# Competition under Incomplete Contracts and the Design of Procurement Policies*

Rodrigo Carril† Andres Gonzalez-Lira‡ Michael S. Walker§

January 19, 2022

[Link to the latest version]

**Abstract**

We study the effects of intensifying competition for procurement contracts. Conceptually, opening contracts up to bids by more participants leads to lower acquisition costs. However, expanding the set of bidders hinders buyers' control over the quality of prospective contractors, potentially exacerbating adverse selection on non-contractible quality dimensions. We study this trade-off in the context of procurement by the U.S. Department of Defense. Our empirical strategy leverages regulation that mandates agencies to publicize contract opportunities whose value is expected to exceed a certain threshold. We find that advertising contract solicitations increases competition and leads to a different pool of selected vendors who, on average, offer lower prices. However, it also worsens post-award performance, resulting in more cost overruns and delays. This negative effect on post-award performance is driven by goods and services that are relatively complex, highlighting the role of contract incompleteness. To further study the scope of this tension, we develop and estimate a model in which the buyer chooses the extent of competition, and the invited sellers decide on auction participation and bidding. We estimate sellers' cost and ex-post quality distributions, as well as buyers' preference parameters over contract outcomes. Simulating equilibrium conditions under counterfactual settings, we benchmark the current regulation design with complexity-tailored publicity requirements, and find that adjustments to publicity requirements could provide savings of 2 percent of spending, or $104 million annually.

JEL Codes: D22, D44, D73, H57, L13, L14

*We want to thank Steve Tadelis, Benjamin Handel, Kei Kawai, and Reed Walker for their dedicated advice since the early stage of the project. We thank our discussants, Paulo Somaini, Decio Coviello, and Alex Arsenault-Morin, for many helpful comments. We also thank Eli Berman, Christopher Campos, Ernesto Dal Bo, Mark Duggan, Liran Einav, Ying Fan, Matt Gentzkow, Phil Haile, Jon Kolstad, Cristobal Otero, Dario Tortarolo, Damian Vergara, Christopher Walters, Heidi Williams, Guo Xu, Roman David Zarate, seminar participants at the Barcelona GSE Summer Forum, DOD (Pentagon) Acquisition Analytics Forum, IIOC, ITAM, NBER Summer Institute, PUC-Chile, PUC-Rio, RIDGE-IO, Stanford, UBC-Sauder, UC Berkeley, Univ. of Michigan and Yale University for helpful comments and suggestions. Finally, we gratefully acknowledge the discussion provided by Ph.D. students of the UCB Research Methods course and the support from the Institute for Business Innovation at the Haas School of Business.

†Universitat Pompeu Fabra and Barcelona School of Economics. Email: rodrigo.carril@upf.edu.

‡Yale University. Email: andres.gonzalez-lira@yale.edu.

§US Department of Defense. Email: michael.s.walker110.mil@mail.mil.

# 1 Introduction

Buyer-seller transactions—concerning everything from standardized goods such as office supplies or fuels, to customized needs such as construction projects or consultancy services—are often governed by competitively-awarded procurement contracts. The pervasive use of competition to assign contracts stems from the notion that competitive bidding can be a powerful tool to reduce procurement prices (Bulow and Klemperer, 1996). Yet, expanding competition for contracts that involve customized obligations and deliverables could allow under-qualified contractors to win, leading to deficient execution ex-post. Therefore, the assessment of competitive procurement mechanisms should account both for potential benefits due to price reductions, as well as for potential adverse effects due to poor execution.

An empirical investigation of this trade-off is complicated, in part due to the need for comprehensive data on contract execution and a compelling research design. In this paper, we aim to make progress on both of these fronts to study the equilibrium effects of enhancing competition for procurement contracts in acquisition price and execution quality. We focus on U.S. Department of Defense (DOD) procurement, a setting of relevance given that it awards $500 billion in procurement contracts per year, representing a sizable fraction of the U.S. economy. Moreover, this setting provides us with policy variation in the degree of contract competition, as well as with detailed administrative data throughout the life-cycle of each DOD contract, from design through execution.

Our empirical strategy exploits regulation that requires agencies to publicize contract opportunities that are expected to exceed $25,000 in value through a centralized online platform.[1] We exploit the discontinuous nature of these requirements to estimate the effect of enhanced contract publicity on four sets of outcomes: (i) the level of competition for the award, (ii) characteristics of the buyer-contractor relationship, (iii) procurement costs, and (iv) post-award contractor performance. By providing evidence on all of these fronts, we comprehensively characterize the consequences of changing the degree of competition for procurement contracts through this advertising channel. Furthermore, we exploit rich heterogeneity in the types of contracts that the DOD awards to assess the role of contract incompleteness in explaining our results.

We start by analyzing the price effects of contract publicity. We do this by investigating the observed contract price densities of publicized and non-publicized contracts.[2] We then estimate the effects of publicizing contract opportunities on three sets of non-price outcomes: the level of competition, the characteristics of the selected vendors, and post-award performance. We do this by implementing a Regression Discontinuity Design (RDD) for contracts with an expected award

[1] From a policy perspective, the volume of contracts impacted by this regulation make its implications economically meaningful. In 2018 alone, the DOD publicized contract solicitations valued at $ 5.56 billion dollars via the online platform `FedBizzOpps.gov`.

[2] Our method is robust to the existence of strategic bunching below the threshold, aimed at avoiding publicizing certain contracts. In fact, we separately quantify the extent of strategic bunching and the price effects of publicizing contracts.

amount close to $25,000. The discontinuity in publicity requirements at this threshold generates a convenient natural experiment for studying the impact of the policy on these non-price outcomes.

We find that contract awards advertised through the government platform see an increase in the number of bids of roughly 60%, confirming that the policy translates into a substantial increase in participation. We show that these marginal participants are competitive, leading to changes in the characteristics of winning firms: awardees of publicized solicitations are, on average, 14% less likely to be small businesses, and are located 60% farther from the buying agency. Furthermore, we find that increased competition leads to contract price reductions: publicized contracts are, on average, awarded at 6% lower prices. However, advertised contracts result in worse ex-post performance: the probability of experiencing cost-overruns and delays in the implementation stage increase by 7% and 8%, respectively. The latter results are driven by service contracts—as opposed to goods purchases—and by contracts that we ex-ante characterize as more complex. These results are robust to different estimation approaches and there is little evidence of buyers bunching at these advertising thresholds which could lead to bias in the RDD results.

Taken together, our results suggest that promoting competition has mixed effects on contract outcomes: while it reduces the winning bid, it leads to worse outcomes at the execution stage. We benchmark price reductions ex-ante with increases in cost-overruns ex-post, and we find that intensifying competition only reduces overall contract costs for "simple" product categories, while increasing overall costs for relatively complex contracts. Overall, we find that suppliers' identity matters for explaining the variation in contract outcomes. Promoting competition hinders buyers' ability to restrict participation to qualified vendors, while attracting new participants who tend to perform poorly ex-post.[3]

Motivated by this evidence, we develop and estimate an equilibrium model of competition for procurement contracts with two general objectives. First, we estimate the underlying firm characteristics, which shape adverse selection in this market. Second, the estimated parameters allow us to study the role of buyer preferences in the promotion of competition, as well as the consequences of counterfactual policies aimed at reducing public spending.

Our model consists of four stages that cover the different phases of a procurement project. First, a buyer decides on the degree of competition by choosing whether to openly publicize the contract, or to invite only specific contractors. Second, firms that receive information about the contract simultaneously decide whether to prepare a bid. They do this by comparing the expected utility of participating with the idiosyncratic cost of preparing the bid. Firms participate if the cost of preparing the bid is sufficiently low. Third, each bidder submits a bid that depends on the realization of a production cost estimate, consisting of a private component and a common component, which accounts for unobserved heterogeneity (Krasnokutskaya, 2011; Haile and Kitamura, 2019).[4] The award mechanism is a first-price, sealed-bid auction. Fourth, the awarded

[3]The main alternative hypothesis is that vendors behave differently ex-post depending on the level of competition ex-ante. We explicitly test and reject this *moral hazard* explanation in favor of our *adverse selection* hypothesis.

[4]Controlling for unobserved heterogeneity is important since, in the procurement setting, bidders likely have more information about the auctioned contracts than the econometrician.

contractor executes the contract. The quality of execution depends on the existence and magnitude of cost-overruns, which stem from an idiosyncratic shock realized ex-post. The model incorporates the potential asymmetry between bidders who are informed directly by the buyer and those who participate only when the contract is openly publicized. Moreover, the model does not impose restrictions on the buyer's preferences over outcomes, and allows for idiosyncratic preferences for certain vendors that are uncorrelated with contract outcomes.

We estimate our model using data on publicity choices, auction participation, contract prices, and observed cost-overruns. Our estimates highlight an asymmetry between the sellers whom the buyer would invite directly in the absence of publicity and those who bid only when the solicitation is openly publicized. The added bidders have slightly lower production costs and substantially lower participation costs, which makes them more likely to participate *ceteris paribus*. They are also considerably more prone to experience cost-overruns in the execution stage. On the other hand, buyers show a preference for lower prices, lower cost-overruns, and incumbent suppliers.

We then use our model to estimate the effects of promoting competition through publicity under the current regulation, as well as under alternative policy scenarios. Overall, our findings are consistent with the estimated reduced-form effects. Increasing competition has heterogeneous effects, leading to cost reductions when the transaction unit is relatively simple. However, competition backfires when the contract involves a complex product category, as increases in cost-overruns exceed price reductions.

Our results show that imposing regulation to promote bidder participation involves a risk of allowing under-qualified firms to bid. An alternative policy design is to rely on buyers (i.e. each local agency) to decide on whether to publicize each contract. As emphasized by the vast literature on the allocation of authority within organizations (Aghion and Tirole, 1997), delegating this decision to the buyer involves a trade-off. On the one hand, more discretion allows the buyer to tailor decisions, mitigating the potential risks of intensifying competition. On the other hand, the buyer could use this added discretion opportunistically, restricting competition to favor specific contractors. We use our model to simulate equilibrium outcomes under a deregulated setting in which the buyer decides whether to publicize each contract. We find that delegating this decision to the buyer is welfare-enhancing when the transaction unit is complex: on average, the buyer achieves better outcomes than in regulated settings with either zero or full publicity. However, when the transaction unit is relatively simple, imposing full-publicity rules is convenient as the risks at the execution stage are minor.

We next use our model to identify improvements to the current policy design, which depart from uniform publicity requirements. Policies that regulate competition in most public procurement settings—including the one we study—are strikingly simple: they do not differ depending on whether the transaction involves a commodity or a highly customized service. This mismatch between unsophisticated policies and a highly diverse set of transactions suggests meaningful room for improvement in policy design. We study the effects of introducing publicity requirements that are tailored to the complexity of the purchase, thus leveraging the benefits of intense competition

for simple products, while limiting its adverse effects on complex products. We find that the cost-minimizing level of publicity for commodity-type products is 100%, whereas more complex product categories should require low use of publicity. We find that this reduces average defense procurement costs by 2 percent, or $104 million annually.

This paper contributes to several branches of the existing literature. First, it contributes to the literature examining transactions under incomplete contracting (Williamson, 1976; Goldberg, 1977; Hart and Moore, 1988).[5] This literature is mostly theoretical, with only a small number of empirical papers studying the interplay between competitive mechanisms and contract outcomes (Spulber, 1990; Bajari, McMillan, and Tadelis, 2009; Decarolis, 2014). Our paper departs from existing work along two relevant margins: first, existing papers focus on different award mechanisms (e.g., auctions versus negotiations), while our empirical framework keeps the award mechanism fixed but exploits variation in the set of relevant potential sellers; second, unlike existing literature that concentrates on studies of specific product categories, our sample includes a wide range of product categories purchased by the DOD. This allows us to provide a comprehensive picture of the implications of promoting competition, arriving at different conclusions for different types of purchases.

Our paper also contributes to the growing literature that evaluates policies aimed at promoting (or restricting) bidders' participation in procurement settings. This literature emphasizes that expanding the pool of potential bidders may not necessarily translate into lower contract prices if bidders' participation is endogenous, as their equilibrium bidding behavior can become less aggressive (Athey et al., 2011, 2013; Li and Zheng, 2009, 2012; Krasnokutskaya and Seim, 2011; Marmer et al., 2013; Bhattacharya et al., 2014; Sweeting and Bhattacharya, 2015).[6] We leverage variation in the number of potential bidders that stems from exogenous changes in publicity requirements. We model entry and bidding decisions and find that incumbents are less likely to participate when they anticipate fiercer competition. However, in our setting, the effect of competition from new entrants dominates that of less aggressive bidding by incumbents, reducing the winning bid as a result. The source of variation in the number of potential bidders is closely related to Coviello and Mariniello (2014), who study a similar policy in Italy.

Third, this paper contributes to the growing literature that examines a buyer's role as an agent affecting market outcomes. In particular, this literature considers the fact that buyers' actions can be motivated by objectives other than simple contract cost reductions (Bandiera et al., 2009; Liebman and Mahoney, 2017; Coviello and Gagliarducci, 2017; Best et al., 2017; Decarolis et al., 2020; Carril, 2019; Szucs, 2020). In particular, this paper relates to Kang and Miller (2017), who study buyers' competition promotion for IT contracts in the United States. We depart from existing papers by comparing buyer preferences for ex-ante and ex-post outcomes, with idiosyncratic preferences for specific vendors.

[5] More recent developments include Laffont and Tirole (1990); Tirole (1999); Chakraborty et al. (2020)

[6] These ideas were initially introduced by Samuelson (1985); Levin and Smith (1994). Li and Zheng (2009) provide an empirical framework highlighting that increasing the number of potential bidders within the independent private value (IPV) setting has ambiguous effects, as the equilibrium behavior interacts two opposite effects: "competition effect" with "entry effect." The former tends to reduce prices, while the latter tends to increase them.

The rest of the paper is organized as follows. Section 2 provides background on the U.S. DOD procurement system and the data we use for our analysis. In Section 3, we provide evidence on the effects of contract publicity on a range of relevant outcomes. In Section 4, we develop and estimate an equilibrium model of procurement competition, which we then use to study outcomes under counterfactual environments in Section 5. Section 6 concludes.

# 2 Setting and Data

## 2.1 US Federal Procurement and Publicizing Requirements

Public procurement is a large component of the US economy. In fiscal year 2019, federal contract awards totaled $926 billion. Contracts are awarded at highly decentralized levels, with more than 3,000 different contracting offices that are part of an executive or independent agency.[7] The workforce in charge of public contracting is made up of over 35,000 contracting officers whose primary role is to plan, carry out, and follow-up on purchases made by their units. Contracting officers' scope of action is defined and limited by the Federal Acquisition Regulation (FAR). The FAR lays out policy goals and guiding principles, as well as a uniform set of detailed policies and procedures to guide the procurement process. Our analysis leverages a specific section of the FAR—Part 5 (*Publicizing Contract Actions*)—as a convenient source of quasi-experimental variation to study the effect of information diffusion.

FAR Part 5 requires publicizing contract opportunities in order to "increase competition", "broaden industry participation", and "assist small businesses (and other minority businesses) in obtaining contracts". Since October 1, 2001, contract actions that exceed $25,000 must be publicized in an online government-wide platform which we will refer to as FedBizOpps (or FBO).[8] This implies uploading a request for quotes with a full description of the good or service being requested, and the instructions to submit the bids. We will refer to this synopsis document as a contract *solicitation*. Most of the contracts in this dollar range are awarded to the lowest price quote that is technically acceptable according to the specifications.

Officers with contracts that are not expected to exceed this threshold are not required to publicize in FedBizOpps; however, they are still free to use it if they want to increase contract visibility.[9] The regulation allows for exemptions to the requirement above the threshold, if doing so "compromises national security", if "the nature of the file does not make it cost-effective or practicable", or if "it is not in the government's interest". Therefore, while this policy discretely

[7] Executive agencies are headed by a Cabinet secretary, like the Department of Defense, the Department of State, or the Department of Health and Human Services. Independent agencies, which are not part of the Cabinet, include the Central Intelligence Agency, the Environmental Protection Agency, and the Federal Trade Commission.

[8] Throughout our period of analysis, this online platform—designated as the "government point of entry" by the FAR—was called Federal Business Opportunitites (FBO) available at: `fedbizopps.gov`. In late 2019 (after our sample period ends), the government point of entry migrated to `beta.sam.gov`, featuring minor changes to the user interface.

[9] Procurement officers with contracts with expected values below the threshold are only required to advertise the solicitation "by displaying [it] in a public place." This includes, for example, a physical bulletin board located at the contracting office.

affects the likelihood of publicized contracts around the threshold, we anticipate that compliance may be far from perfect, given the voluntary nature of the rule below this value and the availability of exceptions above. Appendix C.1 describes the details of the policy and the website.

### 2.2 Data

We use two complementary sources of data. The first consists of the historical files from FedBizOpps, which provides detailed information on pre-award notices (i.e. solicitations) posted on the platform. The second is the Federal Procurement Data System - Next Generation (FPDS-NG), which tracks federal contracts from the time of their award and includes all follow-on actions, such as modifications, terminations, renewals, or exercises of options.

We merge awards from FPDS-NG to notices on FedBizOpps using the solicitation number. Note, however, that while FPDS-NG contains the universe of federal awards, FedBizOpps only has the notices posted on the website. From this matching process, we construct a dummy variable that is equal to 1 if we are able to merge a contract with any pre-award notice on FedBizOpps, in which case we say the contract was *publicized*. Appendix Figure A1 describes the typical timeline of events surrounding the life-cycle of a contract, and the appropriate data source that records that information.

In addition, we observe detailed information for each contract award, including the dollar value of the funds obligated, a four-digit code describing the product or service, codes for the agency, sub-agency, and contracting office making the purchase, the identity of the private vendor, the type of contract pricing, the extent of competition in the award, characteristics of the solicitation procedure, the number of offers received, and the applicability of a variety of laws and statutes. For additional details on the construction of the dataset, see Appendix C.2.

The analysis sample consists of all competitively awarded and definitive contracts[10] with award values between $ 10,000 and $ 40,000, awarded in fiscal years 2015 through 2019 by the Department of Defense (DOD), for products and services other than Research and Development (R&D).[11] Table B.1 presents summary statistics of the sample. In total, there are roughly 86,000 contracts awarded by 597 contracting offices to almost 30,000 firms. Contract durations are expected to be 54 days on average and are awarded on a fixed-price basis. A noteworthy feature of this setting is that competition is limited; an average contract receives 3.5 offers, with one out of four contracts receiving a single offer.[12] The Department of the Navy and the Army each account for more than

[10]Federal contracts can be broadly categorized into two types: definitive contracts (DCs) and indefinite delivery vehicles (IDVs). DCs are stand-alone, one-time agreements with a single vendor for the purchase of goods or services under specified terms and conditions. See Carril (2019) for more details. We simplify the analysis by focusing exclusively on DCs, which are well-defined requirements involving a bilateral relationship between a single government unit and a private firm.

[11]The Department of Defense represents 58% of overall federal spending and more than 60% in the restricted sample. We exclude R&D awards because are subject to a unique set of acquisition rules, see FAR Part 35.

[12]More than half of the awards are set aside for a particular type of firm (typically, small business). Set-asides are a major factor of acquisition strategy in the DOD; contracting offices are required to meet specific set-aside goals. Even though they affect contract competition, we abstract away from that feature as we do not condition nor restrict our sample

40% of the contracts, with the rest being mostly awarded by the Air Force. Winning vendors are often geographically close to the contracting offices, with both located in the same state in 2 out of every 3 contracts. Finally, 75% of suppliers are characterized as small businesses.

We also observe rich information about the type of good and service that is contracted upon. Each award is classified into one of 1,479 possible standardized 4-digit alphanumeric codes. These can be aggregated into 101 broader 2-digit product categories, 77 goods and 24 services. Table B.2 shows the top 10 most common 2-digit good and service categories. The most common product categories are ADP Equipment Software, Medical Equipment and Supplies and Maintenance and Repair Equipment.

# 3 The Effect of Competition on Contract Outcomes

In this section, we study the effects of publicizing procurement solicitations on contract outcomes. As described in Section 2.1, federal regulation introduces a publicity requirement at $25,000. Using different approaches, we exploit this discontinuity to provide evidence of the effects of publicity on contract award price and other contract outcomes. These results will serve as the basis for the development of our model in Section 4.

## 3.1 Preliminaries

For each contract in our data, we observe agencies' decisions to publicly solicit a contract in *FedBizzOpps.gov* prior to its award (decision that we denote as $D \in \{0, 1\}$). We leverage the variation introduced by the regulation, which discontinuously affects the likelihood of public solicitation at an arbitrary threshold ($\bar{p} = 25,000$) depending of the contract's *expected* award price ($\tilde{p}$). We do not observe ex-ante estimated prices $\tilde{p}$, but only ex-post realized prices $p$, which entails two empirical challenges. First, contracting officers know the policy threshold, so it may generate incentives to modify the purchase in a way that makes the ex-ante estimate fall below $\bar{p}$. This behavior would result in bunching on the distribution of ex-ante prices, generating an excess amount of contracts estimated to be at or slightly below $\tilde{p} = \bar{p}$. Second, since prices ex-post may differ from prices ex-ante, estimating effects at the discontinuity may be subject to measurement error biases. In our case, publicity may affect prices due to enhanced competition; thus, the error distribution may differ depending on the publicity status of the contract.

These empirical challenges are discussed in detail by Carril and Gonzalez-Lira (2021), who propose a method that uses the distribution of observed awards $p$ and publicizing decisions $D$ to (nonparametrically) recover information about the distribution of $\tilde{p}$, the distribution of the effects of publicity on price, and the extent of "manipulation".[13] Intuitively, the method hinges

based on that margin.

[13] By manipulation we mean any decision ex-ante that modifies the requirement with the sole purpose of arriving at a different price estimate. The term follows the literature on Regression Discontinuity, which refers to this as manipulation of the running variable. However, it is noteworthy that this behavior need not involve any wrongdoing.

on comparing the observed empirical distributions of award prices and estimated counterfactual distributions stripped of the confounding influence of bunching and competitive price effects. Carril and Gonzalez-Lira (2021) combine existing density analyses approaches aimed at recovering behavioral parameters (Kleven, 2016), estimating policy effects (Chernozhukov et al., 2013; Jales and Yu, 2017) and correcting for measurement error (Schennach, 2020). We use this framework to estimate price effects and correct (and bound) RDD estimates on non-price outcomes accounting for the aforementioned confounds. In Supplementary Material G, we provide discussion about about estimation method.

Section 3.2 discusses the price effects of publicity that are obtained from the density analysis approach. Section 3.3 describes the estimation of publicity effects on non-price outcomes relying on corrected RDD methods. Sections 3.4, 3.5 and 3.6 provide interpretation of the estimated effects.

### 3.2 Inference on Contract Price Distributions

Let $p_t(\tilde{p}_t|D_t)$ be the *potential* log-price that we would have observed for contract $t$, as a function of ex-ante estimates $\tilde{p}_t$ and a publicity decision of $D_t \in \{0, 1\}$. We assume that publicizing solicitations leads to a (log-)linear random price effect for contract $t$: i.e., $p_t(\tilde{p}_t|D_t = 1) = p_t(\tilde{p}_t|D_t = 0) - \gamma_t$, with $\gamma_t \sim F_\gamma(\cdot)$.

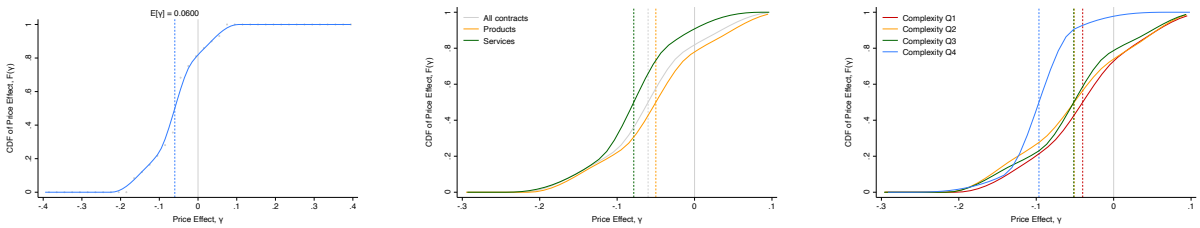
We estimate $E[\gamma_t]$ from the observed distributions of publicized and non-publicized contracts. The intuition of our method is based on three observations. First, relative to a counterfactual with no price effects of publicity, the observed distribution of publicized contracts should be "shifted" by $E[\gamma_t]$. Second, the distribution of non-publicized contracts is not affected by $\gamma_t$. Third, we expect the counterfactual price distribution of the *total* number of contracts (both with and without publicity) to be smooth around the discontinuity, even though each conditional distribution is not. These three observations motivate our method. We pick a value for $\widehat{E[\gamma_t]}$ and "undo" the price effects of publicity by shifting the distribution of publicized contracts, which we then add to the non-publicized contracts. The "right" value of $\widehat{E[\gamma_t]}$ will satisfy smoothness of the overall distribution and an integration constraint.

Carril and Gonzalez-Lira (2021) further show that if $F_\gamma$ is symmetric, then the logic can be extended to nonparametrically identify the full CDF given the observed distributions of realized prices conditional on publicity status, $f(p_t|D_t = 0)$ and $f(p_t|D_t = 1)$. Moreover, the analysis is robust to having strategic bunching in the distribution of non-publicized awards, and the extent of this behavior is also identified using similar arguments. The key is that strategic bunching affects only the distribution of non-publicized awards, so that price effects and bunching are separately identified from the two observed distributions.

Figure 1(a) depicts the (nonparametric) estimate of the CDF of $\gamma_t$, along with a local polynomial smoothing. We find that publicity leads to an average reduction in award price of 0.06 log-points (SE: 0.02), equivalent to \$ 1,456 at the discontinuity. From the full distribution, we see that publicizing contract opportunities reduces award prices for 83% of the contracts. Table B.3 in

Figure 1: Distribution of Price Effects

(a) Full Distribution (b) Good vs. Service (c) Quartile of Complexity




Notes: Notes: This figure presents the estimated CDF of the price effect parameters $\gamma$. The panel (a) shows the cumulative distribution function (CDF) of all contracts in the sample. Every gray dots show actual point estimates given a discretization of the support of $\gamma$. The blue line corresponds to a kernel fit. This estimation procedure builds upon comparing the empirical densities to a counterfactual distribution of publicized contracts assuming no price effect. The counterfactual distribution is generated from the interpolation of a polynomial of degree 5. The dashed vertical line corresponds to the estimated mean effect. The panel (b) shows the CDF of price effects separating contracts of goods and services. The panel (c) describes the CDF of price effects by quartile of complexity.

the Appendix provides more details about the mean and variance of price effects and displays sub-group analyses. We find that price effects are higher for services, and the effects are larger for contracts that are more complex. [14]

Figure A4 shows the density distributions of both publicized and non-publicized contracts, stripped down from price effects and strategic bunching responses. From the distribution of non-publicized awards (Panel (a)), we can directly compute the excess bunching below the threshold, explained by agencies' desire to avoid the publicity mandate. We estimate that the excess mass right below the discontinuity equals 12% of the value of the density at the threshold. This magnitude will be used to account for the effects of this manipulation on our RDD estimates in Section 3.3. However, we can already infer that, since the extent of bunching is arguably modest, its impact on our estimates will be limited as well.

Finally, in Panel (b), we compare the empirical distribution to the sharply discontinuous distribution of publicized awards that would be observed if $\gamma_t = 0$ for all $t$. It is evident how the distribution of $\gamma_t$ smooths out the discontinuity in the density of publicized contracts. As noted by existing literature, observing the assignment variable with error biases the estimated effects towards zero in the RDD setting (Lee and Lemieux, 2010; Davezies and Le Barbanchon, 2017; Pei and Shen, 2017). We leverage the estimated distribution of $\gamma_t$ to correct for this factor in Section 3.3.

### 3.3 Regression Discontinuity Design: Estimating Effects on Non-price Outcomes

In this section, we leverage the discontinuous nature of the publicity requirements to gauge the effects of publicity on a set of other relevant outcomes, including the level of competition, characteristics of the winning bidder, and post-award contractor performance. We use the estimates of price effects and bunching to adjust the RDD estimates accounting for these factors.

[14] By complexity, we refer to the average cost-overrun for all contracts in the product category valued under $20,000. This is discussed later in the paper.

### 3.3.1 Empirical Framework

Consider specifications of the following form:

$$Y_t = \alpha + \beta \cdot D_t + g(\tilde{p}_i) + X_t'\delta + \epsilon_t \quad , \tag{1}$$

the coefficient of interest is $\beta$, the effect of publicizing a solicitation on contract outcome $Y_t$. In the standard Regression Discontinuity Design (RDD), we obtain an estimate of $\widehat{\beta}_{IV}$ by instrumenting $D_t$ with the discontinuity in publicity requirements. The first-stage of this IV procedure is of the form:

$$D_t = \lambda + \gamma \cdot \mathbf{1}[\tilde{p}_t > \bar{p}] + g(\tilde{p}_t) + X_t'\eta + \nu_t \quad , \tag{2}$$

for some smooth function $g(\cdot)$. A key advantage of this approach is that it is possible to provide compelling evidence on the existence of an effect by graphically showing the reduced form of this model, i.e.:

$$Y_t = \mu + \phi \cdot \mathbf{1}[\tilde{p}_t > \bar{p}] + g(\tilde{p}_t) + X_t'\kappa + \xi_t \quad . \tag{3}$$

Again, our key challenge is that we observe $p_t$, but not $\tilde{p}_t$ directly, and that the mapping between these variables is affected by price effects and possible bunching responses. However, with the procedure discussed above we can recover significant information about the latent distribution of $\tilde{p}_t$, which allows us to successfully exploit the discontinuity for estimating causal effects.

Consider first a naive RDD, described by versions of Equation (1), Equation (2), and Equation (3), where we simply replace ex-ante prices $\tilde{p}_t$ by realized observed prices $p_t$. The estimates obtained from this naive RDD will be identical to the true RDD if there are neither price effects ($\gamma_t = 0$ for all $t$) nor bunching responses. The larger these effects are, the more the estimates from the naive RDD will differ from the true parameters. Given this, we take the naive RDD as our baseline and sequentially implement corrections to account for price effects and bunching responses, to transparently show how these elements affect the estimation.

In Supplementary Material G.1.6, we describe in detail the first of such corrections, namely a method to recover the causal parameters of interest in the presence of price effects $\gamma_t$. The key result is that, under our modelling assumptions, we can write the conditional expectation of contract outcomes given *observed* prices $E\left[Y_t|p_t\right]$ as an explicit linear function of the causal parameters that we seek to recover, plus objects that we can directly observe or estimate. This function depends on observed prices $p_t$, observed treatment probabilities $\pi_D$, and moments of the distributions of price effects $F_\gamma$ (which we obtained from the density analysis). We then use this result to estimate the causal parameters using OLS.[15]

[15] We also show in Supplementary Material G.1.6 that this logic can be easily extended to accommodate measurement error in ex-ante prices, so that $\tilde{p}$ is only an unbiased but not necessarily perfect forecast of $p^0(\tilde{p})$. That is, we show that it

On the other hand, we can account for the effect of bunching responses by following the results from Gerard, Rokkanen, and Rothe (2020). These authors derive sharp bounds on treatment effects for the RDD in the presence of bunching. The simple argument is that, if one can estimate the extent of "manipulation in the running variable", which in our case corresponds to the excess mass below the threshold among untreated units (non-publicized contracts), then one can derive bounds on treatment effects by assuming that these units are the ones with either the highest or the lowest values of the outcome variable $Y_t$. Intuitively, these are computed under the "worst" and "best" case scenarios in terms of how selection can influence RDD estimates. In Appendix Section G.1.8, we explain in detail how to derive these bounds in our setting, and how to calculate them using our estimate of excess bunching obtained in our density analysis.

### 3.3.2 Effects on Non-Price Outcomes

**Naive RDD Results.** We start with the naive RDD results, and then sequentially apply corrections to account for the specific issues present in our setting. We estimate specifications Equation (1), Equation (2), and Equation (3), assuming that $\tilde{p}_t = p_t$. In our baseline specifications, we use a simple linear fit for $g(\cdot)$ and no controls $X_t$, but also present results from the robust local polynomial approach proposed by Calonico, Cattaneo, and Titiunik (2014). We present these naive RDD results visually, by plotting binned scatters of Equation (2) and Equation (3). In the next section we explicitly assess how these baseline estimates change as we consider the impact of price effects and (or) bunching responses.[16]

The results for the first stage Equation (2) are presented graphically in Figure 2(a). We see that the use of FedBizOpps jumps sharply past the $25,000 threshold of award amounts. The share of contracts that are publicly solicited in the government platform increases from roughly 30% at or slightly below $25,000, to 50% right above this threshold.

The reduced form specifications (Equation (3)) are estimated on three sets of outcomes: the intensity of competition, winning vendor characteristics (including its relationship with the awarding office), and post-award performance. Most of the existing literature has studied these variables independently.[17] By studying them jointly, we can generate a comprehensive understanding of the mechanisms and implications of policies oriented to enhance competition.

Figure 2(b) shows how posting solicitations on FedBizOpps impacts the number of offers that a contract receives around the threshold. Contracts right above $25,000 (which are more likely to be publicly solicited), receive roughly 0.4 more bids. The magnitude of the increase in the number of offers is considerable given that the policy only changes the likelihood of a publicized solicitation

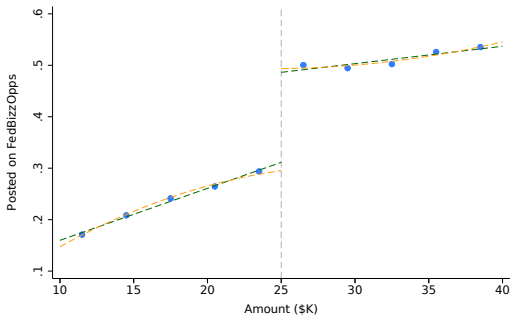
is possible to recover causal parameters even if we drop assumption **A2**.

[16] Appendix Figure A5 presents RDD plots for baseline variables. We find that baseline contract design characteristics are balanced around the threshold, with the exception of goods vs. services. There are more services right above the threshold. The difference is noisy and against possible selection patterns. All of our baseline estimates are robust to the inclusion of a service dummy as control.

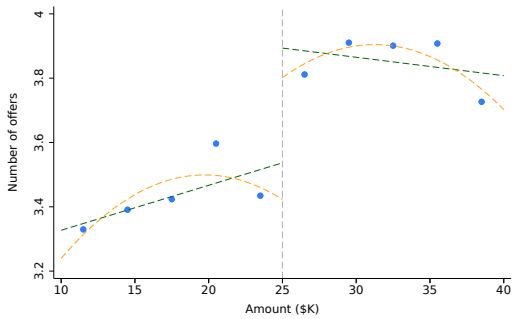
[17] See, for example, Athey (2001); Li and Zheng (2009) (competition), Macleod and Malcomson (1989); Bajari et al. (2009); Malcomson (2012) (relations), and Bajari et al. (2014); Decarolis et al. (2020); Ryan (2020) (ex-post renegotiation and performance).

Figure 2: Publicizing requirement and intensity of competition

(a) Use of FedBizOpps




(b) Number of offers received




Notes: Panels (a) shows the fraction of contracts posted on FedBizOpps. Panel (b) shows the number of offers received. Each graph includes the averages by bins of award amounts (blue dots), as well as linear and quadratic fits at each side of $25,000. The data sources are FBO.gov and the Federal Procurement Data System-Next Generation. The sample consists of competitive, non-R&D, definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

by around 20 p.p.

These results indicate that encouraging the public posting of solicitations leads to the stated goal of increasing competition by attracting additional bids. However, it does not necessarily imply that these new offers affect the equilibrium allocation of the contract, since new marginal bidders may not be competitive. Figure 3 shows that this is not the case. In Panel (a), we see that publicized contracts are awarded to vendors that are relatively larger, as measured by a reduction of the probability of awarding the contract to a small firm.[18] This "penalty" for small businesses is interesting because it goes against the stated goals of the publicity regulation (FAR Part 5). Panel (b) and Panel (c) show that publicized contracts are more likely to be awarded to foreign firms or firms that are located geographically more distant from the contracting office location. These results suggest that marginal entrants attracted by the public solicitation do win awards with a positive probability.

To measure the impact on post-award contract performance, we use two measures that are commonly used in the literature: cost overruns and delays (e.g. Decarolis, 2014; Kang and Miller, 2017; Decarolis et al., 2020; Carril, 2019). Because the data contain the total sum of payments and completion date expected at the time of the award for each contract, we can construct measures of cost overruns and delays by comparing these expectations to the realized payments and duration. These measures have been used by recent studies as performance proxies.[19]

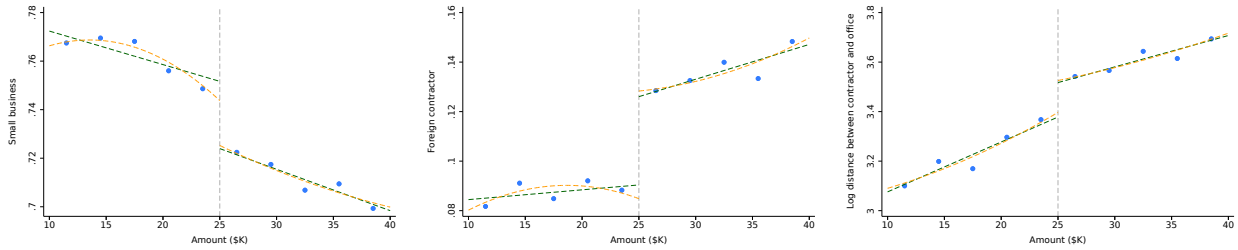
Figure 4 presents the results. We find that the share of contracts with overruns and the share of contracts with delays increase by 2 p.p. and 1.5 p.p., respectively. These differences are statistically

[18]The Small Business Administration (SBA) defines size standards by NAICS Industry. These standards depend on the number of employees and/or annual revenue. As a reference, the revenue standard for *building cleaning services* (NAICS code 561720 ), a common category in the sample, is $ 19.5 million per year.

[19]The FPDS data records whether the modifications are in or out of contract scope. Our analysis does not restrict a specific type of renegotiation, although *out-of-scope* modifications are extremely uncommon in our sample.

Figure 3: Publicity and the characteristics of the winning firm

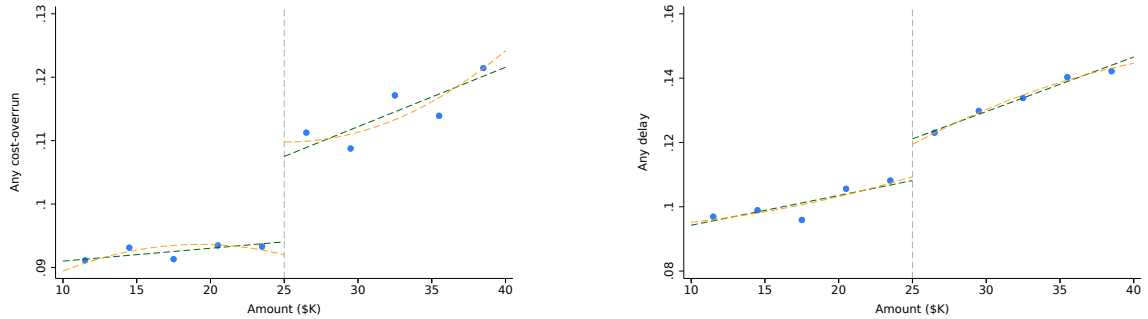
(a) Contractor is a small business (b) Foreign firm (c) Distance to the office




Notes: This figure presents four binned scatter plots, which depict an average outcome by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The outcome in each Panel is as follows: (a) indicator equal to one if the awarded contractor is a small business (based on SBA); (b) an indicator equal to one if the contract is awarded to a foreign vendor; (c) the natural logarithm of the distance (in miles) from the contracting office's location and the vendor location. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D, definitive, and competitively awarded contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

Figure 4: Publicity and post-award contract performance

(a) Share of contracts experiencing cost-overruns (b) Share of delayed contracts




Notes: This figure presents four binned scatter plots, which depict an average outcome by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The outcome in each Panel is as follows: (a) the share of contracts that the actual obligated contract dollars exceed expected total obligations at the time of the award (i.e., cost-overruns); (b) the share of contracts whose actual days of contract duration exceed the expected days of duration at the time of the award (i.e., delays). The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D, definitive, and competitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

and economically considering the magnitude of the first stage. These results show that the execution of publicized contracts tends to result in poorer performance outcomes, including ex-post costs. Figure A6 in the Appendix shows effects on additional performance-related variables; the number of post-award contract modifications, cost-overrun dollars as a share of the original award; and days of delay relative to expected schedule. These results align with the findings presented in Figure 4: publicized contracts experience more problems during the execution stage. In Supplementary Material F.1 we provide figures illustrating how these effects vary by agency and type of product requested.

**Adjusted RDD Results.** In this section, we present a series of refinements to our naive RDD results. First, we explore robustness to our baseline linear specification with the estimator proposed by Calonico, Cattaneo, and Titiunik (2014), which uses robust local polynomial fits. Second, building upon the results of our density analysis in Section 3.2, we adjust the baseline RDD estimates to account for the observed running variable (award price) being subject to both treatment effects (price effects of publicity) and potential manipulation (bunching). The price effect correction is related to existing methods that account for measurement error in the RDD framework (Pei and Shen, 2017). The key advantage of our setting is that we express the conditional expectation of contract outcomes given *observed* prices $E[Y_t|p_t]$, as an explicit function of magnitudes that we estimated in the density analysis, plus the causal parameters that we seek to recover. On the other hand, we follow Gerard, Rokkanen, and Rothe (2020)'s approach to account for the potential effect of bunching responses. They show that given an estimate of the extent of bunching —which we obtained in Section 3.2—, we can bound the estimated treatment effects under "worst" and "best" case scenarios of how selection influences the RDD estimates. The width of these bounds shrinks as the extent of bunching decreases, converging to the point estimates in a case with no manipulation.

Table 1 presents reduced-form estimates for each relevant outcome variable. The first column shows the coefficient of our naive linear RDD using ordinary least squares (OLS). These results replicate the RDD plots discussed earlier. Column (2) presents Calonico et al. (2014)'s local polynomial estimates with robust bias-corrected standard errors. Overall, non-linear estimates are similar in magnitude and significance to simple OLS estimates. The third column present estimates that account for price effects in the treatment group (i.e. publicized contracts), following the method explained in Appendix G.1.6. The correction for price effects is relatively modest, and in most cases tends to amplify the naive results. This is consistent with the fact that the price effects smooth-out the discontinuity for the treatment group. Thus, under naive estimation, some publicized contracts are observed below the threshold when their original (ex-ante) price was above it.

The next two columns present partial identification estimates that account for bunching responses. Column (4) shows lower and upper bounds without accounting for price effects, while the fifth column shows bounds that do adjust for price effects. Notably, since the magnitude of bunching is modest in our context, the bounds presented are relatively narrow, which tells us that bunching does not pose a serious threat to the interpretation of our results. Interestingly, the lower bounds in Column (5) tend to be very close to our baseline estimates. This implies that

Table 1: Reduced-form RDD Estimates and Corrections

| Dependent Variable | OLS (1) | CCT (2) | Price Effect Adjustment (3) | Manipulation Bounds (4) | Price Effect + Manip. Bounds (5) |
|---|---|---|---|---|---|
| Number of offers | 0.3569 (0.0677) | 0.5447 (0.1053) | 0.3526 | [ 0.2762 , 0.5344 ] | [ 0.3073 , 0.4506 ] |
| One offer | -0.0191 (0.0064) | -0.0235 (0.0108) | -0.0204 | [ -0.0272 , 0.0052 ] | [ -0.0248 , -0.0070 ] |
| Log distance firm-office | 0.1392 (0.0481) | 0.1199 (0.0817) | 0.1909 | [ 0.0290 , 0.2688 ] | [ 0.1304 , 0.2619 ] |
| Foreign firm | 0.0357 (0.0045) | 0.0508 (0.0078) | 0.0375 | [ 0.0328 , 0.0520 ] | [ 0.0358 , 0.0465 ] |
| New firm | 0.0175 (0.0075) | 0.0185 (0.0126) | 0.0247 | [ 0.0023 , 0.0350 ] | [ 0.0164 , 0.0344 ] |
| Small business | -0.0277 (0.0065) | -0.0295 (0.0110) | -0.0265 | [ -0.0523 , -0.0195 ] | [ -0.0399 , -0.0220 ] |
| Any cost-overrun | 0.0135 (0.0045) | 0.0246 (0.0077) | 0.0144 | [ 0.0103 , 0.0263 ] | [ 0.0127 , 0.0216 ] |
| Cost-overruns (relative dollars) | 0.0095 (0.0058) | 0.0161 (0.0100) | 0.0127 | [ 0.0053 , 0.0179 ] | [ 0.0104 , 0.0174 ] |
| Any delay | 0.0130 (0.0047) | 0.0151 (0.0080) | 0.0143 | [ 0.0094 , 0.0270 ] | [ 0.0123 , 0.0222 ] |
| Delays (days) | 2.3262 (2.0388) | 4.0361 (3.4935) | 2.7491 | [ 1.2278 , 5.3554 ] | [ 2.1492 , 4.4660 ] |
| Number of modifications | 0.0375 (0.0173) | 0.0619 (0.0300) | 0.0395 | [ 0.0204 , 0.0926 ] | [ 0.0300 , 0.0701 ] |

Notes: This table shows Regression Discontinuity Design (RDD) estimates of the reduced-form relationship between a series of outcome variables and an indicator of whether a contract award price exceeds $25,000. Coefficients in column (1) use a linear fit above and below the discontinuity. Coefficients in column (2) correspond to the robust local polynomial method proposed by Calonico, Cattaneo, and Titiunik (2014). Column (3) applies a correction to the estimates in column (1), accounting for the existence of price-effects, following the method proposed in Appendix G.1.6. Column (4) shows bounds on the reduced-form coefficient in column (1), accounting for the possibility of "running variable manipulation" (i.e. bunching), following the method proposed in Appendix G.1.8. Column (5) shows bounds on the adjusted reduced-form coefficient in column (4), accounting for both the existence of price-effects and the possibility of "running variable manipulation" (i.e. bunching). Standard errors for the coefficients in columns (1) and (2) are shown in parentheses.

the downward bias introduced by price effects on the naive estimates of Column (1) is of similar magnitude than the worst-case upward bias introduced by bunching responses.

Taken together, these results imply that the strong visual evidence presented in Figures 2(a) through 4 is robust to fully accounting for the potentially confounding influence of price effects and strategic bunching by the buyer.

### 3.4 The Role of Contract Complexity

Our analysis includes a wide variety of transactions, from standardized goods to customized services. A procurement contract aims to regulate the nature of the expected transaction. Nevertheless, specifying possible contingencies is easier if the purchase involves a commodity-type

product rather than an *ad-hoc* service. Thus, the more difficult it is to specify the need in a contract, the more variable will be the post-award performance. This explains why some product categories rarely experience performance issues ex-post, while others experience implementation issues in most of the contracts. Similarly, the effects of expanding competition on award prices are also likely to vary depending on the good or service's underlying complexity. For example, if bidders of relatively complex products are more heterogeneous in production costs, additional offers would lower contract prices more than when contractors are homogeneous. Thus, there are reasons to believe that the degree of contract complexity affects both prices ex-ante and ex-post.

To assess these mechanisms more directly, we leverage the rich heterogeneity of our data, which features 1,918 distinct product categories, and we proxy the degree of contract complexity based on the baseline level of post-award performance, which we define as the average cost-overrun experienced by all contracts below $20,000 for the same product category.[20] Table B.5 describes the top and bottom product categories based on average delays and cost-overruns. Contracts for services experience substantially more issues in the implementation stage than goods. Using the measure of mean cost-overruns, we divide the contracts categories in our sample into quartiles of complexity, and re-estimate both price effects and RDDs on performance, separately for each of the four groups. We also consider the more simple heterogeneity of effects between goods and services.

Table B.3 shows estimates for the mean and standard deviation of price effects $\gamma_t$, separately for the full sample (column 1), goods versus services (columns 2 and 3), and each of the four quartiles of complexity (columns 4 through 7). Similarly, Figure 1 shows the CDFs of price effects for each of these groups. Although estimates become noisier as we divide the sample, we see suggestive evidence that large price effects are more concentrated among the most complex contracts. Our point estimates indicate that, on average, publicity reduces the prices of goods by 5% and of services by 7.8%. This effect corresponds to 4% for the least complex quartile, versus 9.6% for the top quartile of complexity.

The results are qualitatively similar for the impact of publicity on post-award performance. Figure 5 shows that the increase in overruns and delays that we reported in Figure 4 is driven by goods and services in the *top* quartile of complexity. We are unable to reject the null for the lower three quartiles. Overall, we see the effects in overruns outweighs the price reductions ex-ante for complex contracts. However, when the unit is "simple" the price reductions, although modest, exceed increases in cost-overruns ex-post.[21]

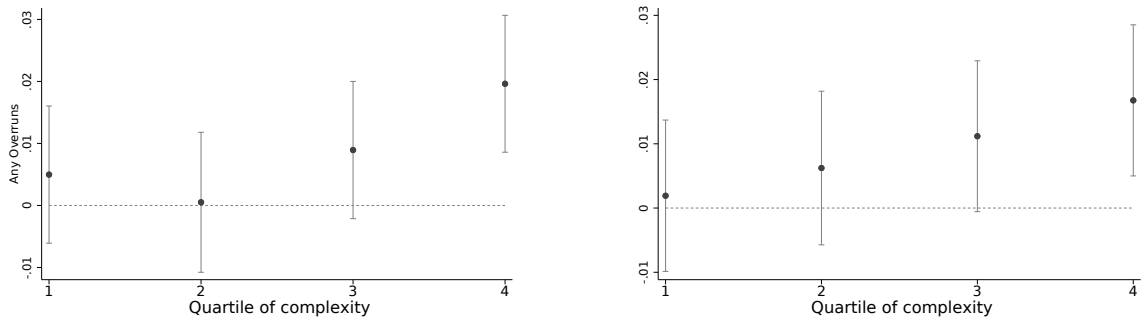
[20] There are multiple ways of characterizing product complexity. We tried different approaches, e.g., using performance's standard deviation, indexing multiple variables, counting the number of words in the solicitation's description, etc. These classifications lead to roughly the same rank of products categories, and thus varying the definition does not threaten the general results (see correlations in Figure A7). We use the mean of cost-overruns because it is transparent and easy to interpret. We get around the issue of classifying based on an outcome by focusing on contracts below $20,000.

[21] Figure A8 in the Supplementary Material shows RD plots for cost-overruns separating for goods and services. Note that cost-overruns increase for both types of contracts. However, both the baseline level and the magnitude of the jump are substantially larger for services.

Figure 5: Performance effects by sub-group

(a) Reduced-Form Effects on Cost-Overruns by Complexity Quartile

(b) Reduced-Form Effects on Delays by Complexity Quartile




Notes: This figure shows four regression coefficients and their 95% confidence intervals. Each coefficient is a estimate of a RD reduced-form Equation (3) per sub-group estimated using (interacted) OLS. The dependent variable of Panel (a) and (b) are indicators for any positive cost-overruns and delays, respectively. The subgroups are determined by the four quartiles of a proxy of contract complexity. The contract complexity proxy is constructed at the product category level and is defined as the average cost overruns for contracts with awards below $20,000 in that category. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019.

### 3.5 Evidence of Adverse Selection

Our results show that increasing the pool of bidders through publicity generates changes to contract prices and the subsequent contract execution. Overall, there are two classes of explanations through which we can rationalize the connection between publicity and contract outcomes.[22] The first explanation is that contractors modify their behavior depending on the publicity status of the contract (i.e. *moral hazard*).[23] The second explanation is that publicity allows the participation of suppliers that are "different," an that their performance capacity is unrelated to the identity of the buyer or contract's advertising (i.e. *adverse selection*).

We combine features of our setting with empirical methods used to detect asymmetric information (e.g., Chiappori and Salanie, 2002), in order to elucidate these mechanisms. In particular, we leverage that buyers often require the same product categories repeatedly over time; indeed, we observe multiple contracts for the same buyer-product combination, with variation in the size of the award and other dimensions. Moreover, from the supplier's side, we observe most contractors executing more than one contract, for one or more different buyers. This variation allows us to test how much of the observed variation is due to contractor's "types," relative to variation "within" contractor. To do this, we re-estimate the RDD analysis on post-award

[22] Publicizing contracts in FBO.gov impacts neither the contract's design nor the mechanism of selection. Appearing in FBO.gov solely affects the diffusion of information.

[23] This could be rationalized if suppliers behaved differently depending on the buyer. For example, if a vendor receives contract information directly from the buyer, she could decide to absorb potential overruns to make sure she gets direct information again in the future.

performance including contractor fixed-effects. The fixed effects demean contractors' performance, allowing us to test how much variation remains. In Appendix Table B.4, we show that the changes in performance disappear once we include contractor fixed effects. This implies that most of the variation on contract performance introduced by publicity is explained by variation across contractors, as opposed to "within" contractors, suggesting that *adverse selection* is the main feature driving these results. The Appendix section E.1 describes how we can classify contractors depending on their usual source of information about contract solicitations.

### 3.6 Discussion

This section provided evidence that promoting vendor participation through publicity increases contract competition, as the average number of offers received rises substantially. The added competition translates into reductions in contract prices. Leveraging detailed information about contracts' implementation, we also found that publicized contracts result in more cost-overruns and delays. Taken together, our results show that promoting contract competition for contracts that are (at least partially) incomplete involves a trade-off: it reduces contract award prices at the cost of attenuating adverse selection by awarding contractors that are unable to execute contracts at the desired quality. Thus, the desirability of promoting competition depends on which of these effects dominate. Furthermore, we find that this trade-off is heterogeneous, with both price effects and performance effects depending on the degree of contract complexity.

While this policy analysis is informative on the effects of promoting competition on contract outcomes, several questions about the underlying market structure that shape bidders' adverse selection remain unanswered. In particular, our reduced-form analysis does not allow us to evaluate equilibrium conditions under alternative policy designs considering buyers' preferences over specific contractors in a setting of vast heterogeneity on complexity across contracts. To make progress on these fronts, we now present an estimate an equilibrium model of competition promotion, firms' participation, and bidding decisions.

## 4 A Model of Competition Promotion, and Firms' Participation and Bidding Decisions

We develop and estimate an equilibrium model of publicity selection, firm participation, and bidding decisions in the public procurement setting. The ultimate goal is to estimate the model's primitives and study the implications of policy counterfactuals. We make modeling assumptions based on the setting's key features, aiming to transition from a theoretical model to an empirical one that can be estimated using the data available. Section 4.1 introduces the theoretical model and discusses the auction's entry and bidding equilibrium strategies. Section 4.2 illustrates the empirical implementation of the model. In Section 4.3 we discuss key variation in the data to identify model parameters. The results are discussed in Section 4.5.

## 4.1 Model

A buyer offers a single and indivisible contract to $N$ potential contractors. Each potential contractor $j$, must incur an entry cost $\omega_j > 0$ to learn her private cost to complete the task $c_j \in [\underline{c}, \overline{c}] \subset \mathbb{R}_+$ and, hence, bid for the contract. Both, $\omega_j$ and $c_j$ are assumed independent random draws from specific distributions. We model the potential bidders' choices in two stages; first, knowing the number of potential competitors $N$, each potential bidder decides whether to incur the entry cost. After the entry stage, the $n \leq N$ firms that incurred entry costs learn their costs of completing the job and submit their bids. The awarding mechanism is a first-price sealed-bid auction; the contract is awarded to the bidder that submits the lowest offer. The quality of the contract execution $q_j$ is observed once the contract is finished.

Our analysis considers asymmetry between potential contractors. In particular, there are two types of firms, locals ($L$) and non-locals ($NL$). These firms differ in their distributions of entry costs, $G^k_\omega$, the cost of completing the project, $F^k_c$, and the execution quality $F^k_q$, where $k \in \{L, NL\}$ and $k(j)$ denotes group affiliation of bidder $j$. We assume that project and bid preparation costs are private information of each firm and are distributed independently across all firms and identically within group.

Departing from existing literature, our model allows for the set of potential contractors to be chosen endogenously, i.e., the buyer decides whether to publicize the contract taking into account the expected price, the quality of execution and the likelihood that a contractor of each group is awarded the contract. The publicity decision determines the set of potential participants as follows; if the contract solicitation is publicized openly, both locals and non-local contractors learn about the contract solicitation. Conversely, if the contract is not advertised, only the local contractors receive the information.

### 4.1.1 Equilibrium in the Bidding Stage

We start by characterizing the bidding stage and then use the results to analyze the participation stage. Our analysis focuses on a group-symmetric equilibrium where bidders of group $k$ follow the same bidding strategy, $\beta_k(\cdot)$, mapping project cost, $c_j$, into a bid $b_j$. Where $c_j$ is drawn independently from a type-specific continuous distribution $F^k_c(\cdot)$, with density $f^k_c(\cdot)$ and common support $[\underline{c}, \overline{c}] \subset \mathbb{R}_+$. The distributions of entry and production costs, and the number of potential bidders of each type are common knowledge. Nevertheless, we assume that bidders do not observe the number of actual competitors of each group $n^k_t$ (Li and Zheng, 2009).

Our setting considers two possible scenarios; if the contract solicitation is publicized, then both, *local* and and *non-local* firms could participate. In this case, the expected utility of bidder $j$ with cost realization $c_j$ and group membership $k(j)$ depends on the number of bidders of each group:

$$\mathbb{E}\left[\pi_j(c_j)\right] = (b_j - c_j)\left(\sum_{l=2}^{N^{k(j)}} \rho_l^{k(j)}\left(1 - F_c^{k(j)}\left(\beta_{k(j)}^{-1}(b_j)\right)\right)^{l-1}\right)\left(\sum_{l'=1}^{N^{-k(j)}} \rho_{l'}^{-k(j)}\left(1 - F_c^{-k(j)}\left(\beta_{-k(j)}^{-1}(b_j)\right)\right)^{l'}\right)$$

where $\rho_l^{k(j)}$ is the probability that the number of actual bidders is equal to $l$, and $-k(j)$ denotes the other group of potential contractors. The optimal bidding requires solving a system of differential equations corresponding to the first order conditions for both types of bidders as follows:[24]

$$\begin{aligned}\frac{1}{b_j - c_j} &= f_c^{k(j)}\left(\beta_{k(j)}^{-1}(b_j)\right)\frac{\partial \beta_{k(j)}^{-1}}{\partial b_j}\left[\frac{\sum_{l=2}^{N^{k(j)}} \rho_l^{k(j)}(l-1)\left(1 - F_c^{k(j)}\left(\beta_{k(j)}^{-1}(b_j)\right)\right)^{l-2}}{\sum_{l=1}^{N^{k(j)}} \rho_l^{k(j)}\left(1 - F_c^{k(j)}\left(\beta_{k(j)}^{-1}(b_j)\right)\right)^{l-1}}\right] \\ &+ f_c^{-k(j)}\left(\beta_{-k(j)}^{-1}(b_j)\right)\frac{\partial \beta_{-k(j)}^{-1}}{\partial b_j}\left[\frac{\sum_{l'=1}^{N^{-k(j)}} \rho_{l'}^{-k(j)} l'\left(1 - F_c^{-k(j)}\left(\beta_{-k(j)}^{-1}(b_j)\right)\right)^{l'-1}}{\sum_{l'=1}^{N^{-k(j)}} \rho_{l'}^{-k(j)}\left(1 - F_c^{-k(j)}\left(\beta_{-k(j)}^{-1}(b_j)\right)\right)^{l'}}\right] \end{aligned} \tag{4}$$

If a contract solicitation is not publicized, only *local* firms can bid, i.e., the number of potential non-local contractors is zero. In this case, the bidding problem is symmetric as there is only one group involved. Local suppliers observe contracts' publicity status and hence the number of potential competitors.

### 4.1.2 Equilibrium of the Entry Stage

Firms compare the *ex-ante* expected profit conditional on entry to their entry cost $\omega_j$. Where $\omega_j$ is independently drawn from a type-specific continuous distribution $G_\omega^k(\cdot)$ with common support $[\underline{\omega}, \overline{\omega}]$. Firms with entry costs below their expected profit decide to incur the entry fee to learn about their cost of completing the project. The *ex-ante* (expected) profits from participating are given by:

$$\bar{\pi}^k(\varphi^k, \varphi^{-k}) = \int_{\underline{c}}^{\overline{c}}\left(\sum_{n_k-1, n_{-k} \subset N_k-1, N_{-k}} \pi^k(c|n^k - 1, n^{-k}) Pr\left(n^k - 1, n^{-k}|N^k, N^{-k}\right)\right) dF_c^k(c) \tag{5}$$

where $\varphi^k$ and $\varphi^{-k}$ are the entry probabilities of each group. Because entry decisions are made simultaneously, the equilibrium condition is characterized by a group-specific entry cost threshold

[24] As noted by previous research on asymmetric auctions (Lebrun, 1999; Bajari, 2001; Maskin and Riley, 2003a,b), the Lipschitz conditions are not satisfied in this case. The bidding strategies cannot be solved analytically but require numerical methods. Campo, Perrigne, and Vuong (2003) and Brendstrup and Paarsch (2003) discuss non-parametric identification of cost functions in this setting.

$\bar{\omega}_k$, i.e., firms whose entry cost is below their group's threshold participate.[25] Finally, when the contract is not publicized, only locals could participate, and thus the participation problem becomes symmetric. Thus, for a given contract $t$, the local's participation threshold differs depending on whether the contract was publicized.

### 4.2 Empirical Model

Based on equilibrium conditions of the general model, we proceed to describe its implementation based on the empirical setting. A contract solicitation $t$ is characterized by $(x_t, z_t, u_t, N_t)$, where $x_t, z_t$ are observable characteristics, e.g., type and complexity of the product required, location, acquiring agency, etc, and $u_t$ is the unobserved project heterogeneity that captures project attributes that are not included in the data but impact firms' bidding behavior. Finally, $N_t = (N_t^L, N_t^{NL})$ denotes the number of potential contractors of each group.[26] The model thus proceeds in four stages depicted in the Figure 6:

T = 0: *Publicity Decision.* In the first stage the buyer observes $(x_t, z_t, u_t, N_t)$ and decides whether to publicize the contract to maximize expected utility. Contract publicity status determines the set of potential bidders.

T = 1: *Entry Decision.* In the second stage, each firm that learns about the contract observes $(x_t, z_t, u_t, N_t)$. They draw individual and private realizations of entry cost, and they simultaneously decide whether to participate.

T = 2 *Bid Decision.* Active bidders draw a realization of the production cost and decide the magnitude of their bid. The contract price equals the lowest bid submitted.

T = 3 *Execution Stage.* The implementation quality is realized once the contract is finished based on a publicly observed quality shock.

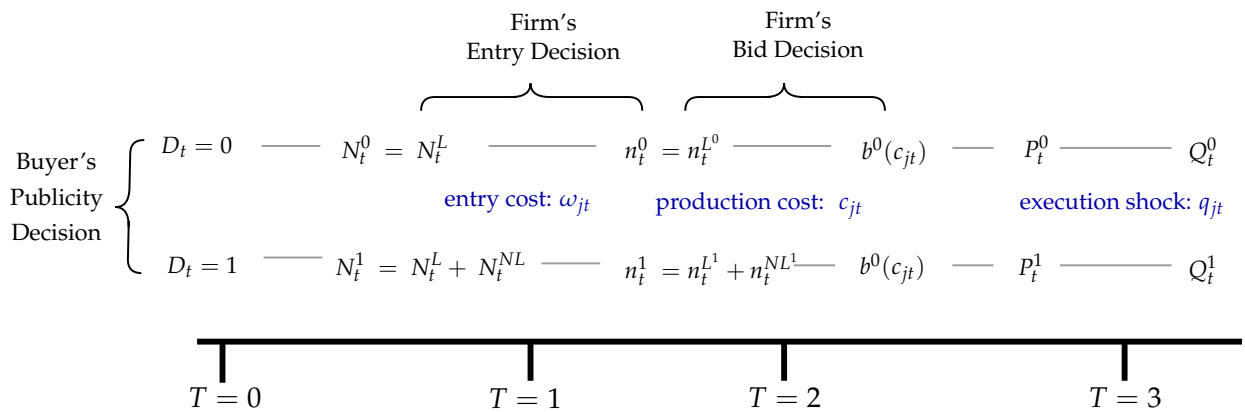
[25] In equilibrium, the entry probabilities are defined by the system of equations:

$$\begin{aligned} \varphi^L &= G_\omega^L\left[\bar{\omega}_L(\varphi^L, \varphi^{NL})\right] \\ \varphi^{NL} &= G_\omega^{NL}\left[\bar{\omega}_{NL}(\varphi^L, \varphi^{NL})\right] \end{aligned}$$

Group-specific equilibrium exist by Brouwer's Fixed Point Theorem. We numerically verified uniqueness of the equilibrium entry probabilities within our estimation routine (Krasnokutskaya, 2011; Roberts, 2013). Espin-Sanchez et al. (2021) discuss sufficient condition for equilibrium uniqueness in entry games with private information.

[26] Identifying the potential number of bidders is not trivial (Athey, Levin, and Seira, 2011; Krasnokutskaya and Seim, 2011; Mackay, 2018). We combine two methodologies: First, using the procedure described in Appendix section E.1, we classify and count the suppliers that ever won a contract for every buyer-product combination. The second method considers the maximum number of actual bidders for buyer-product auctions. This method is discussed by Athey, Levin, and Seira (2011); Roberts (2013). It is rooted in the theoretical idea that if all potential bidders decide whether to enter simultaneously, with enough observations, the maximum number of observed bidders across observations will be equal to the total number of potential bidders. The maximum number of bidders of auctions that weren't publicized informs about the number of potential local bidders. In contrast, the maximum number of bidders of advertised contracts approximates the sum of local and non-local potential bidders. Finally, we define the number of potential bidders for every buyer-product as the maximum of both approaches. Combining these two methods alleviates potential weaknesses of each of them. The median number of potential local and non-local bidders is six and three, respectively.

Figure 6: Timing of the Model




Now, we outline specific modelling assumptions of each stage of the model. In sections 4.3 and 4.4, we discuss the model's identification and estimation strategy, respectively, to recover the parameters of the underlying distributions from available data.

**Publicity Decision.** We assume the buyer is risk-neutral, form unbiased beliefs, and derives utility on expected contract outcomes linearly:

$$\begin{aligned} U_t^D &= U(\bar{P}_t^D, \bar{Q}_t^D, \bar{L}_t^D) \\ &= \beta^P \bar{P}_t^D + \beta^Q \bar{Q}_t^D + \beta^L \bar{L}_t^D + x_t'\zeta + \varepsilon_t^D, \end{aligned}$$

where $D_t \in \{0, 1\}$ denotes contract's publicity status, and $\bar{P}_t^D$, $\bar{Q}_t^D$ and $\bar{L}_t^D$ are the expected awarding price, implementation quality and likelihood that a local wins, conditional on $(x_t, z_t, u_t, N_t)$, and $\varepsilon_t^D$ is an idiosyncratic utility shock.

$$D_t = 1 \iff U(\bar{P}_t^1, \bar{Q}_t^1, \bar{L}_t^1) + \eta \mathbb{1}(\bar{P}_t^0 > 25) \geq U(\bar{P}_t^0, \bar{Q}_t^0, \bar{L}_t^0)$$

The publicity regulation kicks in when the expected award price is higher than \$25,000. It introduces a utility shift $\eta$, which translates into a discrete jump of the probability of advertisement at the threshold.

**Entry and Bidding Decision.** The bidder's $j$ cost for contract $t$ is multiplicative: $c_{jt} = \tilde{c}_{jt} \cdot u_t$, $\tilde{c}_{jt}$ is a firm-specific cost component that is private information of firm $j$, while $u_t$ represents a common cost component that is known to all bidders but is unobserved to the researcher (Haile and Kitamura, 2019). The distribution of the firm-specific cost component for group-k firms is given by $F_{\tilde{c}}^k(\cdot|x_t)$, and is independent conditional on observables. The unobserved project heterogeneity is given by $u_t \sim H(\cdot)$, is independent from project characteristics and the number of potential bidders.

We assume bidders are risk neutral. Thus, the Bayes-Nash equilibrium bid function for group $k$ is multiplicative: $\beta_k(c_{jt}|x_t, z_t, u_t, N_t) = u_t \cdot \tilde{\beta}_k(\tilde{c}_{jt}|x_t, z_t, N_t)$.[27] Each bidder submits a bid of $b_{jt} = \tilde{b}_{jt} u_t$

[27] This relation is discussed by Krasnokutskaya (2011), Proposition 1. It shows that, when the cost function is multiplicative to unobserved heterogeneity, a Bayes-Nash equilibrium bidding strategies are also multiplicative.

where $\tilde{b}_{jt} = \tilde{\beta}_k(\tilde{c}_{jt}|x_t, N_t)$ represents the bid for bidder $j$ when $u_t$ is one. Therefore, $ln(b_{jt}) = ln(u_t) + ln(\tilde{b}_{jt})$, the log of the unobserved heterogeneity component acts as an additive mean shifter to the conditional distribution of log bids. The contract is awarded using a first-price auction to the bidder that submits the lowest bid.

Finally, we assume the entry costs $\omega_{jt}$ are independent conditional on observed project characteristics $x_t$. In equilibrium, the firms' participation behavior is characterized by group-specific thresholds, $\bar{\omega}_t^{k*}$. Thus, the number of actual bidders $n_t^k$ from group $k \in \{L, NL\}$ distributes binomial with an individual entry probability of $\varphi^{k*}(x_t, z_t, u_t, N_t)$ and $N_t^k$ trials. Where $\varphi^{k*}(x_t, z_t, u_t, N_t) = G_\omega^k\left(\bar{\omega}^{k*}(x_t, z_t, u_t, N_t)\right)$. Our model considers entry shifters $z_t$, which capture market-level conditions that affect entry decisions.

**Contract Execution.** The execution is observed *ex-post*, after the contract is performed by the selected contractor. The observed execution is $q_{jt}$, and it corresponds to a draw from group-specific distribution $F_q^k(\cdot|x_t)$. In this section, we focus on one direct measure of performance that is the existence and magnitude of cost overruns. This variable is convenient as it can be directly benchmarked with the contract's dollar value.[28]

Equilibrium is characterized by the buyer choosing contract's publicity status that maximizes her expected utility; informed potential contractors entering if expected profits exceed entry costs and bidding optimally in the market mechanism. Finally, the contract execution is revealed once the contract concludes.

### 4.3 Identification

We aim to identify the type-specific distributions of $\omega_{jt}$, $\tilde{c}_{jt}$ and $q_{jt}$, the distribution of the unobserved heterogeneity, $u_t$, and the parameters that govern buyer's utility function. For every contract $t$, we observe in the data a set model *inputs* $(x_t, z_t, N_t)$ and *outputs* $(D_t, n_t, P_t, Q_t, L_t)$. The model is identified based on three main assumptions:

(i) Contract and market characteristics $(x_t, z_t, N_t)$ are exogenous.

(ii) Idiosyncratic component of entry cost shocks are (conditionally) independent from production cost and quality shocks, i.e., $\omega_{jt} \perp (c_{jt}, q_{jt})|x_t$.

(iii) Unobserved heterogeneity $u_t$ is independent with $\mathbb{E}[u_t|x_t, z_t, N_t] = \mathbb{E}[u_t] = 1$.

The model identification involves pinning down primitive distributions of the two types of bidders (locals and non-locals). In our setting, the contract's publicity status determines the composition of participating bidders; if a contract is not publicized, only locals could participate, while if it's advertised, both types can compete for the contract. Thus, data of unpublicized contracts inform about the distributions of locals, while non-locals are only observed on publicized

[28]This approach abstracts away from other (context-specific) execution costs. For example, Lewis and Bajari (2011) study the welfare gains associated with reducing delays on high-way construction.

contracts. Importantly, identifying type-specific distributions from the variation on contract publicity choices can only be achieved if the publicity decision is exogenous. We leverage the discrete nature of the publicity threshold to obtain quasi-experimental variation on publicity adoption and thus get exogenous variation in the publicity decision to identify type-specific distributions separately (in the spirit of RDD discussed before).[29] Below, we discuss identification of the different components of the model in separate steps:

**Bidding.** The empirical challenge involves separately identifying $F_{\tilde{c}}$ from $H_u$. The identification argument follows Mackay (2018) and builds upon exogenous variation on the number of bidders.[30] In our setting, bidders observe auction characteristics and the set of *potential* competitors, $N_t$, but do not know the set *actual* competitors, $n_t$. Equilibrium bidding strategies depend on the information they have in hand; bidders of auctions that lead to forming the same beliefs about the competitive environment would set the same bidding strategies. Thus, the number of *actual* competitors $(n_t|x_t, z_t, N_t)$ would depend on realizations of (random) individual entry cost shocks.

Exogenous variation in the number of entrants allows for identifying $N-1$ expected order statistics of the bidding distribution for each $(x_t, z_t, N_t)$ combination. Since $u_t$ is assumed independent, one additional competitor under the same bidding strategy is equivalent to one additional draw from the distribution of normalized bids; $G^k_{\tilde{b}}(\cdot)$. Restrictions over expected order statistics approximate the quantiles of $G^k_{\tilde{b}}(\cdot)$, and if $N \to \infty$, $G^k_{\tilde{b}}(\cdot)$ is exactly identified. The underlying cost distribution $F^k_{\tilde{c}}(\cdot)$ is pinned down from the distribution of $G^k_{\tilde{b}}(\cdot)$ (Guerre et al., 2000; Campo et al., 2003). See Appendix D.1 for more details and proofs.

**Entry.** Potential bidders set a threshold for realizations of entry costs. They pay $\omega_{jt}$ and enter auction $t$ only if the realization $\omega_{jt}$ is smaller than the expected profit of participating in the auction, i.e., $\omega_{jt} < \bar{\omega}^k_t$ . Since the probability of participating enters into the expected utility function which defines the cutoff, the (fixed-point) equilibrium entry cutoff is characterized by a type-specific entry probability $\varphi^{k*}_t = G^k_\omega(\bar{\omega}^{k*}_t)$. Identifying $G^k_\omega(\cdot|x_t)$ from the data entails three steps: First, we observe the *realized* fraction of potential bidders that decide to enter each auction $t$, with enough observations per combination of $(x_t, z_t, N_t)$, we can estimate $\varphi^{k*}(x_t, z_t, N_t)$. Then, we use equation (5) to back-up the expected utility of entering, conditional on $(x_t, z_t, N_t)$. The final step, leverages variation in $(z_t, N_t)$ to construct combinations of $(\varphi^{k*}, \bar{\omega}^{k*}|x_t)$ to pin down $G^k_\omega(\cdot|x_t)$

[29] On what follows we omit the distinction depending $D_t \in \{0, 1\}$, because it is given to the bidders. To ease notation, when a contract $t$ is publicized, the set of bidders has two dimensions: i.e., $N^1_t = (N^L_t, N^{NL}_t)$ and $n^1_t = (n^L_t, n^{NL}_t)$

[30] Alternative strategies to identify models with unobserved heterogeneity involve either stringent assumptions on auction participation or observing the full distribution of bids. Compiani et al. (2020) assumes the number of active bidders can be characterized by an (equilibrium) reduced-form relation, $n_t = \eta(x_t, z_t, u_t, N_t)$ that is weakly increasing in $u_t$, thus a realization of $n_t$ inform about (unobservable) realizations of $u_t$. Roberts (2013) provides a similar identification argument but leveraging variation in auctions' reserve price. Alternatively, Krasnokutskaya (2011) follows a measurement error approach and builds upon deconvolution methods to separately identify the distribution of unobserved heterogeneity and individual costs functions. The latter requires observing at least two bids per auction.

**Execution.** The contractor's execution (cost overrun) distribution stems from the realizations of individual execution shock. We assume these shocks are independent (conditional on observables). Thus, the observed distribution of cost-overruns informs directly about $F_q^k(\cdot|x_t)$.

**Buyer's Preference Parameters.** The buyer's taste parameters for price, overruns, and local contractors are identified from the exogenous nature of contract and market characteristics $(x_t, z_t, N_t)$. In particular, variation on the set of potential bidders determines the effects of publicity on price and on having a local winning the auction. The degree of complexity of the transaction helps pin down the potential scope for overruns ex-post. Intuitively, keeping other factors fixed, if a transaction involves a well-defined (commodity-type) product, there would be no difference in performance ex-post, which shuts down that factor in the decision.

Section 4.4 describes the estimation approach, which hinges on parametric assumptions of primitives' distributions. These distributions can be identified nonparametrically based on distributional assumptions discussed above. Nonparametric identification, along with the reduced-form results and robustness checks, suggests that the estimated distributions' features are not driven by specific functional forms.

### 4.4 Estimation Approach

In estimation, we make functional form assumption to characterize the distributions of interests.[31] We assume that the log of individual bids $log(\tilde{b}_{jt})$ are distributed normal with mean $\mathbb{E}[\tilde{b}_t^k|x_t, N_t] = [x_t, N_t]'\alpha^k$, and variance: $\mathbb{V}[\tilde{b}_t^k|x_t] = \left(exp(x_t'\nu^k)\right)^2$. We further assume $ln(u_t)$ is distributed normal with mean zero and variance $\sigma_u^2$.[32] In equilibrium, the entry decision is characterized by a type-specific probability, $\varphi_t^k(x_t, z_t, N_t)$, which depends on a type-specific entry-cost distribution (Krasnokutskaya and Seim, 2011; Athey et al., 2011, 2013). We assume $\varphi_t^k(x_t, z_t, N_t) = \Phi\left([x_t, z_t^e, N_t]'\tau^k\right)$, where $\Phi(\cdot)$ denotes the cumulative distribution of the standard normal distribution, and $z_t$ are entry cost shifters. The number of participating bidders $n_t^k(\cdot)$ is distributed binomial with $N_t^k$ independent draws with probability of success $\varphi_t^k(\cdot)$.[33]

The quality of the implementation is captured by the existence and magnitude of cost overruns. Given that most contracts stay right on budget, the observed distribution of cost overruns is censored at zero. We assume that $ln(q_t^k)$ is the latent distribution, while we only observe $Q_t^k = max\{0, ln(q_t^k)\}$, where $ln(q_t^k)$ distributes normal with mean $\mathbb{E}[q_t^k|x_t^q] = [x_t^q]'\gamma^k$, and variance

[31] Our parametric assumptions are linked to related literature. Overall, our results are not sensitive to adding additional covariates or variations to the functional form. Our data provide enough variation for identifying these distributions independent from the specific functional form.

[32] These parametric assumptions follow existing literature (Krasnokutskaya and Seim, 2011; Hong and Shum, 2002; Porter and Zona, 1993). Moreover, Krasnokutskaya (2011) indicates that the distribution of firm-specific components and unobserved heterogeneity closely resembles log-normality.

[33] Related papers either assume parametric distributions on the entry costs, which, paired with the expected utility of entering, map into well-defined group-specific entry probabilities (Krasnokutskaya and Seim, 2011; Mackay, 2018), or make functional form assumptions on the entry probabilities, which, combined with expected utilities, allow recovering entry costs (Athey et al., 2011, 2013). We follow the latter approach.

$\mathbb{V}[q_t^k|x_t] = \left(exp\{x_t^{q\prime}\xi^k\}\right)^2$. [34]

Finally, an important feature of this model is that we aim to recover buyer's preferences combining estimated parameters with observed publicity decisions. Buyer's utility is decomposed into observed and unobserved parts; the observed part is assumed to be a linear combination of expected outcomes, in particular, we define $\left(\tilde{\bar{P}}_t, \tilde{\bar{Q}}_t, \tilde{\bar{L}}_t\right)$ as the change in the expected outcome under publicity and no publicity, leaving expected outcome without publicity as the omitted category. We further assume the buyer has idiosyncratic preferences distributed standard normal. This way, $Pr(D_t = 1|\cdot) = \Phi\left(\beta^P\tilde{\bar{P}}_t + \beta^Q\tilde{\bar{Q}}_t + \beta^L\tilde{\bar{L}}_t + \eta\mathbb{1}(\bar{P}_t^0 > 25) + x_t'\zeta\right)$, where $\beta^P$, $\beta^Q$, and $\beta^L$ capture the relevant taste parameters for the price, quality, and idiosyncratic preference for awarding local vendors.[35] We control for observable characteristics as well as $\mathbb{1}(\bar{P}_t^0 > 25)$ to indicate whether the expected price without publicity is above the regulation threshold.[36]

**Observables.** We specify the mean of log bids as a linear function of the product's degree of complexity, an indicator of service, and the numbers of potential bidders of each group. We interact the complexity with a dummy for non-local bidders to allow the effects of most of these covariates to differ by bidder's group. The variance of log-bids depends on whether the group is a good or service. The probability of entry depends on the same covariates, and we add a dummy to indicate the solicitation was required at the end of the fiscal year as an entry cost shifter. The cost-overruns' shock is distributed log-normal; its mean depends on the product's degree of complexity, an indicator of service, interactions with non-local dummies, and an indicator if the expected contract duration exceeds the its mean.[37] Finally, the buyer's decision to publicize a contract depends on the expected differences (in logs) of price and overruns, as well as the predicted likelihood that a local wins. In addition to these variables, we include agency fixed effects and a dummy if the contract's expected value without publicity is over $ 25,000.

**Sample.** The data used to estimate the model is the same as those used in previous sections with the exception that to classify local and non-local vendors, we require buyer-product combinations that appear at least four times in the full database between 2013-2019, with at least one, but not all, appearances in FBO. This restriction rules out products that are purchased less often. Table B.7 compares the descriptive statistics of relevant variables between this selected sample and the full sample used in Section 3. Overall, given that the sample selection involves the buyer contracting the same product multiple times, the selected sample includes contracts for categories that are, on average, less durable, i.e., over-represent services. This sample selection does not affect the main results presented in Section 3. Finally, and consistent with the rest of the analysis, we estimate the model using contracts around the regulation threshold, i.e., between 10 and 40 thousand dollars.

[34] Given the structure of the model, we are assuming that the cost-overruns capture excess of the cost, which the vendor can entirely transmit to the buyer. Thus, there's no strategic behavior on behalf of the contractor nor the buyer ex-post. The fact that the observed contractor's overruns are drawn from group-specific distributions speaks about differences in cost prolixity and knowledge about buyer's context. We assume these differences are exogenous.

[35] Our estimation does not restrict the set of values for these parameters; however, in general, we should expect that buyers dislike paying higher prices or experiencing overruns, so we expect $\beta^P$ and $\beta^Q$ to be negative.

[36] Intuitively, the larger the utility shift, $\eta$, that results from from regulation, the higher will be the observed jump of publicity adoption at the threshold.

[37] The mean duration is calculated using contracts under $20,000 to remove the influence of the threshold.

**Estimation.** Our empirical model yields predictions about equilibrium conditions for suppliers' participation, bidding, and ex-post execution with and without publicity. Also, we characterize the buyer's publicity decision. Our estimation strategy proceeds using simulated method of moments (Mcfadden, 1989; Pakes and Pollard, 1989). That is, we choose a vector of parameters $\theta$ to generate simulation-based moments that closely resemble key moments from the data. Using parametrized primitives discussed previously, we simulate four empirical objects: participation decisions, bidding strategies, quality delivered and publicity decisions and create a set of moments conditions to be matched with data.

Our simulation procedure starts with a set of size $T$ of data inputs $(x_t, z_t, N_t)$, then from every observation we generate $S$ random draws of $u_t$. Finally, our setting contemplates that the buyer decides based on expectations, these expectations are formed conditional on $(x_t, z_t, N_t)$ and $u_t$, integrating over Monte Carlo simulated distributions of price, quality, and the likelihood of a local winning. This method, although computationally involved, is useful to circumvent integrating over potentially non-linear functions, and provides enough flexibility to match theoretical moments functions that cannot be evaluated directly.

Formally, denote the target moments by $m_n$ as a vector of moments from the data. The analogous moments generated by simulating observations are denoted by $m_s(\theta)$. Note that this vector depends on the parameters $\theta \in \Theta \subset \mathbb{R}^P$. The estimator minimizes the standard distance metric:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left(m_n - m_s(\theta)\right)' W_n \left(m_n - m_s(\theta)\right) \tag{6}$$

Where $W_n$ is the weighting matrix, which is chosen using the standard two-step approach; the quasi-optimal weight matrix $W_n$ is derived in the first stage, and the parameters are estimated in the second stage (Gourieroux, Monfort, and Renault, 1993). The vector of parameters corresponds to: $\theta = (\alpha^k, \nu^k, \tau^k, \gamma^k, \xi^k, \vec{\beta}, \zeta, \sigma)$.

We use three sets of target moments. The first set of moments are a vector of first and second-order moments of the relevant variables as well as it's interaction with it's relevant covariates. The relevant outcome variables are the auction price, the number of bidders, local winner, the magnitude of cost-overruns, an indicator of any cost-overrun, and publicity choices. The second set of moments consist of means of the same outcome variables conditional on partitions of the domain of contract prices. These moments capture the relation between these outcome variables over the domain of prices and are estimated separately for goods and services. Finally, the third set of moments are a vector of normalized frequencies on the relevant window of contract prices. Stacking together these three vectors, we obtain the vector $m_n$ of 357 moments that seek to match with the model. We use the stochastic optimization algorithm *Differential Evolution* (Storn and Price, 1997) to perform the objective minimization.[38] The details of the estimation procedure are discussed in the Appendix E.

[38]This algorithm performs a (parallel) direct search approach; it does not rely on gradient methods for minimizing possibly nonlinear and non-differentiable continuous space functions.

## 4.5 Estimation Results

Estimation of the model proceeds in two steps. In the first step, we estimate the model's parameters of entry, bidding, execution and publicity selection. We then combine the estimates with model equilibrium conditions to recover the primitive distribution of production and entry costs for locals and non-locals. These estimates are inputs to the policy counterfactuals in section 5.

### 4.5.1 Estimates

To facilitate the interpretation of coefficients, table 2 shows the marginal effects for the set of coefficients associated with the bidders and the buyers. The Appendix table B.8 displays our estimates with the corresponding standard errors.[39]

Several findings are worth highlighting. First, bidders are less prone to participate if the contract involves a service or a relatively complex product. Thus, auctions for these types of products are less competitive. In line with the evidence presented in Section 3.5, non-local contractors are 72 p.p. more likely to participate than locals. This is consistent with the fact that bidders' reduce the probability of entering if they observe more potential non-local competitors; one additional non-local reduces the probability of participating by 7.4 p.p.

Second, bids from non-locals are slightly lower than locals; they bid 4 p.p. lower prices. Another relevant feature is that unobserved heterogeneity is important in our data. Most of the variation in bidding is explained by common factors instead of variation between bidders within auction. The standard deviation of (log) unobserved heterogeneity is 27 times larger than the bids' standard deviation when $log(u_t) = 0$.

Third, the quality shock depends on the transaction product; the mean of log-quality shocks is substantially higher for services and complex product categories. In line with reduced form results, the difference between locals and non-locals in execution quality is substantial; non-locals have a mean shock that is 23 p.p. higher. Interestingly, the difference between these two groups is relatively stable over different degrees of product complexity.

Finally, as discussed in previous sections, contract publicity allows non-local bidders to participate in auctions, which leads to different contract outcomes. Panel C of Table 2 shows that the buyers choose to publicize 2.4 p.p. more if they anticipate that publicity leads to a 10% reduction in awarding price. A 10% increase in cost overruns reduces the probability of advertising by 1 p.p. Buyers have a preference for local vendors; if they anticipate a 10 p.p. reduction in the likelihood of a local winning, buyers reduce the probability of publicity by 2.3 p.p. Finally, predicting that the price without advertising exceeds $25,000 increases the likelihood of publicity by 32 p.p. The latter is in line with the increase in probability estimated in Section 3. These estimates depart from the standard assumption that the buyer *only* aims to minimize price and provides valuable inputs to

[39] Although the model is estimated altogether, tables B.8 and 2 present estimates in different columns to facilitate visual interpretation. The procedure to estimate standard errors is discussed in Appendix E.2.1.

Table 2: Marginal Effects Model Estimates

| | | | Entry | Bidding | Execution | |
|---|---|---|---|---|---|---|
| | $\bar{x}$ | $\Delta x$ | $\Delta\varphi/\Delta x$ | $\Delta\tilde{b}/\Delta x$ | $\Delta q/\Delta x$ | $\Delta(q>0)/\Delta x$ |
| *Panel A: Marginal Effects* | | | | | | |
| Service | 0.38 | 1 | -0.024 | 0.000 | 0.075 | 0.013 |
| Degree of Complexity | 0.09 | 0.1 | -0.028 | 0.000 | 0.163 | 0.030 |
| Non-Local | | 1 | 0.721 | -0.040 | 0.235 | 0.039 |
| Non-Local×Complexity | | 0.1 | 0.001 | -0.003 | 0.009 | 0.001 |
| Last Month | 0.25 | 1 | -0.333 | | | |
| Exp. Duration > Median | 0.50 | 1 | | | 0.127 | 0.021 |
| $N^L$ | 6.08 | 1 | 0.000 | -0.002 | | |
| $N^{NL}$ | 3.34 | 1 | -0.074 | 0.010 | | |
| *Panel B: Standard Deviation* | | | | | | |
| Estimated ($\hat{\sigma}$) | | | | 0.076 | 1.035 | |
| Unob. Het. ($\sigma_u$) | | | | 2.168 | | |
| | | | Publicity | | | |
| *Panel C: Marginal Effects* | | $\Delta x$ | $\Delta Pub/\Delta x$ | | | |
| Exp. Price | | 0.1 | -0.0243 | | | |
| Exp. Cost-Overruns | | 0.1 | -0.0095 | | | |
| Exp. Local Winning | | 0.1 | 0.0228 | | | |
| Above $25K | | 1 | 0.3263 | | | |
| Number of Observations | 24,135 | | | | | |

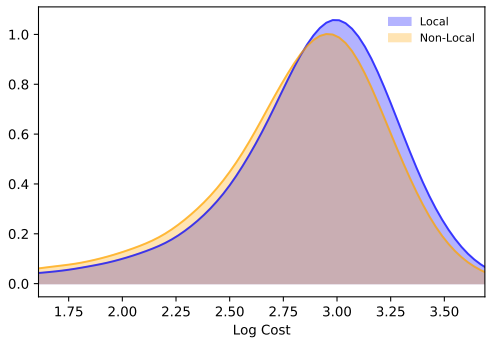
Notes: Panel A shows the marginal effects of main on different dependent variables related to bidders' actions. The Marginal Effects are computed at the mean of each covariate described in the second column. The third column shows the change used to estimate the effect. The dependent variables are the probability of entry, the bid level, and the quality shock in terms of levels and the probability of over zero. Panel B shows the empirical models' estimated standard deviation and the estimated standard deviation of the unobserved heterogeneity component. Panel C displays the marginal effects of expected price, cost-overruns, local winning, and being above $25K on the probability of publicizing the contract solicitation in *FBO.gov*. These coefficients are jointly estimated using Simulated Method of Moments (SMM).

evaluate policy counterfactuals.

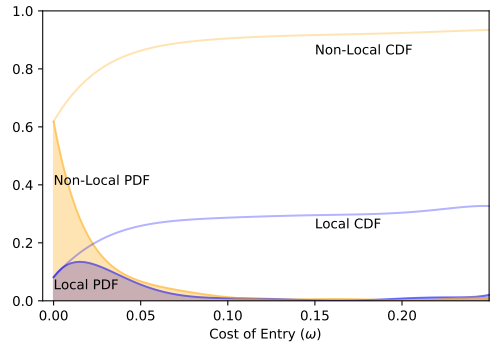
**Model Fit.** Overall, the model closely replicates the key empirical patterns in the estimation sample. We examine model fit using the estimated parameters to simulate equilibrium outcomes and compare simulated to observed outcomes. Our simulations, and the ones discussed later, build upon the estimating dataset by drawing simulations of the unobserved heterogeneity and the quality, entry, and bidding-cost shocks. These draws, combined with estimated parameters, allow us to simulate market-level equilibrium. Figure A9, in the Appendix, compares the the distribution of model-simulated outcome variables with actual data. The simulated data replicates closely publicity choices, actual bidders, and the share of contracts assigned to locals. Panels (e) and (f) separate cost-overruns by products and services. We find that, for services, the model slightly underpredicts the probability of having any (positive) cost-overrun but overpredicts the magnitude of cost-overruns. This dichotomy suggests that buyers' may face frictions when

Figure 7: Estimated Cost Distributions

(a) Production Cost Distribution




(b) Entry Cost Distributions




Notes: The panel (a) shows the PDF of log production costs for locals and non-locals. The panel (b) shows the estimated entry cost probability and cumulative density distributions. The CDFs are estimated comparing the entry probability to different levels of predicted (log) utilities. The PDFs capture the differences of the CDFs within bin widths of size 0.025. Both plots are estimated under average covariates and $log(u) = 0$. The plotted distribution of log costs in panel (a) is smoothed using a kernel. The raw distributions of entry costs in panel (b) are smoothed with a 7-degree polynomial fit. By nature, zero is the minimum expected utility of entering.

introducing contract modifications ex-post that our model does not account for.

### 4.5.2 Recovering the Project Cost Distribution

We recover the distribution of project costs leveraging the methodology introduced by Guerre, Perrigne, and Vuong (2000), and Campo, Perrigne, and Vuong (2003). This method combines the first-order conditions (equation 4) —subject to boundary conditions— with estimated equilibrium bids to estimate the inverse bid function. In our setting, the actual number of bidders is unknown, the first-order condition depends on the probabilities of different combinations of local and non-local bidders. These probabilities are formed from simulations based on model's participation parameters. Finally, strict monotonicity between the bid and the inverse bid functions enables us to to obtain an estimate of project costs from estimated distribution of bids.

Figure 7(a) shows the probability density function of log costs ($log(\tilde{c}_{jt})$) of both groups. Local bidders have slighly higher costs than non-locals. Figure A11 displays the mean $log(\tilde{b}_j(\tilde{c}))$ as a function of the log cost. As expected, markups decrease with higher the cost draws.

### 4.5.3 Recovering Entry Costs

We recover the group-specific entry-costs using the equilibrium conditions for optimal entry behavior discussed in section 4.1.2. A potential bidder compares the draw from the entry-cost distribution $G^k_\omega$ with the expected utility of entering, i.e., $\varphi^k(x_t, z_t, N_t) = G^k_\omega\left(\bar{\pi}^k(x_t, N_t, z_t)\right)$. Our estimated cost distributions $F^k_{\tilde{c}}(c)$ allow us to estimate the (*ex-ante*) predicted utility of participating (equation 5) and benchmark it to observed entry behavior (Athey, Levin, and Seira, 2011).

Figure 7(b) displays estimated entry cost distributions. It shows that the distributions of the two groups differ substantially. On the one hand, roughly 60% of non-locals face zero entry cost (i.e., they enter with probability one), and 90% enter if the entry cost is less than 0.1 log units. On the other hand, local contractors face substantially higher entry costs. Indeed, with a 60% chance, they would *not* enter for any of the values included in the estimated range of existing expected utilities. The estimated entry-cost asymmetry shapes the composition of actual bidders and, subsequently, the winning bid conditions.

Figure A10 in the Appendix shows how the composition of actual competitors (and the identity of the winner) depends on the number of potential non-locals. It shows how the number of actual bidders decreases as the number of potential non-locals increases. This is consistent with the fact that reductions on predicted utility due to increased competition discourage locals' participation, making it substantially more likely that a non-local wins.

### 4.6 Effects of Increasing Competition through Publicity

Having estimated the primitives of the model as a function of observable characteristics, we can replicate the policy evaluation discussed in section 3 and, evaluate contract's outcomes with and without publicity, not only for contracts around the threshold, but throughout the values included in our sample.

Figure 8 displays the variation on contract outcomes around the regulation threshold as a function of the expected price without publicity and simulate the results in a setting without publicity thresholds. These results are in line with our reduced-form analysis discussed previously. Publicizing contract solicitations allows the participation of non-local bidders, which are more prone to experience overruns and have substantially lower participation costs discouraging locals' participation. Thus, enhancing contract participation through publicity reduces prices ex-ante due to increased competition; however, it increases prices *ex-post*. We propose the following definition of the final price that takes into account both effects:

$$p^D_{F,t} = p^D_{I,t}(1 + q^D_t)$$

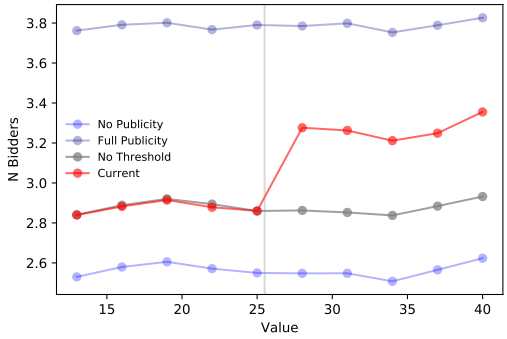
where $D_t \in \{0, 1\}$ denotes contract $t$'s publicity status, $p^D_{I,t}$ is the log awarding price, and $q^D_t$ is the realized share of of cost-overruns *ex-post*. Thus, $p^D_{F,t}$ denotes contract $t$'s log final price.

Figure 9, compares the consequences of publicizing contracts at different levels of complexity. Auctions for complex contracts have a higher variance of bid functions and face lower participation levels. The former increases the support of possible price reductions by added bidders. Adding one bidder to an auction with lower participation levels has more effect than adding one on an auction that already has many competitors. Thus, these two market features contribute to more extensive effects of added competition on award prices.

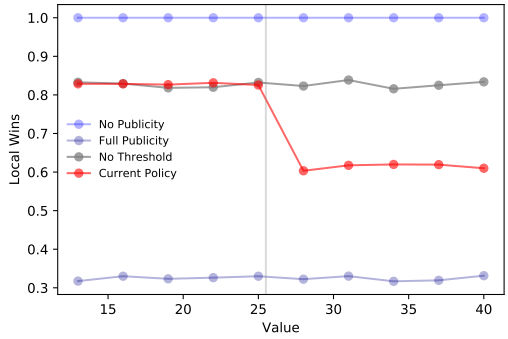
Our results show substantial asymmetry between local and non-local vendors when executing

Figure 8: Policy Evaluation Using Model Parameters

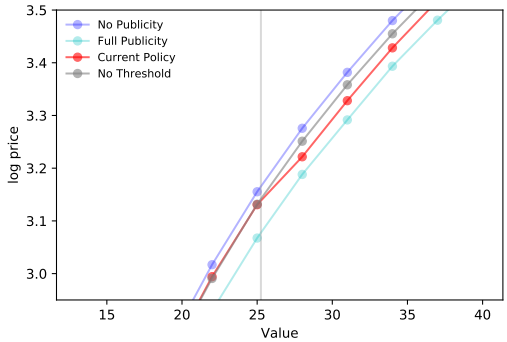
(a) Number of Bidders



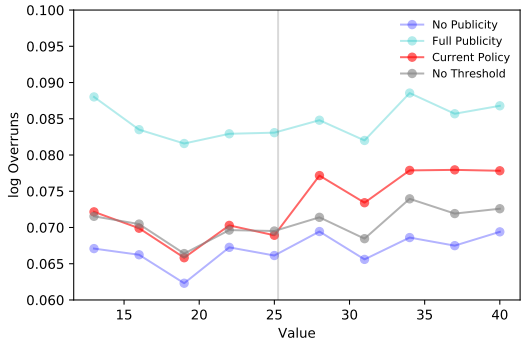

(b) Local Wins




(c) Log Prices




(d) Overruns

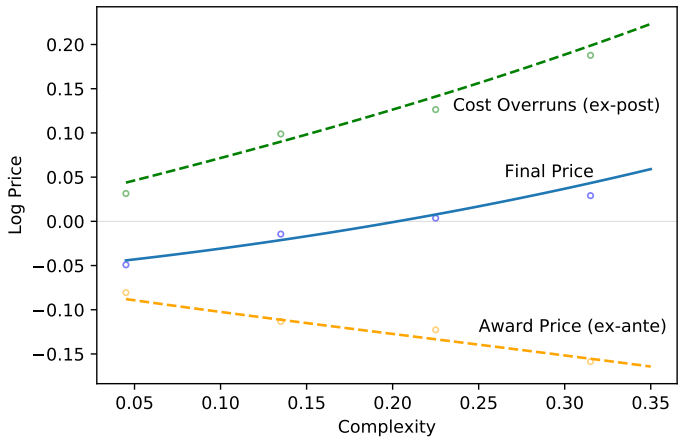



Notes: These figures show how different outcomes vary around the threshold. Panel (a) shows the number of bidders, panel (b) the probability of awarding a local, panel (c) the log of price, and panel (d) the log overruns. The x-axis of every graph is the value of the contract without publicity, which us the relevant dimension affected by the regulation threshold. As a reference, in every graph, we benchmark the current policy design (red line) with a situation of full (and zero) publicity. Also, we include the simulated event of no threshold as a counterfactual scenario.

contracts. Non-locals experience cost-overruns considerably more often than locals, i.e., contract publicity leads to higher cost-overruns ex-post. Therefore, the net of these opposing effects depends on the degree of complexity of the transaction; for relatively complex contracts, the increase in cost-overrun exceeds *ex-ante* price reductions, i.e., more competition leads to higher final contract prices. Conversely, there's little to no rise in cost-overruns associated with publicity for simple contracts. Thus publicizing contracts leads to reductions in final contract costs.

These findings align (and extend) our reduced form results and formalize the idea introduced by seminal papers on incomplete contracts: There exist a degree of transaction complexity beyond which promoting competition may backfire. When there's a high number of possible contingencies during the execution stage, assuring proper performance *ex-post* may be more important than reducing prices *ex-ante* through more competition. (Williamson, 1976; Bajari and Tadelis, 2001; Bajari et al., 2014; Bolotnyy and Vasserman, 2019).

Figure 9: Final Price by Product Complexity




Notes: This figure shows the effect of publicity on (log) of award price, on (log) of cost overruns, and (log) of the final price for different degrees of product complexity. The omitted category is the variable without publicity. The circles represent the mean effect per complexity bin. Each line corresponds to a flexible polynomial fit. The degree of complexity is defined as the log of the product's average overruns, and it is calculated on all contracts for the same product category below $20,000.

# 5 Counterfactual Analysis

We leverage estimated model parameters to evaluate the implications of counterfactual policies on award prices and ex-post cost overruns. Overall, there are two approaches to improving outcomes in a principal-agent setting; (1) delegating the publicity decision to the buyer relying on the buyer's knowledge of the local market and context, or (2) impose rules that restrict the agent's span of possible actions. Our counterfactual exercises aim to benchmark these two approaches. In particular, we evaluate what would be the set of actions that a buyer would take in a deregulated setting or as a result of alternative regulation designs.

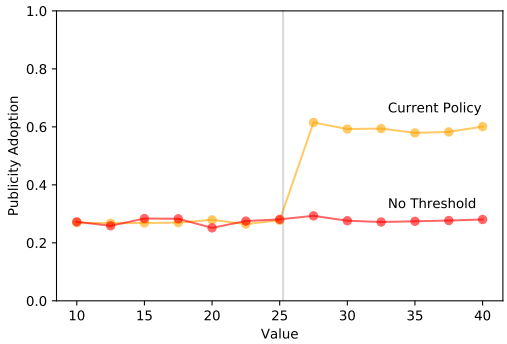
## 5.1 The Strategic Value of Delegating Competition Promotion to the Buyer

What are the implications of allowing the buyer to choose contracts' publicity status relative to tighter publicity rules that prescind the buyer's decision? This trade-off pertains to the more general problem of the delegation of authority within organizations (Aghion and Tirole, 1997), and it's a frequent theme of debate in procurement policy discussion (Kelman, 1990).

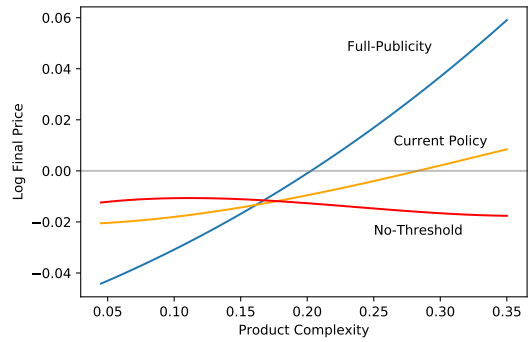
Conceptually, the publicity requirement design works as a discontinuous jump in the cost of not publicizing; below the threshold, the cost of not advertising is zero, whereas above the threshold is positive and involves filing additional paperwork. Therefore, below the threshold, buyers publicize their desired publicity with full discretion, while above the threshold, they are forced to advertise more than desired (due to the added cost of not publicizing). Using the estimated model parameters, we back out the buyer's hypothetical decisions in a full discretion setting and evaluate

Figure 10: Final Price by Product Complexity

(a) Publicity Adoption




(b) Log Final Price Effect




Notes: The panel (a) shows the share of publicized contracts depending on the value of the contracts without publicity. The orange line displays the current policy with a regulation threshold at $25,000. The red line describes the share of adoption in the absence of the regulation threshold. The panel (b) shows the effect on the log of the final price $(p^F_{m,t} - p^F_{0,t})$ for different degrees of product complexity. The blue line is the effect if all contracts are publicized, the orange line is the effect of the current policy (with a threshold at 25,000), the red line shows the effect in the absence of regulation threshold. The omitted category is the log of the final price in the absence of publicity. Each line was constructed using flexible polynomial fits. The degree of complexity is defined as the log of the product's average overruns, and it is calculated on all contracts for the same product category that are below the regulation threshold ($25,000).

its implications relative to the current policy design.

Figure 10 evaluates the implications of current publicity requirements for different levels of contract complexity. Panel (a) shows how the fraction of publicized contracts increases at the threshold relative to a counterfactual scenario without the requirement. Panel (b) illustrates how the price-complexity relation shifts as a result of increased publicity.

The effectiveness of allowing buyers to decide whether or not to publicize contracts depends on the product's complexity. Buyers with full discretion obtain lower prices than what they would get with no publicity over all the complexity spectrum. However, if contract complexity is lower than 0.18, a full publicity rule achieves lower average prices. Hence, providing discretion effectively reduces costs if the transaction unit is complex, whereas forcing publicity is more effective if the unit of transaction is relatively simple.

The current regulation combines these two scenarios; it provides discretion below the threshold and nudges higher levels of publicity above the threshold. As a result, relative to a regulation-free setting, the current regulation involves higher levels of advertising, which is effective when the purchase is simple; however, it backfires when the transaction is complex. In section 5, we discuss the effects of alternative policy tools to improve contract outcomes.

### 5.1.1 The Role of Buyer's Preferences

In our setting, the buyer decides the degree of contract competition motivated by interests that are not necessarily the same as that of the organization. The agency problem hinges on the degree

of *misalignment* between the buyer's (agent) and the organization (principal) objectives. Thus, the issue disappears if their objectives coincide since the buyer, by maximizing her utility, would be, too, maximizing the organization's welfare. An extensive theoretical literature studies the design of compensation mechanisms to "align" agent's objectives (Laffont and Tirole, 1993). This literature often builds upon rationality and completeness of contract menus to improve outcomes.

We study the extent to which contract outcomes depend on specific preference parameters. To do so, we leverage the estimated parameters, varying the degree of "alignment" on the buyer's preferences. We set two benchmark situations: First, the buyer has "Cost-Oriented Preferences", i.e., puts equal weight on price reductions ex-ante and ex-post and has no idiosyncratic preference for local contractors. Second, the buyer has "Local-Oriented Preferences", i.e., preferences oriented to favor local contractors with no emphasis on costs. The specific preference parameters under each scenario are described in Table 3. It is worth mentioning that these two benchmark preference parameters are based on the estimated coefficients –"shutting down" specific taste parameters. Therefore, they should be seen as reference points of policies oriented to affect buyers' motives.

Table 3: Buyer's Preferences

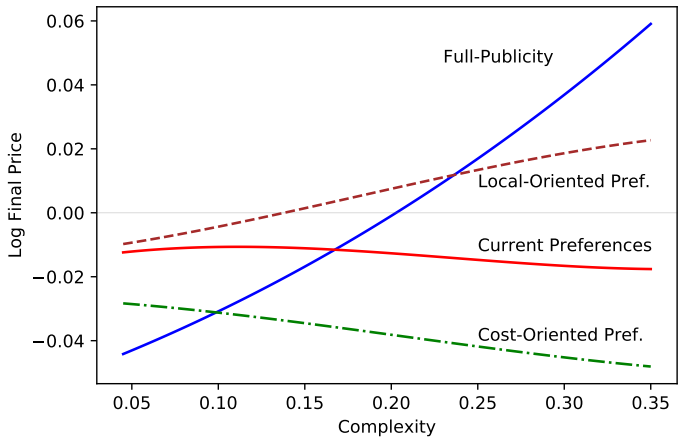
| | Estimated | Benchmarks | |
|---|---|---|---|
| | Preference Parameters | Cost-Oriented Preference | Local-Oriented Preference |
| $\beta^P$ | -0.636 | -0.636 | 0 |
| $\beta^Q$ | -0.245 | -0.636 | 0 |
| $\beta^L$ | 0.588 | 0 | 0.588 |
| Mean Pub. | 0.274 | 0.408 | 0.263 |

Notes: The first column shows the estimated preference parameters for the price ex-ante, overruns ex-post and awarding local contractors. The second column shows the preference parameters associated with a buyer with cost-oriented preferences, i.e., no idiosyncratic preference for locals. The third column shows the preference parameters for a buyer that is fully oriented to locals, i.e., does not have a preference for price or quality but only for favoring local contractors. The last row, 'Mean Publicity,' describes the average adoption of publicity under each of these types of preferences.

Figure 11 shows (log) final price effects depending on the level of complexity of the purchase. Buyers with "Cost-Oriented Preferences" decide to publicize to reduce costs, so, and perhaps not surprisingly, relative to other preference schemes, they generate savings all across the spectrum of product complexity. "Local-Oriented" agents seek to benefit local contractors; they publicize infrequently, and, as a result, the outcome prices are worsen than the situation without any publicity for most of the complexity spectrum. Interestingly, full publicity achieves higher prices than "Local-Oriented" buyers when contracts are sufficiently complex because *favored* local contractors tend to be better at executing these contracts, reducing cost-overruns.

Existing literature on rules vs. discretion (Aghion and Tirole, 1997; Carril, 2019; Bosio, Djankov, Glaeser, and Shleifer, 2020) emphasizes that regulation can be an effective antidote to waste and abuse whenever these are pervasive. Still, it can backfire if agents are virtuous in exercising discretion. Our findings contribute to the existing literature by highlighting that this trade-off

Figure 11: Final Price by Product Complexity




Notes: This figure shows the effect of publicity on the log of the final price for different preference parameters. The blue lines represent the price effect if all contracts are publicized. The red line is the price effect under current preference parameters. The green and brown dashed lines represent the effects of cost-oriented and local-oriented buyers' preferences, respectively. All lines are flexible polynomial fits. The degree of complexity is defined as the log of the product's average overruns, and it is calculated on all contracts for the same product category that are below the regulation threshold ($25,000).

depends, too, on the level of contract complexity: Publicity regulation can be detrimental even when agents are misaligned, as favoring local vendors has positive spill-over on cost-overruns. Moreover, we find that there is room for improving outcomes when buyers are aligned through strict publicity requirements when the transaction unit is simple.

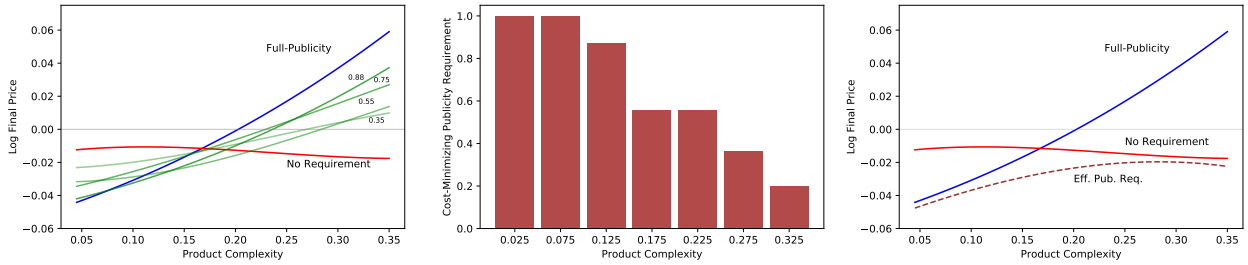
## 5.2 Complexity-Based Publicity Requirements

We now follow a more tactical approach and take the estimated buyer's preferences as given and vary the regulation design because the effects of publicizing depend on the level of contract complexity. The proposed design contemplates identifying the "right level" of publicity requirements depending on the degree of contract complexity. In this exercise, we refer to publicity requirements as the minimum fraction of contracts that buyers must publicize. We proceed in three steps; first, we simulate contract outcomes under different levels of product-specific publicity requirements that buyers are mandated to meet. Second, we estimate the final price under each of these requirements. Finally, we identify the publicity requirement that yields the lowest price at each complexity level. As a result, the proposed regulation imposes publicity requirements that are specific to each product category depending on the complexity level.[40]

[40]The proposed design involves a higher degree of regulation sophistication as it requires fixed fractions of publicized contracts per complexity level. We believe that can be implemented smoothly given the set of rules included in the Federal Acquisition Regulation. As a reference, the current version of chapter 5 (Publicity Requirements) allows buyers to apply for exemptions if they prefer not to publicize a contract. The proposed policy design could be implemented by simply varying the set of exemptions that different product categories are allowed to invoke. For example, if the contract solicitation involves a simple product, whose optimal level of publicity requirement is 100%, then there would be no exemption to be invoked. Conversely, if the solicitation requires a relatively complex product, the buyer could have more

Figure 12: Final Price by Product Complexity

(a) Publicity Adoption (b) Log Final Price Effect (c) Log Final Price Effect




Notes: The panel (a) shows the log of price effect with full publicity (blue line), without regulation requirements (red line), and for different levels of publicity requirements (green lines). The panel (b) describes the publicity requirement that yields the minimum price for varying levels of complexity. Overall, the publicity requirements that minimize costs decrease with product complexity. The panel (c) shows the price effect when imposing efficient publicity requirements at different levels of complexity. As a result, the price effect of this policy (brown-dashed line) corresponds to the lower contour of the panel (a). The omitted category is the log of the final price in the absence of publicity. Each line was constructed using flexible polynomial fits. The degree of complexity is defined as the log of the product's average overruns, and it is calculated on all contracts for the same product category that are below the regulation threshold ($25,000).

Figure 12 summarizes this procedure: Panel (a) shows the price-complexity relation at different publicity requirements. Panel (b) illustrates the publicity requirement that minimizes price at different complexity levels. Panel (c) shows the price-complexity relation that would stem from a "tailored" publicity requirements. Note that the latter (brown-dashed line in panel (c)) corresponds to the lower contour of final prices at different requirements (panel (a)). These tailored publicity requirements alter the span of the buyer's actions. In particular, when the unit of purchase is simple, it removes the buyer's discretion entirely to leverage the benefits of enhanced competition; however, it provides more choice when contracts are more complex to attenuate the negative consequences on contract implementation ex-post.

## 5.3 Comparing Policy Counterfactuals

Table 4 brings together the visual evidence provided Figures 10, 11 and 12, and compares the overall mean price effect under each of these scenarios. The current policy design that introduces publicity requirements at $ 25,000, reduces, on average, the final price by 1.6% relative to a no-publicity scenario. If the publicity choice were delegated to the buyer, this reduction would be, on average, 1.3%. A uniform full-publicity rule would reduce contract costs by 3%. However, the latter could be improved if the publicity requirements were tailored to the purchase's degree of complexity. The latter policy reduces prices by 3.6%. The 2% cost difference between the current and the complexity-based design corresponds to $104 million — competitively awarded— defense contracts, annually.[41]

Finally, we benchmark the consequences of these policy designs with the hypothetical situation

(or total) discretion to file exemptions.

[41] This amount is calculated using defense contracts competitively awarded in 2018 with values between $10,000 and $150,000 (Simplified Acquisition Threshold).

Table 4: Effects of Counterfactual Scenarios

| | Mean | 95 Perc C. I. |
|---|---|---|
| Current Policy | -0.0164 | [ -0.0195 , -0.0133 ] |
| *Delegate the Decision to the Buyer* | | |
| Current Preferences | -0.0136 | [ -0.0161 , -0.0111 ] |
| Buyer with Price-Oriented Pref. | -0.0304 | [ -0.0335 , -0.0271 ] |
| Buyer with Local-Oriented Pref. | -0.0068 | [ -0.0092 , -0.0043 ] |
| *Alternative Regulation Designs* | | |
| Full Publicity | -0.0306 | [ -0.0357 , -0.0256 ] |
| Complexity-Based Publicity Req. | -0.0367 | [ -0.0412 , -0.0323 ] |

Notes: The table reports the mean equilibrium effects to log final prices under different scenarios. In the first row, describes the estimated effect under the current policy, with threshold that increases publicity at $ 25,000. In the second row, the mean effects without publicity requirements. The third and fourth row show the mean price reduction if under cost-oriented and local-oriented buyers, respectively. The fifth row shows the mean price reduction under full publicity. The sixth row shows the average price reduction implementing efficient publicity requirements depending on the level of complexity of the product. These effects are constructed relative to the situation without publicity. Confidence intervals are constructed via bootstrap.

in which the buyer has cost-minimizing preferences. We find that the cost-minimizing publicity requirement achieves better outcomes than Cost-Oriented Buyers. From a policy standpoint, this is significant as arguable modest improvements to the regulation design could mitigate most of the concerns associated with misaligned buyers and achieve, on average, better outcomes than any compensation mechanism that aims to align buyers' objectives.

# 6 Conclusion

This paper studies the relationship between competition and procurement contract outcomes. Even though procurement contracts represent a key component of the economy, there is minimal evidence of the implications of policies oriented to expand competition, considering not only the award price but also the quality of the contract execution. We provide extensive evidence of the effects of enhancing competition through publicity, using the U.S. Department of Defense contracting market as a setting.

Our identification strategy leverages a regulation that generates quasi-experimental variation in the extent to which contract opportunities are broadly advertised to potential suppliers. We find that contract publicity increases contract competition. The added competitive pressure results in lower acquisition prices; however, broader dissemination leads to a different pool of vendors, who perform worse ex-post. We further explore the implications of the key trade-off between price

reductions ex-ante and worsen contract execution ex-post. Our analysis shows that the degree of contract complexity determines the scope of this trade-off. Promoting competition reduces contract costs only for simple transactions, as relatively complex ones are exposed to cost overruns and delays in the execution stage.

Motivated by this evidence, we develop and estimate an equilibrium model of competition for procurement contracts, with two general objectives. This model allows us to estimate the underlying firms' characteristics that shape adverse selection in this market and buyers' objectives when promoting competition through advertising. Our estimates allow us to evaluate two relevant counterfactual policies. The first we assess the implications of delegating competition promotion to the buyer. The second aims to propose a welfare-enhancing regulation design that accounts for the vast heterogeneity in the degree of complexity of procurement transactions.

Our counterfactual analysis shows that delegating competition promotion to the buyer is only welfare-enhancing when the transaction unit is complex: on average, the buyer achieves better outcomes than in regulated settings with either zero or full publicity. However, when the transaction unit is relatively simple, imposing full-publicity rules is preferred as the risks at the execution stage are minor. Moreover, we use our model to engineer improvements to the current policy design by introducing publicity requirements tailored to the degree of complexity of the purchase. We find that departing from a uniform regulation will significantly reduce procurement costs. Notably, while our analysis is carried out using data from the Department of Defense, we believe that the general conclusions apply broadly to private and public organizations' transactions.

## References


Aghion, P. and J. Tirole (1997). Formal and real authority in organizations. *Journal of Political Economy 105*(1), 1–29.

Athey, S. (2001). Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica 69*(4), 861–889.

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

Athey, S., J. Levin, and E. Seira (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *Quarterly Journal of Economics 126*(1), 207–257.

Bajari, P. (2001). Comparing competition and collusion: A numerical approach. *Economic Theory 18*(1), 187–205.

Bajari, P., S. Houghton, and S. Tadelis (2014). Bidding for Incomplete Contracts: An Empirical Analysis. *American Economic Review 104*(4), 1288–1319.

Bajari, P., R. McMillan, and S. Tadelis (2009). Auctions versus negotiations in procurement: An empirical analysis. *Journal of Law, Economics, and Organization 25*(2), 372–399.

Bajari, P. and S. Tadelis (2001). Incentives versus Transaction Costs: A Theory of Procurement Contracts. *RAND Journal of Economics 32*(3), 387–407.

Bandiera, O., A. Prat, and T. Valletti (2009). Active and passive waste in government spending. *American Economic Review 99*(2009 Sept), 1278–1308.

Best, M. C., J. Hjort, and D. Szakonyi (2017). Individuals and Organizations as Sources of State Effectiveness, and Consequences for Policy.

Bhattacharya, V., J. W. Roberts, and A. Sweeting (2014). Regulating bidder participation in auctions. *The RAND Journal of Economics 45*(4), 675–704.

Bolotnyy, V. and S. Vasserman (2019). Scaling Auctions as Insurance: A Case Study in Infrastructure Procurement.

Bosio, E., S. Djankov, E. Glaeser, and A. Shleifer (2020). Public Procurement in Law and Practice.

Brendstrup, B. and H. J. Paarsch (2003). Nonparametric Estimation of Dutch and First-Price , Sealed-Bid Auction Models With Asymmetric Bidders. (319).

Bulow, J. and P. Klemperer (1996). Auctions Versus Negotiations. *American Economic Review*.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica 82*(6), 2295–2326.

Campo, S., I. Perrigne, and Q. Vuong (2003). Asymmetry in first-price auctions with affiliated private values. *Journal of Applied Econometrics 18*(2), 179–207.

Carril, R. (2019). Rules Versus Discretion in Public Procurement.

Carril, R. and A. Gonzalez-Lira (2021). Estimating regression discontinuity designs using realized assignment variables.

Chakraborty, I., F. Khalil, and J. Lawarree (2020). Competitive Procurement with Ex Post Moral Hazard. *RAND Journal of Economics*.

Chernozhukov, V., I. Fernandez-Val, and B. Melly (2013). Inference on Counterfactual Distributions. *Econometrica 81*(6), 2205–2268.

Chetty, R., J. Friedman, and E. Saez (2013). Using Differences in Knowledge Across Neighborhoods to Uncover the Impacts of the EITC on Earnings : Web Appendix Sample and Variable De ... nitions. *American Economic Review 103*(7), 2683–2721.

Chiappori, P. A. and B. Salanie (2002). Testing Contract Theory: A Survey of Some Recent Work.

Compiani, G., P. Haile, and M. Sant'Anna (2020). Common Values, Unobserved Heterogeneity, and Endogenous Entry in US Offshore Oil Lease Auctions. *Journal of Political Economy 128*(10), 3872–3912.

Coviello, D. and S. Gagliarducci (2017). Tenure in office and public procurement. *American Economic Journal: Economic Policy 9*(3), 59–105.

Coviello, D. and M. Mariniello (2014). Publicity requirements in public procurement: Evidence from a regression discontinuity design. *Journal of Public Economics 109*, 76–100.

Davezies, L. and T. Le Barbanchon (2017). Regression discontinuity design with continuous measurement error in the running variable. *Journal of Econometrics 200*(2), 260–281.

Decarolis, F. (2014). Awarding Price, Contract Performance, and Bids Screening: Evidence from Procurement Auctions. *American Economic Journal: Applied Economics 6*(1 A), 108–132.

Decarolis, F., L. M. Giuffrida, E. Iossa, V. Mollisi, and G. Spagnolo (2020). Bureaucratic Competence and Procurement Outcomes. *The Journal of Law, Economics, and Organization No. 24201*.

Espin-Sanchez, J.-A., A. Parra, and Y. Wang (2021). Equilibrium Uniqueness in Entry Games with Private Information.

Gerard, F., M. Rokkanen, and C. Rothe (2020). Bounds on treatment effects in regression discontinuity designs with a manipulated running variable. *Quantitative Economics 11*(3), 839–870.

Goldberg, V. P. (1977). Competitive Bidding and the Production of Precontract Information. *The Bell Journal of Economics 8*(1), 250–261.

Gourieroux, C., A. Monfort, and E. Renault (1993). Indirect inference. *Journal of Applied Econometrics 8*(1 S), S85–S118.

Guerre, E., I. Perrigne, and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica 68*(3), 525–574.

Haile, P. A. and Y. Kitamura (2019). Unobserved Heterogeneity in Auctions. *The Econometrics Journal 22*(1), C1–C19.

Hart, O. and J. Moore (1988). Incomplete Contracts and Renegotiation. *Econometrica 56*(4), 755–785.

Hong, H. and M. Shum (2002). Increasing Competition and the Winner ' s Curse : Evidence from Procurement. *Review of Economic Studies 69*(4), 871–898.

Jales, H. and Z. Yu (2017). Identification and estimation using a density discontinuity approach. *Advances in Econometrics 38*, 29–72.

Kang, K. and R. A. Miller (2017). Winning by Default: Why is there so Little COmpetition in Government Procurement?

Kelman, S. (1990). *Procurement and Public Management: The Fear of Discretion and the Quality of Government Performance*. AEI Press.

Kleven, H. (2016). Bunching. *Annual Review of Economics*.

Kleven, H. J. and M. Waseem (2013). Using Notches To Uncover Optimization Frictions And Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics 128*(December), 669–723.

Krasnokutskaya, E. (2011). Identification and Estimation of Auction Models with Unobserved Heterogeneity. *Review of Economic Studies 78*(1), 293–327.

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

Laffont, J.-J. and J. Tirole (1990). Adverse Selection and Renegotiation in Procurement. *The Review of Economic Studies 57*(4), 597.

Laffont, J.-J. and J. Tirole (1993). *A Theory of Incentives in Procurement and Regulation.* MIT press.

Lebrun, B. (1999). First price auctions in the asymmetric N bidder case. *International Economic Review 40*(1), 125–142.

Lee, D. S. and T. Lemieux (2010). Regression Discontinuity Design in Economics. *Journal of Economic Literature 20*(1), 281–355.

Levin, D. and J. L. Smith (1994). Equilibrium in Auctions with Entry. *American Economic Review 84*(3), 585–599.

Lewis, G. and P. Bajari (2011). Procurement Contracting with Time Incentives Theory and Evidence. *Quarterly Journal of Economics*.

Li, T. and X. Zheng (2009). Entry and competition effects in first-price auctions: Theory and evidence from procurement auctions. *Review of Economic Studies 76*(4), 1397–1429.

Li, T. and X. Zheng (2012). Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics 168*(1), 29–46.

Liebman, J. B. and N. Mahoney (2017). Do Expiring Budgets Lead to Wasteful Year-End Spending? Evidence from Federal Procurement Faculty Research Working Paper Series. *American Economic Review 107*(11), 3510–3549.

Mackay, A. (2018). Contract Duration and the Costs of Market Transactions.

Macleod, W. B. and J. M. Malcomson (1989). Implicit Contracts, Incentive Compatibility, and Involuntary Unemployment. *Econometrica*.

Malcomson, J. M. (2012). Relational incentive contracts. *The Handbook of Organizational Economics* (508), 1014–1065.

Marmer, V., A. Shneyerov, and P. Xu (2013). What model for entry in first-price auctions? A nonparametric approach. *Journal of Econometrics 176*(1), 46–58.

Maskin, E. and J. Riley (2003a). Asymmetric Auctions. *Review of Economic Studies 67*(3), 413–438.

Maskin, E. and J. Riley (2003b). Equilibrium in Sealed High Bid Auctions. *Review of Economic Studies 67*(3), 439–454.

Mcfadden, D. (1989). A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration. *Econometrica 57*(5), 995–1026.

Pakes, A. and D. Pollard (1989). Simulation and the Asymptotics of Optimization Estimators. *Econometrica 57*(5), 1027–1057.

Pei, Z. and Y. Shen (2017). The devil is in the tails: Regression discontinuity design with measurement error in the assignment variable. *Advances in Econometrics 38*(October), 455–502.

Porter, R. H. and J. D. Zona (1993). Detection of Bid Rigging in Procurement Auctions. *Journal of Political Economy 101*(3), 518–538.

Roberts, J. W. (2013). Unobserved heterogeneity and reserve prices in auctions. *RAND Journal of Economics 44*(4), 712–732.

Ryan, N. (2020). Contract Enforcement and Productive Efficiency: Evidence From the Bidding and Renegotiation of Power Contracts in India. *Econometrica 88*(2), 383–424.

Samuelson, W. F. (1985). Competitive Bidding with Entry Costs. *Economics Letters 17*, 53–57.

Schennach, S. M. (2020). *Mismeasured and unobserved variables*, Volume 7. Elsevier B.V.

Spulber, D. F. (1990). Auctions and Contract Enforcement. *The Journal of Law, Economics, and Organization 6*(2).

Storn, R. and K. Price (1997). Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces. *Journal of Global Optimization*, 341–359.

Sweeting, A. and V. Bhattacharya (2015). Selective Entry and Auction Design. *International Journal of Industrial Organization 43*, 189–207.

Szucs, F. (2020). Discretion and Corruption in Public Procurement.

Tirole, J. (1999). Incomplete Contracts: Where do we Stand? *Econometrica 67*(4), 741–781.

Williamson, O. E. (1976). Transaction-Cost Economics : The Governance of Contractual Relations. *Journal of Law & Economics 22*(2), 233–261.

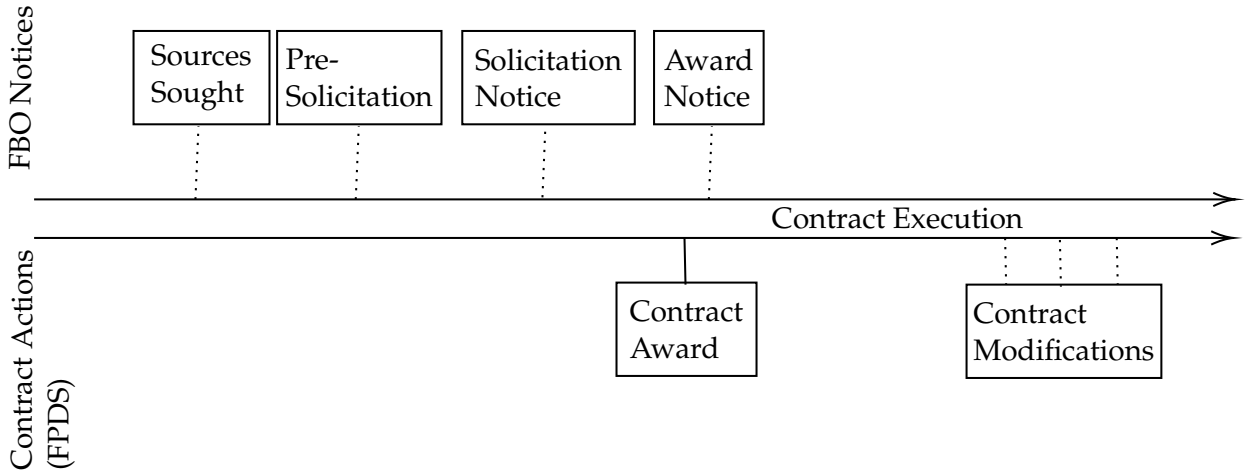
# Online Appendix

## Competition Under Incomplete Contracts and the Design of Procurement Policies

*Rodrigo Carril, Andres Gonzalez-Lira, and Michael S. Walker*

## A Additional Figures

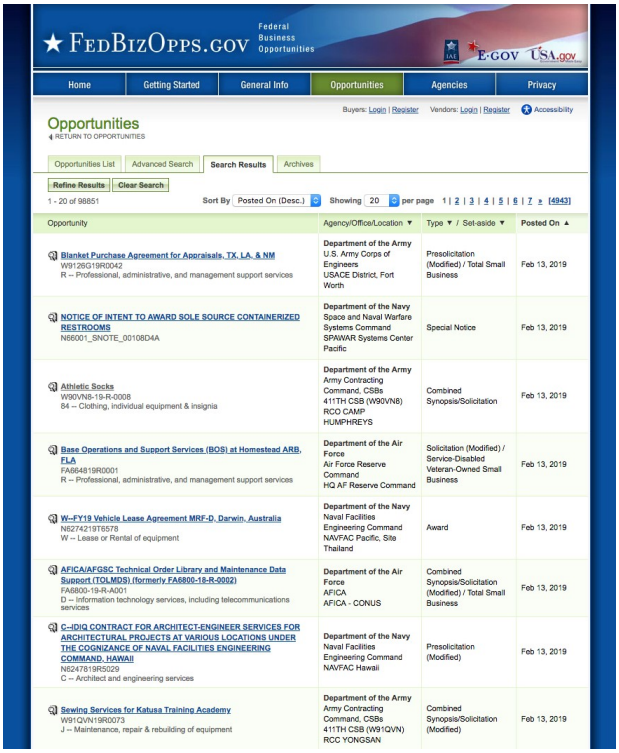
Figure A1: Contract Timeline and Data Sources



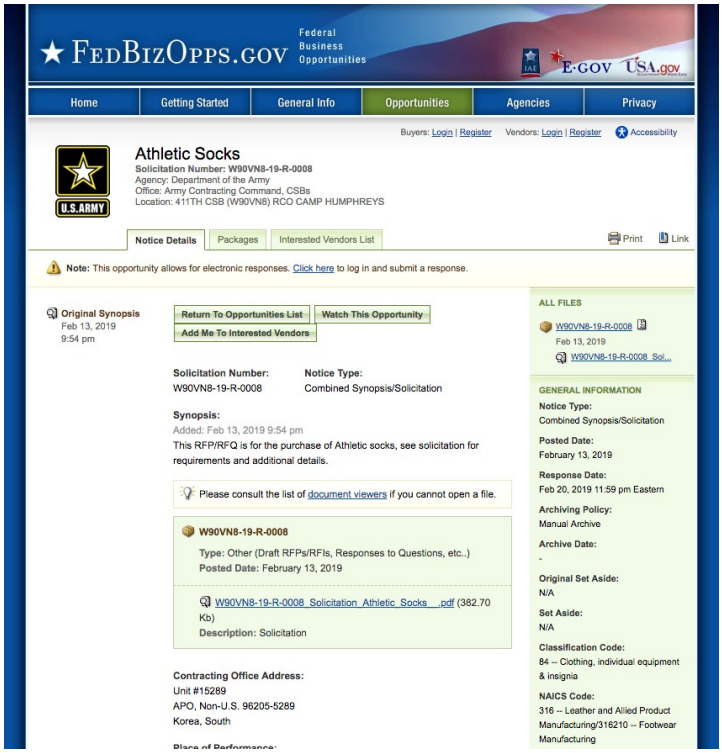

Notes: This figure presents a timeline of events associated with a typical contract. Milestones located above the arrows correspond to notices that are published on the government's point of entry (fedbizopps.gov). Milestones below the arrows generate information that is recorded on the Federal Procurement Data System (FPDS) - Next Generation.

Figure A2: FedBizzOpps

(a) List of Opportunities

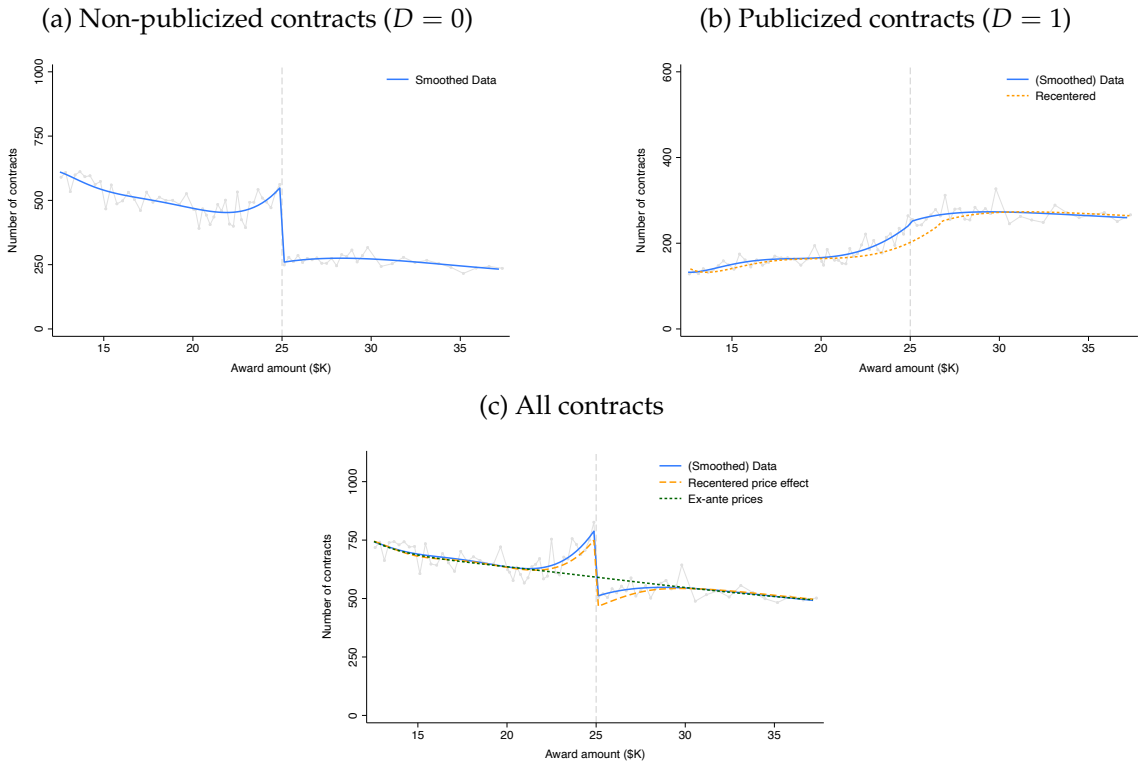



(b) Example Solicitation




Notes: These figures show screenshots to FBO.gov. Panel (a) the list of opportunities and search alternatives. Panel (b) shows a particular solicitation for athletic socks, required by an Army procurement office. These screenshots were captured on Feb 13, 2019.

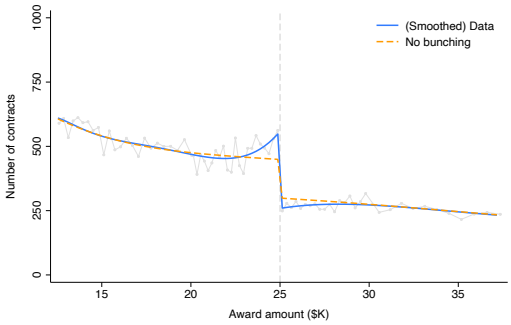
Figure A3: Estimating mean price effects and ex-ante prices



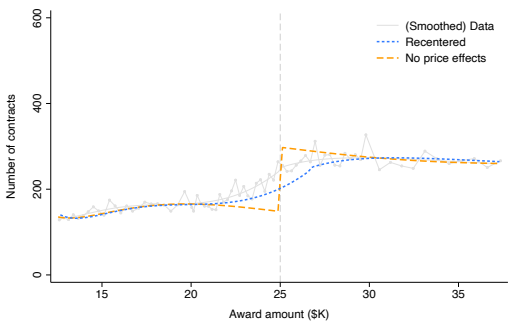

Notes: This figure shows the empirical distribution of the number of contracts at different price bins. Panel (a) shows the distribution of non-publicized contracts ($D = 0$). Panel (b) shows the distribution of publicized contracts ($D = 1$). Panel (c) displays the overall distribution, i.e., the sum of publicized and non-publicized contracts at every price. The blue line corresponds to a polynomial fit of degree five. The orange dashed lines in panels (b) and (c) represent the distribution of contract prices after re-centering publicized contracts by their price effect. The green dashed line in panel (c) represents the corresponding overall interpolation in the absence of price effects and bunching.

Figure A4: Estimating ex-ante prices

(a) Non-publicized contracts ($D = 0$)

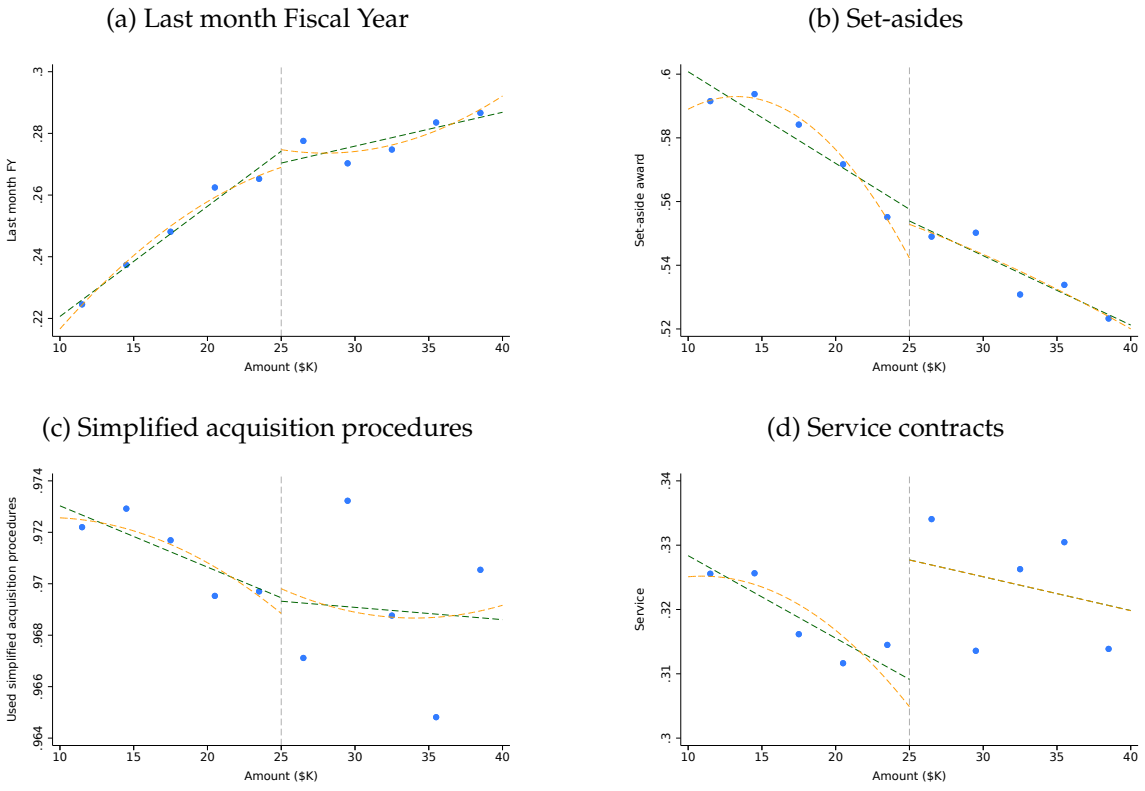



(b) Publicized contracts ($D = 1$)




Notes: This figure shows the empirical distribution of the number of contracts at different price bins. Panel (a) shows the distribution of non-publicized contracts ($D = 0$). Panel (b) displays the distribution of publicized contracts ($D = 1$). The blue line corresponds to a polynomial fit of degree five. The orange dashed lines in panels (b) and (c) represent the counterfactual distributions in the absence of price effects and bunching. The counterfactual distributions stem from the proposed framework. In panel (a), The comparison between the solid blue and the dashed orange lines provide a visual interpretation of the mass of bunched contracts. The comparison between the dashed blue and the dashed orange lines in panel (b) inform visually about the distribution of price effects.

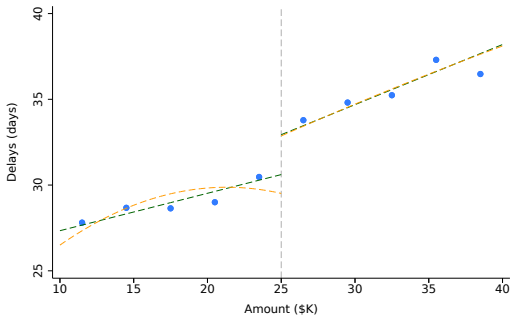
Figure A5: Pre-award characteristics around the threshold



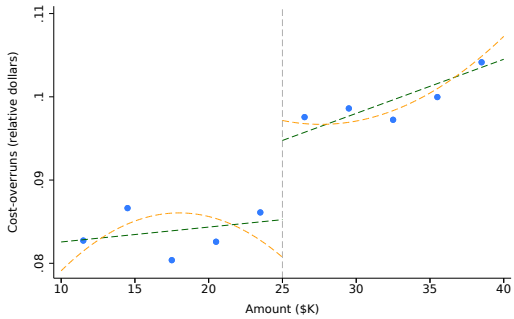

Notes: This figure presents four binned scatter plots, which depict an average pre-award characteristic by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The pre-award characteristic in each Panel is as follows: (a) an indicator equal to one if the contract was solicited the last month of the fiscal year (September); (b) an indicator equal to one if the contract was set-aside for a preferential group (e.g. small businesses); (c) an indicator equal to one if the contract was awarded using simplified acquisition procedures; (d) an indicator equal to one if the award is for a service contract. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

Figure A6: Publicity Effects on Post-Award Contract Performance

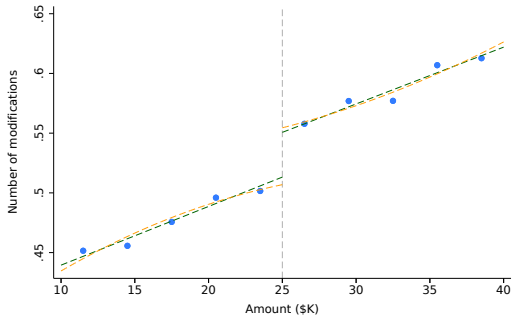
(a) Delays (days)




(b) Cost-Overruns (Share of Award Value)



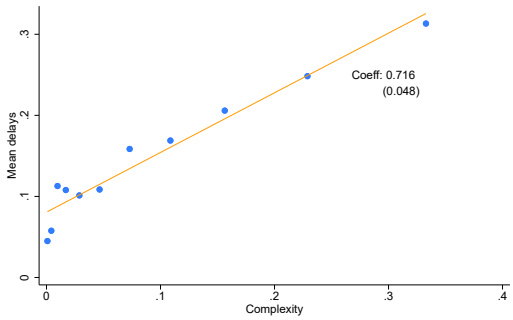

(c) Number of Contract Modifications (Ex-Post)



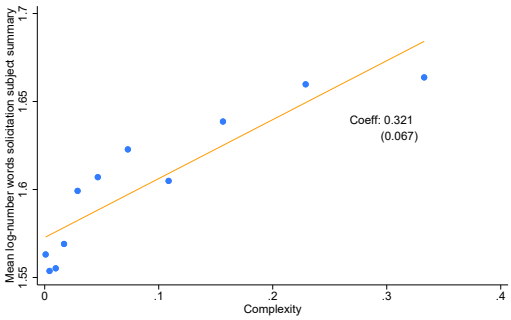

Notes: This figure presents four binned scatter plots, which depict an average post-award characteristic by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The pre-award characteristic in each Panel is as follows: (a) number of days of contract implementation delays; (b) cost-overruns as a share of award value; (c) number of modification to the original contract. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

Figure A7: Correlation Complexity Degree with Other Variables

(a) Delays (days)

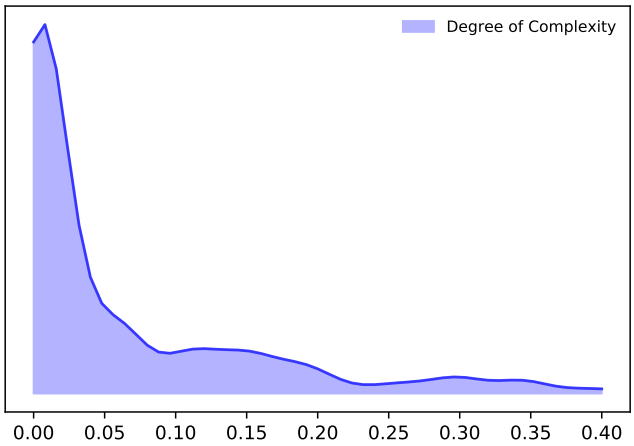



(b) Number of Words Solicitation



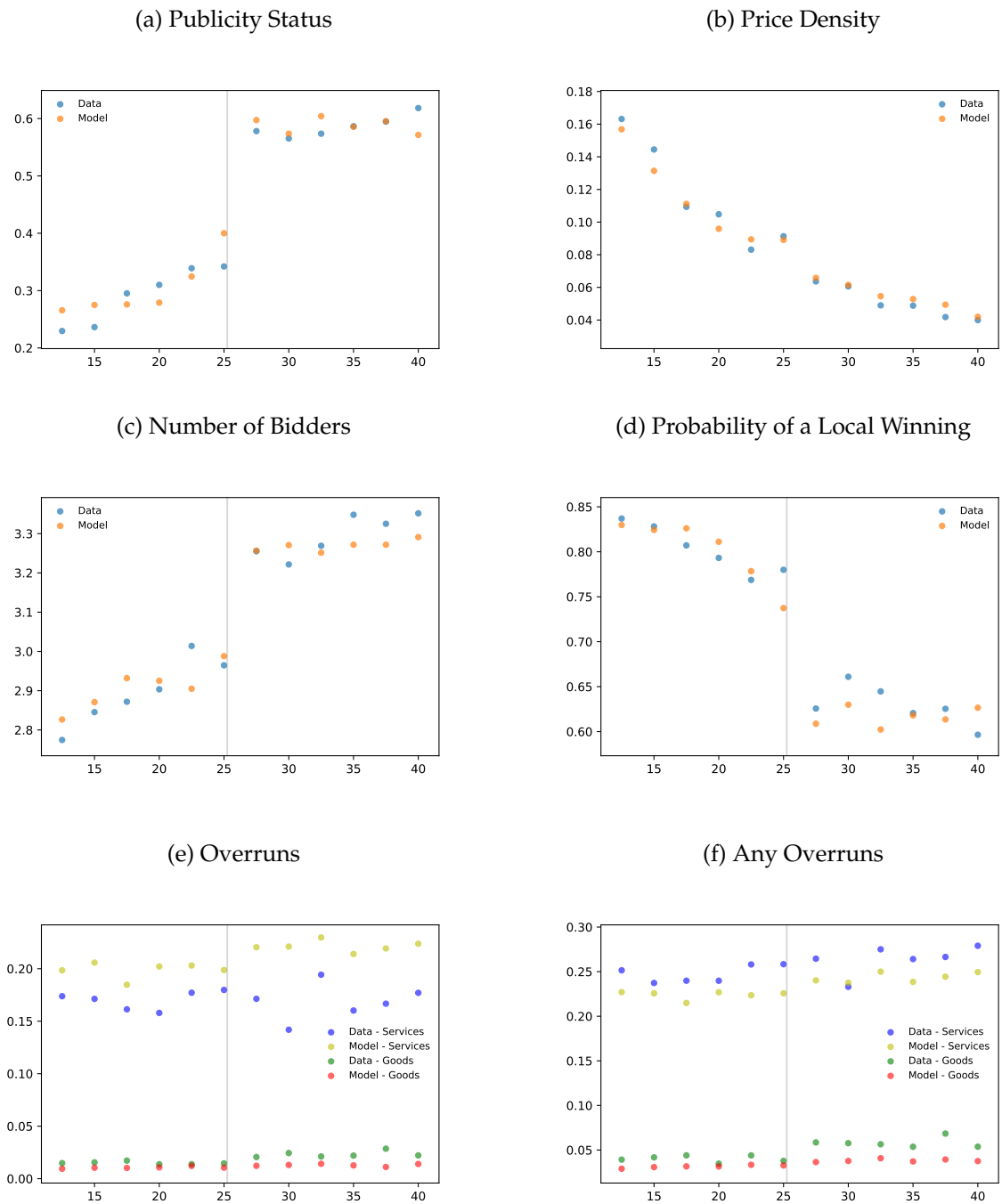

Notes: These graphs display the correlation between our measure of complexity (i.e., product-level average cost-overruns for contracts under $20K) with the product-level average delays (Panel (a)) and the product-level (log) the average number of words contract synopsis subject in FBO. The number of words variable was residualized from office, type of solicitation, and year fixed effects because the text often contains information specific to the office and the solicitation type. Every dot represents the mean of the Y-axis variable at a quantile. The orange line provides a (linear) regression fit at the product level. The slope coefficient (and SE) are presented in the graph.

Figure A8: Complexity Distribution




Notes: This figure presents the probability density function (PDF) of product complexity. Even though there's wide heterogeneity on the degree of complexity, the bulk of contracts in our sample have relatively low levels of complexity. The degree of complexity is defined as the log of the product's average overruns, and it is calculated on all contracts for the same product category that are smaller than the regulation threshold ($25,000). The plotted distribution of log costs is smoothed using a kernel.

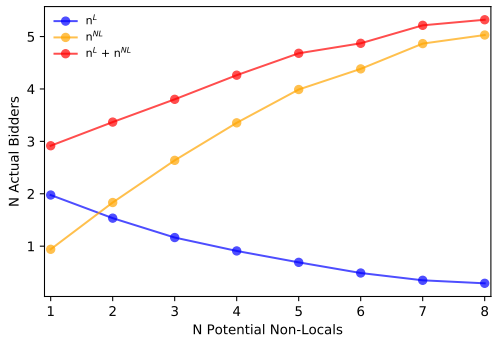
Figure A9: Model Fit



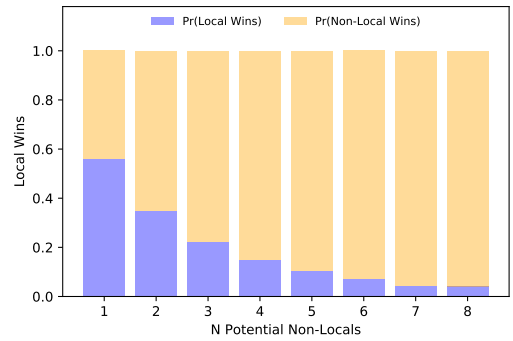

Notes: This figure presents the model fit, based on a simulated method of moments estimation. In each panel, relevant outcome variables relative to the awarding price. Actual data are presented in blue, while model-based simulated data is presented in orange. Panel (a) presents the density of contract prices, Panel (b) the fraction of publicized contracts, Panel (c) the number of actual bidders, Panel (d) fraction awarded to local contractors, Panel (e) and (f) show the average overrun and the chances of having any overrun. The last two panels separate goods from services. The simulated outcomes simulate unobservables building upon actual data. The model is estimated using 24,135 observations, the simulated methods expand each observation multiple times.

Figure A10: Participation

(a) Composition of Bidders

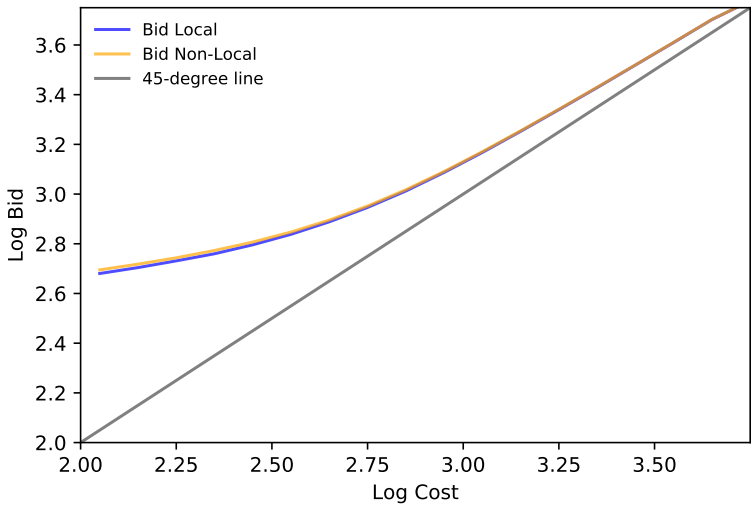



(b) Identity of the Winner




Notes: This figure presents the participation decisions and the subsequent winner identity as a function of the number of potential bidders. Panel (a) the number of actual bidders of each group, the panel (b) displays the average probability of awarding local bidders. The higher the number of potential non-locals, the less likely that locals participate and win. These features connect directly with the fact that locals have substantially higher participation costs; thus, in equilibrium, reductions on predicted utility due to increased competition discourage their participation. Both figures were generated keeping constant (at the mean) the number of potential locals.

Figure A11: Bidding Function




Notes: This figure displays the bidding function of local and non-local contractors. This plot is estimated under average covariates and $log(u) = 0$. The plotted distribution of log bids is smoothed using a kernel.

# B Additional Tables

Table B.1: Summary statistics

| | Mean |
|---|---|
| *Contract Characteristics* | |
| Award Amount | 20,807 |
| Expected Duration (days) | 54.10 |
| Fixed-Price Contract | 0.999 |
| Set Aside Award | 0.571 |
| Simplified Procedure | 0.971 |
| *Competition* | |
| Number of Offers | 3.542 |
| One Offer | 0.239 |
| *Contracting Office Characteristics* | |
| Navy | 0.378 |
| Army | 0.441 |
| Air Force | 0.150 |
| Other | 0.031 |
| *Awarded Firm Characteristics* | |
| Foreign | 0.099 |
| Within-State Firm | 0.690 |
| Small Business | 0.752 |
| Womam Owned Business | 0.188 |
| *Sample* | |
| No. of Contracts | 85,661 |
| No. of Contracting Offices | 597 |
| No. of Firms | 29,641 |

Notes: This table presents summary statistics. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. An observation is a contract, defined by aggregating all contract *actions* (initial award, modification, termination, etc.) associated with the same contract ID.

Table B.2: Top product and service categories

| | Goods | | Services | |
|---|---|---|---|---|
| Rank | Name | N Contracts/year | Name | N Contracts/year |
| 1 | ADP Equipment and Software | 3,005 | Maintenance/Repair of Equipment | 2,430 |
| 2 | Medical Equipment and Supplies | 2,998 | Support Services (Professional) | 1,187 |
| 3 | Laboratory Equipment | 1,643 | Utilities And Housekeeping | 1,096 |
| 4 | Electrical Equipment Compontents | 1,593 | Transport, Travel, Relocation | 854 |
| 5 | Communication/Coherent Radiation | 1,202 | ADP and Telecommunications | 806 |
| 6 | Furniture | 810 | Lease/Rent Equipment | 753 |
| 7 | Power Distribution Equipment | 697 | Maintenance of Real Property | 688 |
| 8 | Ship And Marine Equipment | 574 | Education And Training | 560 |
| 9 | Hardware And Abrasives | 530 | Construct Of Structures/Facilities | 335 |
| 10 | Construction And Building Material | 459 | Social Services | 286 |

Notes: This table presents average annual counts of contracts in the most common product categories. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. An observation is a contract, defined by aggregating all contract *actions* (initial award, modification, termination, etc.) associated with the same contract ID. A 4-digit alphanumeric code (PSC) is observed for each contract. The categories listed are constructed by aggregating PSC codes to two-digits for goods, and to a single digit (letter) for services.

Table B.3: Estimated Price Effect

| Estimate / Sample | All | Goods | Services | Complexity | | | |
|---|---|---|---|---|---|---|---|
| | | | | Q1 | Q2 | Q3 | Q4 |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Mean ($\mu_\gamma$) | 0.0595 | 0.0498 | 0.0782 | 0.0397 | 0.0505 | 0.0510 | 0.0962 |
| | (0.0201) | (0.0622) | (0.0596) | (0.0475) | (0.1692) | (0.1908) | (0.0920) |
| Standard Deviation ($\sigma_\gamma$) | 0.0643 | 0.0670 | 0.0534 | 0.0669 | 0.0739 | 0.0680 | 0.0369 |
| | (0.0075) | (0.0084) | (0.0202) | (0.0140) | (0.0760) | (0.0295) | (0.0280) |

Notes: This table shows the estimates corresponding to the effect of publicity on contract prices. The estimates result from analyzing the observed contract price density distribution relative to a counterfactual distribution. The observed densities are generated using bins of width $250. The counterfactual distribution stems from a polynomial interpolation of degree 5. The standard deviation is calculated over the non-parametric distribution of $\gamma$. The standard errors are calculated through bootstrap. The subgroup analysis is performed independently for each group.

Table B.4: Effect on Cost-Overruns Controlling for Firm-FE

| | Share with Cost Overruns | | | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Estimate | 0.076 | 0.099 | 0.022 | 0.044 |
| S. E. | (0.027) | (0.042) | (0.029) | (0.035) |
| Estimation Method | IV | CCT | IV | CCT |
| Firm Fixed Effect Included | No | No | Yes | Yes |
| Number of Observations | 69296 | 69296 | 69296 | 69296 |

Notes: This table shows the instrumental variable estimates for appearing in FedBizzOpps. The dependent variable is an indicator of having any positive cost-overrun. Columns 1 and 3 are linear IV estimates. Columns 2 and 4 show fuzzy RD estimates using robust-local polynomial regression (CCT - Cataneo et al. 2014). Columns 3 and 4 include firm fixed effects to control for the average performance of the contractors. The sample consists of observations from contractors that appear more than once in the data; otherwise, they are naturally dropped from fixed effect regression. Column 4 is estimated using a residualized outcome variable. Cost overruns are computed as the difference between actual obligated contract dollars and expected total obligations at the award time. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019.

Table B.5: Expected Ex-Post Adaptations by Product Category

| | Goods | | | Services | | |
|---|---|---|---|---|---|---|
| Rank | Name | Average Cost-Overrun | Average Delay | Name | Average Cost-Overrun | Average Delay |
| Low | | | | | | |
| 1 | Fuels, Lubricants, Oils, Waxes | -0.003 | 0.009 | Transport, Travel, Relocation | 0.016 | 0.029 |
| 2 | Musical Inst/Phonograph/Home Radio | -0.001 | 0.016 | Construct Of Structures/Facilities | 0.026 | 0.131 |
| 3 | Valves | -0.000 | 0.016 | Installation Of Equipment | 0.027 | 0.090 |
| High | | | | | | |
| 1 | Chemicals And Chemical Products | 0.037 | 0.062 | Operation Of Govt Owned Facility | 0.758 | 0.703 |
| 2 | Ammunition And Explosives | 0.034 | 0.110 | Utilities And Housekeeping | 0.343 | 0.320 |
| 3 | Office Mach/Text Process/Visib Rec | 0.030 | 0.045 | Medical Services | 0.270 | 0.269 |

Notes: This table presents the top and bottom 3 product categories in terms of average cost-overruns and average delays. The data source is the Federal Procurement Data System-Next Generation. The sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 5,000 and $ 20,000, awarded by the Department of Defense in fiscal years 2015 through 2019. We define the fraction of cost-overrun as the final price, including all modifications, minus the award price divided by the award price. The average delay is the final contract duration minus the original contract duration divided by the original duration. These statistics are constructed based on all contracts for the same contract ID. A 4-digit alphanumeric code (PSC) is observed for each contract. The categories listed are constructed by aggregating PSC codes to two-digits for goods and to a single digit (letter) for services.

Table B.6: Summary Statitistics Local vs. Non-Locals

| | Local | Non-Local | Diff |
|---|---|---|---|
| log Distance | 3.471 | 4.554 | -1.083 |
| Located in the Same State | 0.695 | 0.501 | 0.194 |
| Overruns | 0.078 | 0.236 | -0.158 |
| Delays | 0.130 | 0.275 | -0.145 |
| Number of Modifications | 0.548 | 0.880 | -0.332 |

Notes: This table presents summary statistics for distance and execution variables for contracts performed by locals and non-locals. The sample includes contracts between 10,000 and 40,000 dollars, and buyer-product combinations that appeared at least four times between 2013 and 2019. The need for observing multiple buyer-product observations stem from the way we categorize these contractors. The variables "Overruns" and "Delays" are relative to the original cost and duration, respectively. The difference between these groups is significant at 1% for all variables.

Table B.7: Variable Description Model

| | Model Sample | Full Sample | Diff |
|---|---|---|---|
| *Variables:* | | | |
| Publicized in FBO | 0.373 | 0.274 | 0.099 |
| Award Amount | 21.178 | 20.627 | 0.551 |
| Number of Offers | 3.002 | 3.098 | -0.096 |
| Overruns (relative) | 0.117 | 0.088 | 0.029 |
| Service | 0.375 | 0.308 | 0.067 |
| Mean Overruns Prod Cat | 0.089 | 0.071 | 0.018 |
| Awarded in September | 0.249 | 0.262 | -0.013 |
| log Duration | 3.976 | 3.811 | 0.165 |
| *Bidders' Classification* | | | |
| Local is Awarded | 0.754 | - | - |
| N Potential Local Bidders | 6.078 | - | - |
| N Potential Non-Local Bidders | 3.339 | - | - |
| Number of Observations | 24,135 | 103,899 | |

Notes: This first column describes the mean of the variables included in the model estimation. The second column shows variables mean but over the full sample. The third column shows the differences between these two means. The model's sample corresponds to the subset of contracts over which we could identify the number of potential local and non-local bidders. We restrict the analysis to buyer-product combinations that meet two conditions: at least four contracts were awarded between 2013 and 2019, and not all nor none were publicized.

Table B.8: Estimated Parameters of Entry, Bidding and Execution

| | | Entry (Probit) | | Bid Distribution (Log Normal) | | Execution (Log Normal) | |
|---|---|---|---|---|---|---|---|
| | $\bar{x}$ | Coeff | S.E. | Coeff | S.E. | Coeff | S.E. |
| *Panel A: Coefficients* | | | | | | | |
| Constant | | 0.0048 | ( 0.00057 ) | 3.0517 | ( 0.00025 ) | -1.7041 | ( 0.00038 ) |
| Service | 0.375 | -0.0598 | ( 0.00014 ) | 0.0000 | ( 0.00030 ) | 0.0751 | ( 0.00026 ) |
| Degree of Complexity | 0.089 | -0.7367 | ( 0.00071 ) | -0.0005 | ( 0.00037 ) | 1.6284 | ( 0.00066 ) |
| Non-Local | | 2.1651 | ( 0.00040 ) | -0.0402 | ( 0.00037 ) | 0.2354 | ( 0.00069 ) |
| Non-Local×Complexity | | 0.0299 | ( 0.00026 ) | -0.0308 | ( 0.00063 ) | 0.0868 | ( 0.00082 ) |
| Last Month | 0.249 | -0.8826 | ( 0.00077 ) | | | | |
| Exp. Duration > Median | 0.5 | | | | | 0.1273 | ( 0.00025 ) |
| $N^L$ | 6.078 | 0.0002 | ( 0.00010 ) | -0.0023 | ( 0.00027 ) | | |
| $N^{NL}$ | 3.339 | -0.1876 | ( 0.00028 ) | 0.0099 | ( 0.00024 ) | | |
| *Panel B: Standard Deviation* | | | | | | | |
| Contant | | | | -2.6132 | ( 0.00081 ) | -0.4102 | ( 0.00048 ) |
| Service | | | | 0.0996 | ( 0.00031 ) | 1.1854 | ( 0.00048 ) |
| S.D. Unob. Het. ($\sigma_u$) | | | | 2.1683 | ( 0.00087 ) | | |
| *Panel C: Buyer Preferences* | | | | | | | |
| | | Publicity Choice (Probit) | | | | | |
| Constant | | -0.2584 | ( 0.00059 ) | | | | |
| Exp. Price ($\beta^P$) | | -0.6361 | ( 0.00043 ) | | | | |
| Exp. Cost-Overruns ($\beta^Q$) | | -0.2457 | ( 0.00058 ) | | | | |
| Exp. Local Winning ($\beta^L$) | | 0.5879 | ( 0.00069 ) | | | | |
| Above $25K | | 0.8542 | ( 0.00037 ) | | | | |
| Number of Obs. | 24,135 | | | | | | |

Notes: The table displays the coefficients and corresponding standard errors. Panel A describes the coefficients corresponding to the entry choice mean of (log) bids and the mean of (log) quality shocks. Panel B displays information estimates for the standard deviation of (log) bids, unobserved heterogeneity, and (log) quality. Panel C shows the coefficients associated with the publicity choice by the buyer. Agency and year fixed effects are omitted in this table. These coefficients are estimated altogether using Simulated Method of Moments (SMM). Log-bids and the log of the unobserved project heterogeneity are assumed to be normally distributed. The entry and publicity choices distribute Probit. The standard deviation of log-bids and log-quality shocks are estimated as $\sigma = exp(b_0 + b_1 \mathbb{1}(Service))$, where $\mathbb{1}(Service)$ indicates a contract for service

# C Additional Details on the Setting

## C.1 FedBizzOpps

FedBizOpps.gov (FBO) has been designed as a single government point of entry (GPE) for Federal buyers to publish and for vendors to find posted Federal business opportunities across departments and agencies. The FAR (part 5) regulates the publicity of contract actions. The goals of publicity policy (FAR 5.002) are (a) increase competition, (b) broaden industry participation in meeting Govt requirements (c) assist small businesses (and VO, VOSD, WO, HUBZone, etc.) in winning contracts and subcontracts. The FAR requires that contract actions expected to exceed $25,000 must be *synopsized* in the GPE. Contract actions under $25,000 must publicize "by displaying in a public place, or by any appropriate electronic means." The contracting officer is exempted to advertise in GPE (FAR 5.102(a)5 and 5.202), when "disclosure compromises national security, " "nature of the file (e.g., size) does not make it cost-effective or practicable," the "agency's senior procurement executive makes a written determination that it is not in the Government's interest," and several other special cases (see FAR 5.202).

Figure A2 displays screenshots to the website. Panel (a) shows the list of opportunities, Panel (b) includes the information contained a specific solicitation :

### C.1.1 Types of FBO Notices

There are two broad types of FBO notices: *pre-award* and *post-award* notices. The *pre-award* notices are divided into four actions:[42]

- **Presolicitation:** The pre-solicitation notice makes vendors aware that a solicitation may follow. Vendors may add themselves to the Interested Vendors List, if the posting agency has enabled this feature. This helps government agencies determine if there are qualified vendors to perform the work scope and allows the contracting office to gather information on the interested vendors.
- **Combined Synopsis/Solicitation:** Most opportunities classified this way are open for bids from eligible vendors. These opportunities include specifications for the product or service requested and a due date for the proposal. The notice will specify bidding procedures in the details of the solicitation.
- **Sources Sought:** The Sources Sought notice is a synopsis posted by a government agency seeking possible sources for a project. It is not a solicitation for work or a request for proposal. For more information, see FAR 7.3 and OMB Circular A-76.

[42] Here we omit uncommonly used actions: *Sale of Surplus Property*, *Justification and Approval (J&A)*, *Fair Opportunity / Limited Sources Justification*, *Foreign Government Standard*, and *Intent to Bundle Requirements (DoD-Funded)*.

- **Special Notice:** Agencies use Special Notices to announce events like business fairs, long-range procurement estimates, pre-bid/pre-proposal conferences, meetings, and the availability of draft solicitations or draft specifications for review.

The post-award notices are essentially *award notices*:

- **Award Notice:** When a federal agency awards a contract in response to a solicitation, they may choose to upload a notice of the award to allow the interested vendors to view the vendor receiving the awarded contract, and amount agreed upon.

Figure A1 describes the life-cycle of a project and how different stages are linked to FBO actions.

## C.2 Dataset Details

Our analysis combines data from two sources: Federal Procurement Data System - Next Generation (FPDS-NG) and data scrapped directly from FedBizzOpps.gov (FBO).

**FPDS-NG.** The FPDS-NG tracks the universe of federal awards that exceed $5,000.[43] The Federal Acquisition Regulation (FAR) requires Contracting Officers (COs) must submit complete reports on all contract actions. Thus, every observation corresponds to a contract action, representing either an initial award or a follow-on action, e.g., modification, termination, renewal, or exercise of options. For each observation, we observe detailed information, such as the dollar value of the funds obligated by the transaction; a four-digit product category code (PSC); six-digit Industry (NAICS) code; identification codes for the agency, sub-agency, and contracting office making the purchase; the identity of the private vendor (DUNS); the type of contract pricing (typically, fixed-price or cost-plus); the extent of competition for the award; characteristics of the solicitation procedure; the number of offers received; and the applicability of a variety of laws and statutes. We collapse all actions by contract ID. As a reference, 80% of awarded contracts are smaller than $50,000.

Our analysis contemplates overruns in terms of cost and time of completion. We define contract delays and cost overruns based on related literature (**?**). We exclude outliers on both variables as they are likely associated with data entry issues. We cross-checked dates and amounts for contract award notices that appeared in FBO and found that mismatches are uncommon.

**FBO Data.** We use daily archives of all information posted in FBO. Every data row corresponds to a different notice action. Each action is associated with a unique URL. The two primary IDs to match FBO data with other datasets are "solicitation number" and "contract award number. The former identifies pre-award actions, whereas award notices are identified using "contract award number." A relevant fraction of the award-notices are not linked with any of the pre-award notices. FPDS data contain both IDs. Roughly, an annual database contains 300,000 notices.

The data preparation consists in three steps; first, we clean IDs and classify different actions associated with each ID. Second, we merge with FPDS data using contract number, then update

[43] The data can be downloaded from `usaspending.gov`

solicitation number when both exist, finally merge and append unmatched observations using solicitation number. The last step is to collapse the data at the FPDS contract ID level. So the resulting dataset contains all the contract ids that also appeared in FBO.

We define that a contract appeared in FBO (treatment indicator) if the contract award has a solicitation number associated with at least one of the FBO pre-award actions described above.

# D Model Identification

**Lemma 1.** *The expected k-th order statistic of B with n draws can be written in terms of the expected k-th and $(k+1)$-th order statistics with $n+1$ draws*

*Proof.* The probability density function of $B$ is $g_B(b)$, then the $k$-th order statistic of $B$, $g_{B_k^{(n)}}(b)$, is:

$$\begin{aligned} g_{B_k^{(n)}}(b) &= k\binom{n}{k} g_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n-k} \\ &= \frac{n!}{(n-k)!(k-1)!} f_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n-k} \end{aligned}$$

Thus, the difference is expected $k$-th order statistics with $n$ and $n+1$ actual competitors is expressed as follows:

$$\begin{aligned} &\mathbb{E}[B_k^{(n)}] - \mathbb{E}[B_k^{(n+1)}] = \int_{\underline{b}}^{\bar{b}} b g_{B_k^{(n)}}(b) db - \int_{\underline{b}}^{\bar{b}} b g_{B_k^{(n+1)}}(b) db \\ &= \int_{\underline{b}}^{\bar{b}} bk\binom{n}{k} g_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n-k} db - \int_{\underline{b}}^{\bar{b}} bk\binom{n+1}{k} g_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n+1-k} db \\ &= \int_{\underline{b}}^{\bar{b}} \left( \frac{n!(n+1-k) - (n+1)! \left[1 - G_B(b)\right]}{(k-1)!(n+1-k)!} \right) b g_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n-k} db \\ &= \int_{\underline{b}}^{\bar{b}} \left( \frac{(n+1)! G_B(b) - n!k}{(k-1)!(n+1-k)!} \right) b g_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n-k} db \\ &= \int_{\underline{b}}^{\bar{b}} \frac{(n+1)!}{(k-1)!(n+1-k)!} b g_B(b) G_B(b)^{k} \left[1 - G_B(b)\right]^{n-k} db \\ &\quad - \int_{\underline{b}}^{\bar{b}} \frac{n!k}{(k-1)!(n+1-k)!} b g_B(b) G_B(b)^{k-1} \left[1 - G_B(b)\right]^{n-k} db \\ &= \frac{k}{(n+1-k)} \left( \mathbb{E}[B_{k+1}^{(n+1)}] - \mathbb{E}[B_k^{(n)}] \right) \end{aligned}$$

Rearranging the terms, we get the expected $k$-th order statistic of $n$ draws can be expressed as a simple weighted average of the $k$-th and $k+1$-th order statistic under $n+1$ draws:

$$\mathbb{E}[B_k^{(n)}] = \frac{k}{n+1} \mathbb{E}[B_{k+1}^{(n+1)}] + \frac{n+1-k}{n+1} \mathbb{E}[B_k^{(n+1)}] \tag{7}$$

□

## D.1 Identification under Unobserved Heterogeneity

Below we show that identification can be achieved when only the winning bid and the number of (symmetric) bidders are observed as long as the number of bidders is exogenous. In particular, in our setting, bidders define bidding strategies without knowing the actual number of bidders, $n$, but based on beliefs about market conditions. Thus, $n$ is exogenous conditional on $(N, \varphi)$. We leverage variation in actual bidders to separately identify the private and the common cost components' distributions. To ease notation, we omit $(N, \varphi)$ as conditions for exogeneity of $n$.

**Proposition 1.** *First price auctions with unobserved heterogeneity can be identified when only the winning bid and the number of bidders are observed as long as the number of active bidders is exogenous.*

*Proof.* The ratio of first-order statistics is identified by comparing observed winning bids for different values of $n$:

$$\frac{\frac{1}{T_{n,N}}\sum_t (B_{1,t}|n_t = n)}{\frac{1}{T_{n',N}}\sum_t (B_{1,t}|n_t = n')} \to \frac{\mathbb{E}[B_{1:n}]}{\mathbb{E}[B_{1:n'}]} = \frac{\mathbb{E}[\tilde{B}_{1:n} \cdot u]}{\mathbb{E}[\tilde{B}_{1:n'} \cdot u]} = \frac{\mathbb{E}[\tilde{B}_{1:n}] \cdot \mathbb{E}[u]}{\mathbb{E}[\tilde{B}_{1:n'}] \cdot \mathbb{E}[u]} = \frac{\mathbb{E}[\tilde{B}_{1:n}]}{\mathbb{E}[\tilde{B}_{1:n'}]} \tag{8}$$

where $(B_{1,t}|n_t = n)$ is auction's $t$ observed winning bid with $n$ active bidders. $\mathbb{E}[\tilde{B}_{1:n}]$ is the expected first order statistic normalized based on $u_t = 1$. Finally, $u$ is assumed independent of the number of bidders and cancels out in the last identity. The normalization $\mathbb{E}[u] = 1$ pins down the scale of the first order statistics.

By contradiction; assume $(\hat{G}_{\tilde{b}}, \hat{H}_u)$ provide the same distribution observed in the data,

$$\tilde{B}_{1:n} u \overset{d}{=} \hat{\tilde{B}}_{1:n} \hat{u}$$
$$\tilde{B}_{1:n'} u \overset{d}{=} \hat{\tilde{B}}_{1:n'} \hat{u}$$

Construct $\tilde{b}^*_{n'}, \hat{\tilde{b}}^*_{n'} \tilde{u}^*$, and $\tilde{u}^*$ as random variables that are independent of and have the same conditional distributions as their asterisk-free counterparts. Then it follows that

$$\begin{aligned} (\tilde{B}_{1:n} u) \cdot \left(\hat{\tilde{B}}^*_{1:n'} \hat{u}^*\right) &\overset{d}{=} \left(\hat{\tilde{B}}_{1:n} \hat{u}\right) \cdot \left(\tilde{B}^*_{1:n'} u^*\right) \\ \implies \tilde{B}_{1:n} \cdot \hat{\tilde{B}}^*_{1:n'} &\overset{d}{=} \hat{\tilde{B}}_{1:n} \cdot \tilde{B}^*_{1:n'} \end{aligned} \tag{9}$$

Taking expectations on both sides:

$$
\begin{aligned}
\mathbb{E}[\tilde{B}_{1:n}] \cdot \mathbb{E}[\hat{\tilde{B}}_{1:n'}] &= \mathbb{E}[\hat{\tilde{B}}_{1:n}] \cdot \mathbb{E}[\tilde{B}_{1:n'}] \\
\frac{\mathbb{E}[\tilde{B}_{1:n}]}{\mathbb{E}[\tilde{B}_{1:n'}]} &= \frac{\mathbb{E}[\hat{\tilde{B}}_{1:n}]}{\mathbb{E}[\hat{\tilde{B}}_{1:n'}]}
\end{aligned}
$$

If $(\hat{G}_{\tilde{b}}, \hat{H}_u)$ rationalizes the data, it has a normalized distribution with the same ratio of first order statistics. Using, order statistic's recurrence relation (Lemma 1), we have that $\mathbb{E}[B_{1:n-1}] = \frac{1}{n}\mathbb{E}[B_{2:n}] + \frac{n-1}{n}\mathbb{E}[B_{1:n}]$, we can link together these ratios when $n' = n - 1$:

$$
\begin{aligned}
\frac{\mathbb{E}[\tilde{B}_{1:n}]}{\mathbb{E}[\tilde{B}_{1:n'}]} &= \frac{\mathbb{E}[\hat{\tilde{B}}_{1:n}]}{\mathbb{E}[\hat{\tilde{B}}_{1:n'}]} \\
\frac{\mathbb{E}[\tilde{B}_{1:n}]}{\frac{1}{n}\mathbb{E}[\tilde{B}_{2:n}] + \frac{n-1}{n}\mathbb{E}[\tilde{B}_{1:n}]} &= \frac{\mathbb{E}[\hat{\tilde{B}}_{1:n}]}{\frac{1}{n}\mathbb{E}[\hat{B}_{2:n}] + \frac{n-1}{n}\mathbb{E}[\hat{\tilde{B}}_{1:n}]} \\
\frac{\mathbb{E}[\tilde{B}_{1:n}]}{\mathbb{E}[\tilde{B}_{2:n}]} &= \frac{\mathbb{E}[\hat{\tilde{B}}_{1:n}]}{\mathbb{E}[\hat{\tilde{B}}_{2:n}]}
\end{aligned}
$$

$\hat{G}_{\tilde{b}}$ has the same ratio of second-order statistics. With sequential values of $n \in \{2, \ldots, N\}$, we can iterate forward from the identified first-order and second-order statistics using the recursive relation between order statistics from Proposition 1. Therefore, $G_{\tilde{b}}$ and $\hat{G}_{\tilde{b}}$ are identical up to the first $N$ order statistics from $\tilde{B}$ □

**Corollary 1.** *The distribution of the unobserved heterogeneity, $H_u$ is obtained once $G_{\tilde{b}}$ is identified.*

*Proof.* By Independence of $\tilde{B}$ and $u$, leveraging basic properties of characteristic functions we can write $\psi_{log(B_{1:n})} = \psi_{log(\tilde{B}_{1:n})}\psi_{log(u)}$, where $\psi_{log(B_{1:n})}$ is the characteristic function of the log of observed winning bids under $n$ active bidders. We can construct this characteristic function for different values of $n$. Once the characteristic function of $G_{\tilde{b}}$ is obtained, we can pin down $H_u$ □

**Corollary 2.** *The distribution of normalized private costs, $F_{\tilde{c}}$ is identified once $G_{\tilde{b}}$ and equilibrium entry probabilities are obtained.*

This corollary follows from Guerre et al. (2000). If the distribution of $G_{\tilde{b}}$ is recovered, and the equilibrium entry probabilities are observed from entry choices. Then, we can use the first order and the boundary conditions to recover the latent distribution $F_{\tilde{c}}$.

# E Model Estimation Details

## E.1 Classifying Contractors' *Types*

Based on the patterns of contractor's participation, we identify two separate groups of firms: contractors who win awards without relying on publicity —which we refer to as *locals*—, and contractors that *only* win when contract solicitations are publicized —which we label *non-locals*. The logic is that, if a contractor wins without publicity, this indicates that the buyer informed her directly (e.g. through email or a phone call). The existence of direct communication reveals a buyer's preference for these contractors. Conversely, if a contractor requires a FedBizzOpps announcement to participate (and win), this suggests that there is no specific preference from that buyer for that contractor. This distinction came up frequently in conversations with procurement officers from several organizations.

To classify contractors empirically, we restrict the analysis to buyer-product combinations that are observed at least 4 times between 2013 and 2019, and which had at least one —but not all— contracts publicized.[44] Table B.6 compares buyer-contractor distance and performance for contracts performed by local and non-locals. The third column shows the mean difference of performance between these two groups. As a reference, if the information source is irrelevant, locals and non-locals would have similar outcomes. However, we observe that contracts executed by *non-local* contractors experience 16 percentage points (200%) more cost-overruns and 14.5 percentage points (110%) more delays than locals.

## E.2 Estimation

Denote the target moments by $m_n$ as a vector of moments from the data. The simulated moments are denoted by $m_s(\theta)$. The depends on the parameters $\theta \in \Theta \subset \mathbb{R}^P$. The estimator minimizes the standard distance metric:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmin}} \left(m_n - m_s(\theta)\right)' W_n \left(m_n - m_s(\theta)\right)$$

Where $W_n$ is the weighting matrix, which is chosen using the standard two-step approach. Letting $M_s(\theta)$ be the $(P \times J)$ Jacobian matrix of the vector of simulated moments; under standard regularity assumptions, we have:

$$\sqrt{n}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} N\left(0, \left(1 + \frac{1}{s}\right)(M'WM)^{-1}M'W\Omega W'M(M'WM)^{-1}\right) \tag{10}$$

[44] We noted that the Federal Procurement Data System (FPDS) sometimes missclassifies local buyers, assigning the same code to different branches that depend on a single (higher-level) office. This contrasts with the nature of most procurement officers' job, who typically contract within a particular area, leveraging their local market knowledge. We address this misclassification by defining a buyer based on the office code *and* the Metropolitan Statistical Area (MSA) of the purchase. As before, the definition of a product category is given by the 4-digit PSC code.

where $W$ is the probability limit of $W_n$, $M$ is the probability limit of $M_s(\theta_0)$, and $\Omega$ is the asymptotic variance of $m_n$ (Pakes and Pollard, 1989). The vector of parameters is: $\theta = (\alpha^k, \nu^k, \tau^k, \gamma^k, \xi^k, \vec{\beta}, \zeta, \vec{\sigma})$.

#### E.2.1 Standard Errors

We compute standard errors using the asymptotic variance formula given by (10). The variance-covariance matrix of $\hat{\theta}$ is:

$$V(\hat{\theta}) = \frac{1}{n}\left(1 + \frac{1}{s}\right)(\hat{M}'W\hat{M})^{-1}\hat{M}'W\hat{\Omega}W'\hat{M}(\hat{M}'W\hat{M})^{-1}$$

Where $\hat{\Omega}$ is estimated via bootstrap: re-sampling contracts with replacement from the original data, and recompute the smoothed vector of moments, repeating this process 500 times. $\hat{\Omega}$ is the sample variance of these 500 vectors. $\hat{M}$ is the numeric derivative of the SMM objective function (6) evaluated at $\hat{\theta}$.

#### E.2.2 Minimization

We keep constant the underlying random draws throughout the minimization of the objective function. Nonetheless, the simulated objective is not continuous with respect to $\theta$. Thus, We leverage the stochastic optimization algorithm *Differential Evolution* (Storn and Price, 1997) to perform the objective minimization. This algorithm does not rely on gradient methods, and given its heuristic approach for minimizing possibly nonlinear and non-differentiable continuous space functions, it is robust to poorly behaved objectives.

### E.3 Moments

We use three sets of target moments.

- First set of moments,
  - $\vec{m}_{11} = \mathbb{E}[x_t^{(y)\prime} y_t]$ and $\vec{m}_{12} = \mathbb{E}[x_t^{(y)\prime} y_t^2]$, where $y_t =$ *log winning bid, number of bidders, wins local, log overruns,* and *contract is publicized*, and $x_t^{(y)} = (1, \vec{x}_t^{(y)})$ = covariates associated with outcome variable $y$
- Second set:
  - $\vec{m}_2 = \mathbb{E}[y_t | B_t \in (B^l, B^{l+1})]$, for $l \in \{1, \ldots, L-1\}$, where $y_t =$ *number of bidders, wins local, log overruns,* and *contract is publicized*. We separate these moments based on goods and services, and partition the domain of contract prices in bins of width \$1,000
- Third set of moments:

- $\vec{m}_3 = \mathbb{E}[\mathbb{1}\{b_t \in (b^l, b^{l+1})\}]$, for $l \in \{1, \ldots, L-1\}$. This set of moments correspond to the normalized frequencies on the relevant window of contract prices. The bin width is \$1,000.

As a result we use 357 moments to estimate 37 parameters.

# Supplementary Material for:

## Competition Under Incomplete Contracts and the Design of Procurement Policies

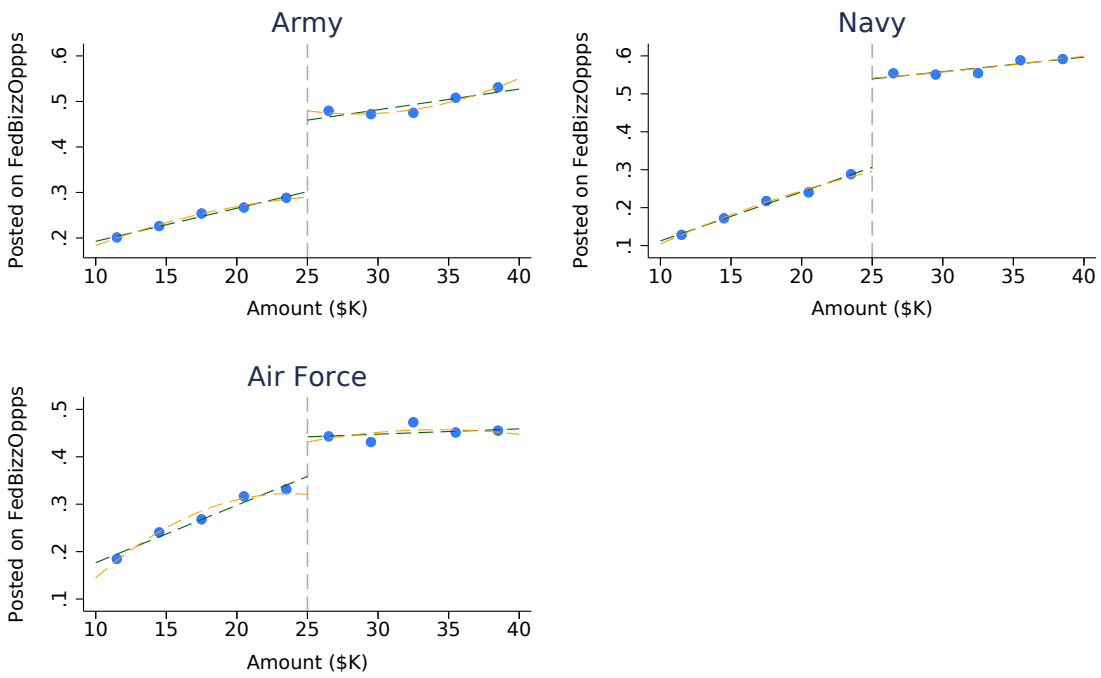
**(Not for Publication)**

*Rodrigo Carril, Andres Gonzalez-Lira, and Michael S. Walker*

## F Additional Figures

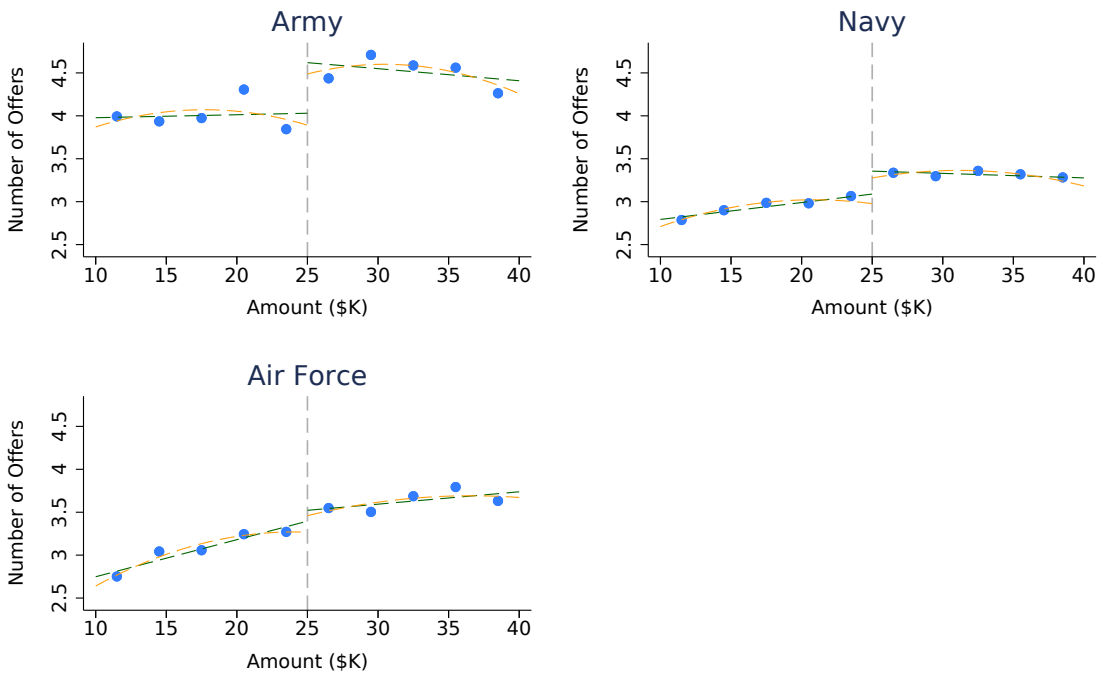
### F.1 Heterogeneous Effects of Publicity

Figure A1: Heterogeneous publicity adoption by major departments



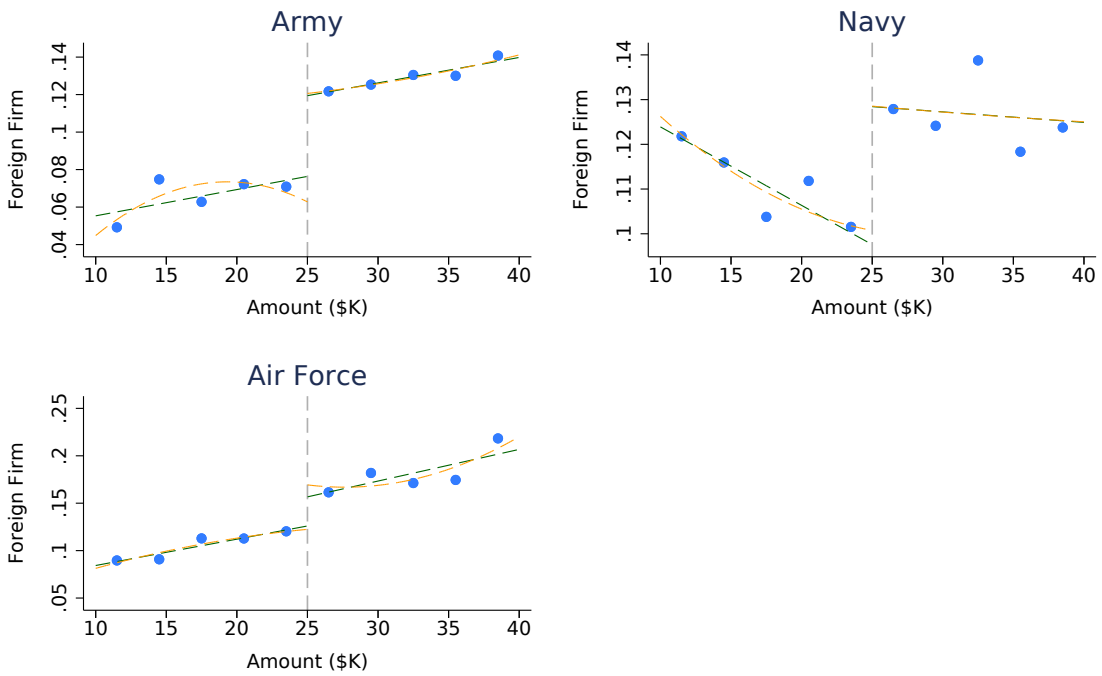

Notes: This figure presents three binned scatter plots, which depict the share of contracts publicized in FedBizzOpps by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 5,000 and $ 45,000, awarded by the Department of Defense in fiscal years 2011 through 2017. Panel (a) restricts the sample to awards made by the Army. Panel (b) restricts the sample to awards made by the Navy. Panel (c) restricts the sample to awards made by the Air Force. Award amounts are discretized into right-inclusive bins of $2,500 dollars length.

Figure A2: Heterogeneous effects on competition by major departments



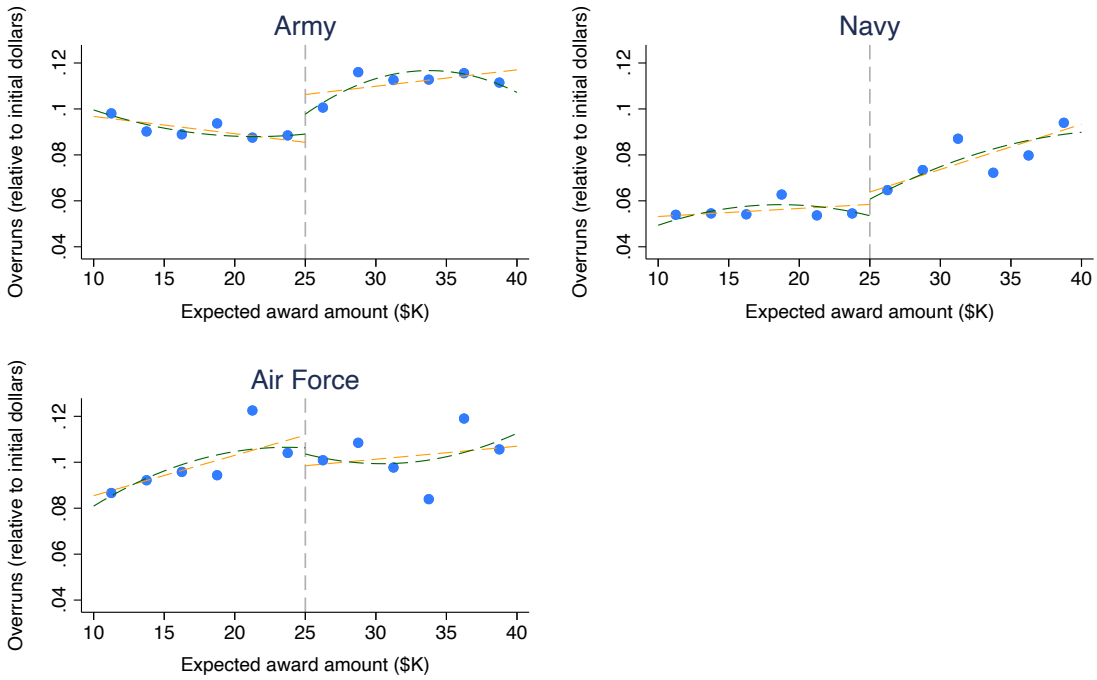

Notes: This figure presents three binned scatter plots, which depict the average number of offers received by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 5,000 and $ 45,000, awarded by the Department of Defense in fiscal years 2011 through 2017. Panel (a) restricts the sample to awards made by the Army. Panel (b) restricts the sample to awards made by the Navy. Panel (c) restricts the sample to awards made by the Air Force. Award amounts are discretized into right-inclusive bins of $2,500 dollars length.

Figure A3: Heterogeneous effects on winner characteristics by major departments



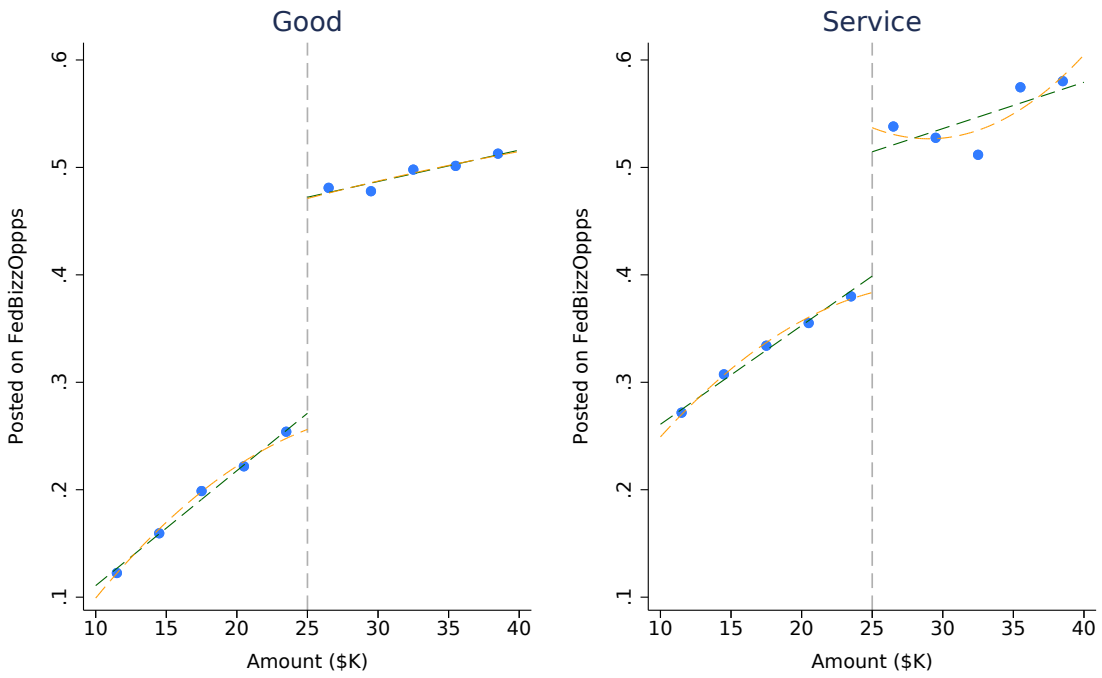

Notes: This figure presents three binned scatter plots, which depict the share of contracts awarded to a foreign firm by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 5,000 and $ 45,000, awarded by the Department of Defense in fiscal years 2011 through 2017. Panel (a) restricts the sample to awards made by the Army. Panel (b) restricts the sample to awards made by the Navy. Panel (c) restricts the sample to awards made by the Air Force. Award amounts are discretized into right-inclusive bins of $2,500 dollars length.

Figure A4: Heterogeneous effects on performance by major departments



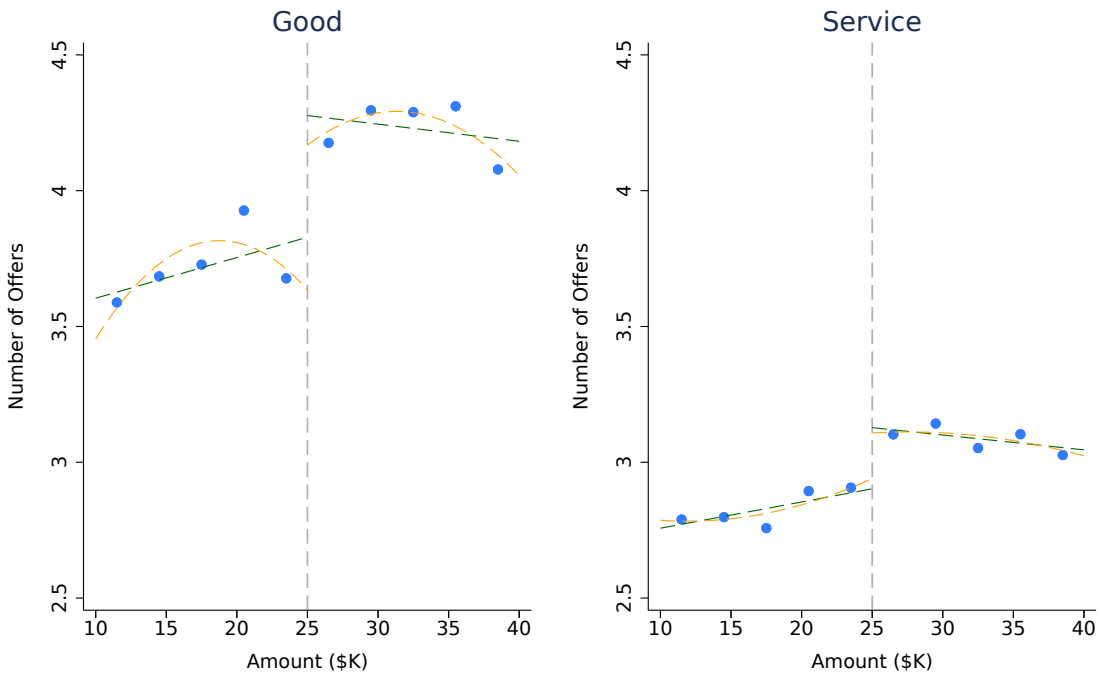

Notes: This figure presents three binned scatter plots, which depict average cost overruns by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. Cost overruns are computed as the difference between actual obligated contract dollars and expected total obligations at the time of the award, divided by expected obligations. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 5,000 and $ 45,000, awarded by the Department of Defense in fiscal years 2011 through 2017. Panel (a) restricts the sample to awards made by the Army. Panel (b) restricts the sample to awards made by the Navy. Panel (c) restricts the sample to awards made by the Air Force. Award amounts are discretized into right-inclusive bins of $2,500 dollars length.

Figure A5: Heterogeneous publicity adoption: goods versus services



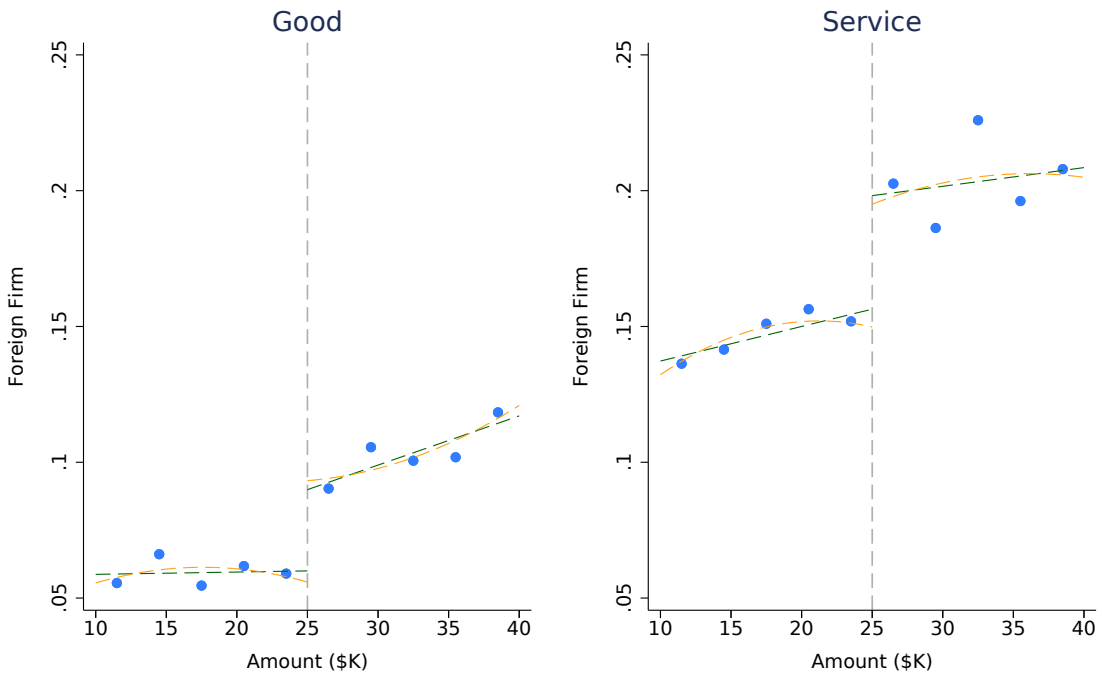

Notes: This figure presents two binned scatter plots, which depict the share of publicized contracts by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Panel (a) restricts the sample to awards for goods, while Panel (b) restricts the sample to service contracts. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

Figure A6: Heterogeneous effects on competition: goods versus services



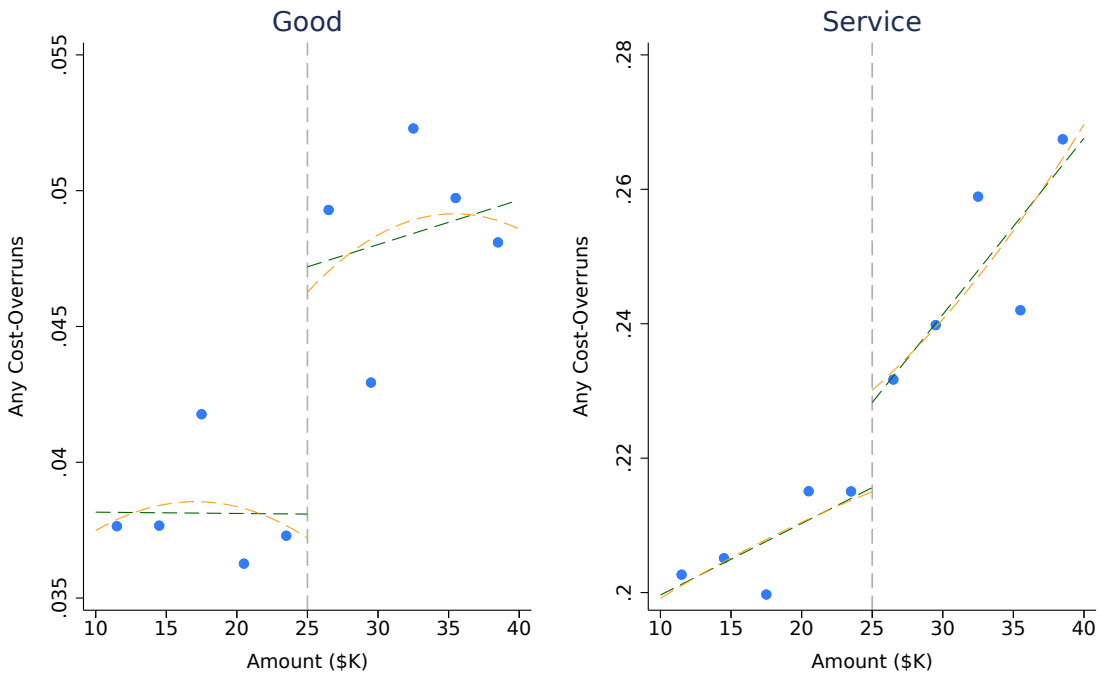

Notes: This figure presents two binned scatter plots, which depict the average number of offers received by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Panel (a) restricts the sample to awards for goods, while Panel (b) restricts the sample to service contracts. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

Figure A7: Heterogeneous effects on winner characteristics: goods versus services




Notes: This figure presents two binned scatter plots, which depict the share of contracts awarded to a foreign firm by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Panel (a) restricts the sample to awards for goods, while Panel (b) restricts the sample to service contracts. Award amounts are discretized into right-inclusive bins of $2,500 dollars length.

Figure A8: Heterogeneous effects on performance: goods versus services




Notes: This figure presents two binned scatter plots, which depict share of contracts with cost overruns by bins of award amounts, as well as linear and quadratic fits at each side of $25,000. Cost overruns are computed as the difference between actual obligated contract dollars and expected total obligations at the time of the award, divided by expected obligations. The data source is the Federal Procurement Data System-Next Generation. The full sample consists of non-R&D definitive contracts and purchase orders, with award values between $ 10,000 and $ 40,000, awarded by the Department of Defense in fiscal years 2015 through 2019. Panel (a) restricts the sample to awards for goods, while Panel (b) restricts the sample to service contracts. Award amounts are discretized into right-inclusive bins of $3,000 dollars length.

# G Empirical Framework for Estimating the Effects of Publicity on Contract Outcomes

This Appendix presents a detailed exposition of the empirical framework introduced in Section 3. Section G.1 presents our theoretical framework and the set of results that motivate the density analysis. Section G.1.5 explains the density analysis in detail, including all implementation details. Section G.1.6 discusses how to correct naive RDD estimates to account for price effects and potential measurement error. Section G.1.8 explains how we account for potential bunching responses in the RDD framework.

## G.1 Empirical Model

### G.1.1 Preliminaries

Consider a series of observed contract awards $t \in \{1, .., T\}$. Let $\tilde{p}_t$ be the *ex-ante award price* of contract $t$, which corresponds to the agency's estimate of what the contract price will be. Let $p_t$ be the *observed award price* of contract $t$. $\tilde{p}_t$ and $p_t$ are normalized relative to a policy threshold of \$25,000 and measured in logs. Therefore, negative (positive) values of $\tilde{p}_t$ and $p_t$ are said to be below (above) the threshold for the purpose of the policy described below.

Prior to the award, the buyer decides whether to publicize the solicitation ($D_t = 1$) or not ($D_t = 0$). Let $p_t^d(\tilde{p}_t)$ be the potential price that we would observe for contract $t$, given an ex-ante estimate of $\tilde{p}_t$ and a publicity decision $D_t = d$, for $d \in \{0, 1\}$. There is a policy that encourages buyers to choose $D_t = 1$ for awards expected to exceed the threshold (i.e. for $\tilde{p}_t > 0$).

The buyer may choose to *strategically bunch* ($B_t = 1$), which means that she modifies the characteristics of the initial purchase, in order to obtain an award price equal to $p_t^B(\tilde{p}_t)$, choosing $D_t = 0$ without being affected by the policy. $p_t^B(\tilde{p}_t)$ is equal to, or slightly below 0.

Therefore, observed prices can be written as:

$$p_t = p_t^0(\tilde{p}_t) + D_t \cdot \left[p_t^1(\tilde{p}_t) - p_t^0(\tilde{p}_t)\right] + B_t \cdot (1 - D_t) \cdot \left[p_t^B(\tilde{p}_t) - p_t^0(\tilde{p}_t)\right]$$

We assume the following:

**A1** $\tilde{p}_t$ are i.i.d. draws from a distribution with smooth density $f_{\tilde{p}}(\cdot)$.

**A2** $p_t^0(\tilde{p}_t) = \tilde{p}_t + \xi_t$, with $\xi_t \sim F_\xi(\cdot)$, $E[\xi_t] = 0$, and $\xi_t \perp \tilde{p}_t$.

**A3** $p_t^1(\tilde{p}_t) = \tilde{p}_t + \gamma_t$, with $\gamma_t \sim F_\gamma(\cdot)$, $\gamma_t \perp \tilde{p}_t$, and $\gamma_t \perp \xi_t$.

**A4** $\Pr(D_t = 1|\tilde{p}_t) \equiv \tilde{\pi}_D(\tilde{p}_t) = \tilde{\pi}_D^*(\tilde{p}_t) + \delta \cdot \mathbf{1}[\tilde{p}_t > 0]$, for a continuous function $\tilde{\pi}_D^*(\cdot)$.

**A5** There exist $p_H > 0$ such that $B_t = 0$ for all $\tilde{p}_t > p_H$.

Note that here we present a slightly more general version of the model that in Section 3. In particular, **A2** allows for measurement error in agencies' ex-ante estimates.

### G.1.2 Discretizing award values

Consider the division of the range of possible (normalized) award values into a set of equally-sized and right-inclusive bins around the threshold $b \in \{-R, (-R+1), ..., -1, 0, 1, ..., (R-1), R\}$. Note that bin $b = 0$ includes awards right at, or slightly below, the policy threshold.

Let $\{n_b^d\}_{b=-R}^{R}$ be the frequency distribution of observed awards conditional on treatment (publicity) status $D_t = d$, for $d \in \{0, 1\}$, so that $n_b^d$ denotes the number of contracts with treatment status $d$ and observed award value $p_t \in b$. Likewise, let $\{\tilde{n}_b^d\}_{b=-R}^{R}$ represent the (unobserved) frequency distribution of latent ex-ante prices. We also denote the distribution of *all* awards (both publicized and non-publicized) by simply omitting the superscript. That is, $n_b = n_b^0 + n_b^1$, and $\tilde{n}_b = \tilde{n}_b^0 + \tilde{n}_b^1$.

Consider also a *shifted* distribution of publicized contracts $\{n_b^{1,s}(\bar{\gamma})\}_{b=-R}^{R}$, which is obtained by subtracting a mean price effect $\bar{\gamma}$ to every publicized ($D_t = 0$) contract. That is, $n_b^{1,s}(\bar{\gamma})$ denotes the number of *publicized* contracts with award value $p_t$ such that $(p_t + \bar{\gamma}) \in b$.

Finally, let $\Delta$ denote the discrete change in the number of publicized contracts at the discontinuity. Given **A4**, note that this is defined as $\Delta = \delta \cdot \sum_b n_b$.

### G.1.3 Propositions

We now make a series of propositions that motivate our estimation method that we label "density analysis" in Section 3.

**Proposition 2.** *There exist some* $(\underline{b}^1, \bar{b}^1)$ *such that* $E[\tilde{n}_b^1] = E[n_b^{s,1}(\bar{\gamma})]$, *for* $\bar{\gamma} = E[\gamma_t]$, $b < \underline{b}^1 < 0$ *and* $b > \bar{b}^1 > 0$. *That is, far enough from the threshold, the distribution of* realized *award prices, appropriately shifted to cancel out mean price effects, coincides with the distribution of* ex-ante *award prices for publicized contracts.*

**Proposition 3.** *There exist some* $(\underline{b}^0, \bar{b}^0)$ *such that* $E[\tilde{n}_b^0] = E[n_b^0]$, *for* $b < \underline{b}^0 < 0$, *and* $b > \bar{b}^0 > 0$. *In other words, far enough from the threshold, the distributions of* ex-ante *and* realized *award prices for non-publicized contracts coincide.*

**Corollary 3.** $E[\tilde{n}_b] = E[n_b^0 + n_b^{s,1}(\bar{\gamma})]$, *for* $\bar{\gamma} = E[\gamma_t]$, $b < \underline{b} = \min\{\underline{b}^0, \underline{b}^1\} < 0$ *and* $b > \bar{b} = \max\{\bar{b}^0, \bar{b}^1\} > 0$.

**Proposition 4.** $\sum_{b \leq 0}(\tilde{n}_b - n_b) = \sum_{b>0}(n_b - \tilde{n}_b)$. *This means that the excess mass below the threshold equals the missing mass above the threshold.*

**Proposition 5.** $\Delta \cdot F_{\gamma'}(x) = E[n_{b_x}^{1,s}(\bar{\gamma}) - \tilde{n}_{b_x}^1]$, *for* $x \in b_x$, $b_x \leq 0$, *and* $\gamma' = \gamma - \bar{\gamma}$.

### G.1.4 Convolution of densities

The key to our propositions stems from characterizing the distribution of observed prices $p_t$, given the distributions of ex-ante estimates, price effects, and measurement error. Throughout this section, we normalize the price of publicized contracts by subtracting the mean of the price effects. This is for convenience, so that we deal with a mean-zero price effect, but is without loss of generality, as the propositions appropriately adjust for $\bar{\gamma}$ when appropriate.

Consider first the density of publicized contracts, $h_p^1$. Because observed prices are given by the sum of two independent random variables, ex-ante estimates and price effects (see **A3**), their density is given by the convolution of the densities $f_{\tilde{p}}^1 \equiv f_{\tilde{p}|D=1}$ and $f_\gamma$. That is:

$$h_p^1(p_t) = \int_{-\infty}^{\infty} f_{\tilde{p}}^1(p_t - \gamma_t) f_\gamma(\gamma) d\gamma \tag{11}$$

On the other hand, using Bayes' rule:

$$f_{\tilde{p}}^1(\tilde{p}_t) = \frac{\tilde{\pi}_D(\tilde{p}_t) \cdot f_{\tilde{p}}(\tilde{p}_t)}{\Pr(D_t = 1)} \tag{12}$$

So that (11) and (12) imply:

$$\begin{aligned} h_p^1(p_t) &= \int_{-\infty}^{\infty} \frac{\tilde{\pi}_D(p_t - \gamma) \cdot f_{\tilde{p}}(p_t - \gamma) \cdot f_\gamma(\gamma)}{\Pr(D_t = 1)} d\gamma \\ &= \int_{-\infty}^{\infty} \frac{(\tilde{\pi}_D^*(p_t - \gamma) + \delta \cdot \mathbf{1}[p_t - \gamma > 0]) \cdot f_{\tilde{p}}(p_t - \gamma) \cdot f_\gamma(\gamma)}{\Pr(D_t = 1)} d\gamma \\ &= \int_{-\infty}^{\infty} \frac{\tilde{\pi}_D^*(p_t - \gamma) \cdot f_{\tilde{p}}(p_t - \gamma) \cdot f_\gamma(\gamma)}{\Pr(D_t = 1)} d\gamma + \int_{-\infty}^{p_t} \frac{\delta \cdot f_{\tilde{p}}(p_t - \gamma) \cdot f_\gamma(\gamma)}{\Pr(D_t = 1)} d\gamma \end{aligned}$$

Or,

$$h_p^1(p_t) \equiv \int_{-\infty}^{\infty} f_{\tilde{p}}^{1*}(p_t - \gamma) \cdot f_\gamma(\gamma) \cdot d\gamma + \int_{-\infty}^{p_t} \Delta(p_t - \gamma) \cdot f_\gamma(\gamma) \cdot d\gamma \tag{13}$$

Consider $p_t << 0$, so that $f_\gamma(p_t) \approx 0$. In words, consider a price sufficiently below the threshold, so that the probability that the ex-ante estimate for this contract was above the threshold is negligible. In this case, the second term in Equation (13) is zero. On the other hand, $f_{\tilde{p}}^{1*}(p_t - \gamma) = f_{\tilde{p}}^1(p_t - \gamma)$ when $p_t < 0$, so that the first term is the convolution between the densities of $\tilde{p}$ and $\gamma_t$. If the former is sufficiently smooth, then adding a mean-zero price effect has no effect on the observed density, and $h_p^1(p_t) = f_{\tilde{p}}^1(p_t)$. It follows that the expected number of contracts with observed price $p_t$ equals the expected number of contracts with ex-ante price estimate equal to $p_t$. Abandoning the normalization to allow for non-zero average price effects implies that this equality of expectations holds only once observed publicized prices are adjusted by adding the mean of $\gamma$. The first part of Proposition 2 follows: for sufficiently low $p_t \in \underline{b}$, $E[\tilde{n}_b^1] = E[n_b^{s,1}(\bar{\gamma})]$, for all $b \leq \underline{b}$.

As we move closer to the threshold from below, the second term in Equation (13) becomes

positive. This corresponds to the excess mass of contracts, relative to the counterfactual density of the first term. Intuitively, this term is given by the mass of contracts with ex-ante estimate to the right of the threshold that receive a sufficiently high price effect so as to end up at the left of it. This is what allows us to identify $F_\gamma$ in Proposition 5. Consider $p_t = x$ closely below the threshold, so that $\Delta(x - \gamma) \approx \Delta$. With a constant $\Delta$, it immediately follows that $\Delta \cdot F_\gamma(p_t) = h_p^1(p_t) - f_{\tilde{p}}^1(p_t)$.

A symmetric argument can be given for $p_t$ closely above the threshold. In this case, the second term becomes the missing mass of the observed density $h_p^1(p_t)$, relative to the counterfactual density of $\tilde{p}$. Once we get to a high enough value of $p_t >> 0$, once again $f_\gamma(p_t)$ goes to zero and this missing mass disappears. Observed and counterfactual densities converge, which completes Proposition 2: for sufficiently high $p_t \in \underline{b}$, $E[\tilde{n}_b^1] = E[n_b^{s,1}(\bar{\gamma})]$, for all $b > \overline{b}$.

The argument for non-publicized contracts is directly analogous. Observed awards are the sum of unobserved ex-ante estimates $\tilde{p}$ and a mean-zero error term $\xi$. This error term only generates a discrepancy between $h_p^0$ and $f_p^0$ when the latter is not smooth, which happens only at the threshold. Proposition 3 follows: for $p_t << 0$ and $p_t >> 0$, the two densities coincide.

All this discussion ignored the potential effect of bunching responses. However, strategic bunching does not affect any of the aforementioned results. This is because of **A5**: bunching responses occur only within a window around the threshold. Therefore, all of our arguments remain unchanged, as long as $b_H \leq \overline{b}$, where $p_H \in b_H$.

Finally, Proposition 4 follows directly from the fact that our model assumes no extensive margin responses. Contracting officers can avoid the mandate via bunching responses, but still need to complete the purchase. We think this assumption is natural for this setting, so that the overall number of observed and counterfactual contracts needs to coincide.

#### G.1.5 Density Analysis: Estimation of Price Effects and Counterfactual Densities

We know explain our density analysis estimation method in detail, building on the Propositions of the previous section.

**Step 1**

Our method starts from the observation that, relative to ex-ante prices, linear price effects will impact the distribution of publicized contracts in two ways: (i) they will shift the full distribution to the left by $E[\gamma_t]$; and (ii) they will smooth out the discontinuity in the distribution around the threshold, because of $V(\gamma_t)$ (see Figure **??** (d)).

Suppose that we knew the true value of mean price effects $E[\gamma_t] \equiv \bar{\gamma}$. From the observed frequency distribution of publicized contracts $\{n_b^1\}$, we can simply undo the first impact of price effects by shifting this distribution back to the right. That is, we construct the *shifted* distribution $\{n_b^{1,s}(\bar{\gamma})\}$, which is obtained by adding the value of $\bar{\gamma}$ to the price award of every publicized contract. If the number of contracts is large, the shifted distribution should coincide with the

unobserved distribution of ex-ante prices $\{\tilde{n}_b^d\}$, except near the threshold.

On the other hand, a similar argument can be made for non-publicized contracts, given the assumption that bunching responses are local to the threshold (**A4**). Except for a window around the threshold where bunching responses manifest, the observed distribution $\{n_b^0\}$ should coincide with the unobserved distribution $\{\tilde{n}_b^d\}$ (see Figure **??** (c)).

This intuition is supported by Propositions 2 and 3. Once we get "far enough" from the threshold, the distribution of non-publicized awards and the *appropriately shifted* distribution of publicly solicited awards should coincide with the latent distributions of ex-ante prices. In particular, we have that: $n_b^0 + n_b^{1,s}(\bar{\gamma}) \approx \tilde{n}_b^0 + \tilde{n}_b^1 = \tilde{n}_b$ for $b$ sufficiently far from 0. On the contrary, close to the threshold we have $n_b^0 + n_b^{1,s}(\bar{\gamma}) \neq \tilde{n}_b$ due to the effects of bunching and the variance in price effects.

Finally, because we know that the unobserved distribution $\{\tilde{n}_b\}$ should be smooth everywhere due to **A1**, we can use a standard bunching estimation procedure (Chetty et al., 2013; Kleven and Waseem, 2013) to infer the shape of it around the threshold. This means fitting a polynomial function through our constructed distribution $\{n_b^0 + n_b^{1,s}(\bar{\gamma})\}$, ignoring the contribution of the bins close to the threshold.

More concretely, we estimate the following specification:

$$\left[n_b^0 + n_b^{1,s}(\widehat{\bar{\gamma}})\right] = \sum_{x=0}^{Q} \alpha_x \cdot b^x + \sum_{j=\underline{b}}^{\bar{b}} \gamma_j \cdot \mathbf{1}[b = j] + \nu_b, \qquad \text{for } b = \{-R, ..., R\} \tag{14}$$

and obtain fitted values:

$$\widehat{n_b} = \sum_{x=0}^{Q} \widehat{\alpha_x} \cdot b^x \qquad \text{for } b = \{-R, ..., R\}.$$

Now, this discussion started by *assuming* that we knew the value of the mean price effect $\bar{\gamma}$. Yet, in practice, this is the main unknown parameter that we seek to recover. So in order to estimate it, we rely on the integration constraint of Proposition 4: $\sum_{b=-R}^{R}(n_b^0 + n_b^{1,s}(\bar{\gamma})) = \sum_{b=-R}^{R} \widehat{n_b}$. As the intuition from Appendix Figure A3 shows, the integration constraints will bind only when we shift the distribution of publicized contracts according to the right value of $\bar{\gamma}$. We, therefore, start from an initial guess of $\widehat{\bar{\gamma}}$, and iterate until we find a value such that the constraint is satisfied.

For the implementation, we choose the following parameters. We use a fifth-degree polynomial, i.e. $Q = 5$. We use bins of constant width of 0.01 log-points. This implies bins of roughly \$250 at the discontinuity. Indeed, bin $b = 0$ includes all contracts with price greater than \$24,751[45] and smaller than or equal to \$25,000. Our estimation is performed on a total set of 150 bins centered around zero, from -0.75 to 0.75. In dollar terms, this corresponds to contracts between \$11,809 and \$52,925. The excluded window for step 1 is symmetric, excluding 12 bins below zero and 12 bins above. In

[45] $\log(x) - \log(25,000) = 0.01 \iff x = 25,000 \cdot \exp(-0.01)$

dollar terms, the excluded window consists of contracts between \$22,173 and \$28,187.

**Step 2**

The second step seeks to estimate separate counterfactual distributions by publicity status, i.e. $\{\widehat{\tilde{n}_b^0}\}$ and $\{\widehat{\tilde{n}_b^1}\}$. To do this, we can go back to the intuition from **??**, assuming that there are neither price effects nor bunching responses, so that the distributions of ex-ante prices and observed realized prices coincide. In this case, the distributions for treated and control units should be continuous, except at the threshold, where we should see a discontinuous jump in publicized contracts mirrored by a discontinuous dip in non-publicized contracts. Suppose that we knew the size of this change, which we denote as $\Delta$. Knowledge of $\Delta$ would allow us to undo these discontinuities by shifting the right part of each distribution vertically. Indeed, the distributions $\{n_b^0 + \Delta \cdot \mathbf{1}[b > 0]\}$ and $\{n_b^1 - \Delta \cdot \mathbf{1}[b > 0]\}$ should be continuous.

In the presence of bunching and price effects, these vertical shifts will not make the observed distributions continuous. However, just as in the discussion above, price effects and bunching should only affect the distributions within some window around the threshold. So we use this logic again and use a polynomial interpolation to estimate the counterfactual distributions around the threshold.

First, we construct distributions that are vertically shifted above the threshold: $\{n_b^0 + \Delta \cdot \mathbf{1}[b > 0]\}_{b=-R}^{R}$ and $\{n_b^{1,s}(\widehat{\tilde{\gamma}}_b) - \Delta \cdot \mathbf{1}[b > 0]\}_{b=-R}^{R}$. We then apply the same interpolation method as before for each of the two distributions. That is, we separately estimate the following two specifications:

$$\left(n_b^0 + \Delta \cdot \mathbf{1}[b > 0]\right) = \sum_{x=0}^{Q} \alpha_x^0 \cdot b^x + \sum_{j=\underline{b^0}}^{\overline{b^0}} \gamma_j^0 \cdot \mathbf{1}[b = j] + \nu_b^0, \quad \text{for } b = \{-R, ..., R\} \tag{15}$$

$$\left(n_b^{1,s}(\widehat{\tilde{\gamma}}_b) - \Delta \cdot \mathbf{1}[b > 0]\right) = \sum_{x=0}^{Q} \alpha_x^1 \cdot b^x + \sum_{j=\underline{b^1}}^{\overline{b^1}} \gamma_j^1 \cdot \mathbf{1}[b = j] + \nu_b^1, \quad \text{for } b = \{-R, ..., R\} \tag{16}$$

and compute fitted values ignoring the contribution of the bins within the excluded window:

$$\widehat{n_b^{*0}} = \sum_{x=0}^{Q} \widehat{\alpha_x^0} \cdot b^x, \quad \text{for } b = \{-R, ..., R\}$$

$$\widehat{n_b^{*1}} = \sum_{x=0}^{Q} \widehat{\alpha_x^1} \cdot b^x, \quad \text{for } b = \{-R, ..., R\}$$

Finally, our estimates of the counterfactual distributions do incorporate the discontinuous effect of the policy. We estimate these by re-adding the shift that we originally removed:

$$\widehat{n_b^0} = \widehat{n_b^{*0}} - \Delta \cdot \mathbf{1}[b > 0] \quad \text{for } b = \{-R, ..., R\}$$

$$\widehat{n_b^1} = \widehat{n_b^{*1}} + \Delta \cdot \mathbf{1}[b > 0] \quad \text{for } b = \{-R, ..., R\}$$

Again, this exposition assumes that we know the value of $\Delta$. Since, in practice, this is not directly observed, our method iterates over guesses of $\widehat{\Delta}$. The convergence criterion in this case is based on the fit of the interpolations outside the excluded window. Indeed, if the vertical shift we guess is too low or too high, the polynomial interpolation will fit poorly just outside of the excluded area. Appendix Figure A4 shows this intuition graphically.

So, given a guess of $\widehat{\Delta}$, we compute the residuals for each of the two regressions (15) and (16). We then search over $\widehat{\Delta}$ to minimize:

$$W\left(\widehat{\Delta}\right) = 0.5 \cdot \sum_{b \neq Z^0} \widehat{\nu_b^0}\left(\widehat{\Delta}\right)^2 + 0.5 \cdot \sum_{b \neq Z^1} \widehat{\nu_b^1}\left(\widehat{\Delta}\right)^2 \quad ,$$

where $Z^0 = \{\underline{b^0}, ..., \overline{b^0}\}$ and $Z^1 = \{\underline{b^1}, ..., \overline{b^1}\}$ correspond to the excluded regions.

For step two, we keep the polynomial degree fixed, binning and range fixed as in step 1. However, we change the excluded region for the specification using non-publicized contracts (15). The justification of this is that we expect bunching to be concentrated closely below the threshold. Concretely, we choose 5 bins below the threshold and 12 bins above for $Z^0$ and keep the symmetric window of 12 bins above and below for $Z^1$.

**Step 3**

In step 3 we rely on the formula from Proposition 5 and use our estimates from above to compute:

$$\widehat{F_{\gamma'}}(x) = \frac{n_{b_x}^{1,s}(\widehat{\tilde{\gamma}}) - \widehat{\tilde{n}_{b_x}^1}}{\widehat{\Delta}}$$

for $x \in b_x$, $b_x \in \{\underline{b}^1, ..., 0\}$, and $\gamma' = \gamma - \widehat{\tilde{\gamma}}$. This is straightforward given implementation of steps 1 and 2. We obtain the $F_{\gamma'}$ evaluated at each bin on the lower half of the excluded region $Z^1$. For values $x < \underline{b}^1$, we impose $F_{\gamma'} = 0$, since below the excluded region there is no longer any influence of price effects. Finally, we then obtain estimates for the rest of the CDF by imposing symmetry, so that $F_{\gamma'}(x) = 1 - F_{\gamma'}(-x)$.

For all of our estimates, we compute standard errors via bootstrap. We sample with replacement from the original distribution of contracts, and implement steps 1 through 3, obtaining a set of estimates $\hat{\theta}$. We repeat this process $H$ times. The standard errors correspond to the empirical standard deviation of $\hat{\theta}^{(h)}$, for $h = \{1, 2, ..., H\}$.

### G.1.6 RDD Correction for Price Effects and Measurement Error

Consider again the model described in Section G.1. Observed prices as a function of ex-ante prices are given by:

$$p_t = \tilde{p}_t + (1 - D_t) \cdot \xi_t + D_t \cdot \gamma_t \tag{17}$$

where $p_t$ are observed normalized (i.e. logged and re-centered around 0) award prices, $\tilde{p}_t$ are normalized ex-ante prices, $D_t \in \{0, 1\}$ are publicity decisions, $\gamma_t$ is the *price effect* of publicity, and $\xi_t$ is *measurement error*. Let $\gamma_t \sim F_\gamma(\cdot)$, with $E[\gamma_t] = \mu_\gamma$ and $V[\gamma_t] = \sigma^2_\gamma$. Let $\xi_t \sim F_\xi(\cdot)$, with $E[\xi_t] = 0$ and $V[\xi_t] = \sigma^2_\xi$. Assume $\gamma_t \perp \xi_t \perp \tilde{p}_t$.

To assess the causal impact of $D_t$ on outcomes of interest $y_t$, we assume a piece-wise linear relationship between expected outcomes and latent ex-ante prices. In particular:

$$E[y_t|\tilde{p}_t] = \mathbf{1}(\tilde{p}_t \leq 0) \cdot (\alpha_0 + \beta_0 \cdot \tilde{p}_t) + \mathbf{1}(\tilde{p}_t > 0) \cdot (\alpha_1 + \beta_1 \cdot \tilde{p}_t) \tag{18}$$

For simplicity, we focus on this reduced form relationship, but it would be straightforward to extend it to a two-equation model with a structural equation relating $y_t$ and $D_t$, and a first-stage equation relating $D_t$ and $\tilde{p}_t$. Our parameters of interest are $(\boldsymbol{\alpha}, \boldsymbol{\beta}) = (\alpha_0, \alpha_1, \beta_0, \beta_1)$. In particular, we focus on $(\alpha_1 - \alpha_0)$, the reduced form effect at the discontinuity.

The problem we face is that we do not observe a sample analog of $E[y_t|\tilde{p}_t]$, but rather of $E[y_t|p_t]$. Our "naive RDD" coefficients correspond to an estimate of $\left(\lim_{p\to 0^+} E[y_t|p] - \lim_{p\to 0^-} E[y_t|p]\right)$, which in general will not be equal to $(\alpha_1 - \alpha_0) = \left(\lim_{\tilde{p}\to 0^+} E[y_t|\tilde{p}] - \lim_{\tilde{p}\to 0^-} E[y_t|\tilde{p}]\right)$. Here we propose an alternative estimator of $(\alpha_1 - \alpha_0)$ based on the following proposition.

**Proposition 6.** *Expected outcomes conditional on observed award prices $E[y_t|p_t]$ can be expressed as an explicit linear function of the structural parameters $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, as well as other variables that we can directly observe or estimate. In particular:*

$$E[y_t|p_t] = \alpha_0 \cdot \psi_1(p_t) + \beta_0 \cdot \psi_2(p_t) + \alpha_1 \cdot \psi_3(p_t) + \beta_1 \cdot \psi_4(p_t) \quad ,$$

*where $\psi_k(\cdot)$, $k \in \{1, 2, 3, 4\}$ are explicit functions of observed prices $(p_t)$, observed treatment probabilities at a given price $(\pi_D(p_t))$, and moments of the distributions of price effects and measurement error evaluated at a given price $(F_\gamma(p_t), F_\xi(p_t))$.*

Below we derive the explicit expressions for each $\psi_k$. We then compute these using our data and the estimate $\widehat{F_\gamma}(p_t)$ that we obtained from the density analysis. We also assume no measurement error, so that $\xi_t = 0$ for all $t$. However, the formulas we derive are general, allowing for any arbitrary distribution of measurement error. Once we compute these estimates $\widehat{\psi_k}(p_t)$, we use the equation in Proposition 6 to estimate $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ by OLS. We are particularly interested in $(\hat{\alpha}^{OLS}_1 - \hat{\alpha}^{OLS}_0)$, which we then directly compare to the "naive RDD" reduced form coefficients.

### G.1.7 Proof of Proposition 6

We now derive the explicit expression for $E[y_t|p_t]$. First, we use the Law of Total Probability to write:

$$E[y_t|p_t] = \underbrace{E[y_t|p_t, \tilde{p}_t \leq 0])}_{\Lambda_1} \cdot \underbrace{\Pr(\tilde{p}_t \leq 0|p_t)}_{\Lambda_2} + \underbrace{E[y_t|p_t, \tilde{p}_t > 0]}_{\Lambda_3} \cdot \underbrace{\Pr(\tilde{p}_t > 0|p_t)}_{\Lambda_4} \tag{19}$$

For each $\Lambda_k$, $k \in \{1, 2, 3, 4\}$, we find an expression that depends only on magnitudes that we can directly observe or estimate.

We start with $\Lambda_2$:

$$\begin{aligned}
\Lambda_2 &= \Pr(\tilde{p}_t \leq 0|p_t) \\
&= \Pr(\tilde{p}_t \leq 0|p_t, D_t = 0) \cdot \Pr(D_t = 0|p_t) + \Pr(\tilde{p}_t \leq 0|p_t, D_t = 1) \cdot \Pr(D_t = 1|p_t) \\
&= \Pr(p_t - \xi_t \leq 0|p_t) \cdot [1 - \pi_D(p_t)] + \Pr(p_t - \gamma_t \leq 0|p_t, D_t = 1) \cdot \pi_D(p_t) \\
&= \left[1 - F_\xi(p_t)\right] \cdot [1 - \pi_D(p_t)] + [1 - F_\gamma(p_t)] \cdot \pi_D(p_t) \\
&\equiv \Lambda_2(p_t, \pi_D(p_t), F_\gamma(p_t), F_\xi(p_t), \boldsymbol{\alpha}, \boldsymbol{\beta})
\end{aligned} \tag{20}$$

Similarly for $\Lambda_4$:

$$\begin{aligned}
\Lambda_4 &= \Pr(\tilde{p}_t \geq 0|p_t) \\
&= \Pr(\tilde{p}_t \geq 0|p_t, D_t = 0) \cdot \Pr(D_t = 0|p_t) + \Pr(\tilde{p}_t \geq 0|p_t, D_t = 1) \cdot \Pr(D_t = 1|p_t) \\
&= \Pr(p_t - \xi_t \geq 0|p_t) \cdot [1 - \pi_D(p_t)] + \Pr(p_t - \gamma_t \geq 0|p_t, D_t = 1) \cdot \pi_D(p_t) \\
&= F_\xi(p_t) \cdot [1 - \pi_D(p_t)] + F_\gamma(p_t) \cdot \pi_D(p_t) \\
&\equiv \Lambda_4(p_t, \pi_D(p_t), F_\gamma(p_t), F_\xi(p_t), \boldsymbol{\alpha}, \boldsymbol{\beta})
\end{aligned} \tag{21}$$

For $\Lambda_1$ and $\Lambda_3$, the analysis is slightly more complicated. First, observe that:

$$\begin{aligned}
\Lambda_1 &= E[y_t|p_t, \tilde{p}_t \leq 0]) \\
&= E[\alpha_0 + \beta_0 \cdot \tilde{p}_t|p_t, \tilde{p}_t \leq 0] \\
&= \alpha_0 + \beta_0 \cdot E[\tilde{p}_t|p_t, \tilde{p}_t \leq 0] \\
&= \alpha_0 + \beta_0 \cdot \{E[\tilde{p}_t|p_t, \tilde{p}_t \leq 0, D_t = 1] \cdot \Pr(D_t = 1|p_t, \tilde{p}_t \leq 0) \\
&\qquad + E[\tilde{p}_t|p_t, \tilde{p}_t \leq 0, D_t = 0] \cdot \Pr(D_t = 0|p_t, \tilde{p}_t \leq 0)\} \\
&= \alpha_0 + \beta_0 \cdot \{E[p_t - \gamma_t|p_t, \tilde{p}_t \leq 0, D_t = 1] \cdot \Pr(D_t = 1|p_t, \tilde{p}_t \leq 0) \\
&\qquad + E[p_t - \xi_t|p_t, \tilde{p}_t \leq 0, D_t = 0] \cdot \Pr(D_t = 0|p_t, \tilde{p}_t \leq 0)\} \\
&= \alpha_0 + \beta_0 \cdot \{(p_t - E[\gamma_t|\gamma_t \geq p_t, p_t]) \cdot \Pr(D_t = 1|p_t, \tilde{p}_t \leq 0) \\
&\qquad + (p_t - E[\xi_t|\xi_t \geq p_t, p_t]) \cdot \Pr(D_t = 0|p_t, \tilde{p}_t \leq 0)\}
\end{aligned}$$

$\iff$

$$\begin{aligned}\Lambda_1 = \alpha_0 + \beta_0 \cdot p_t + \beta_0 \cdot \{E[\gamma_t|\gamma_t \geq p_t, p_t]) \cdot \Pr(D_t = 1|p_t, \tilde{p}_t \leq 0) \\ - E[\xi_t|\xi_t \geq p_t, p_t]) \cdot \Pr(D_t = 0|p_t, \tilde{p}_t \leq 0)\}\end{aligned} \tag{22}$$

Now, applying Bayes' rule to $\Pr(D_t = 0|p_t, \tilde{p}_t \leq 0)$:

$$\begin{aligned}\Pr(D_t = 0|p_t, \tilde{p}_t \leq 0) &= \frac{\Pr(\tilde{p}_t \leq 0|D_t = 0, p_t) \cdot \Pr(D_t = 0|p_t)}{\Pr(\tilde{p}_t \leq 0|p_t)} \\ &= \frac{\Pr(\tilde{p}_t \leq 0|D_t = 0, p_t) \cdot \Pr(D_t = 0|p_t)}{\Lambda_2} \\ &= \frac{\Pr(p_t - \xi \leq 0|p_t) \cdot [1 - \pi_D(p_t)]}{\Lambda_2} \\ &= \frac{\left[1 - F_\xi(p_t)\right] \cdot [1 - \pi_D(p_t)]}{\Lambda_2}\end{aligned} \tag{23}$$

And, therefore,

$$\begin{aligned}\Pr(D_t = 1|p_t, \tilde{p}_t \leq 0) &= 1 - \Pr(D_t = 0|p_t, \tilde{p}_t \leq 0) \\ &= \frac{[1 - F_\gamma(p_t)] \cdot \pi_D(p_t)}{\Lambda_2}\end{aligned} \tag{24}$$

Combining (22), (23) and (24) implies:

$$\begin{aligned}\Lambda_1 &= \alpha_0 + \beta_0 \left[p_t + \frac{E[\gamma_t|\gamma_t \geq p_t, p_t] \cdot [1 - F_\gamma(p_t)] \cdot \pi_D(p_t) - E[\xi_t|\xi_t \geq p_t, p_t] \cdot \left[1 - F_\xi(p_t)\right] \cdot [1 - \pi_D(p_t)]}{\Lambda_2}\right] \\ &\equiv \Lambda_1(p_t, \pi_D(p_t), F_\gamma(p_t), F_\xi(p_t), \boldsymbol{\alpha}, \boldsymbol{\beta})\end{aligned} \tag{25}$$

Analogous calculations yield the following expression for $\Lambda_3$:

$$\begin{aligned}\Lambda_3 &= \alpha_1 + \beta_1 \left[p_t + \frac{E[\gamma_t|\gamma_t \leq p_t, p_t] \cdot F_\gamma(p_t) \cdot \pi_D(p_t) - E\left[\xi_t|\xi_t \leq p_t, p_t\right] \cdot F_\xi(p_t) \cdot [1 - \pi_D(p_t)]}{\Lambda_4}\right] \\ &\equiv \Lambda_3(p_t, \pi_D(p_t), F_\gamma(p_t), F_\xi(p_t), \boldsymbol{\alpha}, \boldsymbol{\beta})\end{aligned} \tag{26}$$

Finally, combining (19), (20), (21), (25), and (26), we obtain:

$$E[y_t|p_t] = \alpha_0 \cdot \psi_1(p_t) + \beta_0 \cdot \psi_2(p_t) + \alpha_1 \cdot \psi_3(p_t) + \beta_1 \cdot \psi_4(p_t)$$

where:

$$\psi_1(p_t) = \left[1 - F_\xi(p_t)\right] \cdot [1 - \pi_D(p_t)] + [1 - F_\gamma(p_t)] \cdot \pi_D(p_t)$$
$$\psi_2(p_t) = \psi_1(p_t) \cdot p_t + E[\gamma_t | \gamma_t \geq p_t, p_t] \cdot [1 - F_\gamma(p_t)] \cdot \pi_D(p_t) - E[\xi_t | \xi_t \geq p_t, p_t] \cdot \left[1 - F_\xi(p_t)\right] \cdot [1 - \pi_D(p_t)]$$
$$\psi_3(p_t) = F_\xi(p_t) \cdot [1 - \pi_D(p_t)] + F_\gamma(p_t) \cdot \pi_D(p_t)$$
$$\psi_4(p_t) = \psi_3(p_t) \cdot p_t + E[\gamma_t | \gamma_t \leq p_t, p_t] \cdot F_\gamma(p_t) \cdot \pi_D(p_t) - E\left[\xi_t | \xi_t \leq p_t, p_t\right] \cdot F_\xi(p_t) \cdot [1 - \pi_D(p_t)]$$

#### G.1.8 Accounting for Bunching

A standard test for the validity of the RDD framework consists on verifying the continuity of the density of the running variable around the threshold. If the running variable is not distributed smoothly around the cutoff, then it is said to be "manipulated". In recent work, Gerard, Rokkanen, and Rothe (2020) show that, while point identification of causal effects is infeasible in this case, it is possible to obtain sharp bounds on the effects of interest.

In their model, the extent of manipulation can be quantified as the excess bunching in the density of the running variable below the threshold. While one cannot identify which are the units below the threshold that are manipulating, the excess bunching $\pi_B$ tells us what share of the observed units are in this group. Bounds on treatment effects are then computed by excluding a share $\pi_B$ of the observations below the threshold, in ways that yield the most extreme values for the estimate.

This process can be quite involved in general, since one does not know the treatment assignment of the units that manipulate. This transforms the computation of the bounds in an optimization problem, searching for the worst- and best-case scenarios in terms of how outcomes are distributed across treatment groups below the threshold.

However, our setting allows us to make a behavioral assumption that tremendously simplifies the problem. In particular, our model assumes that all units that manipulate the ex-ante price to bunch below the threshold, successfully avoid the publicity mandate.[46] Therefore, our model implies that the share $\pi_B$ of units that manipulate all belong to the control group ($D_t = 0$). Bounds on treatment effects are straightforwardly obtained in this case, by simply chopping the tails of the distribution of outcomes $Y_t$ below the threshold for units in the control group.

In practice, we implement this procedure as follows. For each bin $b$ closely below the threshold:

1. Compute the excess bunching in the control group, as $BUNCH_b = (n_b^0 - \widehat{\tilde{n}_b^0})$, obtained from our density analysis.
2. Sort control units according to the outcome variable $Y_b^0$.
3. Drop the $BUNCH_b$ units with the highest value of $Y_b^0$. Compute treatment effects. This yields the lower bound.

[46] This corresponds to a special case of their more general model. The authors explicitly discuss this special case in their Appendix C.3.

4. Drop the $BUNCH_b$ units with the lowest value of $Y_b^0$. Compute treatment effects. This yields the upper bound.