Set asides and value thresholds in Brazilian public procurement from small businesses*

Eduardo P.S. Fiuza† Carlos H.L. Corseuil‡ Athos V.C. Carvalho§
Pedro H.H. Coimbra¶

August 7, 2023

Abstract

Access of small business to public procurement is a customary goal of economic policy targeting this important segment of economy. Among the measures and approaches used to support Small and Medium Enterprises in public procurement, the second most cited by surveyed OECD countries (OECD, 2018b) is the division of contracts into lots. However, impact assessments of such policies are still scarce. The present article purports to fill this void by studying the use of contract value thresholds in Brazil, below which Micro and Small Enterprises (MSEs) should be the only group allowed to bid. Our first finding suggests that lots to the left and right of the threshold have similar characteristics. Hence we use the latter as a counterfactual to what would have happened to those lots with values below the threshold had they not been allocated exclusively to MSEs. Fuzzy RDD estimations – both with the optimal bandwidth by Calonico et al. (2020) and with asymmetrical bandwidths – point out that set-asides have been successful in attracting more and younger MSEs to bid. Moreover, both the participation ratio of MSEs and the award success rate were significantly boosted by the set aside policy. The greater number of MSEs, however, came apparently at the expense of a net reduction in the total number of bidders. Probably related to this effect, both a remarkable and significant rise in the price level (and a corresponding drop in the discounts off the reserve price) and a decrease in the dispute at the second (open cry) stage of the hybrid auction were observed. These movements in prices may reflect a deterioration in competition below the threshold, as

*This research was mostly developed while Eduardo Fiuza was serving Ipea. Pedro Coimbra gratefully acknowledges financial support from the Economic Commission for Latin America and Caribbean (ECLAC) – Contract No. 2500236324. Athos Carvalho gratefully acknowledges financial support from Ipea and from ILO (Contract No. 40410717). Competent research assistance by Jonathan Oliveira is also gratefully acknowledged. Finally, we are extremely grateful to Ronaldo Correa, Viviane Mafissoni and Felipe Ansaloni for sharing their time and providing valuable discussion of the preliminary statistics. The usual disclaimer applies.

†Conselho Administrativo de Defesa Econômica (CADE); eduardo.fiuza@cade.gov.br.

‡Institute for Applied Economic Research (IPEA); Av. Pres. Vargas, 730, 16o andar – Torres 3 e 4 - Rio de Janeiro-RJ 20.071-900, Brazil; carlos.corseuil@ipea.gov.br.

§Institute for Applied Economic Research (IPEA)

¶Université Paris-Sinclay, INRAE, AgroParisTech, UMR Ecosys, 91120 Palaiseau, France; pedrohhc@gmail.com

two important collusive markers – the coefficient of variation of bids and the skewness of the bid distribution – significantly dropped, thus signalling in favour of some suspicion of collusion.

JEL classification: D44, H57, L73, Q23

Keywords: public procurement, small business, favouritism

1 Introduction

Small business account for large shares of total employment (average 60%) and value added (between 50% and 60%) in OECD countries (OECD, 2018a). In the world, public procurement accounted for 12% of GDP in 2018 (Bosio et al., 2022), while the corresponding share tops 15% in low-income countries (Djankov et al., 2016). It is thus unsurprising that the access of small business to public procurement is a customary goal of economic policy targeting this important segment of economy.

Among the measures and approaches used to support Small and Medium Enterprises in public procurement, as listed by OECD (2018b), the second most cited by the surveyed OECD countries is the division of contracts into lots. However, impact assessments of such policies are still scarce. One reason is a methodological challenge for the correct identification, as lots are usually non-randomly allocated as set-asides. Therefore, most of the previous studies prefer not to conduct such evaluation relying on recent econometric techniques developed for impact evaluation, and rather involve an explicit modelling of the decision to participate in the auction. We aim to use the Brazilian experience to analyse their set-aside policy in public procurement.

In the particular setting of Brazilian auctions, a value threshold was created by legislation, below which only Micro and Small Enterprises (MSEs) are allowed to bid. The decision of the procuring entities to apply these set-aside lots, however, has been of incomplete compliance. It follows that above the contract value threshold no restriction to entry is applied, but below it not all lots are set aside. This non-compliance behavior is explained by loopholes in the legislation that enable different interpretations regarding the relevant value to be taken as reference, as well as by explicit opt-out provisions and by inconsistent settings in the platform utilized.

By coupling a definite threshold with an imperfect compliance, it is natural to use this imperfect separation for generating a fuzzy Regression Discontinuity Design (*Fuzzy RDD*). We have also dismissed concerns about gaming the system by procuring entities who would set values below or above the relevant threshold. This gives us confidence that our strategy relies on the exogeneity of the auction lot value across the vicinity of the threshold.

The present article, thus, brings a fresh approach to the use of set asides based on contract value thresholds in Brazil, aiming at a better understanding of the mechanisms of participation and win that affect MSEs, as well as an accurate assessment of the net effects of this program.

The next section brings background notes that contextualize the search for support instruments targeting MSEs. The third section makes a brief review of the literature on preference programs. The fourth section brings methodological and data issues, while the fifth section collects and discusses descriptive statistics and preliminary results. The sixth and last section concludes.

2 Background notes

The 2020-2021 Covid-19 pandemic was a moment of severe stress for small businesses all around the world. State financial aid was poured to assist them with employment maintenance, in the forms of: tax, social security, rent, and utility payment time extensions or waivers; waivers or reductions of financing fees and interest; debt payment moratoria; thus providing liquidity and preventing a more pronounced wave of bankruptcies in the segment (OECD, 2021b). Wage subsidies, permission to temporarily suspend labor contracts, funding for leaves of absence due to Covid-19, corporate tax and social security payment delay, purchase of Small and Medium Enterprises' debts were other strategies adopted (OECD, 2021a). Other policies targeted to SMEs included structural support measures, aiming at digitalisation, upskilling and reskilling, finding new markets, innovation and technology development (OECD, 2021a).

In Brazil, medium-sized business have not been object of public policies in the past decades, at least since the Federal Constitution of 1988, but rather the micro and the small enterprises (MSEs)[1]. Recent programs aimed at them include the National Support Program for Micro and Small Enterprises (Pronampe), which was launched on May 19, 2020, and has provided guarantees to financial agents in order to facilitate access to credit for MSEs. The program has been operated by the National Bank for Economic and Social Development (BNDES) with support from the Investment Guarantee Fund (FGI). The amount allocated for investment and working capital would correspond to 30% of the annual gross income reported in 2019, except for firms younger than that (Dini and Heredia Zurita, 2021). The interest rate applied was considerably lower than the market average rate. An impact assessment conducted by the Federal Court of Accounts (TCU[2]) found that Pronampe, Fampe (a pre-existing credit program targeting small business) and three other emergency credit programs – two of them also targeted to MSEs –, which provided from 80% to 100% guarantee of the loan amount, successfully increased employment levels and the firms' access to credit as a whole. The public auditors, however, recommended that the public funds should cover a lower share of the loans, so as to be coupled with private funds.

It is true that the disproportionate impact of the Covid-19 pandemic on SMEs is due not only to their financial frailty, but also to the great share of this segment in the sectors most affected by the crisis, such as accommodation and food services, real estate, wholesale and retail trade, professional services and other personal services (an average 75% across OECD countries) – other factors were weaker supply chains, a digital gap, and lack of managerial skills (OECD, 2021b).

Despite all this turmoil for SMEs, it is worth noting that firms selling to government in general

[1] For more details on the legal and statistical definitions of MSEs, and programs targeting them, we refer to Fiuza et al. (2023).

[2] See Brasil (2021a,b).

did not face the same difficulties as those in personal services – except, of course, travel and accommodation firms.

It seems therefore that public procurement plays a key role in backing survival and growth of small and medium business (SMB). In fact, according to OECD (2018b):

> Governments are increasingly taking steps both to give SMBs better access to public markets and to remove barriers preventing SMBs from winning public contracts. Engaging SMBs in public procurement is beneficial both for the companies and for the public sector. On the one hand, public procurement contracts give SMBs better access to markets and help them strengthen their own capacities. On the other, the public sector can better meet its procurement needs by working with innovative, responsive and flexible SMBs.(p.15)

The same report adds, however, that "*specific characteristics of public procurement, such as the complexity of procedures, administrative burden and high technical and financial capacity requirements, disproportionately affect SMBs and hamper their access to the market.*". In fact, among the main constraints on SMB access to public contracts in OECD countries, responses to a survey received from 26 OECD countries indicated as very relevant (p.39):

- Difficulties relating to the size of contracts;
- Too high administrative burden;
- Access to relevant information;
- Late payments or lengthy terms of payment;
- Quality and understanding of the information provided

Another report on the participation of SMEs in public procurement, by EBRD (2017), reinforces that, while small business account for large shares of employment and GDP in developed countries, they do not enjoy an *"equal measure in business-to-business and business-to-government transactions"*.

OECD (2018b) also point out that many recent public procurement reforms in OECD countries have been seeking to facilitate SME access to public procurement opportunities and level the playing field, for example by ensuring that (p.39):

- The size of tenders do not unjustifiably discourage SME participation;
- Processes and documents are not unnecessarily complex, and are simplified according to the value and risk of the procurement object;
- The financial capacity required of SMEs is set at a proportionate level and that SMEs' participation in public procurement markets does not excessively limit their financial

conditions; and

- The use of information and communication technologies in public procurement improves SMEs' access to public procurement.

Therefore, the main measures and approaches used by OECD countries to support SMEs in public procurement have been, in order of importance, according to OECD (2018b) (p.51):

1. Encourage the use of e-procurement;
2. Encourage the division of contracts into lots;
3. Encourage joint bidding/consortia rules so as not to discourage SME participation;
4. Simplify processes and documentation requirements for SMEs;
5. Encourage prime contractors to subcontract with SMEs and/or make subcontracting arrangements to facilitate/encourage SME participation;
6. Arrange timely and efficient payment terms for SMEs (or for lower-value contracts).
7. Accord SMEs preferential financial treatment for SMEs, e.g. waving of fees.

In Brazil, most of these measures have been adopted in a way or another (for details, see Fiuza et al. (2023)):

1. **E-Procurement:** The use of e-procurement started in 2000, when the Government of the State of São Paulo created the E-Procurement Exchange, followed by the Federal Government, with the Provisional Law 2026/2000 (converted into a permanent Law 10520 by the Congress only in 2002) that provided for an electronic version of the hybrid auction named *pregão*, made up of a sealed bid stage followed by an open bid session. Two versions were regulated by Decree: face-to-face (*pregão presencial*), where the sealed bid is handed on paper, and the open bids are orally submitted; and e-auction (*pregão eletrônico*), where both sealed and open bids are submitted electronically and with identities of the bidders concealed before the end of the open bid session. Unfortunately for our purposes, this is the only modality of procurement where all bidders' identities remain archived in the logbook system, so it is possible to find out whether more MSEs are awarded contracts, but not how, neither how many (MSEs, medium and large) firms have competed for each of them. Therefore we are not able to trace back in time the evolution of participation before and after the introduction of the two hybrid auctions, or across different modalities. What we do know is that hybrid auctions – and within this group, the e-auctions – have become the prevalent modality of procurement in number of tenders and in number of lots, but not necessarily in value purchased.

2. **Simplifying qualification requirements:** Around the world, in order to prevent the participation of phantom bidders, tax evasion, bid manipulation, etc., governments ordi-

narily take stringent measures to ascertain that firms submitting bids for public contracts are regularly registered and compliant with their legal, tax and labor obligations. This should pose a disproportionate administrative burden on MSEs – in fact, in 2019 Brazil was the top country in average number of hours spent to prepare and pay taxes in a year, with a stunning record of 1,501 hours, against a median time of 207 hours across a sample of 237 countries surveyed by the World Bank[3]. The countervailing remedies for such competitiveness asymmetry in Brazil have been twofold: (i) by unifying taxes and labor payments into a single contribution with the Simples Nacional tax regime, the Brazilian Law reduced the number of certificates to be submitted, without actually waiving them; (ii) MSEs were given more time to hand the required documentation: SL 123/2006 gave them two extra working days as compared to non-MSEs, and this time extension was increased by SL 147/2014 to five working days, with two extra extensions of the same size.

The impact of *Simples Nacional* is impressive: a recent study (Levy-Carciente, ed, 2022) found that Brazilian MSEs enjoyed the smallest administrative burden in Latin America. However, these benefits were not granted homogeneously to all MSEs: some of them were not eligible for the Simples Nacional regime according to their economic activity. A careful compilation of the activity classes and a treatment effect appraisal is due in another article in the near future.

3. **Subcontracting**: The same Supplementary Law 123, art. 48, provides for lots where MSEs **may** be subcontracted. Before the amendment undertaken by S.L. 147/2014, subcontracting was limited to a maximum 30% of the contract's value. The amendment extinguished the maximum share, but at the same time it restricted subcontracting to construction and services. This strategy should be an important measure for fostering the indirect participation of MSEs in public contracts, yet in Federal records (DW Siasg) the presence of subcontracting is approximately zero. We conjecture that this sort of information is either very rarely recorded in the main repository, and is rather left as a hard (non-machine-readable or, at least, described in a non-structured way) detail lost in the proceedings of the procurement, or (most likely) this instrument is in fact rarely utilized.

4. **Preferential financial treatment:** In Brazil, MSEs are not waived from financial obligations when participating in procurement auctions, and no bid subsidy has ever been approved for this segment of firms. However, they benefit from an ingenious arrangement provided by legislation: the so-called *fictitious tie*, which turns out to be a right of final undercut, after the open cry stage is ended (or, in case of pure sealed bid auctions, after the opening of the bid envelopes). Obviously this mechanism is only available in lots of open access to both MSEs and large firms. In a word, it works as follows: after the last

[3] Source: `https://data.worldbank.org/indicator/IC.TAX.DURS`.

bid is submitted or opened, and the bidding session is considered ended, the auctioneer (or the system, automatically) checks if the front runner is a MSE. If it is not, and the runner-up or any bidder as high as 5% (or 10% in pure sealed bid auctions) above the front-runner is an MSE, the author of the best bid in either of these conditions is called to exercise their right to undercut the front runner, even if it is by a single cent. It is important to notice that from 2005 to 2018 the bidders were allowed to bid above the front runner along the whole open cry stage (and lower only their own bid), thus being able to behave less aggressively in *pregões*; they only have to stay close enough to the leading bid, as long as the latter has not been submitted by another MSE – something the MSE cannot be assured of, as the bidders' identities are not disclosed before the auction closure.

In the next subsection, we shall focus on the division of contracts into lots, which provides us the natural experiment to assess the set-aside treatment.

2.1 Division of contracts into lots

Around the world the definition of a micro, small or medium business varies enormously, as different countries adopt different criteria – such as employment, sales or investment (Ayyagari et al., 2007). As from the 1990s, Brazilian legislation aimed at reducing the statutory burden of tax, social security and labor benefits levied on micro and small enterprises (but no word about medium enterprises), not only on their revenues, but also especially on their cost of compliance. As a matter of fact, the definition of micro and small enterprises in Brazil underwent changes along time. For public procurement purposes, it is enough to remark that:

- As from the Supplementary Law 123 in 2006, a firm is classified as micro if its turnover is lower than 360,000 BRL and as small if its turnover is strictly greater than BRL 360,000 and lower than BRL 3.6 millions.
- In 2014 the SL 123 was amended, and the upper threshold for small firms was adjusted upwards, to BRL 4.8 million[4].

Now, regarding the use of division of contracts into smaller lots: the same Supplementary Law 123, art.48 listed two situations for the creation of set-aside lots:

- Item I (contract threshold): purchases below 80,000. A set aside lot **might** take place for MSEs if the reservation total price was below BRL 80,000, or for a quota up to 25% of the annual object purchases (merchandise or service) procured.
- Item III (set-aside share): a set aside lot **might** take place for MSEs in 25% of the total

[4] This threshold applies to public procurement policies, but regarding the unification of Federal and state taxes the previous threshold still applies.

value of divisible goods and services procured along one year.

In 2014, a change in regulation was implemented (Supplementary Law 147, which amended S.L. 123), according to which the set asides **must** be set up for MSEs. Since 2006, no differentiated treatment has been applied in case of less than three eligible MSE competitors, or in case the treatment would not be advantageous.

It is worth noting that, as from 2006, the treatment would not take place if it was not expressly provided for in the bidding notice, or if the purchasing agency was able to replace the tender with a direct purchase (the legislation provides for a great number of such cases, related to patents, market exclusivity, emergency and so on), but one such case of direct purchase was excluded in 2014: small value — in this case, the purchase **must** be made from an MSE, as long as the price is compatible with the market price or previous government purchases.

According to Figures 1, 2, 3 and 4, set-aside lots based on value below the BRL 80,000 threshold have apparently become the main channel to benefit MSEs if we account for the number of lots, even though the value committed for payment still accounts for a very tiny fraction of the total. These figures however, may be deceitful, because:

- All bid waivers are classified as having no MSE, and, as we saw above, bid waivers based on small value purchases have been set aside for MSEs since 2014.
- MSEs may still be awarded in "no-benefit" auctions with the help of 'fictitious ties'.

For the first case, we can estimate the extra share of benefits enjoyed by MSEs with the amount committed in the budget and the number of lots related to the bid waivers justified by small value thresholds (Law 8666, art. 24, items I and II). For the second case, the estimate is a bit trickier: we have to isolate lots where the initial front runner was not an MSE and was not awarded, while an MSE made an undercut after the open bid stage was closed. It is worth noting that we can only observe these timestamps in hybrid auction (both electronic and face-to-face)[5].Fiuza et al. (2023) then constructed a proxy for the subset of hybrid auctions (*pregões*)– indicating whether any bid was submitted after the closure of the open bid session **and** the corresponding bidder have been awarded the contract. With this more stringent approach, we find a gap of 65% in value and 50% in number of lots where the benefit was ultimately granted. More precisely, instead of nine percent of contracts in value and two percent in number, successful fictitious ties yielded only three and one percent, respectively.

[5] An indicator of the fictitious tie is available, but it only informs if a tie-breaker was called for, not if indeed some MSE has been awarded the contract thanks to this instrument, nor even if the MSE did submit a last offer to undercut the front-runner.

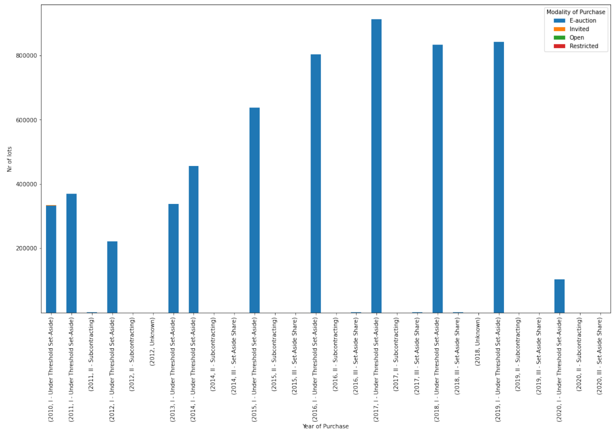

Figure 1: Number of lots with preferences

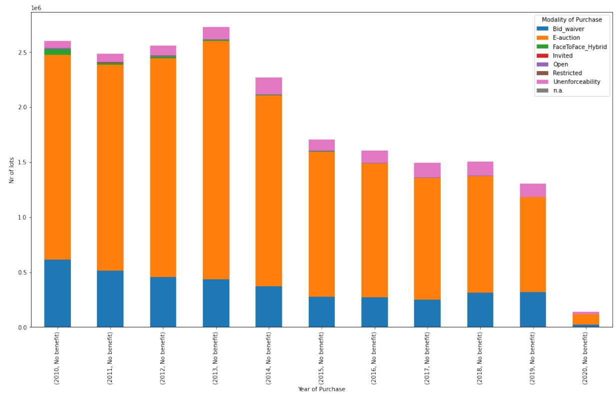

Figure 2: Number of lots without preferences

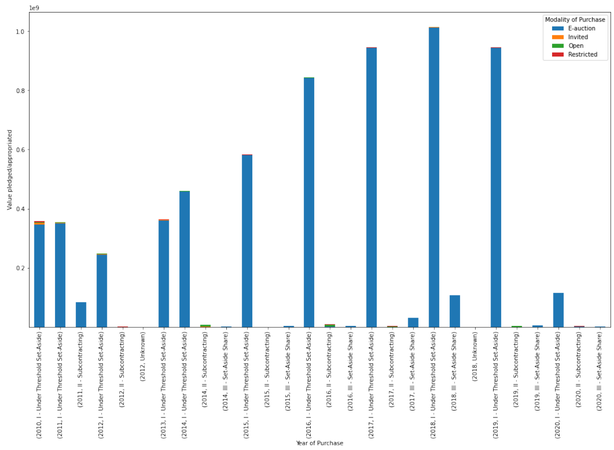

Figure 3: Awarded value with preferences

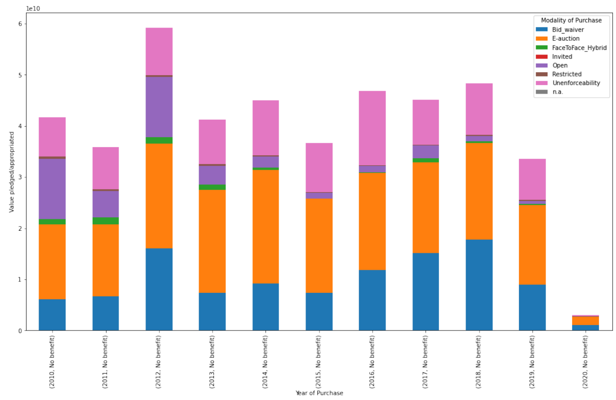

Figure 4: Awarded value without preferences

3 Econometric references in the public procurement literature

3.1 Evidence of the impact of public procurement policies on small business

To the best of our knowledge, academic articles that endeavored to assess the impact of public procurement on small business have concentrated on two major groups of policies: (i) set-aside auctions; (ii) bid subsidies. In the first group, procurement officers set specific auctions or lots to be restricted to bids originated from businesses classified as small (or medium, or micro) enterprises. In the second group, the small businesses' bids enjoy a premium of $x\%$ over the bids of greater firms, in the sense that the bids of preferred (small) firms are scaled by a discount factor for the purposes of evaluation only. So, for example, if a non-small business submits the lowest bid b^*, the small business may bid $b^*\left(1+\frac{x}{100}\right)$ minus 1 cent and still win the lot. Such asymmetry may entail three effects: first, preferred firms may inflate their bids, yet still win the auction; second, non preferred firms may bid more aggressively than in the absence of preferences; third, the preference policy can affect participation (Hubbard and Paarsch, 2009).

Bid preferences have been analyzed by Marion (2007); Hubbard and Paarsch (2009) and Krasnokutskaya and Seim (2011). The first article collects data from California auctions for road construction contracts, where small businesses enjoy a five percent bid subsidy in projects funded by the state, while no preferential treatment is applied on projects using federal aid. A structural auction model is estimated nonparametrically following Guerre et al. (2000). The author finds that procurement costs are 3.8 percent higher on auctions using preferences, because of the reduced participation by lower cost large firms, which more than compensates the reduction of firm profits. It is worth noting that the author does not model participation explicitly.

Hubbard and Paarsch (2009) also resort to a structural Independent Private Cost auction model, which is estimated numerically with an MPEC (Mathematical Programming with Equilibrium Constraints) approach, but on simulated data. Endogenous entry is modelled following Samuelson (1985). The simulation leads to two major effects: (i) an important preference effect, meaning preferred bidders inflate their bids; (ii) an important competitive effect, meaning non preferred bidders behave more aggressively than under equal treatment of bids in an effort to counteract the preferential policy. The importance of (i) is found small and dependent on the distribution of costs. Under four different distributions often employed in structural-econometric research, the authors find that preference policies are not cost inefficient, and can lead to cost savings for the government, however they are not the best approach to increase participation from under-represented classes of bidders. For two out of the four distributions, the policy actually reduced entry by preferred bidders.

Finally, Krasnokutskaya and Seim (2011) also pursue an auction structural model in two stages, such that the decision to participate is accounted for, following not only Samuelson (1985), but also Levin and Smith (1994). Thus, in the first stage, parameters of the bid distribution and the distribution of entry costs are estimated without imposing the full set of equilibrium restrictions. In the second stage, the distribution of project costs is recovered from the equilibrium bidding first order conditions following the procedure described by Guerre et al. (2000), which requires computing moment conditions – the authors rely then on the Method of Simulated Moments (MSM). Whereas bids, entry costs and unobserved heterogeneity are estimated parametrically, these distributions can be identified from their data nonparametrically. The authors recover nonparametrically the marginal distributions of the firm-specific cost components and the distribution of unobserved heterogeneity conditional on the numbers of actual bidders, as demonstrated by Krasnokutskaya (2011). This model is then applied to the California Small Business Preference program, and the authors conclude that: (i) the program generates only small increases in procurement costs; (ii) for the set of projects in their sample, a higher discount (15%) would be needed to reach the allocative target; the usual range of discounts could actually result in significant cost increases for at least a subset of projects; (iii) the government should optimally employ a more nuanced preferential treatment, tailoring the discount rate to the project type, in a similar fashion to the approach taken by the FCC; (iv) a lump-sum entry fee is more effective than a bid discount at reducing the cost of procurement. Finally, the authors claim that their results demonstrate that a preference program evaluation depends critically on capturing firms' participation responses to the policy.

The literature on set-aside auctions is the other vein of relevant empirical articles on the impact of preference programs for MSEs. We highlight next the studies of Brannman and Froeb (2000); Athey et al. (2013); Nakabayashi (2013)[6]. The first two articles analyze the set aside program held by US Forest Service timber auctions, albeit using different datasets. According to this program, a fraction of the auctions is reserved for exclusive participation of small mills or loggers. The former specify a Vickrey (second-price) model of bidding competition for oral auctions, without considering endogenous entry, and they use a within-auction estimator, based on the difference between losing bids, to recover the bidders' value distribution. Using the empirical value distribution, the authors simulate the effects of various policy interventions. In particular, they find that eliminating the set aside program would increase government revenues by 15% and that a policy of granting bidding preferences to small and more distant bidders would raise revenue by less than one-tenth of one percent.

Athey et al. (2013) estimate a structural model using only data from unrestricted sealed bid sales, and assess its performance by comparing the out-of-sample predictions for small business set-asides with the actual outcomes in the data. They claim a good performance, as the predicted prices and entry rate are within five percent of observed values, and equality of the predicted and observed bid distributions cannot be rejected. They find that the set aside program induces losses both in terms of revenue (5%) and efficiency (17%). However, by assuming endogenous entry the authors show that the losses are mitigated in set aside auctions due to the entry of small firms; without this effect the losses would be 30% in terms of revenue and 28% in terms of efficiency. In addition, they use the model to calculate the effect of implementing a bidder subsidy program (applied to all sales) in lieu of direct set-asides for a subset of sales, and find that, with a six percent bid subsidy, small firms would win the same proportion of auctions, the price would be four percent higher, and the efficiency would increase by two percent.

Nakabayashi (2013) studies the impacts of the Japanese small business preference program. His dataset covers construction contracts awarded by the national government through score auctions. In this dataset very few lots have both SMEs and large firms competing with each other. The estimation follows three steps:

1. A nonparametric estimation for asymmetric first-price sealed-bid auctions with affiliated

[6] Other interesting references are Reis and Cabral (2015) and Rosa (2019).

private values (APV) identifies the bidders' costs from observed bids.

2. Expected payoffs (profitability times the probability of winning) are regressed on firm size.

3. An entry static model predicts how many SMEs would drop out because of large firm entry into a market that was previously reserved exclusively to SMEs under the set-aside program.

The results obtained in the second step point out to a cost difference of 1.4% between large firms and SMEs. They also indicate that an SME would win 4.7 percent less frequently than a large firm if they competed one-on-one. These small differences in costs and winning probability translate into a nontrivial difference in profitability between the two groups of bidders. The expected payoff of an SME would be 46 percent lower than that of a large firm when both competed in the same auction. Because of the disadvantage in profitability, the participation of SMEs would drop by 36% on average if set-asides were removed.

Two opposite effects would follow: the quality-adjusted prices of the originally set-aside projects would fall because of the entry of cost-efficient large firms, whereas the quality-adjusted prices of remaining projects that are open only to large bidders because of their complexity would rise, because of a decline of approximately 37 percent in the number of large firms. As counterfactual studies suggest that the latter effect dominates the former, the program should decrease the effective procurement costs by 0.22 percent. Last but not least, the net rent of the large firms is 0.02% with the set aside program and 0.64% without it. It is worth noting that in the dataset there is no clear threshold or procuring entity's attribute to rely on as defining which lot is going to be a set aside; the government has only a target share of set aside lots, but not necessarily related to the size of the work or budget, or complexity.

Hoekman and Taş (2020) study the impact of division of lots on the probabilities of SME participation and win in European tenders. The dataset covers tenders from almost all European Union members, recorded in the TED (Tenders Electronic Daily) database. The model is quite simple: a probability model estimated on the pooled data (logit, OLS, IV-GMM, and HB-IV-GMM), and controlling for an index of quality of the country's public procurement regulation and dummies for open procedure, value above the EU threshold, and divided lots. The authors find that countries with higher public procurement (PP) regulation quality scores are associated with a larger ratio of SME participation and higher probability that SMEs win contracts. Also controlling for PP regulation quality, the use of open competitive tendering methods render larger contracts more likely to be awarded to large firms. Regarding subdivision of contracts into smaller lots, the authors find that it is associated with greater SME participation, but not with an increase in the probability of SMEs winning contracts. Threshold regression analysis suggests that the absolute value of lot sizes matters: for lots below 25,000 €, the likelihood that

SMEs win procurement contracts increases.

3.2 RDD applied to public procurement value thresholds

Thresholds are abundant in public procurement legislations around the world, and so has become the econometric literature trying to take advantage of them for exercises of regression discontinuity designs. The European Union, for example, has a threshold for procurement values above which the procedures must abide to the European Directives, while values below them may be regulated by the Member States' own local legislations (Tas, 2022) In Brazil, not only there is a value threshold for set asides, but also for bid waivers (Fazio, 2022) and for the modalities of sealed bid procedures that have existed since 1986 and prevailed until the mid-2000s, when the reverse hybrid auctions (especially the e-auctions) became the majority.

In the literature, the papers by Coviello and Mariniello (2014) and Coviello et al. (2018) explore differences in entry originated by differences in rules below versus above legal value thresholds, after either rejecting the hypothesis of manipulation or dropping observations where such manipulation might have taken place. Papers by Palguta and Pertold (2017), Szucs (2023), Tas (2022) and Fazio (2022), on the other hand, focus on findings of manipulation, be it apparently perverse (the first three cases) or benign (the last author).

Coviello and Mariniello (2014) study whether, and how a requirement for increased publicity affects the number of potential entrants in Italian public procurement auctions, by using a discontinuity in the legal requirements for publicity: auctions with a reserve price exceeding 500,000 euros are required by law to be publicized more broadly in the Regional Official Gazette and in two provincial newspapers, while those below the threshold may be publicized only on the notice board in the premises of the public administration. They find an effect of increased publicity on the number of bidders participating in the auctions by 9.3%, and on the winning rebate by 7%, notwithstanding the unconventional feature that the awarding criterion of the auctions studied is the Average Bid – as described by Decarolis (2014, 2018) – but in consistency with the theoretical results obtained by Conley and Decarolis (2016). Last, but not least, they show that publicity increases the probabilities that the winner comes from outside the region and that it is a large company, as well as it increases the number of repeated winners. It is worth noting, first of all, that the authors rule out the possibility of perfect manipulation of an auction's value (reserve price) around the discontinuity threshold, using graphical and statistical tests discussed by McCrary (2008) and Lee (2008), and they also verify that no other policy, such as a change in the awarding mechanism, exists that could confound the estimates of the causal effect of publicity.

Another threshold in the Italian procurement legislation is explored by Coviello et al. (2018). They discuss the theoretical and empirical literature, as well as the (heated) political debate on the pros and cons of giving discretion to the procurement officer, and then explore a value threshold below which the procurement officer enjoys some discretion to choose the contractor for a public work in Italy, without being mandated to run an open auction[7]. Again, they test whether the reserve price of the project is not perfectly manipulated around the discontinuity threshold, using graphical and statistical tests by McCrary (2008) and Lee (2008), and they focus on the sample of projects for construction works that "*do not show sorting around the threshold*". They drop from their sample "*roadworks where bunching around the threshold appears to be a problem*" and further select an interval around the discontinuity threshold of their sample by using an optimal bandwidth method developed by Imbens and Kalyanaraman (2012), the main reference in optimal bandwidth by that time. Their main finding is that increased discretion causes a significant increase in the probability of a firm being repeatedly awarded contracts by the same procuring agency. This might entail different interpretations, such as networking, undue favoritism to friends, lower setup costs, or blunt bribing. By checking other procurement outcomes (number of bidders, rebates, size of the winners, distance from the procuring agency, duration of works, value amendments to the contracts) in a close neighbourhood of the discontinuity threshold, they do find evidence of dominance of the positive effects: lower total duration of the works; awards to larger (incorporated) firms – usually able to implement better quality control systems –; and reduction of the number of bids, which may save administrative costs associated to bid screening. Other outcomes, such as the winning rebate, cost overruns, and the probability of a local firm being awarded, are not found significantly affected by the degree of discretion.

Palguta and Pertold (2017) study a reform in the Czech public procurement legislation that gave more autonomy to procurement officers below a certain (estimated) threshold. A value threshold is set below which the procurement officer may preselect any five contractors for the bidding process, thus releasing them from running purely open auctions. They quantify the extent of manipulation using two empirical strategies. In the first, they use the methodology presented in Chetty et al. (2011) to estimate the cross-sectional counterfactual distributions of procurement value. The second strategy, an extension of their own, employs the distributions of procurement value before the reform as counterfactual. Both methods find a substantial impact of the policy reform on the extent of manipulation, albeit with lower levels in the second method. Manipulation, on its turn, is associated with a threefold increase in the probability that contracts are awarded to firms with anonymous owners. This should be an important warning, because the anonymity of company owners may hide shareholders with conflicts of interest, such as the procurement officers themselves, thus raising concerns about corruption and lack of transparency. However, due to data limitations, they are not able to provide evidence of overall losses in welfare.

[7] To be more precise, "*works with a value above the threshold have to be awarded through an open auction in almost all cases. Works below the threshold can more easily be run through a restricted auction, where the buyer has discretion in terms of who (not) to invite to bid*" –Coviello et al. (2018), p.716.

Szucs (2023) investigates the determinants and consequences of buyer discretion in public procurement, in the context of a Hungarian policy reform which allows a "high-discretion" procedure for contracts with estimated value below a certain threshold. At the threshold, he documents "*large discontinuities in procurement outcomes, but (..) also (...) a discontinuity in the density of contract values*", thus indicating manipulation of values to avoid auctions. He exploits the time variation of the policy reform and finds that discretion increases prices and results in the selection of less productive contractors; with a structural model, the amounts are respectively 2 percent and 1.6 percent. Further simulations of the structural model suggest that the optimal threshold would be about a third of the one actually employed. Finally, he shows that high discretion benefits firms with connections to the party of the central government, and that the procuring agents are "*willing to sacrifice more contract value to increase their discretion if more connected firms are operating in the market*".

Tas (2022) is yet another author to study manipulation of contract values in order to enjoy more discretion, this time using procurement data from 5.3 million awards in the European Economic Area, Switzerland, and North Macedonia. He also employs a RDD manipulation test to identify authorities that have higher probabilities of bunching estimated values below EU thresholds. He then examines this manipulative scheme on public procurement outcomes. His findings point out that 10-13% of the authorities have very high probabilities of manipulating estimated costs, and those who employ the bunching scheme are less likely to adopt competitive procurement procedures like open procedure (first price auctions). Finally, procurement prices are significantly (10 to 18%) higher in tenders conducted by bunching authorities.

Fazio (2022) also studies a value threshold in Brazilian legislation below which procurement officers enjoy high discretion and are able to award a contract directly without any auction. Differently from the previous articles, however, he documents an efficiency-quality trade-off: products purchased with higher discretion are 23 percent more expensive than those purchased via auctions; however, at least half of this overpricing is explained by the purchase of higher-quality products (as rated by consumer association tests). This trade-off is more pronounced when the value is manipulated to be set below the threshold. Finally, although enjoying more quality in some goods may be seen as an increase of welfare of the officers themselves, a particular set of goods has a broader welfare effect: public hospitals that use discretion to procure better essential medicines have experienced decreased inpatient mortality.

4 Methodology

4.1 The set-aside program in Brazil

As mentioned in section 2, the Brazilian legislation fostering MSEs opted for two strategies of setting small lots for exclusive bidding by this segment. The first one is based on a value threshold of BRL 80 thousand, below which only MSEs are allowed to participate. The second one is a 25% share of the yearly purchase value. Still, before an amendment to the MSE General Law (Supplementary Law 123) in 2014, procuring entities were not required to label all applicable lots as set-asides. Therefore, by that time the use of set-asides was completely voluntary.

As we show in subsection 4.3, the shift from "may-have to "must-have" set-asides after 2014 is significant, but not complete. The explanations for this are threefold[8]:

1. While the wording in the SL 123 (section 48) became clearer in the amendment that the threshold is applied to "procurement item value" instead of the "procurement value" (a lot is called "item" in Brazilian legislation"), some legal scholars are still able to diverge whether the relevant value refers to the lot's reserve price or the whole batch's reserve price (i.e., the sum of the reserve prices for all lots in the batch);

2. Section 49 in the same law provides for two ways out from the set-asides, one of them being that the procurement officers may claim that "*the differentiated and simplified treatment for MSEs [whatever preference instrument is utilized] is not advantageous for the public administration or represents damage to the aggregate or to the complex of the object to be contracted*";

3. The other way for the procurement officer to opt-out is to claim that "*there is not a minimum of three competitive suppliers classified as MSE based locally or regionally and capable of meeting the requirements established in the invitation to tender*".

In fact, in contact with some procurement practitioners, we have become aware of a specific phenomenon: the head of procurement in a major agency that procures on behalf of university hospitals remarked that a great number of lots initially labeled as "under threshold set-asides" did not produce an award: either no qualified bidder showed up, or the winning bid exceeded the reference (reservation) price[9]. These lots had to be reopened as "open access" lots, and therefore the agency decided to rule out set-asides and produced this justification in written to be taken into consideration by upcoming audits that may question the absence of set asides. They conjecture that pharmaceuticals may have greater barriers to entry for smaller wholesalers, and this might bias the non-compliance rate in our sample, but we have not noticed any outstanding difference between the graphs including and excluding pharmaceuticals, so we have kept them in the final sample[10]. Table 1 shows the number and relative frequency of reopened lots for the

[8] Yet another escape for the officer would be the quite long set of situations that justify a bid waiver (e.g. an even lower reserve value, below the threshold for bid waiver) or unenforceable tender (basically situations of monopoly), but in these cases the purchase is not recorded as a tender, rather they are classified as direct awards, and hence they do not enter into the sample (moreover, such direct awards do not identify possible substitute bidders).

[9] This anecdotal account was later reinforced by our estimates below, which suggest that the number of bidders is indeed reduced below the value threshold, and prices rise.

[10] An interim estimation of our own found that "reopened" lots jumped from 0.57% of the set-asides in 2014,

years 2007-2021.

Year	Total reopened lots	Total set-asides	Total opened lots	Ratio of reopened and set-aside lots	Ratio of reopened and total lots
2007	94	26,787	469,604	0.35	0.02
2008	452	96,255	482,774	0.47	0.09
2009	559	101,781	506,919	0.55	0.11
2010	379	113,925	476,240	0.33	0.08
2011	317	126,550	451,778	0.25	0.07
2012	249	93,588	533,596	0.27	0.05
2013	749	142,971	552,924	0.52	0.14
2014	1,025	179,498	447,333	0.57	0.23
2015	2,015	246,008	314,484	0.82	0.64
2016	3,666	317,577	256,684	1.15	1.43
2017	3,339	344,725	229,521	0.97	1.45
2018	5,209	417,945	247,249	1.25	2.11
2019	7,286	476,976	270,432	1.53	2.69
2020	7,736	633,590	235,085	1.22	3.29
2021	12,274	755,483	312,219	1.62	3.93

Table 1: Number of reopened lots, for the years 2017-2021.

4.2 Data

The Brazilian Federal Government has been consolidating since the turn of the century a repository of data for multiple stages of the procurement cycle, named Comprasnet. Our primary source is then this repository, which comprises two major databases:

1. Data Warehouse (DW) Siasg, which comprises attributes and metrics from purchasing units and their agents, registered bidders, auctions and auction lots, physical-financial schedules, contracts, inspections (in case of construction services) and payments[11].

2. Data Warehouse Comprasnet covers specifically additional features of the hybrid reverse auctions (both electronic and face to face), including bid protests, and detailed timestamps for each bid and every incident recorded from the beginning of the auction until the confirmation of the award. It also records the self declaration of the bidders as small or micro enterprise for each auction[12].

before the reform, to 1.62% in 2021; when calculated with respect to the total number of open-access lots, the share of reopened lots rose from 0.23% to 3.93% along the same period. This behavior reflects the increased use of set-asides in Federal procurement as a whole, while apparently also a great deal of failure in attracting competitive MSEs to these same lots.

11 More recently, since 2020, all contractual and financial operations have been moved to a new module named *Comprasnet Contratos*.

12 It is worth noting that the firm may become ineligible for the MSE preferential treatment along the year, and the auctioneer may be able to find out whether the firm has already earned a turnover above the legal threshold by consulting a repository of payments made by the Government to each and every supplier along the year. This reassures us that the flag for MSEs in this dataset is much more reliable than the current status kept in the Suppliers' Registry (Sicaf) − which is fed into DW Siasg − or even the datasets from the Labor Ministry or the Federal Revenue Service, which only collect the MSE status valid at the end of the year.

We restrict our analysis to goods, thus excluding procurement for services. We also exclude observations with missing supplier data and auctions with international suppliers. Our data ranges from 2009 to 2018 and covers bidder and bid information for e-auctions, for a selection of goods. The subset of products we choose for analysis are some of the most frequently bought essential inputs, presented in Table 2, as well as **all** pharmaceutical drugs. We select these materials due to their relative homogeneity and good description (very standardized, in the case of pharmaceuticals; much less so for the other categories), which facilitates comparisons between auctions.

Observations are divided into sets of goods classified with the same code. For example, a drug code describes the active ingredient (or the combination of them), its strength(s), and pharmaceutical form (solid oral, liquid oral, injectable, topic). Regarding the measure units, it is worth noting that solid oral drugs are typically quoted and bidden in single units (mainly one tablet or one capsule). Liquid orals, topics and injectables are less standardized: for example, an injectable, after controlling for the substance and strength, may be sold in ampoules or syringes of different sizes.

Other merchandise – such as nails, paint, gasoline, thinner and solvent – may be sold in yet more diverse pack sizes, and the list of such supply units is open and interminable, subject to infinite variation of wording and numeric formats (including blunt typos and misspellings). This required from us an effort of standardization, so as to reach a degree of homogeneity and enable us to replicate the strategy of Fazio (2022), who considered as "product" a combination of good and supply unit.

Category
Acrylic Paint
Adhesive label
Construction Nail
Construction Wood
Diesel
Enamel Paint
Fluorescent Lamp
Galactosamine
Gasoline
Liquid Solvent
Paint
Paint Thinner
Plastic Bags
PVC tubes
Stamp
Water

Table 2: List of selected essential input categories

A secondary source we utilize is firm level data taken from the *Relação Anual de Informações Sociais* (RAIS), a federal annual survey of Brazilian employment and companies conducted by the Brazilian Ministry of Labor. RAIS provides, *inter allia*, starting dates, as well as the location, of all suppliers that placed bids in procurement auctions. It is worth noting that the data are self declared by the firms, but the submission of information is mandatory and subject to a monetary penalty in case of failure; as the penalty is the same regardless of the firm's size, underreporting is expected to be much lower for MSEs[13].

We merge the data from DW Siasg with RAIS, to determine company age and the distance between suppliers and purchasing units.

We tabulate the supplier database to analyze auction data, resulting in a final database of 570,000+ auctions. Section 5 provides descriptive statistics for the final data set.

For a better adjustment of price regressions, we followed Fazio (2022) and did a thorough standardization of the supply units, so that prices would be controlled for a same product (defined as the category of material in a same pack size) and quarter[14].

4.3 Empirical Strategy

The parameter we want to identify is the effect of the set aside policy favoring MSEs on some characteristics of the auction. In particular, we will analyze whether the designation of exclusive lots for MSEs changes the degree of competitiveness of the auction. In this regard, we will investigate whether, and to what extent, the number of participating companies, the number of participating MSEs, the discount (relative to the auction reserve price) and the probability of the auction going to the open bidding stage are altered when the auction is either preferentially or exclusively aimed at MSEs.

The simplest and most intuitive identification strategy would be based on comparisons of set aside lots and other lots. However, this strategy may result in a biased estimate for the set aside effect, as this naive comparison incorporates other factors that would confound the real effect of the exclusive lot. This occurs to the extent that the decision of a procuring entity to allocate a lot exclusively to MSEs may be related to issues that also interfere with the configuration of the auction and the winning bid. For example, the purchase unit can opt for an exclusive lot in order to give preference to small local suppliers. Or it may happen that less structured

[13] It is worth noting that firms that closed along the year or just before the deadline for reporting are ordinarily missing in the data.

[14] Standardization of supply units is practically nihil in DW Siasg. The corresponding field is open, thus allowing any officer to entry a different pack size, and/or, worse, with a different writing of abbreviations, misspelling, etc. A better standardization was, however, undertaken for pharmaceuticals by the Ministry of Health. In particular, notice that solid oral pharmaceuticals are typically quoted in pills, capsules or tablets, while injectables may have different ampoule and syringe sizes. Table 3 displays a sample of supply units for selected product categories. In our dataset there are a total of 14399 unique products.

Product	
Description	Unit
Acrylic Paint	3.6 Liter Can
Adhesive label	Unit
Construction Nail	1 kg
Construction Wood	Sq. Meter
Diesel	Liter
Dipyrone	Dragee
Enamel Paint	3.6 Liter Can
Fluorescent Lamp	Unit
Galactosamine	Bottle
Gasoline	Liter
Ibuprofen	Capsule
Liquid Solvent	Liter
Nicotine	Tablet
Paint	3.6 Liter Can
Paint Thinner	Liter
Paracetamol	Pill
Plastic Bags	100 Units
PVC tubes	Unit
Stamp	Unit
Water	Bottle

Table 3: Examples of Products

procuring entities, which tend to make public notices of poorer quality, are eventually more concentrated among those that (do not) carry out set aside auctions.

In short, the probability of a set aside lot does not vary randomly across lots in our database. However, we can say that there is indeed an exogenous variation in this probability for a sub-sample of lots comprised by those whose values are in the neighborhood of BRL 80,000. Intuitively, we will explore the fact that the probability of having a set aside lot increases substantially when we compare lots with values slightly lower than 80 thousand BRL with lots valued slightly above 80 thousand BRL.

The hypothesis behind our strategy is that the procuring entities holding lots with values around BRL 80,000 are homogeneous in all other aspects; such as preferences for local suppliers, or ability to structure notices and auctions. This would ensure that the set aside probability actually grows exogenously around BRL 80K.

In formal terms we employ the **(partially-)fuzzy regression discontinuity design** based on a **local TSLS** regression:

$$y_a = \alpha + T_a\beta + f(X_a - X_0) + \epsilon_a, \tag{1}$$

where a indexes auctions, y is the outcome variable, T is the treatment (set-aside) dummy, $f(X_a - X_0)$ is a polynomial function of the re-centered reserve price of the batch ($X_0 = 80kBRL$) and ϵ is a mean-zero disturbance term.

Our interest relies on the estimation of β. As discussed above the probability of a set aside lot does not vary exogenously in general, which means that T_a andϵ_a would be correlated, biasing OLS estimates of β. In order to circumvent this problem we restrict the sample to lots whose reserve prices are close to $80kBRL$ and use the set-aside eligibility dummy ($D_a = 1\{X_a \leq 80k\}$) as an instrumental variable to the treatment dummy.

The strategy represented by the model above fits into the category of regression with fuzzy discontinuity.

Estimation is based on the widely employed **local polynomial parametric regression**, for which the implementation requires to deal with some extra details. We use the bias-corrected, robust estimator proposed by Calonico et al. (2014) with their corresponding **bandwidth choices** procedure described in Calonico et al. (2020). The function $f(X_a - X_0)$ is specified as a linear or a quadratic **polynomial**; where we take the first as the benchmark. We also use distinct weighting schemes for kernels having the triangular one as our preferred specification.

5 Results

5.1 Preliminary (lack of) evidence on non-random selection

As mentioned in the previous section, a crucial identification hypothesis that we rely on is that the procuring entities holding lots with values around BRL 80,000 are homogeneous in all other aspects. There is no direct test to validate this assumption as it refers to non-observable characteristics of either lots or purchasing units. However we may provide two types of indirect evidence following standard procedures in the applied literature. First we may inspect for bunching patterns of lots around the eligibility threshold in either side. Such pattern would indicate possible manipulation by purchasing units to either have their lots available (if on the left) or not (if on the right) to set aside policy. If that is the case, then we may have non-comparable lots below and above the threshold value of 80k BRL; even within its neighborhood; as some procuring entities may have manipulated the reference price aiming at the eligibility for the set aside regime. That would invalidate our identification hypothesis mentioned above.

To validate this hypothesis, we test our sample for manipulation of the running variable using local polynomial density estimation, as presented in Cattaneo et al. (2018). Figures 5 and 6 present the results graphically, while Table 4 shows the used specification and values. The test failed to reject the null hypothesis of no manipulation into treatment for both analyzed periods.

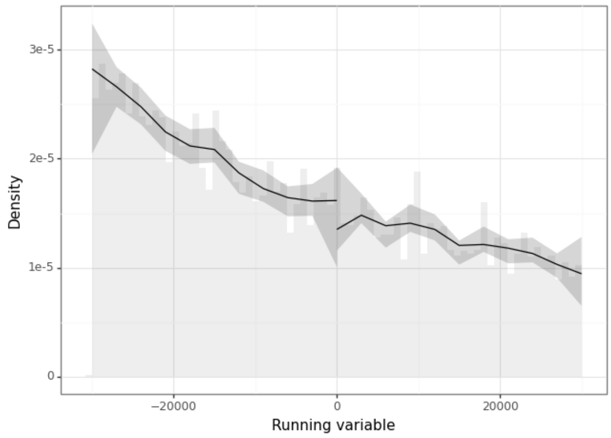

Figure 5: Manipulation test - 2012-2014.

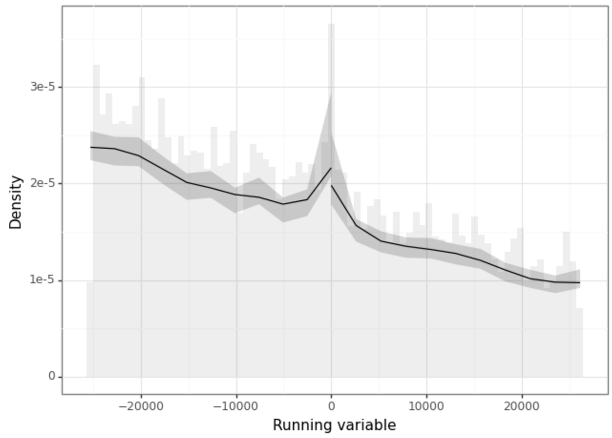

Figure 6: Manipulation test - 2016-2018.

Model:	unrestricted			
Kernel:	triangular			
BW method:	estimated			
VCE:	jackknife			
	2012-2014		2016-2018	
c = 0	Left of c	Right of c	Left of c	Right of c
Number of obs:	5898	3622	7791	4524
Eff. number of obs:	1782	1873	1976	1621
Order est. (p):	2	2	2	2
Order bias. (q):	3	3	3	3
BW est.	11266.7654	13735.1941	8431.7965	8716.7869
Method:	T	P > |T|	T	P > |T|
Robust	-0.3288	0.7423	-0.9736	0.3302

Table 4: Manipulation Tests for all periods

For additional evidence of balanced samples on either side of the threshold, we ran a regression with four different subsamples, comprising the distribution of procuring entities by four major government activities; namely education, health, public administration, and army. If there is no evidence of manipulation, then being left or right of the threshold would not influence the probability of a tender being from a certain procuring entity. Our model is as follows:

$$y_i = \beta_1 runvar_i + \beta_2 less80k_i + \mu_p + \lambda_t + \epsilon_i, \tag{2}$$

Where y is a procuring entity sectoral dummy; $runvar$ in the distance to the threshold; $less80k$ is a dummy which is 1 for reference values under 80k and 0 otherwise; μ are product fixed effects; λ are trimester fixed effects; and ϵ is the error term. Tables 5 and 6 show our results. Again, we find no evidence of manipulation, shown by the lack of significance in the $less80k$ dummy.

In sum, there is no strong evidence of lots been non-randomly selected to the left or right of the

VARIABLES	(1) Education	(2) Health	(3) Army	(4) Public Administration
runvar	-3.14e-07	1.16e-06**	-5.23e-07	2.26e-07
	(7.54e-07)	(4.98e-07)	(7.41e-07)	(8.00e-07)
less_80k	0.00621	0.0188	-0.00775	-0.0119
	(0.0214)	(0.0142)	(0.0211)	(0.0227)
Observations	6,252	6,252	6,252	6,252
R-squared	0.000	0.001	0.000	0.000

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 5: Manipulation in subsamples - 2012-2014

VARIABLES	(1) Education	(2) Health	(3) Army	(4) Public Administration
runvar	3.65e-07	6.75e-07*	-1.99e-06***	-4.95e-07
	(6.62e-07)	(4.04e-07)	(6.39e-07)	(6.79e-07)
less_80k	0.0102	0.0181	-0.0326*	-0.0134
	(0.0183)	(0.0111)	(0.0176)	(0.0187)
Observations	8,491	8,491	8,491	8,491
R-squared	0.000	0.000	0.002	0.000

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6: Manipulation in subsamples - 2016-2018

80k BRL threshold. Therefore in what follows we will assess the effects of the set aside regime comparing those lots with values just above and just below the 80k BRL threshold, as done in the sharp RDD strategy.

5.2 Set aside and auction characteristics

Figures 7 and 8 show a stable number of participants close to seven firms in both periods, 2012-2014 (figure 7) and 2016-2018 (figure 8). Such stable pattern appears in particular in the 80k BRL threshold neighborhood.

Figures 9 and 10 on the other hand show an increase in number of MSEs for lots with reference value just below 80k BRL. Therefore, firms size composition must have changed towards an increased share of MSEs for these lots.

Has this change in firms' composition affected the degree of competition for these lots? In one dimension the answer is apparently no. Figures 11 and 12 indicate that the share of auctions

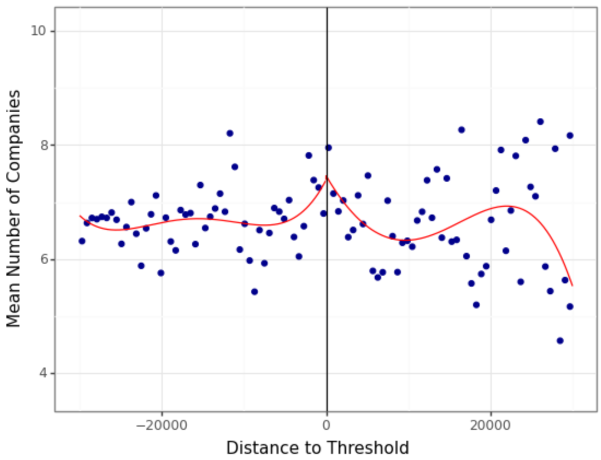

Figure 7: Number of companies by lot 2012-2014.

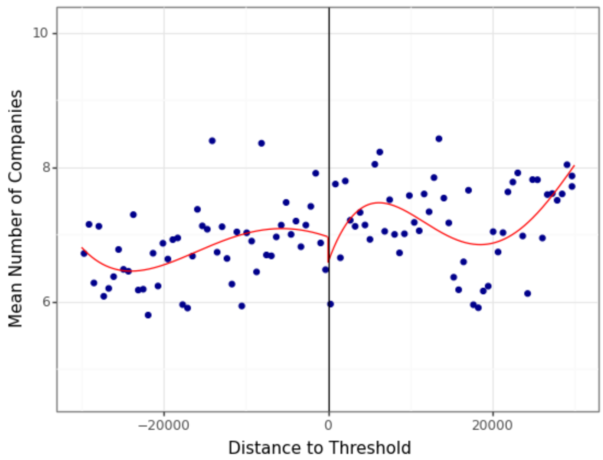

Figure 8: Number of companies by lot 2016-2018.

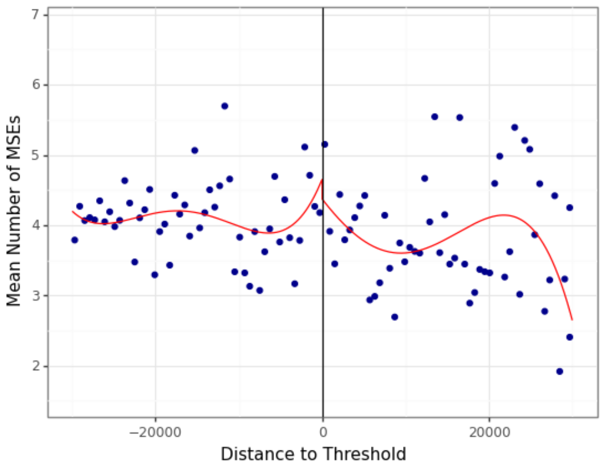

Figure 9: Number of MSEs by lot 2012-2014.

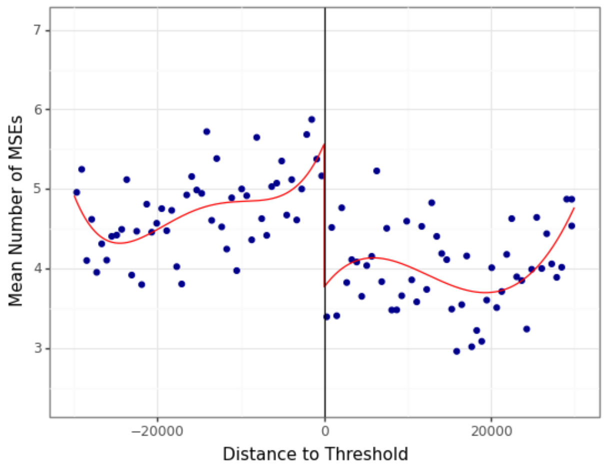

Figure 10: Number of MSEs by lot 2016-2018.

reaching the open bid stage remained stable.

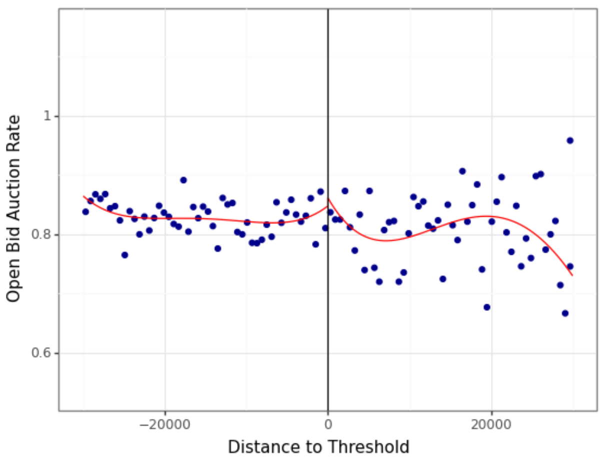

Figure 11: Open Bid Auction Rate 2012-2014.

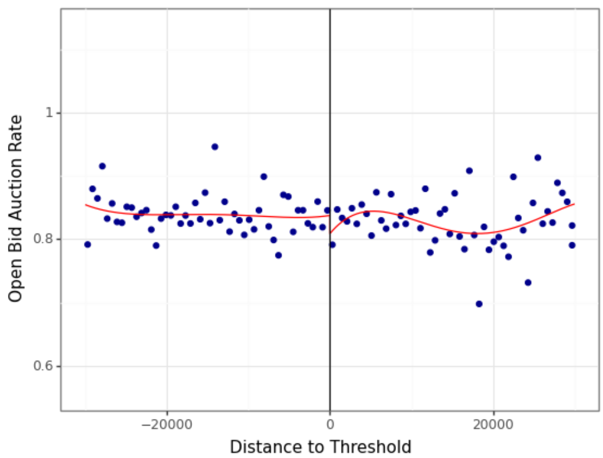

Figure 12: Open Bid Auction Rate 2016-2018.

On efficiency grounds figures 13 and 14 provide information on discount, defined as the difference between winner bid and reference price, relative to the later. One can see that we have no evidence on effects of set aside on discounts as their values remain stable within the threshold neighborhood; both for the 2012-2014 (figure 13) and 2016-2018 (figure 14) periods.

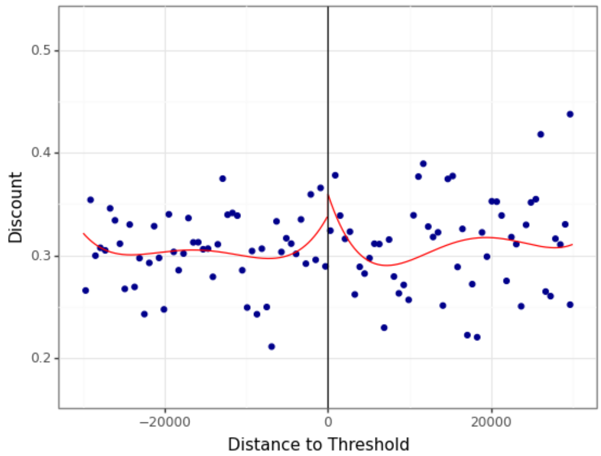

Figure 13: Mean discount rate 2012-2014.

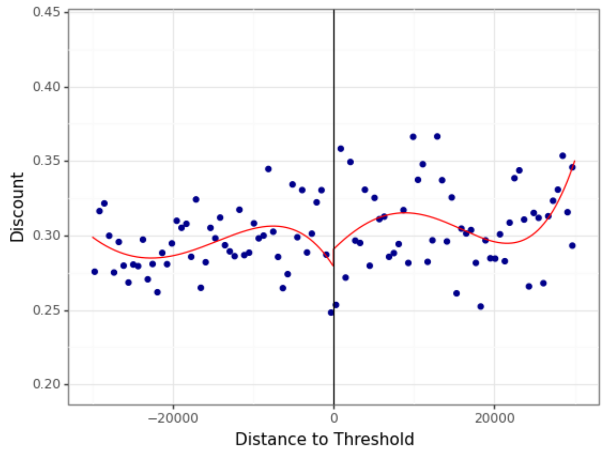

Figure 14: Mean discount rate 2016-2018.

5.3 Main results: set-aside impacts

5.3.1 Outcomes studied

In this subsection we analyse set-aside impacts on the outcomes discussed above through the fuzzy regression discontinuity framework. We first show estimates of set-aside impacts for the 2016-2018 period, when both the eligibility criterion and the obligatoriness of set asides became more clear in the law.

The first outcomes we study relate to participation and composition of the bidders: (i) total number of bidders in the lot, regardless of size; (ii) number of MSEs bidding in the lot; (iii) ratio of MSEs bidding in the lot; (iv) ratio of MSEs awarded; (v) the mean age of all bidders in the lot; (vi) a dummy indicating whether the bidders do lower their bids when the open cry stage is opened (they may refrain from proceeding, and, in case the front runner's bid is compatible with the reserve price, the auctioneer may close the auction and call the front runner to produce its documentation); (vii) the level of the discount obtained off the reserve price; (viii) distance of the supplier to the farthest purchasing agency (as the dataset includes frame contract awards).

Next we assess the effects of the set asides on prices in other dimensions: (ix) the price itself (in natural logarithms); (x) the ratio between price and the reserve price; and a few collusive markers, based on statistics of the bid distributions, which are largely utilized by Antitrust investigators around the world in proactive screening activities:

(xi) **Coefficient of variation of prices**:

$$CV_t = \frac{s_{b_t}}{\bar{b}_t},$$

where $\bar{b}_t$ is the average of all bids, and s_{b_t} is their standard deviation. Normalization by the mean renders this statistic invariant to scale and therefore amenable to comparisons across lots of different product categories.

(xii) **Skewness**:

$$Sk_t = \frac{n_t}{(n_t - 1)(n_t - 2)} \sum_{i=1}^{n_t} \left(\frac{b_{it} - \bar{b}_t}{s_{b_t}} \right)^3 \tag{3}$$

where n_t denotes the number of bidders in tender t. If the differences in the covering bids are small and if the differences in the first and second lowest bids are significant, the asymmetry will be more evident. Therefore, a greater negative asymmetry it is expected to be found for the (biased) distribution of bids within the collusive tenders.

(xiii) **Kurtosis**:

$$Kurt_t = \frac{n_t (n_t + 1)}{(n_t - 1)(n_t - 2)(n_t - 3)} \sum_{i=1}^{n_t} \left(\frac{b_{it} - \bar{b}_t}{s_{b_t}} \right)^4 - \frac{3 (n_t - 1)^2}{(n_t - 2)(n_t - 3)}, \tag{4}$$

where n_t denotes the number of bidders in tender t. Thus, if there is a transformation in the distribution of bids, and the kurtosis statistic reaches higher values, an evidence of bid convergence is inferred, probably arising from communication among the bidding ring members.

(xiv) **Percentage difference** between first and second bids, **after discarding disqualified bids**[15]. It seems reasonable to assume that bidding cartels manipulate the difference between the first and second lowest bids to determine which cartel company will be the designated winner. According to Pesendorfer (2000), to ensure that the designated winner of the cartel actually wins the bid, the companies in collusion maintain a certain distance between the first and second lowest bids. To analyze this difference, Imhof (2018) proposes the following formula:

$$Diff.Perc_t = \frac{(b_{2t} - b_{1t})}{b_{1t}} \tag{5}$$

where b_{1t} is the lowest bid and b_{2t} is the second lowest bid in bid t.

[15] Note that qualification documents in hybrid auctions are opened only after the closure of the open bid session, and only of the front runner. If they are rejected, the runner up is summoned to present their own documents, and so on.

(xv) **Relative distance** between first and second bids, **after discarding disqualified bids**. This indicator was first proposed by Abrantes-Metz et al. (2006) and also utilized by Athey et al. (2004); Harrington-Jr and Chen (2006); Jiménez and Perdiguero (2012); Imhof et al. (2018), as well as applied in successful screenings of bid-rigging schemes in Switzerland:

$$DR_t = \frac{b_{2t} - b_{1t}}{s_{\breve{b}_t}, (6)}$$

where $s_{\breve{b}_t}$ is the standard deviation of losing bids. If DR_t is greater than one, the difference between first and second bids exceeds the differences among losing bids. In this case, the bid distribution is negatively skewed. If $DR_t = 1$, there are not any significant differences in the bid distribution. However, if $DR_t < 1$, this indicates that the differences between the two lowest bids is small, thus signalling a high level of competition among bidders. Therefore, for tenders affected by a cartel, $DR_t > 1$ is expected.

(xvi) Another statistical marker adopted by Huber and Imhof (2019) is obtained from the normalization of the difference between the two smallest bids $(b_{2t} - b_{1t})$, dividing them by the average of the differences between pairs of adjacent bids (b_{jt}, b_{it}). The **normalized distance** ND_t is given by equation 7 below:

$$DN_t = \frac{(b_{2t} - b_{1t})}{\frac{\left(\sum_{i=1,j=i+1}^{n-1} b_{jt} - b_{it}\right)}{n-1}}, \tag{7}$$

Thus, bids are arranged in ascending order. When $ND_t > 1$, the difference between the second and first lowest bids is greater than the average difference between all adjacent bids in the auction.

(xvii) **Bid spread**, proposed by Huber and Imhof (2019); Huber et al. (2022); Wallimann et al. (2022):

$$SPD_t = \frac{b_{max_t} - b_{min_t}}{b_{min_t}}, \tag{8}$$

where b_{min_t} is the lowest bid, and b_{max_t} is the greatest bid in tender t.

Notice that markers based on the differences between lowest bids have been developed in Switzerland after observing that road contractors had to be at a distance of the designed winner so as not to be a threat to the arranged outcome (tenders were actually score auctions, and some detail in the project description might reverse the agreed order of bids), and at the same time the bidders did not want to signal a too high cost. In the present case of merchandise supply, this concern may be diluted, and such markers as ND_t, RD_t, $Perc.Diff_t$ and SPD_t may be useless. (Excess) kurtosis, skewness, and CV, in turn, are typically good signals of the existence of cover bidding.

5.3.2 Estimates obtained

The results of local regressions without fixed effects indicate that set-aside lots in fact attract more MSEs. The first column of Table 7, suggests an impact of three more MSEs when lots are allocated exclusively for them. Such increase in MSEs participation associated to set-aside is mainly driven by young firms, so much so that the average age of participating firms has a statistically significant drop of more than 4 years (52 months). In the present setting, no significant change is observed in the total number of bidders.

Taking these results by face value, there are two possible consequences of the increase in MSEs participation. First, there can be more competition among MSEs. However in the third column we see that at least one indicator of the degree of competition – the probability of having bids in the open cry stage – is not affected by set asides. Second, having an MSE as the winner may favour local firms and compromise efficiency. As shown in the last two columns of the same table, neither effect seems to take place.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	-0.119	3.025	0.513	0.613	-53.968	0.001	-0.007	-65.202	0.958	0.015
Observations	[5769 ; 3721]	[4809 ; 3347]	[4975 ; 3408]	[7179 ; 4315]	[6516 ; 4044]	[8262 ; 4694]	[5885 ; 3765]	[5784 ; 3721]	[6403 ; 4001]	[4804 ; 3346]
Robust 95% CI	[-1.486 ; 1.392]	[1.679 ; 4.361]	[0.442 ; 0.602]	[0.497 ; 0.699]	[-74.066 ; -33.732]	[-0.098 ; 0.079]	[-0.094 ; 0.066]	[-304.403 ; 146.886]	[0.414 ; 1.560]	[-0.131 ; 0.165]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	0.628	0.587	0.035	0.043	8.728	0.038	0.033	98.275	0.246	0.065
Conventional p-value	0.850	0.000	0.000	0.000	0.000	0.978	0.837	0.507	0.000	0.823
Robust p-value	0.949	0.000	0.000	0.000	0.000	0.838	0.726	0.494	0.001	0.823
Order Loc. Poly (p)	1	1	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2	2	2

Table 7: 2016-2018 - Optimal Bandwidth. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020) We use a triangular kernel with a linear local polynomial and quadratic bias.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	0.472	1.767	0.265	0.040	-103.645	0.095	0.011	-307.909	0.718	-0.049
Observations	[5223 ; 3376]	[5977 ; 3644]	[6209 ; 3720]	[6004 ; 3651]	[4648 ; 3147]	[7218 ; 4022]	[2570 ; 2043]	[6340 ; 3764]	[5857 ; 3600]	[5576 ; 3499]
Robust 95% CI	[-10.144 ; 6.363]	[-5.435 ; 8.316]	[-0.251 ; 0.740]	[-0.695 ; 0.710]	[-276.853 ; 1.406]	[-0.451 ; 0.578]	[-0.799 ; 0.459]	[-1383.307 ; 1004.367]	[-2.488 ; 4.350]	[-1.198 ; 0.924]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	3.068	2.570	0.189	0.266	52.950	0.196	0.246	445.885	1.300	0.405
Conventional p-value	0.878	0.492	0.160	0.882	0.050	0.629	0.966	0.490	0.581	0.904
Robust p-value	0.653	0.681	0.334	0.982	0.052	0.808	0.596	0.756	0.594	0.800
Order Loc. Poly (p)	1	1	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2	2	2

Table 8: 2012-2014 - Optimal Bandwidth. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use a triangular kernel with a linear local polynomial and quadratic bias.

Therefore, in principle the change in law was quite effective to achieve the goal of attracting more MSEs to auctions in public procurement. Table 8 displays analogous results for the 2012-2014 period, but estimates are not statistically significant different from zero at conventional levels.

All the results above were derived using the optimal symmetrical bandwidth proposed by Calonico et al. (2020). We have also checked the robustness of our main results to this choice. Tables 9 and 10 provide analogous estimates for both periods considered, but using an asymmetrical bandwidth. As one can see, our main conclusions remain valid under this alternative choice of bandwidth. The magnitudes are also similar: approximately three more MSEs come up to bid in the lot on average, and the mean age's drop is only slightly less than four years, according to the results from 2016 on, while the number of MSEs before 2015 was not significantly affected and the mean age's drop was still significant, but with a much greater variance. However, a novel result comes up regarding the price level: a one hundred percent rise.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	-0.214	2.950	0.501	0.621	-52.759	-0.004	-0.012	-80.494	1.076	0.041
Observations	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]	[4809 ; 4524]
Robust 95% CI	[-1.791 ; 1.473]	[1.540 ; 4.467]	[0.466 ; 0.639]	[0.501 ; 0.728]	[-78.931 ; -32.643]	[-0.093 ; 0.125]	[-0.113 ; 0.054]	[-322.067 ; 187.201]	[0.285 ; 1.603]	[-0.087 ; 0.232]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	0.606	0.539	0.032	0.042	8.538	0.040	0.032	93.923	0.243	0.060
Conventional p-value	0.724	0.000	0.000	0.000	0.000	0.921	0.697	0.391	0.000	0.491
Robust p-value	0.849	0.000	0.000	0.000	0.000	0.773	0.490	0.604	0.005	0.372
Order Loc. Poly (p)	1	1	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2	2	2

Table 9: 2016-2018 - 20-30k Asymmetric Bandwidth. Estimates based on local regressions implemented with rdrobust package. We use an asymetric bandwidth of (20000, 30000). We use a triangular kernel with a linear local polynomial and quadratic bias.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	1.106	2.026	0.278	0.079	-97.250	0.185	0.140	-131.658	0.817	-0.051
Observations	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]	[3466 ; 3622]
Robust 95% CI	[-9.188 ; 9.472]	[-5.639 ; 11.211]	[-0.204 ; 1.012]	[-0.666 ; 1.020]	[-321.295 ; -13.015]	[-0.558 ; 0.781]	[-0.425 ; 0.716]	[-1299.084 ; 1668.280]	[-3.176 ; 5.107]	[-1.480 ; 0.965]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	3.235	2.897	0.213	0.298	53.152	0.238	0.201	518.127	1.450	0.437
Conventional p-value	0.732	0.484	0.192	0.792	0.067	0.437	0.488	0.799	0.573	0.907
Robust p-value	0.976	0.517	0.193	0.681	0.034	0.744	0.617	0.807	0.648	0.680
Order Loc. Poly (p)	1	1	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2	2	2

Table 10: 2012-2014 - 20-30k Asymmetric Bandwidth. Estimates based on local regressions implemented with rdrobust package. We use an asymetric bandwidth of (20000, 30000). We use a triangular kernel with a linear local polynomial and quadratic bias.

These results persist when we run a local regression using Epanechnikov kernel – see Tables 11 and 12 or a quadratic point estimator – see Tables 13 and 14.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	-0.230	2.989	0.503	0.602	-54.910	-0.002	-0.004	-67.178	1.020	0.001
Observations	[5461 ; 3622]	[4649 ; 3259]	[4723 ; 3293]	[6315 ; 3956]	[8726 ; 4834]	[7909 ; 4560]	[5143 ; 3476]	[5305 ; 3548]	[5669 ; 3694]	[4441 ; 3139]
Robust 95% CI	[-1.626 ; 1.286]	[1.604 ; 4.305]	[0.430 ; 0.591]	[0.477 ; 0.685]	[-73.981 ; -36.080]	[-0.104 ; 0.074]	[-0.101 ; 0.069]	[-314.036 ; 145.798]	[0.466 ; 1.637]	[-0.150 ; 0.154]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
Conventional Std. Error	0.631	0.586	0.035	0.044	8.111	0.038	0.034	99.638	0.251	0.067
Conventional p-value	0.715	0.000	0.000	0.000	0.000	0.954	0.896	0.500	0.000	0.990
Robust p-value	0.819	0.000	0.000	0.000	0.000	0.741	0.716	0.473	0.000	0.982
Order Loc. Poly (p)	1	1	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2	2	2

Table 11: Epanechnikov Kernel with Optimal Bandwidth - 2016-2018. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use an Epanechnikov kernel with a linear local polynomial and quadratic bias.

Moving to the collusive markers, we initially do not find any suspicious result; they have not been significantly affected by the threshold or its enforcement either – Tables 15 and 16 – **except** for a non-negligible p-value (0.051) in the CV regression.

This motivated us to pursue a finer regression of the price level. For this purpose, we did the standardization of supply units for each product category, as previously described in Section 4.2 and, in particular on Table 3, and included both quarter and product fixed effects. We

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	0.724	2.119	0.270	0.036	-103.536	0.099	0.065	-323.648	0.162	-0.090
Observations	[4938 ; 3256]	[5351 ; 3420]	[4712 ; 3178]	[7622 ; 4129]	[4079 ; 2873]	[7491 ; 4096]	[2135 ; 1833]	[5961 ; 3640]	[9851 ; 4635]	[4995 ; 3285]
Robust 95% CI	[-10.494 ; 6.076]	[-5.810 ; 8.254]	[-0.409 ; 0.697]	[-0.564 ; 0.661]	[-279.432 ; 5.223]	[-0.391 ; 0.594]	[-0.777 ; 0.513]	[-1376.818 ; 1001.450]	[-1.771 ; 3.556]	[-1.240 ; 0.965]
Kernel Type	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov	Epanechnikov
Conventional Std. Error	3.067	2.615	0.208	0.229	54.004	0.187	0.252	442.705	1.016	0.418
Conventional p-value	0.813	0.418	0.195	0.875	0.055	0.595	0.796	0.465	0.874	0.830
Robust p-value	0.601	0.733	0.610	0.876	0.059	0.685	0.688	0.757	0.511	0.807
Order Loc. Poly (p)	1	1	1	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2	2	2	2

Table 12: Epanechnikov Kernel with Optimal Bandwidth - 2012-2014. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use an Epanechnikov kernel with a linear local polynomial and quadratic bias.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	0.142	3.056	0.529	0.597	-53.361	0.021	-0.009	-49.853	0.961	0.018
Observations	[6830 ; 4180]	[7369 ; 4371]	[7704 ; 4495]	[7941 ; 4575]	[8106 ; 4636]	[6727 ; 4132]	[7184 ; 4316]	[7556 ; 4444]	[8909 ; 4898]	[8728 ; 4834]
Robust 95% CI	[-1.421 ; 1.870]	[1.550 ; 4.468]	[0.455 ; 0.630]	[0.483 ; 0.717]	[-76.576 ; -29.816]	[-0.073 ; 0.152]	[-0.097 ; 0.075]	[-281.324 ; 226.256]	[0.226 ; 1.553]	[-0.121 ; 0.196]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	0.754	0.665	0.040	0.052	10.412	0.051	0.039	115.536	0.288	0.071
Conventional p-value	0.850	0.000	0.000	0.000	0.000	0.685	0.810	0.666	0.001	0.799
Robust p-value	0.789	0.000	0.000	0.000	0.000	0.493	0.803	0.832	0.009	0.640
Order Loc. Poly (p)	2	2	2	2	2	2	2	2	2	2
Order Bias (q)	3	3	3	3	3	3	3	3	3	3

Table 13: Quadratic Point Estimator with Optimal Bandwidth - 2016-2018. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use a triangular kernel with quadratic local polynomials and cubic bias.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
RD_Estimate	0.545	1.843	0.291	0.059	-122.245	0.114	-0.039	-269.285	1.085	-0.091
Observations	[12902 ; 5131]	[13644 ; 5218]	[12597 ; 5086]	[11456 ; 4930]	[11977 ; 5006]	[10569 ; 4767]	[3761 ; 2733]	[18867 ; 5699]	[9530 ; 4587]	[11568 ; 4944]
Robust 95% CI	[-10.366 ; 6.853]	[-5.855 ; 8.807]	[-0.385 ; 0.741]	[-0.839 ; 0.781]	[-302.519 ; -18.213]	[-0.710 ; 0.599]	[-1.018 ; 0.379]	[-1477.937 ; 843.892]	[-3.116 ; 5.129]	[-1.283 ; 1.069]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	3.307	2.884	0.221	0.318	55.992	0.258	0.286	460.491	1.641	0.467
Conventional p-value	0.869	0.523	0.188	0.853	0.029	0.659	0.893	0.559	0.509	0.846
Robust p-value	0.689	0.693	0.536	0.944	0.027	0.869	0.370	0.592	0.632	0.858
Order Loc. Poly (p)	2	2	2	2	2	2	2	2	2	2
Order Bias (q)	3	3	3	3	3	3	3	3	3	3

Table 14: Quadratic Point Estimator with Optimal Bandwidth - 2012-2014. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use a triangular kernel with quadratic local polynomials and cubic bias.

VARIABLES	(1) Coefficient of Variation	(2) Skewness	(3) Kurtosis	(4) Percentage Difference	(5) Relative Distance	(6) Normalized Distance	(7) Bid Spread
RD_Estimate	-0.136	-0.219	-0.070	0.132	-0.007	-0.365	-394.679
Observations	[6431 ; 4006]	[4728 ; 3294]	[5654 ; 3687]	[271 ; 247]	[93 ; 94]	[478 ; 448]	[4645 ; 3255]
Robust 95% CI	[-0.264 ; 0.000]	[-0.538 ; 0.105]	[-1.107 ; 1.034]	[-2.351 ; 0.197]	[-0.356 ; 0.275]	[-1.201 ; 0.268]	[-1085.266 ; 2125.364]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	0.057	0.140	0.467	0.230	0.119	0.290	312.727
Conventional p-value	0.017	0.117	0.881	0.566	0.954	0.207	0.207
Robust p-value	0.051	0.188	0.947	0.097	0.800	0.213	0.525
Order Loc. Poly (p)	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2

Table 15: Collusive Price Markers - 2016-2018. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use a triangular kernel with a linear local polynomial and quadratic bias.

then found the optimal bandwidths with the transformed (demeaned) log-prices, and finally ran the IV-HDFE (High Dimensional Fixed Effects with Instrumental Variables) for all of the outcomes again.

Controlling for these fixed effects brings to light effects that were previously under cover. According to Table 17, the bid discount does go down ten percentage points, the average (de-

VARIABLES	(1) Coefficient of Variation	(2) Skewness	(3) Kurtosis	(4) Percentage Difference	(5) Relative Distance	(6) Normalized Distance	(7) Bid Spread
RD_Estimate	-0.074	0.032	1.170	3.919	18.265	-1910.215	11.401
Observations	[5977 ; 3646]	[6165 ; 3709]	[5587 ; 3507]	[187 ; 197]	[75 ; 95]	[2871 ; 2225]	[239 ; 245]
Robust 95% CI	[-1.016 ; 0.480]	[-1.854 ; 1.622]	[-4.911 ; 6.756]	[-36.240 ; 16.305]	[-55.179 ; 50.584]	[-6284.251 ; 2029.357]	[-1044.734 ; 627.027]
Kernel Type	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular	Triangular
Conventional Std. Error	0.277	0.644	2.136	7.887	24.736	1938.270	410.354
Conventional p-value	0.791	0.961	0.584	0.619	0.460	0.324	0.978
Robust p-value	0.482	0.896	0.757	0.457	0.932	0.316	0.624
Order Loc. Poly (p)	1	1	1	1	1	1	1
Order Bias (q)	2	2	2	2	2	2	2

Table 16: Collusive Price Markers - 2012-2014. Estimates based on local regressions implemented with rdrobust package. Optimal Bandwidth obtained following Calonico et al. (2020). We use a triangular kernel with a linear local polynomial and quadratic bias.

meaned) price goes up 29.2 percent, and the total number of bidders is reduced: almost two bidders less. It is worth noting that the previous results of significantly greater MSE participation ratio and MSE win ratio, as well as significantly lower age, are maintained.

As regards the collusive markers, in the new 2SLS regressions (IV-HDFE) we do notice that after the 2014 reform, when set asides were more strictly enforced, the CV decreased and Skewness became more negative in that group, while kurtosis were not significantly different. Regarding the price markers related to distance among bids, they were either not significant (RD_t and NDt) or significant in the "wrong direction" (i.e., pointing out to lower levels).

In other words, blocking entry from larger firms apparently reduced competition by lowering the number of bidders (that is, more MSEs came at the expense of the exit of a larger number of larger bidders, so that the net balance was negative). This new situation favoured a skewer and more concentrated distribution of bids, and an overall rise in price level, which reflected in lower discounts off the reserve price and signalled in favour of a suspicion of collusion among bidders under the threshold.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
set_aside	-1.960***	0.687*	0.377***	0.541***	-53.24***	-0.0814**	-0.107*	29.67	0.292***	0.0834
	(0.484)	(0.385)	(0.0263)	(0.0386)	(8.588)	(0.0378)	(0.0599)	(97.07)	(0.0753)	(0.0665)
runvar	1.94e-05***	1.19e-05**	-3.71e-07	-4.19e-07	1.65e-05	-1.53e-08	-1.80e-06	7.84e-05	-2.17e-07	-2.16e-06**
	(6.42e-06)	(4.86e-06)	(3.10e-07)	(3.78e-07)	(0.000108)	(5.53e-07)	(3.00e-06)	(0.000971)	(9.82e-07)	(9.29e-07)
runvar2	-2.10e-10	1.01e-10	0**	-0	-8.46e-10	-5.15e-11*	7.47e-10	-2.67e-08	0	0
	(3.12e-10)	(2.22e-10)	(0)	(0)	(4.96e-09)	(0)	(6.12e-10)	(3.31e-08)	(0)	(0)
Observations	8,411	8,887	9,621	11,777	8,886	7,573	2,074	11,473	8,585	7,933
R-squared	0.070	0.005	0.370	0.297	0.108	0.012	0.008	-0.001	0.019	0.011

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 17: 2SLS Optimal Bandwidth - 2016-2018. Estimates based on local regressions using 2sls method implemented with the ivreghdfe package, controlling for product and quarter fixed effects. Optimal bandwidths obtained following Calonico et al. (2020). as implemented in the rdrobust package, using data after demeaning product fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	Number of Bidders	Number of MSEs	MSE Ratio	Ratio of MSEs Awarded	Mean Company Age	Open Auction Ratio	Discount	Distance to Supplier	ln(price)	ln(ratio)
set_aside	-1.404	0.394	0.140	0.145	-22.31	0.109	0.107	-354.4	-0.591	-0.00887
	(2.108)	(1.796)	(0.135)	(0.210)	(34.31)	(0.174)	(0.194)	(449.9)	(0.471)	(0.378)
runvar	7.85e-06	5.19e-06	-1.03e-07	-6.48e-07	2.46e-05	8.24e-07	1.38e-06	-0.000457	-7.18e-07	-2.12e-06*
	(7.30e-06)	(6.71e-06)	(4.31e-07)	(6.42e-07)	(7.66e-05)	(5.19e-07)	(8.75e-07)	(0.00142)	(1.66e-06)	(1.22e-06)
runvar2	-6.55e-10*	-3.30e-10	0*	-0	3.10e-09	-0	-0	-3.19e-08	-0	0
	(3.43e-10)	(3.34e-10)	(0)	(0)	(2.49e-09)	(0)	(5.22e-11)	(5.11e-08)	(8.10e-11)	(0)
Observations	5,868	5,460	6,587	7,110	9,743	7,352	4,451	6,829	5,718	6,543
R-squared	0.007	0.004	0.069	0.036	0.010	-0.007	-0.003	-0.007	-0.051	0.003

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 18: 2SLS Optimal Bandwidth - 2012-2014. Estimates based on local regressions using 2sls method implemented with the ivreghdfe package, controlling for product and quarter fixed effects. Optimal bandwidths obtained following Calonico et al. (2020). as implemented in the rdrobust package, using data after demeaning product fixed effects.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	Number of Bidders	Number of MSEs	MSE Ratio	Ratio of MSEs Awarded	Mean Company Age	Open Auction Ratio	Discount	Distance to Supplier	ln(price)	ln(ratio)
set_aside	-2.070***	0.507	0.361***	0.519***	-54.03***	-0.0666*	-0.0680**	57.47	0.258***	0.0899
	(0.475)	(0.385)	(0.0268)	(0.0423)	(8.621)	(0.0349)	(0.0329)	(105.8)	(0.0744)	(0.0648)
runvar	1.85e-05***	9.48e-06*	-7.31e-07**	-9.30e-07*	-1.59e-05	3.39e-07	6.94e-07	0.000768	-8.57e-07	-1.89e-06**
	(6.23e-06)	(5.06e-06)	(3.52e-07)	(5.56e-07)	(0.000113)	(4.58e-07)	(4.33e-07)	(0.00139)	(9.77e-07)	(8.51e-07)
runvar2	1.02e-10	2.72e-10	0**	0	-2.94e-09	-0	0	-6.96e-08	0	-0
	(2.49e-10)	(2.02e-10)	(0)	(0)	(4.51e-09)	(0)	(0)	(5.54e-08)	(0)	(0)
Observations	8,491	8,491	8,491	8,491	8,491	8,491	8,491	8,491	8,491	8,491
R-squared	0.070	0.004	0.328	0.253	0.090	0.010	0.010	-0.002	0.019	0.011

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 19: 2SLS Asymmetric Bandwidth - 2016-2018. Estimates based on local regressions using 2sls method implemented with the ivreghdfe package, controlling for product and quarter fixed effects. We use an asymmetric bandwidth of (20000, 30000).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
VARIABLES	Number of Bidders	Number of MSEs	MSE Ratio	Ratio of MSEs Awarded	Mean Company Age	Open Auction Ratio	Discount	Distance to Supplier	ln(price)	ln(ratio)
set_aside	-0.911	0.187	0.0300	0.00374	-23.95	0.0226	0.0610	-157.0	-0.598	0.170
	(2.194)	(1.808)	(0.136)	(0.221)	(44.23)	(0.185)	(0.173)	(451.9)	(0.481)	(0.371)
runvar	1.04e-05	4.01e-06	-8.52e-07*	-1.50e-06*	3.22e-05	4.92e-07	8.95e-07	0.000555	-1.08e-06	-6.62e-07
	(7.82e-06)	(6.45e-06)	(4.84e-07)	(7.87e-07)	(0.000158)	(6.59e-07)	(6.17e-07)	(0.00161)	(1.72e-06)	(1.32e-06)
runvar2	-2.56e-10	-1.26e-10	0	-0	7.07e-10	-0	-0	-4.23e-09	-0	-0
	(2.75e-10)	(2.27e-10)	(0)	(0)	(5.55e-09)	(0)	(0)	(5.67e-08)	(6.04e-11)	(0)
Observations	6,252	6,252	6,252	6,252	6,252	6,252	6,252	6,252	6,252	6,252
R-squared	0.007	0.002	0.018	0.005	0.009	0.000	0.000	-0.000	-0.047	-0.004

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 20: 2SLS Asymmetric Bandwidth - 2012-2014. Estimates based on local regressions using 2sls method implemented with the ivreghdfe package, controlling for product and quarter fixed effects. We use an asymmetric bandwidth of (20000, 30000).

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
VARIABLES	Coefficient of Variation	Skewness	Kurtosis	Percentage Difference	Relative Distance	Normalized Distance	Bid Spread
set_aside	-0.141**	-0.327**	-0.232	-431.1**	61.27	-3.061	-1,180***
	(0.0629)	(0.144)	(0.526)	(198.7)	(69.56)	(4.098)	(449.5)
runvar	9.95e-07	-3.50e-07	6.38e-06	-0.0126	0.00102	-0.000936	-0.0332
	(1.05e-06)	(2.13e-06)	(8.13e-06)	(0.00903)	(0.00279)	(0.000659)	(0.0270)
runvar2	-0	-6.09e-11	-3.19e-10	-1.84e-06	-2.10e-07	-9.87e-09	-9.25e-06
	(6.73e-11)	(1.15e-10)	(4.70e-10)	(1.65e-06)	(4.20e-07)	(3.17e-07)	(7.17e-06)
Observations	6,474	7,473	7,086	2,286	2,598	708	1,689
R-squared	0.019	0.011	0.003	-0.030	-0.007	-0.012	-0.040

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 21: Collusive marker estimates of a two stage least squares approach (using the ivreghdfe package and controlling for product and quarter fixed effects) with the optimal bandwidth given by the rdrobust package after demeaning the data to control for product fixed effects - 2016-2018

VARIABLES	(1) Coefficient of Variation	(2) Skewness	(3) Kurtosis	(4) Percentage Difference	(5) Relative Distance	(6) Normalized Distance	(7) Bid Spread
set_aside	0.0595	0.282	1.493	250.4	-989.8	-8.168	-965.7
	(0.227)	(0.469)	(2.299)	(219.5)	(856.3)	(45.07)	(6,517)
runvar	1.96e-07	2.80e-07	2.81e-06	0.000442	-0.00389	0.00100	-0.00309
	(6.26e-07)	(9.41e-07)	(7.56e-06)	(0.000822)	(0.00461)	(0.00222)	(0.0244)
runvar2	0	-0	-6.89e-10**	-2.72e-08	-3.90e-07	5.11e-07	-3.23e-07
	(0)	(0)	(2.85e-10)	(4.22e-08)	(2.95e-07)	(1.08e-06)	(1.22e-06)
Observations	7,890	11,067	6,417	5,399	3,708	400	5,448
R-squared	-0.002	-0.008	-0.017	-0.026	-0.043	-0.005	-0.000

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 22: Collusive marker estimates of a two stage least squares approach (using the ivreghdfe package and controlling for product and quarter fixed effects) with the optimal bandwidth given by the rdrobust package after demeaning the data to control for product fixed effects - 2012-2014

6 Concluding remarks

Preference programs for Small Business in public procurement are a widespread instrument for fostering participation of this segment in the market for public contracts and for the awards of these contracts themselves. One important type of preference program is the use of set-aside auctions, where only the preferred bidders are allowed to submit bids. The regulation for set-asides may just target a share of the contracts' value to Small and Medium Enterprises (or, in the Brazilian case, Micro and Small Enterprises), and give some discretion to the procuring entity.

In Brazil the legislation allows for two situations where the procuring entity should use set-asides. One rule is quite simple, and therefore is the most utilized: bidding for contracts below a value threshold should be restricted for MSEs. Still, even such a simple mandatory rule has been customarily evicted by procurement officers, on three legal (alone or in combination) grounds: (i) the ambiguous wording in the legislation, which used to refer to the procurement value, but would leave doubt whether this value referred to the lot or to the whole batch − though this ambiguity was formally reversed by a reform enacted in 2014 and implemented in 2015 −; (ii) a written justification explaining that the set-aside would be disadvantageous for the public administration; (iii) a demonstration that the local or regional market does not include three or more suppliers classified as MSEs. In fact, a test made by auditors in the platform system (Brasil, 2023) showed that the system perfectly allows the user to register lots worth less than the BRL 80k ceiling and label them as open access (no preference).

The statutory threshold coupled with a feeble compliance by the agents suggests that a fuzzy Regression Discontinuity Design should be the best model to assess the impacts of this preference program. We only had to be careful with the change in the compliance rate from 2014 to 2016 following an amendment to the MSE General Law that turned the program from voluntary to compulsory (albeit still with the aforementioned opt-out mechanisms).

Before proceeding to the RDD, however, we made sure that (i) there was no bunching neither before nor after the 2014 amendment; (ii) the surge in the number of lots abiding to the set aside regime with values just below the threshold after the 2014 reform did not happen at the expense of the frequency of lots valued just above the threshold. In sum, that there is no strong evidence of lots been non-randomly selected to the left or right of the 80k BRL threshold. Manipulation tests focusing on density estimation assert that we cannot reject the null hypothesis of no manipulation. This is further demonstrated in the results of an auxiliary regression on subsamples of our data, showing that there is also no evidence of manipulation within major purchasing agency groups.

Visual inspection indicates that (i) the share of auctions reaching the open bid stage remained stable; (ii) there is no evidence on effects of set aside on discounts as their values remain stable within the threshold neighborhood; both for the 2012-2014 and 2016-2018 periods.

Following the validation of our sample, we have run a Fuzzy RDD model in two subsamples: one before the 2014 reform and another after the reform was formally implemented by a Decree in 2015 (thus starting in 2016). In fact, the set-aside mechanism did not produce palpable results before the reform in attracting MSEs to the set-aside lots, while the results are clearly positive after the reform, leading to a greater number and share of MSEs in those lots, a greater win ratio, and a much lower mean age of the bidders as a whole, as compared to the open-access lots above the BRL 80k value threshold.

Depending on the bandwidth utilized, however, price increases came up in the estimates. Moreover, when we control the regressions for both time (quarter) and product fixed effects, we find that the greater number of MSEs bidding in the lots under the value threshold was not enough to compensate the dropout rate of larger firms, as the total number of bidders in fact decreased significantly by approximately two bidders per lot. Moreover, not only the winning bid discount off the reserve price is found to have gone down in ten percent and the average (demeaned) price went up by 29 percent, but also two widely accepted collusive price markers – CV and Skewness – signalled in favor of a greater suspicion of collusion. Last but not least, participation in the open cry stage decreased, which reinforces the picture that the price dispute was cooled down by the set aside policy.

It is interesting to compare this price rise to the findings by Fazio (2022). There, in contrast to our setting, the value threshold separates direct contracting (i.e., a negotiated procedure without prior publication) from competitive procedures. In that case, a search for higher quality (as well as lower transaction costs) may be claimed to justify discretion in choosing the supplier. Here, in turn, the public buyer – had he some flexibility to value the purchase below or above the threshold – faces the choice between having full competition or a restricted competition. And competition is based on price only. There is no reason to believe that a restricted competitive procedure would give any advantage to the public buyer or its staff in terms of quality or convenience. The price increase that we observe here sounds much more as an undesired – albeit not unpredictable – side effect of giving precedence to a group of bidders at the expense of other larger – and presumably more cost-efficient – suppliers.

Of course, on the mind of the policymakers, this extra cost should be fully justified by the boost given to small business, which are meant to generate more employment than larger firms. Since Brazilian MSEs alone do not account for as great a share in employment or value added as SMEs do in OECD countries (Fiuza et al., 2023), this favourable cost-benefit analysis should be scrutinized.

References

Abrantes-Metz, Rosa M, Luke M Froeb, John Geweke, and Christopher T Taylor, "A variance screen for collusion," *International Journal of Industrial Organization*, 2006, *24* (3), 467–486.

Angrist, Joshua D. and Jorn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Athey, Susan, Dominic Coey, and Jonathan Levin, "Set-Asides and Subsidies in Auctions," *American Economic Journal: Microeconomics*, 2013, *5* (1), 1–27.

_ , Kyle Bagwell, and Chris Sanchirico, "Collusion and price rigidity," *The Review of Economic Studies*, 2004, *71* (2), 317–349.

Ayyagari, Meghana, Thorsten Beck, and Asli Demirguc-Kunt, "Small and medium enterprises across the globe," *Small business economics*, 2007, *29* (4), 415–434.

Bosio, Erica, Simeon Djankov, Edward Glaeser, and Andrei Shleifer, "Public procurement in law and practice," *American Economic Review*, 2022, *112* (4), 1091–1117.

Brannman, Lance and Luke M Froeb, "Mergers, cartels, set-asides, and bidding preferences in asymmetric oral auctions," *Review of Economics and Statistics*, 2000, *82* (2), 283–290.

Brasil, CGU, "Relatório de Avaliação sobre o Tratamento Diferenciado concedido a ME/EPP em Compras Públicas," Technical Report, Controladoria-Geral da União – Secretaria de Gestão 2023.

Brasil, TCU, "Programas emergenciais de concessão de crédito na pandemia da COVID-19: Análise da eficácia e sustentabilidade," Technical Report, Tribunal de Contas da União 2021.

_ , "Relatório de Políticas e Programas do Governo," Technical Report, Tribunal de Contas da União 2021.

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

_ , **Matias D. Cattaneo, and Rocio Titiunik**, "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," *Econometrica*, 2014, *82* (6), 2295–2326.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma, "Manipulation Testing Based on Density Discontinuity," *The Stata Journal*, 2018, *18* (1), 234–261.

Chetty, Raj, John N Friedman, Tore Olsen, and Luigi Pistaferri, "Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records," *The quarterly journal of economics*, 2011, *126* (2), 749–804.

Conley, Timothy G and Francesco Decarolis, "Detecting bidders groups in collusive auctions," *American Economic Journal: Microeconomics*, 2016, *8* (2), 1–38.

Coviello, Decio and Mario Mariniello, "Publicity requirements in public procurement: Evidence from a regression discontinuity design," *Journal of Public Economics*, 2014, *109*, 76–100.

_ , **Andrea Guglielmo, and Giancarlo Spagnolo**, "The effect of discretion on procurement performance," *Management Science*, 2018, *64* (2), 715–738.

Decarolis, Francesco, "Awarding price, contract performance, and bids screening: Evidence from procurement auctions," *American Economic Journal: Applied Economics*, 2014, *6* (1), 108–132.

_ , "Comparing public procurement auctions," *International Economic Review*, 2018, *59* (2), 391–419.

Dini, Marco and Andrea Heredia Zurita, "Analysis of policies to support SMEs in confronting the COVID-19 pandemic in Latin America," 2021.

Djankov, Simeon, Asif Islam, and Federica Saliola, "How large is public procurement in developing countries," *Peterson Institute for International Economics*, 2016.

EBRD, *Public procurement: Is open competition good for small and medium-sized enterprises?*, European Bank for Reconstruction and Development, 2017.

Fazio, Dimas, "The Two Sides of Discretion: Assessing Efficiency and Quality in Government Purchases," *Available at SSRN: https://ssrn.com/abstract=4212969*, 2022.

Fiuza, Eduardo P.S., Mauro O. Nogueira, Jonathan A.S. Oliveira, and Leandro P. Rocha, "Micro and small enterprises in Brazil: Regulatory framework and public policies aimed at them," 2023. Ongoing unpublished manuscript.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong, "Optimal nonparametric estimation of first-price auctions," *Econometrica*, 2000, *68* (3), 525–574.

Harrington-Jr, Joseph E. and Joe Chen, "Cartel pricing dynamics with cost variability and endogenous buyer detection," *International Journal of Industrial Organization*, 2006, *24* (6), 1185–1212.

Hoekman, Bernard and Bedri K. O. Taş, "Procurement policy and SME participation in public purchasing," *Small Business Economics*, 2020, pp. 1–20.

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

Huber, Martin and David Imhof, "Machine learning with screens for detecting bid-rigging cartels," *International Journal of Industrial Organization*, 2019, *65*, 277–301.

_ , _ , and Rieko Ishii, "Transnational machine learning with screens for flagging bid-rigging cartels," *Journal of the Royal Statistical Society Series A*, 2022, *185* (3), 1074–1114.

Imbens, Guido and Karthik Kalyanaraman, "Optimal bandwidth choice for the regression discontinuity estimator," *The Review of economic studies*, 2012, *79* (3), 933–959.

Imhof, David, "Empirical methods for detecting bid-rigging cartels." PhD dissertation, Bourgogne Franche-Comté 2018.

_ , Yavuz Karagök, and Samuel Rutz, "Screening for bid riggingâdoes it work?," *Journal of Competition Law & Economics*, 2018, *14* (2), 235–261.

Jiménez, Juan Luis and Jordi Perdiguero, "Does rigidity of prices hide collusion?," *Review of industrial organization*, 2012, *41*, 223–248.

Krasnokutskaya, Elena, "Identification and estimation of auction models with unobserved heterogeneity," *The Review of Economic Studies*, 2011, *78* (1), 293–327.

_ and Katja Seim, "Bid preference programs and participation in highway procurement auctions," *American Economic Review*, 2011, *101* (6), 2653–86.

Lee, David S, "Randomized experiments from non-random selection in US House elections," *Journal of Econometrics*, 2008, *142* (2), 675–697.

_ , Justin McCrary, Marcelo J. Moreira, and Jack Porter, "Valid t-ratio Inference for IV," *American Economic Review*, 2022, *112* (10), 3260–90.

Levin, Dan and James L Smith, "Equilibrium in auctions with entry," *The American Economic Review*, 1994, pp. 585–599.

Levy-Carciente, Sary, ed., *Índice de Burocracia en América Latina 2022*, Arlington, VA: Atlas Network, Center for Latin America, 2022.

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

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

Moreira, Humberto and Marcelo J. Moreira, "Optimal two-sided tests for instrumental variables regression with heteroskedastic and autocorrelated errors," *Journal of Econometrics*, 2019, *213* (2), 398–433.

Nakabayashi, Jun, "Small business set-asides in procurement auctions: An empirical analysis," *Journal of Public Economics*, 2013, *100*, 28–44.

OECD, *Financing SMEs and Entrepreneurs 2018: An OECD Scoreboard*, Organisation for Economic Co-operation and Development, 2018.

_ , *SMEs in Public Procurement: Practices and Strategies for Shared Benefits*, Organisation for Economic Co-operation and Development, 2018.

_ , "An in-depth analysis of one year of SME and entrepreneurship policy responses to COVID-19," *OECD SME and Entrepreneurship Papers*, 2021, (25).

_ , "One year of SME and entrepreneurship policy responses to COVID-19: Lessons learned to 'build back better'," *OECD Policy Responses to Coronavirus (COVID-19)*, 2021.

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

Pesendorfer, Martin, "A study of collusion in first-price auctions," *The Review of Economic Studies*, 2000, *67* (3), 381–411.

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

Rosa, Benjamin V, "Resident bid preference, affiliation, and procurement competition: Evidence from new mexico," *The Journal of Industrial Economics*, 2019, *67* (2), 161–208.

Samuelson, William F, "Competitive bidding with entry costs," *Economics Letters*, 1985, *17* (1-2), 53–57.

Staiger, Douglas and James H Stock, "INSTRUMENTAL VARIABLES REGRESSION WITH WEAK INSTRUMENTS," *Econometrica*, 1997, *65* (3), 557–586.

Szucs, Ferenc, "Discretion and Favoritism in Public Procurement," *Journal of the European Economic Association*, 03 2023. jvad017.

Tas, Bedri Kamil Onur, "Bunching below thresholds to manipulate public procurement," *Empirical Economics*, 2022, pp. 1–17.

Wallimann, Hannes, David Imhof, and Martin Huber, "A machine learning approach for flagging incomplete bid-rigging cartels," *Computational Economics*, 2022, pp. 1–52.

A First-stage Regressions

As mentioned in subsection 4.3, our identification strategy consists in using a dummy of lot values lower than 80k BRL as an instrumental variable to an endogenous indicator of set aside. Such strategy relies not only on the exogenous definition of the lot value in the neighborhood of the 80k BRL; but also in the power of the instrument to predict the endogenous variable. It is well documented that a weak instrument causes inefficiency and may also bias the estimate (Angrist and Pischke (2009)). Therefore we devote this appendix to present the results of the first stage regressions for the 2SLS, and some results on weak instrument tests. The F-test on the joint significance of estimated coefficients in the first stage regression is the most standard procedure to test the weak instrument hypothesis. According to Staiger and Stock (1997) a test statistic greater than 10 should suffice to reject the null hypothesis of weak instrument. Such procedure is under scrutiny as it relies on other properties of the non observable variables which should not be taken as given. Lee et al. (2022) provides more details on this topic and derives the value on 147 as an alternative value to be compared with the first stage F-test with no further worries on non-observable properties. The authors also reinforce the result established earlier by Moreira and Moreira (2019), which points out to the Anderson-Rubin test statistics as an optimal weak instrument test when the model has one endogenous variable and one instrument, just as in our specification.

Motivated by this recent development in the weak instrument literature, we will show both Anderson and Rubin (AR); and first stage F test statistics for our main regression specifications.

Table 23 shows results for first stage regression using the sample for the 2016-2018 period. One can see that all coefficients have the expected positive sign, meaning that batch values lower than BRL 80k increases the probability of set aside. As for the weak instrument tests, both of them (AR and F) suggest that we can reject such hypothesis in nine out of ten samples (considering here the more stringent 147 critical value for the F test).

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
Coef	0.3684	0.3637	0.3545	0.3441	0.3638	0.3684	0.4193	0.3476	0.3676	0.3728
F test	500.33	523.91	569.28	662.03	523.67	467.10	139.40	648.72	509.86	482.29
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR F test	15.52	3.16	137.12	145.27	34.79	4.60	3.19	0.09	14.85	1.56
P-val	0.0001	0.0754	0.0000	0.0000	0.0000	0.0320	0.0744	0.7598	0.0001	0.2110
AR Chi-sq test	18.09	3.67	158.16	165.05	40.37	5.41	4.21	0.11	17.28	1.83
P-val	0.0000	0.0554	0.0000	0.0000	0.0000	0.0200	0.0402	0.7443	0.0000	0.1758

Table 23: First stage estimates of a two stage least squares approach (using the ivreghdfe package and controlling for product and quarter fixed effects) with the optimal bandwidth given by the rdrobust package after demeaning the data to control for product fixed effects - 2016-2018

In principle the first stage results should not vary for regression models that differs only according to the dependent variable. But as we use the optimal bandwidths, the sample changes according to the dependent variables, and so does the estimated coefficients and weak instrument test statistics.

Table 24 reports analogous results for the 2012-2014 period; when the instrument is clearly much weaker. Point estimates of the instrument coefficient are much lower, and both AR and F tests do not reject the weak instrument hypothesis for all considered samples.

These results combined attest that the change in legislation in 2014 was effective in enforcing the set aside procedure for purchases below 80k BLR.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
Coef	0.0937	0.0943	0.0848	0.0853	0.0927	0.0862	0.0947	0.0842	0.0946	0.0844
F test	67.28	63.92	76.03	82.19	95.67	84.40	53.23	78.83	65.31	76.42
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR F test	0.44	0.05	1.01	0.46	0.42	0.39	0.30	0.63	1.65	0.00
P-val	0.5062	0.8268	0.3141	0.4962	0.5176	0.5313	0.5811	0.4292	0.1991	0.9813
AR Chi-sq test	0.53	0.06	1.21	0.55	0.49	0.46	0.37	0.74	1.99	0.00
P-val	0.4659	0.8098	0.2714	0.4585	0.4861	0.4958	0.5405	0.3883	0.1587	0.9795

Table 24: First stage estimates of a two stage least squares approach (using the ivreghdfe package and controlling for product and quarter fixed effects) with the optimal bandwidth given by the rdrobust package after demeaning the data to control for product fixed effects - 2012-2014

We also show first stage results together with weak instrument test statistics for the model specification with predetermined asymmetrical bandwidth. Results confirm the pattern highlighted above, namely of an instrument with great power for the 2016-2018 period but that can be considered a weak one in the 2012-2014 period.

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
Coef	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680	0.3680
F test	628.58	628.58	628.58	628.58	628.58	628.58	628.58	628.58	628.58	628.58
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR F test	18.07	1.72	130.11	119.26	36.26	3.61	4.24	0.30	11.85	1.92
P-val	0.0000	0.1891	0.0000	0.0000	0.0000	0.0574	0.0395	0.5864	0.0006	0.1664
AR Chi-sq test	21.10	2.01	151.92	139.25	42.34	4.22	4.95	0.35	13.83	2.24
P-val	0.0000	0.1559	0.0000	0.0000	0.0000	0.0400	0.0260	0.5566	0.0002	0.1348

Table 25: First stage estimates of a two stage least squares approach (using the ivreghdfe package and controlling for product and quarter fixed effects) with an asymmetric bandwidth of (20000, 30000) - 2016-2018

VARIABLES	(1) Number of Bidders	(2) Number of MSEs	(3) MSE Ratio	(4) Ratio of MSEs Awarded	(5) Mean Company Age	(6) Open Auction Ratio	(7) Discount	(8) Distance to Supplier	(9) ln(price)	(10) ln(ratio)
Coef	0.0878	0.0878	0.0878	0.0878	0.0878	0.0878	0.0878	0.0878	0.0878	0.0878
F test	92.45	92.45	92.45	92.45	92.45	92.45	92.45	92.45	92.45	92.45
P-val	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
AR F test	0.17	0.01	0.05	0.00	0.29	0.01	0.12	0.12	1.62	0.21
P-val	0.6787	0.9179	0.8266	0.9865	0.5899	0.9026	0.7245	0.7282	0.2038	0.6449
AR Chi-sq test	0.21	0.01	0.06	0.00	0.35	0.02	0.15	0.14	1.94	0.25
P-val	0.6502	0.9101	0.8104	0.9852	0.5552	0.8935	0.6996	0.7036	0.1642	0.6139

Table 26: First stage estimates of a two stage least squares approach (using the ivreghdfe package and controlling for product and quarter fixed effects) with an asymmetric bandwidth of (20000, 30000) - 2012-2014