
Surgical mask	Box 100 units	33.29
Tracheal tube	Unit	10.12

Note: The table presents a list of homogeneous products selected through text analysis of item descriptions. Each product in the table specifies the most commonly used unit. For further explanation, refer to Section 1.3.1. Expenditure represents the cumulative value of all items containing the product's name in the description, from purchases made between 2019 and 2020. Prices below the 1th percentile and above the 99th percentile were excluded from the analysis.

Table 1.A.4: Procurement Methods

Purchasing method	Competitive	Characteristics	Contract size
Reverse auction (*Pregão*)	Yes	Reverse auction, open to any interested firm. Online or in-person. Off-the-shelf goods. Multiples bids per participant.	Any value
Waiver (direct contracting)	No	Small purchases.	Up to 17,600 BRL
Invitation to tender (*Convite*)	Yes	Participants are invited. Minimum of 3 bidders. Uninvited firms are allowed to participate. One bid per participant.	Up to 176,000 BRL
Competitive bidding (*Concorrência*)	Yes	Open to any interested bidder. One bid per participant.	Any value
Submission of prices (*Tomada de preços*)	Yes	Bidder must be previously registered. One bid per participant.	Up to 1,430,000 BRL
Direct contracting	No	There is only one supplier.	-
Contest	Yes	Artistic, scientific or technical work.	-

Notes: Contract size refers to the purchases of products and services other than construction (see Federal Decree 9,412, of June 18, 2018, for more details). Thresholds for construction are different (33,000 BRL). The maximum contract size for direct contracting changed in 2018 from 8,000 BRL to 17,600 BRL. Table from Dahis et al. (2022).

2
Set-Aside Policy for SMEs and Competition in Brazilian Procurement

Nathalia Sales
PUC-Rio

Thiago Scot
World Bank

Abstract

Understanding the impact of policies targeting SMEs in public procurement is essential for fostering competition and improving the efficient allocation of public resources. In this paper, we apply a regression discontinuity design to evaluate the effect of a set-aside policy for SMEs in Brazilian auctions. We find that the use of set-asides decreases with contract size and that purchasing entities exercise substantial discretion in applying the policy. However, there is no evidence of manipulation in contract values to exploit policy thresholds. We observe that set-asides reduces competition, with a decline in participation from larger bidders and no statistically significant increase in SME entry. In sectors with traditionally low SME participation, the policy leads to a significant increase in small bidders, outweighing the exit of larger firms. We further investigate the impact of reduced competition on prices, discuss policy compliance, and assess the validity of the instrument within a restricted product sample.

2.1
Introduction

Worldwide, governments use their acquisition of goods and purchases (i.e. public procurement) as a tool to advance certain policy goals. While some argue that procurement practices should aim exclusively at assuring value-for-money, policy-makers often enact rules that leverage purchases to encourage local employment or encourage the development of specific technologies, even if at the expense of price and/or quality considerations. These goals are recognized by international development agencies, where lead voices have argued that "Government purchasing should be more than just a transactional business process that helps increase the efficiency of spending and free up fiscal space.

Public procurement must be a strategic tool for socioeconomic change that uses government purchasing decisions and technology more strategically". [2.1]

One specific goal often pursued by government is the awarding of procurement contracts to small- and medium-enterprises (SMEs). World Bank Group (2017) discusses several possible rationales behind those policies, including that i) encouraging SME participation might benefit procurement by increasing competition and fostering innovation; ii) SMEs face particular barriers to participating in procurement, such as lack of information or finance, so encouraging them is "leveling the playing field"; and iii) targeting SMEs might achieve other policy goals such as increasing participation of women-owned firms. The policy tools available to foster SME participation range from broad rules that foster transparency and competition to specific policies that provide training to SMEs or simply establish quotas or set-asides for their participation.

In this paper we study one such policy in Brazil, which allows public procurement entities to set-aside part of their purchases exclusively for SMEs. Similar policies exist in several countries, including Indonesia, Colombia, India and South Korea (World Bank Group, 2017), where contracts below a certain threshold or specific goods/services can be set-aside for SMEs.

The policy we study in Brazil allows public procurement entities to reserve certain lots in competitive auctions for exclusive participation by SMEs, conditional on estimated value being below R$ 80,000 (approximately USD 35,000[2.2]). Even though the law initially establishes that lots below that price should be reserved for SMEs, it also allows for discretion by procuring entities: they can forego the use of the set-aside in cases where it is not advantageous for the public administration or if there are less than three available SME suppliers in the region.

In the first part of the paper, we document the use of the set-aside policy. In the period we study, the share of eligible lots (that is, below R$ 80,000) set-aside to SMEs increased from about 20% in 2013 to almost 60% by 2019. We document that this increase was driven both by an extensive margin, where some entities that never used the policy start using it over time, but also on the intensive margin, meaning that most purchasing units start to use the set-aside policy more often as time goes by. We also show that, consistent with the broad exceptions allowed under the law, purchasing entities exercise discretion and vary substantially in their use of the set-aside for lots under the eligibility threshold.

[2.1] https://blogs.worldbank.org/voices/hidden-1-trillion-halting-waste-public-procurement.

[2.2] Using 2019 PPP conversion rate of 1 USD = BRL 2.28.

In the second part of the paper, we tackle the question of whether the use of the set-aside policy causally impacts the level of competition observed. We exploit the fact that the policy only allows the use of set-asides for lot values below R$ 80,000 and implement a regression discontinuity-design (RDD), comparing lots arbitrarily close but on either side of the threshold to evaluate how outcomes of interest change around that value. We test and reject the hypothesis that procurement officers manipulate the value of contracts to stay on either side of the threshold.

Our key finding is that the use of set-asides for SMEs reduce competition in the auctions. We document that, precisely around the eligibility threshold, the average number of bidders in an auction increases from about 8.2 - when set-asides are allowed - to 8.5 immediately above. Combined with the fact that only 15% of lots are reserved to SMEs below the threshold, we estimate that the use of set-asides decreases the number of bidders by about 2 - a large effect given the average number of 8.2 bidders below the threshold. We also assess whether the set-aside affects the number of lots with non single-bidders, another indicator of competition, and we cannot reject the null hypothesis of zero effects.

Overall, we find that the decline in competition is driven by the exit of large bidders, while there is no evidence of SME entry. On the other hand, the set-aside policy increases the likelihood of an SME winning. Our estimates show that the probability of a lot being awarded to small firms (EPPs henceforth) increases by 29 percentage points, and to micro-enterprises (MEs henceforth) by 30 percentage points. Interestingly, this increase is not matched by a rise in local supplier winners, suggesting that the policy may not fully achieve its intended goal of promoting local development (Brasil, 2023).

Regarding the null effect on SME entry, we argue that this may be attributed to the already high level of SME participation in certain sectors. Our analysis indicates that in sectors with low SME participation in unrestricted lots, the policy increases SME entry by 2.8 bidders, with MEs accounting for 1.7 of those. Additionally, the probability of winning significantly increases for MEs and EPPs. However, in sectors with high SME participation in unrestricted lots, the number of SMEs decreases by 3.8 bidders, with EPPs experiencing a decline of 3 bidders.

Although the exit of SME bidders may seem counterintuitive, a potential explanation is that in sectors where they are predominant, there might already be an informal market division, with each firm having its specific niches. The restriction imposed by the policy can strengthen this division, resulting in fewer firms competing directly in each lot. For this group, we also observe a decrease of 8 percentage points in the probability of having a new SME winner —that is, an SME winning a lot for the first time within our study period. Furthermore, our estimates point to a tendency toward contracting with firms that have a history of winning, although this effect is not statistically significant.

Finally, we select a sample of goods commonly used in procurement literature to investigate whether the observed decrease in competition translates into higher prices. Despite the reduction in competition, we can not reject the null hypothesis of zero effects on prices. This finding suggests that, for this group of items, other factors may be mitigating the impact of restricted competition on pricing.

This paper contributes to two branches of the literature. First, it contributes to the literature on incentives in procurement, in which several papers study ex-ante and ex-post procurement outcomes when there is manipulation of awarding mechanism or flexibility of the contract structure - see Carril et al. (2022) for a comprehensive discussion. Second, our paper contributes more specifically to the literature that evaluates policies aimed at promoting (or restricting) bidders' participation in procurement. In general, this literature analyzes the effects of these policies on competition and prices. As we discuss in the following paragraphs, it is not clear whether promoting participation for some group, will deliver more competition or lower prices.

In a study looking at preference margins applied to bids from small firms for road construction in California, Marion (2007) finds that procurement costs are 3.8 percent higher on auctions using preferences. This difference is explained by reduced participation of large firms and the changed composition of auction winners in preference auctions, including the shift of some contracts to higher-cost firms. Alternatively, Krasnokutskaya and Seim (2011) analyze the same bid preference program and find that the upward pressure on procurement costs is somewhat mitigated by the aggressive bidding from larger firms to win against favored smaller firms and by the overall increase in competition and participation of firms[2.3]. As a result, the aggregate cost of the program is only 1.4 percent higher than the aggregate cost under no preferential treatment. Moreover, the program induces substantial changes in small and large firms' participation and probabilities of winning, resulting in a redistribution of profits from large to small firms.

[2.3]Under preference margins or bid subsidies, non-favored firms might behave more aggressively, bidding closer to their costs to compete with favored firms. Hubbard and Paarsch (2009) refers to this as the competitive effect. In our setting, this effect is ruled out once larger firms are prohibited from competing. Consequently, the impact on costs depends on the preference effect— whether firms receiving preferential treatment inflate their bids and still win the auction—and the participation effect, which reflects firms' entry decisions based on their costs and the adopted preferential policy.

Studies examining the Brazilian context yield mixed results. In an early investigation covering the period between 2007 and 2010, Szerman (2012) provides a starting point to study the set-aside policy targeting SMEs in Brazil. He finds that restricting participation of large firms has little effect on prices, while it increases the incentives of small firms to participate, which more than compensates the reduction in the number of larger bidders. More recently, Fiuza et al. (2023) analyze the same set-aside policy and find that, from 2016 to 2018, the policy increased both the participation and success rate of SMEs. However, the greater number of small firms did not compensate the exit of larger bidders, leading to a decrease of approximately two bidders per lot. Additionally, the study reports a rise in price levels and suggests potential collusion among bidders below the threshold.

Another interesting reference is Reis and Cabral (2015), which examines service contracts from four Brazilian federal government agencies between 2003 and 2012. They evaluate the broader impact of the General Micro and Small Enterprise Law, which includes not only the reservation of lots below R$ 80,000 but also a quota up to 25% of divisible goods in contracts above this threshold and a preference mechanism allowing SMEs to match the lowest bid if their price is up to 5% higher than those of non-favored firms. The authors find that these preference programs increase SME participation in public auctions without affecting prices. However, they also observe that SMEs are more likely to have their contracts terminated due to poor performance.

Although we look to the same policy as Szerman (2012) and Fiuza et al. (2023), our paper differs in several ways. First, in comparison to Szerman (2012), we cover a more recent period, from 2013 to 2019, which reduces the impact of the ambiguous criteria interpretation that was in place before the legislative changes in 2014.[2.4] Second, our dataset includes all available products, whereas Szerman (2012) focuses on a restricted sample of seven commonly purchased items and Fiuza et al. (2023) primarily examines pharmaceutical drugs.[2.5] We only restrict our analysis to commonly purchased products when evaluating the effect on prices. This broader dataset allows us to gain a more comprehensive understanding of the policy's impact on competition. Additionally, we examine the differential effects of the policy on MEs and EPPs, which vary significantly in size and revenue, as discussed later.

The remainder of the paper is organized as follows. Section 3.2 provides

[2.4]Before 2014, there was uncertainty about whether the set-aside applied to individual lots or to groups of lots (tenders) below the eligibility threshold. Section 2.2.2 provides details on this change.

[2.5]Their analysis focuses on frequently purchased essential inputs as well as all pharmaceutical drugs.

a summary of the institutional context, including the rules for a firm to be classified as a SME and the main features of the preferential treatment policy. In Section 3.3 and Section 3.4, respectively, we present some descriptive statistics and preliminary evidence on set-aside policy. Section 3.5 presents our empirical strategy. Finally, Sections 2.6 and 2.7, discuss the main results and final remarks.

2.2 Institutional Context

During the period studied in this paper, the general rules for public procurement in Brazil were regulated by Federal Law 8.666/93, which has experienced numerous amendments and changes over the years.[2.6] While the rules apply to all public sector levels (federal, state and municipal), we restrict our analysis to federal public procurement, which represents 5% of the GDP or 50% of total procurement amount (Thorstensen and Giesteira, 2021b). For the purposes of this study, we focus our analysis on auctions and framework agreements.

The specific set-aside policy for SMEs was introduced by Complementary Law 123 in 2006, also known as the Statute of Micro and Small Enterprises. Beyond public procurement, the law also established a set of rules to promote the development of micro and small firms in Brazil, including simplified taxation[2.7], access to credit facilities, and incentives for formalization. In subsections 2.1 and 2.2, we give more details of SMEs definition and how set aside policy works, respectively.

2.2.1 SME Definition

The Law No. 123/2006 sets out clear guidelines for classifying companies as either a micro-enterprise or a small enterprise based on their annual gross revenue and number of employees. To qualify as a micro-enterprise (ME), the company must have annual gross revenue of up to R$ 360,000 and less than 9 (service) or 19 employees (manufacturing); while a small enterprise (EPP) must have annual gross revenue between R$ 360,000 and R$ 4.8 million, and between 10 and 49 (service) or between 20 and 99 (manufacturing) employees.

Another category that can benefit from preferential treatment in tenders is the individual microentrepreneur (MEI). According to the Federal Law No. 128/2008, the annual gross revenue must be under R$ 81,000 to qualify as a

[2.6]In 2021 it was replaced by Law 14.133/21.

[2.7]The law created the Simples Nacional, a simplified tax regime that allows micro and small businesses to pay their taxes in a simplified way, reducing bureaucracy and tax costs.

MEI. In addition, it is important to note that these criteria have changed over time. The last update was made in 2018, increasing the possibility for more firms to participate in the policy. Table 2.1 summarizes the revenue criteria and the changes of 2018.

Table 2.1: SME criteria and 2018 changes

SIZE	Before 01/01/2018	After 01/01/2018
MEI	R$ 60,000	R$ 81,000
MICRO	R$ 360,000	R$ 360,000
SMALL	R$ 3,600,000	R$ 4,800,000

Among the benefits of SME's law is that these entities can opt for the Simplified Tax Regime, called SIMPLES. This regime allows firms to file several tax and contribution obligations – including corporate income tax, social security contributions, and local sales tax – in one tax form, where the tax liability ranges from 4- 30% of gross revenue, depending on firm sector and size (World Bank, 2021). The information on whether firms adhere to the SIMPLES is available in the matched employer-employee database (RAIS) and serves as an initial indicator for identifying SMEs. However, some SMEs may not opt for this regime, which could lead to an underestimation of their presence. To address this, we classify firms as SMEs using registry information from *Receita Federal*. This classification is updated annually based on the gross revenue criteria from the previous year, as detailed in Table 2.1. This is the same requirement to be able to apply for preferential treatment according to the Law No. 123/2006.[2.8]

2.2.2 Set-Aside Policy

The set-aside policy introduced in 2006 states that purchasing entities "may carry out a bidding process exclusively for the participation of micro and small businesses in contracting whose value is up to R$ 80,000". Furthermore, it also established that "in auctions for the acquisition of goods and services of a **divisible** nature, a quota of up to 25% (twenty-five percent) of the object shall be exclusive for the contracting of micro and small businesses".

In 2014, the law was amended in two relevant ways. First, it replaced the expression "**may** carry out a bidding process exclusively for the participation

[2.8] By merging data from RAIS and Receita Federal, we find that 79% of SME suppliers used the SIMPLES regime from 2013 to 2019.

of SMEs" with "**must** carry out (...)". That is stronger language, implying buying entities should not use discretion when deciding whether to make tenders exclusive to SMEs or not. In practice, however, the law still allowed for exceptions in cases where the set-aside to SMEs is not beneficial for the public administration. This includes cases where there are fewer than three available SME suppliers in the region. Second, the 2014 amendment also clarified that the policy applies to *lots* valued at less than R$ 80,000.[2.9] Before that, there was uncertainty of whether the policy applied to *lots* or to the group of lots being sold together (what we refer to as *tender*) (Szerman, 2012). The change in law, if anything, should make the set-aside policy weakly more likely to be used: if two lots priced at R$ 45,000 each were sold together, before 2014 there was uncertainty if that *tender* priced at R$ 90,000 could be set-aside for SMEs, while after that was clearly allowed. Those changes are particularly important to our study, since we focus only on the exclusivity created by the threshold of R$ 80,000 and our analysis period goes from 2013 to 2019.

2.3 Data

The main data source used in this study is public information on the universe of federal purchases available through the Transparency Portal.[2.10] Our sample covers the 2013 - 2019 period. The data is comprehensive and includes information on each tender, purchasing entity, method of purchase, suppliers, prices, dates, and quantities. We complement this data with information from Compras Dados, an API from the federal government allowing users to extract additional information related to public purchases.[2.11] We extract from Compras Dados an indicator of whether each lot was set-aside for SMEs, as well as the unit prices of items.

We clarify here some of the naming conventions we use throughout the paper. In the purchasing modalities we consider in our sample (auctions and framework agreements), competition happens at the **lot level**. A lot is a group of items - of the same product or service - sold together. This is the most disaggregated level we observe, and bidders/winners are connected to one lot. These lots are often bundled together in a **tender**. Competition still happens at the lot level, but we often observe the same firms winning contracts of different lots on a single tender. The set-aside policy for SMEs is defined at

[2.9] A lot consists of some indivisible quantity of a good or service. In section 3 we clarify the terminology used throughout the paper.

[2.10] Available at https://portaldatransparencia.gov.br/download-de-dados/licitacoes

[2.11] Available at https://compras.dados.gov.br/docs/home.html the **lot level**: if the buying entities decides to set a lot aside for SMEs, all items in that lot will be exclusively sold by SMEs.

In addition, we gather firms' registry data from *Receita Federal* to identify SME firms and map their status on a yearly basis. We also access the number of employees for each firm from the *Relação Anual de Informações Sociais* (RAIS), an administrative longitudinal data set, provided by the Brazilian Ministry of Labor, that covers all formal firms and workers in Brazil. We use this variable as an indicator to analyze differences in firm size between the target and non-target groups.

2.3.1 Descriptive Statistics

In Table 2.1 we present descriptive statistics for different samples of our data. In column (1) we use the entire, unrestricted sample for the period 2013-2019, while in column (2) we restrict the sample to reverse auctions (RA) and framework agreements (FA). In the other two columns, we additionally restrict the sample for goods and materials (excluding services) and to estimated prices around the R$ 80,000 threshold for SME set-aside.

In the entire sample, we observe over 9 million lots purchased in 1.1 million tenders. Over 80% are classified as goods and materials. While almost 30% of lots are set-aside for SMEs, less than 6% of the total lot value is reserved for SME competition. The average value of the lots is $\approx$ R$ 28,000, with 5.4 competitors participating in the tenders on average. Even though SMEs account for an average of 83% of participants, they win approximately 77% of the lots. When the sample is restricted to materials purchased around the R$ 80,000 threshold under FA-RA (Column 4), we observe a lower fraction of lots set-aside to SME, a larger number of participants, and a smaller share of SME winners - consistent with the fact that we are excluding lots of lower-value where SMEs are more likely to win. This is the main sample used for the estimates in this paper, comprising 209,806 lots in 47,751 tenders.

In Table 2.1 we also document that there are approximately 3,200 buying entities making purchases in the data. When ranking largest buyers in terms of number of tenders, Table 2.2 shows that eight of the top ten buyers are federal universities.[2.12] We also present statistics on the most commonly purchased goods in Figure 2.1. According to the data, the product on which the government spends the most is basic food, followed by electrical mechanical parts for automotive vehicles. All items are classified using the CATSER

[2.12] Results are quite different when we consider monetary volume of purchases: the largest buyers are the Army, Navy and Health-related units. We present these results of top purchasing entities by volume in Table 2.A.1.

(catalog for services), and the CATMAT (catalog for goods) and are presented at 5 digits level aggregation in Figure 2.1.

Table 2.1: Descriptive Statistics

	(1) All	(2) FA-RA	(3) FA-RA only materials	(4) FA-RA between only materials and 50,000 and 110,000
Share products	82.2	87.9	100.0	100.0
Share SME set-aside (# lots)	29.9	41.5	44.7	16.0
Share SME set-aside (lot value)	5.7	7.4	10.0	13.5
Avg estimated value	28237.9	30445.0	23341.6	73028.3
Avg # participants	5.4	6.9	6.7	8.4
Avg # SME participants	4.5	5.9	5.8	6.5
Avg # EPP participants	2.7	3.5	3.5	3.8
Avg # ME participants	1.9	2.4	2.3	2.7
Share SME win	77.8	81.9	82.9	64.2
Share EPP win	48.2	54.5	55.7	41.9
Share ME win	29.6	27.5	27.2	22.4
Share SME new bidder	3.3	2.1	1.5	1.8
Avg # workers	72.3	46.6	26.4	59.9
Avg # wins (2 years prior)	104.3	111.0	120.6	63.7
N lots	9,556,832	6,895,762	6,064,399	209,806
N tenders	1,188,590	222,159	150,688	47,751
N buyer entities	3,226	2,644	2,446	2,059

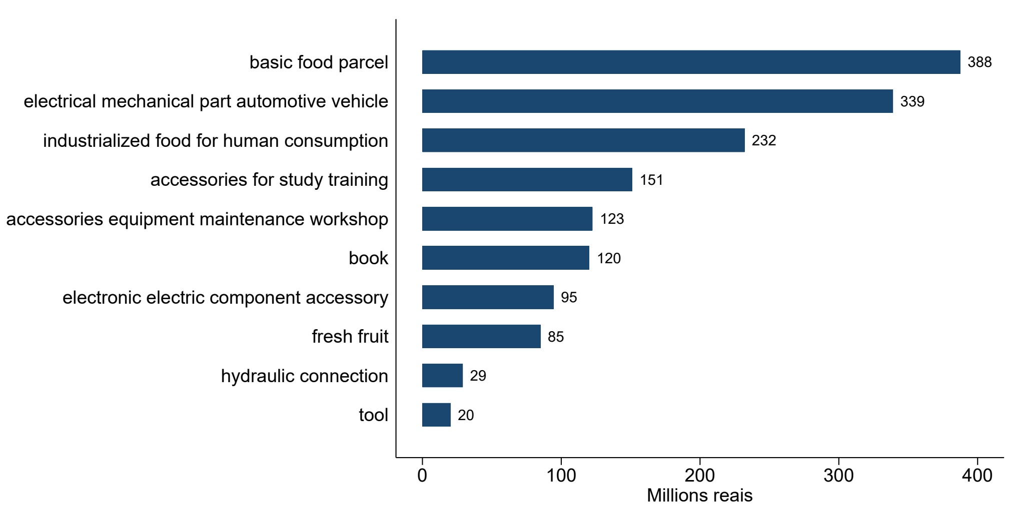
Note: This table shows statistics for four different samples from 2013-2019. Column 1 includes all purchases of goods and services made by federal buyer entities during this period, while Column 2 is restricted to purchases made through reverse auction and framework agreement. Column 3 is further restricted to purchases made through reverse auction and framework agreement for goods only. Column 4 maintains the same restrictions as Column 3 but adds the condition that the estimated value of the purchases falls between R$50,000 and R$110,000.

Table 2.2: Top 10 Entities Ranked by Tenders

	Number of tenders
universidade federal do rio grande do sul	24,122
universidade federal do para	13,062
universidade federal do parana	10,161
universidade federal de pernambuco	9,214
fundacao universidade de brasilia - fub	8,588
universidade federal de santa catarina	8,322
universidade federal do rio grande - furg	6,217
comissao nacional de energia nuclear-ipen	5,923
universidade federal de goias	5,695
departamento de logistica em saude - dlog	5,665

Note: This table presents the total number of tenders of the top 10 entities between 2013-2019.

Figure 2.1: Top 10 Product Ranked by Volumes



Note: This figure presents the average of year's volume of the top 10 product (5 digits classification) between 2013-2019.

2.4 Preliminary Evidence on Set-Aside Policy

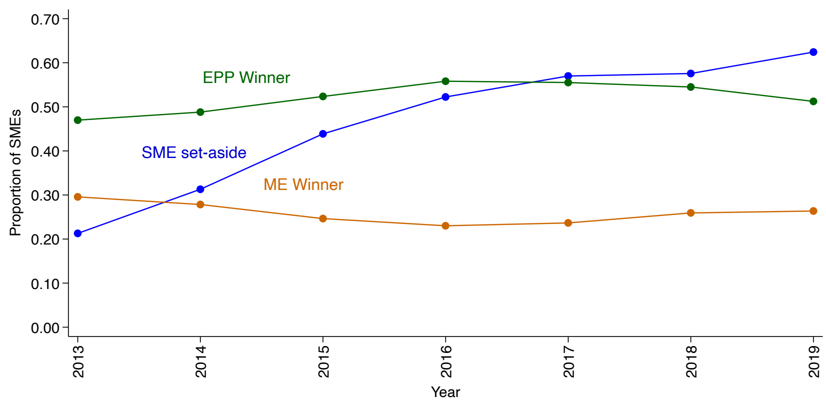
In this section, we start discussing the empirical evidence on the use of the set-aside policy. In Figure 2.1 we present the share of eligible lots, that is below R\$80,000, that were set aside for SMEs in each year and the share whose winners were SMEs. First, we note that the share of lots set-aside for small and medium enterprises has substantially increased since 2013, from about 20% to around 60% by 2019. Second, the share of lots being awarded to SMEs has also increased in the period, but starting from a much higher level, around 75% in 2013, and rising to 80% by 2018, before decreasing slightly to 78% in 2019. This means that, even when the share of set-asides was only one in five lots, SMEs were already winning 7 in 10 lots. Furthermore, these contracts were predominantly awarded to EPPs. On average, EPPs won 50% of the lots, while MEs won 25%.

In terms of monetary value, Figure 2.A.1 shows that in 2013, 15% of the total value of lots under R\$ 80,000 was reserved for SMEs, exceeding 45% in 2019. This suggests that the rise in preferential treatment was more significant for smaller values, well below the cutoff. By winning half of the lots, EPPs secured an average of 45% of the total amount, while MEs accounted for 25%. The remaining 30% was allocated to larger companies.

What explains the substantial increase in use of the set-aside policy over time? Since buying entities always had some discretion in the use of set-asides, we first consider the "extensive" margin presented in Figure 2.2. We document a constant increase in the share of buyers who use the set-aside policy at least once in each year. In 2013, approximately 66% of the more than 3,000 entities used the set-aside policy in any lot. That number steadily increases over time, plateauing at approximately 83% in the period 2017-2019.

At the same time we observe a substantial increase in the share of entities using the set-aside policy, we also observe increased use of set-asides within institutions over time. In Figure 2.4 we aggregate our buying entities at the ministerial-level and show that, out of twenty-two ministries, twenty-one increased the share of items that are set-aside for SMEs between 2013 and 2019. The two largest buyers at the ministerial level, Defense and Education ministries, increased their share of set-aside from approximately 25% and 15%, respectively, to over 60% by 2019. The broad-based increase in the use of set-asides is also observed at the level of agencies, a more disaggregated category of buyers. In Figure 2.5, we observe that out of more than one-hundred and fifty agencies, less than forty decreased the use of set-asides in their purchases. As can be seen by the size of the markers in the figure, those entities decreasing the use of set-asides are particularly small, with medium and large entities almost all increasing the use, albeit at very different rates. [2.13]

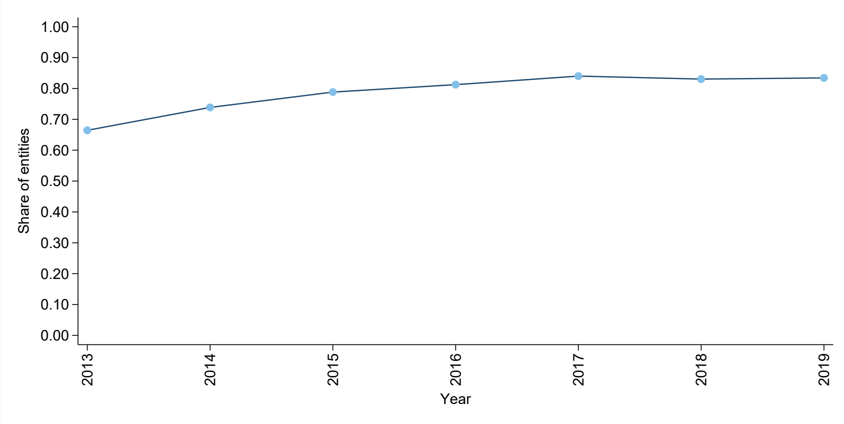
Figure 2.1: Share SME Set-Aside vs. Winners



Note: This figure presents the proportion of lots that used the set-aside benefit and the percentage of SMEs that won an item-tender process over the year, divided by MEs and EPPs. The data is based on reverse auction and framework agreement between 2013-2019. Additionally, data is restricted to lot values below R$ 80,0000.

[2.13] Purchases happen at the entity level, which are mapped to Agencies and Top Agencies. To illustrate, one purchasing entity might be the "Rio de Janeiro Campus" of the Federal University of Rio de Janeiro, mapped to the "Federal University of Rio de Janeiro" agency and the "Ministry of Education" top agency.

Figure 2.2: Share of Entities Using Set-Aside Policy



Note: This figure displays the annual percentage of entities that utilized the set-aside benefit at least once a year in reverse auction and framework agreement between 2013-2019.

Figure 2.3: Use of Set-Aside Over Time (2019 vs. 2013) Across Purchasing Entities

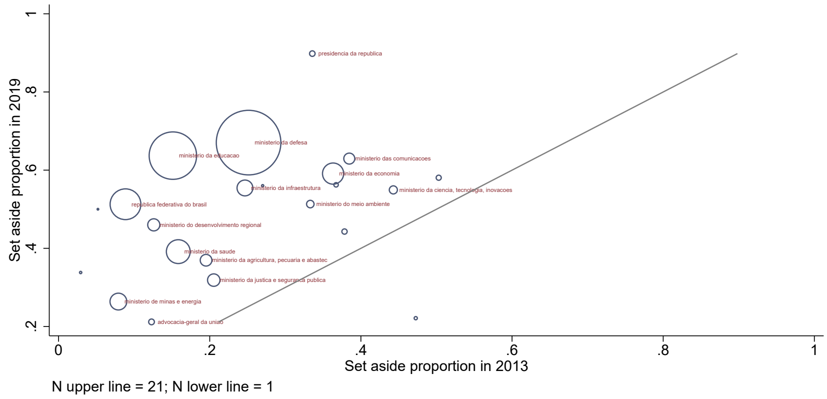


Figure 2.4: Ministry-level

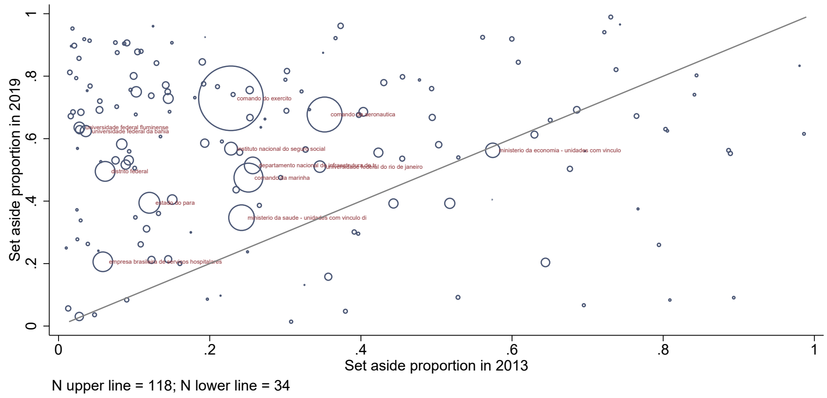


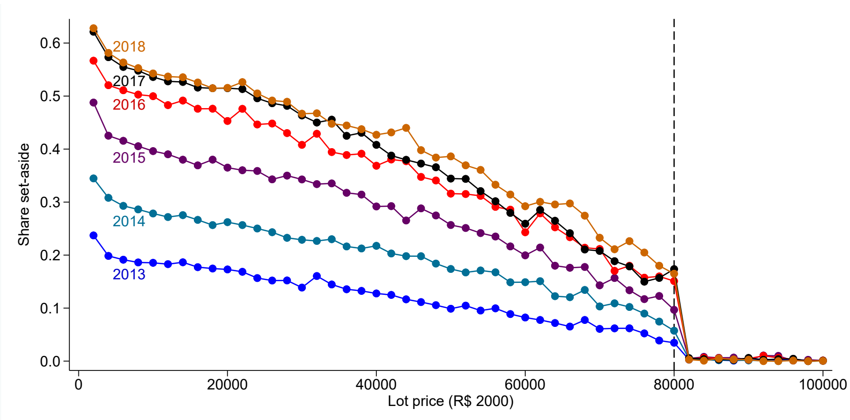
Figure 2.5: Agency-level

Note: This figure presents a scatter plot of the share of set-aside number of lots set side in 2019 vs. 2013 by ministry and agency level. The size of markers is proportional to the total estimated volume of the organs in 2019. The dots labeled are only the top 15 volumes. The data is restricted to reverse auction and framework agreements.

So far we have seen that the use of set-aside policies have substantially increased over time, both within and between purchasing entities. Now we turn to investigate how the use of set-aside changes across the distribution of estimated lot prices.

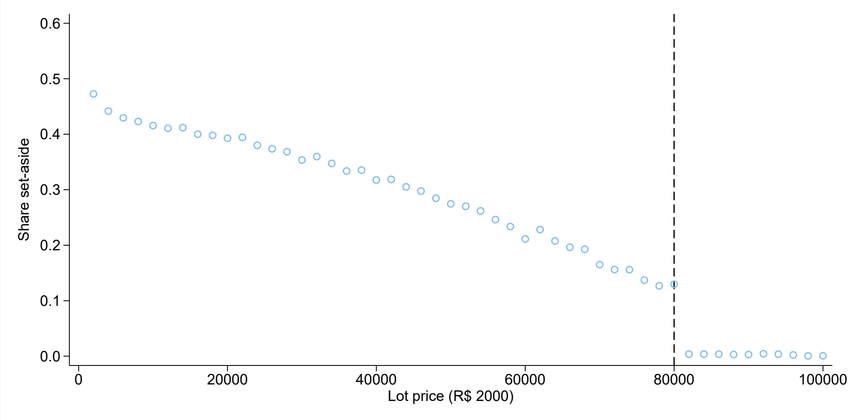
The set-aside policy for SMEs is restricted to lot prices estimated at less than R$ 80,000. The data confirms this is followed in practice: in Figure 2.6 we present the share of lots set-aside for SMEs, across the distribution of lot sizes and years. We note three stylized facts from these data. First, even though the only restriction is that lot prices are below R$ 80,000, the use of set-side is highly dependent on prices even below that level. The downward slope of the curves in each year shows that set-asides are much more common for lots of small prices than those with higher prices but still below the threshold. Second, the aggregate increase in the use of set-aside over time is mirrored across the distribution: starting in 2013, the use of set-asides increase across all levels of prices, until it stabilizes in the period 2016-2018. At those years, almost 60% of lots with prices below R$2,000 are set-aside for SMEs, while for those close but below the cutoff the rate is closer to 20%. Finally, we note that in all years the R$ 80,000 limit for the use of set-asides is almost universally binding, given that the rate of set-aside uses above those levels is virtually zero in all periods. We pool all years together in Figure 2.7 and show that discontinuity more clearly. While 15-20% of lots slightly below R$ 80,000 is set-aside for SMEs, that rate falls to zero above that threshold.

Figure 2.6: Proportion of Set-Aside Benefit by Year



Note: This figure presents the proportion of lots that used the set-aside benefit in each bin of R$2000 of lot value by year. The data is restricted to reverse auction and framework agreement between 2013-2019.

Figure 2.7: Proportion of Set-Aside Benefit



Note: This figure presents the proportion of lots that used the set-aside benefit in each bin of R$2000. The data is restricted only to reverse auction and framework agreement between 2013-2019.

2.5 Empirical Strategy

In this section, we explain our empirical strategy to assess whether the use of SME set-aside affects procurement outcomes, such as the number of bidders and the characteristics of suppliers. A naive approach would be to estimate an OLS model where we simply regress outcomes of interest Y_i on an indicator B_i of whether the lot was set-aside for SMEs and a range of controls for observable characteristics of the lot:

$$Y_i = \alpha + \beta_n B_i + \theta X_i + \epsilon_i \tag{2.1}$$

The coefficient β_n presents the differences in average outcomes of interest between lots that were reserved for SMEs and those not reserved. As we document in the previous section, there is a wide variation in the adoption of the set-aside policy between agencies, within agencies and over time. If the decision to concede the benefit is correlated with other unobservable factors that also determine outcomes of interest (i.e. if $E(B_i \epsilon_i | X_i) \neq 0$), then the coefficient of this regression does not represent the causal effect of setting-aside lots for SMEs. It is reasonable to assume this is likely the case, e.g. if specific auctioneers are more likely to use the set-aside policy and also put less effort to attract bidders, decreasing average competition in their tenders, then the coefficient β_n will also capture that correlation and bias our estimate of the causal effect of interest.

To address the problem of endogeneity related to the decision to set lots aside for SMEs, we implement a regression discontinuity design (RDD), in which we compare lots arbitrarily close to the threshold of R$ 80,000 but in differences sides of the cutoff. Since the adoption of the set-aside policy below the cutoff is not mandatory, and we document that in practice adoption around the cutoff is approximately 15%, we adopt a *fuzzy RDD* approach, in which being below the cutoff is an instrument to having an SME set-aside.

Formally, let Z be the indicator of assignment to treatment; V the lot reserved price; and c the threshold to be eligible.

$$Z_i = \begin{cases} 1 & \text{if } V_i \leq c \\ 0 & \text{if } V_i > c \end{cases}$$

In addition, let $Pr(B_i|V_i = v)$ be the probability of receiving the set-aside benefit B, given the lot reserved price V. Then, we can estimate the regression as a standard IV estimation. In this context, B is endogenous and $Z_i = \mathbb{I}\{V_i < c\}$ is the instrument, which affects Y_i only through B_i.

We estimate a local linear regression around the threshold, following Calonico et al. (2020):

$$Y_i = \alpha + \beta_{IV} B_i + \delta V_i + \lambda_t + \epsilon_i \tag{2.2}$$

Where Y_i refers to the procurement outcome, such as the number of bidders in lot i; B is the endogeneous adoption of the set-aside policy, and λ_t is a vector of year-quarter fixed effect.[2.14] Our coefficient of interest is β_{IV}, the coefficient of the instrumented regression. We instrument the adoption of the set-aside using the first-stage:

$$B_i = \alpha' + \gamma Z_i + \delta' V_i + \lambda_t' + \varepsilon_i$$

In all specifications, the sample is restricted to lots whose value is between R$ 50,000 and R$ 110,000.

One key test to assess the possible validity of our empirical strategy is whether there is manipulation of prices around the cutoff. If lot prices can be manipulated in order to be at specific sides of the cutoff, outcomes can be different not because of the availability of the set-aside policy but due to differential sample selection on either side.

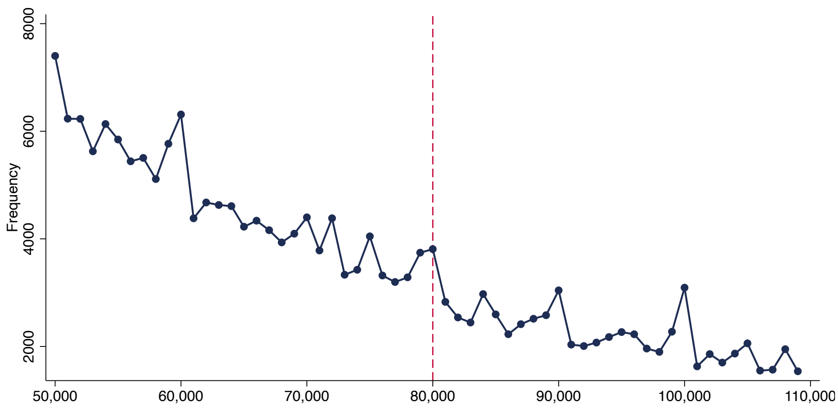
The literature on contract value manipulation in public procurement is extensive. Studies by Palguta and Pertold (2017b), Szucs (2023), Carril et al. (2021), Coviello et al. (2022), and Fazio (2022) identify bunching of procurement contracts just below regulatory thresholds designed to bypass open competition, and examine the implications for procurement outcomes. Manipulation can often be subtle, taking forms beyond simply adjusting estimated costs;

[2.14] In alternative specifications we also include state and agency fixed effects (Table 2.A.3 and Table 2.A.5). In Table 2.5 we also include fixed effects for the interaction between year-quarter and product categories.

for example, agencies may split the procurement of similar products into separate contracts rather than bundling them into a single process (Fazio, 2022). In the case of Brazil's set-aside policy, however, the manipulation of lot values below the threshold is less likely, as the cutoff amount is relatively high. For certain products with significant SME competition, the R$ 80,000 threshold may be high enough to discourage contract fragmentation. As for manipulation above the threshold, low compliance with the policy suggest that purchasing entities may bypass the policy without needing to inflate the value above R$ 80,000.[2.15] Still, sectoral differences could play an important role: in sectors where justifying the non-compliance is more challenging, the likelihood of manipulation above the threshold may increase.

In Figure 2.1, we present the distribution of lots value around the threshold. As we can note, there is no evidence of bunching near the cutoff, alleviating concerns about manipulation. As a robustness check, we also do the analysis by year, and Figure 2.2 presents a similar trend. Finally, we apply the test developed by McCrary (2008) to examine whether there is a discontinuity in the density of the assignment variable. Figure 2.3 presents these results graphically, indicating that there is no evidence of sorting around the threshold.

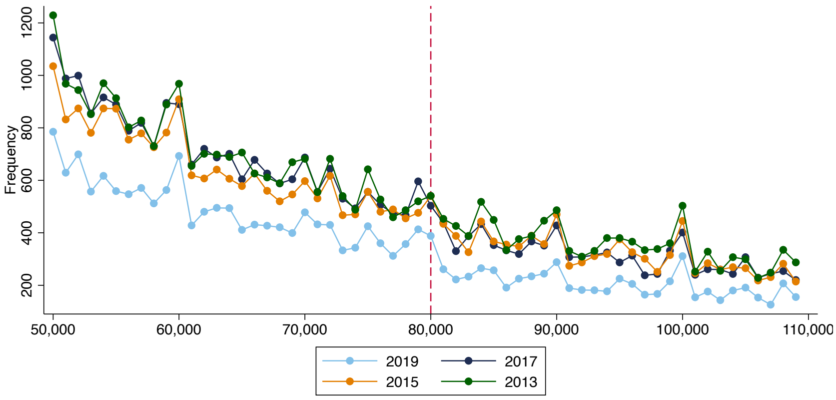
Figure 2.1: Distribution of Lot Values



Note: This figure presents the number of lots aggregated by bins of R$1000 of lot value. The data is restricted to goods purchased under reverse auction and framework agreement between 2013-2019 with an estimated value between R$ 50,000 and R$ 110,000.

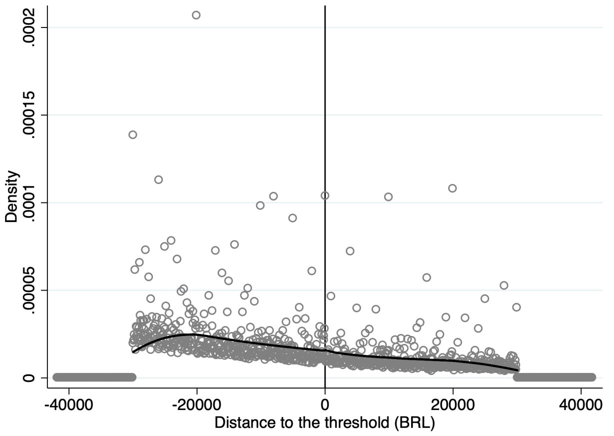
[2.15] This results from specific exceptions allowed by legislation, as discussed in Section 2.2.2.

Figure 2.2: Distribution of Lot Values by Year



Note: This figure presents the number of lots aggregated by bins of R$1000 of lot value and by year. The data is restricted to goods purchased under reverse auction and framework agreement between 2013-2019 with an estimated value between R$ 50,000 and R$ 110,000.

Figure 2.3: McCrary Discontinuity Test

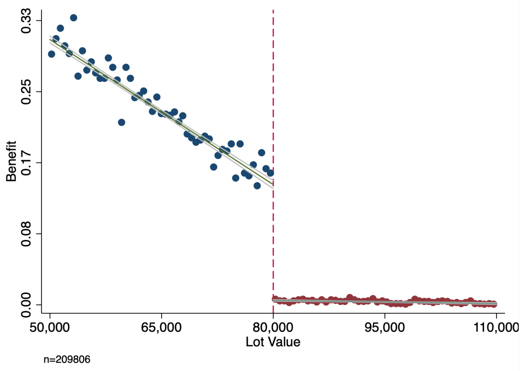


Note: The running variable is the difference between the lot value and the R$ 80,000 threshold. Confidence intervals are at the 95% level. In this overall sample, the discontinuity test is 0,024 and the standard-error 0,0157, suggesting that the null hypothesis of no sorting can not be rejected.

2.6 Results

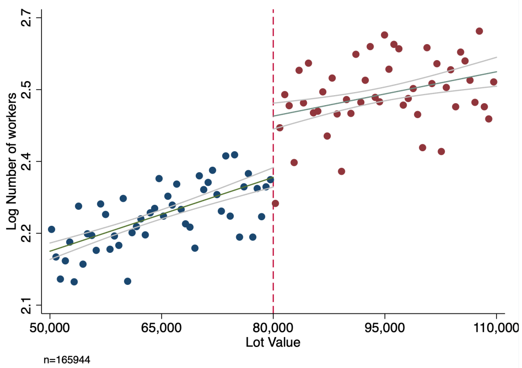
We start this section by graphically presenting our first-stage, the change in probability of set-aside adoption across the R$ 80,000 threshold. Consistent with the raw data presented in Figure 2.7, we document in Figure 2.1 that the adoption of the policy is decreasing in contract size and, precisely at the threshold, it decreases from approximately 15% to zero above the threshold. In addition, as shown in Figure 2.2, our analysis confirms that firms receiving the set-aside benefit also tend to have considerably fewer employees. This finding provides further evidence of the intended targeting of the policy towards this group.

Figure 2.1: Fraction of Lots with Set-Aside



Note: The figure shows the fraction of lots with preferential treatment. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The grey lines denote the confidence intervals plotted for fitted lines at the 95% level.

Figure 2.2: Workforce Size in Winning Firms



Note: The Figure shows the log transformation in the number of workers of the winning firms. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The grey lines denote the confidence intervals plotted for fitted lines at the 95% level.

Having documented the magnitude of the change in adoption precisely at the threshold, we now turn to present visual evidence of the reduced-form effect of the policy: how does competition change around the R$ 80,000 threshold? We present this evidence in Figure 2.3. We first document that the number of bidders seem to significantly change precisely above the threshold: while the average number of bidders below the threshold is around 8.2, for lots slightly above that threshold the average number is around 8.5 - a 3.5% increase in the average number of bidders (Panel A). That is, we observe a higher number of bidders when the set-aside policy is no longer available above the threshold. On the other hand, Panels B and C show that the number of SME bidders remains unchanged, while the number of large bidders increases above the R$ 80,000 threshold.

Another common measure in public procurement competitiveness is whether a specific lot has more than one bidder, that is, if there is effective competition for the lot. We replicate the same exercise in Panel D of Figure 2.3, showing how the share of non single-bidder lots vary across the threshold. The results indicate that below the threshold, the percentage of lots with more than one bidder is slightly higher. However, at the threshold, this percentage remains constant. Despite a decrease in the overall number of bidders, the policy does not appear to increase the incidence of auctions with only one bidder.

We summarize these findings in Table 2.1. We begin by estimating the 'naive OLS' model in Panel A, in which we simply regress one outcome of interest, the number of bidders, in a dummy for lots that include the SME set-aside benefit, controlling for year-quarter dummies. The estimated coefficient is negative and significant - it suggests that lots with the set-aside policy have on average 1.6 less bidders than those without the set-aside, a decrease of 19%. As we previously discussed, this cannot be interpreted as the causal effect of the policy since the decision to set lots for SMEs might be correlated with other unobservables that also influence competition.

Panel B of Table 2.1 presents the main results of our fuzzy RDD estimates. In this specification, we use the non-parametric optimal bandwidth procedure proposed by Calonico et al. (2014a). As in Panel A, we control for year-quarter dummies. The estimates suggest that the number of bidders decrease by 1.9 (-23%). That is, lots just below the cutoff that are set aside exclusively for SMEs have, on average, fewer bidders than lots just above the cutoff - not set aside for SMEs. We also presents the estimates for SME and Non-SME (large) bidders, showing that the decline in the number of bidders is primarily driven by reduced participation from larger firms. On the other hand, there is no evidence that the policy increased the entrance of SMEs.

To ensure our results are not being driven by a specific choice of bandwidth, Table 2.A.2 presents some robustness checks. We compare the regressions using the optimal bandwidth with variations using half or double bandwidth. We also vary the kernel specification and results remain similar. Additionally, Table 2.A.3 incorporates state and agency fixed effects into the main specification. By adding agency fixed effects, the reduction in the number of bidders becomes even more pronounced, averaging around 3 participants.

Given the significant heterogeneity among SME firms, with annual revenues ranging from R$ 81,000 to R$ 4.8 million, we distinguish between MEs and EPPs in Columns 4 and 5. The OLS estimate is negative, indicating that the number of ME bidders decreases by 0,06 bidders (2,5%). However, the RDD coefficient for ME participants is positive, suggesting an increase of 0.6 bidders (23%). This is observed across all RDD specifications presented in Table 2.A.2, even though only the regression using Uniform Kernel is statistically significant. On the other hand, the coefficients for EPPs are consistently negative, although they are statistically significant only when controlling for state or agency fixed effects, which helps explain the greater reduction in competition discussed in the previous paragraph. This is consistent with the following story: if there is a positive effect on entry, it should be higher for MEs, given that they have a greater advantage by not competing with larger firms.

Regarding the non-single-bidder outcome, the positive coefficient indicates that the policy increases the probability of having more than one bidder by 19 percentage points, thereby enhancing competition at this margin. Similar coefficients are presented in alternative specifications in Table 2.A.2 and Table 2.A.3. However, as documented in Figure 2.3 (Panel D), the percentage of lots with more than one bidder is high and nearly constant around the threshold. The high coefficient is therefore due to low compliance near the threshold.

In addition to analyzing participation outcomes, examining the characteristics of the winning firms provides valuable insights. Table 2.2 presents the same regressions outlined in Equations 2.1 and 2.2, focusing specifically on the results related to these firms. As shown in Column 1, being below the threshold of 80,000 is associated with a significant decrease in the number of workers. According to the RDD estimates (Panel B), this means that firms with set aside benefit have, on average, 75% fewer employees. Our estimates also show that reserved lots increase the winning rate for both MEs and EPPs. The likelihood of winning rises by 30 percentage points for MEs and by 29 percentage points for EPPs. Moreover, we find no evidence that reserved lots are being awarded to new SME winners—those who are winning for the first time during our analysis period.

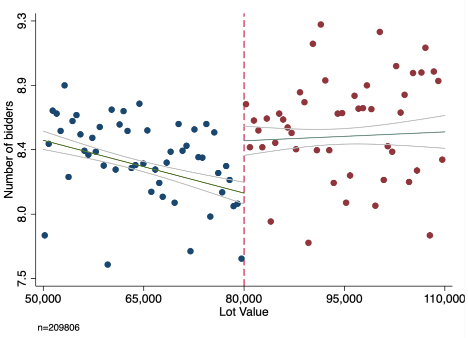
Additionally, we investigate whether lots with preferential treatment are more likely to be awarded to firms located in the same municipality or microrregion as the agency, considering that one explicit objective of the policy is to encourage local and regional development (Brasil, 2023).[2.16] However, we do not find evidence of lots being more awarded by local suppliers below the threshold. Actually, both OLS and RDD coefficients indicate the opposite sign. Similar results are supported by Table 2.A.4 across various bandwidths and

[2.16] One of the requirements for setting aside a lot is the presence of at least three locally based SME suppliers, although this does not imply that all three must participate in the auction.

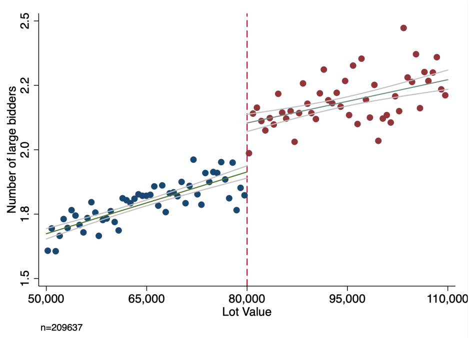
kernel specifications, and by Table 2.A.5 when including state or agency fixed effect.

In summary, it appears that the main effect of the policy is to increase the success rate of SMEs rather than increase their participation, measured by the number of bidders. The reduction in competition is largely attributed to the exit of larger firms, with no corresponding increase in the entry of SMEs. Surprisingly, some specifications also reveal a decline in the number of EPP participants, further contributing to the reduced competition. Equally important, we show that the policy does not appear to increase the hiring of local firms.

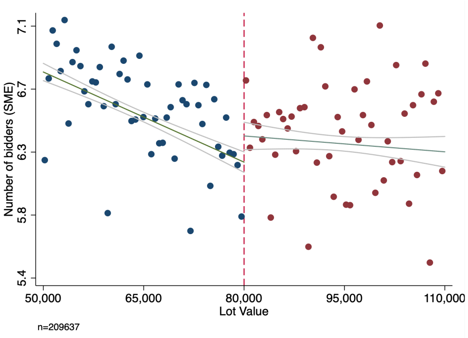
Figure 2.3: Competition Around Threshold



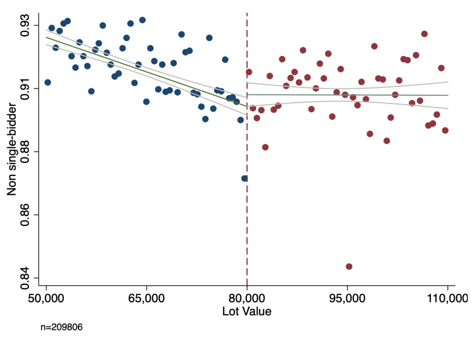
A: Number of Bidders



B: Number of Large Bidders



C: Number of SME Bidders



D: Non Single-Bidder

Note: Panels A, B, and C show the number of bidders, the number of large bidders, and the number of SME bidders per lot, respectively. Panel D show the fraction of lots with non-single bidders. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The grey lines denote the confidence intervals plotted for fitted lines at the 95% level.

Table 2.1: The Impact of Set-Aside on Competition

	N bidders	N large bidders	N bidders (SME)	N bidders (ME)	N bidders (EPP)	Non single-bidder
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: OLS						
Benefit	-1.669***	-1.616***	-0.056	-0.065***	0.009	0.016***
	(0.041)	(0.013)	(0.038)	(0.019)	(0.022)	(0.002)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.013	0.075	0.006	0.007	0.008	0.003
N	209,806	209,637	209,637	209,637	209,637	209,806
Panel B: RDD						
Benefit	-1.933	-1.750***	-0.197	0.590	-0.775	0.202***
	(1.190)	(0.341)	(1.076)	(0.511)	(0.646)	(0.054)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	7,154	8,642	7,349	8,108	7,299	7,814
N	209,806	209,637	209,637	209,637	209,637	209,806

Note: Data used in this table is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

Table 2.2: The Impact of Set-Aside on Winning Firms' Characteristics

	Log(N Workers)	ME Winner	EPP Winner	SME New Winner	Log(N Prev Win)	Same Municipality	Same Microrregion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: OLS							
Benefit	-1.101***	0.109***	0.233***	0.006***	0.072***	-0.028***	-0.016***
	(0.011)	(0.003)	(0.003)	(0.001)	(0.009)	(0.003)	(0.003)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
r2	0.059	0.011	0.035	0.007	0.051	0.007	0.006
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322
Panel B: RDD							
Benefit	-1.413***	0.300***	0.290***	-0.029	0.178	-0.048	-0.015
	(0.249)	(0.072)	(0.073)	(0.020)	(0.256)	(0.069)	(0.076)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	12,341	8,594	10,938	12,745	7,538	10,871	10,236
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322

Note: Data used in this table is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

2.6.1 Heterogeneity Across Market Structure

The lack of effect on SME entry presented in the previous section is somewhat surprising. At first glance, one would expect SMEs entering more when the auction is reserved and the competitive threat from larger firms is removed. In such cases, SMEs could update their probability of success and reconsider participating in the auction.

A potential explanation for not observing an increase in SME entry is the already high number of SMEs competing in lots without preferential treatment. The average number of SME participants is 6.5 below and 6.3 above the R$ 80,000 threshold.[2.17] This is consistent with the fact that SMEs constitute 99% of establishments in Brazil, with many of these firms having relatively high annual revenues.

In this Section, we investigate the extent to which the prior level of participation among SMEs drives our results. We examine the heterogeneous effects of the set-aside policy across different market structures. By market structure, we refer to the average participation rate of SMEs within various sectors, which is calculated as the number of SME participants divided by the total number of participants in each lot without the set-aside benefit. Each sector is identified by its 2-digit code, provided by the CATMAT catalog.[2.18] In total, we have 76 distinct sectors. The median sector has 80% of its bidders as SMEs. We then classify the sample into sectors with high SME participation (above the median) and those with low SME participation (below the median). For instance, items categorized as "Medical, dental, and veterinary supplies" are classified as having low SME participation, whereas items categorized as "Food Products" are classified as having high SME participation.

Table 2.3 summarizes the results of the regressions, showing that competition decreases in sectors with typically high SME participation and increases in sectors with typically low SME participation. As expected, there is a reduction in the number of large bidders for both groups. However, unlike the general null effect on SME entry, Panel A indicates that the policy increases the number of SME bidders by 2.8 in sectors with low SME participation, with the most significant effect coming from MEs (+1.6 bidders). This indicates that low SME participation was not due to a lack of available small firms but rather because they struggled to compete with larger companies. Once the lot is restricted, these firms are more inclined to participate.

[2.17] Considering the unrestricted sample (Column 3 from Table 2.1), the average is 5.6 participants below and 6.5 participants above the cutoff.

[2.18] The CATMAT is the official government catalog used to classify items with varying levels of disaggregation, ranging from 2-digit to 6-digit definitions.

On the other hand, Panel B shows that in sectors with high SME participation, the number of SMEs actually decreases by 3.8 bidders, with EPPs seeing a drop of 3 bidders. While this reduction might seem counterintuitive, it could be due to existing informal market division in these sectors, where firms have established specific niches. The set-aside policy may reinforce these divisions, leading to fewer firms competing directly for each lot. It is also worth noting that lots above the threshold may receive other preferential treatment for SMEs, such as quotas for divisible goods and the opportunity to match the lowest bid if their price is up to 5% higher than those of non-favored firms. Consequently, the set-aside benefit may not be a decisive factor for the participation of EPPs.

Regarding the characteristics of the winning firms, Table 2.4 confirms the results from the previous section. Lots under the set-aside policy are awarded to firms with fewer workers in both low and high SME participation sectors. There is also an increase in the probability of a lot being awarded to an SME, particularly in sectors with low participation. Panel A shows that the likelihood of a ME winning an auction increases by 33.6 percentage points, while for an EPP, it increases by 39.2 percentage points. In sectors with high SME participation, the only significant effect is for MEs, increasing by 20 percentage points. This indicates that in sectors in which SME participates the most, the success rate of EPPs is already high. However, the policy still plays a key role in improving the chances for MEs. In these cases, reserved lots tend to go to firms with a track record of winning contracts, rather than to new winners, as shown in Columns 4 and 5.

Table 2.3: The Impact of Set-Aside on Competition by SMEs Participation

	N bidders	N large bidders	N bidders (SME)	N bidders (ME)	N bidders (EPP)	Non single-bidder
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Low SME Participation						
Benefit	1.421	-1.362**	2.815**	1.662**	1.150	0.261***
	(1.524)	(0.577)	(1.336)	(0.646)	(0.803)	(0.092)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	8,504	8,735	8,492	8,293	8,609	8,829
Mean Above Cutoff	6.900	2.608	4.290	1.659	2.635	0.844
N	103,354	103,263	103,263	103,263	103,263	103,354
Panel B: High SME Participation						
Benefit	-5.292***	-1.535***	-3.802***	-0.694	-3.091***	0.027
	(1.301)	(0.299)	(1.188)	(0.585)	(0.734)	(0.035)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	10,719	12,219	10,531	11,595	10,183	11,135
Mean Above Cutoff	10.092	1.632	8.457	3.613	4.838	0.969
N	106,452	106,374	106,374	106,374	106,374	106,452

Note: This table uses data from reverse auctions and framework agreements, limited to products and covering the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

Table 2.4: The Impact of Set-Aside on Winning Firms' Characteristics by SMEs Participation

	Log(N Workers) (1)	ME Winner (2)	EPP Winner (3)	SME New Winner (4)	Log(N Prev Win) (5)	Same Municipality (6)	Same Microrregion (7)
Panel A: Low SME Participation							
Benefit	-1.696***	0.336***	0.392***	0.023	-0.005	-0.189*	-0.118
	(0.435)	(0.094)	(0.114)	(0.027)	(0.365)	(0.111)	(0.119)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	8,740	8,065	8,618	10,975	8,216	8,522	8,371
Mean Above Cutoff	2.956	0.128	0.270	0.015	3.081	0.273	0.329
N	89,466	98,654	98,654	98,654	96,722	97,851	97,851
Panel B: High SME Participation							
Benefit	-0.687**	0.200**	0.120	-0.084***	0.379	-0.002	-0.015
	(0.314)	(0.091)	(0.093)	(0.031)	(0.260)	(0.094)	(0.094)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	11,622	11,476	12,721	12,712	11,272	11,824	12,576
Mean Above Cutoff	1.931	0.300	0.502	0.029	2.738	0.358	0.438
N	76,478	96,596	96,596	96,596	98,546	95,471	95,471

Note: Data used in this table is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

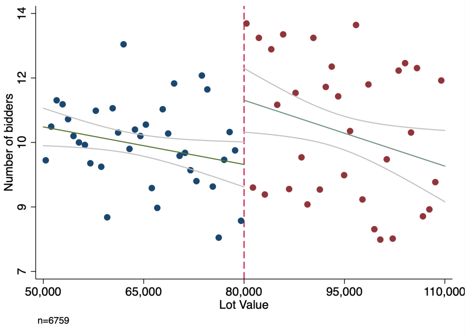
2.6.2 Price Effects

In previous sections, we observed that lots reserved for SMEs tend to have fewer bidders because larger firms can not compete, and the number of small participants does not increase. This raises the question of whether the reduced competition affects prices. To explore this, we restrict the analysis to a subset containing some of the most frequently bought products. We select the products based on their relative homogeneity and good specification. To improve the comparisons between auctions, we define a product as the combination of its 6-digits code from CATMAT and its unit of measure (Fazio, 2022; Fiuza et al., 2023). Table 2.A.6 provides a list with selected products and their most common unit measure.

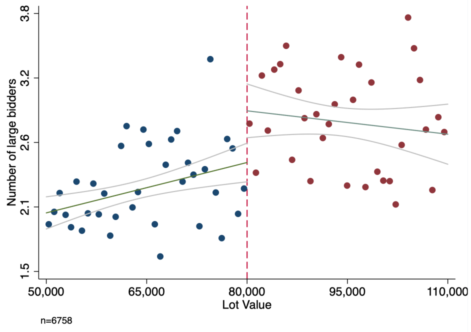
The final sample includes 184,163 lots from 31,103 tenders, purchased by 2,178 buyer entities and awarded to 10,145 distinct firms between 2013 and 2019. To handle outliers, we winsorized observations where the product's unit price fell below the 5th percentile or above the 95th percentile.

Figure 2.4 illustrates that at the threshold, the number of bidders increases by 15%, rising from approximately 10 to 11.5. The number of large bidders increases with contract size, rising from 2.5 to nearly 3 bidders around the cutoff before stabilizing above this point. Meanwhile, the number of SME bidders shows an interesting pattern: while it decreases with lot value, it increases precisely at the threshold, going from 7.5 to around 9 participants per lot, on average. For instance, some SMEs may prefer to compete in slightly more expensive lots even when a set-aside is in place, as long as they can remain competitive on price. As a result, their participation increases around the threshold but declines again for more expensive lots.

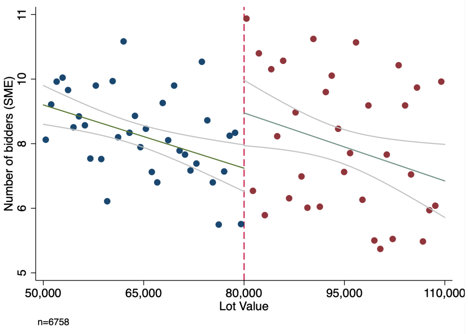
Figure 2.4: Competition Around Threshold - Restricted Sample



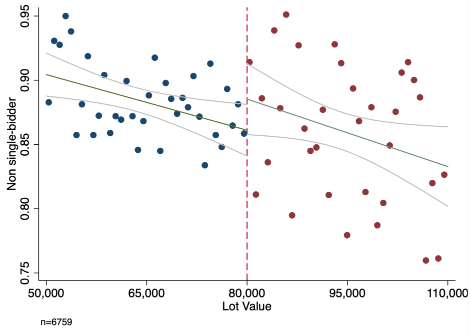
A: Number of Bidders



B: Number of Large Bidders



C: Number of SME Bidders



D: Non Single-Bidder

Note: Panels A, B, and C show the number of bidders, the number of large bidders, and the number of SME bidders per lot, respectively. Panel D show the fraction of lots with non-single bidders. The data is restricted to reverse auction and framework agreement, includes only products listed in Table 2.A.6, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The grey lines denote the confidence intervals plotted for fitted lines at the 95% level.

Table 2.5 exhibits the results of the OLS and RDD regressions using this new sample and including the log transformation of the item unit price as an additional outcome variable. The specifications compare the same products around the threshold, incorporating product-year-quarter fixed effects. These fixed effects also control for any variations in the overall demand for each product during that period.

Similarly to the general case, we find that lots with the set-aside benefit have, on average, less bidders than those without the set-aside. The OLS coefficient suggests that reserved lots have, on average, 3.6 fewer bidders, representing a 33% decrease. In this case, in addition to the smaller number of large firms participating, there is also a reduction in the participation of SME firms by 1.7 bidders (-21%). Despite that, we can not reject the hypothesis of null effects on prices.

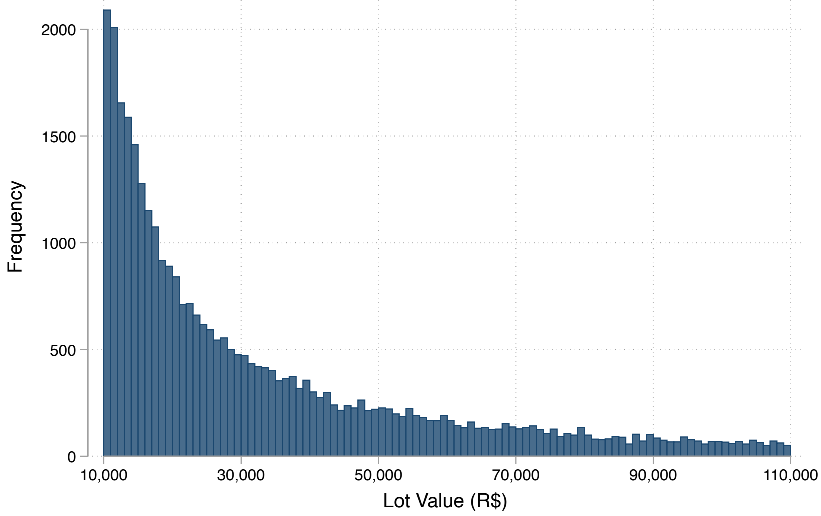
Panels B and C of Table 2.5 shows coefficients that align with the overall trend. However, the RDD coefficients are unreliable high. We raise two reasons for this. First, there is a limited number of observations within the bandwidth, even when applying double bandwidth selection. Figure 2.5 shows that lot values are predominantly concentrated below R$ 30,000, with few observations around the R$ 80,000 threshold. As a result, only 1,500 out of the 184,163 lots fall within the optimal bandwidth. A potential solution would be to expand the sample to include a broader range of cleaned products with higher lot values. This approach does come with the trade-off of potentially increasing selection bias.

Table 2.5: The Impact of Set-Aside on Competition - Restricted Sample

	N bidders (1)	N large bidders (2)	N bidders (SME) (3)	N bidders (ME) (4)	N bidders (EPP) (5)	Non single-bidder (6)	Log(Unit Price) (7)
Panel A: OLS							
Benefit	-3.671***	-1.970***	-1.696***	-0.787***	-0.909***	-0.064**	0.112
	(0.760)	(0.460)	(0.462)	(0.212)	(0.263)	(0.029)	(0.104)
Year-Quarter*Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	10.809	2.845	7.964	3.633	4.331	0.860	2.797
r2	0.635	0.590	0.647	0.623	0.620	0.366	0.855
N	6,691	6,690	6,690	6,690	6,690	6,691	6,440
Panel B: RDD - Optimal Bandwidth							
Benefit	-10.291**	-1.503	-8.953***	-4.360**	-4.651**	-0.338***	0.176
	(4.041)	(1.295)	(3.404)	(1.961)	(1.945)	(0.118)	(0.478)
Year-Quarter*Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	7,209	7,795	7,119	7,314	7,061	7,050	8,898
Eff Obs Left	811	882	805	820	796	797	1,009
Eff Obs Right	608	663	593	619	588	587	717
Mean	11.622	3.002	8.702	4.025	4.623	0.880	2.790
N	6,759	6,758	6,758	6,758	6,758	6,759	6,507
Panel B: RDD - Double Bandwidth							
Benefit	-12.312**	-2.770**	-9.912**	-6.180**	-3.711	-0.051	0.564
	(5.148)	(1.084)	(4.591)	(2.422)	(2.411)	(0.121)	(0.374)
Year-Quarter*Product FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	14,417	15,591	14,239	14,628	14,122	14,100	17,797
Eff Obs Left	1,791	1,950	1,771	1,818	1,752	1,752	2,189
Eff Obs Right	1,180	1,284	1,167	1,206	1,156	1,156	1,371
Mean	11.622	3.002	8.702	4.025	4.623	0.880	2.790
N	6,759	6,758	6,758	6,758	6,758	6,759	6,507

Note: Data used in this table is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

Figure 2.5: Distribution of Lot Value - Restricted Sample



Note: This figure presents the distribution of lot value, in Brazilian reais, for the products listed in Table 2.A.6.

Second, compliance near the threshold is very low, with only 11% of lots having set-aside status, compared to the 15% observed in the first stage when considering all products. If lot value is not a strong predictor of receiving treatment near the threshold, the instrument becomes weak. Therefore, our findings suggest that RDD estimates do not provide a clear picture of the local average treatment effect in this restricted sample. Nevertheless, the effect on price remains statistically insignificant.

2.7 Conclusion

This paper provides insights into the use and impact of set-aside policies for SMEs in public procurement auctions in Brazil. Using comprehensive data from public sources, including the Transparency Portal and Compras Dados, we gather information from federal purchases from 2013 to 2019. First, we find that the adoption of the policy is decreasing in contract size and that purchasing entities exercise substantial discretion in its use. On the other hand, we reject the hypothesis that procurement officers manipulate the value of contracts to stay on either side of the threshold.

Our key finding is that the use of set-asides for SMEs reduces competition in the auctions - there is a reduction in participation of larger firms with no corresponding entrance of SMEs. Our regression discontinuity design shows that the average number of bidders in an auction decreases by about 1.6 to 3 bidders, due to the usage of set-asides, depending on the specification - a large effect given the average number of participants below the threshold. However, this effect varies depending on the market structure. In sectors with low SME participation, the policy does result in a higher number of small bidders, which more than compensate the exit of larger firms.

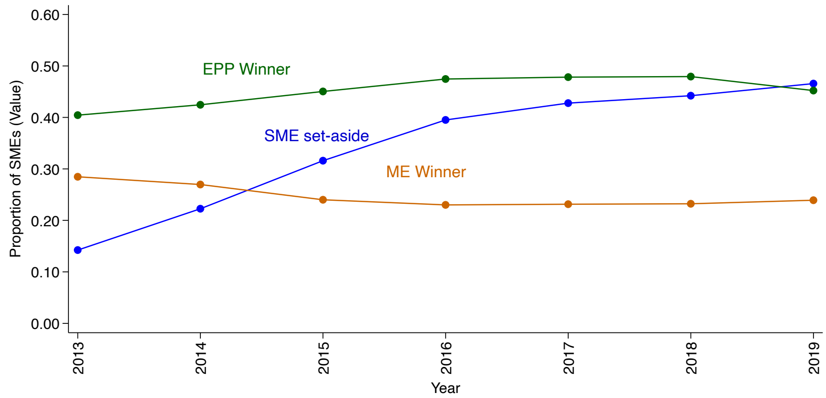
Finally, when focusing on a sub-sample of frequently purchased products, we find that, despite the observed low competition, there is no statistically significant effect on price increases. Still, given the limited number of observations around the optimal bandwidth and the low level of compliance near the cutoff, a more careful analysis of these effects is necessary.

As future research, it would be interesting to investigate the competition between MEs and EPPs on reserved lots. Our findings indicate that 50% of the lots below the threshold are awarded to EPPs, while only 25% are awarded to MEs. According to the CGU (Brasil, 2023), EPPs often benefit from advantages that they may not necessarily need, as they are capable of competing and winning a significant portion of auctions without such benefits. This suggests that there is potential for public policy to better differentiate between micro and small firms.

Future research is also needed to better understand the factors behind the low compliance with the set-aside policy. Despite legislative clarifications after 2014, we observe significant variation across agencies, with many showing compliance rates below 50%. The compliance is even worse the closer we get to R$ 80,000. It remains unclear how much of this low compliance is due to a shortage of available SME suppliers versus issues related to discretionary decision-making, lack of justification, or system difficulties. Understanding these dynamics is crucial for improving the effectiveness of the policy.

2.A Appendix

Figure 2.A.1: Share SME set-aside vs. winners in terms of value



Note: This figure presents the proportion of lots that used the set-aside benefit and the percentage of SMEs that won an item-tender process over the year, divided by MEs and EPPs. The data is based on reverse auction and framework agreement between 2013-2019. Additionally, data is restricted to lot values below R$ 80,0000.

Table 2.A.1: Top 10 Entities Ranked by Volume

	total volume estimated
comando do exercito	84,203,092,789
comando da aeronautica	23,246,437,302
ministerio da saude - unidades com vinculo di	14,943,732,148
comando da marinha	14,636,050,155
estado do para	7,070,805,683
empresa brasileira de servicos hospitalares	6,769,428,150
distrito federal	6,337,138,718
empresas de energia	6,291,738,846
fundacao oswaldo cruz	5,883,242,885
departamento nacional de infraestrutura de tr	5,757,446,520

Note: This table presents the total volume of the top 10 entities between 2013-2019.

Table 2.A.2: The Impact of Set-Aside on Competition Across Different Bandwidths and Kernels

	N bidders	N large bidders	N bidders (SME)	N bidders (ME)	N bidders (EPP)	Non single-bidder
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Half Bandwidth						
Benefit	-0.527	-1.869***	1.169	1.226	-0.197	0.292***
	(1.791)	(0.505)	(1.626)	(0.769)	(0.981)	(0.085)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	3,577	4,321	3,674	4,054	3,650	3,907
N	209,806	209,637	209,637	209,637	209,637	209,806
Panel B: Double Bandwidth						
Benefit	-1.586*	-1.509***	-0.062	0.514	-0.623	0.116***
	(0.819)	(0.239)	(0.738)	(0.353)	(0.440)	(0.035)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	14,308	17,284	14,698	16,215	14,599	15,628
N	209,806	209,637	209,637	209,637	209,637	209,806
Panel C: Epanechnikov Kernel						
Benefit	-2.110*	-1.731***	-0.378	0.507	-0.830	0.187***
	(1.130)	(0.323)	(1.021)	(0.486)	(0.611)	(0.050)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
BW	7,154	8,642	7,349	8,108	7,299	7,814
N	209,806	209,637	209,637	209,637	209,637	209,806
Panel D: Uniform Kernel						
Benefit	-1.962*	-1.616***	-0.485	0.968**	-0.911	0.184***
	(1.066)	(0.297)	(0.959)	(0.427)	(0.572)	(0.047)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
BW	7,154	8,642	7,349	8,108	7,299	7,814
N	209,806	209,637	209,637	209,637	209,637	209,806

Note: This table presents regressions using different set of bandwidths and kernels. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

Table 2.A.3: The Impact of Set-Aside on Competition - Additional Fixed Effects

	N bidders (1)	N large bidders (2)	N bidders (SME) (3)	N bidders (ME) (4)	N bidders (EPP) (5)	Non single-bidder (6)
Panel A: RDD (YQ and State FE)						
Benefit	-2.310**	-1.664***	-0.591	0.554	-1.164**	0.124***
	(1.030)	(0.303)	(0.929)	(0.446)	(0.551)	(0.044)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Agency FE	No	No	No	No	No	No
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	9,593	10,817	9,805	10,390	9,949	10,435
N	207,755	207,586	207,586	207,586	207,586	207,755
Panel B: RDD (YQ and Agency FE)						
Benefit	-3.031***	-1.467***	-1.568*	0.057	-1.663***	0.128***
	(1.053)	(0.342)	(0.918)	(0.439)	(0.571)	(0.048)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	No	No	No	No	No
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	8,934	8,598	9,492	10,323	8,887	9,316
N	209,806	209,637	209,637	209,637	209,637	209,806

Note: This table presents regressions using different set of bandwidths and kernels. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. The total number of bidders comprises both large bidders and SME bidders. The number of SME bidders is the sum of ME and EPP bidders. P-values: * 0.10 ** 0.05 *** 0.01

Table 2.A.4: The Impact of Set-Aside on Winning Firms' Characteristics Across Different Bandwidths and Kernels

	Log(N Workers) (1)	ME Winner (2)	EPP Winner (3)	SME New Winner (4)	Log(N Prev Win) (5)	Same Municipality (6)	Same Microrregion (7)
Panel A: Half Bandwidth							
Benefit	-1.374***	0.390***	0.296***	-0.033	1.000**	-0.077	0.045
	(0.353)	(0.112)	(0.110)	(0.030)	(0.402)	(0.105)	(0.116)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	6,171	4,297	5,469	6,373	3,769	5,436	5,118
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322
Panel B: Double Bandwidth							
Benefit	-1.255***	0.206***	0.288***	-0.015	0.006	-0.068	-0.060
	(0.184)	(0.049)	(0.051)	(0.014)	(0.173)	(0.049)	(0.053)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	24,683	17,188	21,876	25,490	15,076	21,742	20,472
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322
Panel C: Epanechnikov Kernel							
Benefit	-1.407***	0.290***	0.287***	-0.029	0.064	-0.044	-0.022
	(0.240)	(0.068)	(0.068)	(0.019)	(0.242)	(0.066)	(0.072)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BW	12,341	8,594	10,938	12,745	7,538	10,871	10,236
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322
Panel D: Uniform Kernel							
Benefit	-1.290***	0.288***	0.272***	-0.031*	-0.094	-0.054	-0.046
	(0.233)	(0.062)	(0.064)	(0.018)	(0.223)	(0.061)	(0.065)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BW	12,341	8,594	10,938	12,745	7,538	10,871	10,236
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322

Note: This table presents regressions using different set of bandwidths and kernels. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. P-values: * 0.10 ** 0.05 *** 0.01

Table 2.A.5: The Impact of Set-Aside on Winning Firms' Characteristics - Additional Fixed Effects

	Log(N Workers)	ME Winner	EPP Winner	SME New Winner	Log(N Prev Win)	Same Municipality	Same Microrregion
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: RDD (YQ and State FE)							
Benefit	-1.300***	0.270***	0.259***	-0.036*	-0.001	-0.064	-0.058
	(0.256)	(0.064)	(0.067)	(0.020)	(0.213)	(0.064)	(0.065)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Agency FE	No	No	No	No	No	No	No
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW	11,976	10,624	13,222	12,670	10,400	12,274	13,098
N	164,293	193,322	193,322	193,322	193,674	193,322	193,322
Panel B: RDD (YQ and Agency FE)							
Benefit	-1.101***	0.232***	0.220***	-0.039*	0.148	-0.078	-0.080
	(0.270)	(0.059)	(0.066)	(0.022)	(0.221)	(0.065)	(0.068)
Year-Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State FE	No	No	No	No	No	No	No
Agency FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
BW	10,170	12,011	13,510	10,786	9,713	12,179	12,178
N	165,944	195,250	195,250	195,250	195,268	193,322	193,322

Note: This table presents regressions using different set of bandwidths and kernels. The data is restricted to reverse auction and framework agreement, includes only products, and covers the period from 2013 to 2019. In addition, we consider lots with estimated value between R$ 50,000 and R$ 110,000. P-values: * 0.10 ** 0.05 *** 0.01

Table 2.A.6: List of Selected Materials

Product	**Unit**
A4 paper	Package/Ream
Alcohol	1 Liter
Ballpoint Pen	Unit
Battery	Unit
Broom	Unit
Clip	Box
Compressed Gas	1 Cubic Meter
Disposable Coffee Cup	Package
Diesel	1 Liter
Envelope	Unit
External HD	Unit
File Folder	Unit
Flexible Electric Cable	1 Meter
Fluorescent Lamp	Unit
Fresh Peas	1 Kilogram
Fuel Filter	Unit
Gloves for Non-Surgical Procedure	Box
Gloves for Surgical Procedure	Pair
Gasoline	1 Liter
Highlighter Pen	Unit
Nail	1 Kilogram
Printer Toner Cartridge	Unit
Sand	1 Cubic Meter
Sodium Chloride	Bottle
Stapler	Unit
Sugar	1 Kilogram
Water	20 Liter Gallon
Whiteboard Pen	Unit
Switch	Unit

3
Reelection Incentives and Corruption: New Data and an Assessment of the Literature

Nathalia Sales
PUC-Rio

Martin Mattsson
NUS

Ricardo Dahis
Monash University

Abstract

Reliable data on corruption is notoriously hard to find. We extend previous manual attempts at classifying corruption audit reports with a Large Language Model (LLM) to encode reports from 2,197 Brazilian municipalities. We find that the correlations between the LLM's assessment of the extent of corruption and manual assessments are similar to the correlations between different manually encoded datasets. We then apply our extended corruption data to reassess key findings in the literature on the effect of reelection incentives on corruption. We find some evidence that reelection incentives reduce corruption, corroborating previous findings. However, the effect sizes are smaller in the extended data and the effects are only statistically significantly different from zero for one of the three outcome variables. We introduce alternative explanations to the empirical findings.

3.1
Introduction

Corruption is a serious problem across the world and is especially severe in low- and middle-income countries (Svensson, 2005). Several approaches have been proposed to combat corruption, including increasing the wages of potentially corrupt officials (Van Rijckeghem and Weder, 2001), reducing corruption opportunities by decreasing regulation and thus removing opportunities for corrupt behavior (Rose-Ackerman, 1998), improving both top-down and local-level monitoring (Olken, 2007; Björkman and Svensson, 2010), and implementing new technologies such as "e-governance" (Banerjee et al., 2020). To the extent that voters value honesty, reelection incentives would also be a strong force against corruption (Ferraz and Finan, 2008, 2011).

However, one of the main problems in evaluating anticorruption measures is the limited availability of reliable corruption data (Olken, 2007; Olken and

Pande, 2012).[3.1] To overcome the concerns associated with this approach, more direct measures of corruption have been developed. One popular data source is the audit reports generated by the Random Audits Anti-Corruption Program for Brazilan Municipalities, introduced in 2003 by the *Controladoria Geral da União* (CGU). These reports have been extensively used in the literature to quantify corruption (Ferraz and Finan, 2008, 2011; Brollo et al., 2013; Avis et al., 2018; Colonnelli and Prem, 2022; Ash et al., 2020).

A drawback of using audit reports is that it is often not feasible to manually read numerous documents and quantify the amount of corruption uncovered by each report. In this paper, we use a Large Language Model (LLM) to read 2,197 reports (~175,000 pages) and extend previous manual readings of the reports. To make them more legible to the LLM, we employ Retrieval Augmented Generation (RAG) to extract pertinent contextual information from each report. This information is then fed into OpenAI's GPT-4. By bringing the appropriate information and inserting it into the model prompt, the LLM is able to answer specific questions about the audits' findings. The LLM assesses whether irregularities associated with corruption were found in each report, the share of resources audited where any corruption was found, and the number of corruption cases.

We construct a new dataset on corruption spanning all reports from 1,914 Brazilian municipalities audited between 2003 to 2015. We compare these data with three existing manually coded datasets: 2003-2005 audit reports coded by Ferraz and Finan (2011), 2003-2009 audit reports coded by Brollo et al. (2013), and a dataset managed by the CGU containing all irregularities found by auditors from 2005 to 2015. When comparing the results from the LLM's reading of the audit reports with the manually coded data we find correlations of the share of audited resources where corruption was found between 0.44 and 0.47 . These correlations are similar to the correlations between the manually coded datasets that range from 0.48 and 0.71 .[3.2]

We then combine our data with the empirical strategy from Ferraz and Finan (2011) to estimate the effect of reelection incentives on corruption. Mayors only face reelection incentives in their first term, as they are only allowed to be reelected once. Therefore, we compare mayors in their first term with mayors in their second term controlling for a wide range of observable charac-

[3.1] A common attempt to measure corruption is through perception-based indicators, such as the Transparency International's Corruption Perceptions Index (CPI). However, surveys based on perception may be considerably biased by respondents' beliefs and characteristics (Olken, 2009).

[3.2] While we use manually encoded reports to validate our methodology, we do not rely on previous reports to "train" the LLM. Thus our methodology is not reliant on audit reports having been manually encoded and there is no in-sample vs. out-of-sample validation needed.

teristics. We find some evidence that reelection incentives reduce corruption, corroborating previous findings from the literature. However, the effect sizes are smaller and more precisely estimated than in the original Ferraz and Finan (2011) paper. Furthermore, the effects are only statistically significantly different from zero for one of the three measures of corruption.[3.3] The reduced effect is also supported by estimates derived from alternative datasets such as those from Brollo et al. (2013) and CGU.

Differences in estimates that arise from different encodings of the data can be explained to a large extent by differences in the estimated effect over time. Focusing on the period 2001-2004, the period considered by Ferraz and Finan (2011), the estimates from the four encodings of the data confirm that reelection incentives reduce corruption. However, for the two subsequent electoral terms (2005-2012) we find an effect close to zero. Large effects are estimated again for the period 2013-2015.

This raises the question: what could explain the results not extending beyond 2005 but then reappearing in 2013? We consider three hypotheses. First, newly elected mayors in 2000 could have been particularly honest due to a change in election legislation increasing political competition in the 2000 election. Second, mayors may have started entering state or national politics at a higher rate and put more weight on their reputation even when they were no longer eligible to be reelected as mayors. However, the empirical evidence does not support either of these two hypotheses.

Another possibility is that something made first-term mayors more corrupt over time, compared to second-term mayors. For example, a new political party that grows in strength, recruits a large number of new mayoral candidates, and therefore reduces the screening of candidates could explain this pattern. To address this hypothesis, we investigate whether the rise of the Workers' Party (PT) in the 2004 and 2008 elections played a role. While we observe that first-term mayors are generally less corrupt, we find that first-term mayors from the PT party between 2005 and 2012 were more corrupt than other first-term mayors. This suggests that the rise of PT and higher corruption levels among their first-term mayors partly explain the decrease in the difference between first and second-term mayors during the 2005-2012 period. However, this can only partially explain the differences over time, and the PT party's first-term mayors are only statistically significantly more corrupt in the data encoded by Brollo et al. (2013).

A further possibility, suggested by the results in Avis et al. (2018), is that

[3.3]For one of the three outcomes we can reject the point estimate from the original paper at the 95% confidence level.

legal actions against corrupt public officials may have decreased the importance of reelection incentives in curbing corruption. Such legal actions increased substantially from 2004 to 2012. As both first- and second-term mayors are subject to these legal actions, this may have decreased the difference in corrupt behaviors between the two groups.

Our paper contributes to two strands within the literature on corruption. First, we add to the efforts to measure corruption in general (Olken, 2007), and particularly through random audits. The literature up to now has employed manual methods to encode audit reports (Ferraz and Finan, 2008, 2011; Brollo et al., 2013; Avis et al., 2018; Colonnelli and Prem, 2022).[3.4] We introduce a novel approach of employing an LLM to encode audit reports. Our methodology is automatic, cheap, scalable, and broadly applicable in other settings.

Second, we contribute to the literature on the political determinants of corruption. Ferraz and Finan (2011) argue that mayors with reelection incentives misappropriate 27 percent fewer resources than mayors without reelection incentives. Bobonis et al. (2016) show that corruption is lower in municipalities with audits right before elections, but find no sustained effects in subsequent audits.[3.5] We highlight how different forces determining measured corruption may gain importance as times change, potentially masking the effects of electoral incentives.[3.6]

The rest of the paper proceeds as follows. Section 3.2 provides background information on Brazil's anti-corruption program and reviews data from previous encodings of audit reports. Section 3.3 presents the methodology used to construct our corruption measures and discusses inherent challenges. Section 3.4 compares our measures with previous manual classifications and present key summary statistics. Section 3.5 presents the identification strategy for the

[3.4]A recent literature employs automated methods to detect corruption *out-of-sample*. Ash et al. (2020) applied machine learning methods to detect corruption using municipal budgets and outcomes from Brollo et al. (2013). They found that compared to random audits, a machine-guided targeted policy could detect almost twice as many corrupt municipalities for the same audit rate. Colonnelli et al. (2022) used data based on CGU encoding of audits reports plus municipal characteristics, including financial development, human capital, and local politics, to predict corruption. The authors concluded that measures of private sector activity, financial development, and human capital are the strongest predictors of corruption, while public sector and political features play a secondary role. Instead, our goal is to create new *ground truth* data.

[3.5]A related literature studies the direct effects of corruption audits and information transparency more broadly. Ferraz and Finan (2008) found that exposing audit results significantly impacted the electoral performance of incumbent mayors in 2004, leading to a 17% reduction in the likelihood of reelection for candidates with higher indications of corruption. Avis et al. (2018) showed that being audited increased the likelihood of subsequent legal action by 20 percent and reduced later corruption by 8 percent.

[3.6]For evidence on the policy consequences of reelection incentives, see Besley and Case (1995), Besley and Case (2003), Alt et al. (2011), and List and Sturm (2006).

empirical exercise, as well as the main findings and alternative explanations for varying effects over time. Section 3.6 concludes with our final remarks.

3.2 Background

3.2.1 The Random Audits Anti-Corruption Program

In 2003 the Brazilian federal government created the *Controladoria Geral da União* (CGU), tasked with promoting transparency, preventing corruption, and enforcing integrity in public administration. As the primary oversight body, the agency is responsible for monitoring and auditing the utilization of public funds across various Government agencies.

An important initiative introduced shortly after the CGU's creation was the *Programa de Fiscalização por Sorteios Públicos*. This initiative involved randomly selecting municipalities with populations under 500 thousand inhabitants to audit their use of federal funds. Over 13 years, the program conducted 40 lotteries and 2,199 audits across 1,910 municipalities. After 2015, the program was reformulated to include both random and non-random audits. For this reason, we restrict our analysis to the first 40 lottery rounds held between 2003-2015.

Once a municipality was selected to be audited, the CGU gathered information on all federal funds transferred to the municipal Government in that political term and in some cases in the previous term. CGU auditors were then sent to the municipality to examine accounts and documents, as well as to evaluate the existence and quality of public infrastructure projects and the provision of public services.

The detailed inspections conducted by CGU auditors resulted in comprehensive reports detailing the extent of corruption and mismanagement. The reports range from 30 to 200 pages and are on average 85 pages long. These reports were submitted to the CGU headquarters after approximately one week of inspections and a few months later the summaries of the main findings were made available to the public online on CGU webpage.[3.7]

3.2.2 Previous Encodings of Audit Reports

Given the limited availability of data on corruption, CGU's audit reports quickly became a popular source of data on corruption. In quantitative

[3.7] The summaries can be found at https://auditoria.cgu.gov.br/

social science the audits were first used by Ferraz and Finan (2008). In this and subsequent papers, the reports were turned into quantitative data by manually encoding each report. Subsequently, a series of other papers used this classification as a basis and/or developed their own corruption classification (Ferraz and Finan, 2011; Brollo et al., 2013; Avis et al., 2018; Colonnelli and Prem, 2022; Ash et al., 2020).

To validate the corruption data encoded by LLM, we use the data from Ferraz and Finan (2011) (henceforth, FF) and Brollo et al. (2013) (henceforth, Brollo et al.).[3.8] The former manually classifies reports from 2003 to 2004 (lotteries 2 to 11), covering the period from 2001 to 2004, whereas the latter covers reports from 2003 to 2009 (lotteries 2 to 29) covering the period from 2001 to 2009. See Table 3.1 for an overview of the data used.

FF's main measure of corruption is the total amount of resources where some corruption was found, expressed as a share of the total amount of resources audited. We follow this convention and use this variable as our main benchmark when validating the encoding generated by the LLM. Additionally, we use two other variables provided by FF — a binary variable for if any corruption was found and the number of corruption cases. Similarly, Brollo et al. construct a continuous indicator, the ratio between the funds involved in the irregularities and the total amount audited, and a binary variable, whether any irregularity was found or not. The reports do not formally describe if an irregularity should be considered evidence of corruption or not. Therefore, Brollo et al. divide potential corruption cases into general (broad) irregularities, that could also be interpreted as bad administration rather than as overt corruption, and severe (narrow) irregularities, where there is clearer evidence that an act of corruption took place.

One caveat of these reports is the possibility to audit resources transferred in the preceding political term, especially when the audit was held at the beginning of the current term. For instance, an audit held in 2005 may contain audits of resources transferred to the municipality in 2004. Therefore, we exclude the first two audits in the 2005 and 2010 Mayoral terms from our analysis in Section 3.4. [3.9]

In addition to the two mentioned data sources, we also gather data from the CGU, provided under the Law on Access to Public Information (LAI). This dataset contains a list of all identified irregularities for each municipality be-

[3.8] Data from FF is available through the replication package, while data from Brollo et al. can be found on Brollo's website. We appreciate their efforts in making this data accessible.

[3.9] Brollo et al.'s classify corruption by municipality-term instead of by municipality-audit. This is important to keep in mind when comparing Brollo et al. data and data encoded by the LLM, as further explained in Appendix 3.A.3.

tween 2006 and 2015 (lotteries 20 to 40). They are classified as administrative, medium, or serious irregularities. However, even the serious irregularity definition considers a more comprehensive classification of corruption than those from FF and Brollo et al. Beyond the corruption categories considered by them, CGU also codes cases of mismanagement as serious irregularities .[3.10]

3.3 Classifying Corruption Audit Reports with LLMs

3.3.1 The LLM Framework

Despite being very useful, previous attempts to classify corruption based on audit reports span different time periods and use different manual encoding methods. Consequently, we lack a unified classification approach for all reports conducted across the 40 lotteries. This is not surprising given the number and length of the reports. Across the 13 years, 185,000 pages of reports have been published from the 2,197 audits. Analyzing all reports manually would require a substantial time investment and probably involve more than one person, potentially leading to increased variability in interpretations of what should be classified as corruption. In this Section, we present an alternative way to read these reports.

To construct our corruption measures, we employ a Retrieval Augmented Generation (RAG) process to extract pertinent contextual information from texts and then apply it to OpenAI's GPT-4, an LLM at the current publicly available frontier of the technology. When building a question-answering (QA) system without RAG, models can only draw upon data that existed when they were trained. On the other hand, with RAG, models can leverage provided context for more informed responses.

This framework works as follows. First, all PDFs are transformed into text files and split in smaller "chunks", always keeping a chunk overlap to preserve the context between two chunks. Subsequently, all pieces of text are transformed into vectors (embeddings), which are numerical representations of the text, and then stored in a vector dataset.[3.11] Texts containing semantically similar content exhibit similar vectors in the embedding space. When a question is posed, the algorithm compares the embeddings of the query with those of the chunks in the vector store, creating a similarity score between

[3.10] For instance, the lack of creation of the Municipal Commission for the Eradication of Child Labor is considered a serious irregularity. In another example, the absence of mapping/diagnosis of areas of risk and social vulnerability is considered a serious irregularity.

[3.11] We use Chroma, an open-source database, to store the embeddings (https://docs.trychroma.com).

them. Finally, the highest-scoring chunks are used as context to generate the responses.

To carry out this entire process we rely on LangChain — an open-source framework for building LLM applications.[3.12] The operational mechanism underlying this process can be summarized in the following four steps:

1. Write the question

2. Transform the question into embeddings

3. Compare the embeddings of the question with all vectors stored in the vector dataset.

4. Select the n most similar vectors [3.13]

Following the definitions of corruption from FF, we asked GPT-4 five different questions about each audit report.[3.14] The first three questions address the value found in each category of corruption: diversion of funds, overinvoicing, and procurement irregularities. The fourth question asks about the total number of cases across all three categories. All questions were asked in Portuguese, and their English translations are provided in Appendix 3.A.1. For each answer from questions 1 to 3, we extract the value associated with corruption and discard duplicate values to prevent double counting.[3.15] We then add the values from the responses to get the total corruption amount for each report.

The binary variable denoting the presence of corruption is assigned a value of one if the total corruption amount exceeds zero, and zero otherwise. The count of corruption cases is derived straightforwardly from question number four. Finally, to calculate the share of resources where some corruption was found, we ask one additional question regarding the total amount of federal funds audited by the inspectors in each audit. The response provided by this question is the denominator of our main variable - the share of audited resources where corruption was found.

$$\text{Share corrupt LLM}_{m,l} = \frac{\text{Values from Q1} + \text{Values from Q2} + \text{Values from Q3}}{\text{Value from Q5}}$$

[3.12] We used the version 0.0.349, available in January 2024. The documentation for LangChain can be found in https://python.langchain.com/docs/get_started/introduction.

[3.13] We chose the two most similar vectors.

[3.14] We set GPT-4's "temperature" to zero, ensuring direct responses and avoiding creative interpretations.

[3.15] For example, the same case may be classified as a procurement irregularity and a case of overpricing and appear in both answers.

3.3.2 Challenges in Using LLMs

Although LLMs provide a prominent framework to transform text into data, there are still some challenges with building a question-answering (QA) system, especially over large documents. The first obstacle lies in the token limits imposed on LLMs, which constrain the amount of context that can be provided. For instance, many documents exceed the capacity of 4-8k token contexts offered by GPT models. As a result, the standard practice consists of splitting the document into chunks, calculating a similarity score between those chunks and the query using embeddings, and then use the highest-scoring chunks as context for the query. This is the RAG process described in the previous section.

Another significant challenge arises from the fact that documents, such as PDFs, are naturally structured with different pages, tables, sections, and text indentation.[3.16] Therefore, there is great difficulty in creating robust prompts and chunking strategies that responds well to the variability among documents.

More specifically, building a RAG system involves determining the ideal chunk size for the documents processed by the retriever. The determination of the ideal chunk size involves considering various factors, including the characteristics of the data, the limitations of the retriever model, and the computational resources available (Farenas, 2024). In addition to size, splitting the document into chunks entails additional decisions. For instance, determining the degree of overlap between them and selecting the number of top scoring chunks to be considered by the LLM.

The study by Liu et al. (2024) offers insights on how well LLMs use longer context. The research evaluated a range of open and closed-source language models, including OpenAI's GPT-3.5-Turbo, across two distinct tasks: multi-document question answering and key-value retrieval. They observe that optimal performance is often achieved when relevant information is located at the beginning or end of the input context. However, performance significantly degrades when models need to access relevant information embedded within the middle of long contexts. This suggests that current language models struggle to leverage information within extended input contexts.

In order to improve the output generated by the model, we conduct a series of manual verifications following a set of rules. We elaborate on these manual checks in the Appendix 3.A.2, along with illustrative examples.

[3.16] This is very clear in the case of audits reports. Over the years and across municipalities, there is a lot of variation in how the reports are structured.

3.4 Comparing LLM and Manual Classifications

In this Section, we discuss in more detail our encoding of the audit reports and how it compares to three data sources presented in Section 3.2.2. Table 3.1 offers a comprehensive comparison of the coverage provided by the LLM data in contrast to the other datasets. Our LLM encoded dataset is the only one covering the complete period from 2003 to 2015. The other datasets cover only partial periods, limiting the scope of the comparison. Additionally, it is worth noting that the frequency of draws was higher initially, with seven lotteries occurring in the first year, while in the last three years, only one lottery took place annually.

As in FF, we define corruption as any irregularity associated with diversion of funds, over-invoicing of goods and services, or illegal procurement practices. The primary variable of interest, the "share of corruption", represents the aggregate value of resources where at least one of these irregularities were identified, divided by the total audited amount. Besides being compared to FF, this variable can also be compared to Brollo et al.'s measures. Unfortunately, the Government data lacks information regarding the value associated with each irregularity. Consequently, we are unable to construct the share of corruption using this dataset. Nonetheless, we can analyze the frequency of irregularities per lottery. Then, the number of irregularities related to corruption is comparable to both FF and Government data but not to Brollo et al.'s data.

Table 3.1: Summary Data

	FF	Brollo	Gov	LLM
Lotteries 2-11 (2003-2004)	✓	✓		✓
Lotteries 12-20 (2004-2006)		✓		✓
Lotteries 20-29 (2006-2009)		✓	✓	✓
Lotteries 30-40 (2009-2015)			✓	✓
Corruption amount as a share of audited resources	✓	✓		✓
Number of irregularities related to corruption	✓		✓	✓
Dummy whether any irregularity was found	✓	✓	✓	✓

Note: This table provides an overview of the coverage of various sources of corruption data utilized in this paper. Additionally, the table displays the availability of each measure of corruption. For each range of available lotteries and variables, we mark them with a check. Variables from Brollo et al. (2013) are divided into broad and narrow irregularities as explained in Section 3.2.2.

Table 3.2 presents the correlation matrix between the LLM encoded measures and each comparable variable. Our analysis reveals a positive correlation of 0.47 between the share of corruption from LLM and the share of corruption from FF. The correlation with Brollo et al.'s narrow and broad definitions are 0.45 and 0.44 , respectively. This suggests that there is consistency between the three datasets. The correlation between the variable derived by the LLM and the manually coded datasets is similar to the lower bound correlation observed among the manually coded data, which is 0.48 , although smaller than the upper bound correlation of 0.71 . The fact that the manually coded data has correlations far below one, show that there is a substantial amount of subjective evaluation in the reading of the audit reports. This highlights the difficulty inherent in classifying corruption within these reports.

Table 3.2: Data Correlations

	Share corrupt FF	Share broad	Share narrow	Share corrupt LLM	Any corruption FF	Any broad corruption	Any narrow corruption	Any corruption LLM	Some serious irregularity
Share corrupt FF	1.00	0.71	0.48	0.47	.	.	.	.	.
Share broad	0.71	1.00	0.65	0.44	.	.	.	.	.
Share narrow	0.48	0.57	1.00	0.45	.	.	.	.	.
Share corrupt LLM	0.47	0.44	0.45	1.00	.	.	.	.	.
Any corruption FF	.	.	.	.	1.00	0.59	0.31	0.35	.
Any broad corruption	.	.	.	.	0.59	1.00	0.57	0.25	0.31
Any narrow corruption	.	.	.	.	0.31	0.57	1.00	0.20	0.29
Any corruption LLM	.	.	.	.	0.35	0.25	0.20	1.00	0.07
Some serious irregularity	.	.	.	.	.	0.31	0.29	0.07	1.00

Note: This tables presents the pair-wise correlations between the LLM variables and manually coded variables from Ferraz and Finan (2011), Brollo et al. (2013), and Government. Additionally, we present the pair-wise correlations between all variables used as reference. These variables span different periods, as illustrated in Table 3.1. For that reason, we can not calculate the correlation between "Some serious irregularity" and "Any corruption FF".

We present the summary statistics for LLM's variables in Table 3.3. Overall, the summary statistics of the LLM encoded data is similar to the summary statistics of the manually encoded data. According to the LLM encoding, 77% of the reports have at least one case of corruption. When considering only the first 11 lotteries, this percentage is 70%, which is in between the 79% reported by FF and the 67% reported by Brollo et al.'s measure of broad corruption. Comparing our findings to all the audits analyzed by Brollo et al., we observe a similar percentage to that of broad corruption (75% versus 78%). Additionally, when compared to Government data, the percentage is slightly higher (82% compared to 78%).

Moreover, we find that, on average, 2.3%-3.4% of the audited resources were subject to diversion, overpricing, or procurement fraud, depending on the time period considered. The data encoding by FF finds that 6.3% of resources were subject to corruption in the 2001-2004 time period while Brollo et al.'s broad measure find a share of 5.2% for the 2001-2009 time period, while their narrow measure is 2.1% for the same time period. Regardless of what time period is considered, the percentage in our LLM-encoded data is always within the range of the percentages in the manually encoded datasets.

Another attempt to gauge the intensity of corruption is through the count of corruption cases. Among the manually encoded datasets, this variable is available in the FF and Government encoded data. In our dataset, this variable is derived from an independent inquiry to GPT-4, as detailed in the Appendix 3.A.1. We found that the average number of cases per report is 1.21 from 2001 to 2015. Comparing equivalent audits, the mean is lower than that reported both by FF (0.96 versus 1.93) and Government data (1.32 versus 7.07). In the last case, the high number of irregularities confirms a point discussed in Section 3.2.2: the classification of serious occurrences is quite broad, sometimes categorizing cases of mismanagement as serious irregularities.

Table 3.3: Summary Statistics

	2001-2015	2001-2004				2001-2009			2005-2015	
	LLM	LLM	FF	Broad	Narrow	LLM	Broad	Narrow	LLM	Gov.
Any corruption	0.765	0.694	0.786	0.675	0.389	0.744	0.783	0.470	0.814	0.783
	(0.424)	(0.461)	(0.411)	(0.469)	(0.488)	(0.436)	(0.412)	(0.499)	(0.389)	(0.412)
Observations	2,128	490	476	489	489	1,537	1,401	1,401	1,168	1,112
Share corrupt	0.027	0.034	0.063	0.059	0.026	0.029	0.052	0.021	0.023	
	(0.061)	(0.075)	(0.102)	(0.116)	(0.069)	(0.065)	(0.102)	(0.064)	(0.050)	
Observations	2,127	490	476	483	483	1,536	1,336	1,337	1,167	
Number of cases	1.209	0.957	1.931			1.283			1.318	7.068
	(2.258)	(1.998)	(1.707)			(2.395)			(2.410)	(9.377)
Observations	1,823	444	476			1,309			971	1,112

Note: This table presents the mean and the standard deviation of all available variables. "Broad" and "Narrow" refer to definitions from Brollo et al. The first column span all years, including all available audits (lotteries 2 to 40). The other columns are divided into different time periods to align with manually coded versions. From 2001 to 2004, we report the statistics from audits conducted within the first 11 lotteries, comparing these to the findings of both FF and Brollo et al. From 2001 to 2009, we report the statistics from audits conducted within the first 29 lotteries and compare it to Brollo et al.'s findings. Finally, from 2005 to 2015 we report the statistics from audits conducted from lottery 20 onwards, comparing these to CGU variables. The reduced number of observations in the "Number of cases" from the LLM is due to missing data—cases where the algorithm does not provide an exact number. Lotteries held in the first six months of each term were excluded because most of the audited resources refers to the preceding political term. For Brollo et al. and Government variables, we restrict the analysis to observations within the same term, excluding corruption associated to resources transferred in previous political terms. Additionally, the data is restricted to municipalities where we have information on whether the mayor is serving their first or second term.

3.5 Reelection Incentives and Corruption

3.5.1 Empirical Strategy

In this Section, we apply our extended corruption data to reassess a key finding in the literature, which is that mayors facing reelection are less corrupt than those not eligible for reelection. Following Ferraz and Finan (2011), we test whether reelection incentives affect the level of political corruption in a municipality using the following OLS regression:

$$Corruption_{ml} = \alpha_s + \alpha_l + \beta FirstTerm_{ml} + \gamma Z_{ml} + \varepsilon_{ml} \tag{3.1}$$

where $Corruption_{ml}$ is a measure of corruption in municipality m as reported in an audit from lottery l. $FirstTerm_{ml}$ indicates whether the mayor is the first term, while Z_{ml} represents a set of controls accounting for the mayor's observable characteristics. The terms α_s and α_l respectively denote state and lottery fixed effect, and ε_{ml} is the error term.

We also estimate a close election regression to account for any unobserved municipal determinants of corruption that may differ between first and second-term mayors. We compare municipalities where incumbent mayors barely won the election, thus serving as second-term mayors in the following term, to municipalities where the incumbent barely lost the election and thus was replaced by a new mayor. This setting provides a quasi-random assignment of municipalities with a first- versus second-term mayor. The main hypothesis supporting the use of the RD is that if an election is competitive enough, who wins it is as good as random.

To estimate the effect of reelection on corruption, we subset the data, including only mayors associated with a vote margin whose absolute value is sufficiently close to zero. The optimal distance to use as bandwidth is defined according to the minimum squared error (MSE) criteria (Calonico et al., 2014b). The following local linear regression is then used:

$$Corruption_{mt} = \alpha_s+\alpha_l+\tau FirstTerm_{mt}+\lambda_0 MV_{mt}+\lambda_1 FirstTerm_{mt}MV_{mt}+\gamma Z_{mt}+\varepsilon_{mt} \tag{3.2}$$

where $Corruption_{mt}$ is the corruption outcome, $FirstTerm_{mt}$ is the indicator for first-term mayors, and Z_{mt} is a vector of mayors characteristics, as before. The terms α_s and α_l, as before, respectively denotes state and lottery dummies. The term MV_{mt} represents the candidate's margin of victory in municipality m for the election corresponding to political term t. It is specified as the difference between the vote share of incumbent mayor minus the vote share of the challenger receiving the largest number of votes. This measure is therefore less than zero in municipalities where the incumbent was not reelected and a new mayor was elected, and greater than zero otherwise.

3.5.2 Results

The main finding from FF's paper is that mayors with reelection incentives misappropriate 27 percent fewer resources than those without reelection incentives. We investigate if this result extends to later years. Table 3.1 and Figure 3.1 present the estimates from the OLS regression specified in Equation 3.1, spanning the complete period from 2001 to 2015. This period encompasses resources from four political terms audited along thirty-five lotteries. The pilot lottery, which audited only five municipalities, was excluded as in FF. Additionally, we excluded four lotteries conducted within the first six months of each political term, as documented in the Appendix 3.A.3. All estimates include controls for mayoral characteristics, such as education level, gender, and age, as well as indicators for lotteries and state, accounting for any state-specific or lottery-specific unobservables that might have affected political corruption.

Table 3.1 shows the first set of results using LLM data. Our preferred specification is presented in the even columns. Column 2 suggests that municipalities where mayors are eligible for reelection exhibit a 4.8 percentage point decrease in the likelihood of having a case of corruption detected when compared to municipalities with mayors in their second term. Surprisingly, this effect is not observed in either on the number of cases or the share of corruption, the main outcome presented by FF. Our estimates show negative but statistically insignificant results.

In Figure 3.1 we visualize a comparison of the estimated effect size using all four encodings of the data. We conducted the regression described in Equation 3.1 on a variable that is normalized so that the coefficients can be interpreted as a percentage change over the variable mean. The FF data spans the 2001-2004 time period, the Brollo et. al. data spans the 2001-2009 period, and the Government data spans the 2005-2015 period. Only the LLM variables span the complete 2001-2015 period. The figure shows that the coefficients from the LLM's encoding of the number of corruption cases and the indicator for if any corruption was found are similar in direction and magnitude to the coefficients generated using FF's encoding of the data. When using FF's encoding of the data we successfully replicate the findings of their paper. However, the coefficient of "Share corrupt", is considerably smaller and statistically significantly different from the original FF estimate.

In addition, the coefficients from the Government's and Brollo et al.'s data deviate significantly from those estimated using FF data, with the exception of the variable "Any narrow corruption" which is the only variable for which we find a negative and statistically significant effect in these two encodings of the data.

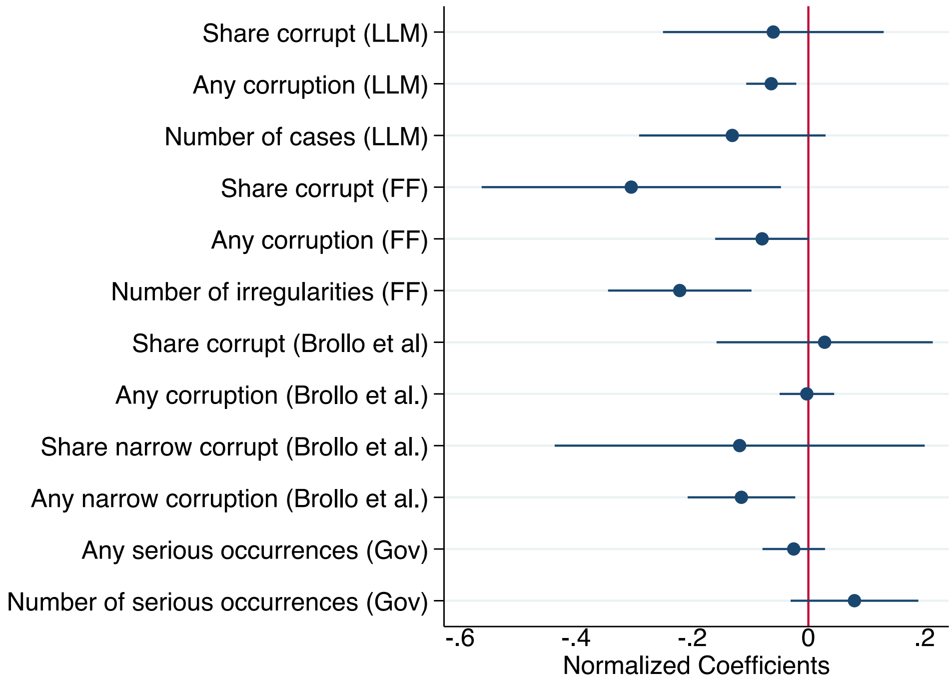
As a robustness test, we conduct an additional exercise. We estimate the same regressions restricting the samples to include only reelected mayors. As pointed out by FF, if elections serve to select the most able politicians, and ability and corruption are positively correlated, we need to compare second-term mayors with the set of first-term mayors who are reelected in the subsequent election — those presumed to have greater political skills. The normalized coefficients are displayed in Figure 3.A.2. Overall, the coefficients do not exhibit substantial differences from those presented in Figure 3.1.

Table 3.1: The Effect of Reelection Incentives on Corruption (2001-2015)

	Any corruption LLM		Share corrupt LLM		Number of cases LLM	
	(1)	(2)	(3)	(4)	(5)	(6)
Mayor in first term	-0.0356*	-0.0493**	-0.0024	-0.0016	-0.0282	-0.1554
	(0.0200)	(0.0202)	(0.0030)	(0.0031)	(0.1182)	(0.1158)
Constant	0.7919***	0.4338	0.0285***	0.0058	1.1957***	1.9030**
	(0.0158)	(0.5231)	(0.0024)	(0.0293)	(0.0980)	(0.8080)
Mayor Characteristics	No	Yes	No	Yes	No	Yes
Lottery Dummies	No	Yes	No	Yes	No	Yes
State Dummies	No	Yes	No	Yes	No	Yes
Party Dummies	No	Yes	No	Yes	No	Yes
r2	0.0016	0.1406	0.0003	0.0729	0.0000	0.1285
N	1,894	1,885	1,893	1,884	1,626	1,620

Note: This table presents the impact of reelection incentives on three corruption metrics: the probability of finding a corruption case, the proportion of audited resources associated with corruption, and the number of detected corruption cases. Each column displays the results from the OLS regression presented in Equation 3.1, where the respective corruption measure is regressed on an indicator variable denoting whether the mayor is in their first term. The even numbered columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. This estimate includes municipalities audited from lotteries 2 to 40, with the exception of lotteries 15, 16, 28, and 38 —excluded due to their occurrence within the first six months of a political term (See Section 3.A.3 for further explanation). The period from 2001 to 2015 spans the four political terms with audited resources. The last term, however, does not consider the final year (2016), as the last available lottery (40) was conducted in 2015. Robust standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

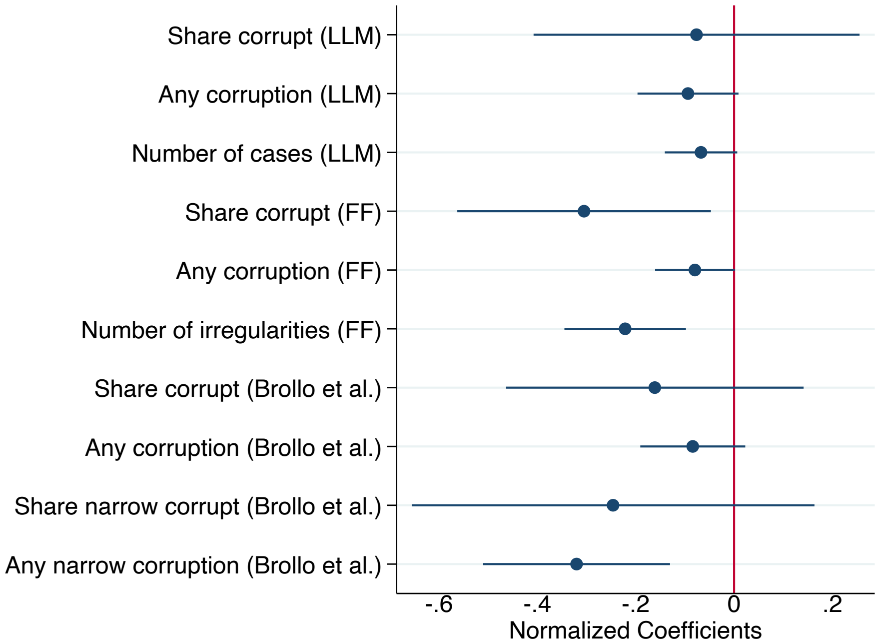
Figure 3.1: Overall Effects of Reelection Incentives on Corruption (All Available Years)



Note: This figure depicts the coefficients from the OLS regression outlined in Equation 3.1. All coefficients are normalized by the mean. We estimate the impact of reelection incentives on corruption using all available measures obtained from LLM, FF and Brollo et al.'s dataset. All regressions include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. For each variable, we plot its source in parentheses. The regressions include different time coverage. Estimates using LLM variables includes data from 2001 to 2015 (lotteries 2 to 40). We exclude lotteries 15, 16, 28, and 38 due to their occurrence within the first six months of a political term (See Section 3.A.3 for further explanation). Brollo et al.'s variables includes data from 2001 to 2009 (lotteries 2-29), while Government variables includes data from 2005 to 2015 (lotteries 20 to 40). For both Brollo and Government data, the analysis is restricted to observations within the same term. The coefficients from FF are the same to Figure 3.2. Confidence intervals are displayed at the 90% level.

We then narrow our analysis to examine the impact of reelection incentives on corruption, focusing exclusively on audits conducted between 2003 and 2004. In that case, we investigate whether the estimates using alternative measures of corruption generate results in line with FF's original findings for the 2001-2004 term. Figure 3.2 presents the normalized coefficients restricted to that period and shows that we successfully replicate the primary finding presented in FF when using the same audit reports. While certain variables lack statistical significance, all coefficients point in the expected direction, indicating that first-term mayors were indeed less corrupt during the 2001 term. Thus, the differences in the results for different encoding of the data presented in Figure 3.1 can to a large extent be explained by differences in the estimated effect across different time periods.

Figure 3.2: The Effect of Reelection Incentives on Corruption (2001-2004)

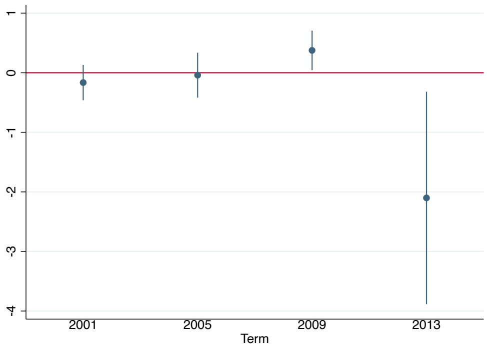


Note: This figure depicts the coefficients from the OLS regression outlined in Equation 3.1. All coefficients are normalized by the mean. We estimate the impact of reelection incentives on corruption using all available measures obtained from LLM, FF and Brollo et al.'s dataset. The regressions are restricted to lotteries 2 to 11, thus including only the political term from 2001 to 2004, the first in which reelection was allowed at the municipal level. All regressions include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. For each variable, we plot its source in parentheses. The Government variables are not included in this figure due to the absence of data prior to lottery 20 (See Table 3.1 for detailed information on data coverage).Confidence intervals are displayed at the 90% level.

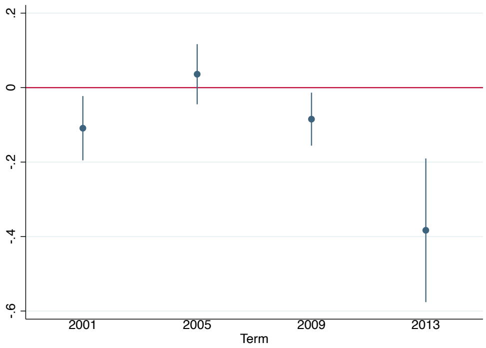
Figure 3.3 presents evidence on how the magnitude of the difference between first- and second-term mayors change over time. During the 2001 term, two of the three LLM encoded variables indicate that first-term mayors were associated with less corruption than second-term mayors. This effect is also apparent in the 2013 term. However, we observe no differential effect between first and second-term mayors from 2005 to 2012. A similar pattern is observed for Brollo et al.'s corruption measures, as shown in Figure 3.A.1 in the Appendix.

Existing literature indicates that the effects of term limits can differ across periods. Building on their 1995 findings on the fiscal impact of gubernatorial term limits in the U.S., Besley and Case (2003) reveal that these effects have shifted significantly over time. Initially, they found that governors tended to spend and tax more when they could not stand for reelection. However, with data extended to the mid-1990s, this effect weakened and even reversed. Alt et al. (2011) explain this shift as a "competence effect", emerging from changes in the structure of term limits across states, as many states moved from single-term to two-term limits, allowing voters to retain more competent incumbents. In Section 3.5.3, we discuss and test alternative explanations for the observed changes over time in our context.

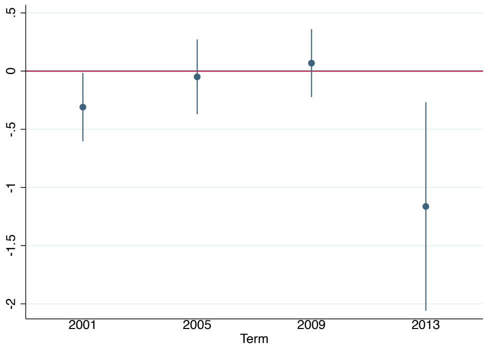
Figure 3.3: The Effect of Reelection Incentives on Corruption Over Time (LLM)



A: Share corruption



B: Any corruption



C: Number of cases

Note: This figure presents the coefficients from the OLS regression specified in Equation 3.1. All regressions include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. The results are broken down by term. Estimates includes data from 2001 to 2015 (lotteries 2 to 40). We exclude lotteries 15, 16, 28, and 38 due to their occurrence within the first six months of a political term (See Section 3.A.3 for further explanation). Each term spans a four-year period. The last term, however, does not consider the final year (2016), as the last available lottery (40) was conducted in 2015. Confidence intervals are displayed at the 90% level.

Finally, we assess the effects of reelection incentives using elections in which the incumbents won or lost by a narrow margin. The RD outlined in Equation 3.2 provides quasi-random assignment of first-term and second-term mayors across these competitive elections, eliminating potential confounds. The sample is conditioned on the incumbents who ran for reelection in each election. Table 3.2 presents the point estimates for the LLM's measures spanning all terms. All columns are estimated using the MSE optimal bandwidth. In the Appendix 3.A.4, we also provide robustness using half and double the optimal bandwidth. All specifications include the controls used previously: mayor's characteristics, party affiliation, state, and lottery fixed effects.

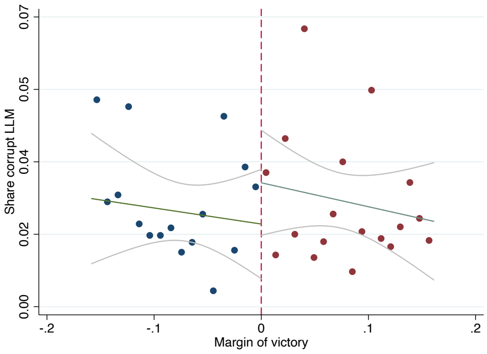
The coefficients estimated using the RD specification are all negative but small and statistically in indistinguishable from zero. However, results are also statistically indistinguishable from the main results presented in Table 3.1. In Figure 3.4, we depict the results graphically. Similar results, using the data encodings by FF and Brollo et. al., are shown in Appendix Table 3.A.4 and Appendix Figures 3.A.3 and 3.A.4.

Table 3.2: The Impact of Reelection Incentives on Corruption, RD

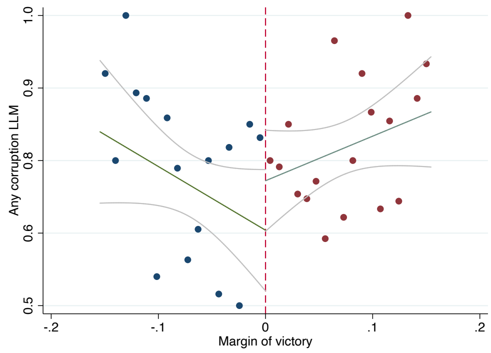
	Share corrupt (1)	Any corruption (2)	Number of cases (3)
Mayor in first term	-0.005 (0.012)	-0.060 (0.077)	-0.087 (0.332)
Robust 90% CI	[-.025 ; .034]	[-.222 ; .221]	[-1.066 ; .875]
Kernel Type	Triangular	Triangular	Triangular
BW Type	CCT	CCT	CCT
BW	0.161	0.155	0.226
Observations	1026	1027	878

Note: This table presents the coefficients from the RD regression specified in Equation 3.2. We evaluate the impact of reelection incentives on three corruption metrics: the probability of finding a corruption case, the proportion of audited resources associated with corruption, and the number of detected corruption cases. All columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. We include municipalities audited from lotteries 2 to 40 if the mayor ran for reelection. As in the previous cases, we excluded lotteries 15, 16, 28, and 38 due to their occurrence within the first six months of a political term (See Section 3.A.3 for further explanation). The BW Type indicates that the MSE optimal bandwidth was used (CCT). The BW parameter reports the respective bandwidth for each regression. Standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

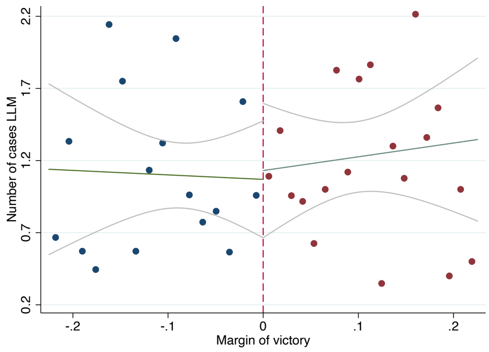
Figure 3.4: The Effect of Reelection Incentives on Corruption (LLM), RD



A: Share corruption



B: Any corruption



C: Number of cases

Note: The figure shows the proportion of audited resources associated with corruption (Panel A), the indicator for detected corruption (Panel B) and the number of detected corruption cases (Panel C) by the margin of victory for incumbents who ran for reelection. The grey lines denote the confidence intervals for fitted lines at the 90% level. All regressions use the optimal bandwidth according to the minimum squared error (MSE) criteria (Calonico et al., 2014b). We restrict the observations, such that only mayors associated with a vote margin within the interval of the optimal bandwidths are considered.

3.5.3 Testing for Alternative Explanations

What factors might explain the large negative effects during the 2001 and 2013 mayoral terms, and the absence of an effect during the 2005 and 2009 mayoral terms, as shown in Figure 3.3? We test and discuss potential explanations that could be acting to change the results in the two terms from 2005 to 2012. We start by leveraging two hypotheses that may have contributed to making first-term mayors more susceptible to corruption over time, compared to second-term mayors.

One hypothesis follows from a change in cohorts. The Random Audits Program started in 2003, thus auditing resources from the political term that began in 2001.[3.17] Coincidentally, this cohort was the first generation of second-

[3.17] There are few cases related to the 1997 term, according to Ferraz and Finan (2011).

term mayors, as the 2000 election was the first to allow reelection at the municipal level. Naturally, barring any irregularities with the electoral court, all mayors who held office from 1997 to 2000 were eligible to run again. This means that newly elected politicians faced a high probability of running against an eligible politician, creating an environment of higher political competition. If increased competition correlates with the quality of elected officials, first-term mayors elected in 2000 may have been of higher quality and potentially less corrupt than those elected in subsequent elections. To test this hypothesis, we estimate the following equation:

$$Corruption_{ml} = \alpha_s + \alpha_l + \beta FirstTerm_{ml} + \theta(FirstTerm_{ml} \times Incumbent\ Eligible_{ml}) + \gamma Z_{ml} + \varepsilon_{ml} \tag{3.3}$$

The term of interest lies in the interaction between $FirstTerm_{ml}$ and $Incumbent\ Eligible_{ml}$. As in previous equations, $FirstTerm_{ml}$ is an indicator variable for whether the mayor in their first-term in municipality m during lottery l. The term $Incumbent\ Eligible_{ml}$ is an indicator variable for if the incumbent was eligible to run again for the office when the current mayor was elected. In the case of the 2000 election, all of the incumbent mayors were eligible since they were all serving their first term. In the subsequent elections, only those who were in their first-term were eligible for reelection. The results from these regressions are shown in Table 3.3.

A second hypothesis is that a new political party grew in strength, launching a lot of new mayors, and loosening the party's screening for the quality of new mayoral candidates. To address this hypothesis, we investigate whether the rise of the Workers' Party (PT) in the 2004 and 2008 elections played a role. The performance of the PT in the 2002 national election, where they successfully elected their presidential candidate, raised its profile in local elections. The popularity shock doubled PT's number of mayors, as illustrated in Figure 3.A.5. To meet local demand for PT mayors, the party may have rushed to supply new candidates, who were potentially more corrupt.[3.18] To examine this hypothesis, we proceed to estimate the following equation:

$$C_{ml} = \alpha_s + \alpha_l + \beta FirstTerm_{ml} + \theta(FirstTerm_{ml} \times PT_{ml}) + \delta PT_{ml} + \gamma Z_{ml} + \varepsilon_{ml} \tag{3.4}$$

Here the main coefficient of interest is θ, measuring the difference in the effect on corruption of being a first-term mayor from PT and from other parties.

[3.18] Around the same time, in 2005, a major corruption scandal involving PT in Congress came to public attention. The *Mensalão* scandal revealed a systemic corruption scheme involving bribery payments to congressmen in exchange for political support. This scandal significantly diminished the party's reputation and resulted in the conviction of several officials.

The term PT_{ml} is an indicator variable, assigned the value one if the mayor was from the Workers' Party in municipality m during lottery l. The results from these regressions are presented in Table 3.3. This table also displays the results of a joint test considering both interactions.

In Table 3.3, we report the results for the "Any corruption LLM" dependent variable, which exhibited significant results in the OLS regressions. Columns 1 to 3 of this table display the results from Equation 3.1, separated by each period. As discussed in the preceding section, the effect is substantial in magnitude and statistically significant for the 2001-2004 term,[3.19] as well as when considering all terms (2001-2015). However, it is notably smaller and not statistically significant between 2005 and 2012.

In Columns 4 through 6 of Table 3.3, we present the results from Equation 3.3, Equation 3.4, and their combined effect for the entire period (2001-2015), respectively. Contrary to expectations, the coefficients for first-term mayors elected when the incumbent was eligible are positive, albeit close to zero, and the coefficients for first-term mayors from the PT are negative. If our hypothesis holds, we should expect observing first-term mayors being less corrupt when the incumbent was eligible and more corrupt when affiliated with the Workers' Party.

Considering that the party's growth primarily occurred during the 2004 and 2008 elections, we restrict the sample to the 2005-2012 period in Columns 7 and 8. In this case, the coefficients for PT first-term mayors are positive and of considerable magnitude, although not statistically significant. In the Appendix, we present similar regressions results using Brollo et al.'s "Any narrow corruption" measure (Table 3.A.5). Coefficients for PT first-term mayors are substantially higher and statistically significant, suggesting that changes in party composition among first-term mayors may partially explain the lack of effect between 2005-2012.

[3.19] The number of observations for the 2001-2004 term is higher than those reported in Table 3.3 when using lotteries 2-11 because there were three additional lotteries held within this term.

Table 3.3: The Impact of Changes in Cohort and Workers' Party Growth on Reelection Incentives

	Any corruption LLM							
	2001-2004	2001-2015	2005-2012	2001-2015	2001-2015	2001-2015	2005-2012	2005-2012
Mayor in first term	-0.0809**	-0.0512**	-0.0129	-0.0556**	-0.0497**	-0.0539*	-0.0195	-0.0260
	(0.0362)	(0.0201)	(0.0261)	(0.0277)	(0.0208)	(0.0283)	(0.0273)	(0.0320)
First x Incumbent eligible				0.0086		0.0084		0.0117
				(0.0270)		(0.0271)		(0.0297)
First x Workers' Party					-0.0250	-0.0258	0.0762	0.0756
					(0.0807)	(0.0810)	(0.0902)	(0.0902)
Mayor Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lottery Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Party Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.6987	0.7687	0.8086	0.7680	0.7687	0.7680	0.8086	0.8084
r2	0.1968	0.1388	0.1189	0.1405	0.1388	0.1405	0.1195	0.1195
N	667	1,894	1,118	1,884	1,894	1,884	1,118	1,117

Note: This table presents the coefficients from OLS regressions specified in Equations 3.1, 3.3, and 3.4. All columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. Columns 1 to 3 are derived from Equation 3.1 and cover different time periods: only the first political term, all terms, and middle terms. Columns 4 and 5 are derived from Equations 3.3 and 3.4, respectively, while Column 6 includes both interactions together. These regressions consider all terms, except for the final year (2016), as the last available lottery (40) was conducted in 2015. Finally, in Columns 7 and 8, we estimate Equation 3.4 and the combination of 3.3 and 3.4, respectively, only for the middle terms. Standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

So far we have raised two hypotheses related to changes in reelection incentives for first-term mayors, but it may be that something has occurred to change incentives for second-term mayors or parties. For instance, if mayors started to enter national level politics more frequently, this may have increased reelection incentives during their second term, consequently decreasing their motivation to engage in corruption. In Tables 3.A.6 and 3.A.7, we present the percentage of second-term mayors holding or running for some political office over time. The term "2 years later" refers to the first national/state election following the municipal election in which the mayor was reelected, indicating that if elected to a higher position, the mayor does not complete their term. On the other side, the term "6 years later" refers to the second national/state election following the municipal election in which the mayor was reelected.[3.20]

The Tables show that 1,5% of the 2000 second-term mayors ran for a higher position within two years of their tenure, with only 0.7% successfully elected. Additionally, 12,7% of them ran for a higher position after finishing their term, of which 4,8% were elected. However, we do not observe an increase

[3.20] Municipal and national elections in Brazil occur every four years, but they are staggered, with a two-year difference between them. Along with national elections, state elections also take place.

in second-term mayors holding or running for political office either two years or six years later. In fact, the percentage of second-term mayors running in national elections remains relatively stable between 2000 and 2004, declining in 2008, while the percentage of those elected for higher positions decreased in 2004, 2008 and 2009.

We also investigate if there was any change in reelection incentives from the perspective of political parties. One way to test this is to assess whether party turnover in subsequent elections increased significantly over time. Table 3.A.8 reveals that the percentage of second-term mayors whose parties remained in power in the subsequent election does not change between 2000 and 2004, and decreases in 2008.

Finally, alongside the various factors that could have influenced reelection incentives over the years, it is possible that our data simply contains noise. As detailed in Section 3.3.1, LLMs offer a prominent framework to transform text into data, yet building a QA system encounters challenges, particularly with large documents. In addition to their large size, CGU's audit reports exhibit different structures in the way the information is displayed over time, including variations in sections, tables, and text indentation. This variability naturally imposes challenges in creating robust prompts and chunking strategies, such as determining the optimal chunk size and overlap degree.

3.6 Conclusion

In this paper, we extend previous manual attempts and re-encode corruption audit reports using a LLM. We construct a new dataset on corruption that includes all Brazilian municipalities audited between 2003 and 2015. Our method measures corruption through random audits using a new methodology that is automatic, cost-effective, scalable, and broadly applicable to other settings. When comparing our data to existing manually encoded datasets, we find similar, albeit low, correlations between the main variables. We show that manually encoded data have correlations significantly below one, indicating a degree of subjective evaluation in interpreting the audit reports and highlighting the difficulty inherent to classifying corruption from text.

Our dataset allows us to reassess the impact of reelection incentives on corruption. We find some evidence that reelection incentives reduce corruption, corroborating FF's findings. However, the effect sizes are smaller and more precisely estimated. This reduced effect is further supported by estimates using alternative datasets, showing consistency across both OLS and RD estimates.

Our results reveal interesting time heterogeneity, with estimates resembling those of FF initially, declining in subsequent electoral terms, and increasing again in the last term. We find that the rise of the Worker's Party and higher corruption levels among their first-term mayors during the 2005-2012 period may partially explain the decrease in the difference between first and second-term mayors during that period. Alternatively, we explore two hypotheses: newly elected mayors in 2000 may have been of particularly higher quality due to high electoral competition, and mayors may have become more interested in entering state or national-level politics, valuing reputation when they were no longer eligible for reelection as mayors. However, we do not find empirical evidence supporting them. We find some evidence for the hypothesis that new first-term Workers' Party mayors are more corrupt.

This paper leaves avenues for future research open. First, we emphasize the potential of LLMs to read audit reports, thus creating more reliable data on corruption. As the LLM technology continues to advance, there is an opportunity for future studies to explore and apply more sophisticated models capable of better handling long and unstructured text, such as audit reports. Second, it is important to better understand how different forces determining corruption may evolve over time, potentially changing the effects of electoral incentives.

3.A Appendix

3.A.1 Corruption Definition and LLM Queries

Regarding the definition of corruption, both Ferraz and Finan (2008) and Brollo et al. (2013) consider corruption to be cases in which there is diversion of funds, over-invoicing of goods and services or illegal procurement practices. Specifically, diversion of resources may be any irregularity involving the embezzlement of public funds, such as resources that simply "disappear" from municipal bank accounts or incomplete service (unfinished construction, for example) and goods that were supposedly paid for but not delivered. In turn, over-invoicing are classified when there is evidence that goods and services were purchased at a value above market price. Finally, irregularities related to procurement involve any manipulation of the procurement process, simulation of the call for bids, use of fake receipts, and contracts being awarded to a friendly/politically connected firm or non-existing firms.

Following these definitions, we asked the LLM three different questions to identify the value of corruption in each category (questions 1-3). We also asked about the number of cases across all categories — question 4. Additionally, we asked a fifth question to determine the total amount of resources audited to calculate the share of corruption in each audit. All questions were asked in Portuguese, and their English translations are provided below:

1. Does the report mention cases of diversion of funds? If yes, what are the amounts diverted? End the response with: "The total diverted was" followed by the corresponding value.

2. Does the report mention cases of overpricing or excessive billing? If yes, what are the amounts? End the response with: "The total overpriced was" followed by the corresponding value.

3. Does the report mention cases of fraud or serious irregularities in procurement processes? If yes, what is the value of the fraud? End the response with: "The total fraud was" followed by the corresponding value.

4. How many cases of diversion of funds, overpricing, or fraud in the procurement process are mentioned in the report? End the response with "Total cases:" followed by the corresponding value.

5. What is the total value, in R$, of the audited resources? End the response with 'The total audited was' followed by the corresponding value.

As the information on total audited value is typically presented within the initial pages and does not require much interpretation, we employ a slightly different algorithm than the one detailed in Section 3.3.1. In order to extract the values, we simply transform the pdfs into text files and split only the first eight pages into smaller chunks. Then, we input these chunks directly into the GPT-4 prompt and ask the question.

3.A.2 Manual Verifications Based on LLM Responses

To enhance the quality of our data, we conducted several manual verifications on Share corrupt $\text{llm}_{m,l}$. We created four rules to flag the values we should check carefully:

1. The denominator falling below the 1st percentile and above the 99th percentile of the distribution.
2. The fraction falling above the 99th percentile of the distribution.
3. The fraction is equal to zero.
4. The fraction is not equal to zero but the total corruption value is less than R$ 500,00.

Overall, instances where Question 5 did not accurately capture the total audited value mainly occurred due to formatting issues. For instance, in the case of Peritiba, SC, in lottery 5, there was a table indicating the total audited resources as R$ 1,271,260.02, whereas our algorithm only captured R$ 1.27. Another example is observed in São João das Missões, MG, in lottery 2, where the returned value was 0 because this information appeared beyond the defined 8-page interval in the algorithm, within a figure on page 12. Additionally, there were cases where the algorithm failed to return the total value because it was not explicitly stated in the report. However, we managed to obtain it by summing up the audited values for each program. In total, we identified 44 observations where the values obtained via LLM fell below the 1st percentile or above the 99th percentile. We manually inspected all these observations, and 25 cases required corrections.[3.21]

Regarding the remaining rules, within 2197 observations, we have 22 cases flagged by rule number 2, 913 cases flagged by rule number 3 and 56 cases flagged by rule number 4. After we inspected the denominator using the

[3.21] We also expanded the interval to look at more cases (those above 98th percentile) and all the additional values were correctly obtained by our algorithm.

first rule, we went to analyze the numerator. Given that we have a large number of cases flagged under these rules, we developed a method to investigate them, which is completely based on LLM responses.

Our code extracts values from responses that follow the expressions 'The total diverted was', 'The total overpriced was' and 'The total fraud was'. However, there are answers in which the value appears as undefined but they indicate suspicious cases. The example below illustrates this.[3.22]

Example of LLM response

"The report mentions a case of irregularity in procurement, where medicines purchases were made through tender waiver in an amount higher than that set in Law No. 8,666/93. However, the report does not specify the exact value of this irregularity. Therefore, I cannot conclude with 'Total fraud was' followed by a value as the report does not provide that information." (Onça de Pitangui-MG, lottery 6)

Although the model does not provide a specific value, it does mention a case where the procurement law was not respected. In such cases, we conduct a search for associated keywords in the reports to retrieve the value. On the other hand, if all the answers are generic such as: "The report does not mention any case of diversion of funds. The total is R$ 0,00.", we accept the zero value. By doing this analysis, we reduced the number of observations with zero corruption from 913 to 512.

We used the same logic to investigate observations flagged by rules 2 and 4. High percentages of corruption may arise from errors in the denominator, what we deal with the investigation of very low audited resources, or from overestimation of corruption value. In these cases, we checked whether the values mentioned in the responses were consistent with the values present in the report. From the 22 observations falling above the 99th percentile, we fixed 11.

Finally, we investigated corruption values under R$ 500,00. Some of these cases were wrongly assigned by our algorithm, particularly in cases related to overpricing. In the response below, the algorithm captured the R$ 5.55, which is the number that follows the expression 'The total overpriced was'. However, this is the unitary price. In total, we fixed 44 out of 56 cases where the values were being underestimated.

Example of LLM response

[3.22] The response was translated to English.

"Yes, the report mentions a case of overpricing. The company FERCONST Material de Construção Ltda requested and was granted a "price realignment" for the 50 kg cement bag item, from R$ 16.80 to R$ 22.50. According to the current price list for the state of Minas Gerais from Revista Informador das Construções, as of January 2003, the average price surveyed was R$ 16.95. At that time, the Municipality was paying a price 32.74% higher than the average price suggested by the magazine. The total overcharged amount was R$ 5.55 per cement bag." (Recreio-MG, lottery 11)

3.A.3
Addressing Inconsistencies in Data

Corruption Measures

A difference between our corruption measures and those manually coded by FF and Brollo, as well as data provided by the Government, lies in the level of aggregation used to identify corruption. In their data, corruption is identified by lottery and political terms. The political term is given by the year in which the resources associated with the corruption were transferred, instead of the audit year.

A limitation of our algorithm is the inability to identify the year in which the resources involved in corruption were transferred. For example, if an audit from lottery 15, held in 2005, found a corruption case involving resources transferred in 2004, our model can not attribute it to either 2004 or 2005 — two different political terms. Occasionally, information about the year of resource transfer is presented as a tag named “extension of exams”. However, in some cases, this information is provided as a range, such as from January/2004 to August/2005 (See Figure 3.A.7). While precise information may occasionally be present within the text, extracting it would introduce additional complexity and noise into our model’s inquiries. In addition to asking about the corruption value we would need to ask about the year the referred resource was transferred. We opt then to identify the corruption at the level of municipality-lottery.

In order to compute the correlation between LLM and Brollo et al.’s variables, we aggregate the share of broad and narrow corruption by lottery and municipality. Similarly, for the indicator variables “Any Broad” and “Any Narrow”, we take the maximum by lottery and municipality. The same logic is applied to “Some serious irregularity”, generated from CGU data.

Regarding the data generated from the LLM, we excluded lotteries where audits occurred within the first six months of each term, as most of the resources in such cases may refer to the preceding political term. This encompasses four lotteries: 15, 16, 28, and 38. Unfortunately, we can not rule out the possibility that corruption is wrongly associated even when the audit occurs later in the term. Brollo et al.'s findings indicate that 63% of detected corruption cases occur within the same term, with the average share of corruption being higher in such instances (5.4% compared to 3.1%).

Elections Data

Regarding elections data, we find a correlation of 0.97 between our variable and FF indicator for mayors in their first term. This near-perfect correlation is slightly reduced due to the poor quality of the 1996 election data[3.23] The differences in this variable arise from two main sources: missing candidate names or identifiers in 1996, which makes it difficult to determine their status in 2000, and cases where the mayor did not complete their term (e.g., due to death) and the vice mayor took over and was subsequently re-elected.

If no information was available for 1996 but the 2000 mayor ran for re-election in 2004, we assumed they were in their first term in 2000. If no information was available for either 1996 or 2004, we conducted Google searches to verify the information. Lastly, in cases where the mayor did not complete their term, we categorized these mayors as serving a second term.

[3.23] The main source was the TSE harmonized data provided by *Base dos Dados*, but we also supplemented it with TSE original data for the 1996 election.

3.A.4
Tables and Figures

Table 3.A.1: The Effect of Reelection Incentives on Corruption using Brollo et al.'s data (2001-2009)

	Any corruption		Any narrow corruption		Share corrupt		Share narrow corrupt	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Mayor in first term	0.0119	-0.0020	-0.0437	-0.0548**	-0.0010	0.0015	-0.0053	-0.0026
	(0.0232)	(0.0224)	(0.0279)	(0.0268)	(0.0059)	(0.0061)	(0.0038)	(0.0042)
Constant	0.7753***	-0.2138	0.4980***	-0.2092	0.0531***	-0.0757	0.0247***	0.0044
	(0.0188)	(0.2025)	(0.0225)	(0.1920)	(0.0047)	(0.0490)	(0.0032)	(0.0309)
Mayor Characteristics	No	Yes	No	Yes	No	Yes	No	Yes
Lottery Dummies	No	Yes	No	Yes	No	Yes	No	Yes
State Dummies	No	Yes	No	Yes	No	Yes	No	Yes
Party Dummies	No	Yes	No	Yes	No	Yes	No	Yes
r2	0.0002	0.2480	0.0018	0.2524	0.0000	0.1522	0.0016	0.1084
N	1,401	1,392	1,401	1,392	1,336	1,327	1,337	1,328

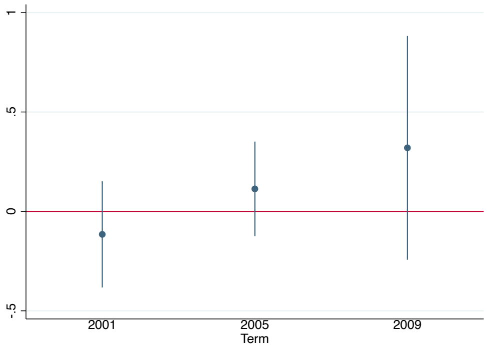
Note: This table reports the effects of reelection incentives on the probability of finding a corruption case and the share of resources found to involve corruption. Broad corruption includes irregularities that could also be interpreted as bad administration rather than as overt corruption, and narrow corruption includes severe irregularities. We regress each corruption measure on an indicator variable for whether the mayor is in his first term, as specified in Equation 3.1. The even numbered columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercept. Mayor's characteristics include the age, gender, education, and party affiliation. We restrict the analysis to observations within the same term, excluding corruption associated with resources transferred in previous political terms (See Section 3.A.3 for further explanation). The period from 2001 to 2009 indicates the years with audited resources. Robust standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

Table 3.A.2: The Effect of Reelection Incentives on Corruption using Government data (2005-2015)

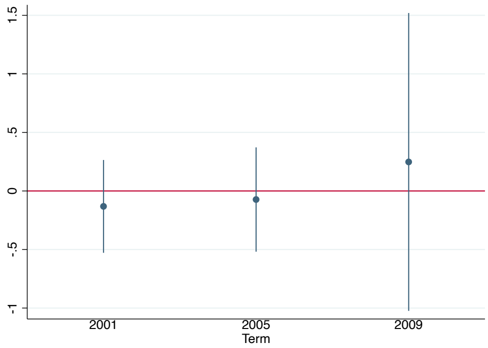
	Number of serious occurrences		Any serious occurrences	
	(1)	(2)	(3)	(4)
Mayor in first term	0.7717	0.5642	-0.0019	-0.0203
	(0.5628)	(0.4982)	(0.0261)	(0.0258)
Constant	6.5710***	20.0658	0.7855***	0.9721*
	(0.4314)	(18.2706)	(0.0213)	(0.5135)
Mayor Characteristics	No	Yes	No	Yes
Lottery Dummies	No	Yes	No	Yes
State Dummies	No	Yes	No	Yes
Pary Dummies	No	Yes	No	Yes
r2	0.0015	0.3886	0.0000	0.1870
N	1,117	1,117	1,117	1,117

Note: This table reports the effects of reelection incentives on the number of irregularities associated with corruption and on the probability of finding a serious occurrence. We regress each corruption measure on an indicator variable for whether the mayor is in his first term, as specified in Equation 3.1. The even numbered columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercept. Mayor's characteristics include the age, gender, education, and party affiliation. We restrict the analysis to observations within the same term, excluding corruption associated with resources transferred in previous political terms (See Section 3.A.3 for further explanation). The period from 2005 to 2015 spans the three political terms with audited resources. The last term, however, does not consider the final year (2016), as the last available lottery (40) was conducted in 2015. Robust standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

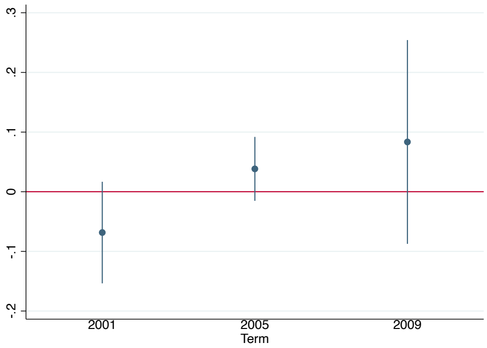
Figure 3.A.1: The Effect of Reelection Incentives on Corruption Over Time (Brollo et al.)



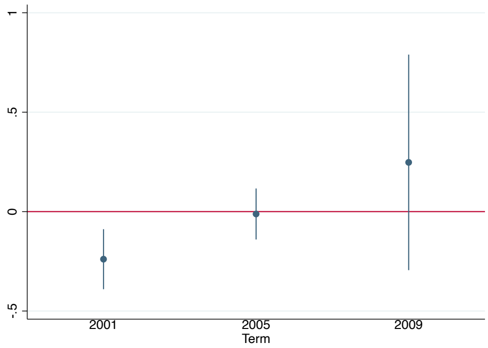
A: Share broad corruption



B: Share narrow corruption



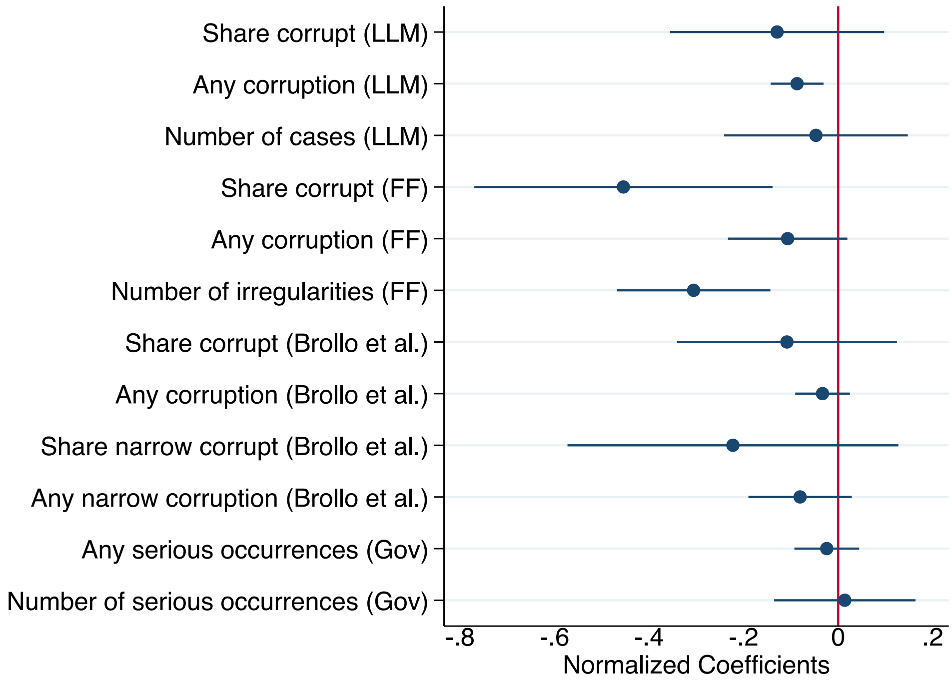
C: Any broad corruption



D: Any narrow corruption

Note: This figure presents the coefficients from the OLS regression specified in Equation 3.1. All regressions include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. The results are broken down by term. We restrict the analysis to observations within the same term, excluding corruption associated with resources transferred in previous political terms (See Section 3.A.3 for further explanation). Estimates includes data from 2001 to 2009 (lotteries 2-29). Each term spans a four-year period, with the exception of 2009, which includes only the first year due to the last available lottery (29) being conducted in that year. Confidence intervals are displayed at the 90% level.

Figure 3.A.2: Overall Effects of Reelection Incentives on Corruption
Only Reelected Mayors (All Available Years)



Note: This figure depicts the coefficients from the OLS regression outlined in Equation 3.1. All coefficients are normalized by the mean. We estimate the impact of reelection incentives on corruption using all available measures obtained from LLM, FF and Brollo et al.'s dataset. We restrict the sample to consider only reelected mayors. All regressions include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. For each variable, we plot its source in parentheses. The regressions include different time coverage. Estimates using LLM variables includes data from 2001 to 2015 (lotteries 2 to 40). We exclude lotteries 15, 16, 28, and 38 due to their occurrence within the first six months of a political term (See Section 3.A.3 for further explanation). Brollo et al.'s variables includes data from 2001 to 2009 (lotteries 2-29), while Government variables includes data from 2005 to 2015 (lotteries 20 to 40). For both Brollo and Government data, the analysis is restricted to observations within the same term. Confidence intervals are displayed at the 90% level.

Table 3.A.3: The Impact of Reelection Incentives on Corruption, RD Robustness

	Share corrupt		Any corruption		Number of cases	
	(1)	(2)	(3)	(4)	(5)	(6)
Mayor in first term	0.004	-0.005	-0.017	-0.097*	-0.066	-0.084
	(0.015)	(0.009)	(0.109)	(0.059)	(0.460)	(0.251)
Robust 90% CI	[-.035 ; .046]	[-.032 ; .018]	[-.344 ; .273]	[-.235 ; .092]	[-1.104 ; 1.237]	[-.716 ; .717]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW Type	.5CCT	2CCT	.5CCT	2CCT	.5CCT	2CCT
BW	0.081	0.322	0.077	0.310	0.113	0.451
Observations	1026	1026	1027	1027	878	878

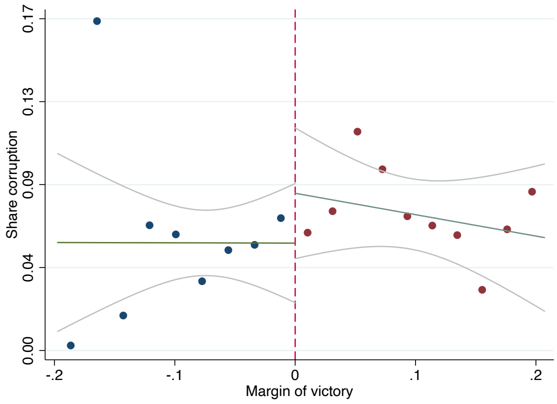
Note: This table presents the coefficients from the RD regression specified in Equation 3.2. We evaluate the impact of reelection incentives on three corruption metrics: the probability of finding a corruption case, the proportion of audited resources associated with corruption, and the number of detected corruption cases. All columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. We include municipalities audited from lotteries 2 to 40 if the mayor ran for reelection. As in the previous cases, we excluded lotteries 15, 16, 28, and 38 due to their occurrence within the first six months of a political term (See Section 3.A.3 for further explanation). The BW Type specifies whether half of the MSE optimal bandwidth (.5CCT) or twice the MSE optimal bandwidth (2CCT) was used. The BW parameter reports the respective bandwidth for each regression. Standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

Table 3.A.4: The Impact of Reelection Incentives on Corruption, RD (FF and Brollo et al.)

	Share corrupt	Any corruption	Share corrupt	Share narrow corrupt	Any corruption	Any narrow corruption
	(1)	(2)	(3)	(4)	(5)	(6)
	FF	FF	Brollo et al.	Brollo et al.	Brollo et al.	Brollo et al.
Mayor in first term	-0.023	-0.082	-0.000	0.004	0.007	-0.072
	(0.024)	(0.089)	(0.017)	(0.011)	(0.082)	(0.084)
Robust 90% CI	[-.067 ; .065]	[-.264 ; .18]	[-.038 ; .033]	[-.014 ; .029]	[-.204 ; .237]	[-.25 ; .218]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
BW Type	CCT	CCT	CCT	CCT	CCT	CCT
BW	0.207	0.232	0.171	0.186	0.160	0.223
Observations	318	318	752	753	782	782

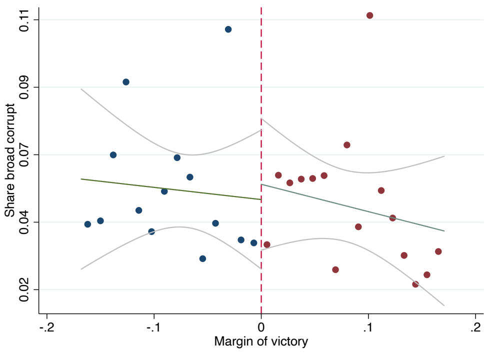
Note: This table presents the coefficients from the RD regression specified in Equation 3.2. We evaluate the impact of reelection incentives on the share of audited resources associated with corruption and on the probability of finding a corruption case. Columns 1 and 2 refer to FF measures and include data from 2001 to 2004 (lotteries 2-11), while Columns 3-6 refer to Brollo et al. measures, thus including data from 2001 to 2009 (lotteries 2-29). Our analysis is restricted to mayors who pursued reelection. All columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. The BW Type indicates that the MSE optimal bandwidth was used (CCT). The BW parameter reports the respective bandwidth for each regression. Standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

Figure 3.A.3: The Effect of Reelection Incentives on Corruption (FF)

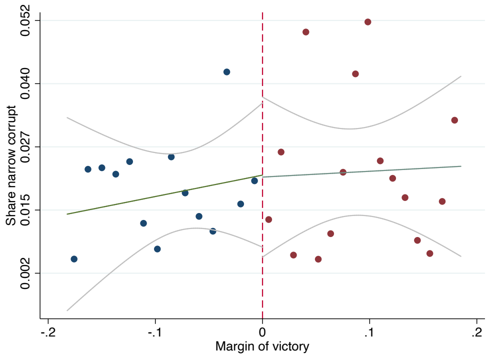


Note: The figure shows the share of audited resources involving corruption by the margin of victory for incumbents who ran for reelection in 2000. The grey lines denote the confidence intervals plotted for fitted lines at the 90% level. The regression used the optimal bandwidth according to the minimum squared error (MSE) criteria (Calonico et al., 2014b). We restrict the observations, such that only mayors associated with a vote margin within the interval of the optimal bandwidths are considered.

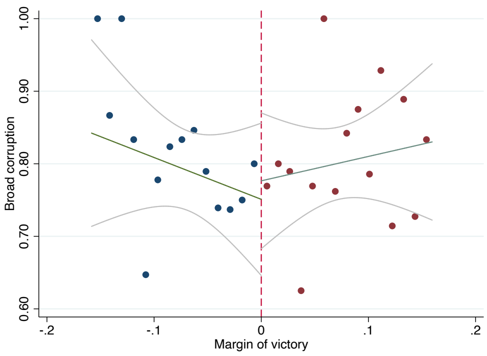
Figure 3.A.4: The Effect of Reelection Incentives on Corruption (Brollo et al.)



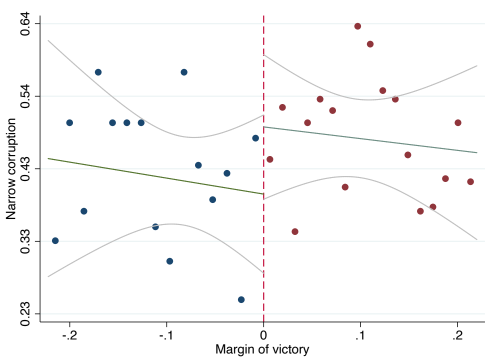
A: Share broad corruption



B: Share narrow corruption



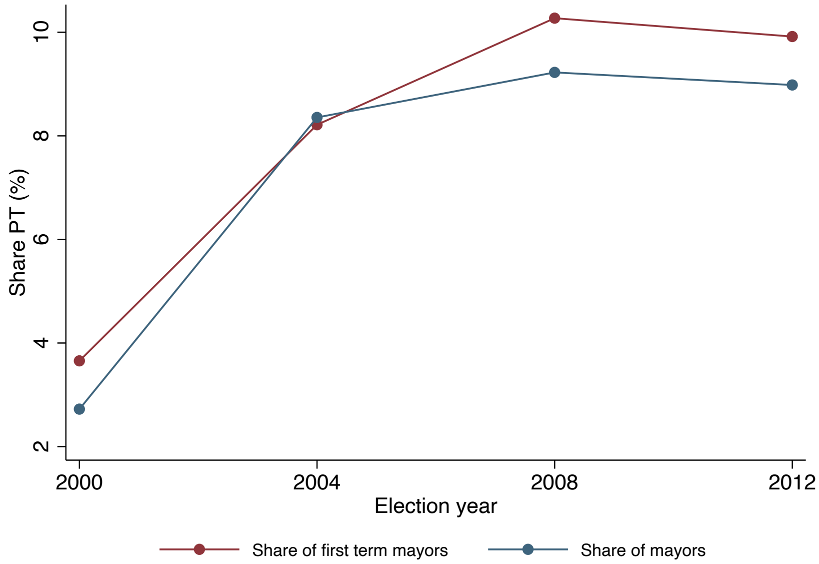
C: Any broad corruption



D: Any narrow corruption

Note:The figure shows the share of audited resources involving corruption (Panels A and B) the indicator for detected corruption (Panels C and D) by the margin of victory for incumbents who ran for reelection. The grey lines denote the confidence intervals for fitted lines at the 90% level. All regressions use the optimal bandwidth according to the minimum squared error (MSE) criteria (Calonico et al., 2014b). We restrict the observations, such that only mayors associated with a vote margin within the interval of the optimal bandwidths are considered.

Figure 3.A.5: Percentage of Worker's Party Mayors Over Time



Note: This figure depicts the percentage of elected mayors and first-term mayors affiliated with the Worker's Party over time. The x-axis represents the municipal election years.

Table 3.A.5: The Impact of Changes in Cohort and Workers' Party Growth on Reelection Incentives (Brollo et al.)

	Any narrow corruption							
	2001-2004	2001-2009	2005-2009	2001-2009	2001-2009	2001-2009	2005-2009	2005-2009
Mayor in first term	-0.0998***	-0.0540**	-0.0051	-0.0225	-0.0731***	-0.0442	-0.0408	-0.0379
	(0.0383)	(0.0266)	(0.0387)	(0.0375)	(0.0276)	(0.0381)	(0.0407)	(0.0441)
First x Incumbent eligible				-0.0454		-0.0420		-0.0155
				(0.0360)		(0.0357)		(0.0382)
First x Workers' Party					0.3533***	0.3512***	0.4006***	0.4067***
					(0.0937)	(0.0936)	(0.1099)	(0.1103)
Mayor Characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Lottery Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Party Dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mean	0.4234	0.4697	0.5108	0.4702	0.4697	0.4702	0.5108	0.5101
r2	0.2312	0.2514	0.3326	0.2552	0.2568	0.2605	0.3423	0.3485
N	659	1,401	742	1,391	1,401	1,391	742	741

Note: This table presents the coefficients from OLS regressions specified in Equations 3.1, 3.3, and 3.4. All columns include controls to Mayor's characteristics and party affiliation, as well as state and lottery intercepts. Mayor's characteristics include age, gender and education. Columns 1 to 3 are derived from Equation 3.1 and cover different time periods: the first political term, all available data, and only the middle term. Columns 4 and 5 are derived from Equations 3.3 and 3.4, respectively, while Column 6 includes both interactions together. These regressions consider all available data. Finally, in Columns 7 and 8, we estimate Equation 3.4 and the combination of 3.3 and 3.4, respectively, only for the middle term. Standard errors are displayed in parenthesis. P-values: * 0.10 ** 0.05 *** 0.01

Table 3.A.6: Mayors Holding a Political Office After Second Term

	2000 mean/sd	2004 mean/sd	2008 mean/sd	2012 mean/sd	2016 mean/sd
Second term mayors elected 2 years later	0.007	0.006	0.005	0.003	0.004
	(0.085)	(0.079)	(0.070)	(0.056)	(0.065)
Second term mayors elected 6 years later	0.048	0.033	0.030	0.023	0.043
	(0.213)	(0.180)	(0.172)	(0.150)	(0.203)
Observations	2074	1285	2037	1254	1167

Note: This table presents the percentage of second-term mayors who were subsequently elected to higher office. The years listed in columns indicate the election year when the mayor was reelected. The term “2 years later” refers to the first national/state election following the municipal election in which the mayor was reelected. The term “6 years later” refers to the second national/state election following the municipal election in which the mayor was reelected.

Table 3.A.7: Mayors Running for Political Office After Second Term

	2000 mean/sd	2004 mean/sd	2008 mean/sd	2012 mean/sd	2016 mean/sd
Second term mayors running 2 years later	0.015	0.017	0.010	0.009	0.013
	(0.123)	(0.130)	(0.101)	(0.093)	(0.113)
Second term mayors running 6 years later	0.127	0.125	0.095	0.097	0.143
	(0.333)	(0.330)	(0.294)	(0.296)	(0.350)
Observations	2074	1285	2037	1254	1167

Note: This table presents the percentage of second-term mayors running for higher office in subsequent elections. The years listed in columns indicate the election year when the mayor was reelected. The term “2 years later” refers to the first national/state election following the municipal election in which the mayor was reelected. The term “6 years later” refers to the second national/state election following the municipal election in which the mayor was reelected.

Table 3.A.8: Parties Remaining in Office

	2000 mean/sd	2004 mean/sd	2008 mean/sd	2012 mean/sd	2016 mean/sd
Percentage of second term mayors whose parties stayed in power in the next election	0.209	0.209	0.197	0.176	0.230
	(0.407)	(0.407)	(0.398)	(0.381)	(0.421)
Percentage of second term mayors whose parties returned to power two elections later	0.152	0.213	0.251	0.189	.
	(0.359)	(0.409)	(0.434)	(0.392)	(.)
Observations	2098	1294	2063	1264	1177

Note: This table presents the percentage of second-term mayors whose parties remained in power in the subsequent election or returned to power 4 years later. The years listed in the columns indicate the election year when the mayor was reelected.

Figure 3.A.6: Example of Total Audited Amount

PRESIDÊNCIA DA REPÚBLICA
CONTROLADORIA-GERAL DA UNIÃO
SECRETARIA FEDERAL DE CONTROLE INTERNO

RELATÓRIO DE FISCALIZAÇÃO Nº 08/2003

1. Trata o presente Relatório dos resultados das **84 ações de fiscalização** realizadas em decorrência do 3º sorteio do Projeto de Fiscalização a partir de Sorteios Públicos, no qual foi sorteado o Município de **Irauçuba-CE.**

2. As fiscalizações tiveram como objetivo analisar a aplicação dos recursos federais aplicados no Município sob a responsabilidade de órgãos federais, estaduais, municipais ou entidades legalmente habilitadas, bem como, avaliar a atuação dos Conselhos Municipais responsáveis pelo acompanhamento dos referidos Programas de Governo.

3. Os trabalhos foram realizados "in loco" no Município, no período de 30/06/2003 a 04/07/2003, sendo utilizados em sua execução: inspeções físicas, análises documentais, entrevistas, aplicação de questionários e registros fotográficos, em observância ao que foi estabelecido nas Ordens de Serviço expedidas pelas Coordenações-Gerais das Diretorias da Secretaria Federal de Controle Interno, responsáveis pelo acompanhamento da execução dos Programas de Governo fiscalizados.

4. Os Programas de Governo que foram objeto das ações de fiscalização, estão apresentados no quadro a seguir, por Ministério Supervisor, discriminando, a quantidade de fiscalizações realizadas e os recursos aproximados envolvidos, por Programa.

4.1 Recursos recebidos e quantidade de fiscalizações realizadas

Ministério Supervisor	Objeto Fiscalizado	Quantidade de Fiscalizações	Recursos Fiscalizados
Ministério da Fazenda	Financiamento e Equalização de Juros para a Agricultura Familiar – PRONAF	01	-
Ministério da Educação	Veículos para Transporte Escolar	01	50.000,00
	Alimentação Escolar	02	179.916,60
	Garantia de Padrão Mínimo de Qualidade para o Ensino Fundamental de Jovens e Adultos – Recomeço	01	148.270,76

Ministério Supervisor	Programa / Ação	Quantidade de Fiscalizações	Recursos Fiscalizados
Ministério de Minas e Energia	Fiscalização e Controle da Produção Mineral	02	-
	Fiscalização da Distribuição e Revenda de Derivados de Petróleo e Álcool Combutível	01	-
Ministério da Integração Nacional	Financiamento aos Setores Produtivos da Região Nordeste	01	-
	Ações Emergenciais de Defesa Civil – Bolsa Renda	01	1.260,00
	Construção de Açude Público	04	515.338,87
	Ampliação de Açude Público	04	475.802,35
	Construção de Passagem Molhada	05	572.117,37
Ministério do Desenvolvimento Agrário	PRONAF	01	-
	Investimento em Infra-Estrutura Básica para Assentamentos Rurais-Nordeste - Construção de açude	01	102.693,21
Ministério do Turismo	Desenvolvimento da Infra-Estrutura Turística na Região	02	187.500,00
Ministério do Esporte	Implantação de Núcleos de Esporte	01	105.300,00
TOTAL		**84**	**4.877.991,51**

Note: This figure shows a table detailing the total audited amount for the Municipality of Irauçuba-CE, audited during the third lottery.

Figure 3.A.7: Example of Exam Extension Information

Constatações da Fiscalização

1 – Programa: PAB - Fixo
Ação: Atendimento assistencial básico nos municípios brasileiros.
Objetivo da Ação de Governo: Ampliar o acesso da população rural e urbana à atenção básica, por meio da transferência de recursos federais, com base em um valor per capita, para a prestação da assistência básica, de caráter individual ou coletivo, para a prevenção de agravos, tratamento e reabilitação, levando em consideração as disparidades regionais.
Ordem de Serviço: 170386

Controladoria-Geral da União *Secretaria Federal de Controle Interno* 1
Missão da SFC: **"Zelar pela boa e regular aplicação dos recursos públicos."**
18º Sorteio de Unidades Municipais – Apiúna - SC

Objeto Fiscalizado: Transferência de recursos para auxiliar nas despesas na área da saúde.
Agente Executor Local: Prefeitura Municipal de Apiúna
Qualificação do Instrumento de Transferência: repasse direto à prefeitura (Fundo a Fundo)
Montante de Recursos Financeiros: R$ 183.715,00
Extensão dos exames: Analisado o total dos recursos repassados à Prefeitura Municipal no período de janeiro/2004 a agosto/2005.

Note: This figure provides an example of how information on exam extensions is presented. The example is from the Municipality of Apiúna-SC, audited during the eighteenth lottery.

Figure 3.A.8: Example of Fraud

1 - Programa/Ação: Ações Emergenciais de Defesa Civil

Objetivo do Programa/Ação: Construção de passagens molhadas.
Montante Fiscalizado: R$ 243.936,11

1.1) Constatação da Fiscalização:

Fraude em Prestação de Contas

Fato
A Prestação de Contas do convênio nº 376/2000 (SIAFI nº 400999) foi enviada ao Ministério da Integração Nacional em 26/06/2001, indicando que as obras objeto do referido convênio estavam concluídas conforme as especificações previstas. Porém, na ocasião da vistoria "in loco" constatamos que a passagem molhada de Cachoeirinha, na localidade de Mandacaru, não havia sido realizada na vigência do convênio. Verificamos que esta obra se encontrava em execução na data da visita realizada pela equipe de fiscalização. Para a execução da três passagens molhadas, objeto do convênio supracitado, foi contratada a empresa Construtora Riviera Ltda pelo valor global de R$ 119.087,00, sendo R$ 43.812,10 para a passagem de Cachoeirinha, R$ 56.300,56 para a passagem do Riacho do Missi e R$ 18.974,34 para a passagem molhada sobre o Riacho Jurema. Observamos que as notas fiscais nº 002, de 22/12/2000, no valor de R$ 47.634,80 e nº 013, de 24/01/2001, no valor de R$ 71.452,20, não continham referência ao convênio.
Recebemos denúncia de que a construção desta passagem molhada e outras construídas no município foram executadas pelo Sr. Manuel Anastácio Tabosa Braga, e não pelas empresas vencedoras dos certames licitatórios.

Evidência
Prestação de Contas e vistoria "in loco".

Note: This figure illustrates an example of a fraud case detected in the audit report for the Municipality of Irauçuba-CE, audited during the third lottery.

Bibliography

Alt, J., Bueno de Mesquita, E., and Rose, S. (2011). Disentangling accountability and competence in elections: evidence from us term limits. *The Journal of Politics*, 73(1):171–186.

Ash, E., Galletta, S., and Giommoni, T. (2020). A machine learning approach to analyzing corruption in local public finances. *Center for Law & Economics Working Paper Series*, 6.

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

Bandiera, O., Best, M. C., Khan, A. Q., and Prat, A. (2021a). The allocation of authority in organizations: A field experiment with bureaucrats. *The Quarterly Journal of Economics*, 136(4):2195–2242.

Bandiera, O., Bosio, E., Spagnolo, G., et al. (2021b). *Procurement in focus: rules, discretion, and emergencies*. CEPR Press.

Bandiera, O., Prat, A., and Valletti, T. (2009). Active and passive waste in government spending: evidence from a policy experiment. *American Economic Review*, 99(4):1278–1308.

Banerjee, A., Duflo, E., Imbert, C., Mathew, S., and Pande, R. (2020). E-governance, accountability, and leakage in public programs: Experimental evidence from a financial management reform in india. *American Economic Journal: Applied Economics*, 12(4):39–72.

Barros Barbosa, R., dos Santos Barbosa, G., Marques Nojosa, G., and Tomaz de Sousa, D. (2022). Local fiscal status and the response to the pandemic of covid-19: Evidence for brazilian municipalities. *Cadernos de Finanças Públicas*, 22(1).

Besley, T. and Case, A. (2003). Political institutions and policy choices: evidence from the united states. *Journal of Economic Literature*, 41(1):7–73.

Besley, T. J. and Case, A. (1995). Does Electoral Accountability Affect Economic Policy Choices? *Quarterly Journal of Economics*, 110(3):769–798.

Björkman, M. and Svensson, J. (2010). When is community-based monitoring effective? evidence from a randomized experiment in primary health in uganda. *Journal of the European Economic Association*, 8(2-3):571–581.

Bobonis, G. J., Fuertes, L. R. C., and Schwabe, R. (2016). Monitoring Corruptible Politicians. *American Economic Review*, 106(8):2371–2405.

Bosio, E., Djankov, S., Glaeser, E., and Shleifer, A. (2022). Public procurement in law and practice. *American Economic Review*, 112(4):1091–1117.

Brasil (2023). Relatório de avaliação sobre o tratamento diferenciado concedido a me/epp em compras públicas. Technical report, Controladoria-Geral da União – Secretaria de Gestão.

Brollo, F., Nannicini, T., Perotti, R., and Tabellini, G. (2013). The Political Resource Curse. *American Economic Review*, 103(5):1759–1796.

Calonico, S., Cattaneo, M. D., and Farrell, M. H. (2020). Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4):909–946.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326.

Carril, R. et al. (2021). *Rules versus discretion in public procurement*. GSE, Graduate School of Economics.

Carril, R., Gonzalez-Lira, A., and Walker, M. S. (2022). Competition under incomplete contracts and the design of procurement policies. Technical Report 1824, Department of Economics and Business, Universitat Pompeu Fabra. Publication Title: Economics Working Papers.

Colonnelli, E., Gallego, J., and Prem, M. (2022). 16. what predicts corruption? *A Modern Guide to the Economics of Crime*, page 345.

Colonnelli, E. and Prem, M. (2022). Corruption and Firms. *Review of Economic Studies*, 89(2):695–732.

Coviello, D., Guglielmo, A., Lotti, C., and Spagnolo, G. (2022). Procurement with manipulation.

Coviello, D., Guglielmo, A., and Spagnolo, G. (2018). The effect of discretion on procurement performance. *Management Science*, 64(2):715–738.

Dahis, R., Carabetta, J., Scovino, F., Israel, F., and Oliveira, D. (2022). Data basis (base dos dados): Universalizing access to high-quality data. *Available at SSRN 4157813*.

Dahis, R., Ricca, B., Scot, T., Sales, N., and Nascimento, L. (2023). Mides: New data and facts from local procurement and budget execution in brazil.

Decarolis, F., Fisman, R., Pinotti, P., and Vannutelli, S. (2020). Rules, discretion, and corruption in procurement: Evidence from italian government contracting. Technical report, National Bureau of Economic Research.

Farenas, F. (2024). Figuring out the ideal chunk size. https://medium.com/@farenas1/fabian-7d1f90ac4cb4. Accessed: Mai 2, 2024.

Fazio, D. (2022). Rethinking discretion in public procurement: Evidence from brazil. *Available at SSRN*.

Ferraz, C. and Finan, F. (2008). Exposing Corrupt Politicians: The Effects of Brazil's Publicly Released Audits on Electoral Outcomes. *Quarterly Journal of Economics*, 123(2):703–745.

Ferraz, C. and Finan, F. (2011). Electoral Accountability and Corruption: Evidence from the Audits of Local Governments. *American Economic Review*, 101:1274–1311.

Fiuza, E. P., Corseuil, C. H., Carvalho, A. V., and Coimbra, P. H. (2023). Set asides and value thresholds in brazilian public procurement from small businesses.

Gallego, J. A., Prem, M., Vargas, J. F., et al. (2020). Corruption in the times of pandemia. *Available at SSRN*, 3600572.

Hubbard, T. P. and Paarsch, H. J. (2009). Investigating bid preferences at low-price, sealed-bid auctions with endogenous participation. *International Journal of Industrial Organization*, 27(1):1–14.

Kelman, S. (1990). Procurement and public management. Technical report, American Enterprise Institute.

Khemani, S. (2020). An opportunity to build legitimacy and trust in public institutions in the time of covid-19. *World Bank Research and Policy Briefs*, (148256).

Krasnokutskaya, E. and Seim, K. (2011). Bid Preference Programs and Participation in Highway Procurement Auctions. *American Economic Review*, 101(6):2653–2686.

Leeson, P. T. and Sobel, R. S. (2008). Weathering corruption. *The Journal of Law and Economics*, 51(4):667–681.

List, J. A. and Sturm, D. M. (2006). How Elections Matter: Theory and Evidence From Environmental Policy. *Quarterly Journal of Economics*, 121(4):1249–1281.

Liu, N. F., Lin, K., Hewitt, J., Paranjape, A., Bevilacqua, M., Petroni, F., and Liang, P. (2024). Lost in the middle: How language models use long contexts. *Transactions of the Association for Computational Linguistics*, 12:157–173.

Marion, J. (2007). Are bid preferences benign? The effect of small business subsidies in highway procurement auctions. *Journal of Public Economics*, 91(7):1591–1624.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, 142(2):698–714.

Olden, A. and Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*, 25(3):531–553.

Olken, B. A. (2007). Monitoring corruption: evidence from a field experiment in indonesia. *Journal of political Economy*, 115(2):200–249.

Olken, B. A. (2009). Corruption perceptions vs. corruption reality. *Journal of Public economics*, 93(7-8):950–964.

Olken, B. A. and Pande, R. (2012). Corruption in developing countries. *Annu. Rev. Econ.*, 4(1):479–509.

Palguta, J. and Pertold, F. (2017a). Manipulation of procurement contracts: Evidence from the introduction of discretionary thresholds. *American Economic Journal: Economic Policy*, 9(2):293–315.

Palguta, J. and Pertold, F. (2017b). Manipulation of Procurement Contracts: Evidence from the Introduction of Discretionary Thresholds. *American Economic Journal: Economic Policy*, 9(2):293–315.

Querubin, P., Snyder, J. M., et al. (2011). The control of politicians in normal times and times of crisis: Wealth accumulation by us congressmen, 1850-1880. Technical report, National Bureau of Economic Research.

Reis, P. R. and Cabral, S. (2015). Public procurement strategy: the impacts of a preference programme for small and micro businesses. *Public Money & Management*, 35(2):103–110.

Rose-Ackerman, S. (1998). Corruption and development. *Annual World Bank Conference on Development Economics 1997*, pages 35–57.

Svensson, J. (2005). Eight questions about corruption. *Journal of economic perspectives*, 19(3):19–42.

Szerman, D. (2012). *Public procurement auctions in Brazil.* phd, London School of Economics and Political Science.

Szucs, F. (2023). Discretion and favoritism in public procurement. *Journal of the European Economic Association*, page jvad017.

Thorstensen, V. and Giesteira, L. (2021a). Caderno brasil na ocde–compras públicas. *Relatório Institucional.*

Thorstensen, V. and Giesteira, L. F. (2021b). Caderno Brasil na OCDE – Compras Públicas. *Relatório Institucional*, pages 1–49.

Van Rijckeghem, C. and Weder, B. (2001). Bureaucratic corruption and the rate of temptation: do wages in the civil service affect corruption, and by how much? *Journal of development economics*, 65(2):307–331.

World Bank (2021). Subnational Doing Business in Brazil 2021. Publisher: World Bank, Washington, DC.

World Bank Group (2017). Technical Report: Policies that Promote SME Participation in Public Procurement.

Yemeke, T. T., Umaru, F. A., Ferrand, R. A., and Ozawa, S. (2023). Impact of the covid-19 pandemic on the quality of medical products in zimbabwe: a qualitative study based on key informant interviews with health system stakeholders. *BMJ open*, 13(6):e068923.