On the Costs and Benefits of Set Aside Auctions for Small Businesses

Alexsandros Cavgias
UGhent

Luis Meloni
USP

Dimas Fazio
NUS

Vita Titl
UUtrecht

February, 2025

Abstract

This paper investigates the impact of Set Aside (SA) auctions for Small and Medium Enterprises (SMEs) in Brazilian public procurement. Exploiting a legal reform that made SA auctions mandatory for contracts below BRL 80,000, we show that restricting competition to SMEs significantly increases their chances of winning contracts by 24%, but also raises procurement costs by 12.1%, implying an annual fiscal cost of approximately BRL 73 million, given the total procurement spending affected by the policy. This result likely reflects the exclusion of larger, potentially more efficient firms. A trade-off emerges when analyzing the effects of mandatory SA auctions. Using a difference-in-differences approach, we find significant positive effects on employment and earnings for SMEs, with an average increase of 0.93 jobs per firm. Our cost-benefit analysis suggests that, despite the additional procurement expenses, the economic gains for SMEs indicate that SA auctions can be a viable policy tool in certain contexts, with an estimated cost per job created of approximately USD 2,800.

Keywords: Public procurement, preferential treatment policies,
JEL Classifications: H57, D44, J23, L26, C21

1 Introduction

Governments often face a trade-off between maximizing efficiency in resource allocation and pursuing broader societal goals of inclusion and equity. This tension is particularly evident in public procurement, which accounts for approximately 15% of GDP and 28% of public spending globally (Bosio et al., 2022). While traditional procurement frameworks prioritize cost minimization, governments increasingly use their purchasing power to support economic inclusion. Preferential procurement policies, which favor certain underprivileged firms, are a key tool in this effort. Among them, Set-Aside (SA) auctions—where contracts are exclusively reserved for Small and Medium Enterprises (SMEs)—have gained popularity in both developed and developing economies.[1]

The economic effects of SA auctions remain ambiguous, particularly in developing economies. On one hand, restricting competition to SMEs may increase procurement costs by excluding potentially more efficient larger firms. On the other hand, SA auctions can promote SME growth by expanding access to government contracts, allowing firms to scale up, hire more workers, and integrate into formal markets. Smaller firms may also be more sensitive to procurement-driven demand shocks, amplifying employment effects. The net impact of these competing forces remains an open empirical question.

This paper provides a comprehensive cost-effectiveness analysis of SA auctions for SMEs in Brazil. Leveraging unique regulatory features and policy changes, we estimate both the costs and benefits of these preferential procurement policies. Our analysis employs two complementary empirical strategies. First, to quantify fiscal costs, we exploit a regulatory threshold that encourages SA auctions for contracts valued below BRL 80,000. Using a regression discontinuity design (RDD), we compare procurement outcomes for nearly identical products just above and below this threshold. Second, to measure the benefits of SA auctions, we exploit a 2014 reform that made SA auctions mandatory for contracts below BRL 80,000. Using a difference-in-differences (DiD) strategy, we compare outcomes for SMEs that participated in government procurement before the reform versus those that did not, before and after the policy change.

[1] The United States has promoted SA auctions for small businesses since 1953 (Albano et al., 2006, p.284), while China reserves 30% of its procurement budget for small firms (OECD, 2018, p.88). Russia has allocated 15% of annual procurement budgets to small businesses since 2006 (Tkachenko et al., 2019). More broadly, 47% of Central Asian countries, 37% of Sub-Saharan African nations, and 26% of Latin American states implement some form of preferential treatment for SMEs (IFC World Bank Group, 2010).

Our findings reveal a key trade-off in the implementation of SA auctions. On the cost side, our RDD analysis shows that restricting competition to SMEs increases procurement prices by 12.1%, implying an annual fiscal cost of approximately BRL 72.6 million, given that total procurement spending affected by the policy amounts to BRL 600 million per year. Over the seven years covered in our sample, this represents a cumulative cost of BRL 508.2 million. Despite this price premium, SA auctions achieve their primary objective: increasing SME participation in public procurement, as the probability of an SME winning a government contract rises by 17.9 percentage points, a 24% increase relative to the control mean of 75.3%. Moreover, our difference-in-differences estimates indicate that SMEs benefiting from SA auctions experience an average increase of 0.93 jobs per firm, which, given our sample of approximately 33,000 firms, translates into the creation of around 31,000 new jobs. A back-of-the-envelope cost-benefit analysis suggests that the cost per job created is BRL 16,400 (approximately USD 2,800).

While our estimates are subject to potential identification challenges, we take several steps to ensure the robustness of our findings. On the cost side, one key concern is the potential manipulation of procurement values around the regulatory threshold, which could introduce bias in our RDD estimates. To address this, we implement a donut RDD approach that excludes potentially manipulated observations within BRL 2,500 below the threshold. This more conservative specification yields a slightly larger price effect of 13.3%, suggesting that any manipulation may actually lead to an underestimation of the price premium. Another concern is that differences in product quality between SMEs and larger firms could drive the observed price differential. To rule out this possibility, we analyze quality-adjusted prices that control for product brand and lot size, finding a nearly identical effect of 12.3% on quality-adjusted prices.

To further validate these cost-side findings, we complement our RDD estimates with a difference-in-differences strategy that exploits the 2014 reform making SA auctions mandatory below the BRL 80,000 threshold. This approach compares agencies with varying levels of pre-reform SA auction adoption and yields price effects ranging from 7–12% across specifications, closely aligning with our RDD estimates. The consistency between these two distinct empirical strategies, which leverage different sources of variation, strengthens the evidence that the price premium reflects a genuine fiscal cost of the policy rather than an artifact of our research design.

We apply a similarly rigorous approach to ensure the validity of our benefit-side estimates. One concern is that government-supplying and non-supplying SMEs may have followed different trajectories for reasons unrelated to the policy change. We address this by showing parallel pre-trends in employment and earnings from 2011 to 2013, followed by clear divergence after the 2014 reform. Another challenge is that selection into government-supplier status could create compositional differences between treatment and control firms. To mitigate this, we implement a propensity score matching procedure that balances firm characteristics across 567 industry-state strata. Additionally, time-varying regional or industry-specific shocks could differentially affect treatment and control firms. We account for this by progressively adding state-by-year, industry-by-year, and state-industry-year fixed effects, with our estimates remaining remarkably stable across these increasingly demanding specifications. As additional robustness checks, we interact year fixed effects with a rich set of pre-treatment firm characteristics and outcomes. Even in this most demanding specification, which flexibly controls for potential confounds, we find a highly significant employment and wage effects. The stability of our estimates across these specifications provides strong evidence that our findings capture the causal impact of increased access to SA auctions rather than differential trends across regions or firm types.

This paper contributes to several strands of literature. First, we contribute to the broad literature studying the impact of different procurement policies on procurement outcomes. While quantitative analyses of mediators and causal pathways are scarce in the procurement literature (Fazekas and Blum, 2021), we precisely quantify the contribution of supplier selection and auction competitiveness to the treatment effect of SA auctions on procurement outcomes. Second, we add to the literature investigating the economic consequences of preferential policies for SMEs, which offers limited evidence from developing countries (Fazekas and Blum, 2021). Previous research on preference policies in public procurement (e.g., Marion, 2017, 2007, 2009) has typically focused on a single product and geographically restricted markets. We contribute by providing clean evidence about the costs and benefits of SA auctions using data from a large developing country and a multiplicity of products.

Our paper also contributes to the growing literature on SME growth and development. While most work investigating how to promote firm growth has focused on relaxing supply-side constraints, we know much less about which demand-side policies are effective for promoting firm growth (Woodruff, 2018). Our findings shed light on how targeted procurement policies can influence SME development through increased government demand, building on evidence that winning procurement contracts increases employment (Ferraz et al., 2015), firm growth (Gugler et al., 2020), and firm survival (Cappelletti et al., 2024).

The remainder of this paper is organized as follows. Section 2 provides background on Brazil's procurement regulations and the implementation of SA auctions. Section 3 describes our data sources and sample construction. Section 4 outlines our empirical strategy and results of the impact of SA auctions on procurement outcomes and firm performance. Section 5 discusses the cost-effectiveness of the policy. Section 6 concludes with policy implications.

2 Background

2.1 Set-Aside Auctions for Small Businesses

The implementation of Set-Aside (SA) auctions for Small and Medium-sized Enterprises (SMEs) in Brazil underwent a significant regulatory change in 2014. Initially, under Supplementary Law 123 of 2006, SA auctions were optional for purchases below BRL $80,000$, granting bureaucrats full discretion over their enforcement. Moreover, the total value of contracts awarded through SA auctions could not exceed 25% of the annual procurement budget.

This changed with the enactment of Supplementary Law 147 of 2014, which introduced two key modifications. First, it made SA auctions mandatory for purchases below BRL $80,000$, eliminating bureaucratic discretion. Second, it removed the 25% cap on the total value allocated to SA auctions, expanding their scope significantly.

Despite this shift, the law still provided exceptions under which public managers could bypass SA auctions. According to Article 49, exemptions were allowed if SA auctions were deemed disadvantageous for public administration, if they could compromise the supply of goods, or if fewer than three eligible MSE suppliers were available. As a result, even after 2014, SA auctions were not universally enforced, as illustrated in Graph (a) in Figure 1.

Additionally, ambiguity remains regarding whether the BRL $80,000$ threshold applies to individual auctions or groups of auctions (i.e, bundles or tenders). Given the tendency of Brazilian public managers to adopt more conservative interpretations of regulations, they likely apply the rule at the group level, restricting SA auctions only when the total group value does not exceed the threshold.

2.2 Defining Small Business in Brazil

SMEs in Brazil are businesses classified according to Lei Complementar 123/2006 (and its update in Lei Complementar 155/2016) as Microempresas (MEs) and Empresas de Pequeno Porte (EPPs), with annual gross revenue up to BRL 4.8 million. This encompasses both *Microempresas* with annual revenue up to BRL 360,000 and *Empresas de Pequeno Porte* with annual revenue between BRL 360,000 and BRL 4.8 million. Individual microentrepreneurs (MEIs) with revenue up to BRL 81,000 are also included in the broader SME category.

Given that the legal criteria for defining SMEs vary widely across countries (Ayyagari et al., 2007), comparing the type of firms targeted by SA auctions in Brazil with those in the US, EU, and other developing countries is essential to understanding the external validity of our results. While Brazil defines SMEs primarily based on annual revenue, other countries use different criteria, such as the number of employees, turnover, or assets.

The European Union classifies SMEs as firms with fewer than 250 employees, an annual turnover of less than €50 million, or total assets below €43 million (IFC World Bank Group, 2010). In the United States, SME classification depends on industry-specific thresholds for number of employees and annual receipts, as established by the Small Business Administration (Athey et al., 2013).[2]

2.3 The Brazilian Procurement System

Public procurement in Brazil is regulated by Federal Law No. 8,666/1993, which establishes two main procedures: **bid waivers** (Dispensa de Licitação) and **competitive auctions** (Pregão Eletrônico). Each of these mechanisms accounted for nearly half of the total procurement value in 2019 (World Bank, 2012). Most competitive auctions take place

[2]See also https://www.sba.gov/federal-contracting/contracting-guide/size-standards. Manufacturing entities generally qualify as small businesses with up to 500 employees, while many service industries use revenue thresholds ranging from USD 750,000 to USD 38.5 million in annual receipts The SBA reviews and adjusts these standards periodically to reflect changing market conditions.

through electronic bidding, governed by Federal Law No. 10.520/2002.

The procurement process begins with regulatory approval, followed by the publication of a detailed notice specifying the purchase objective (e.g., acquisition of painting materials), the required items (e.g., paint sealer, acrylic paint), and other procedural details.

An **auction** refers to a single item listed in the procurement system, for which we observe the bidding outcomes. A **tender**, on the other hand, is a broader event grouping multiple auctions under a single public notice, typically issued by a purchasing agency (PA). Each tender consists of separate auctions for each item, with all items sharing a common procurement goal. PAs publish procurement notices and related documents at the tender level. A **bundle** is a set of related goods or services grouped together as a single procurement lot, with each lot in the bundle procured in a different auction. PAs typically bundle related products to ensure they are supplied by a single vendor, which can improve compatibility, simplify contract management, and potentially reduce costs.

Once the electronic bidding stage begins in the ComprasNet (CNET) portal, bidders' initial proposals automatically become their first bids. Participants can submit new bids, which must be lower than their previous bid but can be higher than the lowest bid at that moment. Throughout the process, bidders can observe the current lowest bid in real time but do not have access to bid histories or the identities of other participants.

Two key bureaucrats play distinct roles in the public procurement process: the **auctioneer** (*pregoeiro*) and the **procurement manager** (*homologador*). The auctioneer is responsible for conducting bidding sessions, evaluating proposals, negotiating prices, and selecting the winning bid. However, they do not participate in drafting the tender notice. The procurement manager, in contrast, holds an administrative role within the purchasing agency (PA), overseeing tasks such as market price research and determining the appropriate bidding method.

The Brazilian public sector operates under two main administrative systems: one responsible for managing payments and contracts related to public procurement, and another structuring the federal bureaucracy.[3] Since our analysis focuses on procurement outcomes, we primarily rely on organizational units from the procurement management

[3] The financial system used to manage procurement is the *Sistema Integrado de Administração de Serviços Gerais* (SIASG), while the bureaucratic structure follows the *Sistema de Organização e Inovação Institucional do Governo Federal* (SIORG).

system rather than the personnel system. Therefore, we define a purchasing agency (PA) as the lowest hierarchical level within the payment system, typically corresponding to institutions such as public hospitals and schools.[4]

Products procured by public agencies follow a standardized classification system based on an official government catalog.[5] Following Fazio (2024), we define a product as the combination of a cataloged product code and its corresponding unit of measurement. Specifically, we use the variable *PADRAO_MATERIAL*, which represents the second most detailed classification level in the system. Internet Appendix Table A1 provides an example list of products and their corresponding units of measurement, illustrating the variety of items procured — from office supplies to medical and construction materials.

3 Data

3.1 Data sources

We leverage detailed procurement data from *ComprasNet*, the Brazilian federal government's electronic procurement system. This dataset provides a comprehensive record of procurement tenders, containing the full universe of documents used in federal procurement since 2007.

We complement our procurement data with firm-level information from the publicly available *Cadastro Nacional de Pessoas Jurídicas (CNPJ Aberto)*, the Brazilian National Registry of Firms, maintained by the *Receita Federal do Brasil (RFB)*, the country's fiscal authority.[6] This dataset contains rich firm-level characteristics, including registration status, sector classification, and opening and closing dates. Crucially for our analysis, it provides an SME status indicator based on revenue reports to the RFB. Since we use a 2019 version of the dataset, our SME classification follows the revenue thresholds established by Law 155/2016.

To track labor market outcomes, we use data from the *Relação Anual de Informações Sociais* (RAIS) from 2011 to 2021. RAIS is an employer-employee matched panel cover-

[4] In the SIASG, PAs are formally called *Unidade de Administração de Serviços Gerais* (UASG).

[5] The classification system follows the Catálogo de Materiais (CATMAT), maintained by the Brazilian federal government. The most updated version is available at the link.

[6] The most updated version of the CNPJ Aberto is available at the link.

ing the universe of formal employment relationships in Brazil.[7] This dataset allows us to examine firm-level employment dynamics, including workforce size, wages, and job creation patterns.

To account for spatial and regional economic characteristics, we incorporate data from the ESTBAN, a dataset maintained by the Central Bank of Brazil, and the Demographic Census. The ESTBAN dataset provides detailed information on the banking sector at the municipal level, including the density of bank branches and deposit levels. Additionally, we use demographic and economic indicators from the Demographic Census, such as GDP per capita, Gini index, and total population at the municipal level, sourced from the *Instituto Brasileiro de Geografia e Estatística* (IBGE).

3.2 RDD estimating sample

Samples: procurement data We define four samples based on the auction-level procurement data. We use two samples in the estimation of the costs of SA auctions for MSEs: the RDD sample and the DiD sample. **Full sample.** The *full sample* comprises 36.2 million purchase acts recorded on the CNET website between 1997 and 2021. This sample includes purchase acts for 330 thousand distinct product categories (goods and services), organized by 6.7 thousand purchasing agencies into 19.2 million bundles and 5 million tenders.

Filtered sample. We apply four filters to extract our filtered sample from the full sample, which will later serve as input to select the estimating samples of both our RDD and DiD with procurement outcomes. First, as SA auctions for MSEs only occur in non-discretionary procurement, we keep only the 19.27 million auctions using competitive auctions - i.e., MODALIDADE=5. Second, we excluded 2.6 million auctions from 1997–2006, when CNET's auction volume was still rising. Third, to mitigate the impact of potential errors in measurement units, we exclude auctions where the unit price exceeds ten times or falls below one-tenth of the average unit price for the same product category in the same year. Respectively, these filters remove .89 and 1.11 million auctions, around 5.34% and 6.67% of the remaining 16.64 million auctions. Finally, we exclude 1.4 million auctions involving services because the fine-grained service classifications lack sufficient

[7] Access to the identified version of the RAIS was granted through an institutional agreement between the *Ministério do Trabalho e Emprego* (MTE) and the *Universidade de São Paulo* (USP).

precision for cross-product price comparisons. After these adjustments, the filtered sample consists of 13.6 million auctions spanning 118.92 thousand product categories, managed by 4.4 thousand purchasing agencies and structured into $xx.x$ million bundles and 408.8 thousand tenders.

DiD sample. We apply one additional restrictions to extract the DiD sample from the filtered sample. More precisely, we exclude 4.5 million auctions implemented between 2007 and 2010, when the adoption of voluntary SA auctions were voluntary and adoption was rising. Moreover, by implementing such a restriction, we can also match the sample periods of the DiD with procurement outcomes and the DiD with establishment outcomes (2011-2020). After this step, the DiD sample contains 9.53 million auctions in 298.5 thousand tenders procuring 92 thousand distinct product categories organized by 4 thousand purchasing agencies. Of these, 3.76 million are SA auctions designated for MSEs.

RDD sample. We apply two additional restrictions to extract the RDD sample from the filtered sample. First, to estimate the effect of mandatory SA auctions for MSEs, we exclude 3.36 million auctions conducted between 2013 and 2021, during which SA auctions for MSEs were optional for purchases below BRL $80,000$. Second, we restrict the sample to the 1.05 million auctions in bundles valued between BRL $20,000$ and BRL $140,000$, using a symmetric window to avoid including discontinuities caused by changes in the likelihood of discretionary procurement at BRL $17,600$ and the introduction of preference margins for MSEs at BRL $180,000$. After these steps, the RDD sample contains 780 thousand bundles of 107 thousand tenders procuring 30.4 thousand distinct product categories organized by 3.3 thousand purchasing agencies. Of these, 496 thousand are SA auctions designated for MSEs.

Outcome Variables. Our primary procurement outcome is a price index, constructed as the residual from an OLS regression of the logarithm of the unit price on product-year fixed effects (FEs), using a sample that includes all auctions from 2013 to 2021. This price index captures the percentage deviation of the unit price from the average unit price paid by the government for products within the same category and year. Using the logarithm of the unit price directly as the outcome in an RDD regression with product-year FEs is less suitable for cost-effectiveness analysis. Such an approach would introduce greater noise when restricting the sample to auctions within the RD population, leading to less precise measures of deviations from the product category average.

Following influential contributions in public economics Best et al. (2019, 2017); Bandiera et al. (2021), we also compute a measure of *quality-adjusted price* as a primary outcome. This measure is derived as the residual from an OLS regression of the logarithm of the unit price on product-year fixed effects (FEs), brand FEs, and the size of the lot as a control. The quality-adjusted price reflects the percentage variation of the unit price compared to the average unit price paid by the government for products within the same category and year, holding product quality and lot size constant.

Figure 1 presents the main source of variation used in this study to causally identify the effects of SA auctions for SMEs, which stems from the regulatory change introduced in 2014. Graph **(a)** depicts the share of SA auctions for SMEs across tenders of varying values, highlighting the discontinuity at the BRL $80,000$ threshold. Graph **(b)** presents the evolution of SA auction frequency between 2007 and 2021, capturing the sharp increase following the 2014 reform.

Several patterns emerge from Graph **(a)** in Figure 1, with direct implications for identification. First, consistent with the assumption that public managers adopt the most restrictive interpretation of the regulation, we observe a sharp increase in the share of SA auctions for SMEs for tenders below the BRL $80,000$ threshold, but not when using the auction value as the running variable. Second, even after the 2014 regulatory change, SA auctions for SMEs continue to occur in approximately 10% of cases, which we attribute to the exceptions outlined in the regulation. Third, in line with findings from the procurement literature (Fazio, 2024; Decarolis et al., 2020), we detect visual evidence of bunching just below the BRL $80,000$ threshold, although the magnitude appears moderate in our context.

Given these patterns, our empirical strategy, detailed in Section 4, employs a *Donut Fuzzy* RDD using tender value as the running variable, with standard errors (SEs) clustered at the tender level.

Encouraging patterns also emerge from Graph **(b)** in Figure 1. The share of SA auctions increases substantially after 2013, aligning precisely with the moment when SA auctions became mandatory. This sharp change supports the validity of our empirical strategy. Consequently, in subsubsection 4.1.2, we implement a difference-in-differences (DiD) design, defining 2007–2013 as the pre-treatment period and 2014–2021 as the post-treatment period.

3.3 DiD estimating sample

In this section, we describe the process of constructing the establishment-level samples used in our analysis, particularly the final dataset employed in the difference-in-differences (DiD) estimation. We begin by introducing the three key datasets, followed by the criteria applied to filter establishments into our final estimating sample.

We use three establishment-level data samples. The *full sample* is a yearly panel dataset containing the universe of 5.2 million unique establishments with employment information registered in RAIS between 2011 and 2020, of which $56,561$ are Government Suppliers (GSs). The *filtered sample* is a subset of the full sample obtained by applying a series of selection criteria to ensure comparability across establishments, resulting in 1.48 million establishments, of which $33,731$ are GSs. Finally, the *matched sample* is our final estimation sample for the DiD regression model described in Subsection 4.2, containing $66,698$ establishments, evenly split between GSs and non-GSs.

Filtering Criteria. To construct the filtered sample from the full sample, we apply four selection filters designed to isolate the establishments most relevant to the SA auction policy and to ensure robustness in our DiD estimation.

First, to target the population affected by SA auctions, we restrict the sample to SME establishments. Since the official SME classification is based on firm revenues—unavailable in our dataset—we rely on an employment-based proxy, which is consistently observed over time. Specifically, following Colonnelli et al. (2020), we define an establishment as an SME if it falls into one of three size categories: *Micro1* (1–4 employees), *Micro2* (5–9 employees), or *Small* (10–49 employees). We then classify each establishment based on the most frequent (modal) size category it belonged to during the pre-treatment period (2011–2013).

Second, to avoid selection biases stemming from establishments without formal employees before the policy change, we exclude those with zero wages or zero employees in the pre-treatment period. This filter reduces the sample to approximately 4.8 million establishments, of which about 54.8 thousand are GSs.

Third, to ensure that pre-treatment trends can be properly assessed, we exclude establishments that are not continuously observed throughout the entire pre-treatment period (2011–2013). This restriction further refines the sample to approximately 1.5 million establishments, of which about 33.7 thousand are GSs.

After applying these filters, the resulting *filtered sample* contains 1.48 million establishments, of which approximately 33.7 thousand are GSs. This refined dataset serves as the foundation for our final matching procedure, which constructs the *matched sample* used in our DiD estimation.

SME Proxy Selection As discussed in Subsection 3.1, the CNPJ Aberto dataset is only available from 2018 onwards. This limitation prevents us from directly observing SME status during the pre-treatment period (2011–2013), requiring the use of a proxy variable to define the DiD estimating sample. To address this issue, we rely on an SME classification derived from RAIS data, following the methodology proposed by Colonnelli et al. (2020).

Although the official SME definition is based on annual revenue, using an employment-based proxy offers several advantages. First, it ensures a high degree of accuracy, as the correlation between the official SME classification from CNPJ Aberto and our proxy from RAIS exceeds 0.99 in 2018, when both variables are available. Second, using the CNPJ Aberto classification from 2018 would introduce sample contamination, as it could include relatively large firms that held SME status in 2018 but were non-SMEs in 2007–2013, as well as firms that only gained SME status after the 2016 revision of the eligibility criteria. Third, by limiting the sample based on employment size, our approach reduces the skewness of the primary outcome variable, improving the common support assumption in the matching process and minimizing the influence of extreme values in our estimates.

Thus, the employment-based SME proxy provides a reliable alternative for sample selection, enhancing the credibility of our matching procedure and facilitating inference in the DiD analysis.

Matching Algorithm. We implement a three-step matching procedure to construct the matched sample from the filtered sample.

First, we stratify establishments into $21 \times 27 = 567$ industry-state groups, created by the interaction of 21 two-digit industry categories (CNAE 2.0) and the 27 states in which establishments are located.

Second, within each industry-state stratum, we estimate propensity scores (PSs) at the establishment level using a logit model. The covariates include the establishment's average monthly wages (2011, 2012, and 2013), number of employees (2011, 2012, and

2013), total payroll (2011, 2012, and 2013), and the quintile of the total population distribution across municipalities (2011). By restricting the PS estimation to a limited set of pre-treatment variables, we ensure that we can later evaluate the performance of the matching algorithm through balance tests on covariates not included in the PS estimation.

Third, we pair each GS establishment with the closest non-GS establishment based on the estimated PS, using nearest-neighbor matching without replacement. The matching procedure results in minimal attrition, as only around 300 of the 33,700 GS establishments remain unmatched. The final matched sample consists of approximately 66,700 establishments, evenly split between GS and non-GS establishments.

Validation of Matching Algorithm. We take three steps to validate the outcomes of our matching algorithm. First, we motivate the common support assumption by analyzing the distribution of the estimated propensity scores (PSs) of GSs and non-GSs. Both the unscaled and scaled (standardized) PS distributions of GSs and non-GSs show a very high degree of overlap, suggesting that our matching algorithm produces a control group that satisfies the common support assumption.

Second, we evaluate whether our matching algorithm improves the relative level of covariate unbalance between GSs (treatment group) and non-GSs (control group) by comparing the balance-check statistics from the filtered sample with those from the matched sample. For nearly all of our pre-treatment variables, we observe a substantial decrease in the T-statistic and standardized mean difference from the matched sample in comparison to those from the filtered sample.

Third, we evaluate whether the absolute level of covariate unbalance is suitable for conducting a DiD estimation by comparing the joint distribution of balance-check statistics from the matched sample with statistical significance critical values and definitions of small magnitudes of mean differences. Our matching procedure generates sufficient similarity between GSs and non-GSs to conduct causal inference using a DiD regression model based on the parallel trends hypothesis. In particular, we document that pre-treatment outcomes of GSs and non-GSs become balanced in the matched sample in the first difference, showing that pre-treatment outcome trends are parallel within the matched sample. Additionally, while we reject the hypothesis of balance in covariates for more than half of our pre-treatment variables, the standardized differences have a small magnitude in the matched sample.

Relevant to our identification strategy, the parallel trends hypothesis does not require that treatment and control groups be perfectly balanced, making the outcome of the matching algorithm acceptable for our purpose. However, to ensure that such remaining unbalances between GSs and non-GSs do not drive our results, we re-estimate our DiD specification with year fixed effects interacted with pre-treatment variables as controls in Subsection B.2.

Figure B2 provides evidence that our matching algorithm significantly improves balance between GS and non-GS establishments. The balance improves across key covariates, particularly those related to firm size, workforce composition, and local economic conditions. These results indicate that our matching procedure effectively minimizes pre-treatment differences, enhancing the comparability between treated and control establishments in our DiD analysis.

Outcomes. We analyze two main procurement outcomes. First, we evaluate the effect of SA auctions on the procurement price, defined as the log of the final price paid for each auctioned item, capturing potential cost impacts of restricted competition. Second, we examine SME participation, measured by an indicator variable that equals one if the auction was won by an SME, assessing whether the policy effectively increased SME access to public procurement.

Treatment Variable. We construct a continous measures of treatment intensity: the share of non-SA auctions in eligible procurements (i.e., below BRL 80,000) conducted by each agency during the pre-treatment period (2011–2013). These measures leverage heterogeneity in pre-reform procurement practices to identify the causal impact of SA auctions. Agencies that relied more on non-SA auctions before the reform experienced a larger policy shock when SA auctions became mandatory. Our empirical specification interacts treatment intensity with a post-reform indicator, ensuring identification comes from within-agency changes over time.

Control variables. We divided the control variables from the DiD estimating sample, $\mathbf{x}_e$, into three blocks: $[(\mathbf{y}_{2011,e}, \mathbf{y}_{2012,e}, \mathbf{y}_{2013,e}), (\mathbf{x}_{2011,e}, \mathbf{x}_{2012,e}, \mathbf{x}_{2013,e}), \mathbf{x}_{m(e)}]$. Here, e denotes an establishment, and $m(e)$ refers to the municipality where establishment e is located. We precisely define these units in Section 2. The first block, $(\mathbf{y}_{2011,e}, \mathbf{y}_{2012,e}, \mathbf{y}_{2013,e})$, includes 12 pre-treatment outcomes: our four establishment-level outcomes (number of employees, average monthly earnings, and total yearly payroll) observed in each of the three pre-treatment years (2011, 2012, and 2013). The second block, $(\mathbf{x}_{2011,e}, \mathbf{x}_{2012,e}, \mathbf{x}_{2013,e})$, consists of 15 pre-treatment controls: five establishment-level characteristics (employee and employer attributes) recorded for each of the three pre-treatment years. These controls include establishment age, average employee tenure, percentage of female employees, average employee age, and percentage of employees with a university degree. The third block, $\mathbf{x}_{m(e)}$, captures six municipality-level characteristics: GDP per capita (in 2000), the Gini index of GDP per capita (in 2000), total population, the percentage of GDP from the manufacturing sector, bank agency density, and bank deposit density. Table B2 in Appendix B provides detailed definitions of all variables in the DiD estimating sample.

4 Empirical Analysis

4.1 SA Auctions and Procurement Outcomes

This section evaluates the impact of set-aside auctions (SA) for SMEs on procurement outcomes, such as prices. The effect of SA auctions on prices is theoretically ambiguous. On one hand, it could drive up procurement prices by limiting competition and excluding potentially more efficient non-SME firms from the auction. On the other hand, SA auctions it might increase participation of SMEs who would otherwise be deterred from competing against larger firms, potentially leading to lower prices via higher competition. The actual effect on procurement costs thus remains an empirical question.

To evaluate the effects of SA auctions on procurement outcomes, we exploit a regulatory threshold from Brazil's procurement regulations, which encourage SA auctions for auctions valued below BRL 80,000. We exploit two different empirical approaches. First, we analyze the impact on procurement outcomes by comparing auctions around this value threshold using a regression discontinuity design. This allows us to estimate how SA auctions affect prices and SME participation in government contracts. Second, we use a 2014 reform that strengthened the enforcement of SA auctions below the BRL 80,000 threshold. Using a difference-in-differences strategy, we compare outcomes for SMEs that participated in government procurement before the reform (treated) versus those that did not (control), before and after the policy change. This approach enables us to quantify how increased access to SA auctions affects key firm outcomes such as employment and wages. Both of these methodologies and its results are explained in detail below.

4.1.1 Comparing Outcomes Around the SA Auctions Threshold

We first quantify the impact of SA auctions for SMEs on procurement outcomes by comparing prices (within fine-grained product categories) of auctions around the BRL $80,000$ cutoff for SA auctions in a fuzzy RDD setting. Panel A of Figure 1 shows the frequency of set asides around the threshold. One can see that the number of procurements that are set asides below the threshold average 60% while those above the threshold are about 10%. Below the threshold the number of procurements is not exactly equal to 1 for two reasons. First, before 2014 it was not mandated that tenders below 80,000 are set asides. Second, even after 2014, there are exceptions to the rule that allow agencies not to set aside procurements below the threshold. Specifically, agencies can forgo set-asides when: (1) there are fewer than three competitive SME suppliers in the local or regional area capable of fulfilling the contract requirements, or (2) the differential treatment of SMEs would not be advantageous for the public administration or would compromise the overall objective of the procurement. Conversely, agencies retain the discretion to set aside procurements above the threshold, which explains the non-zero frequency of set-asides above BRL 80,000.

Given the imperfect enforcement of mandatory SA auctions for SMEs below the BRL $80,000$ cutoff documented in Figure 1, we estimate the Fuzzy RDD regression model

$$\begin{aligned} Set\_Aside_i &= \delta + \gamma \cdot \mathbf{1}(Value_{j(i)} < 80000) + f(Value_{j(i)}) + \epsilon_i \\ Outcome_i &= \alpha + \beta \cdot Set\_Aside_i \qquad\qquad\quad + f(Value_{j(i)}) + \epsilon_i \end{aligned}$$

where i denotes an auction and j is a tender that includes auction i. $Outcome_i$ is the procurement outcome of auction i and Set_Aside_i is an indicator taking value 1 when auction i is a SA for SMEs. $f(Value_{j(i)})$ is the RD polynomial. $Value_{j(i)}$ is the value of all products in auctions that belong to tender j.

Figure 2 provides graphical evidence using of our results. First, Panel A provides a similar graph as Panel A of Figure 1, showing the procurements below the threshold are more likely to be set aside. As a result of that, Panel B shows that procurements below the threshold are more likely to have SME winners. Panels C and D examine price outcomes, comparing auctions for identical products around the threshold. Both raw prices and quality-adjusted prices exhibit significantly higher levels just below the SA threshold compared to above it. This price differential suggests that restricting competition to SMEs through set-asides imposes a fiscal cost on the government: procurements restricted to SMEs result in higher prices paid by government agencies, even after accounting for potential differences in product quality.

Table 1 quantifies these graphical patterns through our RDD specification. The intention-to-treat estimates show that being below the threshold reduces SME winner likelihood by 9.2 percentage points and prices by 6.2%. The first stage demonstrates strong compliance with the policy - being below the threshold increases SA auction probability by 51.2 percentage points.

The second-stage estimates reveal that SA auctions increase SME winner probability by 17.9 percentage points, raise procurement prices by 12.1%, and increase quality-adjusted prices by 12.3%. These estimates are precisely estimated, with standard errors of 2.4-3.7 percentage points, and highly statistically significant. The results demonstrate a clear trade-off: while SA auctions successfully increase SMEs participation in procurement by about 24% relative to the control mean of 75.3%, they do so at the cost of a 12.1% price premium paid by the government. This price increase likely reflects the exclusion of potentially more efficient non-SMEs firms from the bidding process.

To address potential concerns about manipulation of the running variable, we compare our preferred donut RDD specification to conventional RDD estimates in Table 2. Our specification implements a one-sided donut hole that removes observations between BRL 77,500 and BRL 80,000, effectively excluding data within 2,500 BRL below the threshold where manipulation is most likely to occur. Panel A shows the intention-to-treat estimates, indicating that being below the BRL 80,000 threshold increases the likelihood of an MSE winning by 9.9 percentage points and increases prices by 6.3%. Panel B demonstrates the strength of our first stage, with auctions below the threshold being approximately 48 percentage points more likely to be set aside for MSEs. The second-stage fuzzy RDD estimates in Panel C reveal that SA auctions increase the likelihood of an MSE winning by 20.1 percentage points and raise prices by 13.3%. These results are similar and higher in magnitude to those in Table Table 1, suggesting that any potential manipulation at the threshold is not substantially biasing our findings regarding the impact of SA auctions on procurement outcomes. If anything, our results are slightly stronger when removing these possibly manipulated auctions.

4.1.2 Exploiting a Change in Enforcement of the SA Threshold

To complement our RDD analysis and address potential concerns about manipulation at the threshold, we exploit a reform that made SA auctions mandatory below BRL 80,000 in 2014. As shown in Panel B of Figure 1, this reform substantially increased the frequency of SA auctions, providing an alternative source of variation to identify the effects of restricting competition to SMEs on procurement costs.

Our identification strategy leverages heterogeneity in pre-reform SA adoption across agencies. We construct two measures of treatment intensity: (i) a continuous measure based on each agency's share of non-SA auctions in eligible procurements during 2011-2013, and (ii) a binary indicator for agencies whose pre-reform share of non-SA auctions exceeded the median. This approach allows us to identify effects from agencies that had the most scope to increase their use of SA auctions when they became mandatory.

We estimate event study specifications of the form:

$$y_{i,t} = \alpha_{p(i),t} + \alpha_{a(i)} + \alpha_{b(i)} + \beta \cdot Intensity_{a(i)} \cdot Post_t + \alpha_t \cdot Value_{j(i)} + \epsilon_{i,t}$$

where i denotes a procurement auction, and t time, which is the year. $\alpha_{p,t}$ captures product-year fixed effects (FEs), α_a purchasing agency FEs, and α_b brand FEs. $Intensity_a$ is our continuous treatment variable: the share of non-SA auctions made by agency a in the sample of purchases below BRL 80000 during the pre-treatment (2011-2013), when SA auctions for SMEs where voluntary. $Post_t = \mathbb{1}(t > 2013)$ is an indicator variable that equals one after 2013, the first year before the 2014 reform that made SA auctions for SMEs mandatory for purchases below BRL 80000.

Table 3 presents estimates from increasingly demanding specifications that progressively add fixed effects to control for potential confounders. Panel A shows the reduced-form effects on procurement prices, while Panel B presents the first-stage results on SA auction probability.

In Panel A, across all specifications, we find positive and statistically significant effects of the reform on prices. Starting with the baseline specification in Column (1), which includes only product and year fixed effects, we estimate that agencies with higher pre-reform shares of non-SA auctions experienced a 3.7% increase in prices after 2014 (standard error = 0.017). Adding product-year fixed effects in Column (2) increases the estimate to 5.2%, suggesting that controlling for time-varying product characteristics is important. The effect moderates to 4.6% when we add agency fixed effects in Column (3), and further stabilizes around 4.0-4.1% when we add controls for the running variable interacted with year fixed effects (Column 4) and brand fixed effects (Column 5).

Panel B reveals strong first-stage effects that are stable across specifications. In our most demanding specification (Column 5), agencies that never used SA auctions pre-reform increased their SA auction probability by 57.1 percentage points more than agencies that always used them (standard error = 0.029). This large and precisely estimated first-stage effect indicates that the reform substantially changed agency behavior.

Combining the reduced-form estimate of 4.0% with the first-stage estimate of 57.1% yields a scaled effect of approximately 7.0% (0.040/0.571). This scaling represents the price effect of SA auctions for complier agencies - those that changed their behavior due to the reform. The similarity of this estimate to our RDD results (which found effects of 7-12%) provides strong validation of our main findings through a completely different source of variation.

The stability of these estimates across specifications with increasingly demanding fixed effects suggests our results are robust to controlling for potential confounders like time-varying product characteristics, agency-specific factors, and brand-specific quality differences. The high R-squared values, particularly in Panel A (reaching 0.875 in Column 5), indicate that our specifications explain a large portion of the variation in procurement prices.

Figure 5 presents the dynamic effects, confirming the validity of our research design. Panel (a) shows the reform sharply increased SA auction probability, with parallel pre-trends followed by a roughly 40 percentage point increase by 2019. This translates directly into increased SME participation - Panel (b) demonstrates a similar pattern in SME winning probability. Most importantly, Panels (c) and (d) reveal that prices rose by about 5-7% in affected agencies after the reform, with the effect persistent across specifications with and without brand fixed effects. The consistency between our RDD and DiD estimates strengthens confidence in our core finding that restricting competition to SMEs raises procurement costs. While the RDD leverages variation from the contract value threshold, the DiD exploits differences in pre-reform SA adoption across agencies. Both approaches yield price premiums of 7-12%, suggesting this reflects a real cost of limiting the bidder pool rather than statistical artifacts from either research design.

The magnitude of these effects is economically meaningful. Given annual federal procurement spending of approximately BRL X billion in contracts below BRL 80,000, our estimates imply the mandatory SA policy increased government costs by roughly BRL Y billion per year. This raises the question of whether these fiscal costs are justified by corresponding benefits to SME development, which we examine next.

4.2 SA auctions and real outcomes

Our results show that restricting competition through set-aside auctions comes at a substantial cost to government procurement efficiency, with price premiums of 7-12% across different empirical strategies. This raises an important question: do these increased costs to the government translate into employment benefits for SMEs? While the higher prices represent additional profits for winning SMEs, these firms could potentially capture these gains as income for existing owners rather than expanding employment. Understanding whether and how much these fiscal costs generate employment gains is crucial for evaluating the benefits of the program for SME development, which we examine next.

To quantify the impact of SA auctions for SMEs on the outcomes of the target firms, we compare firm outcomes of government suppliers (GS) before and after the 2014 reform that caused a pronounced increase in the frequency of SA auctions for SMEs. Specifically, we estimate:

$$y_{e,t} = \alpha + \beta_1 GS_e + \beta_2 Post_t + \beta_3 GS_e \cdot Post_t + \epsilon_{e,t}$$

where f denotes firms and t, years. The dependent variable $y_{e,t}$ represents firm outcomes, GS_e indicates whether firm e was a government supplier before 2014, and $Post_t$ indicates post-reform periods $(t > 2013)$. Our treatment group consists of SMEs that participated in at least one procurement auction before 2014 (Government Suppliers or GSs), while our control group comprises SMEs that did not participate in any procurement auction during our sample period (Non-Government Suppliers or NGSs).

We next examine the impact of SA auctions on SME performance using our difference-in-differences strategy. Figure 5 provides graphical evidence of the effects. The figure shows that both employment and earnings exhibit parallel trends before 2014, followed by a clear divergence after the reform. This pattern persists across different specifications, with Panels (a) and (b) showing similar employment trajectories whether we use just establishment fixed effects or add state-industry-time controls. The same holds for average earnings effects in Panels (c) and (d).

Table 4 presents our baseline difference-in-differences estimates with increasingly demanding fixed effects specifications. Panel A focuses on employment effects. The raw specification in Column (1) shows that government suppliers increased employment by 0.831 workers after the reform. The effect increases to 0.939 workers when we include establishment fixed effects in Column (3), suggesting that controlling for time-invariant firm characteristics is important. The estimate remains stable at 0.947 and 0.933 workers as we progressively add state-by-year and industry-by-year fixed effects (Column 4) and state-industry-year fixed effects (Column 5). Relative to the pre-treatment mean of 10.171 workers, our preferred specification in Column 5 represents a 9.2% increase in employment.

Panel B of Table 4 examines earnings effects through the same progression of specifications. The raw difference-in-differences estimate in Column (1) shows an increase of 51.028 BRL in average monthly wages (standard error = 3.268). This effect declines modestly to 48.504 BRL when adding year fixed effects (Column 2) and further to 42.352 BRL with establishment fixed effects (Column 3). The estimate stabilizes around 43 BRL in our most demanding specifications with state-by-year and industry-by-year controls (43.082 BRL in Column 4) and state-industry-year fixed effects (43.178 BRL in Column 5). This final estimate represents a 3.9% increase relative to the pre-treatment mean of 1,102.494 BRL.

Table 5 provides additional robustness checks by flexibly controlling for potential confounds. For employment (Panel A), adding spatial controls interacted with year fixed effects (Column 2) reduces the baseline effect of 0.939 workers to 0.884 workers. The estimate remains stable at 0.896 workers when we add pre-treatment firm characteristics interacted with year effects (Column 3). Our most demanding specification, which includes pre-treatment outcome controls (Column 4), yields an effect of 0.766 workers that remains highly significant (standard error = 0.096).

The earnings effects in Panel B of Table 5 show similar stability. The effect declines from the baseline 42.352 BRL to 36.038 BRL when adding spatial controls (Column 2), then increases slightly to 37.512 BRL with pre-treatment firm controls (Column 3). Our most demanding specification with pre-treatment outcome controls (Column 4) yields an effect of 37.322 BRL that remains precisely estimated (standard error = 2.754).

The stability of these estimates across increasingly demanding specifications suggests they capture real effects of the reform rather than differential trends across regions or firm types. The combination of employment and wage growth indicates that increased access to government contracts through SA auctions enabled SMEs to both expand their workforce and pay higher wages.

5 Cost-Benefit Analysis

On the one hand, our findings indicate that the policy increased the share of SMEs winning public contracts. On the other hand, it did so at a considerable cost. In this section, we conduct a back-of-the-envelope analysis to quantify the associated fiscal burden and evaluate the cost-effectiveness of the policy in terms of employment outcomes.

Our RDD estimates suggest that the implementation of SA auctions led to a 12.1% increase in procurement prices. Given that total annual procurement spending affected by the policy amounts to approximately BRL 600 million per year, this implies an annual fiscal cost of approximately BRL 72.6 million and a total cost of BRL 508.2 million over the 7 years covered in our sample.

On the other hand, our difference-in-differences estimates indicate that SMEs benefiting from SA auctions experienced an increase of approximately 0.93 jobs per firm after the implementation of the policy. Given that our sample includes around 33,000 establishments, this implies the creation of approximately 31,000 new jobs as a direct consequence of the policy.

Dividing the total fiscal cost of R$ 72.6 million by the estimated 31,000 jobs created, we estimate a cost per job of R$ 16,400, equivalent to approximately US$ 2,800. Although this figure might not seem high when compared to other findings, we do not account for potential costs faced by larger companies that lose auctions due to the policy.[8]

6 Conclusion

This paper provides a comprehensive cost-benefit analysis of Set-Aside auctions for SMEs in Brazil's public procurement system. Our findings reveal a clear trade-off between procurement efficiency and SME development objectives. On one hand, restricting compe-

[8] Corbi et al. (2019) finds that a transfer policy in Brazil has a positive effect on the economy and job creation, with costs around US$ 8,000 per job.

tition to SMEs increases procurement prices by 12.1%, leading to an annual fiscal cost of approximately BRL 72.6 million. Over the seven years covered in our sample, this translates to a cumulative cost of BRL 508.2 million. On the other hand, increased access to government contracts generates significant benefits for targeted firms, including a 9.2% increase in employment and a 3.9% increase in average wages. The consistency of our estimates across multiple empirical strategies and robustness checks strengthens confidence in these results. Our cost-benefit analysis suggests that the policy costs approximately BRL 16,400 (US$2,800) per job created, a figure that compares favorably to other job-creation policies.

The implications of our findings extend beyond the Brazilian context. As governments worldwide increasingly adopt preferential procurement policies to support SMEs, understanding their economic consequences is critical for evidence-based policymaking. Our results suggest that while SA auctions successfully increase SME participation in public procurement and generate meaningful employment gains, they do so at a measurable fiscal cost. This underscores the importance of carefully designing and targeting such policies to maximize their effectiveness. Future research could explore alternative mechanisms that achieve similar distributive goals with lower efficiency costs, such as bid preferences or capacity-building programs for SMEs. Additionally, examining longer-term effects on firm productivity, innovation, and market competitiveness would provide a more comprehensive assessment of these policies' broader economic impact. Ultimately, the decision to implement preferential procurement policies involves balancing efficiency and equity objectives, requiring policymakers to weigh economic trade-offs alongside broader social priorities.

References

Albano, G. L., Dimitri, N., Perrigne, I., and Piga, G. (2006). *Fostering participation*, page 267–292. Cambridge University Press.

Athey, S., Coey, D., and Levin, J. (2013). Set-asides and subsidies in auctions. *American economic journal: microeconomics*, 5(1):1–27.

Ayyagari, M., Beck, T., and Demirguc-Kunt, A. (2007). Small and medium enterprises across the globe. *Small Business Economics*, 29(4):415–434.

Bandiera, O., Best, M. C., Khan, A. Q., and Prat, A. (2021). The Allocation of Authority in Organizations: A Field Experiment with Bureaucrats*. *The Quarterly Journal of Economics*, 136(4):2195–2242.

Best, M., Khan, A., Prat, A., Rehman, M., Asad, S. A., Hussain, K., and Majid, A. O. (2017). Motivating bureaucrats: Autonomy vs performance pay for public procurement in pakistan.

Best, M. C., Hjort, J., and Szakonyi, D. (2019). Individuals and organizations as sources of state effectiveness. Working Paper 23350, National Bureau of Economic Research.

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

Cappelletti, M., Giuffrida, L. M., and Rovigatti, G. (2024). Procuring Survival. Technical report.

Colonnelli, E., Gallego, J. A., and Prem, M. (2020). What predicts corruption? *Available at SSRN 3330651*.

Corbi, R., Papaioannou, E., and Surico, P. (2019). Regional transfer multipliers. *The Review of Economic Studies*, 86(5):1901–1934.

Decarolis, F., Fisman, R., Pinotti, P., and Vannutelli, S. (2020). Rules, discretion, and corruption in procurement: Evidence from italian government contracting. Working Paper 28209, National Bureau of Economic Research.

Fazekas, M. and Blum, J. R. (2021). *Improving Public Procurement Outcomes : Review of Tools and the State of the Evidence Base*. The World Bank.

Fazio, D. (2024). The two sides of discretion: Assessing efficiency and quality in government purchases.

Ferraz, C., Finan, F., and Szerman, D. (2015). Procuring firm growth: The effects of government purchases on firm dynamics. Working Paper 21219, National Bureau of Economic Research.

Gugler, K., Weichselbaumer, M., and Zulehner, C. (2020). Employment behavior and the economic crisis: Evidence from winners and runners-up in procurement auctions. *Journal of Public Economics*, 182:104112.

IFC World Bank Group (2010). Scaling-up sme access to financial services in the developing world.

Marion, J. (2007). Are bid preferences benign? the effect of small business subsidies in highway procurement auctions. *Journal of Public Economics*, 91(7-8):1591–1624.

Marion, J. (2009). How costly is affirmative action? government contracting and california's proposition 209. *The Review of Economics and Statistics*, 91(3):503–522.

Marion, J. (2017). Affirmative action exemptions and capacity constrained firms. *American Economic Journal: Economic Policy*, 9(3):377–407.

OECD (2018). *SMEs in Public Procurement*.

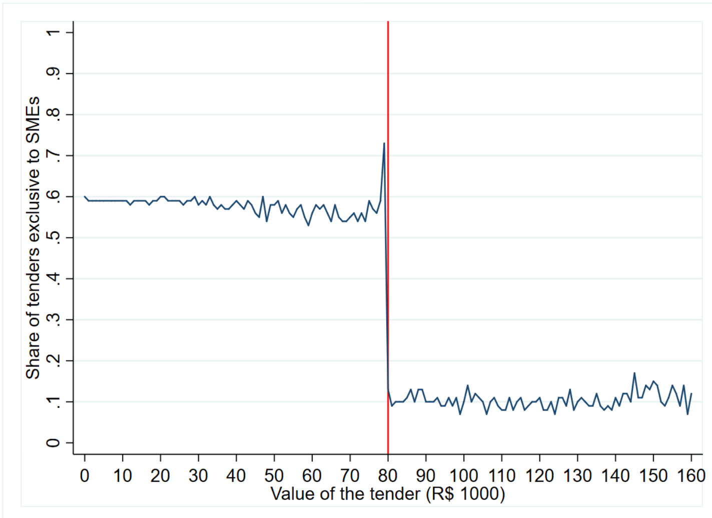
Tkachenko, A., Valbonesi, P., Shadrina, E., and Shagbazian, G. (2019). Efficient design of set-aside auctions for small businesses: an empirical analysis. Technical Report 240, Marco Fanno Working Papers.

Woodruff, C. (2018). Addressing constraints to small and growing businesses 24 october 2018.

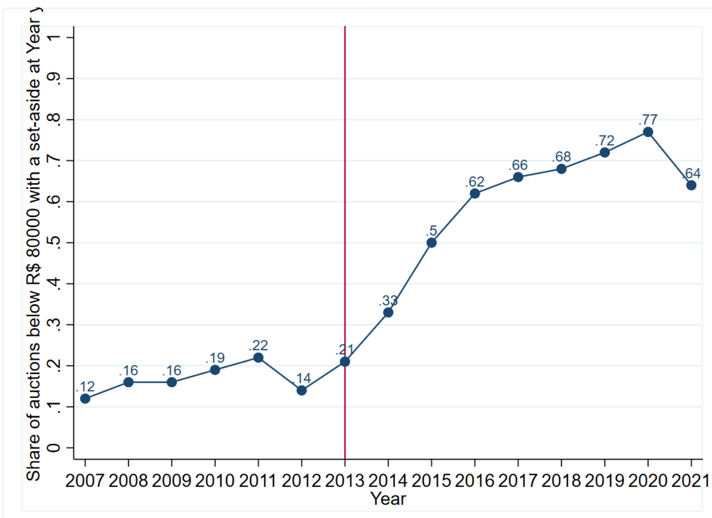
World Bank (2012). *Scaling-up SME access to financial services in the developing world*. World Bank Group, Washington, D.C.

Figure 1: There is a pronounced increase in the frequency of set-aside for MSEs both across auctions and across years

(a) Frequency of set-aside auctions for MSEs as a function of the value of the lot being auctioned



(b) Frequency of set-aside auctions for MSEs across years



Notes: This figure describes the two sources of variation in set-aside policies exploited by this research project. Graph (a) shows the share of set aside (SA) auctions for Micro and Small Enterprises (MSEs) as a function of the value of the lot for the period 2014 and 2021. It documents a pronounced increase in the frequency of auctions with SA for those lots below 80 thousand reais, the cutoff for mandatory SA. For those auctions below the 80 thousand reais cutoff, Graph (b) displays the share of auctions with SA across years between 2007 and 2021. It shows a substantial increase in the share of auctions with SA auctions a

Table 1: Main results from the Conventional RDD: SA auctions for MSEs cause a substantial increase in the likelihood of an MSE winning an auction and a moderate increase in procurement prices

	(1) MSE Winner	(2) Log(Prices)	(3) Log(Quality-Adjusted Prices)
Panel A: Intention-to-Treat			
RD estimate	-0.092	-0.062	-0.063
	[0.013]***	[0.015]***	[0.018]***
Panel B: First-Stage			
RD estimate	-0.512	-0.512	-0.522
	[0.013]***	[0.012]***	[0.013]***
Panel C: Second-Stage			
RD estimate	0.179	0.121	0.123
	[0.024]***	[0.030]***	[0.037]***
Bandwidth	27034	24006	22487
Observations	521922	516578	479351
N below	366293	362563	336639
N above	155629	154015	142712
Num. of clusters	3714	3677	3652
Outcome mean (control)	0.753		

Note: This table displays the treatment effect of SA auctions for Micro and Small Enterprises (MSEs) estimated by a Donut Regression Discontinuity Design (RDD) specification using procurement data from 2014 to 2021 and a donut hole between BRL 78000 and BRL 82600. All specifications report a bias-corrected RDD coefficient estimated using a linear polynomial, triangular kernel, and MSERD bandwidth selection. Panel A displays results from the Intention-to-Treat specification: sharp RDD estimates of the impact of the cutoff indicator on each of the outcome variables specified in the columns. Panel B displays results from the First-Stage specification: sharp RDD estimates of the impact of the cutoff indicator on the treatment variable within each of the RD bandwidths defined in Panel A. Panel C displays results from the Second-Stage specification: fuzzy RDD estimates of the impact of the treatment variable on each of the outcome variables specified in the columns. Column (1) shows RD estimates using the likelihood of an MSE winning a procurement auction as the dependent variable. Columns (2) and (3) replicate Column (1) but use Log(Prices) (demeaned within product-year) and value for money as dependent variables, respectively. We estimate the following *First-Stage* regression model to quantify the effect of the treatment cutoff on the treatment variable $Set\_Aside_i = \delta + \gamma \cdot \mathbf{1}(Value_{j(i)} < 80000) + f(Value_{j(i)}) + \epsilon_i$, where i denotes an auction and j is a tender that includes auction i. Set_Aside_i is an indicator taking value 1 when auction i is a SA for MSEs. $f(Value_{j(i)})$ is the RD polynomial. $Value_{j(i)}$ is the value of all products in auctions that belong to tender j. We estimate the following *Intention-to-Treat* regression model to quantify the effect of the treatment cutoff on our outcome variables $y_i = \alpha + \beta \cdot Set\_Aside_i + f(Value_{j(i)}) + \epsilon_i$, where y_i is the procurement outcome of auction i. The *Second-Stage* RD coefficient in Panel C is the ratio between the *Intention-to-Treat* coefficient in Panel A β and the *First-Stage* coefficient γ in Panel B. We display standard errors (SEs) clustered at the Purchasing Unit (PU) level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

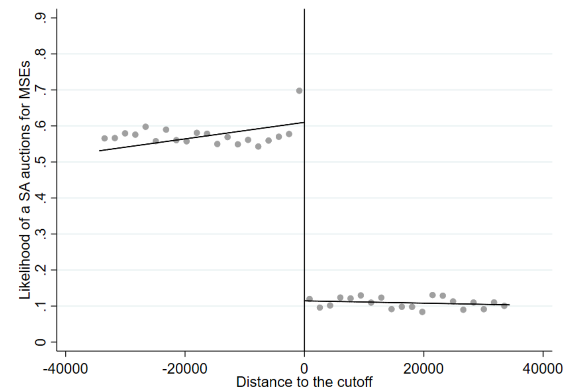
Table 2: The main results from the conventional RDD hold with a one-sided donut RDD

	(1) MSE Winner	(2) Log(Prices)	(3) Log(Quality-Adjusted Prices)
Panel A: Intention-to-Treat			
RD estimate	-0.099	-0.063	-0.052
	[0.012]***	[0.010]***	[0.011]***
Panel B: First-Stage			
RD estimate	-0.490	-0.478	-0.478
	[0.013]***	[0.012]***	[0.012]***
Panel C: Second-Stage			
RD estimate	0.201	0.133	0.109
	[0.023]***	[0.023]***	[0.023]***
Bandwidth	32972	28708	30171
Donut hole	[77500;80000]	[77500;80000]	[77500;80000]
Observations	1002873	992357	932967
N below	784848	776576	731339
N above	218025	215781	201628
Num. of clusters	3071	3006	3028
Outcome mean (control)	0.766		

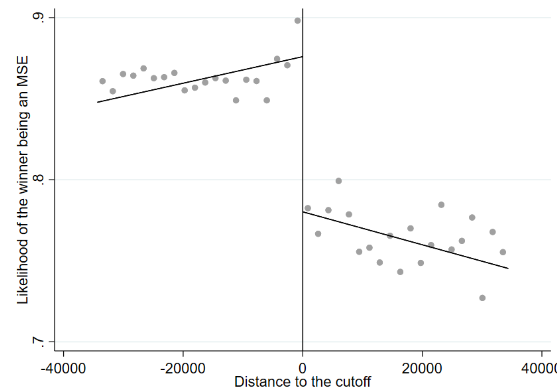
Note: This table displays the treatment effect of SA auctions for Micro and Small Enterprises (MSEs) estimated by a Donut Regression Discontinuity Design (RDD) specification using procurement data from 2014 to 2021 and a one-sided donut hole between BRL 77500 and BRL 80000. All specifications report a bias-corrected RDD coefficient estimated using a linear polynomial, triangular kernel, and MSERD bandwidth selection. Panel A displays results from the Intention-to-Treat specification: sharp RDD estimates of the impact of the cutoff indicator on each of the outcome variables specified in the columns. Panel B displays results from the First-Stage specification: sharp RDD estimates of the impact of the cutoff indicator on the treatment variable within each of the RD bandwidths defined in Panel A. Panel C displays results from the Second-Stage specification: fuzzy RDD estimates of the impact of the treatment variable on each of the outcome variables specified in the columns. Column (1) shows RD estimates using the likelihood of an MSE winning a procurement auction as the dependent variable. Columns (2) and (3) replicate Column (1) but use Log(Prices) (demeaned within product-year) and value for money as dependent variables, respectively. We estimate the following *First-Stage* regression model to quantify the effect of the treatment cutoff on the treatment variable $Set\_Aside_i = \delta + \gamma \cdot \mathbf{1}(Value_{j(i)} < 80000) + f(Value_{j(i)}) + \epsilon_i$, where i denotes an auction and j is a tender that includes auction i. Set_Aside_i is an indicator taking value 1 when auction i is a SA for MSEs. $f(Value_{j(i)})$ is the RD polynomial. $Value_{j(i)}$ is the value of all products in auctions that belong to tender j. We estimate the following *Intention-to-Treat* regression model to quantify the effect of the treatment cutoff on our outcome variables $y_i = \alpha + \beta \cdot Set\_Aside_i + f(Value_{j(i)}) + \epsilon_i$, where y_i is the procurement outcome of auction i. The *Second-Stage* RD coefficient in Panel C is the ratio between the *Intention-to-Treat* coefficient in Panel A β and the *First-Stage* coefficient γ in Panel B. We display standard errors (SEs) clustered at the Purchasing Unit (PU) level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

Figure 2: Conventional RDD results: SA auctions for MSEs cause a substantial increase in the likelihood of an MSE winning an auction and a moderate increase in procurement prices and quality-adjusted prices

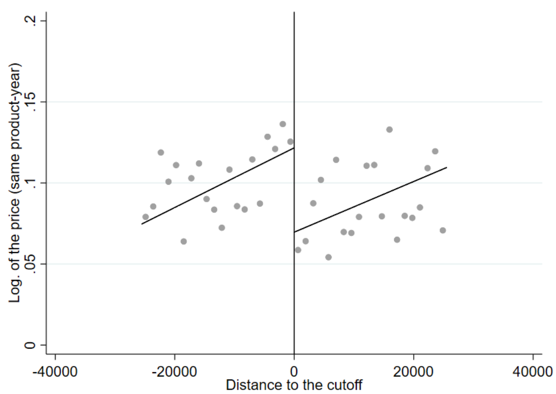
(a) First Stage, Set-Aside Auction



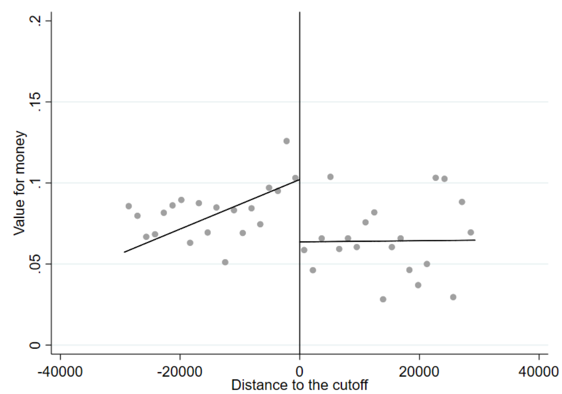
(b) Intention-to-Treat, MSE Winner



(c) Intention-to-Treat, Log-Price

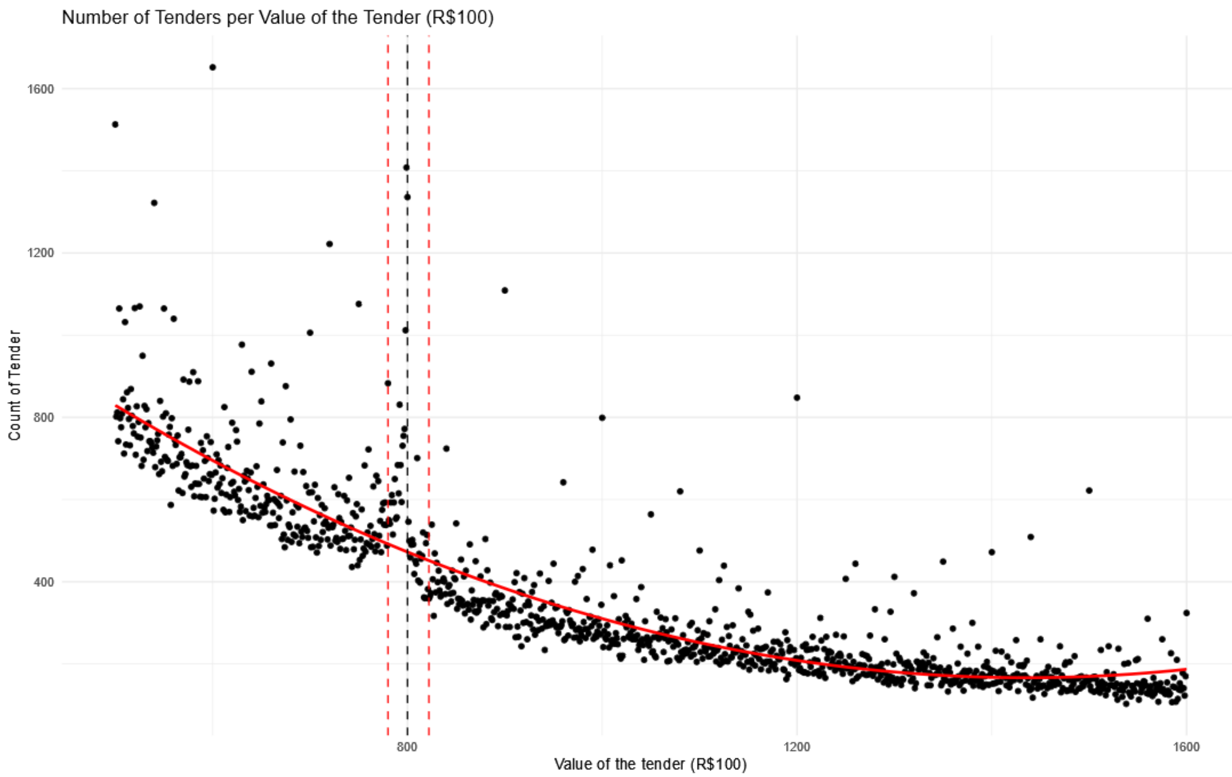


(d) Intention-to-Treat, Quality-Adjusted Prices



Notes: This figure shows four RD graphs using procurement data estimated by a Regression Discontinuity Design (RDD) specification using procurement data from 2014 to 2021. All specifications report a bias-corrected RDD coefficient estimated using a linear polynomial, triangular kernel, and MSERD bandwidth selection. Graph **(a)** shows the RDD graph describing one of our three First-Stage (FS) graphs: sharp RDD estimates of the impact of the cutoff indicator on the treatment variable within the bandwidths defined using the MSE Winner indicator as an outcome. The FS graphs using Log-Price and Quality-Adjusted Prices as outcomes are similar to graph **(a)** and not reported. Graph **(b)** shows the RDD graph describing the Intention-to-Treat (ITT) graph with the MSE Winner indicator as an outcome: sharp RDD estimates of the effect of the cutoff indicator on the likelihood of the auction winner being an MSE. Graphs **(c)** and **(d)** replicate Graph (b) but use Log-Price (demeaned within product-year) and Quality-Adjusted Prices as dependent variables, respectively. We estimate the following FS regression model to quantify the effect of the treatment cutoff on the treatment variable $Set\_Aside_i = \delta + \gamma \cdot \mathbf{1}(Value_{j(i)} < 80000) + f(Value_{j(i)}) + \epsilon_i$, where i denotes an auction and j is a tender that includes auction i. Set_Aside_i is an indicator taking value 1 when auction i is a SA for MSEs. $f(Value_{j(i)})$ is the RD polynomial. $Value_{j(i)}$ is the value of all products in auctions that belong to tender j. We estimate the following ITT regression model to quantify the effect of the treatment cutoff on our outcome variables $y_i = \alpha + \beta \cdot Set\_Aside_i + f(Value_{j(i)}) + \epsilon_i$, where y_i is the procurement outcome of auction i. We cluster the Standard Errors (SEs) at the Purchasing Unit (PU) level.

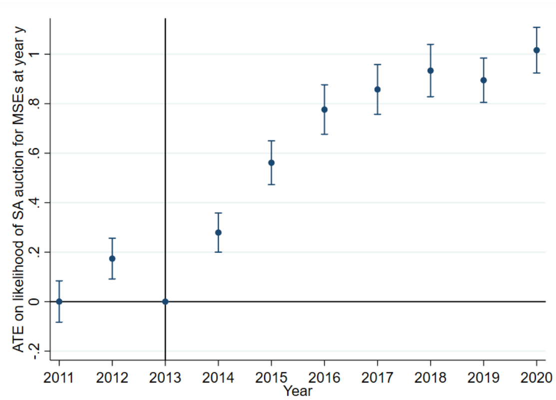
Figure 3: There is a moderate excess of mass below the cutoff for mandatory SA auctions



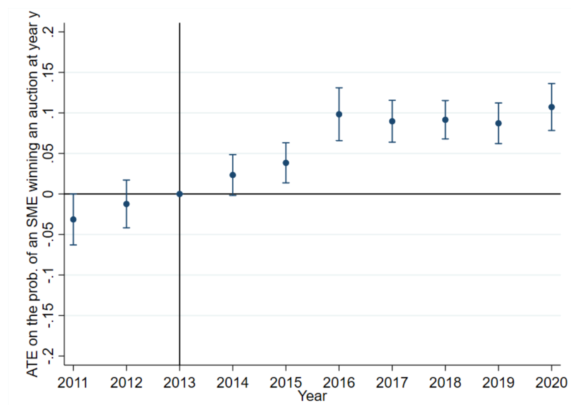
Notes: This figure shows the frequency distribution of procurement tenders by their total value, with particular focus around the BRL 80,000 threshold that determines whether set-aside (SA) auctions for Micro and Small Enterprises (MSEs) are mandatory. The x-axis represents the tender value in hundreds of Brazilian Reals (BRL), while the y-axis shows the frequency count. The vertical black dashed line indicates the BRL 80,000 threshold. The moderate excess mass observed just below the cutoff suggests strategic manipulation of tender values by purchasing agencies to fall below the threshold, potentially to maintain flexibility in procurement methods.

Figure 4: Main results from the DiD specification: A reform significantly increasing the number of SA auctions for MSEs leads to a moderate increase in procurement prices, comparable in magnitude to the estimate obtained from the RDD.

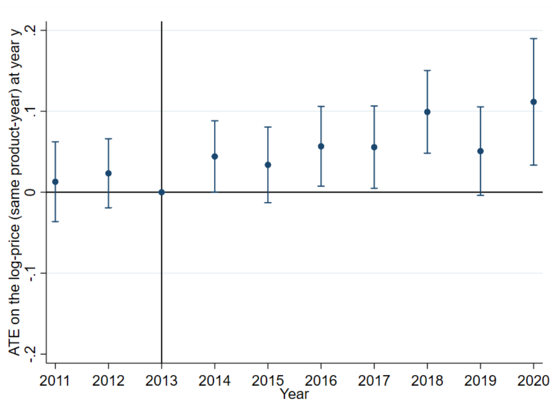
(a) Likelihood of SA auction for MSEs



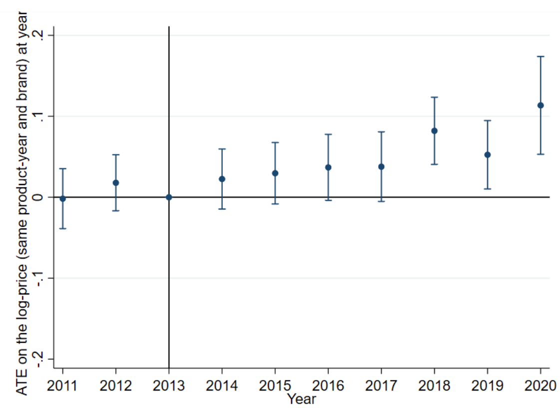
(b) Likelihood of a MSE winning the auction



(c) Log-price, product-year FEs



(d) Log-price, product-year FEs and brand FEs



Notes: This figure displays estimates of the effect of the reform that made SA auctions for MSEs mandatory for purchases below BRL 80000 on the auction-level procurement outcomes using a difference-in-differences (DiD) specification and the DiD sample described in subsubsection 4.1.2. Respectively, Graphs **(a)**, **(b)**, and **(c)** display the effects of the reform on the likelihood of an SA auction for MSEs, the likelihood of an SME winning the auction, and the log-price (same product category) using a DiD specification with product-year fixed effects (FEs) and purchasing agency FEs. Graph **(d)** replicates the specification in Graph **(c)** but adding brand FEs. Then, the DiD regression model reported in Graph **(d)** has the form $y_{i,t} = \alpha_{p(i),t} + \alpha_a + \alpha_{q(i)} + \beta \cdot Intensity_{a(i)} \cdot Post_t + \epsilon_{i,t}$ where i denotes a procurement auction, and t time, which is the year. $\alpha_{p(i),t}$ captures product-year fixed effects (FEs), $\alpha_{a(i)}$ purchasing agency FEs, and $\alpha_{q(i)}$ brand FEs, which we use as a proxy of quality. $Intensity_{a(i)}$ is our continuous treatment variable: the share of non-SA auctions made by agency a in the sample of purchases below BRL 80000 during the pre-treatment (2011-2013), when SA auctions for MSEs where voluntary. $Post_t = \mathbb{1}(t > 2013)$ is an indicator variable that equals one after 2013, the first year before the 2014 reform that made SA auctions for MSEs mandatory for purchases below BRL 80000. $Value_{b(i)}$ represents the value of the bundle that includes auction i, which is the running variable of the RDD described in Subsection 3.2. We display clustered standard errors (SEs) at the establishment level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

Table 3: Main results from the DiD specification: A reform significantly increasing the number of SA auctions for MSEs leads to a moderate increase in procurement prices, comparable in magnitude to the estimate obtained from the RDD.

	(1)	(2)	(3)	(4)	(5)
Panel A: Reduced-Form					
Intensity*Post	0.037	0.052	0.046	0.041	0.040
	[0.017]**	[0.016]***	[0.014]***	[0.015]***	[0.013]***
Observations	3360679	3311868	3311845	3311845	3103146
Num. of clusters	2422	2420	2397	2397	2389
R-squared	0.809	0.838	0.841	0.842	0.875
Product FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Product × Year FEs	No	Yes	Yes	Yes	Yes
Agency FEs	No	No	Yes	Yes	Yes
Run. Var. × Year FEs	No	No	No	Yes	Yes
Brand FEs	No	No	No	No	Yes
Panel B: First-Stage					
Intensity*Post	0.576	0.605	0.568	0.568	0.571
	[0.030]***	[0.029]***	[0.030]***	[0.030]***	[0.029]***
Observations	3361354	3312352	3312323	3312323	3103637
Num. of clusters	2426	2426	2400	2400	2395
R-squared	0.396	0.452	0.549	0.549	0.581
Product FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Product × Year FEs	No	Yes	Yes	Yes	Yes
Agency FEs	No	No	Yes	Yes	Yes
Run. Var. × Year FEs	No	No	No	Yes	Yes
Brand FEs	No	No	No	No	Yes

Note: This table displays estimates of the effect of the reform that made SA auctions for MSEs mandatory for purchases below BRL 80000 on the auction-level procurement outcomes using a difference-in-differences (DiD) specification and the DiD sample described in **??**. Respectively, Panels A and B display the effects of the reform on the log price (same product category) and the likelihood of an SA auction for MSEs. Column (1) presents the results of a DiD specification with product fixed-effects (FEs) and year FEs. Columns (2) to (5) include, respectively, product-year FEs, agency FEs, year FEs interacted with the bundle value, and brabd FEs to the specification in Column (1). The DiD regression model reported in Column (5) has the form $y_{i,t} = \alpha_{p(i),t} + \alpha_{a(i)} + \alpha_{b(i)} + \beta \cdot Intensity_{a(i)} \cdot Post_t + \alpha_t \cdot Value_{j(i)} + \epsilon_{i,t}$ where i denotes a procurement auction, and t time, which is the year. $\alpha_{p,t}$ captures product-year fixed effects (FEs), α_a purchasing agency FEs, and α_b brand FEs. $Intensity_a$ is our continuous treatment variable: the share of non-SA auctions made by agency a in the sample of purchases below BRL 80000 during the pre-treatment (2011-2013), when SA auctions for MSEs where voluntary. $Post_t = \mathbb{1}(t > 2013)$ is an indicator variable that equals one after 2013, the first year before the 2014 reform that made SA auctions for MSEs mandatory for purchases below BRL 80000. $Value_{j(i)}$ represents the value of the bundle j that includes auction i, which is the running variable of the RDD described in Subsection 3.2. We display clustered standard errors (SEs) at the establishment level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

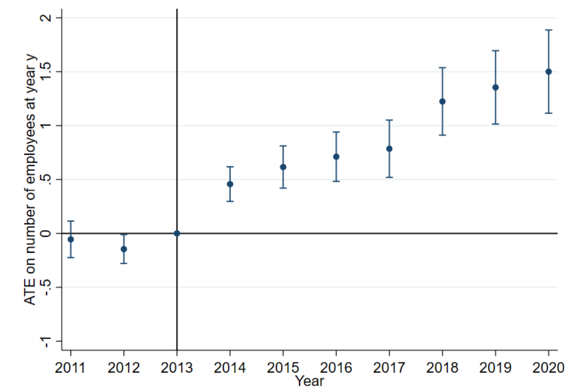
Table 4: Main results: A reform substantially increasing the number of SA auctions for MSEs causes a moderate increase in the number of employees and average monthly earnings of GSs in comparison to non-GSs

	(1)	(2)	(3)	(4)	(5)
Panel A: Number of employees					
Government Supplier*Post	0.831	0.833	0.939	0.947	0.933
	[0.107]***	[0.107]***	[0.100]***	[0.100]***	[0.099]***
Observations	540756	540756	540756	540756	540742
Num. of clusters	63280	63280	63280	63280	63280
R-squared	0.001	0.001	0.600	0.603	0.609
Outcome mean (pre-treatment)	10.171	10.171	10.171	10.171	10.171
Outcome S.D. (pre-treatment)	11.985	11.985	11.985	11.985	11.985
Year FEs	No	Yes	Yes	Yes	Yes
Establishment FEs	No	No	Yes	Yes	Yes
State $\times$ Year FEs	No	No	No	Yes	Yes
Industry $\times$ Year FEs	No	No	No	Yes	Yes
State $\times$ Industry $\times$ Year FEs	No	No	No	No	Yes
Panel B: Avg. monthly earnings					
Government Supplier*Post	51.028	48.504	42.352	43.082	43.178
	[3.268]***	[3.249]***	[2.817]***	[2.764]***	[2.760]***
Observations	540756	540756	540756	540756	540742
Num. of clusters	63280	63280	63280	63280	63280
R-squared	0.115	0.153	0.770	0.774	0.776
Outcome mean (pre-treatment)	1102.494	1102.494	1102.494	1102.494	1102.494
Outcome S.D. (pre-treatment)	466.964	466.964	466.964	466.964	466.964
Year FEs	No	Yes	Yes	Yes	Yes
Establishment FEs	No	No	Yes	Yes	Yes
State $\times$ Year FEs	No	No	No	Yes	Yes
Industry $\times$ Year FEs	No	No	No	Yes	Yes
State $\times$ Industry $\times$ Year FEs	No	No	No	No	Yes

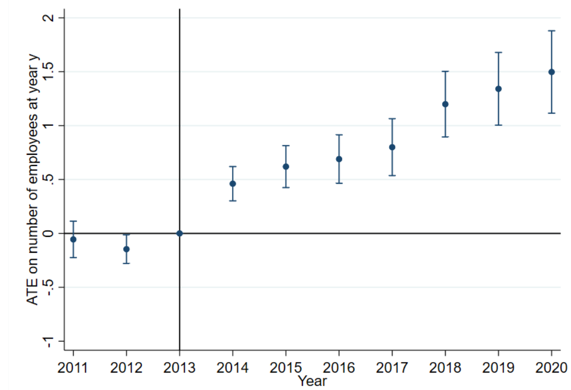
Note: This table displays estimates of the effect of the reform that made SA auctions for MSEs mandatory for tenders below BRL 80000 on the establishment-level outcomes of MSEs using a difference-in-differences (DiD) specification and yearly data between 2011 and 2020 from the matched sample. We described how we generated the matched sample in **??** and Section B. Respectively, Panels A and B display the effects of the reform on the number of employees and average monthly earnings. Column (1) presents the results of a classic DiD specification without fixed effects (FEs). Respectively, Columns (2) to (5) include, respectively, year FEs, establishment FEs, industry-time FEs and state-time FEs, and state-industry-time FEs to the specification in Column (1). The DiD regression model reported in Column (5) has the form $y_{e,t} = \alpha_e + \alpha_{i,s,t} + \alpha_{s,t} + \beta \cdot GS_e \cdot Post_t + \epsilon_{e,t}$ where e denotes an establishment, and t time, which is the year. α_e captures establishment fixed effects (FEs), $\alpha_{s,i,t}$ stete-industry-time FEs. GS_e is an indicator variable equal to if the establishment e participated in a procurement auction in the pre-treatment period (2011-2013). $Post_t = \mathbb{1}(t > 2013)$ is an indicator variable that equals one after 2013, the first year before the 2014 reform that made SA auctions for MSEs mandatory for procurement tenders below BRL 80000. We display clustered standard errors (SEs) at the establishment level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

Figure 5: Main results: A reform substantially increasing the number of SA auctions for MSEs cause a moderate increase in the number of employees and average monthly earnings of GSs in comparison to non-GSs

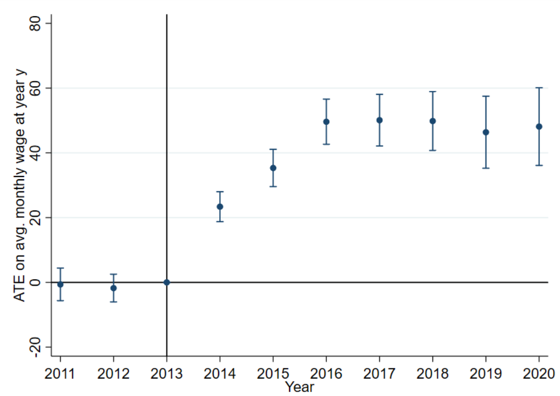
(a) TWFE, Number of employees



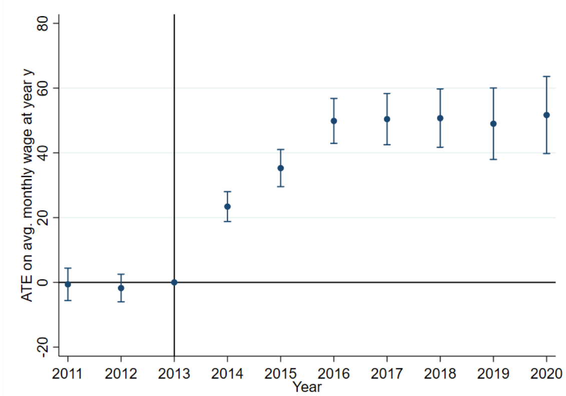
(b) State-Industry-Time FEs, Number of employees



(c) TWFE, Average monthly earnings



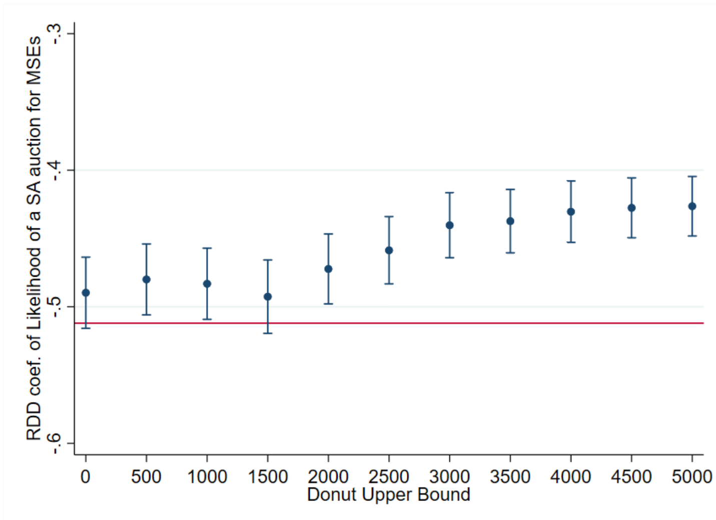
(d) State-Industry-Time FEs, Average monthly earnings



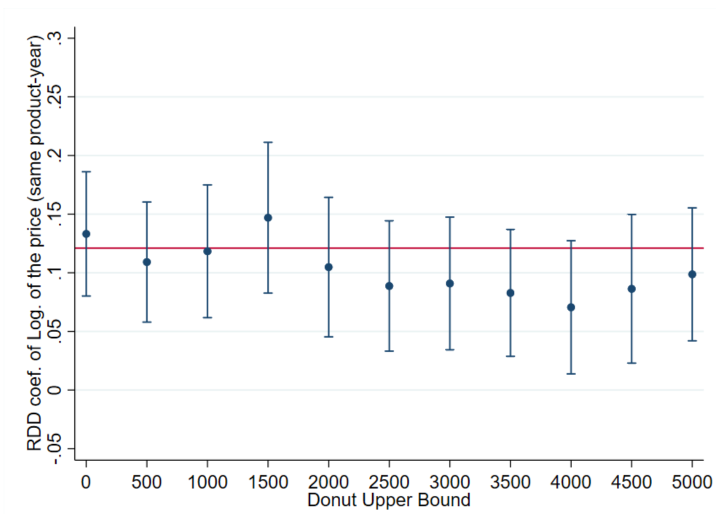
Notes: This figure shows four DiD graphs using procurement data estimated by a TWFE specification using establishment data from 2011 to 2020. Each Graph shows estimates of the yearly effects of the reform that made SA auctions for MSEs mandatory for tenders below BRL 80000 on the establishment-level outcomes of MSEs. We described how we generated the post-matching sample in **??** and Section B. Respectively, Graphs (a) and (b) display the dynamic effects of the reform on the number of employees and average monthly earnings using the TWFE model $y_{e,t} = \alpha_e + \alpha_t + \beta \cdot GS_e \cdot Post_t + \epsilon_{e,t}$ where e denotes an establishment, and t time, which is the year. α_e captures establishment fixed effects (FEs), and α_t time FEs. $GS_{f(e)}$ is an indicator variable equal to one if the establishment e participated in a procurement auction in the pre-treatment period (2011-2013). $Post_t = \mathbb{1}(t > 2013)$ is an indicator variable that equals one after 2013, the first year before the 2014 reform that made SA auctions for MSEs mandatory for procurement tenders below BRL 80000. Respectively, Graphs **(c)** and **(d)**, mirror Graphs **(a)** and **(b)** but adding State-Industry-Time FEs $\alpha_{s(e),i(e),t}$ to the TWFE regression model. We display clustered standard errors (SEs) at the establishment level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

Figure 6: Main results from the Conventional RDD are robust to different donuts

(a) Likelihood of SA auction for SMEs



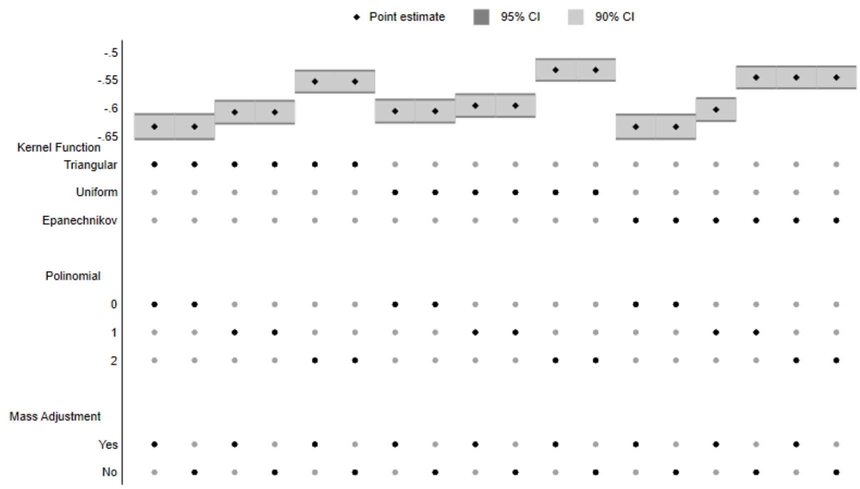
(b) Log-price (same product)



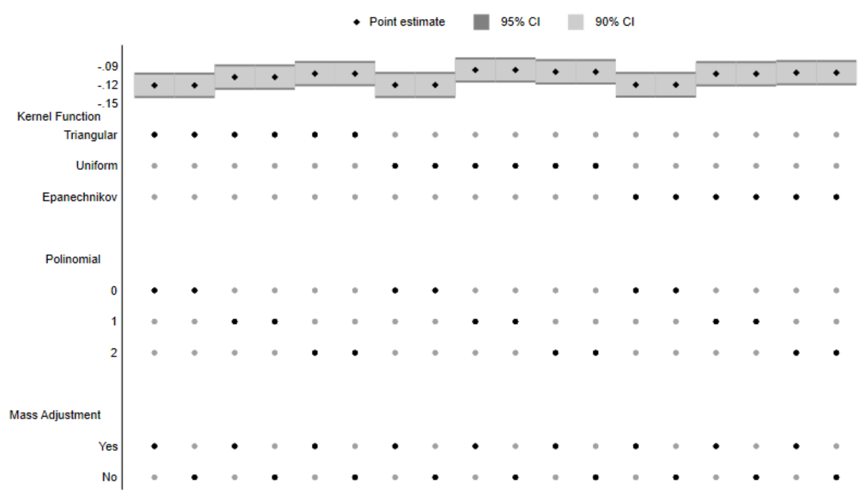
Notes: This figure presents a robustness test for our one-sided Donut RDD specification from Table 2. We assess robustness by systematically removing observations above the threshold, increasing the Donut's upper bound (UB) from BRL 80,000 (one-sided) to BRL 85,000 (two-sided with UB = 5,000) in BRL 500 increments while keeping the lower bound (LB) fixed at BRL 77,500. Graph (a) shows the Donut RDD estimates for the likelihood of SA auctions for SMEs as we adjust the UB, while Graph (b) presents the estimates for the log-price. The x-axis represents the increasing bandwidth of excluded observations above the threshold, measured as the distance from BRL 80,000. The coefficient at x = 0 corresponds to the one-sided Donut RDD estimate from Table 2, while the red line marks the Conventional RDD estimate from Table 1. The stability of the coefficient estimates across specifications suggests that manipulation at the threshold does not drive our results. Each point estimate includes 95% confidence intervals (shown as vertical lines).

Figure 7: This figure demonstrates that the Conventional RDD estimates remain stable across nearly all combinations of methodological choices.

(a) Likelihood of SA auction



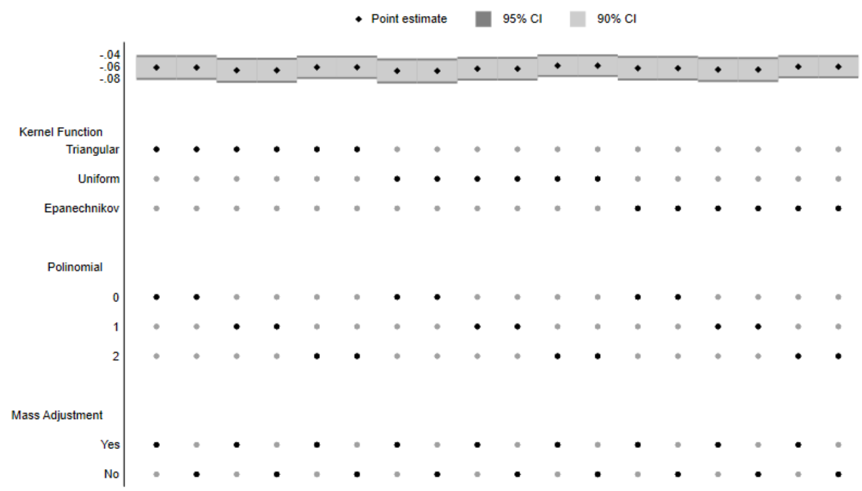
(b) Likelihood of SME Winner



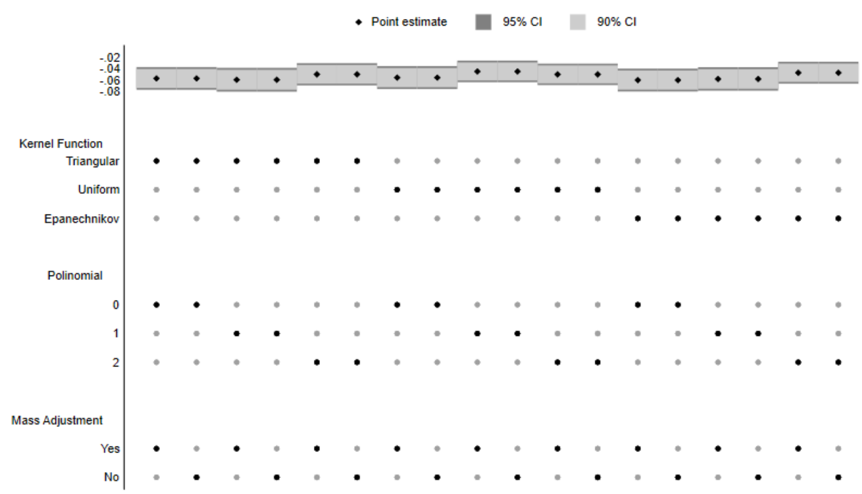
Notes: This figure presents a robustness test for our RD specification from Figure 2. We evaluate robustness by systematically varying key specification characteristics. First, we test three different kernel functions (Triangular, Uniform, and Epanechnikov). Next, we vary the polynomial order used for approximation. Finally, we toggle the mass adjustment on and off. This process results in a total of 18 different estimates, each altering one of these characteristics. Figure 7a displays the robustness results for the likelihood of a set-aside auction, while Figure 7b presents the results for the likelihood of an SME winning the auction.

Figure 8: This figure demonstrates that the Conventional RDD estimates remain stable across nearly all combinations of methodological choices.

(a) Log Price Residual



(b) Value for Money



Notes: This figure presents a robustness test for our RD specification from Figure 2. We evaluate robustness by systematically varying key specification characteristics. First, we test three different kernel functions (Triangular, Uniform, and Epanechnikov). Next, we vary the polynomial order used for approximation. Finally, we toggle the mass adjustment on and off. This process results in a total of 18 different estimates, each altering one of these characteristics. Figure 8a displays the robustness results for the log price residual, while Figure 8b presents the results for value for money.

Table 5: Robustness: Estimates of the impact of the reform that substantially increased the number of SA auctions for SMEs on the number of employees and average monthly wage remain similar with a demanding parametric specifications

	(1)	(2)	(3)	(4)
Panel A: Number of employees				
Government Supplier*Post	0.939	0.884	0.896	0.766
	[0.100]***	[0.101]***	[0.101]***	[0.096]***
Observations	540756	540756	540734	540734
Num. of clusters	63280	63280	63280	63280
R-squared	0.600	0.603	0.603	0.662
Outcome mean (pre-treatment)	10.171	10.171	10.171	10.171
Outcome S.D. (pre-treatment)	11.985	11.985	11.985	11.985
Year FEs	Yes	Yes	Yes	Yes
Establishment FEs	Yes	Yes	Yes	Yes
Spatial controls × Year FEs	No	Yes	Yes	Yes
Pre-treatment controls × Year FEs	No	No	Yes	Yes
Pre-treatment outcomes × Year FEs	No	No	No	Yes
Panel B: Avg. monthly earnings				
Government Supplier*Post	42.352	36.038	37.512	37.322
	[2.817]***	[2.840]***	[2.827]***	[2.754]***
Observations	540756	540756	540734	540734
Num. of clusters	63280	63280	63280	63280
R-squared	0.770	0.773	0.775	0.791
Outcome mean (pre-treatment)	1102.494	1102.494	1102.494	1102.494
Outcome S.D. (pre-treatment)	466.964	466.964	466.964	466.964
Year FEs	Yes	Yes	Yes	Yes
Establishment FEs	Yes	Yes	Yes	Yes
Spatial controls × Year FEs	No	Yes	Yes	Yes
Pre-treatment controls × Year FEs	No	No	Yes	Yes
Pre-treatment outcomes × Year FEs	No	No	No	Yes

Note: This table displays estimates of the effect of the reform that made SA auctions for MSEs mandatory for tenders below BRL 80000 on the establishment-level outcomes of MSEs using a demanding difference-in-differences (DiD) specification with pre-treatment control variables interacted with the time indicators and yearly data between 2011 and 2020 from the post-matching sample. We described how we generated the post-matching sample in **??** and Section B. Respectively, Panels A and B display the effects of the reform on the number of employees and average monthly earnings. Column (1) presents the results from the TWFE regression model with establishment and year FEs displayed in Column (3) of Table 4. Respectively, Columns (2) to (4) include year FEs interacted with pre-treatment outcomes $(\mathbf{y}_{2011,e}, \mathbf{y}_{2011,e}, \mathbf{y}_{2013,e})$, pre-treatment controls $(\mathbf{x}_{2011,e}, \mathbf{x}_{2011,e}, \mathbf{x}_{2013,e})$, and spatial controls $x_{m(e)}$ to the specification in Column (1). The DiD regression model reported in Column (4) has the form $y_{e,t} = \alpha_e + \alpha_t + \beta \cdot GS_{f(e)} \cdot Post_t + \gamma_t \cdot \mathbf{x}_e + \epsilon_{e,t}$, where e denotes an establishment, f firm, and t time, which is the year. α_e captures establishment fixed effects (FEs), α_t year FEs. $GS_{f(e)}$ is an indicator variable equal to if the firm f of which establishment e belongs participated in a procurement auction in the pre-treatment period (2011-2013). $Post_t = \mathbb{1}(t > 2013)$ is an indicator variable that equals one after 2013, the first year before the 2014 reform that made SA auctions for MSEs mandatory for procurement tenders below BRL 80000. $\mathbf{x}_e$ is a vector with our pre-treatment controls. We list the control variables in the vector x_e in **??** and describe them in Table B2 of Appendix B. $\gamma_t \cdot \mathbf{x}_e$ are time-specific slopes controlling for the influence of pre-treatment outcomes, establishments' characteristics, and features of the municipality of establishments' location on the change in their outcomes across time. We display clustered standard errors (SEs) at the firm level between squared brackets. Coefficients significantly different from zero at 99% (***), 95% (**), and 90% (*) confidence levels.

Table A1: List of Items

item description	unit
Petrol	liter
Banana	kg
Watermelon	kg
Mineral water 500ml	unit
Protective eyewear	unit
Nails	kg
USB flash drive, 16GB	unit
Ibuprofen, 600 mg	pill
Rubbing alcohol	liter
Toilet fixture	unit
Nasal Cannula	unit
20W Fluorecent lamp	unit
Bleach	liter
35mm Padlock	unit
Digital Thermometer	unit

Appendix

A Background

Table A2: Government Agencies

Name of Government Agency	Classification
Universidade Federal do Rio Grande do Sul	Education
Universidade Federal do Pará	Education
Universidade Federal de Pernambuco	Education
Hospital Universitario UFSC	Hospitals
Hospital Universitario Antonio Pedro (UFF/RJ)	Hospitals
Hospital Universitario Gaffree e Guinele (UNIRIO)	Hospitals
Grupamento de Apoio de São José dos Campos	Armed Forces
Grupamento de Apoio de Brasilia	Armed Forces
14 Grupo de Artilharia de Campanha	Armed Forces
Comissao Nacional de Energia Nuclear	Other
Governo do Estado do Ceara	Other
Departamento de Logistica em Saude	Other

A.1 Regulation

The identification of both estimation procedures relies on the exogenous policy change regarding the restriction of public auctions below BRL 80,000 for SMEs. Initially, as set by Art. 47 and Art. 48 of Supplementary Law 147 of 2014, set-aside lots could be created for purchases below BRL 80,000 or up to 25% of annual purchases. Moreover, these set-aside lots could be allocated for SMEs up to 25% of the total value procured in one year. The two paragraphs related to these changes are reported below in Figure A1.

Art. 47. Nas contratações públicas da administração direta e indireta, autárquica e fundacional, federal, estadual e municipal, deverá ser concedido tratamento diferenciado e simplificado para as microempresas e empresas de pequeno porte objetivando a promoção do desenvolvimento econômico e social no âmbito municipal e regional, a ampliação da eficiência das políticas públicas e o incentivo à inovação tecnológica. (Redação dada pela Lei Complementar nº 147, de 2014) (Vide Lei nº 14.133, de 2021

Parágrafo único. No que diz respeito às compras públicas, enquanto não sobrevier legislação estadual, municipal ou regulamento específico de cada órgão mais favorável à microempresa e empresa de pequeno porte, aplica-se a legislação federal. (Incluído pela Lei Complementar nº 147, de 2014)

~~Art. 48. Para o cumprimento do disposto no art. 47 desta Lei Complementar, a administração pública poderá realizar processo licitatório:~~

Art. 48. Para o cumprimento do disposto no art. 47 desta Lei Complementar, a administração pública: (Redação dada pela Lei Complementar nº 147, de 2014) (Vide Lei nº 14.133, de 2021

~~I - destinado exclusivamente à participação de microempresas e empresas de pequeno porte nas contratações cujo valor seja de até R$ 80.000,00 (oitenta mil reais);~~

I - deverá realizar processo licitatório destinado exclusivamente à participação de microempresas e empresas de pequeno porte nos itens de contratação cujo valor seja de até R$ 80.000,00 (oitenta mil reais); (Redação dada pela Lei Complementar nº 147, de 2014)

~~II - em que seja exigida dos licitantes a subcontratação de microempresa ou de empresa de pequeno porte, desde que o percentual máximo do objeto a ser subcontratado não exceda a 30% (trinta por cento) do total licitado;~~

II - poderá, em relação aos processos licitatórios destinados à aquisição de obras e serviços, exigir dos licitantes a subcontratação de microempresa ou empresa de pequeno porte; (Redação dada pela Lei Complementar nº 147, de 2014)

~~III - em que se estabeleça cota de até 25% (vinte e cinco por cento) do objeto para a contratação de microempresas e empresas de pequeno porte, em certames para a aquisição de bens e serviços de natureza divisível.~~

III - deverá estabelecer, em certames para aquisição de bens de natureza divisível, cota de até 25% (vinte e cinco por cento) do objeto para a contratação de microempresas e empresas de pequeno porte. (Redação dada pela Lei Complementar nº 147, de 2014)

Figure A1: Supplementary Law 147/2014 – Original Regulation

Later in 2014, these two paragraphs were overturned by Supplementary Law 147 of 2014, which introduced a change that made set-aside auctions **mandatory** for purchases below BRL 80,000 and eliminated the annual restriction on the share of set-aside lots. The paragraphs related to this change are reported below in Figure A2.

"Art. 47. Nas contratações públicas da administração direta e indireta, autárquica e fundacional, federal, estadual e municipal, deverá ser concedido tratamento diferenciado e simplificado para as microempresas e empresas de pequeno porte objetivando a promoção do desenvolvimento econômico e social no âmbito municipal e regional, a ampliação da eficiência das políticas públicas e o incentivo à inovação tecnológica.

Parágrafo único. No que diz respeito às compras públicas, enquanto não sobrevier legislação estadual, municipal ou regulamento específico de cada órgão mais favorável à microempresa e empresa de pequeno porte, aplica-se a legislação federal." (NR)

"Art. 48. Para o cumprimento do disposto no art. 47 desta Lei Complementar, a administração pública:

I - deverá realizar processo licitatório destinado exclusivamente à participação de microempresas e empresas de pequeno porte nos itens de contratação cujo valor seja de até R$ 80.000,00 (oitenta mil reais);

II - poderá, em relação aos processos licitatórios destinados à aquisição de obras e serviços, exigir dos licitantes a subcontratação de microempresa ou empresa de pequeno porte;

III - deverá estabelecer, em certames para aquisição de bens de natureza divisível, cota de até 25% (vinte e cinco por cento) do objeto para a contratação de microempresas e empresas de pequeno porte.

Figure A2: Supplementary Law 147/2014 – Revised Regulation

B Data

B.1 Variables definitions and Descriptive Statistics

Table B1: Description of the variables in the RDD estimating sample

Variable	Description	Source
Panel A: Outcomes		
=1 if MSE winner	Indicator variable equal to 1 if the auction winner is an MSE, and 0 otherwise.	CNET
log-price	Residual from the regression of the logarithm of the unit price on product category indicators.	CNET
quality-adjusted log-price	Residual from the regression of the logarithm of the unit price on product category and product brand indicators.	CNET
Panel B: Treatment		
=1 if auction is SA for MSEs	Indicator variable equal to 1 if the bundle value is below BRL 80,000, and 0 otherwise.	CNET
Panel C: RDD controls		
=1 if auction is below cutoff of mandatory SA for MSEs	Indicator variable equal to 1 if the bundle value is below BRL 80,000, and 0 otherwise.	CNET
bundle value	Value of the bundle to which the auction belongs. The bundle and auction coincide if the bundle contains a single lot.	CNET

Notes: Elaborated by the authors.

Table B2: Description of the variables in the DiD estimating sample

Variable	Description	Source
Panel A: Outcomes		
Number of employees	Number of active job relations at the establishment level at December 31	RAIS
Average monthly earnings	Average monthly earnings of the active employees within an establishment at December 31	RAIS
Total yearly payroll	Total yearly payroll of the establishment	RAIS
Average monthly working hours	Average hours worked by the active employees within an establishment at December 31	RAIS
Panel B: Treatment		
=1 if estab. is a Government Supplier (GS)	Indicator variable equal to one if the establishment participated in a procurement auction during the pre-treatment (2011-2013)	CNET
Panel C: Employee controls		
Percentage of female employees	Percentage of female employees within an establishment among those active at December 31	RAIS
Percentage of employees with a college degree	Percentage of workers with college or more within an establishment among those active at December 31	RAIS
Average tenure of employees	Average monthly tenure of the active employees within an establishment at December 31	RAIS
Average age of employees	Average age of active employees within an establishment level at December 31	RAIS
Panel D: Employer controls		
Age of the establishment	Age of the establishment	CNPJ Aberto
Panel E: Spatial controls		
GDP per capita of 2000	GDP per capita of the municipality in which the establishment is located	IBGE
Gini index of 2000	Gini index of the GDP per capita of the municipality in which the establishment is located	IBGE
Total population	Population of the municipality in which the establishment is located (in 1000s)	IBGE
Perc. of the GDP from manufacturing	Percentage of the GDP of 2000 from manufacturing of the municipality in which the establishment is located	IBGE
Deposit density	Total deposits divided by the area of the municipality in which the establishment is located (in BRL 1000s)	Estban
Bank density	Number of bank agencies divided by the area of the municipality in which the establishment is located	Estban

Notes: Elaborated by the authors.

Table B3: Descriptive statistics from the RDD estimating sample.

	Mean	Std. dev.	Min.	Max.
Panel A: Outcomes				
=1 if auction winner is an MSE	0.85	0.36	0.00	1.00
log-price	0.09	0.84	-3.71	3.29
quality-adjusted log-price	0.15	0.83	-9.41	9.39
Panel B: Treatment				
=1 if auction is a SA for MSEs	0.48	0.50	0.00	1.00
Panel C: RD variables				
bundle value	55453.91	31163.49	20000.00	140000.00
=1 if auction is below cutoff of mandatory SA for MSEs	0.79	0.41	0.00	1.00

Table B4: Descriptive statistics from the DiD estimating sample from the pre-treatment period.

	Mean	S.D.	Min.	Max.
Panel A: Outcomes (2011-13)				
Number of employees	10.17	11.99	1.00	768.00
Average monthly earnings	1102.07	466.60	526.53	3474.99
Average monthly payroll	8290.05	9838.98	0.00	60584.36
Average monthly working hours	43.36	2.50	1.00	44.00
Panel B: Treatment (2011-13)				
=1 if estab. is a Government Supplier (GS)	0.50	0.50	0.00	1.00
Panel C: Employees Characteristics (2011-13)				
Percentage of female employees	35.49	31.36	0.00	100.00
Percentage of employees with a college degree	6.93	16.30	0.00	100.00
Average tenure of employees	29.59	27.51	0.00	467.20
Average age of employees	33.58	6.48	16.00	99.00
Panel D: Establishment Characteristics (2011-13)				
Age of the estab.	6.31	8.04	0.00	54.00
=1 if estab. locates in the N region	0.09	0.28	0.00	1.00
=1 if estab. locates in the NE region	0.18	0.38	0.00	1.00
=1 if estab. locates in the SE region	0.38	0.49	0.00	1.00
=1 if estab. locates in the S region	0.23	0.42	0.00	1.00
=1 if estab. locates in the CO region	0.13	0.34	0.00	1.00
=1 if sector is agriculture or mining	0.04	0.19	0.00	1.00
=1 if sector is low-tech manufacturing	0.08	0.27	0.00	1.00
=1 if sector is high-tech manufacturing	0.05	0.22	0.00	1.00
=1 if sector is construction	0.06	0.24	0.00	1.00
=1 if sector is trade	0.00	0.07	0.00	1.00
=1 if sector is transport or utilities	0.06	0.24	0.00	1.00
=1 if sector is services	0.06	0.23	0.00	1.00
=1 if sector in others	0.64	0.48	0.00	1.00
Panel E: Characteristics of the munic. where the estab. locates				
GDP per capita of 2000	30.83	17.74	3.20	274.08
Gini index of 2000	0.58	0.05	0.30	0.80
Total population (in 1000s)	1633.63	2717.52	1.57	11821.87
Perc. of the GDP (of 2000) from manufacturing	19.47	10.43	0.08	95.08
Deposit density of 2010 (in 1000s)	4509.32	7478.33	0.00	37187.64
Bank density of 2010	0.03	0.04	0.00	0.83
Observations	189561			

Note: This Table displays descriptive statistics from the DiD estimating sample described in Subsection 3.3 from years before the 2014 reform that made SA auctions for MSEs mandatory for tenders below BRL 80000. Table B2 in Appendix B provides a precise description of each variable in the DiD estimating sample. Respectively, Columns (2) to (5) show the sample mean, the standard deviation, the sample minimum, and the sample maximum of the across establishments distribution of each variable displayed in Column (1). Respectively, Panels A to E display outcomes, treatment, employee controls, employer controls, and spatial controls. The variables in panels A to D date from the pre-treatment period (2011-13), while the spatial controls in panel E indicate their dates in Column (1).

Table B5: Balance-check statistics from the DiD estimating sample before the 2014 reform that made SA auctions for MSEs mandatory for tenders below BRL 80000.

	GSs (treatment)		non-GSs (control)		
Variable	Mean	Std. dev.	Mean	Std. dev.	p-value
Panel A: Pre-treatment outcomes (2011-13)					
Number of employees	10.11	11.66	10.23	12.31	0.15
Average monthly earnings	1105.17	455.74	1098.97	477.21	0.07
Average monthly payroll	8339.59	9790.74	8240.51	9886.79	0.17
Average monthly working hours	43.33	2.58	43.40	2.41	0.00
Panel B: Post-treatment outcomes (2014-20)					
Number of employees	10.08	20.11	9.37	14.47	0.00
Average monthly earnings	1582.46	671.05	1525.43	657.06	0.00
Average monthly payroll	11866.42	15675.69	11125.77	15080.70	0.00
Average monthly working hours	43.02	3.35	43.11	3.10	0.00
Panel C: Treatment (2011-13)					
=1 if estab. is a Government Supplier (GS)	1.00	0.00	0.00	0.00	
Panel D: Employee controls (2011-13)					
Percentage of female employees	32.61	28.16	38.37	34.02	0.00
Percentage of employees with a college degree	7.96	16.97	5.90	15.53	0.00
Average tenure of employees	28.68	24.55	30.50	30.16	0.00
Average age of employees	33.66	6.14	33.51	6.80	0.00
Panel E: Employers controls (2011-13)					
Age of the estab.	6.34	7.75	6.29	8.32	0.40
Panel F: Spatial controls (2011-13)					
GDP per capita of 2000	30.97	16.77	30.69	18.66	0.04
Gini index of 2000	0.59	0.05	0.58	0.05	0.00
Total population (in 1000s)	1796.55	2805.32	1470.68	2616.65	0.00
Perc. of the GDP (of 2000) from manufacturing	18.80	9.57	20.15	11.19	0.00
Deposit density of 2010 (in 1000s)	5015.53	7702.52	4003.04	7211.77	0.00
Bank density of 2010	0.04	0.05	0.03	0.04	0.00

Note: This Table displays balance-check statistics from the DiD estimating sample described in Subsection 3.3. Table B2 in Appendix B provides a precise description of each variable in the DiD estimating sample. The second and third columns display the mean and standard deviation for the sub-sample of establishments classified as GSs (treatment group). The third and fourth columns repeat the previous two columns but restrict the sub-sample to establishments classified as non-GSs (control group). In the last column, we plot the p-value associated with the statistical hypothesis $H_0 : \mu_x(GS_e = 1) = \mu_x(GS_e = 1)$, where x is one of our baseline controls described in the first column, and μ_x is the mean of x. We calculate the p-value by regressing each covariate x_e on our treatment indicator GS_e using robust SEs clustered at the establishment level.

B.2 Building the DiD estimating sample

Samples. We use three samples based on the establishment-level data. First, the *full sample* is a yearly panel data set for the universe of 5.2 million unique establishments with employment information registered in the RAIS between 2011 and 2020, of which 56.5 thousand are Government Suppliers (GSs). Second, the *filtered sample* is a subset of the *full sample* used as an input for the matching algorithm that generates the *matched sample*. Third, the *matched sample* is the estimating sample of the DiD regression model described in subsubsection 4.1.2.

FULL SAMPLE →	FILTERED SAMPLE →	MATCHED SAMPLE
5279800 establishments	1482414 establishments	66698 establishments
5279800 firms	1482414 firms	66698 firms
5274917 MSEs	1482414(1480172) MSEs	66698(66616) MSEs
56561 GSs	33731 GSs	33484 GSs

Filters. We implement four filters to select the filtered sample from the full sample. First, to select the population targeted by SA auctions, we keep only the MSE establishments using an MSE proxy based on the number of employees, which is observable yearly. More precisely, using size categories proposed by Colonnelli et al. (2020), we code an establishment as a MSE in if the establishment is Micro1 (1-4 employees), Micro2 (5–9 employees), or small (10–49 employees), and collapse it at the establishment level by computing the mode across the pre-treatment period (2011-2013). Second, to prevent differences between establishments with and without employers before the treatment to drive our results, we keep only establishments with non-zero wages and employees during the pre-treatment period (2011-2013), reducing the sample size to approximately 4.8 million establishments, of which roughly 54.8 thousand are GSs. Third, to properly test whether outcome trends are parallel during the pre-treatment period, we exclude establishments not observed throughout the entire pre-treatment period (2011-2013), further restricting the sample to approximately 1.5 million establishments, of which roughly 33.7 thousand are GSs. After the four steps, the filtered sample contains 1.47 million establishments, of which roughly 32.1 thousand are GSs.

Sample selection using a proxy. As mentioned in Subsection 3.1, the CNPJ Aberto is only available after 2018, implying that we cannot observe the MSE status during our pre-treatment period (2011-2013) and must use a proxy variable to select the DiD estimating sample. Given this restriction, We select our DiD estimating sample using an MSE proxy computed from the RAIS data following a definition proposed by Colonnelli et al. (2020).

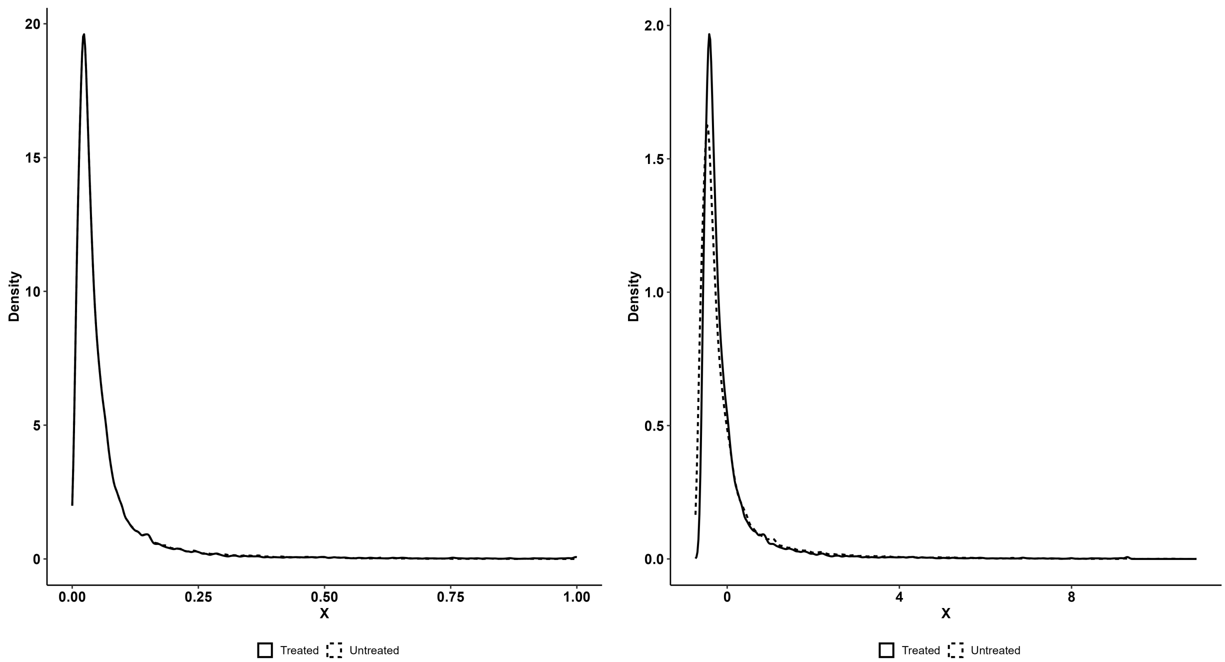
Although the official MSE definition depends on yearly revenue, our sample selection based on the proxy variable likely minimizes contamination in our sample, makes the matching procedure more credible, and facilitates inference in the DiD. First, the correlation between the official MSE status from the CNPJ Aberto and the proxy MSE status from the RAIS is above .99 in 2018, when we observe both variables. Second, if we used the MSE status from the CNPJ Aberto of 2018, we would contaminate the estimating sample by including relatively large firms with MSE status in 2018 but non-MSE status in 2007-2013 and firms with non-MSE status in 2007-2016 that gained such status after the change in the SE criteria in 2016. Third, given that the sample selection using MSE proxy from the RAIS limits the maximum number of employees in the sample, our primary outcome becomes less left-skewed, making the common support assumption of our matching process more plausible and decreasing the weight of the extreme values in the estimates.

Matching algorithm. We implement a three-step matching procedure to select the matched sample from the filtered sample. First, we group establishments according to $21 \times 27 = 567$ industry-state strata generated by the interaction of 21 2-digit industry categories (CNAE 2.0) and the 27 states where establishments are located. Second, for each industry-state stratum, we estimate propensity scores (PSs) at the establishment level by estimating a logit model using the average monthly wages (2011, 2012, and 2013), the number of employees per establishment (2011, 2012, and 2013), the total payroll of the establishment (2011, 2012, 2013), and the quintile of the across municipalities distribution of the total population (of 2011) for the whole country as dependent variables. By using a small subset of the pre-treatment variables to estimate the PS, we can evaluate the performance of our matching algorithm by implementing a balance check in the variables not used in the PS estimation. Third, we match each establishment belonging to a GS establishment to a non-GS establishment according to the nearest PS estimated by the logit model (without replacement). Our matching produces little attrition, as we cannot match around 300 of the 31.6 thousand GSs establishments, generating a matched sample with 63.2 thousand establishments perfectly divided between GSs and non-GSs.

Validating our matching algorithm. We take three steps to validate the outcomes of our matching algorithm. First, we motivate the common support assumption by plotting the distribution of the estimated propensity scores (PSs) of GSs and non-GSs in Figure B1. Second, we evaluate whether our matching algorithm improves the relative level covariate unbalance between GSs (treatment group) and non-GSs (control group) by comparing the balance-check statistics from the filtered sample with those from the matched sample in Figure B2. Third, we evaluate whether the absolute level of

Figure B1: The sample distribution of the Propensity Score (PS) of GSs and non-GSs suggests that our matching algorithm produces a control group that satisfies the common support assumption.

(a) Propensity Score (unscaled) **(b)** Propensity Score (scaled)



Notes: This figure displays the sample Propensity Score (PS) distribution across establishments of both GSs and non-GSs in the matched sample. Respectively, Graph **(a)** and Graph **(b)** plot the unscaled and the scaled PS distribution of GSs (full line) and non-GSs (dashed line). We scaled the PS using a standardization around the sample mean $PS_{i,g} = \frac{PS_{i,g} - \overline{PS_g}}{\widehat{SD_g(PS_{i,g})}}$, where $PS_{i,g}$ represents the PS of observation i of group $g \in \{GS, GS^C\}$, $\overline{PS_g}$ is the sample mean of $PS_{i,g}$, and $\widehat{SD(PS_{i,g})}$ is the estimate of the Standard Deviation (SD) of the $PS_{i,g}$ in group g.

covariate unbalance is suitable for conducting a DiD estimation by comparing the joint distribution of balance-check statistics from the matched sample with statistical significance critical values and definitions of small magnitudes of mean differences in Figure B3.

Common support. Figure B1 displays the sample Propensity Score (PS) distribution across establishments of both GSs and non-GSs in the matched sample. Respectively, Graph **(a)** and Graph **(b)** plot the unscaled and the scaled (standardized) PS distribution of GSs (full line) and non-GSs (dashed line). Graphs **(a)** and **(b)** shows that sample distributions of the PS of GSs and non-GSs have a very high degree of overlap, suggesting that our matching algorithm produces a control group that satisfies the common support assumption.

Relative balance-check. Figure B2 displays balance-check statistics from the *Filtered Sample* and the *Matched Sample*. Graph **(a)** shows the T-statistic for the mean difference between GSs (treatment group) and non-GSs (control group), testing the hypothesis $H_0 : \mu_x(GS = 1) = \mu_x(GS = 0)$, where x represents one covariate, GS is the treatment indicator, and μ_x is the mean of x. Graph **(b)** reports the standardized mean difference $\Delta_x = \frac{\mu_x(GS=1)-\mu_x(GS=0)}{\sigma_x}$, where σ_x is the standard deviation of x. We calculate both statistics by regressing each standardized baseline covariate $\frac{x}{\hat{\sigma}_x}$ on our treatment indicator $GS_f(e)$ using robust SEs clustered at the establishment level.

Results from Figure B2 show that our matching algorithm substantially decreased the degree of observable differences between GSs and non-GSs, suggesting that causal inference is more appropriate in the *Matched Sample*. In Graph **(a)**, for nearly all of our xx pre-treatment variables, we observe a decrease in the T-statistic T_x from the matched sample (blue dots) in comparison to those from the filtered Sample (red dots). In Graph **(b)**, we observe the same visual pattern when inspecting standardized difference Δ_x in both samples.

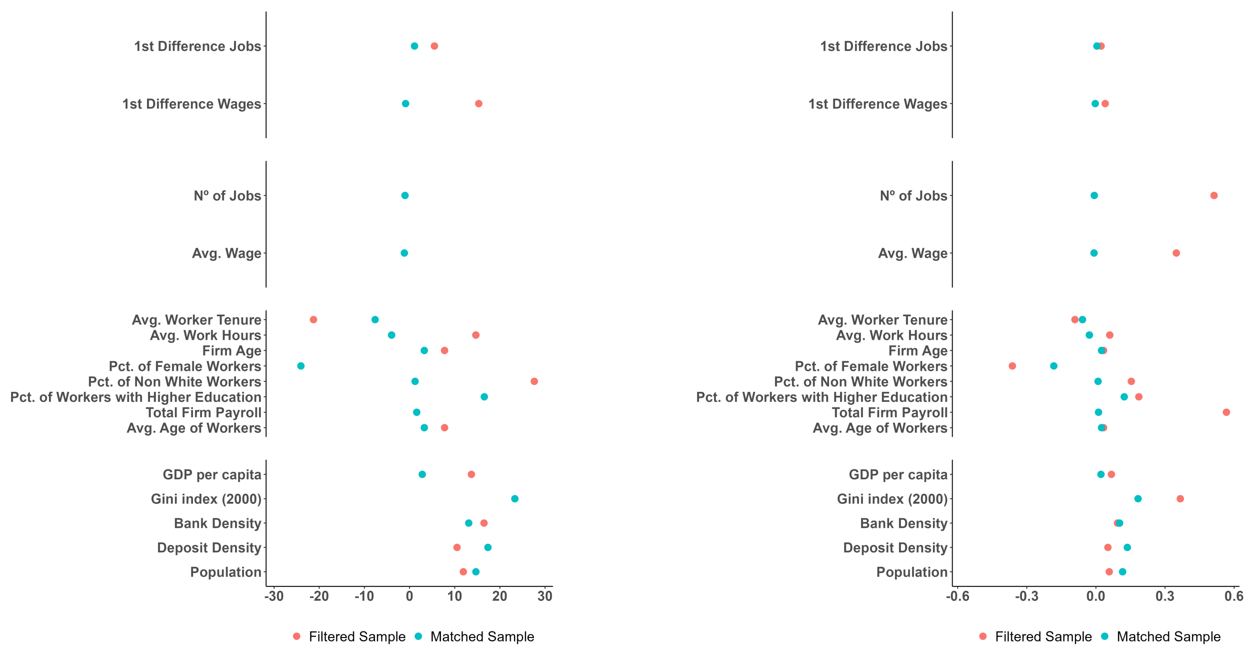
Absolute balance-check. Figure B3 displays balance-check statistics from the the matched sample. Graph **(a)** shows the T-statistic for the mean difference between GSs (treatment group) and non-GSs (control group), testing the hypothesis $H_0 : \mu_x(GS = 1) = \mu_x(GS = 0)$, where x represents one covariate, GS is the treatment indicator, and μ_x is the mean of x. Graph **(b)** reports the standardized mean difference $\Delta_x = \frac{\mu_x(GS=1)-\mu_x(GS=0)}{\sigma_x}$, where σ_x is the standard deviation of x. We calculate both statistics by regressing each standardized baseline covariate $\frac{x}{\hat{\sigma}_x}$ on our treatment indicator $GS_{f(e)}$ using robust SEs clustered at the establishment level. In Graph **(a)**, we plot vertical lines at -1.96 and 1.96 to indicate statistically significant differences at the 5% level. In Graph **(b)**, we plot vertical lines at -0.2 and 0.2 to mark the range of differences considered to be small in magnitude.

Results from Figure B3 show that our matching procedure generates enough similarity between GSs and non-GSs to conduct causal inference using a DiD regression model based on the parallel trends hypothesis. First, in Panel A, we document that pre-treatment outcomes of GSs and non-GSs become balanced in the matched sample in the first difference, showing that pre-treatment outcome trends are parallel within the matched sample. Second, while we reject the hypothesis of balance in covariates for more than half of our pre-treatment variables, the standardised differences have a small magnitude in the matched sample.

Relevant to our identification, the parallel trends hypothesis does not require that treatment and control groups be perfectly balanced, making the outcome of the matching algorithm acceptable for our purpose. However, to ensure that such remaining unbalances between GSs and non-GSs do not drive our results, we re-estimate our DiD specification with year FEs interacted with pre-treatment variables as controls in **??**.

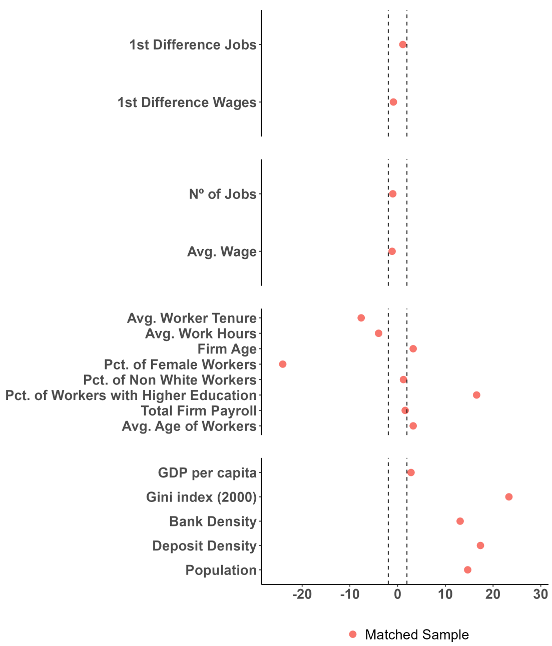
Figure B2: Our matching algorithm considerably decreases differences between GSs and non-GSs.

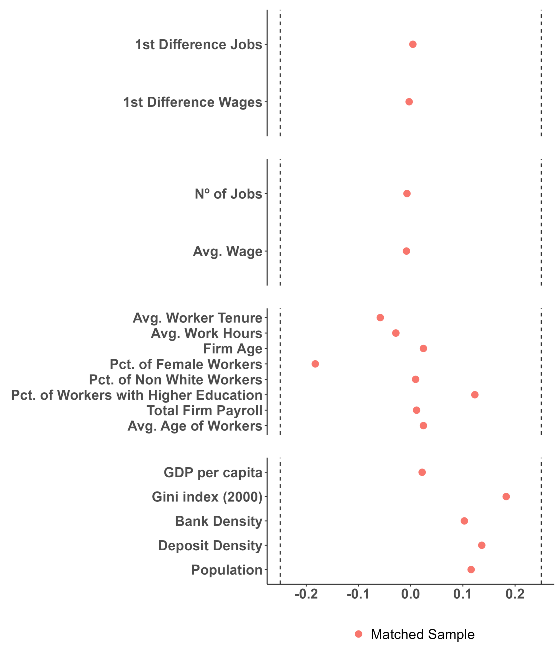
(a) T-Statistic **(b)** Standardized Mean Differences



Notes: This figure displays balance-check statistics from the filtered sample (blue dots) and the matched sample (red dots). Graph **(a)** shows the T-statistic for the mean difference between GSs (treatment group) and non-GSs (control group), testing the hypothesis $H_0 : \mu_x(GS = 1) = \mu_x(GS = 0)$, where x represents one covariate, GS is the treatment indicator, and μ_x is the mean of x. Graph **(b)** reports the standardized mean difference $\Delta_x = \frac{\mu_x(GS=1)-\mu_x(GS=0)}{\sigma_x}$, where σ_x is the standard deviation of x. We calculate both statistics by regressing each standardized baseline covariate $\frac{x}{\sigma_x}$ on our treatment indicator $GS_{f(e)}$ using robust SEs clustered at the establishment level.

Figure B3: Our matching algorithm makes GSs and non-GSs balanced in terms of pre-treatment outcomes in levels and first difference and reduces differences in establishment and location characteristics to small magnitudes.

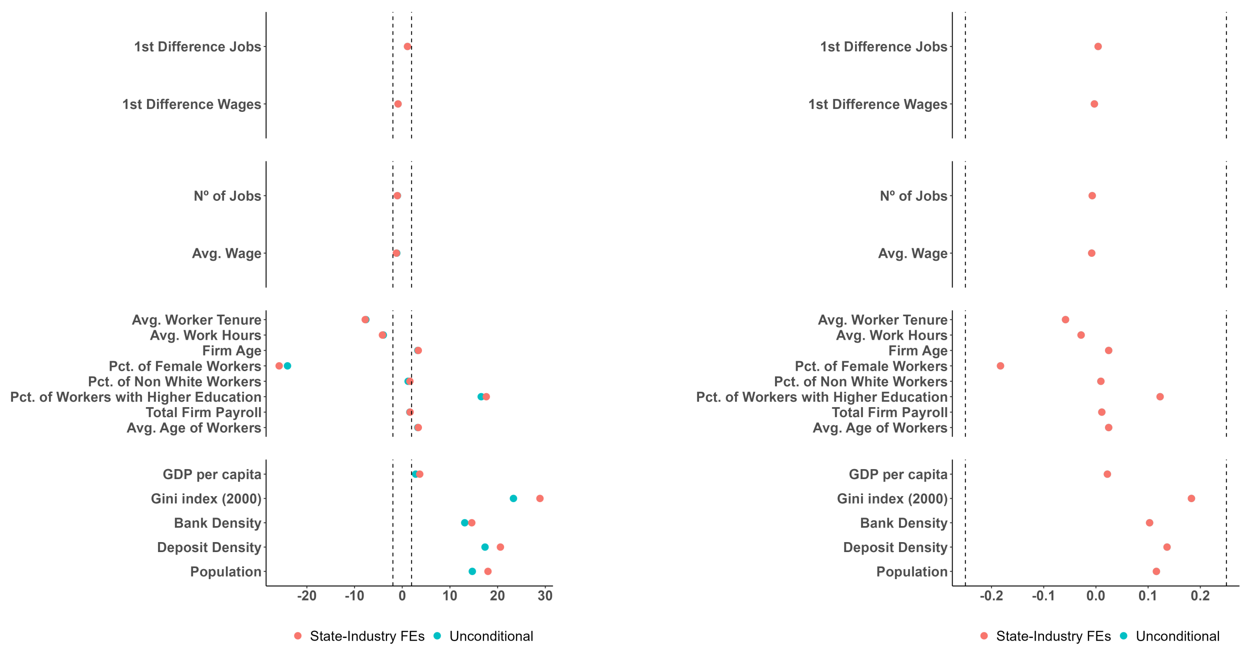




Notes: This figure displays balance-check statistics from the the matched sample. Graph **(a)** shows the T-statistic for the mean difference between GSs (treatment group) and non-GSs (control group), testing the hypothesis $H_0 : \mu_x(GS = 1) = \mu_x(GS = 0)$, where x represents one covariate, GS is the treatment indicator, and μ_x is the mean of x. Graph **(b)** reports the standardized mean difference $\Delta_x = \frac{\mu_x(GS=1)-\mu_x(GS=0)}{\sigma_x}$, where σ_x is the standard deviation of x. We calculate both statistics by regressing each standardized baseline covariate $\frac{x}{\sigma_x}$ on our treatment indicator GS_e using robust SEs clustered at the establishment level. In Graph **(a)**, we plot vertical lines at -1.96 and 1.96 to indicate statistically significant differences at the 5% level. In Graph **(b)**, we plot vertical lines at -0.25 and 0.25 to mark the range of differences considered to be small in magnitude.

Figure B4

(a) T-Statistic **(b)** Standardized Mean Differences



Notes: This figure displays balance-check statistics from the the matched sample. Graph **(a)** shows the T-statistic for the mean difference between GSs (treatment group) and non-GSs (control group), testing the hypothesis $H_0 : \mu_x(GS = 1) = \mu_x(GS = 0)$, where x represents one covariate, GS is the treatment indicator, and μ_x is the mean of x. Graph **(b)** reports the standardized mean difference $\Delta_x = \frac{\mu_x(GS=1)-\mu_x(GS=0)}{\sigma_x}$, where σ_x is the standard deviation of x. Blue points represent unconditional balance-check statistics estimated using all comparisons between GSs and non-GSs. Red points represent conditional balance-check statistics estimated using comparisons between GSs and non-GSs within the same State-Industry stratum. We calculate both statistics by regressing each standardized baseline covariate $\frac{x}{\hat{\sigma}_x}$ on our treatment indicator GS_e using robust SEs clustered at the establishment level. We estimate the statistics in red using a regression model with State-Industry Fixed Effects (FEs). In Graph **(a)**, we plot vertical lines at -1.96 and 1.96 to indicate statistically significant differences at the 5% level. In Graph **(b)**, we plot vertical lines at -0.25 and 0.25 to mark the range of differences considered to be small in magnitude.