Using Bid Rotation and Incumbency to Detect Collusion: A Regression Discontinuity Approach

Kei Kawai*
UC Berkeley

Jun Nakabayashi
Kindai University

Juan Ortner
Boston University

Sylvain Chassang
Princeton University

December 21, 2021

Abstract

Cartels participating in procurement auctions frequently use bid rotation or prioritize incumbents to allocate contracts. However, establishing a link between observed allocation patterns and firm conduct has been difficult: there are cost-based competitive explanations for such patterns. We show that by focusing on auctions in which the winning and losing bids are very close, it is possible to distinguish allocation patterns reflecting cost differences across firms from patterns reflecting non-competitive environments. We apply our tests to two datasets: the sample of Ohio milk auctions studied in Porter and Zona (1999), and a sample of municipal procurement auctions from Japan.

Keywords: procurement, collusion, backlog, incumbency, regression discontinuity, antitrust.

*Corresponding author: kei@berkeley.edu. We thank Frank Wolak for helpful comments on an early draft. We also thank seminar audiences at Bonn, Boston University, Duke, Humboldt University in Berlin, Northwestern, Stanford, the University of Chicago, UPenn, the 2019 Berkeley-Paris Organizational Economics Workshop, the 2019 UCLA IO Miniconference and UCSD, for helpful comments.

1 Introduction

The ability of competition authorities to proactively detect and punish collusion is crucial for achieving the goal of promoting and maintaining competition. Not only do the possibility of detection and prosecution serve as strong deterrents against collusion, they also affect the incentives of firms in existing cartels to apply for leniency programs. Successful identification of cartels thus deters collusive activity and complements enforcement programs.

In the absence of concrete leads, using data-driven screens to flag suspicious firm conduct can be useful for regulators as a first step in identifying collusion. While screens cannot substitute for direct evidence of collusion, such as testimonies and records of communication, they can provide guidance on which markets or firms to focus investigation. A growing number of countries are adopting algorithm-based screens that analyze bidding data from public procurement auctions to flag suspicious behavior.[1] More recently, the U.S. Department of Justice announced the formation of a procurement collusion strike force whose goal includes bolstering "data analytics employment to identify signs of potential anticompetitive, criminal collusion."[2] Imhof et al. (2018) describes an antitrust investigation initiated on the basis of statistical screens and resulting in successful cartel prosecution. The results from screens can be used in court to obtain warrants, or to support civil antitrust litigation as well as private litigation.[3]

Screening cartels can also be useful to stakeholders other than antitrust authorities. For example, screening can help procurement offices counter suspected bidding rings by soliciting new bidders more aggressively or adopting auction mechanisms that are less susceptible to collusion. In large decentralized organizations, collusion may be organized by firm employees against the will of senior management (Sonnenfeld and Lawrence, 1978).[4] In that context,

[1] A report by the OECD (OECD, 2018) gives a brief description of the screening programs used in Brazil, Switzerland and the UK.

[2] Announcement of the Antitrust Division's Procurement Collusion Strike Force, November 22, 2019.

[3] Baker and Rubinfeld (1999) give an overview on the use of statistical evidence in court for antitrust litigation.

[4] See also Ashton and Pressey (2012), who study 56 international cartels investigated by the EU. They find that there is involvement of individuals at the most senior levels of management (CEOs, chairpersons, etc.) in about half of those cases.

screening tools can help internal auditors and compliance officers contain collusive practices initiated by employees.

Because bidding rings often adopt rotation schemes or give priority to incumbents in project allocation, bid rotation and incumbency advantage are very often suggested as indicators of collusion.[5] However, it is well known that there are non-collusive cost-based explanations for these allocation patterns. Bid rotation can arise under competition if marginal costs increase with backlog. Incumbency advantage can be explained by cost asymmetries among competitive firms or by learning-by-doing. Hence, establishing a tight link between these bidding patterns and collusion has been difficult. As Porter (2005) describes, "An empirical challenge is to develop tests that can discriminate between collusive and non-cooperative explanations for rotation or incumbency patterns."

We show that it is possible to discriminate between competitive and non-competitive bid rotation and incumbency patterns using the logic of regression discontinuity design (Thistlethwaite and Campbell, 1960). We compare the backlog and incumbency status of a bidder who wins the auction by a small margin to those of a bidder who loses by a small margin. Although bids are endogenous, we show that under an appropriate notion of competition, the probability that a bidder wins or loses an auction conditional on close bids approaches 50%, regardless of the bidders' characteristics (e.g., the size of backlog, incumbency status, etc). This implies that a firm with high levels of backlog will be a marginal winner half of the time and a marginal loser in the other half, and similarly for a firm with low levels of backlog. In fact, this is true for a firm with any level of backlog or incumbency status. Winning and losing are "as-if-random" conditional on close bids. As a result, under competition, differences in backlog or incumbency status between close winners and close losers should vanish as the bid difference between them approaches zero even if these variables are correlated with costs. If instead, bids are generated by collusive bidding, the differences in these variables between close winners and close losers need not disappear. For example, if the bidding ring always allocates projects to the incumbent bidder, close win-

[5] See, for example, the "Red Flags Of Collusion" report, published by the U.S. DOJ or the pamphlet on bid rigging published by the Canadian Competition Bureau.

ners will be incumbents with significantly higher probability than close losers. Our tests of non-competitive behavior seek to detect discontinuities in the distribution of economically relevant covariates around close winners and close losers.

We illustrate our test using two datasets. First, we consider the sample of Ohio school milk auctions studied by Porter and Zona (1999). Firms located around Cincinnati, Ohio were charged with colluding on hundreds of school milk auctions by allocating markets according to incumbency status (*State of Ohio* v. *Louis Trauth Dairies, Inc. et al*). According to the testimony of the representatives of the colluding dairies, the firms colluded by agreeing not to undercut the bid of the incumbent firm that had served a given school district in the previous year. We test whether or not marginal winners are more likely to be incumbents than marginal losers separately for the set of collusive auctions and the set of non-collusive auctions.[6] We find that for collusive auctions, marginal winners are significantly more likely to be incumbents than marginal losers, rejecting the null of competition. In contrast, we do not find statistically significant differences in incumbency status between marginal winners and marginal losers among non-collusive auctions despite the fact that the sample size is more than 10 times bigger.

Second, we apply our tests to a dataset of public procurement auctions held by municipalities from the Tohoku region of Japan. Firms in this dataset have not been prosecuted for collusion, but there are reasons to suspect that collusion is present in this dataset. Kawai and Nakabayashi (2018) provide evidence that some of the firms in this dataset colluded over procurement contracts let by the Ministry of Land, Infrastructure and Transportation. Chassang et al. (2020) suggest that non-competitive behavior may have been prevalent in auctions held by a different set of Tohoku municipalities.

We first apply our test to the full sample of municipal auctions. As in the case of collusive Ohio school milk auctions, we find that marginal winners are significantly more likely to be incumbents than marginal losers. We also find that backlog is significantly lower for marginal winners compared to marginal losers. These findings suggest the presence of

[6] More precisely, we apply the tests separately for the set of auctions in which all of the bidders were implicated and the set of auctions in which none of the participants were implicated.

collusive agreements using both priority to incumbents and bid rotation. We then split the sample of municipal bids into high and low bid groups depending on whether the bid is above or below the median winning bid for the municipality letting the auction.[7] Because the primary purpose of collusion is to elevate prices, we expect collusion to be more prevalent in the high bid sample than in the low bid sample. Indeed, for the high bid sample, we find that marginal winners are more likely to be incumbents and have on average lower backlog, compared to marginal losers. We find much smaller, and mostly insignificant differences in the characteristics of marginal winners and marginal losers for the low bid sample. These findings suggest that our test is able to discriminate between competition and collusion.

Literature. Our work fits in the industrial organization literature interested in detecting collusion in auctions and markets. Pioneering work in this literature include Hendricks and Porter (1988), Baldwin et al. (1997), as well as Porter and Zona (1993, 1999). Our contribution is particularly related to Porter and Zona (1993) who study the impact of cost shifters such as backlog and proximity to construction sites on the bids and rank order of bidders in road pavement auctions. They find that the losing bids of suspected ring members do not respond to cost shifters, suggesting that those bids are likely to be phantom bids. Although both Porter and Zona (1993) and our paper study the relationship between the rank order of bids and possible cost shifters to screen for collusion, the underlying idea behind the proposed tests are quite different. Porter and Zona (1993) focus on the lack of incentives among losing cartel bidders to bid in ways that reflect their true costs. Hence, their primary focus is on losing bidders.[8] Our primary focus is on differences between winners and losers. The tests we propose are based on the idea that under collusion, close winners and losers need not be statistically similar: under incumbency priority close winners are more likely to be incumbents; under bid rotation close winners are likely to have lower backlog.[9] The tests

[7]We normalize raw bids with each auction's reserve price to make bids more comparable.

[8]Porter and Zona (1993) describe their tests as follows: "... our rank-based test is designed to detect differences in the ordering of higher bids, as opposed to the determinants of the probability of being the lowest bid ...", although parts of their paper analyze the determinants of being the winner.

[9]Interestingly, Marshall and Marx (2007) shows that optimal collusive schemes may require designated losers to bid marginally above designated winners.

of Porter and Zona (1993) and ours are complementary.

Recent work seeking to detect non-competitive behavior includes Bajari and Ye (2003), Ishii (2009), Athey et al. (2011), Conley and Decarolis (2016), Andreyanov (2017), Schurter (2017), Kaplan et al. (2017), and Kawai and Nakabayashi (2018).[10] A complementary literature focuses on known cartels and studies the practical details of collusive arrangements. Pesendorfer (2000) studies bidding rings with and without side-payments. Asker (2010) studies knockout auctions among members of a bidding ring. Clark et al. (2018) analyze the breakdown of a cartel and its price implications. Other contributions (see for instance Ohashi (2009), or Chassang and Ortner (2019)) take a design perspective and document how changes in the auction format affect the ability of bidders to maintain collusion.

The paper shares its emphasis on general information structures with Chassang et al. (2020), but develops a qualitatively different strategy that considerably expands the scope for applications. In previous work, we document that in a significant subset of procurement auctions held in Japan, winning bids are isolated – there are very few close winners and close losers. This pattern, as well as others, can be exploited to obtain lower bounds on the share of non-competitive histories under general information structures. The current paper complements this previous work by focusing on settings where the missing-bids pattern does not arise, i.e. when close wining and losing bids are not rare. Moreover, the tests that we propose in this paper exploit observable bidder characteristics, like incumbency status or backlog. This allows us to extend the analysis to environments with intertemporal linkages such as learning by doing, or increasing marginal costs, which Chassang et al. (2020) excludes. Our framework also lets analysts use any available covariate data to test for collusion, expanding on the study of Kawai and Nakabayashi (2018). Kawai and Nakabayashi (2018) focus on the identity of the lowest bidder when there are multiple rounds of bidding. The current paper applies even when there is no rebidding.

We believe that the tests proposed in this paper are well suited to complement standard antitrust practice, as a tool to target agency attention and effort, or to justify more invasive evidence collection. First, our test formalizes intuitive ideas often mentioned by antitrust

[10]For a survey of the literature up to the mid 2000s, see Porter (2005) and Harrington (2008).

agencies. Second, the test is easy to implement and requires no sophisticated programming. Third, our approach does not require detailed data on project or bidder characteristics because the regression discontinuity design makes it less important to control for auction and bidder heterogeneity. Fourth, our approach naturally extends to other types of auctions such as handicap auctions, scoring auctions and all-pay auctions by appropriately modifying the running variable.[11] Finally, our approach can be easily adapted to exploit other markers of collusion. Imagine a cartel is suspected of using geographic segmentation to allocate projects.[12] With data on the location of firms and project sites, one could assess whether or not close winners are located nearer the project site than close losers. Another possible marker of collusion is the extent of subcontracting and joint bidding.[13] If procurement agencies require the list of subcontractors to be specified at the time of bidding, one can test whether or not marginal winners have more subcontractors than marginal losers.[14] If designated losers of bidding rings do not bother contacting subcontractors for projects that they know they will lose, marginal losers may have significantly fewer subcontractors than marginal winners.

2 Framework

The section specifies our model of dynamic procurement. We describe our test of non-competitive behavior in Section 3, and provide theoretical foundations in Section 4. We turn to data in Section 5.

[11] For instance, one can apply our test to scoring auctions by focusing on auctions in which the first and second scores are close.

[12] Pesendorfer (2000) documents evidence of market division among school milk providers in Texas.

[13] For example, the Department of Justice maintains a document called "Price Fixing, Bid Rigging, and Market Allocation Schemes: What They Are and What to Look For" , in which it states "Subcontracting arrangements are often part of a bid-rigging scheme." Similar statements are found in a report by the OECD (2013). See also Conley and Decarolis (2016) for a discussion of subcontracting and collusion.

[14] For example, "Subletting and Subcontracting Fair Practices Act" (Public Contract Code 4100 et seq.) of California requires that "any person making a bid or offer to perform the work, shall, in his or her bid or offer, set forth ... (T)he name, the location of the place of business, ... of each subcontractor who will perform work or labor or render service to the prime contractor."

Game form. In each period $t \in \mathbb{N}$, a buyer procures a single item from a finite set N of potential suppliers. The procurement contract is allocated through a sealed-bid first-price auction with a public reserve price r, which we normalize to 1. Each potential bidder $i \in N$ decides whether or not to participate in each auction. Bidders incur a cost $k > 0$ for submitting an actual bid $b_{i,t} \in [0,1]$, and may prefer not to participate. Non-participation is denoted by $b_{i,t} = \emptyset$.

We denote by $\mathbf{b}_t = (b_{i,t})_{i\in N}$ the profile of bids, and by $\wedge \mathbf{b}_t$ the lowest bid among participating bidders.[15] This is the winning bid. Ties are broken with uniform probability. We denote by $\mathbf{b}_{-i,t} \equiv (b_{j,t})_{j\neq i}$ bids from firms other than i, and by $\wedge \mathbf{b}_{-i,t} \equiv \min_{j\neq i} b_{j,t}$ the lowest bid among i's participating competitors. Let $\wedge \mathbf{b}_{-i} \succ b_i$ denote the event that bidder i wins the contract, i.e. b_i is the lowest bid and possible ties are broken in favor of bidder i. Bids are publicly revealed at the end of each period.

State transitions. In each period t, before bidding, each bidder observes a state $\theta_t \in \Theta$, with Θ finite, summarizing the state of the industry. The state follows an endogenous Markov chain: θ_{t+1} is distributed according to a probability distribution $F_\Theta(\cdot|\theta_t, w_t^*)$, depending only on the previous state θ_t, and the identity of the winning bidder $w_t^* \in \arg\min_{i\in N} b_{i,t}$

Our model allows for settings in which a bidder's procurement costs depend on backlog or incumbency status through state variable θ_t. For example, θ_t can be a vector that tracks how many auctions each bidder has recently won to capture the effect of backlog on costs. Alternatively, θ_t can be a vector that tracks whether or not a given bidder has won a particular type of auction to capture the effect of learning-by-doing. State θ_t can also capture exogenous auction characteristics such as the distance between the project site and each of the bidders, the scale of the project, or the type of work being procured. Because we do not assume that θ is observed to the econometrician, θ captures both observed and unobserved auction heterogeneity.

[15] If no bidder participates, i.e. $b_i = \emptyset$ for all bidders, then by convention $\wedge \mathbf{b}_t = +\infty$.

Information. In addition to state θ_t, each bidder i privately observes a signal $z_{i,t} \in Z_i$, with Z_i finite. The distribution of signal profile $\mathbf{z}_t = (z_{i,t})_{i \in N} \in Z = \prod_{i \in N} Z_i$ depends only on θ_t but is otherwise unrestricted. Signals may be arbitrarily correlated. We denote by $F_Z(\cdot|\theta)$ the distribution of signals conditional on state θ.

Costs $\mathbf{c}_t = (c_{i,t})_{i \in N} \in \mathbb{R}^N$ are drawn independently conditional on state θ_t, and on each private signal $z_{i,t}$. In particular, we have that

$$c_{i,t}|\theta_t, z_{i,t} \sim c_{i,t}|\theta_t, \mathbf{z}_t, \mathbf{c}_{-i,t}.$$

Bidder i's cost does not provide information about the cost of other bidders beyond the information already provided in state θ_t and private signal $z_{i,t}$. We assume private values, so that each bidder observes her own costs.[16] This class of information structures nests asymmetric independent private values, correlated values, and complete information. We denote by $F_C(\cdot|\theta_t, \mathbf{z}_t)$ the conditional distribution of the profile of costs $\mathbf{c}_t$ given state θ_t, and signals $\mathbf{z}_t$.

The underlying economic environment, denoted by $\mathcal{E}$, corresponds to the tuple $\mathcal{E} = (F_\Theta, F_Z, F_C)$. Overall, firm i's profits at period t are $(b_{i,t} - c_{i,t})\mathbf{1}_{\wedge \mathbf{b}_{-i,t} \succ b_{i,t}} - k\mathbf{1}_{b_{i,t} \in [0,1]}$. Firms discount future payoffs using common discount factor $\delta < 1$.

Observables. We now introduce variables observed by the econometrician. We denote by $x_{i,t} \in X \subset \mathbb{R}^n$, with X finite, the characteristics of bidder i at time t that the econometrician observes. The observables at time t, $x_t = (x_{i,t})_{i \in N}$, can be a subset of θ_t, a coarsening of θ_t, or any variable that is predetermined at the beginning of the period.[17] In our application, $x_{i,t}$ corresponds to measures of a bidder's backlog or incumbency status. Given that bidders in our data work on projects that are not in our dataset (e.g., construction work for other firms), our measures of backlog and incumbency are likely to be imperfect measures of the

[16] Because the signals are allowed to be correlated, $z_{i,t}$ helps bidder i predict the cost of other bidders. The main restriction is that set Z is finite. This ensures that pointwise convergence results established later on hold uniformly over histories.

[17] More generally, $x_{i,t}$ can be any garbling (in the sense of Blackwell (1953)) of bidder i's information at the time of bidding in period t.

backlog and incumbency status that are relevant for bidders' costs. Because observables x can be arbitrarily noisy statistics of θ_t, the actual state variables that matter to the bidders, our framework allows for unobserved heterogeneity and measurement error.

Strategies and solution concepts. Throughout the paper, we focus on Markov strategies and Markov perfect equilibrium (Maskin and Tirole, 2001). A Markov strategy σ_i is a mapping from information $h_{i,t} = (\theta_t, z_{i,t})$ and costs $c_{i,t}$ to bids $b_{i,t} \in [0,1] \cup \{\emptyset\}$. A strategy profile $\sigma = (\sigma_i)_{i \in N}$ is a Markov perfect equilibrium (MPE) if it is a perfect Bayesian equilibrium in Markov strategies. MPE has received much attention in the empirical industrial organization literature studying dynamic oligopolistic competition, starting with Ericson and Pakes (1995). While MPE rules out collusive strategies when the state evolves exogenously, MPE is not a sufficient condition for competition in an environment with endogenous payoff relevant states: it is possible to sustain obviously collusive strategies in MPE. In Section 4 we discuss additional restrictions we impose on MPEs to capture the notion of competition in a dynamic setting.

3 Empirical Strategy

We now delineate our tests of non-competitive behavior and clarify the goal of our theoretical analysis.

Consider the problem of assessing whether or not firms in a given industry are engaging in collusive bid rotation. Empirically, this implies that bidders with low levels of backlog (firms that have not won many auctions in the recent past) are more likely to win than bidders with high levels of backlog. The difficulty is that there may also be competitive reasons for this pattern. Suppose that firms' procurement costs are increasing with backlog. Even if firms are competitive, on average, firms with lower backlog will have lower costs and be more likely to win an auction than firms with higher backlog. In this environment, a test seeking to detect collusive bid rotation by comparing the unconditional backlog of winners and losers would yield false positives.

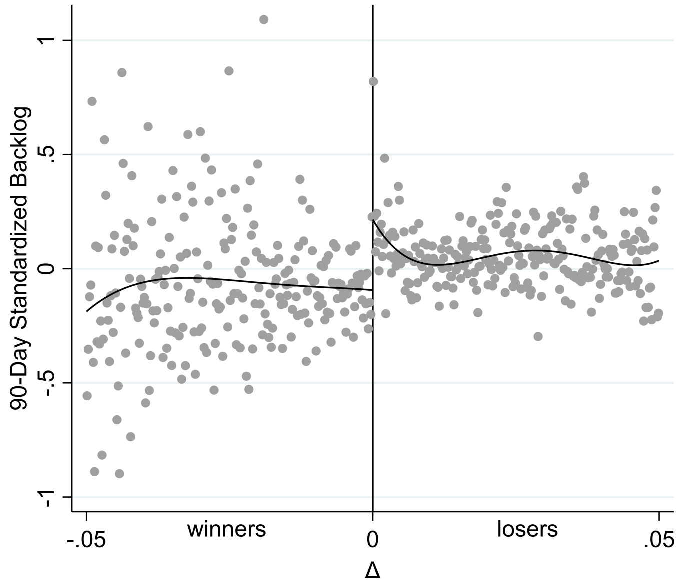
Our proposal is to compare the backlog of a selected group of firms: bidders that win or lose by a small margin. Intuitively, no firm can consistently be a marginal winner or a marginal loser under competition. Winning or losing should be as-if-random conditional on close bids. As a result, close winners and losers should be statistically similar. If instead, close winners have consistently lower levels of backlog than close losers, this is evidence of collusive bid rotation.

We operationalize this idea as follows. Denote by $\Delta_{i,t} \equiv b_{i,t} - \wedge \mathbf{b}_{-i,t}$ the difference between the bid of firm i, and the most competitive alternative bid at time t. If $\Delta_{i,t} < 0$, bidder i wins the auction, if $\Delta_{i,t} > 0$, bidder i loses. Let $x_{i,t}$ be a measure (observed by the econometrician) of firm i's backlog before bidding at time t (alternatively it could be incumbency, or another relevant covariate). We define coefficient β as the difference in average backlog between close losers and close winners:

$$\beta = \lim_{\epsilon \searrow 0^+} \mathbb{E}[x_{i,t} | \, \Delta_{i,t} = \epsilon] - \lim_{\epsilon \nearrow 0^-} \mathbb{E}[x_{i,t} | \, \Delta_{i,t} = \epsilon]. \tag{1}$$

We test the null of $\beta = 0$. When x denotes backlog, we expect β to be strictly positive under bid rotation. When x denotes incumbency status, we expect β to be strictly negative if the cartel allocates market shares according to incumbency. Figure 1 foreshadows the results of Section 5 using a dataset of Japanese procurement auctions. The figure is a binned scatter plot that illustrates the relationship between bidder i's 90-day backlog at time t against $\Delta_{i,t}$ for all i and t for the subset of auctions with above-median bids. The null of $\beta = 0$ is rejected: the average backlog is discontinuous around $\Delta_{i,t} = 0$. Close winners have a significantly lower backlog than close losers.

A heuristic motivation. Constructing a test of competition based on a test of $\beta = 0$ presumes that, under competition, allocation is as-if random conditional on close bids. If this presumption is true, a rejection of the null of $\beta = 0$ implies rejection of competition. In order to gain intuition for why this presumption is reasonable, consider the case in which a bidder's demand conditional on information $h_{i,t} = (\theta_t, z_{i,t})$ is sufficiently smooth.



Note: For each firm i and auction t, the standardized backlog of firm i at t is the Yen denominated amount of work it won in the 90 days prior to auction t, re-expressed in units of standard deviation from the firm's time-series average. The figure is a binned scatter plot of this measure against $\Delta_{i,t}$. See Section 5 for details.

Figure 1: Binned Scatter Plot of Standardized Backlog, Japanese Municipal Auctions.

For all histories $h_{i,t} = (\theta_t, z_{i,t})$ and bids $b \in [0, 1]$, define bidder i's *residual demand* as

$$D_i(b|h_{i,t}) \equiv \text{prob}(\wedge \mathbf{b}_{-i} \succ b|h_{i,t}).$$

$D_i(b|h_{i,t})$ is the probability with which firm i expects to win the auction at history $h_{i,t}$ if she places bid b. The probability that bidder i wins conditional on submitting a close bid satisfies

$$\mathsf{prob}(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| \leq \epsilon) = \frac{D_i(b_{i,t}|h_{i,t}) - D_i(b_{i,t} + \epsilon|h_{i,t})}{D_i(b_{i,t} - \epsilon|h_{i,t}) - D_i(b_{i,t} + \epsilon|h_{i,t})}. \tag{2}$$

It follows that whenever D_i is strictly decreasing and continuously differentiable, then for a bid-difference ϵ small, the probability of winning conditional on close winning and losing bids is approximately 1/2, regardless of history $h_{i,t}$. In other words, winning and losing are as-if random. This is a straightforward consequence of the fact that the numerator on the right-hand side of (2) is approximately $\epsilon D_i'(b_{i,t}|h_{i,t})$ and the denominator is approximately $2\epsilon D_i'(b_{i,t}|h_{i,t})$. Hence, the following result holds (all proofs are in Appendix C).

Lemma 1 (smooth demand)**.** *Assume that $D_i(\cdot|h_{i,t})$ is differentiable, with $D_i'(b_i|h_{i,t})$ strictly negative and continuous in bids $b_i \in [0,1]$. For all $\eta > 0$, there exists $\epsilon > 0$ small enough such that for all histories $h_{i,t}$,*

$$\left|\mathsf{prob}(i \text{ } wins \mid h_{i,t} \text{ } and \text{ } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| \leq \epsilon) - \frac{1}{2}\right| \leq \eta. \tag{3}$$

Lemma 1 implies the following corollary.

Corollary 1. *For all $\eta > 0$, there exists $\epsilon > 0$ small enough such that for all $x \in X$,*

$$|\mathit{prob}\,(x_{i,t} = x \mid \Delta_{i,t} \in (0,\epsilon)) - \mathit{prob}\,(x_{i,t} = x \mid \Delta_{i,t} \in (-\epsilon, 0))| < \eta.$$

In words, the distribution of covariates $x_{i,t}$ observable to the econometrician has to be the same for marginal winners and marginal losers.[18] Whenever X is finite, Corollary 1 implies that the expectation of $x_{i,t}$ conditional on Δ must be continuous around $\Delta = 0$. This is not true in the data illustrated by Figure 1.

We note that while the choice of which covariates x to consider may be driven by a particular theory of collusion, tests built around Corollary 1 are really testing the null of competition. For this reason we don't need to formalize a specific model of collusion. In fact, data that fails such tests may reflect a multiplicity collusive arrangements: some firms may be colluding using bid rotation, while some split local markets according to incumbency.

Why formal foundations are important. The foregoing discussion is based on the premise that demand is smooth. While smooth demand is a feature of competitive equilibrium in some environments (e.g., cost distributions are smooth and there are no dynamics),

[18] In addition, the result continues to hold when we condition on any information that is available to bidders ahead of bidding. As a result, our tests can be applied to subsets of data adapted to bidders' information in the sense of Chassang et al. (2020). In Section 5.2 we leverage this result and apply our tests separately for bids above and below the median winning bid.

and can be justified if bidders bid with small mistakes or trembles (e.g., Quantile Response Equilibrium of McKelvey and Palfrey, 1995), there exist competitive environments in which smooth demand fails. For example, suppose that costs $\mathbf{c}_t$ are public information and bidding cost k is zero. The residual demand faced by bidders conditional on their information is not smooth in this case (see Appendix A for details). The purpose of the next section (Section 4) is to delineate formally the conditions under which Corollary 1 holds. In particular, we show that if bidding is costly ($k > 0$), Corollary 1 holds under a suitable notion of competition.

4 Theoretical Foundations

In this section we provide theoretical foundations for the hypothesis that assignment conditional on close bids should be as-if random under competition. We begin with a discussion of appropriate equilibrium concepts that capture competitive behavior in dynamic auctions in which past behavior and future payoffs are linked. We then show that, under our proposed notion of competitive behavior, assignment conditional on close bids is as-if random. We exploit incentive compatibility constraints specific to first-price auctions to establish this result.

4.1 Competition

We start with the observation that an appropriate notion of competition depends on the underlying economic environment $\mathcal{E}$. For instance, while Markov perfect equilibrium (Maskin and Tirole, 2001) may seem intuitively competitive when the payoff-relevant state θ_t evolves exogenously, this is no longer the case when the payoff-relevant state depends on past actions, as the following example illustrates.

A simple collusive MPE. Consider a special case of our model with three bidders, $N = \{1, 2, 3\}$, who at each time t share the same publicly observed procurement cost: $c_{i,t} = c(\theta_t)$ for all i, t, with $\max_{\theta \in \Theta} c(\theta) < r$. The state θ_t keeps track of: (i) the winner at each of the last three periods; (ii) whether or not a bidder in the past won twice within a three period window; and (iii) whether or not there was an auction in the past without a winner. Bidders incur a small bid preparation cost $k > 0$.

For values of discount factor δ close enough to 1, the following strategy profile constitutes a Markov perfect equilibrium. On the equilibrium path, at each time t the winner at $t-3$ places a bid of $r = 1$, and the other two bidders don't participate (i.e., they bid $\emptyset$). If at any point in the past a bidder won twice in a three-period window, or there was an auction without a winner, bidders revert to static Nash.

This example illustrates that MPE may be uncompetitive in settings with endogenous states. However, this example does not show that Corollary 1 can fail under MPE. Indeed, on path there are no close winners and close losers under this equilibrium. In Appendix A we present an example of a collusive MPE that fails Corollary 1.

Continuation values. Because MPE does not rule out collusive behavior in our setting, we place additional restrictions on MPE to formulate a notion of competition. The restrictions we impose are on the bidders' continuation values.

Consider an environment $\mathcal{E}$ and an MPE σ. For any $i, j \in N$, and state $\theta \in \Theta$, let $W_i(\theta, j)$ denote bidder i's expected continuation value conditional on state θ, and winner j:

$$W_i(\theta, j) = \mathbb{E}_{\mathcal{E},\sigma}\left[\sum_{t=1} \delta^{t-1}\pi_{i,t}\middle|\theta_0 = \theta, w_0^* = j\right],$$

where $\pi_{i,t}$ denotes the profits of bidder i in period t (net of bidding costs).

Let us denote by $\zeta_i \in \{0, 1\}$ bidder i's outcome in the auction (where $\zeta_i = 1$ denotes winning the auction). For any history $h_i = (\theta, z_i)$, winning bid b_w and allocation outcome ζ_i for bidder i, let

$$V_i(\zeta_i, b_w|h_i) \equiv \mathbb{E}_{\mathcal{E},\sigma}[W_i(\theta, w^*)|\mathbf{1}_{b_i \prec \wedge \mathbf{b}_{-i}} = \zeta_i, \wedge\mathbf{b} = b_w, h_i].$$

Value V_i is the expected continuation value of player i depending on whether or not she wins the auction, and the winning bid.

Remark. *Conditional on winning, bidder i's continuation value $V_i(1, b_i|h_i)$ does not depend on her own bid b_i.*

This is driven by the fact that: (i) V_i controls for the current state through history h_i; and (ii) state transitions depend only on the current state and the identity of the winner. In general, continuation values upon losing may depend on the winning bid, since the winning bid may be correlated with the identity of the winning bidder.

For any pair of bids b, b' with $b' \geq b$ and any history h_i, let

$$v_i(0, b, b'|h_i) \equiv \mathbb{E}_{\mathcal{E},\sigma}[V_i(0, \wedge \mathbf{b}_{-i}|h_i)|h_i, b \leq \wedge \mathbf{b}_{-i} \leq b']$$

denote firm i's expected continuation value conditional on losing and on the winning bid being in $[b, b']$.

Competitively enforced equilibria. We propose a property on the bidders' continuation values (sensitivity) that is motivated by strategy profiles that are obviously collusive. We define MPEs that do not exhibit this property as competitive.

Definition 1. *We say that bidding behavior is* sensitive *if there exists h_i such that expected continuation value $v_i(0, b, b'|h_i)$ is not Lipschitz continuous in b, b'.*

When bidder behavior is sensitive, small changes in others' bids have a disproportionate effect on a losing bidder's continuation value. Bidding behavior that is sensitive is inherently suspicious because it suggests that equilibrium is sustained by the threat of significantly lower continuation values. Bidding behavior that is sensitive also goes against the idea that "minor causes should have minor effects," emphasized by Maskin and Tirole (2001) as a rationale for MPE.

We note that bidding behavior in the example above is sensitive: along the path of play, the continuation value of a designated loser falls discretely if the other designated loser undercuts the winning bid. Indeed, if bidder i is a designated loser at h_i, her continuation value $v_i(0, r-\epsilon, r-\epsilon|h_i)$ when the other designated loser deviates and bids $r-\epsilon$ is significantly lower than her on-path continuation value $v_i(0, r, r|h_i)$.

We now present our notion of competitive behavior. We define an MPE to be competitive if bidding behavior is not sensitive.

Definition 2. *We say that a Markov perfect equilibrium σ is* competitively enforced *if bidding behavior under σ is not sensitive.*

We stress that all MPEs are competitively enforced when either (i) the state θ_t evolves exogenously (i.e., $\theta_{t+1} \sim F_\Theta(\theta_t)$); or (ii) the market is large, so bidders do not interact repeatedly with each other. Indeed, in either of these cases, bidders' continuation payoffs conditional on losing $V_i(0, b_w|h_i)$ do not depend on the value of the winning bid, b_w (and hence $v_i(0, b, b'|h_i)$ is Lipschitz continuous in b and b'). We note that these are precisely the environments in which MPEs are intuitively competitive. The definition of competition that we propose strikes a balance between being permissive enough to accommodate competitive environments while at the same time being able to rule out obviously collusive conduct.

4.2 Equilibrium beliefs conditional on close bids

We now show that under a competitively enforced MPE, contract allocation conditional on close bids is as-if random. For the results that follow, we maintain the assumption that bidders incur a strictly positive participation cost (i.e., $k > 0$). This implies that competitive bidders do not participate if they expect to lose with probability close to 1.

Our first main result establishes that conditional on being a close winner or loser, any bidder believes that they win with probability greater than 50%.

Proposition 1 (equilibrium beliefs conditional on close bids)**.** *Consider an environment $\mathcal{E}$ and an MPE σ that is competitively enforced. For all $\eta > 0$ there exists $\epsilon > 0$ small enough such that, for all histories $h_{i,t} = (\theta_t, z_{i,t})$ and bid $b_{i,t} \in (\epsilon, 1 - \epsilon)$,*

$$\textsf{prob}_\sigma(i \textit{ wins} \mid h_{i,t} \textit{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \geq 1/2 - \eta.$$

Proof heuristic: For simplicity, in the main text we consider the special case in which the continuation value of firm i, taking as given the assignment to firm i, does not depend on the winning bid (this would hold in large markets; Appendix C deals with the more general case):

$$V_i(\theta, \mathbf{1}_{b_i \prec \wedge \mathbf{b}_{-i}}, \wedge \mathbf{b}) = V_i(\theta, \mathbf{1}_{b_i \prec \wedge \mathbf{b}_{-i}}).$$

Bidder i's discounted expected payoff from bid $b \in (\epsilon, 1-\epsilon)$ at history $h_{i,t} = (\theta_t, z_{i,t})$ can be written as

$$\begin{aligned} U_i^\sigma(b|h_{i,t}) &= \mathbb{E}_\sigma\left[(b - \kappa_{i,t})\mathbf{1}_{b_{i,t} \prec \wedge \mathbf{b}_{-i,t}} | h_{i,t}\right] + \delta V_i(0|h_{i,t}) - k \\ &= D_i(b|h_{i,t})(b - \kappa_{i,t}) + \delta V_i(0|h_{i,t}) - k \end{aligned}$$

where $\kappa_{i,t} \equiv c_{i,t} - \delta(V_i(1|h_{i,t}) - V_i(0|h_{i,t}))$ is bidder i's cost of winning the auction, including its impact on continuation values. Note that firm i would obtain a payoff of $\delta V_i(\theta_t, 0)$ if it didn't submit a bid. Hence, bidder i's participation constraint implies that $D_i(b_{i,t}|h_{i,t})(b_{i,t} - \kappa_{i,t}) \geq k > 0$, so that $b_{i,t} - \kappa_{i,t} \geq k$.

Since bid $b_{i,t}$ is optimal, for all $\epsilon > 0$ we have that

$$\begin{aligned} & U_i(b_{i,t}|h_{i,t}) \geq U_i(b_{i,t} + \epsilon|h_{i,t}) \\ & \qquad \iff D_i(b_{i,t} + \epsilon|h_{i,t})(b_{i,t} + \epsilon - \kappa_{i,t}) \leq D_i(b_{i,t}|h_{i,t})(b_{i,t} - \kappa_{i,t}), \qquad (4) \\ \text{and } & U_i(b_{i,t}|h_{i,t}) \geq U_i(b_{i,t} - \epsilon|h_{i,t}) \\ & \qquad \iff D_i(b_{i,t} - \epsilon|h_{i,t})(b_{i,t} - \epsilon - \kappa_{i,t}) \leq D_i(b_{i,t}|h_{i,t})(b_{i,t} - \kappa_{i,t}). \qquad (5) \end{aligned}$$

Conditions (2), (4) and (5) imply that

$$
\begin{aligned}
\mathsf{prob}_\sigma(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) &= \frac{D_i(b_{i,t}|h_{i,t}) - D_i(b_{i,t}+\epsilon|h_{i,t})}{D_i(b_{i,t}-\epsilon|h_{i,t}) - D_i(b_{i,t}+\epsilon|h_{i,t})} \\
&= \frac{1 - \frac{D_i(b_{i,t}+\epsilon)}{D_i(b_{i,t})}}{\frac{D_i(b_{i,t}-\epsilon)}{D_i(b_{i,t})} - \frac{D_i(b_{i,t}+\epsilon)}{D_i(b_{i,t})}} \geq \frac{1 - \frac{b_{i,t}-\kappa_{i,t}}{b_{i,t}+\epsilon-\kappa_{i,t}}}{\frac{D_i(b_{i,t}-\epsilon)}{D_i(b_{i,t})} - \frac{b_{i,t}-\kappa_{i,t}}{b_{i,t}+\epsilon-\kappa_{i,t}}} \\
&\geq \frac{1 - \frac{b_{i,t}-\kappa_{i,t}}{b_{i,t}+\epsilon-\kappa_{i,t}}}{\frac{b_{i,t}-\kappa_{i,t}}{b_{i,t}-\epsilon-\kappa_{i,t}} - \frac{b_{i,t}-\kappa_{i,t}}{b_{i,t}+\epsilon-\kappa_{i,t}}} = \frac{1}{2}\frac{b_{i,t}-\kappa_{i,t}-\epsilon}{b_{i,t}-\kappa_{i,t}} \\
&\geq \frac{1}{2}\frac{k-\epsilon}{k} \to \frac{1}{2} \text{ as } \epsilon \searrow 0.
\end{aligned}
$$

Note that the speed of convergence of lower bound $\frac{1}{2}\frac{k-\epsilon}{k}$ is independent of $b_{i,t}$ and $h_{i,t}$. This concludes the proof. ■

Proposition 1 provides a lower bound on firms' winning probability at any given history, conditional on close bids. Because at most one bidder can win, and because there are at least two close bidders conditional on the existence of close bids, it cannot be that firms' winning probability (conditional on their information) is frequently much larger than 1/2. We now make this argument formal. For any $\epsilon > 0$, let ϵ-close denote the event that the winning bid is within ϵ of the second lowest bid. For any environment $\mathcal{E}$, MPE σ and threshold $\epsilon > 0$, let $\mathbb{E}_{\mathcal{E},\sigma}[\cdot | \, \epsilon\text{-}\mathsf{close}]$ denote the expectation over histories h conditional on the event ϵ-close.

Corollary 2 (as-if random bids)**.** *Consider an environment $\mathcal{E}$ and MPE σ that is competitively enforced. For all $\eta > 0$ there exists $\epsilon > 0$ small enough such that*

$$
\mathbb{E}_{\mathcal{E},\sigma}\left[\left|\mathsf{prob}_\sigma(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) - \frac{1}{2}\right| \,\middle|\, \epsilon\text{-}\mathsf{close}\right] \leq \eta. \tag{6}
$$

In words, winning is as-if random conditional on close bids. An implication of Corollary 2 is that Corollary 1 (Section 3) holds whenever equilibrium σ is an MPE that is competitively enforced. Hence, failure to pass our test implies that bidding behavior is either non-Markov, or, if it is Markov, it is sensitive. The following Corollary highlights this.

Corollary 3. *Consider an environment $\mathcal{E}$ and an MPE σ such that for some observable $x \in X$,*

$$\limsup_{\epsilon \searrow 0^+} |\mathit{prob}\,(x_{i,t} = x \mid \Delta_{i,t} \in (0, \epsilon)) - \mathit{prob}\,(x_{i,t} = x \mid \Delta_{i,t} \in (-\epsilon, 0))| > 0.$$

Then, it must be that bidding behavior under $\mathcal{E}$ and σ is sensitive.

Sample implications. Corollary 2 holds under the joint distribution of bids and histories generated under an MPE σ. In empirical applications, however, this distribution is not directly observed and must be replaced by its sample counterpart. In Appendix C we show that if (6) holds under the bidders' beliefs, then it holds asymptotically under the sample joint distribution of bids $\mathbf{b}$ and characteristic $\mathbf{x} \in X^N$ observable to the econometrician. Moreover, we show that the result continues to hold when we restrict attention to any subset of histories that is adapted to the bidders' information at the time of bidding. As a result, our test remains valid when we focus on specific subsets of the data. In Section 5.2 we leverage this result and apply our test separately to bids that are above and below the median winning bid.

The reason such a result holds is that bidders get sufficient feedback about past states and outcomes: in our framework, bidders observe both past states θ, and past bids $\mathbf{b}$. This prevents bidders from making repeated mistakes about realized bidding profiles and characteristics.[19] Expectations must match sample averages with high probability.

5 Empirical Analysis

5.1 Ohio School Milk Auctions

In order to validate our test, we first apply it to the sample of Ohio school milk auctions analyzed by Porter and Zona (1999). Porter and Zona (1999) study bidding on school milk

[19]Bidders do receive feedback from past auctions in our empirical applications. Indeed, municipalities in Japan are usually required to post auction outcomes shortly after each auction, typically within five days.

auctions using data collected by the state of Ohio as part of its efforts to sue dairies for bid rigging. The dataset is an unbalanced panel of milk auctions let by Ohio school districts spanning 11 years between 1980 and 1990 with information on the bids and the identity of the bidders.[20]

Several features of the setting are worth highlighting. First, the auctions are recurring. School districts hold auctions every year, typically between May and August to determine the supplier of milk for the following school year. This allows us to easily track the incumbent firm for a given auction. Second, the dataset includes bids from three bidders located around Cincinnati that were charged for collusion. According to the testimony of the individuals involved, the cartel allocated contracts according to incumbency. Aside from two years (1983 and 1989) during which the cartel broke down, conspirators respected incumbency, with non-incumbents submitting complementary bids. Porter and Zona (1999) show that the bids of non-defendants are consistent with a model of competitive bidding while the bids of the defendants are not.

Table 1 reports summary statistics of the data. Column (1) reports summary statistics for all of the auctions in the sample, column (2) reports those for the subset of auctions in which only the defendant firms participated (Non-competitive) and column (3) reports those for the subset of auctions in which no cartel firm participated (Control).[21] Because the cartel broke down in 1983 and 1989 according to the testimony of the individuals who were involved in collusion, we also report summary statistics for the sample that excludes years 1983 and 1989 for columns (2) and (3). We find that, on average, the number of bidders is about 1.86 for the entire sample, and slightly higher for the non-competitive sample than for the control sample. The winning bid, reported in units of dollars per half-pint of milk, is about $0.131 for the entire sample, and slightly higher in the non-competitive sample. Table 1 also reports the average second lowest and third lowest bids.

[20]We use the dataset constructed by Wachs and Kertész (2019).

[21]One of the findings of Porter and Zona (1999) is that the defendant firms bid more aggressively against non-defendant firms in distant school districts than they did against other cartel firms in the Cincinnati area. Hence, we exclude from the non-competitive sample auctions in which both defendant and non-defendant firms participate. Including these auctions in the control group does not have a meaningful effect on the estimates.

	(1)	(2)		(3)	
	All	Non-Competitive		Control	
	All Years	All Years	Excl 83,89	All Years	Excl 83,89
# Bidders	1.866	1.983	2.058	1.763	1.770
	(0.909)	(0.891)	(0.882)	(0.838)	(0.846)
Winning Bid	0.131	0.136	0.138	0.131	0.131
	(0.013)	(0.015)	(0.015)	(0.013)	(0.013)
2nd-Lowest Bid	0.135	0.142	0.144	0.135	0.135
	(0.013)	(0.015)	(0.014)	(0.012)	(0.013)
3rd-Lowest Bid	0.138	0.147	0.149	0.138	0.137
	(0.013)	(0.016)	(0.014)	(0.012)	(0.012)
Obs.	3,754	235	189	3,267	2,658

Note: The first column corresponds to the set of all auctions, the second column corresponds to the set of auctions in which only the defendant firms bid and the last column corresponds to those in which no defendant firm bid.

Table 1: Summary Statistics of Auctions: Ohio School Milk Auctions.

Table 2 reports summary statistics with respect to incumbency. We define a bidder to be an incumbent for a given school milk auction if the bidder was the winner of the district's auction in the previous year. Column (1) corresponds to the set of all auctions in the dataset, while columns (2) and (3) respectively correspond to auctions in which all participants were defendants and auctions in which none of the participants were defendants. Focusing on the row labeled 1981 in column (1), we find that there are a total of 185 auctions in which an incumbent firm participates. Out of these auctions, the incumbent won 136 of them, or about 74%. Note that we lack the data needed to define incumbency for 1980, which is the first year of the sample. The fraction of auctions in which the incumbent wins is about 80% in column (1), 86% in column (2) if we exclude years 1983 and 1989 (83% if we include those two years), and 81% in column (3). While the fractions are slightly higher in column (2) than in column (3), the differences are quite small. This highlights the general difficulty of using incumbency patterns to detect collusion since both collusive and competitive auctions are

	(1) All			(2) Non-Competitive			(3) Control		
	Win/Inc	Ratio	Total	Win/Inc	Ratio	Total	Win/Inc	Ratio	Total
1980	.	.	249	.	.	4	.	.	230
1981	136/185	0.74	273	6/7	0.86	12	123/162	0.76	235
1982	148/188	0.79	287	9/10	0.90	13	131/161	0.81	252
1983	162/214	0.76	318	7/10	0.70	16	150/187	0.80	274
1984	199/249	0.80	339	18/20	0.90	24	174/215	0.81	293
1985	205/260	0.79	357	18/18	1.00	22	177/226	0.78	314
1986	242/293	0.83	378	16/19	0.84	25	216/255	0.85	332
1987	236/287	0.82	411	18/20	0.90	27	211/255	0.83	358
1988	253/304	0.83	419	18/20	0.90	28	227/263	0.86	359
1989	257/332	0.77	392	13/19	0.68	30	236/289	0.82	335
1990	185/247	0.75	331	17/29	0.59	34	165/211	0.78	285
Obs.		3,754			235			3,267	

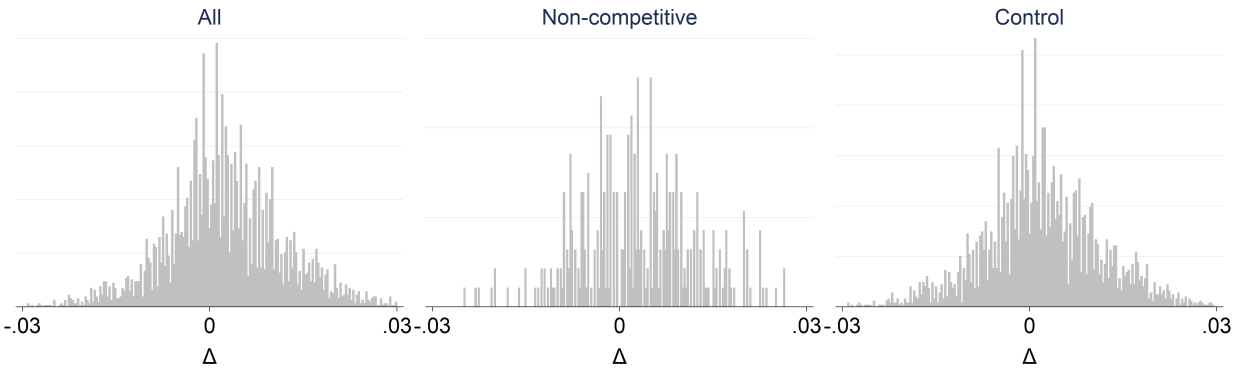
Note: Column (1) corresponds to the set of all auctions, Column (2) corresponds to the set of auctions in which only the defendant firms bid and the Column (3) corresponds to those in which no defendant firm bid.

Table 2: Summary Statistics on Incumbency: Ohio School Milk Auctions.

characterized by high rates of incumbency. As we will show below, the differences between the two samples become pronounced only when we condition on close auctions.

Figure 2 plots the histogram of the running variable, $\Delta_{i,t} = b_{i,t} - \wedge \mathbf{b}_{-i,t}$. A negative value of $\Delta_{i,t}$ implies that bidder i won auction t, and a positive value of $\Delta_{i,t}$ implies that bidder i lost auction t. Values of $\Delta_{i,t}$ close to zero correspond to auctions in which the winner was determined by a very small margin. The left panel of Figure 2 corresponds to the full sample, the middle panel corresponds to the sample of non-competitive auctions and the right panel corresponds to the control sample. There are no obvious differences in the distribution of bid differences $\Delta_{i,t}$ across panels.[22]

[22] This highlights the value-added of considering covariates to detect non-competitive behavior. Tests provided in Chassang et al. (2020) use only the information contained in the distribution of $\Delta_{i,t}$, and would draw similar inference from the different datasets illustrated in Figure 2.



Note: The left panel corresponds to the sample of all auctions, the middle corresponds to the sample of non-competitive auctions and the right panel corresponds to the set of competitive auctions. The horizontal axis is units of dollars.

Figure 2: Histogram of $\Delta_{i,t}$: Ohio School Milk Auctions.

Empirical implementation. Recall the definition of coefficient β,

$$\beta = \lim_{\Delta_{i,t} \searrow 0^+} \mathbf{E}[x_{i,t}|\Delta_{i,t}] - \lim_{\Delta_{i,t} \nearrow 0^-} \mathbf{E}[x_{i,t}|\Delta_{i,t}].$$

We define the variable $x_{i,t}$ as a dummy variable for incumbency status, i.e., $x_{i,t} = 1$ if firm i is an incumbent in auction t, and 0 otherwise. If a cartel allocates contracts to incumbents, we expect β to be strictly negative.

We estimate β using a local linear regression as follows:

$$\widehat{\beta} = \widehat{b_0^+} - \widehat{b_0^-}, \text{ with}$$
$$(\widehat{b_0^+}, \widehat{b_1^+}) = \arg\min \textstyle\sum_{i,t}^T (X_{i,t} - b_0^+ - b_1^+ \Delta_{i,t})^2 K\left(\frac{\Delta_{i,t}}{h_n}\right) \mathbf{1}_{\Delta_{i,t}>0},$$
$$(\widehat{b_0^-}, \widehat{b_1^-}) = \arg\min \textstyle\sum_{i,t}^T (X_{it} - b_0^- - b_1^- \Delta_{i,t})^2 K\left(\frac{\Delta_{i,t}}{h_n}\right) \mathbf{1}_{\Delta_{i,t}<0},$$

where h_n is the bandwidth and $K(\cdot)$ is the kernel. Note that we pool across bidders i and auction t when computing $\widehat{\beta}$.[23] For our baseline estimates, we use a coverage error rate

[23]Corollary 1 is expressed for a particular bidder i and auction t, but it should be obvious that an analogous statement holds when we pool across bidders and auctions.

optimal bandwidth and a triangular kernel with a bias correction procedure as proposed in Calonico et al. (2014). Standard errors are clustered at the level of the school district. We test the null $H_0 : \beta = 0$, against the alternative $H_1 : \beta \neq 0$.

Results. Table 3 presents the results. Panel (A) reports estimates $\widehat{\beta}$ for the sample of auctions in which only the defendant firms participated. In column (1), we use all years between 1980 and 1990 while in column (2), we exclude 1983 and 1989, the two years in which the cartel purportedly broke down. In both columns, we focus on the sample of auctions in which there is an incumbent. We find that the gap β in incumbency rates across close losers and winner is negative (-0.312) and marginally statistically significant ($p = 0.077$) for column (1).[24] The point estimate implies that the marginal winner is about 31.2 percentage points more likely to be an incumbent than the marginal loser. The bandwidth used for estimation is 0.004, or 0.4 cents. In column (2), we find that the estimate is -0.379, and statistically significant at the 5 percent level.

Panel (B) reports findings for the set of control auctions. We find that the regression discontinuity estimate is -0.031 in column (1), which is not statistically different from zero. Because there is no reason to expect 1983 and 1989 to be any different from other years for non-colluding firms, we do not expect any significant differences between column (1) and column (2) for Panel (B). Indeed, the estimate of β in column (2) is -0.068, and statistically indistinguishable from 0.

Overall, the results of Table 3 suggest that our test has reasonable power and size in practice. Figure 3 illustrates the binned scatter plots that correspond to the results in Table 3. The left panel of the figure corresponds to the sample of non-competitive auctions excluding 1983 and 1989, and the right panel corresponds to the sample of all control auctions. The left

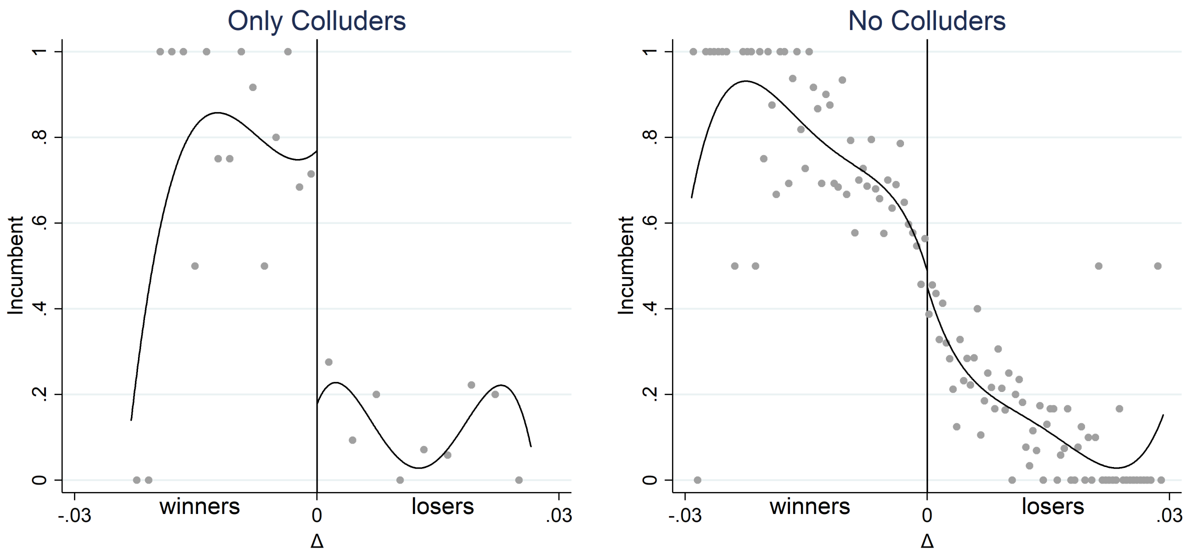
[24]We report two-sided test statistics. Since, specific hypotheses about cartel behavior correspond to specific signs for β (e.g. the term β associated to incumbency rates is negative when cartels allocate contracts to incumbents), one could arguably justify using one-sided tests.

	(1)	(2)
	Incumbency	
	All Years	Exclude 1983 and 1989
Panel (A) : Non-competitive auctions		
$\hat{\beta}$	-0.312^{*}	-0.379^{**}
	(0.177)	(0.181)
h	0.004	0.005
Obs.	309	266
Panel (B) : Control		
$\hat{\beta}$	-0.031	-0.068
	(0.063)	(0.062)
h	0.004	0.005
Obs.	3,053	2,455

Panel (A) corresponds to the sample of auctions in which only the defendant bidders bid. Panel (B) corresponds to the sample of control auctions in which none of the defendant bidders bid. Standard errors are clustered at the level of the school district and reported in parenthesis. The table also reports the bandwidth h used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table 3: Regression Discontinuity Estimates: Ohio School Milk Auctions.

panel of the figure displays a visible discontinuity in incumbency status between marginal winners and marginal losers while the right panel of the figure shows a smooth continuous relationship between $\Delta_{i,t}$ and incumbency status. As it is clear from the figure, incumbents win with high probability even among the competitive sample. It is only by looking at marginal auctions that we find differential rates of incumbency between the two samples.



Note: Left panel corresponds to column (2) Panel (A) of Table 3 and right panel corresponds to column (1), Panel (B) of Table 3. The curves in the figure correspond to 4th order (global) polynomial approximations of the conditional means.

Figure 3: Binned Scatter Plot for Incumbency: Ohio School Milk Auctions

5.2 Public Procurement Auctions in Japan

Our second dataset consists of bids submitted by construction firms participating in auctions for construction projects let by municipalities in the Tohoku region of Japan. Our baseline sample consists of roughly 11,000 procurement auctions let by 16 municipalities between 2004 and 2018. The total award amount for these auctions is about 232 billion yen, or about \$2.3 billion U.S. dollars. No firm has been charged for colluding in any of the auctions in our sample. However, as we note in the Introduction, results in Kawai and Nakabayashi (2018) and Chassang et al. (2020) suggest that some of these auctions are collusive.

The baseline sample consists of auctions from municipalities in which tests of non-competitive behavior exploiting isolated winning bids (Chassang et al., 2020) do not apply.[25] In the Online Appendix, we show that our findings extend to the sample of all of municipal-

[25] In order to choose the set of municipalities to include in our sample, we first compute the density of the running variable for each municipality. The running variable, $\Delta_{i,t}$ is defined as $\Delta_{i,t} = b_{i,t} - \wedge \mathbf{b}_{-i,t}$ where bids are in percentages of the reserve price. We include the municipality in the sample if the following is

ities from which we have obtained data. This is not surprising since it is likely that cartels are operating in the excluded cities.

5.2.1 Data and Empirical Implementation

Data and institutional background. Auctions are first-price sealed bid and the lowest bidder is awarded the project subject to the reserve price. Some of the municipalities use public reserve prices and others use secret reserve prices.[26] For example, in 2012, 7 municipalities used public reserve prices, 8 municipalities used secret reserve prices, and 1 municipality used both, in our sample. The lowest bid was rejected in about 12.5% of the overall sample.[27] Online Appendix B shows that our findings are qualitatively unchanged if we focus on the subset of municipalities with a public reserve price.

Our data includes all bids, the identity of the bidders, and a brief description of the construction project. Column (1) of Table 4 reports the summary statistics of the auctions. On average, the reserve price is 22.26 million yen, or about 222 000 US dollars. The average winning bid is 20.71 million yen. The average ratio of the winning bid to the reserve is about 92.6%. On average, 6.80 bidders participate in each auction. Column (2) reports the summary statistics of the bidders in our sample. Bidders in our sample participate on average in 22.56 auctions and win on average 3.32 times. The table also reports the summary statistics on incumbents and the amount of backlog of the firms. We discuss how we define

---

satisfied:

$$0.7 \times \max_{d \in [-3\%, -0.5\%]} f_\Delta(d) \leq \min_{d \in [-0.2\%, 0\%]} f_\Delta(d),$$

where $f_\Delta(d)$ is the density of Δ. In municipalities with isolated bids, there will be a trough in the density of f_Δ around 0, and the inequality is not satisfied. We also drop municipalities in which f_Δ exhibits a mass at 0. We do so by running a McCrary test (McCrary, 2008) on the running variable and dropping municipalities with p-values less than 0.05. The auctions in these municipalities have binding price floors.

[26] Proposition 1 and Corollary 2 extend as stated to auctions with a secret reserve price.

[27] It is very common in auctions for buyers to retain the option of rejecting the lowest bid when the buyer believes that the price is high. The fraction of auctions in which the low bid is rejected in our sample is comparable to other settings with a secret reserve price. For example, in their study of federal offshore oil and gas drainage lease sales, Hendricks and Porter (1988) report that the most competitive bid was rejected in 7 percent of wildcat tract auctions, and 15 percent of drainage tract auctions.

	(1)		(2)	
	Auctions		Bidders	
	Mean	Std. Dev.	Mean	Std. Dev.
Reserve (Mil. Yen)	22.26	77.14		
Winning Bid (Mil. Yen)	20.71	71.78		
Win Bid/Reserve	0.926	0.083		
# of Bidders	6.80	4.21		
Incumbent Participates (0/1)	0.044	0.204		
# of Auctions Participated			22.56	45.93
# of Wins			3.32	6.97
Raw Backlog (90-Day)			4.11	17.16
Raw Backlog (180-Day)			6.45	22.85
Obs.	11,207		3,377	

Note: The reserve price, winning bid, and backlog measures are reported in units of millions of yen.

Table 4: Summary Statistics of Auctions and Bidders: Municipal Auctions from Japan.

these variables next.

Empirical implementation. Our first covariate of interest is the firms' backlog. We consider both raw backlog and standardized backlog. We define the raw backlog of firm i at auction t as either the 90-day or 180-day cumulative size (measured by the reserve price) of projects won by firm i. We define the 90-day and 180-day backlogs, denoted by $x_{i,t}^{B_{90}}$ and $x_{i,t}^{B_{180}}$, as follows:

$$x_{i,t}^{B_k} \equiv \sum_{\tau \in T_t^k} r_\tau \mathbf{1}_{\wedge \mathbf{b}_{-i,\tau} \succ b_{i,\tau}},$$

where r_τ denotes the reserve price of auction τ and T_t^k denotes the set of auctions in our sample that take place in the $k \in \{90, 180\}$ days prior to auction t. We make sure not to include auction t in T_t^k since its outcome is not in the information set of bidders at time

t. Although the raw backlog is a natural metric for capturing the amount of work recently awarded to a firm, variation in raw backlog captures both intertemporal change in backlog as well as heterogeneity in firm size. In order to construct a measure of backlog that only captures the intertemporal variation, we standardize the raw backlog at the firm level, using its within-firm mean and standard deviation. The 90-day and 180-day standardized backlogs, $x_{i,t}^{\overline{B}_{90}}$ and $x_{i,t}^{\overline{B}_{180}}$, are defined as follows:

$$x_{i,t}^{\overline{B}_k} = \frac{x_{i,t}^{B_k} - \mu_{x_i^{B_k}}}{\sigma_{x_i^{B_k}}}, \tag{7}$$

where $\mu_{x_i^{B_k}}$ is the within-firm mean of $x_{i,t}^{B_k}$ and $\sigma_{x_i^{B_k}}$ is the within-firm standard deviation of $x_{i,t}^{B_k}$.[28] Because standardized backlog is defined relative to the firm's own historical average, $x_{i,t}^{\overline{B}_k}$ is zero if firm i's raw backlog is equal to its time-series average at the time of auction t.

We emphasize that all of our backlog measures are likely to be noisy measures of the firms' true cost-relevant backlog. The number of days we use to define our backlog measures (90 or 180 days) is arbitrary, and most firms are likely to work on projects that are not included in our measures of backlog.[29] This does not invalidate our test. As we discussed in Section 2, variables observable to the econometrician can be imperfect and imprecise. Corollary 1 holds regardless.

Column (2) of Table 4 reports summary statistics of raw backlog in millions of yen. The average 90-day backlog is around 4.11 million yen and the average 180-day backlog is around 6.45 million yen. Standardized backlog averages to zero for each firm by construction.

[28] Our analysis requires covariates $x_{i,t}$ to be observable to bidder i at time t. For simplicity, the measure of standardized backlog defined by (7) implicitly assume that firms know the sample mean and standard deviation of their backlog. Online Appendix B relaxes this assumption: we use rolling sample averages to estimate $\mu_{x_i^{B_k}}$ and $\sigma_{x_i^{B_k}}$ in expression (7). Those statistics are mechanically observed by bidders at the time of bidding and the associated findings are very similar.

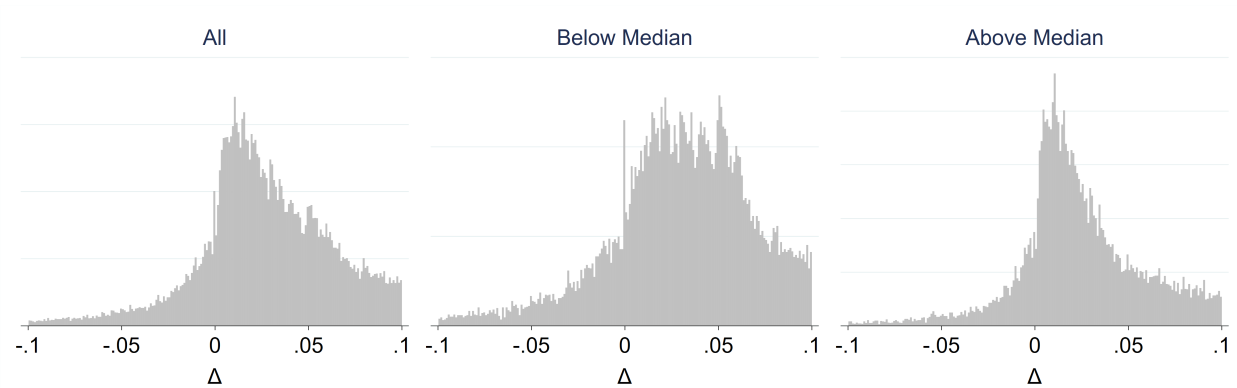
[29] Many bidders who participate in auctions let by municipal governments also participate in auctions that are let by the Ministry of Land Infrastructure and Transportation and prefectural governments. Many firms also do work for other private firms.

Another covariate of interest is whether or not a given firm is an incumbent for a given project. We define a firm to be an incumbent if it is the winner of the previous auction with the same project name let by the same municipality. To give an example, the city of Miyako in Iwate prefecture held procurement auctions with the project name "Restoration of Yagisawa public housing complex" on 3 occasions, November 22, 2011, September 19, 2012, and December 16, 2014. A firm named Kikuchi Painting won each time. We define this firm to be the incumbent in the second and third auctions. We define all other participants in the second and third auctions to be non-incumbents. We do not define incumbency status for any of the bidders in the first auction. Similarly, if there is only one auction for a given project name in a municipality, we do not define incumbency for any bidders. Column (1) of Table 4 reports summary statistics of incumbency status. There is an incumbent bidder in 4.4% of the auctions in our sample.

The running variable is $\Delta_{i,t} = b_{i,t} - \wedge \mathbf{b}_{-i,t}$, where bids are normalized by the reserve price. The left panel of Figure 4 is the histogram of $\Delta_{i,t}$. The distribution is skewed to the right of zero because the average number of bidders is 6.80 ($\Delta_{i,t}$ is negative for only one bidder per auction, and it is positive for all of the losing bidders). Because we report our regression discontinuity results separately for the set of bids above and below the median winning bid for the municipality letting the auction, the next two panels of Figure 4 plot the histogram of $\Delta_{i,t}$ separately for the two sets of auctions. The middle panel corresponds to the sample of bids below the municipal median and the right panel corresponds to bids above the median.

As before, we estimate discontinuities in the expectation of $x_{i,t}$ as a function of $\Delta_{i,t}$ using a local linear regression with a coverage error rate optimal bandwidth and a triangular kernel with a bias correction procedure as proposed in Calonico et al. (2014).[30] Standard errors are

[30]We restrict our sample to auctions in which bid difference $|\Delta|$ is less than 20% of the reserve price: often, bids that are more than 20% lower than the second lowest bid are likely to be misrecorded.



Note: The left panel corresponds to the histogram of $\Delta_{i,t}$ for the entire sample. The middle panel corresponds to the sample of bids below the median winning bid of the relevant municipality. The right panel corresponds to the sample of bids above the median. The histogram is truncated at $\Delta_{i,t} = -0.1$ and $\Delta_{i,t} = 0.1$ for readability.

Figure 4: Histogram of $\Delta_{i,t}$: Municipal Auctions from Japan.

clustered at the auction level.

5.2.2 Results

We first present results for the full sample, and then distinguish between the subset of bids that are above or below the municipality median.

Full sample. Table 5 presents our results. Column (1) reports the estimate of β for the 90-day standardized backlog. We find that the average standardized backlog of marginal losers is higher than that of marginal winners by about 0.136 standard deviations. The estimate is statistically significant at the 1% level. The coverage error rate optimal bandwidth we use is about 0.020, or about 2% of the reserve price. Column (2) reports our estimate of β for the 90-day raw backlog (measured in millions of yen). We find that the 90-day backlog of marginal losers is, on average, 3.78 million yen higher than that of marginal winners. The estimate is statistically significant at the 10% level. Note that the sample size in column (2) is larger than that in column (1) because we can only compute the standardized backlog for firms that win at least one auction while we can compute the raw backlog for all firms.[31]

Columns (3) and (4) report our results for the 180-day backlog.[32] The 180-day backlog of marginal losers is on average 0.147 standard deviations or 6.747 million yen higher than that of marginal winners. Column (5) reports the estimate of β using incumbency status as the outcome variable. We find that marginal losers are 18.4 percentage points less likely to be an incumbent than marginal winners. We only use the set of auctions in which there is an incumbent for estimation in column (5). Correspondingly, the number of observations is much smaller than in columns (1)-(4). While the results for raw backlog are somewhat weaker than those for standardized backlog, overall the results of Table 5 suggest that some of the auctions in our sample are uncompetitive.

	(1)	(2)	(3)	(4)	(5)
	90-Day Backlog		180-Day Backlog		Incumbent
	Standardized	Raw	Standardized	Raw	
$\hat{\beta}$	0.136***	3.782*	0.147***	6.747**	−0.184**
	(0.038)	(2.250)	(0.038)	(3.157)	(0.078)
h	0.020	0.016	0.022	0.015	0.026
Obs.	59,367	63,742	59,413	63,742	2,517

Note: Standard errors are clustered at the auction level and reported in parenthesis. The table also reports the bandwidth h used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

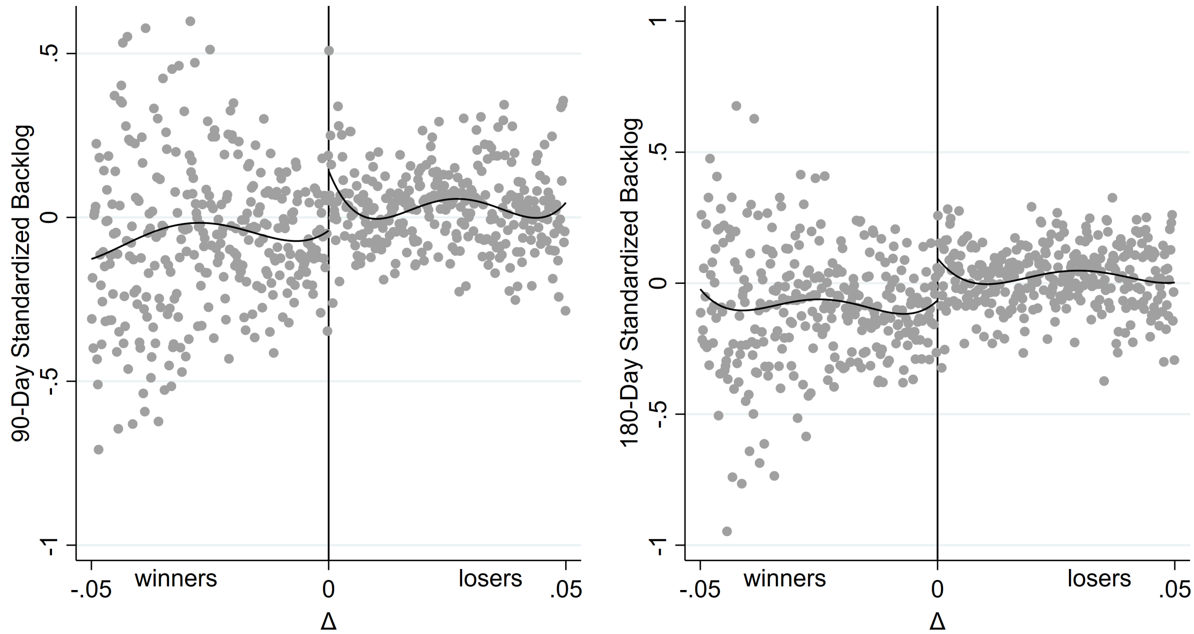
Table 5: Regression Discontinuity Estimates: Municipal Auctions from Japan.

The two panels of Figure 5 are the binned scatter plots of standardized backlog. The

[31]If a firm never wins an auction, $\sigma_{x_{i_k}^B}$ in expression (7) is zero, and $x_{i,t}^{\overline{B}_k}$ is undefined.

[32]Note that the sample size in column (3) is larger than in column (1). Suppose that a firm participates twice in the sample, say, January 1, 2015 and May 1, 2015. Suppose that the firm wins the first auction. According to our 90-day backlog measure, the firm's backlog would be zero for both auctions. Hence, we cannot define the standardized backlog for this firm. However, according to our 180-day backlog measure, the firm has a positive backlog in the second auction. Hence, we can compute the within-firm standard deviation for 180-day backlog, but not for 90-day backlog.

figures correspond to the regression results reported in columns (1) and (3) of Table 5 (we report the binned scatter plots of raw backlog in Appendix B.). The graphs show that there is a discontinuity in the binned averages at $\Delta = 0$.



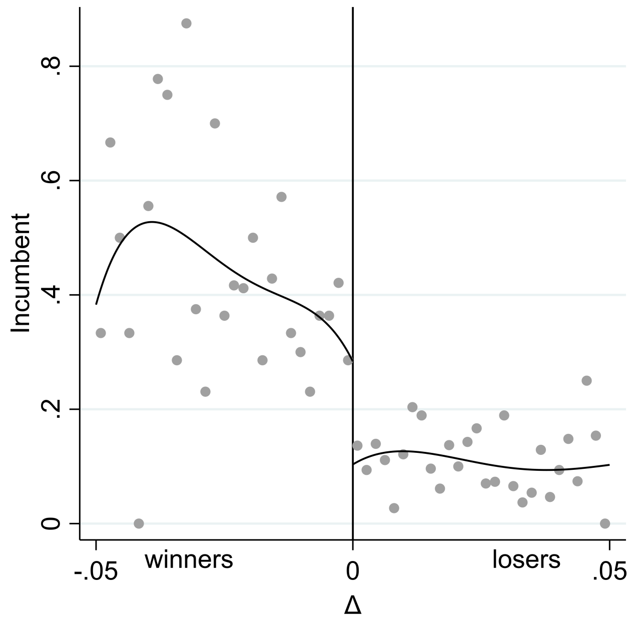
Note: The curves in the figure correspond to 4th order (global) polynomial approximations of the conditional means.

Figure 5: Binned Scatter Plot for 90-Day and 180-Day Standardized Backlog: Municipal Auctions from Japan.

The binned scatter plots for incumbency is reported in Figure 6. The figure shows that marginal winners are more likely to be incumbents than marginal losers as in the case of collusive auctions in the Ohio school milk dataset.

High versus low bids. We now distinguish statistics computed for the subset of bids above the municipality median and statistics computed for the subset of bids below the median.[33] We normalize all of the bids by the reserve price so that the sample of bids above

[33] More precisely, we compute the median winning bid for each municipality. We then categorize bids according to whether or not they are higher or lower than the median.



Note: The curves in the figure correspond to 4th order (global) polynomial approximations of the conditional means.

Figure 6: Binned Scatter Plot for Incumbency: Municipal Auctions from Japan.

the median are those with bids close to 1. As we discussed in Section 3 (see, in particular footnote 18) and Section 4, Corollaries 1 and 2 continue to hold when we focus on the sample of bids below or above a particular threshold. Since the goal of collusion is to elevate prices, we expect the set of bids below the median to be less collusive than the set of bids above the median.

Panel (A) of Table 6 reports the estimates for the sample of bids above the municipality median. Column (1) and (2) correspond to the regression discontinuity estimates for the 90-day backlog. The 90-day backlog of marginal losers is on average 0.24 standard deviations and 5.57 million yen higher than that of marginal winners. The estimates are statistically significant at the 1% and 5% levels, respectively. Columns (3) and (4) report our results for the 180-day backlog. The 180-day backlog of marginal losers is on average 9.54 million yen and 0.22 standard deviations higher than that of marginal winners.

	(1)	(2)	(3)	(4)	(5)
	90-Day Backlog		180-Day Backlog		Incumbent
	Standardized	Raw	Standardized	Raw	
Panel (A) : Above Median					
$\hat{\beta}$	0.244***	5.572**	0.223***	9.544**	−0.260**
	(0.049)	(2.793)	(0.049)	(3.906)	(0.106)
h	0.017	0.016	0.021	0.016	0.034
Obs.	44,945	48,178	44,982	48,178	1,604
Panel (B) : Below Median					
$\hat{\beta}$	−0.045	0.377	−0.017	0.108	−0.324**
	(0.058)	(2.238)	(0.060)	(3.316)	(0.150)
h	0.030	0.024	0.027	0.021	0.026
Obs.	14,438	15,580	14,447	15,580	488

Note: Panel (A) corresponds to the sample of bids that are above the median winning bid. Panel (B) corresponds to the sample of bids below the median winning bid. Standard errors are clustered at the auction level and reported in parenthesis. The table also reports the bandwidth h used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table 6: Regression Discontinuity Estimates: Municipal Auctions from Japan.

Column (5) reports estimates of β using incumbency status as the outcome variable. We find that marginal losers are about 26.0 percentage points less likely to be an incumbent than marginal winners.

Panel (B) reports the results for the sample of bids that are below the median winning bid. In contrast to Panel (A), the estimates of β are not statistically significant for columns (1) - (4) at the 5% level. In the case of incumbency, we find that marginal winners are more likely to be incumbents than marginal losers. While it seems possible that the sample of bids in Panel (B) include some non-competitive bids, the overall results in Table 6 are consistent with the notion that there is less collusion among bidders that submit low bids.

We note that lack of statistical significance in Panel (B) is unlikely to be driven by the smaller sample sizes since Panel (A) and (B) contain the same number of bids for which $\Delta_{i,t}$ is less than zero.[34] Hence, the effective sample sizes are similar. Online Appendix B considers an alternative way of partitioning the sample in which we divide the bids according to whether or not the winning bid of the auction is above or below the median. Partitioning the sample according to the winning bid of the auction allows us to obtain two groups with roughly similar sample sizes. We find similar results as Table 6.

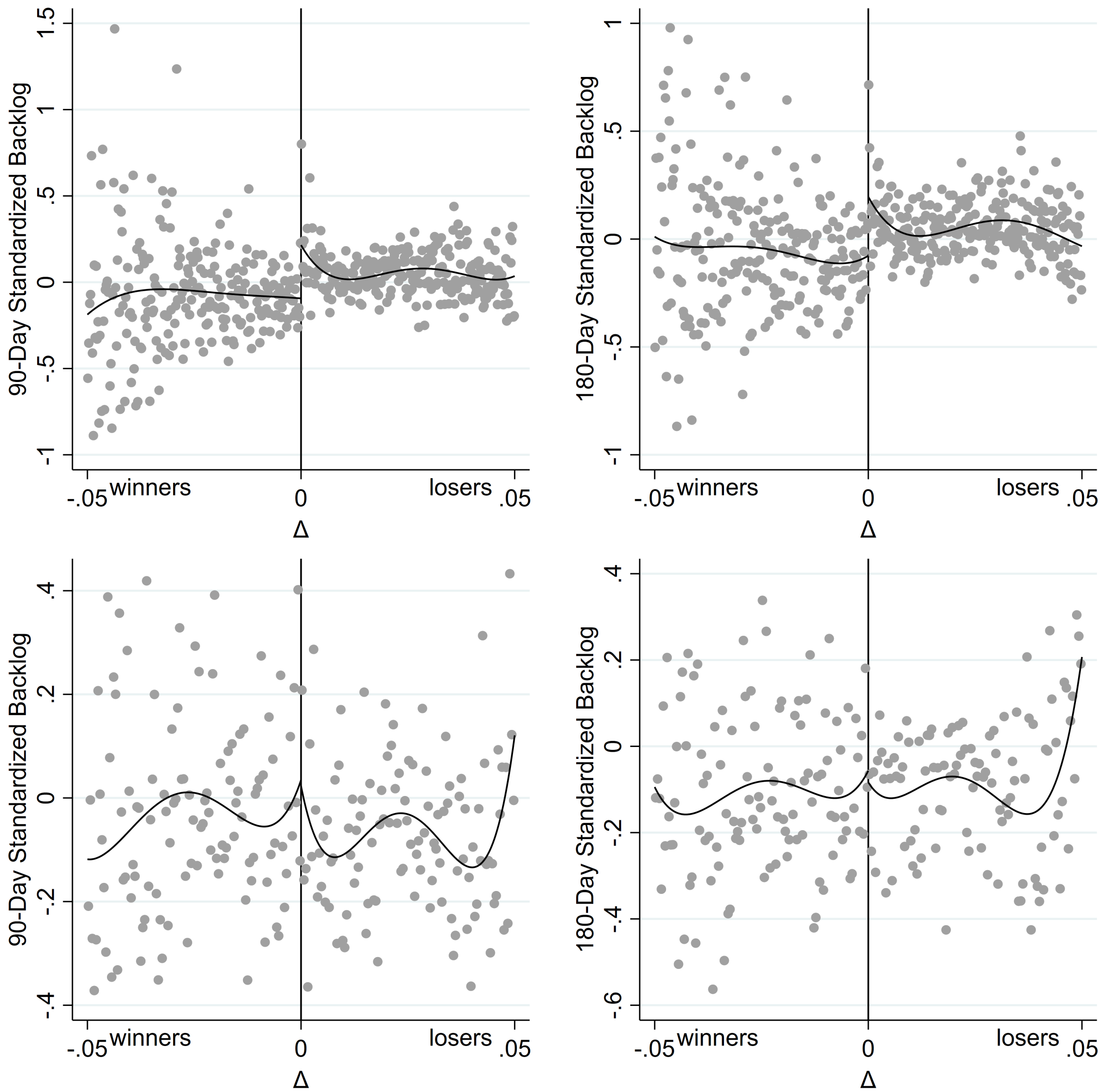
Figure 7 plots the binned scatter plot of the standardized backlog (we report the binned scatter plots of raw backlog in Appendix B). The top panels correspond to the above median sample and the bottom panel corresponds to the below median sample. The left panels are for the 90-day standardized backlog and the right panels are for the 180-day standardized backlog. We find a clear discontinuity at zero in the top panels, but not for the bottom panels, consistent with the results of Table 6. Figure 8 plots the binned scatter plots corresponding to column (5) of Table 6. The discontinuity in the binned averages is visible in the top panel.

The results of Table 5 and 6 suggest that there is a significant amount of non-competitive auctions among the dataset, and that they are concentrated in high-bid auctions. The results also suggest that our tests can discriminate between competitive and non-competitive bids.

5.2.3 A Placebo Test

Because the precise order of the losing bidders is unimportant for allocation by a cartel, it seems plausible that bidding rings would not have specific rules for determining which bidder should bid the second or third lowest. If this is the case, we should not expect significant differences in backlog or incumbency status between marginally second and marginally third place bidders for both competitive and non-competitive auctions. This suggests the following placebo test.

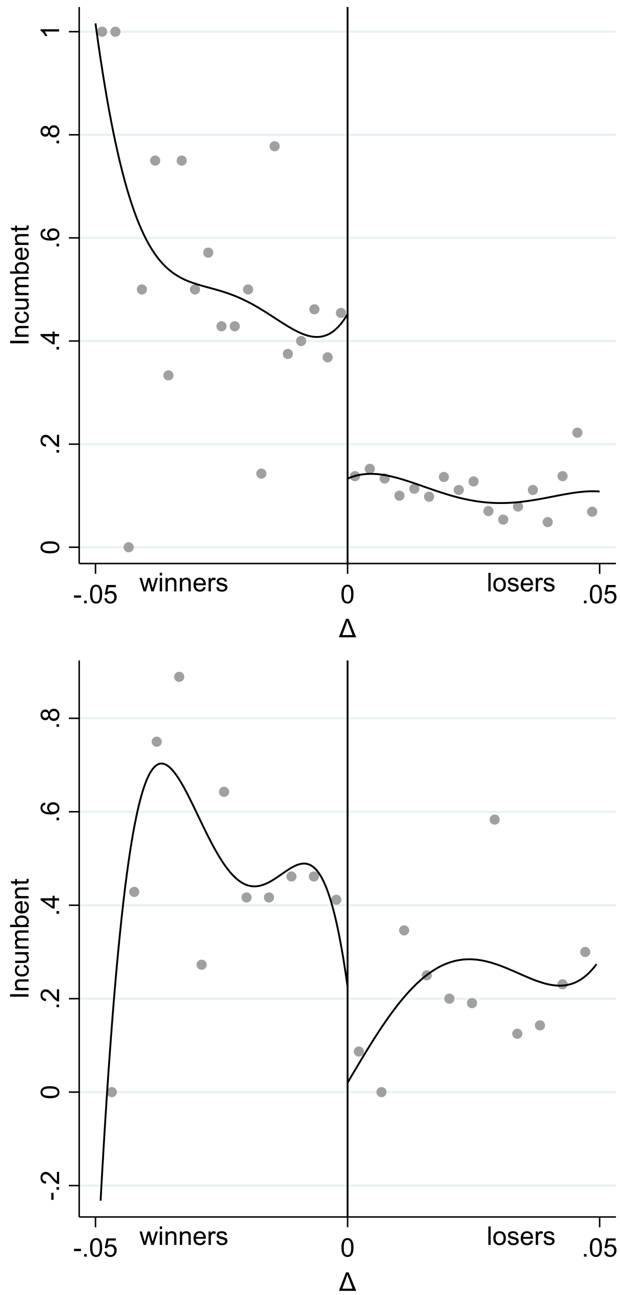
[34] Recall that we partition the sample according to whether or not a bid is above or below the median winning bid.



Note: Top panels correspond to columns (1) and (3) of Panel (A) of Table 6. Bottom panels correspond to columns (1) and (3) of Panel (B) of Table 6. Left panels correspond to 90-day standardized backlog and the right panels correspond to 180-day standardized backlog. The curves correspond to 4th order (global) polynomial approximations of the conditional means.

Figure 7: Binned Scatter Plot for 90-Day and 180-Day Standardized Backlog: Municipal Auctions from Japan.

For any non-winning bidder i, define $\Delta^2_{i,t} \equiv b_{i,t} - \min\{b_{j,t} \text{ s.t. } j \neq i \text{ and } j \text{ loses}\}$. Bid difference $\Delta^2_{i,t}$ is negative for the second lowest bidders and positive for other bidders. We



Note: Top panels correspond to Panel (A) of Table 6 and bottom panels correspond to Panel (B) of Table 6. The curves correspond to 4th order (global) polynomial approximations of the conditional means.

Figure 8: Binned Scatter Plot for Incumbency: Municipal Auctions from Japan.

do not define Δ^2 for the lowest bidder. Even under collusion, we do not expect that there should be systematic differences in the mean backlog and mean incumbency of close second

	(1)	(2)	(3)	(4)	(5)
	90-Day Backlog		180-Day Backlog		Incumbent
	Standardized	Raw	Standardized	Raw	
Panel (A) : Above Median					
$\hat{\beta}$	−0.053	−0.756	−0.016	−1.847	−0.023
	(0.048)	(1.433)	(0.043)	(2.216)	(0.058)
h	0.010	0.016	0.013	0.015	0.026
Obs.	41,473	44,802	41,508	44,802	1,466
Panel (B) : Below Median					
$\hat{\beta}$	0.035	−0.494	0.026	−0.094	0.080
	(0.063)	(1.792)	(0.059)	(2.274)	(0.130)
h	0.025	0.020	0.026	0.020	0.020
Obs.	9,795	10,796	9,805	10,796	317

Note: Panel (A) corresponds to the sample of bid that are above the median winning bid. Panel (B) corresponds to the sample of bids that are below the median winning bid. Standard errors are clustered at the auction level and reported in parenthesis. The forcing variable is Δ^2. The table also reports the bandwidth h used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table 7: Placebo: Regression Discontinuity Estimate with Respect to Δ^2

and third (or fourth, fifth, etc.) bidders.

Table 7 reports estimates of the discontinuity in backlog and incumbency around $\Delta^2_{i,t} = 0$. The top panel corresponds to the sample of bids that are higher than the median winning bid. Unlike in Panel (A) of Table 6, discontinuity estimates are statistically insignificant at the 5% level in Panel (A) of Table 7. The bottom panel corresponds to the sample of bids that are lower than the median. Unsurprisingly, the same holds for Panel (B). Binned scatter plots corresponding to these estimates are given in Online Appendix B.

6 Discussion

This paper proposes a novel method to screen for non-competitive behavior using covariates such as backlog and incumbency status. While many practitioners have advocated using these patterns to screen for collusion, identifying allocation patterns that reflect agreements among cartels from those that simply reflect bidder cost heterogeneity has been difficult. Our contribution is to make this possible by conditioning on auctions that are determined by a close margin. Our approach is easy to implement, requires no sophisticated programming, and is fairly robust to model misspecification. In addition, our approach can easily be adapted to formulate tests of non-competitive behavior exploiting any observed covariate suspected to reflect collusive strategies, such as geographic segmentation, subcontracting, or joint bidding. Our approach can also be extended to other auction formats such as handicap auctions, scoring auctions and all-pay auctions.

We end the paper with a discussion of practical aspects of our tests: (i) the relation between the rejection of the test and collusion and (ii) firms' responses to antitrust oversight.

Rejection of the test and collusion. Section 4 shows that competitively enforced MPE implies the null of $\beta = 0$ under fairly general information structures. Hence, rejection of the null suggests that bidding behavior is either non-Markov, or, if it is Markov, it is sensitive.

While bidding behavior that is sensitive does not immediately imply bidder collusion, the empirical results in Section 5 suggest a correlation between the rejection of our test and non-competitive behavior. Firms bidding for school milk contracts that were charged with collusion fail our tests while firms who were not charged with collusion pass; the sample of bids that are relatively high fail our tests while the sample of bids that are relatively low pass. This suggests that our tests are sufficiently powered to flag potential cartels while keeping the rate of false positives low in practical applications.

Firm response to screening. Screens for collusion are perhaps most useful when firms are unaware of the details of the screening technology. When screens are known to the colluders, they can potentially adapt their behavior to avoid detection. Are screens for collusion still useful if cartels adapt? Are there tests that reduce the incentives of cartels, and don't harm competitive industries? We study these and related issues in our companion paper, Ortner et al. (2020).

We say that a test of collusive behavior is *safe* if the rate of false positives vanishes as the number of observations grows. The tests proposed in the current paper satisfy this property. Ortner et al. (2020) shows that antitrust oversight based on safe tests always reduces the set of enforceable collusive schemes available to cartels. Put differently, even if firms know they are being monitored and adapt their play accordingly, screens based on safe tests always make cartels weakly worse off.

Moreover, as we illustrate in Ortner et al. (2020), adaptive responses by cartels may themselves lead to suspicious bidding patterns that can also be detected. Consider, for example, the test that compares the incumbency status of marginal winners and marginal losers. If the cartel has a rule of allocating projects to incumbents and wishes to maintain this rule, then the cartel needs to have the lowest bidder bid substantially lower than the second lowest bidder to avoid detection. However, this would generate isolated winning bids similar to the pattern documented in Chassang et al. (2020). Hence, avoiding one test may lead cartels to bid in ways that lead to rejection of other tests. Alternatively, the cartel can change its allocation rule so that incumbents are not always guaranteed to win. However, changing the allocation rule may reduce efficiency and increase the cost of coordination. This reduces bidders' incentives to collude.

Online Appendix – Not for Publication

A Examples

A.1 An example of non-smooth demand.

Consider a complete information auction with an incumbent I and an entrant E with respective known costs $c_I < c_E$. Assume that bidding cost k is zero.

Lemma A.1 (non-smooth demand)**.** *In any efficient equilibrium in weakly undominated strategies, the incumbent wins with bid c_E with probability 1. The density of the entrant's bid below c_E is 0. The density of the entrant's bids above c_E is strictly positive and bounded away from 0. Specifically, for all $\epsilon > 0$, the incumbent's demand D_I satisfies $\frac{D_I(c_E+\epsilon)-1}{\epsilon} \leq -\frac{1}{c_E+\epsilon-c_I}$.*

Proof. In an efficient equilibrium in weakly undominated strategies, the incumbent cannot bid above c_E with positive probability: the entrant's optimal bid would win with positive probability.

In turn, the entrant cannot bid below c_E. This implies that the incumbent's optimal bid is c_E. Optimality of c_E implies that for any $\epsilon > 0$,

$$D_I(c_E+\epsilon)(c_E+\epsilon-c_I) \leq D_I(c_E)(c_E-c_I) = c_E - c_I \iff \frac{D_I(c_E+\epsilon)-1}{\epsilon} \leq -\frac{1}{c_E+\epsilon-c_I}.$$

□

A.2 A collusive Markov perfect equilibrium

We now describe an environment and an MPE which satisfy our assumptions, including positive participation costs, but nevertheless supports collusive behavior and fails to pass our tests.

Two bidders $i \in \{1, 2\}$ compete for contracts. Bidder 1 has a publicly observable cost $c_H > 0$ at even periods, and a publicly observable cost $c_L \in (0, c_H)$ at odd periods. Bidder 2 has i.i.d. costs, equal to c_L with probability $q > 50\%$, and equal to c_H with probability $1 - q$. Bidder 2's cost is her private information. Auctions have reserve price $r = 1 > c_H$. Bid preparation cost k is small, satisfying $k < \min\{(1-q)(r-c_H), (c_H-c_L)\}$. Let $\hat{c} = qc_L + (1-q)c_H$ be bidder 2's expected cost. We assume that $\delta < 1$ is sufficiently large, so that $c_H > \max\{r(1-\delta) + (c_L+k)\delta, (r-k)(1-\delta) + \delta\hat{c}\}$.

For simplicity, we expand the bidding space to deal with tied bids. For every bid b, we add bid b^-, equal in value to b, but such that $b^- \prec b$. We also consider a degenerate case where the impact of the state on costs is vanishingly small.[35] The state θ_t keeps track of:

- Is time period $t \in \mathbb{N}$ even or odd (i.e. $t \mod 2$)?
- Has any bidder won the auction both at times $2t$ and $2t+1$ the past?
- Who has won the auction in the last period?

The collusive equilibrium we construct is as follows. If at any point in the past a bidder has won the auction in consecutive even and odd periods, or no player won an auction, players bid according to a static Nash equilibrium.[36] If this is not the case, then:

- If $t \mod 2 = 0$, bidder 1 participates, and bids r; bidder 2 participates only if her cost is c_L, and bids r^-.
- If $t \mod 2 = 1$, and bidder 1 won in the previous period, then only bidder 2 participates, and bids r.
- If $t \mod 2 = 1$, and bidder 2 won in the previous period, then only bidder 1 participates, and bids r.

One can check that, when discount factor δ high enough, this is a Markov perfect equilibrium. Furthermore, bidder 2 is the only close winner, and conditional on being a close winner, has an expected 1-period backlog equal to $1-q$. Bidder 1 is the only close loser, and conditional on being a close loser, bidder 1 has an expected 1-period backlog equal to $q > 1-q$.

We note that bidding behavior under this MPE is sensitive. Indeed, both bidders' expected continuation payoff fall discretely if the winning bid changes from $b_w \leq r$ to $b_w > r$ (i.e., to $b_w = \emptyset$).

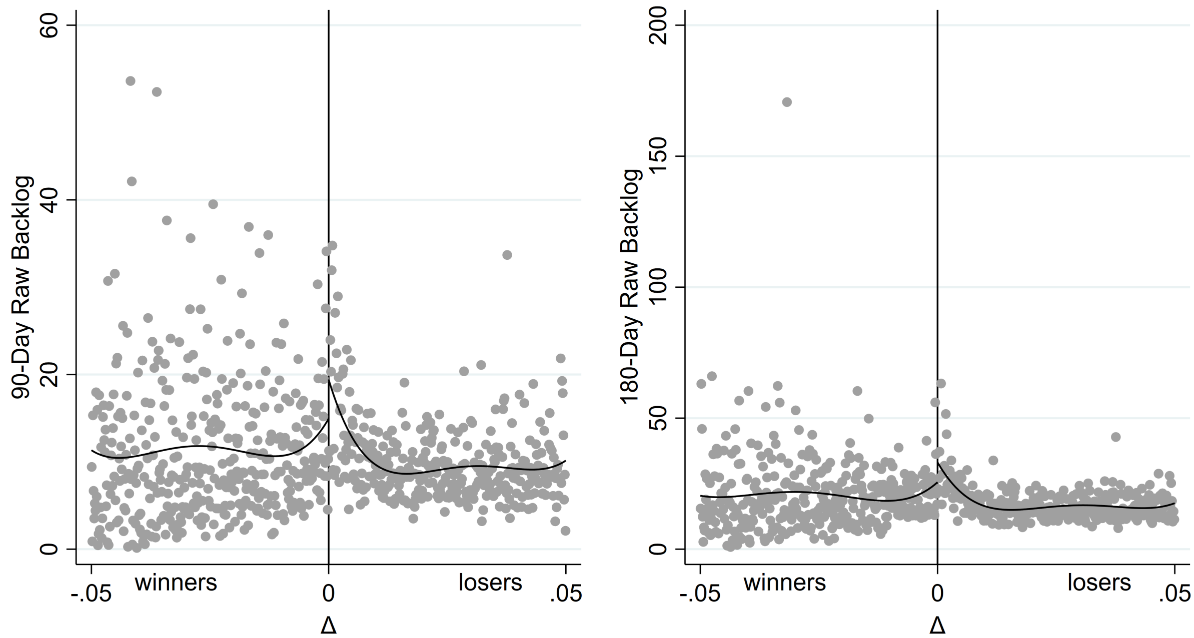
[35] This information can be made payoff relevant in different ways, for instance by shifting costs slightly as a function of the state.

[36] For t even or odd, the stage game has a Nash equilibrium in which bidder 1 randomizes between entering or not, and bidder 2 enters with probability 1 if her cost is c_L (earning profits $c_H - c_L$), and enters with probability 0 if her cost is c_H.

B Further Empirics

In this section, we first present the binned scatter plots corresponding to the regression results in Tables 5, 6 and 7 of Section 5. We next present a series of results that show robustness of the results that we report in Section 5.2. In particular, we present the regression discontinuity estimates when we partition the sample of bids into two depending on whether or not the winning bid of the auction is above or below the median. This alternative way of partitioning equates the sample sizes across the two groups. We next report the results from using an alternative way of standardizing the backlog so that it is measurable with respect to the information of bidder i, $h_{i,t}$. We also report the results when we limit our sample to the municipalities that use public reserve prices for their auctions. Finally, we report findings for the entire sample of auctions for which we have data.

Omitted binned scatter plots for Table 5. Figure B.1 plots the binned scatter plots of 90-day and 180-day raw backlog that correspond to columns (2) and (4) of Table 5. The discontinuities at 0 are quite modest.

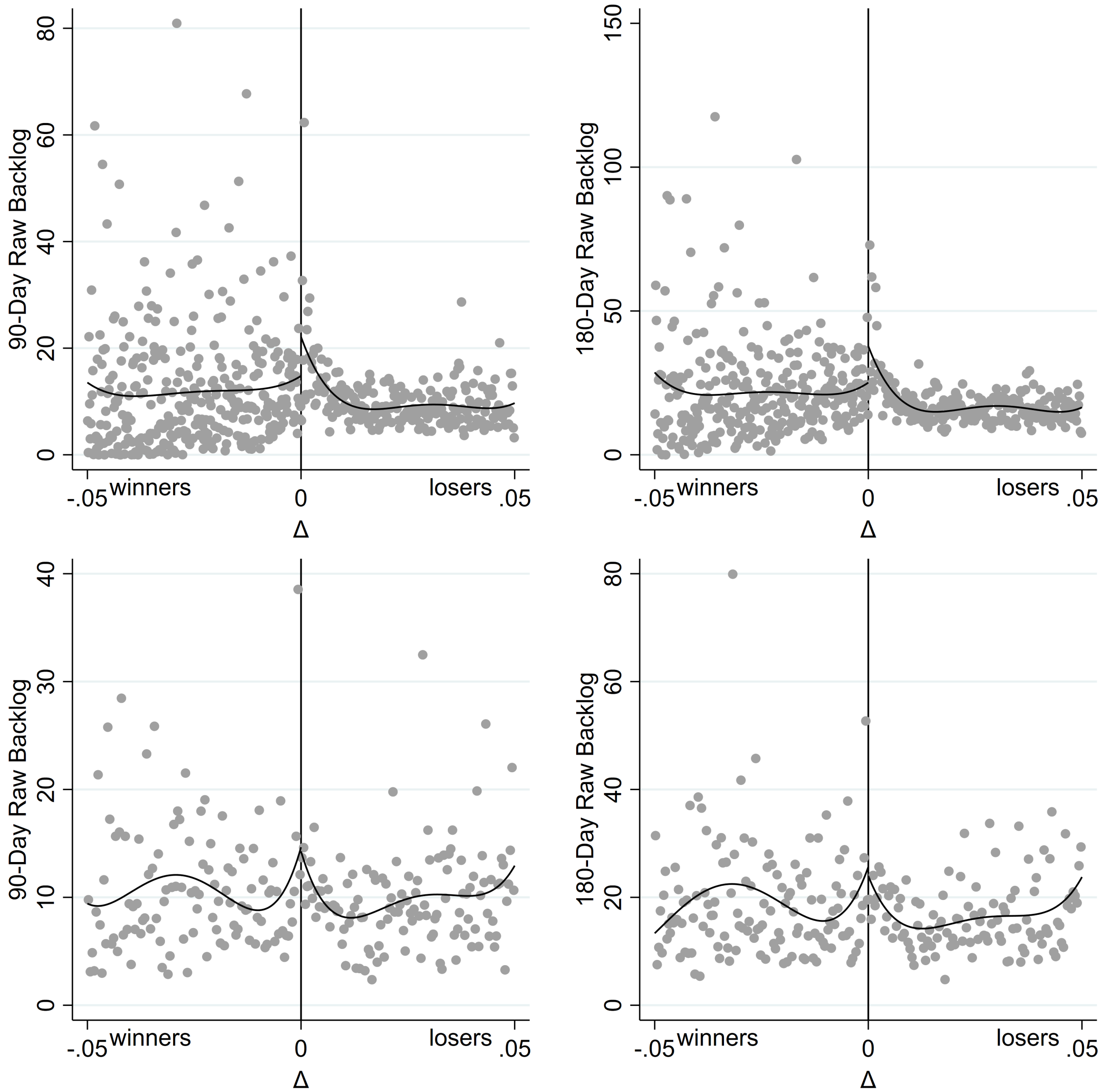


Note: The curves in the figure correspond to 4th order (global) polynomial approximations of the conditional means.

Figure B.1: Binned Scatter Plot for 90-Day and 180-Day Raw Backlog: Municipal Auctions from Japan.

Omitted binned scatter plots for Table 6. Figure B.2 displays binned scatter plots of 90-day and 180-day raw backlog corresponding to the regression results reported in columns (2) and (4) of Table 6. The top panels correspond to the results for the high bid sample (Panel (A)) and the bottom panels correspond to the low bid sample (Panel (B)). The left two panels plot the raw 90-day backlog against Δ and the right two panels plot the raw 180-day backlog against Δ. There is a modest discontinuity in the binned averages at $\Delta = 0$ in the top panels. In contrast, the graphs in the bottom panels, corresponding to columns (2) and (4) of Panel (B), do not exhibit any clear discontinuities at $\Delta = 0$.

Omitted binned scatter plots for Table 7. Figures B.3, B.4 and B.5 are the binned scatter plots corresponding to Table 7. In all of the panels, the horizontal axis corresponds to values $\Delta^2_{i,t} \equiv b_{i,t} - \min\{b_{j,t}, \text{s.t. } j \neq i \text{ and } j \text{ loses}\}$ for losing bidders i. A small negative value of $\Delta^2_{i,t}$ corresponds to a bid that is second lowest, but close to being third lowest. A small positive value of $\Delta^2_{i,t}$ corresponds to a bid that was higher than, but close to the second

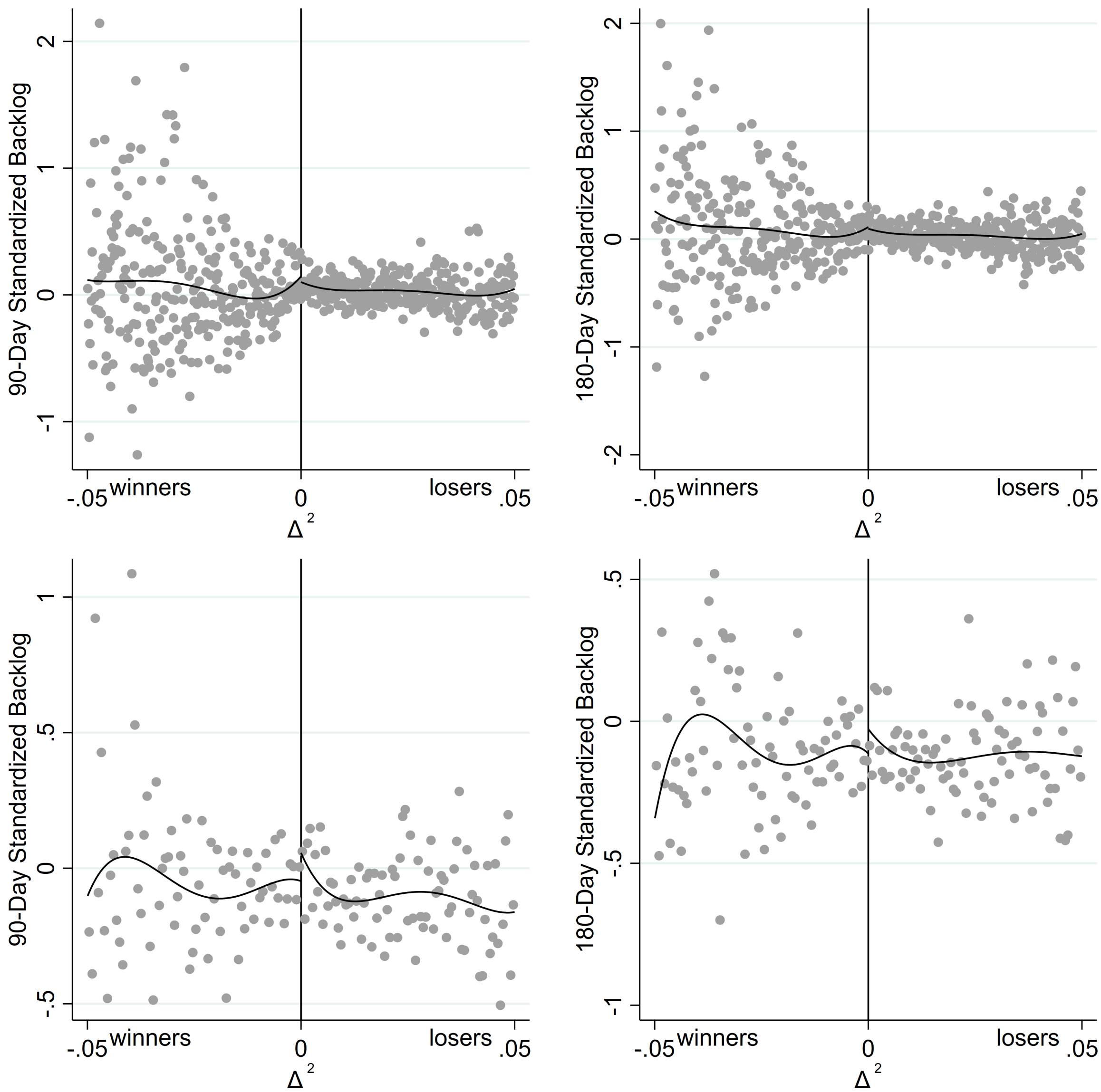


Note: Top panels correspond to Panel (A) of Table 6 and bottom panels correspond to Panel (B) of Table 6. Left panels correspond to 90-day raw backlog and the right panels correspond to 180-day standardized backlog.

Figure B.2: Binned Scatter Plot for 90-Day and 180-Day Raw Backlog: Municipal Auctions from Japan – High vs. Low bids.

lowest bid.

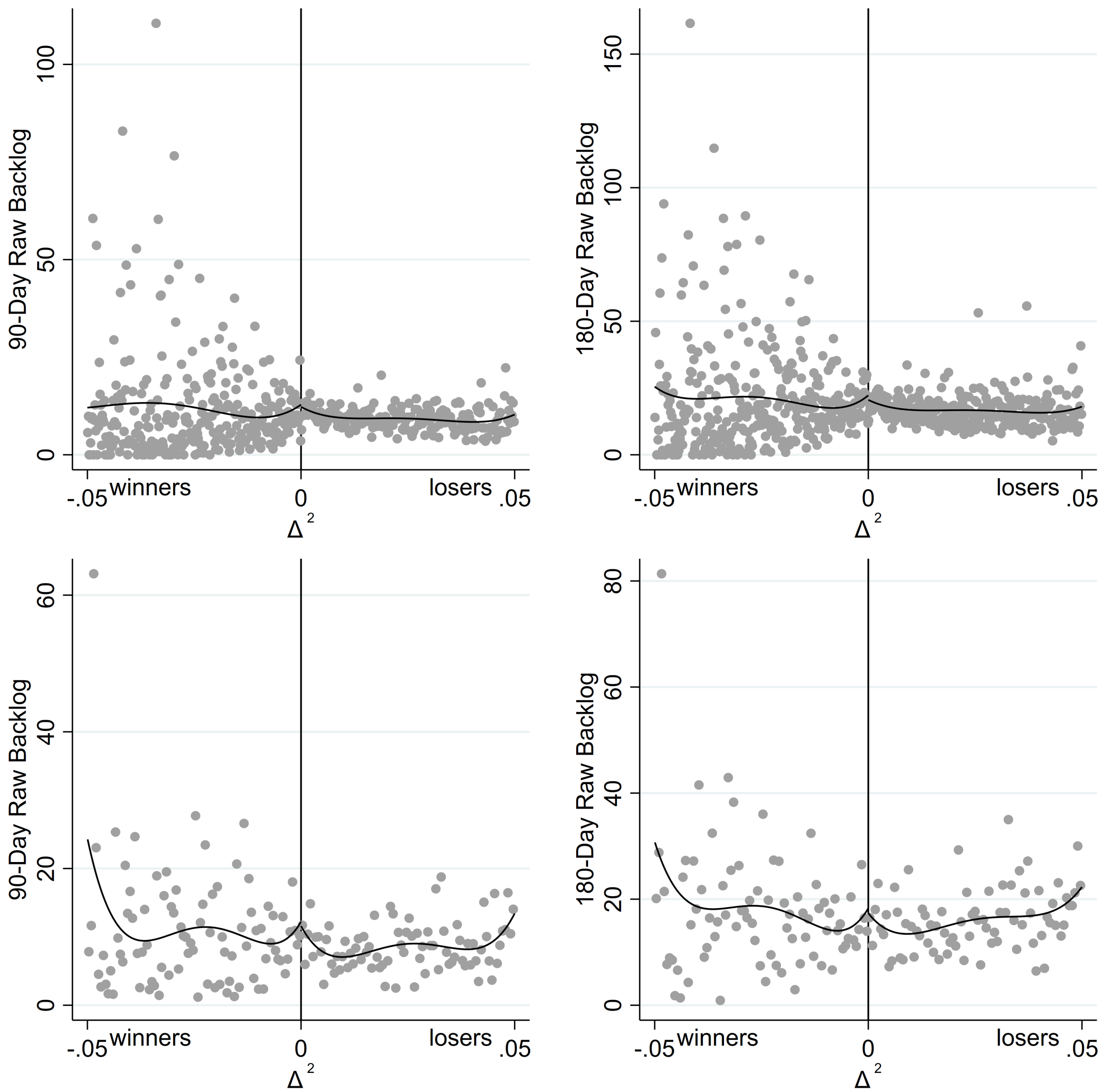
The panels in Figure B.3 are the binned scatter plots that correspond to columns (1) and (3) of Table 7. The panels in Figure B.4 correspond to columns (2) and (4) of Table 7. The panels in Figure B.5 correspond to column (5). The top panels of each figure correspond to



Note: Top panels correspond to columns (1) and (2) of Panel (A) of Table 7. Bottom panels correspond to columns (1) and (2) of Panel (B) of Table 7.

Figure B.3: Binned Scatter Plot for 90-Day Backlog with Respect to Δ^2: Municipal Auctions from Japan.

the sample of bids that are above the municipal median. The bottom panels correspond to the sample of bids that are below the median. Unlike our results for marginal winners and

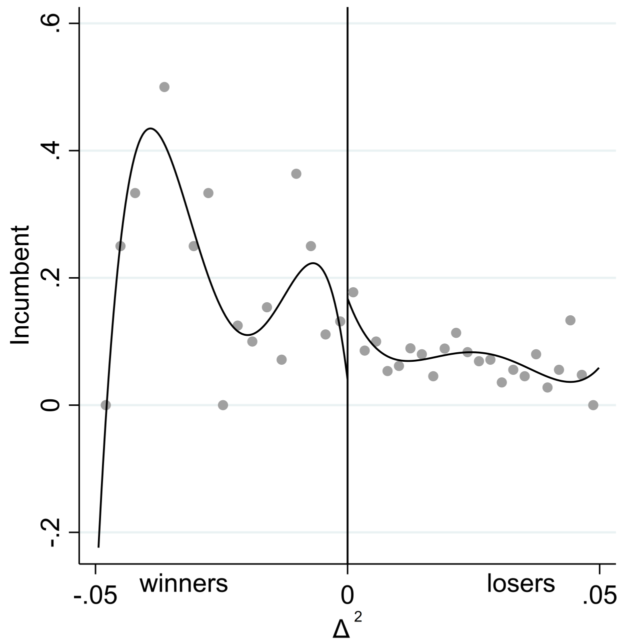


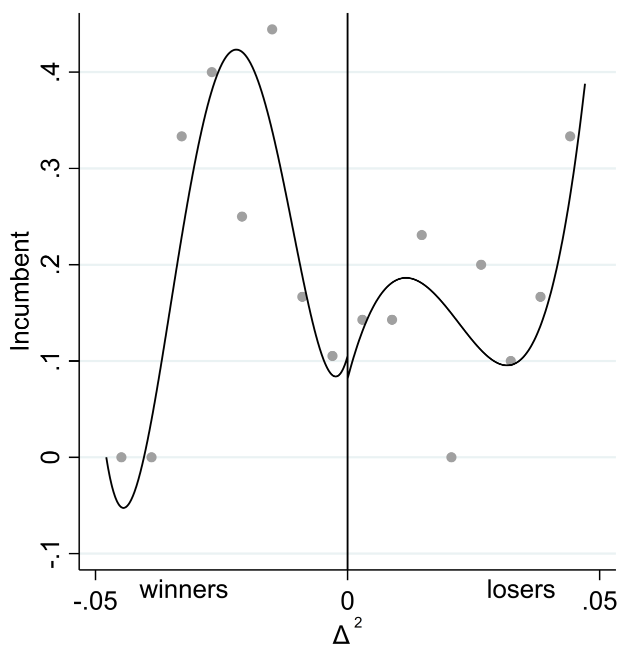
Note: Top panels correspond to columns (3) and (4) of Panel (A) of Table 7. Bottom panels correspond to columns (3) and (4) of Panel (B) of Table 7.

Figure B.4: Binned Scatter Plot for 180-Day Backlog with Respect to Δ^2: Municipal Auctions from Japan.

marginal losers, the figures do not show any discontinuities around $\Delta^2_{i,t} = 0$.

Partitioning auctions by the winning bid. In our main analysis, we partition the sample of bids according to whether or not the bids are above or below the median winning





Note: The top panel corresponds to column (5), Panel (A) of Table 7. The bottom panel corresponds to column (5), Panel (B) of Table 7.

Figure B.5: Binned Scatter Plot for Incumbent with Respect to Δ^2: Municipal Auctions from Japan.

bid. This results in the sample sizes of the two partitions to be unequal. In order to show that our results are not driven by differences in sample sizes, we consider an alternative partitioning in which we divide bids according to whether or not the winning bid of the auction is above or below the median winning bid.[37] This partitioning results in the same number of auctions in the two groups, and hence, roughly the same number of bids.

	(1)	(2)	(3)	(4)	(5)
	90-Day Backlog		180-Day Backlog		Incumbent
	Standardized	Raw	Standardized	Raw	
Panel (A) : Above Median					
$\hat{\beta}$	0.249***	7.155**	0.221***	13.205***	−0.277**
	(0.048)	(3.239)	(0.046)	(4.708)	(0.110)
h	0.019	0.014	0.025	0.013	0.031
Obs.	28,650	30,666	28,665	30,666	1,058
Panel (B) : Below Median					
$\hat{\beta}$	−0.058	−0.290	−0.022	−1.278	−0.268*
	(0.065)	(1.983)	(0.064)	(2.957)	(0.143)
h	0.021	0.027	0.021	0.024	0.027
Obs.	30,739	33,100	30,770	33,100	1,032

Note: We partition auctions into two depending on whether or not the winning bid is above or below the median. Panel (A) corresponds to the bids of auctions that are above the median. Panel (B) corresponds to the bids of auctions that are below the median. Standard errors are clustered at the auction level and reported in parenthesis. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table B.1: Partitioning Sample by Auctions: Municipal Auctions from Japan.

Table B.1 reports the results. The top panel corresponds to the sample of bids submitted in auctions in which the winning bid is higher than the median. We find that the estimate of β is statistically significant for all five regressions. The bottom panel corresponds to the sample of bids submitted in auctions in which the winning bid is below the median. We find that in Panel (B), none of the estimates of β are statistically significant at the 5%. Note that the sample sizes in Panel (A) and (B) are roughly equal. The results of Table B.1 suggests

[37] In particular, for each bid, we consider the winning bid of the auction. We partition the sample of bids depending on whether or not the winning bid is above or below the median.

that sample sizes are not driving our results in the main text.

Alternative standardization of backlog. In the main specification, we define the standardized backlog by subtracting the within-firm mean from the raw backlog and then dividing it by the within-firm standard error. Strictly speaking, standardized backlog defined this way is not measurable with respect to bidders' information at the time of bidding, as required by the theory. In order to define an outcome variable that is perfectly consistent with the theory, we consider an alternative standardization of backlog in which we use the mean and standard error of the rolling backlog as follows. Let

$$\mu'_{x_{i,t}^{B_k}} = \frac{1}{N_{i,t}} \sum_{\tau<t} x_{i,\tau}^{B_k},$$
$$\sigma'_{x_{i,t}^{B_k}} = \sqrt{\frac{1}{N_{i,t}} \sum_{\tau<t} (x_{i,\tau}^{B_k} - \mu'_{x_{i,t}^{B_k}})^2},$$

where $N_{i,t}$ is the number of auctions that firm i participates before auction t, $\mu'_{x_{i,t}^{B_k}}$ is the average of firm i's backlog up to auction t, and $\sigma'_{x_{i,t}^{B_k}}$ is the standard deviation of firm i's backlog up to auction t. We define standardized backlog as

$$x_{i,t}^{\overline{B}_k} = \frac{x_{i,t}^{B_k} - \mu'_{x_{i,t}^{B_k}}}{\sigma'_{x_{i,t}^{B_k}}}.$$

The difference between this definition and the one in the main text is that we now consider only auctions that take place before auction t in the summation ($\tau < t$). Note that the new definition of standardized backlog is measurable with respect to bidders' information at the time of bidding.

We estimate β using a local linear regression as follows:

$$\widehat{\beta} = \widehat{b_0^+} - \widehat{b_0^-}, \text{ with}$$
$$(\widehat{b_0^+}, \widehat{b_1^+}) = \arg\min \textstyle\sum_{i,t}^{T} (x_{i,t}^{B_k} - b_0^+ - b_1^+ \Delta_{i,t})^2 K\left(\frac{\Delta_{i,t}}{h_n}\right) \mathbf{1}_{\{\Delta_{i,t}>0\}\cap\{x_{i,t}^{B_k}\neq 0\}},$$
$$(\widehat{b_0^-}, \widehat{b_1^-}) = \arg\min \textstyle\sum_{i,t}^{T} (x_{i,t}^{B_k} - b_0^- - b_1^- \Delta_{i,t})^2 K\left(\frac{\Delta_{i,t}}{h_n}\right) \mathbf{1}_{\{\Delta_{i,t}<0\}\cap\{x_{i,t}^{B_k}\neq 0\}}.$$

Note that we condition our regression discontinuity estimate on the event $\{x_{i,t}^{B_k} \neq 0\}$. Because this event is measurable with respect to firms' information, Corollary 2 holds.

	(1)	(2)
	90-Day Rolling Backlog	180-Day Rolling Backlog
Panel (A) : Above Median		
$\hat{\beta}$	0.599***	0.354***
	(0.076)	(0.069)
h	0.020	0.018
Obs.	20,343	26,449
Panel (B) : Below Median		
$\hat{\beta}$	0.198*	0.100
	(0.116)	(0.102)
h	0.026	0.025
Obs.	7,596	9,180

Note: Panel (A) corresponds to the sample of bids above the median winning bid. Panel (B) corresponds to the sample of bids below the median. Standard errors are clustered at the auction level and reported in parenthesis. The table also reports the bandwidth h used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table B.2: Alternative Standardization of Backlog: Municipal Auctions from Japan.

Table B.2 reports the results. Panel (A) corresponds to the sample of bids above the median winning bid and Panel (B) corresponds to the sample of bids below the median. The estimates for Panel (A) are statistically significant at the 5% level while the estimates in Panel (B) are not. The results of Table B.2 are similar to the results we report in column (2) and (4) of Table 6.

Results for the sample of auctions with a public reserve price. We now report the results of our analysis when we restrict the sample to auctions let by municipalities using public reserve prices. Table B.3 reports the results. Panel (A) of Table B.3 reports estimation results for the set of bids above the municipal median.[38] Although the estimate of β is not statistically significant for the 90-day raw backlog in column (2), we find statistically significant differences between marginal losers and marginal winners for other measures of

[38]As before, we compute the median winning bid for each municipality and divide the sample according to whether or not a bid is above or below the municipal median.

	(1)	(2)	(3)	(4)	(5)
	90-Day Backlog		180-Day Backlog		Incumbent
	Standardized	Raw	Standardized	Raw	
Panel (A) : Above Median					
$\hat{\beta}$	0.242**	12.375	0.239**	24.760**	−0.025
	(0.117)	(8.449)	(0.120)	(11.607)	(0.276)
h	0.006	0.007	0.006	0.006	0.009
Obs.	10,665	11,432	10,674	11,432	464
Panel (B) : Below Median					
$\hat{\beta}$	−0.002	−1.895	−0.012	−1.166	−0.306
	(0.136)	(6.476)	(0.131)	(8.432)	(0.441)
h	0.022	0.021	0.018	0.019	0.017
Obs.	2,675	2,883	2,676	2,883	61

Note: Panel (A) corresponds to the sample of bids above the median. Panel (B) corresponds to the sample of bids below the median. Standard errors are clustered at the auction level and reported in parenthesis. The table also reports the bandwidth h used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

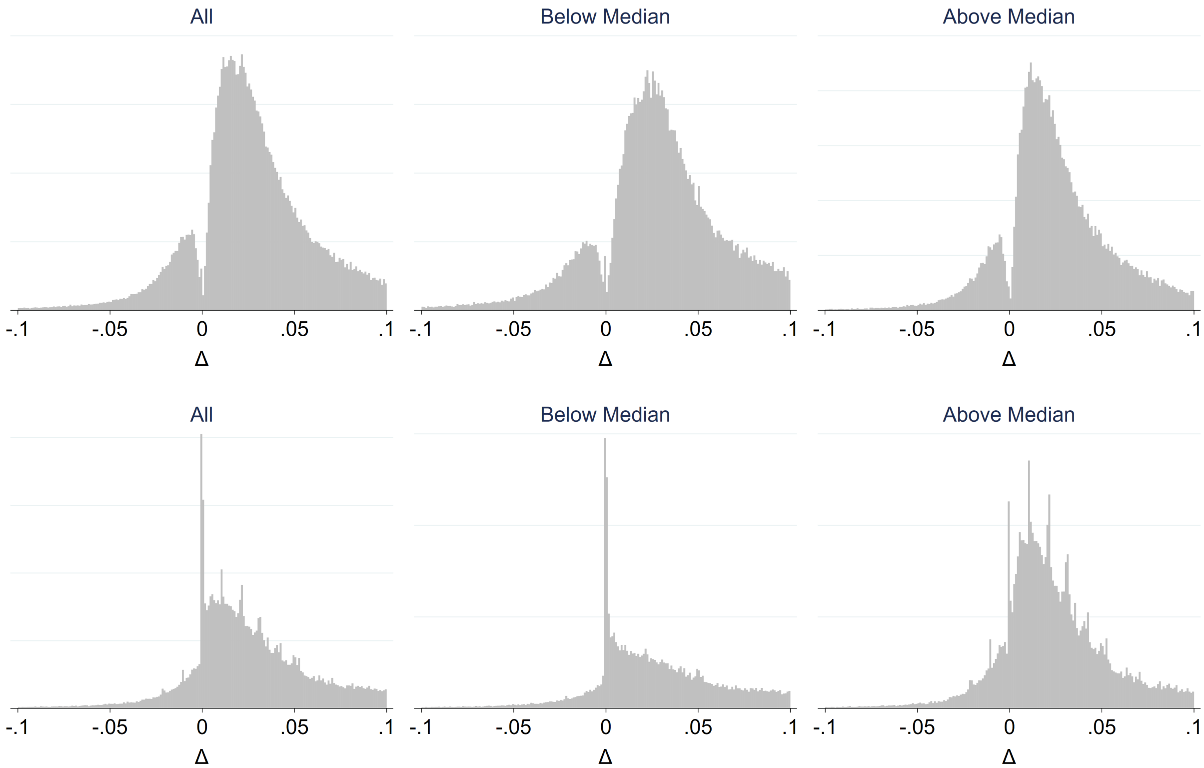
Table B.3: Restricting the Sample to Municipalities with Public Reserve Price.

backlog in columns (1), (3), and (4). These results are qualitatively similar to those reported in Table 6. The results overall strongly suggest that there are non-competitive auctions among the sample of public reserve auctions in which a close winner submits a high bid.

In Panel (B), we report the results for the set of low bids. We find that there are no statistically significant differences between the marginal winner and the marginal loser for this subset, implying that we cannot reject the null of competition.

All municipalities. We now discuss the results of our tests when we include auctions from Japanese municipalities that we drop in our main analysis. There are a total of 109 municipalities for which we have auction data. Recall that, in order to construct the dataset used in Section 5.2, we drop municipalities for which the distribution of Δ has a missing mass at 0 (71 municipalities) and those for which the distribution of Δ has a point mass at exactly 0 (22 municipalities).

Figure B.6 plots the histograms of $\Delta_{i,t}$ for auctions let by the municipalities with missing mass in the distribution of $\Delta_{i,t}$ at 0 (first row) and for those let by municipalities with a mass in the distribution of $\Delta_{i,t}$ at exactly zero (second row). The left two panels correspond to the histogram for all of the auctions let by each of the groups of municipalities. The middle and right panels correspond to the histogram for bids below the municipal median (middle panel) and above the municipal median (right panel).



Note: The top panels correspond to auctions from 71 municipalities with missing mass in the distribution of $\Delta_{i,t}$ at zero. The bottom panels correspond to auctions from 22 municipalities with a mass in the distribution of $\Delta_{i,t}$ at exactly zero. The left panels correspond to all auctions let by each of these groups, the middle panels condition on the winning bids to be below the municipality median and the right panels condition on the winning bids to be above the median.

Figure B.6: Histogram of $\Delta_{i,t}$: Municipal Auctions from Japan.

The missing mass in the distribution of $\Delta_{i,t}$, apparent in the top panels, has previously been documented in Chassang et al. (2020). In that paper, we show that this distinctive pattern in the distribution of $\Delta_{i,t}$ is inconsistent with competitive bidding under fairly general conditions. Because our previous paper specifically focuses on the implications of these patterns, we opted to exclude these municipalities in our baseline analysis.

	(1) Sample with Missing Mass (71 munis)		(2) Sample with Mass at 0 (22 munis)		(3) Baseline Sample (16 munis)	
	Mean	Std.	Mean	Std.	Mean	Std.
Panel A: By Auction						
Reserve (Mil. Yen)	24.03	104.39	20.92	101.08	22.26	77.14
Winning Bid (Mil. Yen)	22.60	97.64	19.09	95.63	20.71	71.78
Win Bid/Reserve	0.940	0.073	0.911	0.078	0.926	0.083
# of Bids	7.44	3.78	8.00	4.64	6.80	4.21
Incumbent	0.064	0.244	0.043	0.202	0.044	0.204
Obs.	44,993		54,153		11,207	
Panel B: By Bidder						
# of Participation	23.80	60.13	32.65	74.45	22.56	45.93
# of Wins	3.20	8.26	4.23	10.17	3.32	6.97
Raw Backlog (90-Day)	4.02	21.15	4.75	19.19	4.11	17.16
Raw Backlog (180-Day)	6.58	35.09	7.18	26.30	6.45	22.85
Obs.	14,065		11,605		3,377	

Note: Column (1) reports summary statistics for the sample of auctions with missing mass in the distribution of $\Delta_{i,t}$ at zero (71 municipalities). Column (2) reports summary statistics for the sample with mass at exactly zero (22 municipalities). Column (3) reports sample statistics for the sample used in Section 5.

Table B.4: Summary Statistics by Auctions and Bidders: All Municipalities.

The distributions of $\Delta_{i,t}$ in the bottom panels have spikes at zero which are the result of binding price floors. Price floors can result in multiple bidders bidding exactly at the price floor. Note that because the spikes are generated by price floors, and because multiple bids at the price floor typically imply that the winning bid of the auction is low, the spike is very pronounced for the middle panel, but mostly disappears in the right panel. The summary statistics of the auctions for each of the groups are reported in Table B.4. Column (1) corresponds to the sample statistics for municipalities with a missing mass at zero, column (2) corresponds to the sample statistics for those with a mass at 0, and column (3) corresponds to the sample statistics for the baseline sample used in Section 5.

We now report the regression discontinuity results for all of the auctions in our sample. Panel (A) of Table B.5 reports the regression discontinuity estimates for bids above the median winning bid. Panel (B) of Table B.5 reports the estimates for bids below the median winning bid. Focusing on Panel (A), we find that marginal losing bidders have about 0.086 higher 90-day standardized backlog (column (1)) and about 3.2 million yen more in terms of 90-day backlog (column (2)) than marginal winners. The estimates are statistically significant at the 5% level. Similarly, we find that marginal losing bidders have higher standardized and raw 180-day backlog (column (3), (4)) than marginal winners, and are less likely to be an incumbent (column (5)) than marginal winners. The coefficients are all statistically significant at the 1% level. These findings lead us to reject the null hypothesis of competition for this sample.

The bottom panel of Table B.5 reports the results for bids below the median. While the regression discontinuity estimate is statistically significant at the 5% level in columns (2), (4), and (5), the estimated differences between marginal winners and losers are smaller than in Panel (A). The results suggest the existence of some collusive bidding among this sample, but likely to a lesser extent than the sample in Panel (A). Overall, the results of Table B.5 suggest that the null of competitive bidding is strongly rejected for the sample of high bids, but that the evidence is less strong for the sample of low winning bids. This is consistent with the expectation that there would be more collusion among auctions with high winning bids than among those with low winning bids.

C Proofs

C.1 Proofs for Section 3

Proof of Lemma 1. We show that for all $\eta > 0$, there exists $\epsilon > 0$ small enough such that for all histories $h_{i,t}$,

$$\left| \mathsf{prob}(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| \leq \epsilon) - \frac{1}{2} \right| \leq \eta.$$

By assumption, $D_i'(b_i|h_i)$ is continuous in $b_i \in [0,1]$ and strictly negative for all histories $h_i = (\theta, z_i)$. Since there are finitely many histories (θ, z_i), it follows that there exists $\nu > 0$ such that $D_i'(b_i|h_{i,t}) \leq -\nu$ for all b_i and all histories $h_{i,t}$. In addition, for all $\widehat{\eta} > 0$, there exists ϵ small enough that for all $\widehat{b}_i \in [b_i - \epsilon, b_i + \epsilon]$, $|D_i'(\widehat{b}_i|h_{i,t}) - D_i'(b_i|h_{i,t})| \leq \widehat{\eta}$.

	(1)	(2)	(3)	(4)	(5)
	90-Day Backlog		180-Day Backlog		Incumbent
	Standardized	Raw	Standardized	Raw	
Panel (A) : Above Median					
$\hat{\beta}$	0.086^{***}	3.218^{**}	0.117^{***}	6.506^{***}	-0.341^{***}
	(0.019)	(1.279)	(0.019)	(2.175)	(0.052)
h	0.012	0.013	0.013	0.009	0.011
Obs.	530,966	558,152	531,357	558,152	19,929
Panel (B) : Below Median					
$\hat{\beta}$	-0.026	1.484^{**}	-0.012	2.366^{**}	-0.091^{***}
	(0.016)	(0.701)	(0.018)	(1.168)	(0.034)
h	0.013	0.015	0.011	0.014	0.012
Obs.	177,304	188,032	177,457	188,032	6,864

In addition to the auctions used in the baseline analysis, we include auctions from 70 municipalities with missing mass in the distribution of $\Delta_{i,t}$ at zero and those from 18 municipalities with mass in the distribution of $\Delta_{i,t}$ at exactly zero. Panel (A) corresponds to the sample of bids above the median. Panel (B) corresponds to the sample of bids below the median. Standard errors are clustered at the auction level and reported in parenthesis. The forcing variable is Δ^1. The table also reports the bandwidth used for the estimation. *, **, and *** respectively denote significance at the 10%, 5%, and 1% levels.

Table B.5: Regression Discontinuity Estimates: All Municipalities.

This implies that for ϵ small

$$\begin{aligned}\left|\mathsf{prob}(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge\mathbf{b}_{-i,t}| \leq \epsilon) - \frac{1}{2}\right| &= \left|\frac{D_i(b_{i,t}|h_{i,t}) - D_i(b_{i,t}+\epsilon|h_{i,t})}{D_i(b_{i,t}-\epsilon|h_{i,t}) - D_i(b_{i,t}+\epsilon|h_{i,t})} - \frac{1}{2}\right| \\ &\leq \left|\frac{-\epsilon D_i'(b_{i,t}|h_{i,t}) - \epsilon\widehat{\eta}}{-2\epsilon D_i'(b_{i,t}|h_{i,t}) + 2\epsilon\widehat{\eta}} - \frac{1}{2}\right| \\ &\leq \left|\frac{\nu - \widehat{\eta}}{2\nu + 2\widehat{\eta}} - \frac{1}{2}\right|.\end{aligned}$$

Lemma 1 follows by taking $\widehat{\eta}$ small enough. ■

Proof of Corrolary 1. Note that, for each $x \in X$,

$$\begin{aligned}\mathsf{prob}\left(x_i = x \,|\, \Delta_{i,t} \in (-\epsilon, 0)\right) =& \mathsf{prob}(x_{i,t} = x | i \text{ wins } \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \\ =& \mathsf{prob}(x_{i,t} = x || b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \frac{\mathsf{prob}(i \text{ wins } | \, x_{i,t} = x \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon)}{\mathsf{prob}(i \text{ wins } || b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon)}\end{aligned}$$

Similarly,

$$\begin{aligned}\mathsf{prob}\left(x_i = x \,|\, \Delta_{i,t} \in (0, \epsilon)\right) =& \mathsf{prob}(x_{i,t} = x | i \text{ loses } \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \\ =& \mathsf{prob}(x_{i,t} = x || b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \frac{\mathsf{prob}(i \text{ loses } | \, x_{i,t} = x \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon)}{\mathsf{prob}(i \text{ loses } || b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon)}\end{aligned}$$

By Lemma 1, we have that

$$\lim_{\epsilon \searrow 0} \mathsf{prob}(i \text{ wins } | \, x_{i,t} = x \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) = \lim_{\epsilon \searrow 0} \mathsf{prob}(i \text{ wins } || b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) = \frac{1}{2},$$
$$\lim_{\epsilon \searrow 0} \mathsf{prob}(i \text{ loses } | \, x_{i,t} = x \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) = \lim_{\epsilon \searrow 0} \mathsf{prob}(i \text{ loses } || b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) = \frac{1}{2}.$$

Hence, for each $x \in X$,

$$\lim_{\epsilon \searrow 0} \left| \mathsf{prob}\left(x_i = x \,|\, \Delta_{i,t} \in (-\epsilon, 0)\right) - \mathsf{prob}\left(x_i = x \,|\, \Delta_{i,t} \in (0, \epsilon)\right) \right| = 0.$$

Since X is finite, for all $\eta > 0$ there exists $\epsilon > 0$ small enough such that for all $x \in X$,

$$\left| \mathsf{prob}\left(x_i = x \,|\, \Delta_{i,t} \in (0, \epsilon)\right) - \mathsf{prob}\left(x_{i,t} = x \,|\, \Delta_{i,t} \in (-\epsilon, 0)\right) \right| < \eta.$$

This completes the proof. ■

C.2 Proofs for Section 4

We now establish Proposition 1. Throughout this section we consider an environment $\mathcal{E}$ and an MPE σ that is competitively enforced. We begin by establishing two intermediary lemmas. Recall that continuation value $V_i(\zeta_i, b_w | h_i)$ does not depend on winning bid b_w when bidder i wins. Hence, we suppress the dependency of V_i on b_w when $\zeta_i = 1$.

Lemma C.1 (minimum demand)**.** *There exists $\nu > 0$ such that for every history $h_i = (\theta, z_i)$*

and bid $b_i \in [0,1]$ in the support of $\sigma_{i|h_i}$, $D_i(b_i|h_i) \geq \nu$. In addition,

$$b_i - c_i + \delta \mathbb{E}_\sigma \left[V_i(1|h_i) - V_i(0, \wedge \mathbf{b}_{-i}|h_i) \,|\, h_i, b_i \prec \wedge \mathbf{b}_{-i} \right] \geq k.$$

Proof. Since firm i chooses to participate, it must be that

$$\begin{aligned}
\mathbb{E}_\sigma \left[\mathbf{1}_{b_i \prec \wedge \mathbf{b}_{-i}} (b_i - c_i + \delta V_i(1|h_i)) + \mathbf{1}_{b_i \succ \wedge \mathbf{b}_{-i}} \delta V_i(0, \wedge \mathbf{b}_{-i}|h_i) \,|\, h_i \right] - k &\geq \mathbb{E}_\sigma \left[\delta V_i(0, \wedge \mathbf{b}_{-i}|h_i) \,|\, h_i \right] \\
\iff \mathbb{E}_\sigma \left[\mathbf{1}_{b_i \prec \wedge \mathbf{b}_{-i}} (b_i - c_i + \delta V_i(1|h_i) - \delta V_i(0, \wedge \mathbf{b}_{-i}|h_i)) \,|\, h_i \right] &\geq k \\
\iff D_i(b_i|h_i) \left(b_i - c_i + \delta \mathbb{E}_\sigma \left[V_i(1|h_i) - V_i(0, \wedge \mathbf{b}_{-i}|h_i) \,|\, h_i, b_i \prec \wedge \mathbf{b}_{-i} \right] \right) &\geq k
\end{aligned}$$

Since $D_i \geq 0$, it must be that both left-hand side factors are strictly positive. In addition, since continuation values are bounded by some constant $\overline{V}$, it follows that $D_i(b_i|h_i) \geq k/(1+2\overline{V})$. Similarly, since demand is bounded above by 1, we have that

$$b_i - c_i + \delta \mathbb{E}_\sigma \left[V_i(1|h_i) - V_i(0, \wedge \mathbf{b}_{-i}|h_i) \,|\, h_i, b_i \prec \wedge \mathbf{b}_{-i} \right] \geq k.$$

This concludes the proof. $\square$

Lemma C.2 (continuous demand)**.** *For every history $h_i = (\theta, z_i)$, residual demand $D_i(b_i|h_i)$ is continuous in b_i over $(0,1)$.*

Proof. The proof is by contradiction. Assume that demand $D_i(\cdot|h_i)$ is discontinuous at bid b_0. There must exist a bidder j and a history $h_j = (\theta, z_j)$ such that firm j bids $b_j = b_0$ with probability $q > 0$. By Lemma C.1, bidder j must win with probability at least $\nu > 0$ when bidding b_0.

Consider a bidder l and a history $h_l = (\theta, z_l)$ such that history h_j has positive probability, and bidder l loses with positive probability against bidder j when bidder j bids b_0. Since the number of histories is finite, there exists $\nu_1 > 0$ such that at any such history h_l bidder j bids b_0 with positive probability ν_1.

Pick $\epsilon > 0$ and consider the payoff of bidder l bidding $b_l \in [b_0, b_0 + \epsilon)$. Bidder l gets payoff (excluding participation costs and payoffs upon non-participation)

$$U_l(b_l|h_l, c_l) = D(b_l|h_l) \left(b_l - c_l + \delta \mathbb{E}_\sigma [V_l(1|h_l) - V_l(0, \wedge \mathbf{b}_{-l}|h_l) \,|\, h_l, b_l \prec \wedge \mathbf{b}_{-l}] \right).$$

We know from Lemma C.1 that

$$b_l - c_l + \delta \mathbb{E}_\sigma \left[V_l(1|h_l) - V_l(0, \wedge \mathbf{b}_{-l}|h_l) \,|\, h_l, b_l \prec \wedge \mathbf{b}_{-l} \right] \geq k.$$

Since bidding behavior is not sensitive, there exists a Lipcshitz constant $L > 0$ such that

$$\mathbb{E}_\sigma \left[V_l(0, \wedge \mathbf{b}_{-l} | h_l) \,|\, h_l, b_l - \epsilon \prec \wedge \mathbf{b}_{-l} \right] \leq \mathbb{E}_\sigma \left[V_l(0, \wedge \mathbf{b}_{-l} | h_l) \,|\, h_l, b_l \prec \wedge \mathbf{b}_{-l} \right] + \epsilon L.$$

Altogether, it follows that for every $\eta > 0$, there exists $\epsilon > 0$ small enough that

$$\begin{aligned} b_l - \epsilon - c_l + &\delta \mathbb{E}_\sigma \left[V_l(1|h_l) - V_l(0, \wedge \mathbf{b}_{-l} | h_l) \,|\, h_l, b_l - \epsilon \prec \wedge \mathbf{b}_{-l} \right] \\ &\geq b_l - c_l + \delta \mathbb{E}_\sigma \left[V_l(1|h_l) - V_l(0, \wedge \mathbf{b}_{-l} | h_l) \,|\, h_l, b_l \prec \wedge \mathbf{b}_{-l} \right] - \eta \geq k - \eta. \end{aligned}$$

Hence, it follows that by bidding $b_l - \epsilon$, bidder l gets a payoff

$$\begin{aligned} U_l(b_l - \epsilon | h_l, c_l) &= D(b_l - \epsilon | h_l) \left(b_l - \epsilon - c_l + \delta \mathbb{E}_\sigma [V_i(1|h_l) - V_i(0, \wedge \mathbf{b}_{-l} | h_l) \,|\, h_l, b_l - \epsilon \prec \wedge \mathbf{b}_{-l}] \right) \\ &\geq U_l(b_l | h_l, c_l) - \eta + \nu_1 (k - \eta) \end{aligned}$$

Since ν_1 is fixed, it follows that for ϵ small enough $U_l(b_l - \epsilon | h_l, c_l) > U_l(b_l | h_l, c_l)$. Hence, there exists ϵ small such that bidder l does not bid in $[b_0, b_0 + \epsilon)$. Since there are only finite histories, this implies that there exists $\epsilon > 0$ such that no bidder l that loses against bidder j bidding b_0 bids in the range $[b_0, b_0 + \epsilon)$. Hence, bidder j would benefit from bidding $b_0 + \epsilon/2$ rather than b_0. This contradicts the assumption that σ is an MPE and concludes the proof. $\square$

Proof of Proposition 1. Consider an environment $\mathcal{E}$ and an MPE σ that is competitively enforced. Fix a history $h_{i,t} = (\theta_t, z_{i,t})$ of firm i. Let $b_{i,t} < r = 1$ denote firm i's bid at this history when her costs are $c_{i,t}$. For any bid b, let $U_i(b | h_{i,t}, c_{i,t})$ denote i's payoff from bidding b at history $h_{i,t}$ when her cost is $c_{i,t}$:

$$U_i(b | h_{i,t}, c_{i,t}) = \mathbb{E}_\sigma \left[\mathbf{1}_{\wedge \mathbf{b}_{-i,t} \succ b} (b - c_{i,t} + \delta V_i(1 | h_{i,t}) + (1 - \mathbf{1}_{\wedge \mathbf{b}_{-i,t} \succ b}) \delta V_i(0, \wedge \mathbf{b}_{-i,t} | h_{i,t}) \,|\, h_{i,t} \right] - k.$$

Since bid $b_{i,t}$ is optimal, for all $\epsilon > 0$ it must be that,

$$\begin{aligned} &U_i(b_{i,t} | h_{i,t}, c_{i,t}) \geq U_i(b_{i,t} + \epsilon | h_{i,t}, c_{i,t}) \\ \iff &(D_i(b_{i,t} | h_{i,t}) - D_i(b_{i,t} + \epsilon | h_{i,t}))(b_{i,t} - \kappa^{+\epsilon}_{i,t}) \geq D_i(b_{i,t} + \epsilon | h_{i,t}) \times \epsilon \end{aligned} \tag{8}$$

where $\kappa^{+\epsilon}_{i,t} \equiv c_{i,t} - \delta \mathbb{E}^\sigma [V_i(1 | h_{i,t}) - V_i(0, \wedge \mathbf{b}_{-i,t} | h_{i,t}) | h_{i,t}, b_{i,t} + \epsilon \succ \wedge \mathbf{b}_{-i,t} \succ b_{i,t}]$. Since $D_i(\cdot | h_{i,t})$ is continuous at $b_{i,t}$ (Lemma C.2), and since $D_i(b_{i,t} | h_{i,t}) > 0$ (Lemma C.1) it must be that $b_{i,t} - \kappa^{+\epsilon}_{i,t} > 0$ for $\epsilon > 0$ small.

Similarly, for all $\epsilon > 0$ it must be that

$$\begin{aligned} & U_i(b_{i,t}|h_{i,t}, c_{i,t}) \geq U_i(b_{i,t} - \epsilon|h_{i,t}, c_{i,t}) \\ \iff & (D_i(b_{i,t} - \epsilon|h_{i,t}) - D_i(b_{i,t}|h_{i,t}))(b_{i,t} - \kappa^{+\epsilon}_{i,t}) \leq D_i(b_{i,t} - \epsilon|h_{i,t}) \times \epsilon \\ & \qquad - (D_i(b_{i,t} - \epsilon|h_{i,t}) - D_i(b_{i,t}|h_{i,t}))(\kappa^{+\epsilon}_{i,t} - \kappa^{-\epsilon}_{i,t}) \end{aligned} \tag{9}$$

where $\kappa^{-\epsilon}_{i,t} \equiv c_{i,t} - \delta \mathbb{E}^\sigma[V_i(1|h_{i,t}) - V_i(0, \wedge \mathbf{b}_{-i,t}|h_{i,t})|h_{i,t}, b_{i,t} \succ \wedge \mathbf{b}_{-i,t} \succ b_{i,t} - \epsilon]$.

Using (8) and (9), together with $b_{i,t} - \kappa^{+\epsilon}_{i,t} > 0$, we have that

$$\begin{aligned} & \mathsf{prob}_\sigma(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \\ & \quad = \frac{D_i(b_{i,t}|h_{i,t}) - D_i(b_{i,t} + \epsilon|h_{i,t})}{D_i(b_{i,t} - \epsilon|h_{i,t}) - D_i(b_{i,t} + \epsilon|h_{i,t})} \\ & \quad = \frac{D_i(b_{i,t}|h_{i,t}) - D_i(b_{i,t} + \epsilon|h_{i,t})}{D_i(b_{i,t} - \epsilon|h_{i,t}) - D_i(b_{i,t}|h_{i,t}) + D_i(b_{i,t}|h_{i,t}) - D_i(b_{i,t} + \epsilon|h_{i,t})} \\ & \quad \geq \frac{D_i(b_{i,t} + \epsilon|h_{i,t})}{D_i(b_{i,t} - \epsilon|h_{i,t}) - (D_i(b_{i,t} - \epsilon|h_{i,t}) - D_i(b_{i,t}|h_{i,t}))\frac{\kappa^{+\epsilon}_{i,t} - \kappa^{-\epsilon}_{i,t}}{\epsilon} + D_i(b_{i,t} + \epsilon|h_{i,t})}. \end{aligned} \tag{10}$$

Since $D_i(\cdot|\theta_t, z_{i,t})$ is continuous on $[0, 1]$, it is uniformly continuous. Since there are finitely many (θ, z_i), for every $\gamma_D > 0$ there exists $\bar{\epsilon} > 0$ such that, for all i, θ, z_i and for all b, b' with $|b - b'| \leq 2\bar{\epsilon}$, $D_i(b|\theta, z_i) - D_i(b'|\theta, z_i) < \gamma_D$.

Moreover, since bidding behavior is not sensitive, and since there are finitely many $h_i = (\theta, z_i)$, there exists a Lipschitz constant $L > 0$ such that, for all $i, \theta, z_i, c_{i,t}$, $\kappa^{+\epsilon}_{i,t} - \kappa^{-\epsilon}_{i,t} \geq -2\epsilon L$.

Using (10), for every $\gamma_D > 0$, there exists $\tilde{\epsilon} > 0$ such that, for all $\epsilon < \tilde{\epsilon}$,

$$\begin{aligned} & \mathsf{prob}_\sigma(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \\ & \quad \geq \frac{D_i(b_{i,t} + \epsilon|h_{i,t})}{D_i(b_{i,t} + \epsilon|h_{i,t}) + \gamma_D + 2\gamma_D L + D_i(b_{i,t} + \epsilon|h_{i,t})} \\ & \quad \geq \frac{\nu - \gamma_D}{\nu + \gamma_D 2L + \nu - \gamma_D}, \end{aligned} \tag{11}$$

where the second inequality uses the inequality $D_i(b_{i,t} + \epsilon|h_{i,t}) \geq D_i(b_{i,t}|h_{i,t}) - \gamma_D \geq \nu - \gamma_D$ (Lemma C.1). Picking γ_D small, we obtain that $\mathsf{prob}_\sigma(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \geq 1/2 - \eta$. ■

Proof of Corollary 2. For each $\epsilon > 0$, let $\mathsf{prob}_\sigma(\cdot|\epsilon\text{-}\mathsf{close})$ denote the distribution over histories conditional on event ϵ-close. Then, for each $i \in N$ and each $\epsilon > 0$, the probability with which firm i wins an auction under σ conditional on event ϵ-close satisfies

$$\mathsf{prob}_\sigma(i \text{ wins}|\epsilon\text{-}\mathsf{close}) = \mathbb{E}_{\mathcal{E},\sigma}\left[\mathsf{prob}_\sigma(i \text{ wins } | \, h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon)\middle| \, \epsilon\text{-}\mathsf{close}\right] \times \mathsf{prob}_\sigma(|b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon | \epsilon\text{-}\mathsf{close}). \tag{12}$$

By Proposition 1, it follows that

$$\forall i \in N, \quad \liminf_{\epsilon \searrow 0} \mathbb{E}_{\mathcal{E},\sigma}\left[\mathsf{prob}_\sigma(i \text{ wins } | \, h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon)\middle| \, \epsilon\text{-}\mathsf{close}\right] \geq \frac{1}{2}. \tag{13}$$

Towards a contradiction, suppose that the result is not true. Hence, there exists a player j and a number $\eta > 0$ such that

$$\limsup_{\epsilon \searrow 0} \mathbb{E}_{\mathcal{E},\sigma}\left[\mathsf{prob}(j \text{ wins } | \, h_{j,t} \text{ and } |b_{j,t} - \wedge \mathbf{b}_{-j,t}| < \epsilon)\middle| \, \epsilon\text{-}\mathsf{close}\right] \geq \frac{1}{2} + \eta. \tag{14}$$

Note that, for each $\epsilon > 0$, we have that

$$\sum_{i \in N} \mathsf{prob}_\sigma(i \text{ wins } |\epsilon\text{-}\mathsf{close}) = 1 \text{ and}$$
$$\sum_{i \in N} \mathsf{prob}_\sigma(|b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon|\epsilon\text{-}\mathsf{close}) = \mathbb{E}_{\mathcal{E},\sigma}[|\{i \text{ s.t. } |b_{i,t} - \wedge \mathbf{b}_{-i,t} < \epsilon\}| \, | \, \epsilon\text{-}\mathsf{close}] \geq 2.$$

Using (12), (13) and (14), we obtain that

$$\begin{aligned} 1 = \limsup_{\epsilon \searrow 0} \sum_{i \in N} \mathsf{prob}_\sigma(i \text{ wins } |\epsilon\text{-}\mathsf{close}) &\geq \frac{1}{2} \limsup_{\epsilon \searrow 0} \sum_{i \in N} \mathsf{prob}_\sigma(|b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon|\epsilon\text{-}\mathsf{close}) \\ &\quad + \eta \mathsf{prob}_\sigma(|b_{j,t} - \wedge \mathbf{b}_{-j,t}| < \epsilon|\epsilon\text{-}\mathsf{close}) \\ &\geq 1 + \eta \limsup_{\epsilon \searrow 0} \mathsf{prob}_\sigma(|b_{j,t} - \wedge \mathbf{b}_{-j,t}| < \epsilon|\epsilon\text{-}\mathsf{close}) > 1, \end{aligned}$$

a contradiction. ■

Proof of Corollary 3. Follows from the fact that Proposition 1 and Corollary 2 imply that Corollary 1 must hold whenever σ is an MPE that is competitively enforced. ■

Sample implications of Corollary 2. We now show that when the sample size is large, Corollary 2 must hold approximately under the sample distribution of bids and characteristics

$\mathbf{b}, \mathbf{x}$.

Data consists of bids and observable characteristics $(\mathbf{b}_t, \mathbf{x}_t)_{t \in \{0,\cdots,T\}}$ for auctions happening at times $t \in \{0, \cdots, T\}$. Let $H = \{h_i\}$ be the set of histories corresponding to data $(\mathbf{b}_t, \mathbf{x}_t)$: i.e., for each data point $(b_{i,t}, x_{i,t})$, history $h_{i,t} \in H$ corresponds to the information that bidder i had at time t, prior to bidding. We denote by $\widehat{\mathsf{prob}}$ the sample joint distribution of bids and characteristics in $(\mathbf{b}_t, \mathbf{x}_t)$.

Definition C.1. *We say that a set of histories H is adapted to the players' information if and only if the event $h_{i,t} \in H$ is measurable with respect to player i's information at time t, prior to bidding.*

A subset H can be thought of as a set of histories that satisfy a certain criteria defined by the analyst. Definition C.1 states that H is adapted if it is possible to check whether $h_{i,t}$ satisfies the criteria needed for inclusion in H using only information available to bidder i at time t, prior to bidding. Consider, for example, the histories in which the bid is above a particular threshold. Because a bidder knows, at the time of bidding, that its bid will be above a given threshold, the set of histories in which a bid is above a given threshold is adapted. Consider next, the histories in which a particular bidder wins. Because a bidder does not know who will win the auction at the time of bidding, the set of histories in which a given bidder wins is not adapted.

As we now show, when Proposition 1 holds and the set of histories H is adapted, sample beliefs satisfy condition (6). This allows us to apply our tests to specific subsets of the data.

Given $\epsilon > 0$ and $x \in X$, we define $B_{x,\epsilon} \equiv \{(i,t) \text{ s.t. } x_{i,t} = x,\ |b_{i,t} - \wedge b_{-i,t}| \leq \epsilon\}$ the subsample of close bids such that the bidders characteristics x_i are equal to x. We denote by $B_\epsilon \equiv \{(i,t) \text{ s.t. } |b_{i,t} - \wedge b_{-i,t}| \leq \epsilon\}$ the sample of close bids. A bidder's sample probability of winning conditional on close bids and type x is denoted by $\widehat{P}_{x,\epsilon}$. Formally, we have,

$$\begin{aligned} \widehat{P}_{x,\epsilon} &\equiv \widehat{\mathsf{prob}}(i \text{ wins} \mid x_i = x,\ |b_i - \wedge b_{-i}| \leq \epsilon) \\ &= \frac{|\{(i,t) \in B_{x,\epsilon} \text{ s.t. } b_{i,t} \prec \wedge b_{-i,t}\}|}{|B_{x,\epsilon}|} \end{aligned} \tag{15}$$

We make the following assumption about data.

Assumption C.1. *There exists $\lambda > 0$ such that for all datasets of interest B, and all $x \in X$,*

$$\frac{\sum_{x' \in X \setminus x} |B_{x',\epsilon}|}{|B_{x,\epsilon}|} \leq \lambda$$

The following result holds:

Proposition C.1 (winning is independent of bidder characteristics)**.** *Suppose H is adapted. For all $\eta > 0$, there exists $\epsilon > 0$ small enough such that with probability approaching 1 as $|B_\epsilon|$ goes to infinity,*

$$\forall x \in X, \quad \left| \widehat{P}_{x,\epsilon} - \frac{1}{2} \right| \leq \eta.$$

Proof. Take $\eta' > 0$ as given. We know from Proposition 1 that for epsilon small enough, for all histories $h_{i,t}$, $\mathsf{prob}(i \text{ wins} \mid h_{i,t} \text{ and } |b_{i,t} - \wedge \mathbf{b}_{-i,t}| < \epsilon) \geq 1/2 - \eta'$.

Fix $x \in X$. We show that with probability approaching 1 as $|B_\epsilon|$ goes to infinity, $\widehat{P}_{x,\epsilon} \geq \frac{1}{2} - 2\eta'$. Observe first that, by Assumption C.1, when $|B_\epsilon|$ grows large, $|B_{x,\epsilon}|$ grows proportionally large:

$$\frac{|B_{x,\epsilon}|}{|B_\epsilon|} = 1 - \frac{\sum_{x' \in X \setminus x} |B_{x',\epsilon}|}{|B_{x,\epsilon}| + \sum_{x' \neq x} |B_{x',\epsilon}|} \geq 1 - \frac{\lambda}{1+\lambda}.$$

We denote by $\{t_1, \cdots, t_n\}$ auctions occurring at times t such that $(i,t) \in B_{x,\epsilon}$, ordered according to the timing of the auction. Since the number N of bidders is finite, n grows large proportionally with $|B_{x,\epsilon}|$. We define $C_k = \{i \in N \text{ s.t. } (i, t_k) \in B_{x,\epsilon}\}$. In equilibrium,

$$H_K \equiv \sum_{k=1}^{K} \sum_{i \in C_k} \mathbf{1}_{b_{i,t_k} \prec \wedge b_{-i,t_k}} - \mathsf{prob}_i(b_{i,t_k} \prec \wedge b_{-i,t_k} | i \in C_k)$$

is a martingale, provided set H is adapted. Indeed note that given the information I_K available at the time of bidding in auction K,

$$\begin{aligned}
\mathbb{E}\left[\sum_{i \in C_K} \mathbf{1}_{b_{i,t_K} \prec \wedge b_{-i,t_K}} \middle| I_K\right] &= \mathbb{E}\left[\sum_{i \in N} \mathbf{1}_{i \in C_K} \mathbf{1}_{b_{i,t_K} \prec \wedge b_{-i,t_K}} \middle| I_K\right] \\
&= \mathbb{E}\left[\mathbb{E}_{C_K}\left[\sum_{i \in N} \mathbf{1}_{i \in C_K} \mathbf{1}_{b_{i,t_K} \prec \wedge b_{-i,t_K}} \middle| I_K\right]\right] \\
&= \mathbb{E}\left[\sum_{i \in N} \mathbf{1}_{i \in C_K} \mathsf{prob}_i(\mathbf{1}_{b_{i,t_K} \prec \wedge b_{-i,t_K}} | i \in C_K) \middle| I_K\right] \\
&= \mathbb{E}\left[\sum_{i \in C_K} \mathsf{prob}_i(\mathbf{1}_{b_{i,t_K} \prec \wedge b_{-i,t_K}} | i \in C_K) \middle| I_K\right].
\end{aligned}$$

Using Proposition 1, this implies that

$$G_K \equiv \sum_{k=1}^{K} \sum_{i \in C_k} \mathbf{1}_{b_{i,t_k} \prec \wedge b_{-i,t_k}} - \frac{1}{2} + \eta'$$

is a submartingale with increments bounded by $|N|$ (the maximum number of bidders in an auction). It follows from the Azuma-Hoeffding Theorem that as n grows large, with probability approaching 1, $G_n \geq -\eta' n$. Since $n \leq |B_{x,\epsilon}|$, this implies that with probability approaching 1,

$$\widehat{P}_{x,\epsilon} \equiv \frac{1}{|B_{x,\epsilon}|} \sum_{k=1}^{n} \sum_{i \in C_k} \mathbf{1}_{b_{i,t_k} \prec \wedge b_{-i,t_k}} \geq \frac{1}{2} - 2\eta'.$$

Since X is finite, with probability approaching 1 as $|B_\epsilon|$ becomes large, we have that for all $x \in X$, $\widehat{P}_{x,\epsilon} \geq \frac{1}{2} - 2\eta'$. In addition, since $\sum_{x' \in X} |B_{x',\epsilon}| \widehat{P}_{x',\epsilon} = |\{(i,t) \in B_\epsilon \text{ s.t. } i \text{ wins }\}|$, it follows that

$$\frac{\sum_{x' \in X} |B_{x',\epsilon}| \widehat{P}_{x',\epsilon}}{\sum_{x' \in X} |B_{x',\epsilon}|} \leq \frac{1}{2}.$$

Hence, with probability approaching 1, we have that

$$|B_{x,\epsilon}| \widehat{P}_{x,\epsilon} \leq \frac{1}{2} |B_{x,\epsilon}| + \sum_{x' \in X \setminus x} |B_{x',\epsilon}| \left(\frac{1}{2} - \widehat{P}_{x',\epsilon} \right)$$

$$\Rightarrow \widehat{P}_{x,\epsilon} \leq \frac{1}{2} + 2\eta' \frac{\sum_{x' \in X \setminus x} |B_{x',\epsilon}|}{|B_{x,\epsilon}|} \leq \frac{1}{2} + 2\eta' \lambda.$$

Hence by selecting η' sufficiently small in the first place, it follows that for any $\eta > 0$, there exists ϵ such that as $|B_\epsilon|$ grows large, $|\widehat{P}_{x,\epsilon} - \frac{1}{2}| \leq \eta$ with probability 1. □

A corollary of Proposition C.1 is that our regression discontinuity design remains valid: conditional on close bids, the sample distribution of covariates is independent of whether the bidder wins or loses the auction.

Corollary C.1 (close winners and losers have similar characteristics)**.** *For all $\eta > 0$, there exists $\epsilon > 0$ small enough such that with probability approaching 1 as $|B_\epsilon|$ goes to infinity,*

$$\forall x \in X, \quad \left| \widehat{\mathsf{prob}}(x_i = x \mid i \text{ wins }, |b_i - \wedge b_{-i}| \leq \epsilon) - \widehat{\mathsf{prob}}(x_i = x \mid i \text{ loses }, |b_i - \wedge b_{-i}| \leq \epsilon) \right| \leq \eta.$$

Proof. Observe that

$$\widehat{\mathsf{prob}}(x_i = x \mid i \text{ wins }, |b_i - \wedge b_{-i}| \le \epsilon) = \widehat{\mathsf{prob}}(x_i = x \mid |b_i - \wedge b_{-i}| \le \epsilon) \frac{\widehat{\mathsf{prob}}(i \text{ wins } \mid x_i = x, |b_i - \wedge b_{-i}| \le \epsilon)}{\widehat{\mathsf{prob}}(i \text{ wins } \mid |b_i - \wedge b_{-i}| \le \epsilon)}$$

$$\widehat{\mathsf{prob}}(x_i = x \mid i \text{ loses }, |b_i - \wedge b_{-i}| \le \epsilon) = \widehat{\mathsf{prob}}(x_i = x \mid |b_i - \wedge b_{-i}| \le \epsilon) \frac{\widehat{\mathsf{prob}}(i \text{ loses } \mid x_i = x, |b_i - \wedge b_{-i}| \le \epsilon)}{\widehat{\mathsf{prob}}(i \text{ loses } \mid |b_i - \wedge b_{-i}| \le \epsilon)}.$$

Therefore,

$$\begin{aligned}
&\left|\widehat{\mathsf{prob}}(x_i = x \mid i \text{ wins }, |b_i - \wedge b_{-i}| \le \epsilon) - \widehat{\mathsf{prob}}(x_i = x \mid i \text{ loses } |b_i - \wedge b_{-i}| \le \epsilon)\right| \\
\le &\left|\frac{\widehat{\mathsf{prob}}(i \text{ wins } \mid x_i = x, |b_i - \wedge b_{-i}| \le \epsilon)}{\widehat{\mathsf{prob}}(i \text{ wins } \mid |b_i - \wedge b_{-i}| \le \epsilon)} - \frac{\widehat{\mathsf{prob}}(i \text{ loses } \mid x_i = x, |b_i - \wedge b_{-i}| \le \epsilon)}{\widehat{\mathsf{prob}}(i \text{ loses } \mid |b_i - \wedge b_{-i}| \le \epsilon)}\right| \\
= &\left|\frac{\widehat{P}_{x,\epsilon}}{\sum_{x' \in X} \frac{|B_{x',\epsilon}|}{|B_\epsilon|} \widehat{P}_{x',\epsilon}} - \frac{1 - \widehat{P}_{x,\epsilon}}{1 - \sum_{x' \in X} \frac{|B_{x',\epsilon}|}{|B_\epsilon|} \widehat{P}_{x',\epsilon}}\right|.
\end{aligned}$$

It follows from Proposition C.1 that for any $\eta' > 0$, there exists ϵ such that with probability 1 as $|B_\epsilon|$ grows large,

$$\frac{\widehat{P}_{x,\epsilon}}{\sum_{x' \in X} \frac{|B_{x',\epsilon}|}{|B_\epsilon|} \widehat{P}_{x',\epsilon}} - \frac{1 - \widehat{P}_{x,\epsilon}}{1 - \sum_{x' \in X} \frac{|B_{x',\epsilon}|}{|B_\epsilon|} \widehat{P}_{x',\epsilon}} \in \left[\frac{1/2 - \eta'}{1/2 + \eta'} - \frac{1/2 + \eta'}{1/2 - \eta'}, \frac{1/2 + \eta'}{1/2 - \eta'} - \frac{1/2 - \eta'}{1/2 + \eta'}\right].$$

By picking η' small enough, this implies that with probability approaching 1,

$$\left|\widehat{\mathsf{prob}}(x_i = x \mid i \text{ wins }, |b_i - \wedge b_{-i}| \le \epsilon) - \widehat{\mathsf{prob}}(x_i = x \mid i \text{ loses }, |b_i - \wedge b_{-i}| \le \epsilon)\right| \le \eta.$$

□

References

Ashton, J. and A. Pressey (2012): "Who manages cartels? the role of sales and marketing managers within international cartels: Evidence from the european union 1990-2009," Tech. rep., Centre for Competition Policy, University of East Anglia, Norwich, UK.

Asker, J. (2010): "A study of the internal organization of a bidding cartel," *The American Economic Review*, 724–762.

Athey, S., J. Levin, and E. Seira (2011): "Comparing open and Sealed Bid Auctions: Evidence from Timber Auctions*," *The Quarterly Journal of Economics*, 126, 207–257.

Bajari, P. and L. Ye (2003): "Deciding between competition and collusion," *Review of Economics and Statistics*, 85, 971–989.

Baker, J. and D. Rubinfeld (1999): "Empirical Methods in Antitrust Litigation: Review and Critique," *American Law and Economics Review*, 1, 386–435.

Baldwin, L. H., R. C. Marshall, and J.-F. Richard (1997): "Bidder collusion at forest service timber sales," *Journal of Political Economy*, 105, 657–699.

Blackwell, D. (1953): "Equivalent comparisons of experiments," *The annals of mathematical statistics*, 265–272.

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

Chassang, S., K. Kawai, J. Nakabayashi, and J. Ortner (2020): "Robust Screens for Non-Competitive Bidding in Procurement Auctions," .

Chassang, S. and J. Ortner (2019): "Collusion in Auctions with Constrained Bids: Theory and Evidence from Public Procurement," *Journal of Political Economy*, 127, 2269–2300.

Clark, R., D. Coviello, J.-F. Gauthier, and A. Shneyerov (2018): "Bid rigging and entry deterrence in public procurement: Evidence from an investigation into collusion and corruption in Quebec," *The Journal of Law, Economics, and Organization*, 34, 301–363.

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

Ericson, R. and A. Pakes (1995): "Markov-perfect industry dynamics: A framework for empirical work," *The Review of economic studies*, 62, 53–82.

Harrington, J. E. (2008): "Detecting cartels," *Handbook of antitrust economics*, 213, 245.

Hendricks, K. and R. Porter (1988): "An Empirical Study of an Auction with Asymmetric Information," *American Economic Review*, 78, 865–883.

Imhof, D., Y. Karagök, and S. Rutz (2018): "Screening for Bid Rigging —- Does It Work?" *Journal of Competition Law & Economics*, 14, 235–261.

Ishii, R. (2009): "Favor exchange in collusion: Empirical study of repeated procurement auctions in Japan," *International Journal of Industrial Organization*, 27, 137–144.

Kaplan, U., M. Vadim, and A. Shneyerov (2017): "Identifying Collusion in English Auctions," .

Kawai, K. and J. Nakabayashi (2018): "Detecting Large-Scale Collusion in Procurement Auctions," *Available at SSRN 2467175*.

Marshall, R. C. and L. M. Marx (2007): "Bidder collusion," *Journal of Economic Theory*, 133, 374–402.

Maskin, E. and J. Tirole (2001): "Markov perfect equilibrium: I. Observable actions," *Journal of Economic Theory*, 100, 191–219.

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

McKelvey, R. D. and T. R. Palfrey (1995): "Quantal response equilibria for normal form games," *Games and economic behavior*, 10, 6–38.

OECD (2013): "Ex officio cartel investigations and the use of screens to detect cartels," .

——— (2018): "Summary of the workshop on cartel screening in the digital era," .

Ohashi, H. (2009): "Effects of transparency in procurement practices on government expenditure: A case study of municipal public works," *Review of Industrial Organization*, 34, 267–285.

Ortner, J., S. Chassang, K. Kawai, and J. Nakabayashi (2020): "Screening Adaptive Cartels," Tech. rep., Boston University.

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

Porter, R. (2005): "Detecting Collusion," *Review of Industrial Organization*, 26, 147–167.

PORTER, R. H. AND J. D. ZONA (1993): "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 101, 518–38.

——— (1999): "Ohio School Milk Markets: An Analysis of Bidding," *RAND Journal of Economics*, 30, 263–288.

SCHURTER, K. (2017): "Identification and inference in first-price auctions with collusion," Tech. rep., working Paper, University of Chicago.

SONNENFELD, J. AND P. R. LAWRENCE (1978): "Why do companies succumb to price fixing," *Harvard Business Review*, 56, 145–157.

THISTLETHWAITE, D. AND D. CAMPBELL (1960): "Regression-discontinuity analysis: An alternative to the ex post facto experiment," *Journal of Educational Psychology*, 51, 309–317.

WACHS, J. AND J. KERTÉSZ (2019): "A Network Approach to Cartel Detection in Public Auction Markets," *Scientific Reports*, 9, 1–10.