# Data Driven Antitrust: Theory and Application to Missing Bids


Sylvain Chassang — New York University

Kei Kawai — U.C. Berkeley

Jun Nakabayashi — Kindai University

Juan Ortner*†

Boston University


March 13, 2020


**Abstract**

We document a novel bidding pattern observed in procurement auctions from Japan: winning bids tend to be isolated, and there is a missing mass of close losing bids. This pattern is suspicious in the following sense: it is inconsistent with competitive behavior under arbitrary information structures. Building on this observation, we develop a class of robust tests of non-competitive behavior exploiting weak equilibrium conditions. We provide an empirical exploration of our tests, and show they can help identify other suspicious patterns in the data.



Keywords: missing bids, collusion, antitrust, procurement.



*Chassang: chassang@nyu.edu, Kawai: kei@berkeley.edu, Nakabayashi: nakabayashi.1@eco.kindai.ac.jp, Ortner: jortner@bu.edu.



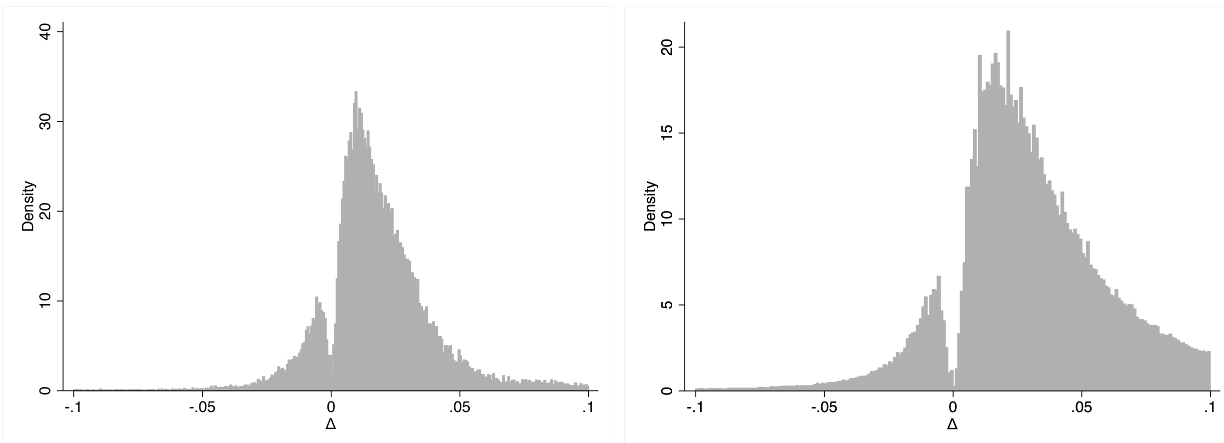
†We are especially indebted to Steve Tadelis for encouragement and detailed feedback. The paper benefited from discussions with Pierpaolo Battigali, Eric Budish, Yeon-Koo Che, Francesco Decarolis, Emir Kamenica, Roger Myerson, Ariel Pakes, Wolfgang Pesendorfer, Andrea Prat, Michael Riordan, Jozsef Sakovics, Larry Samuelson, Andy Skrzypacz, Paulo Somaini, as well as comments from seminar participants at the 2018 ASSA meeting, the 2018 Canadian Economic Theory Conference, Bocconi, the 2017 Berkeley-Sorbonne workshop on Organizational Economics, Boston University, Columbia University, EIEF, the 2018 ESSET meeting at Gerzensee, FGV Rio, Harvard, HKUST, Johns Hopkins, MIT, NYU, the 2017 NYU CRATE conference on theory and econometrics, Penn State, Princeton, PUC Rio, Queen's University, Rochester, the 2018 Stanford SITE conference, Stanford GSB, the 2018 Transparency in Procurement Conference, the Triangle Microeconomics Conference, the University of Arizona, the University of Chicago, the University of Copenhagen, the University of Illinois at Urbana-Champaign, UC Berkeley, the University of Tokyo, the University of Virginia, UCLA, and Yale.

# 1 Introduction

One of the key functions of antitrust authorities is to detect and punish collusion. Although concrete evidence is required for successful prosecution, screening devices that flag suspicious firms help regulators identify collusion, and encourage members of existing cartels to apply for leniency programs. Correspondingly, an active research agenda has sought to build data-driven methods to detect cartels using naturally occurring market data (e.g. Porter, 1983, Porter and Zona, 1993, 1999, Ellison, 1994, Bajari and Ye, 2003, Harrington, 2008). This paper seeks to make progress on this research agenda by developing systematic tests of non-competitive behavior that hold under minimal assumptions on the environment.

We begin by documenting a suspicious bidding pattern observed in procurement auctions in Japan: the density of the bid distribution just above the winning bid is very low; there is a missing mass of close losing bids. These missing bids are related to bidding patterns observed among collusive firms in Hungary (Tóth et al., 2014) and Switzerland (Imhof et al., 2018). We establish that this pattern is inconsistent with competitive behavior under a general class of asymmetric information structures, and arbitrary unobserved heterogeneity. Indeed, when winning bids are isolated, bidders can profitably deviate by increasing their bids. Expanding on this observation, we propose general tests of non-competitive behavior, and use them to identify other suspicious patterns in our data.

Our data come from two sets of public works procurement auctions in Japan. The first dataset contains information on roughly 7,000 city-level auctions held between 2004 and 2018 by 14 different municipalities in the Ibaraki prefecture and the Tohoku region of Japan. The second dataset, analyzed by Kawai and Nakabayashi (2018), contains data on approximately 78,000 national-level auctions held between 2001 and 2006 by the Ministry of Land, Infrastructure and Transportation. We are interested in the distribution of bidders' margins of victory (or defeat). For every (bidder, auction) pair, we compute $\Delta \equiv \frac{\text{own bid} - \min(\text{other bids})}{\text{reserve}}$, the difference between the bidder's own bid and the most competitive bid among this bidder's opponents, divided by the reserve price. When $\Delta < 0$, the bidder won the auction. When $\Delta > 0$ the bidder lost. The finding motivating this paper is summarized by Figure 1, which plots the distribution of bid-differences $\Delta$ for city-level, and national-level auctions. In both cases, there is a striking missing mass around $\Delta = 0$. Our main goal is to clarify the sense in which this gap — and other patterns that could be found in the data — are suspicious.




(a) city auctions (b) national auctions

Figure 1: Distribution of bid-differences $\Delta$ over (bidder, auction) pairs.

We analyze our data within a fairly general framework. A group of firms repeatedly participates in first-price procurement auctions. Players can observe arbitrary signals about one another, and bidders' costs and types can be arbitrarily correlated within and across periods. Importantly, intertemporal linkages between actions and payoffs are ruled out. Behavior is called *competitive*, if it is stage-game optimal under the players' information.

Our first set of results establishes that in its more *extreme forms*, the pattern of missing bids illustrated in Figure 1 is not consistent with competitive behavior under any information structure. We exploit the fact that in any competitive equilibrium, firms must not find it profitable in expectation to increase their bids. This incentive constraint implies that with high probability the elasticity of firms' *sample residual demand* (i.e., the empirical probability of winning an auction at any given bid) must be bounded above by -1. This condition is not satisfied in our data: because winning bids are isolated, the elasticity of the sample residual demand is close to zero. In addition, we are able to derive bounds on the minimum number of histories at which non-competitive bidding must happen. These results show how to exploit the fact that firms maximizing short-term profits must price in the elastic part of their residual demand curve in settings with arbitrary incomplete information.

Our second set of results generalizes this test to derive sharper bounds on the extent of non-competitive bidding in our data. These tests have the property that they are passed with probability one whenever firms are behaving competitively under *some* information structure. We refer to these tests as *safe* tests for two reasons. First, up to transparent economic constraints imposed on the underlying environment, they yield no false positives.[1] Second, as we show in a companion paper (Ortner et al., 2019), antitrust policy based on safe tests cannot be exploited by cartels to enhance collusion. This addresses a concern over the equilibrium effects of data driven antitrust raised by Cyrenne (1999) and Harrington (2004). For these reasons safe tests strike us as a healthy starting point for data-driven regulation.

Our third set of results takes our tests to the data. We delineate how different moment conditions (i.e. different deviations) uncover different non-competitive patterns. While missing bids suggest that a small increase in bids is attractive, we show that a small drop in bids (on the order of 2%) may also be attractive to bidders: it yields large increases in demand. In addition, we show that upward and downward deviations can be more informative together than separately. Finally, we show that the outcomes of our tests are consistent with other proxy evidence for competitiveness and collusion. Histories such that bids are close to the reserve price are more likely to fail our tests than histories where bids are low relative to the reserve price. Histories before an industry is investigated for collusion are more likely to fail our tests than histories after it is investigated for collusion. Altogether this suggests that,

[1] Imhof et al. (2018) emphasizes the importance of keeping false positives to a minimum due to the fact that formal investigations triggered by data-analysis can have significant adverse consequences for affected parties.

although our tests are conservative, they still have bite in practice.

Our paper relates primarily to the literature on cartel detection.[2] Porter and Zona (1993, 1999) show that suspected cartel and non-cartel members bid in statistically different ways. Bajari and Ye (2003) design a test of collusion based on excess correlation across bids. Porter (1983) and Ellison (1994) exploit dynamic patterns of play predicted by the theory of repeated games (Green and Porter, 1984, Rotemberg and Saloner, 1986) to detect collusion. Conley and Decarolis (2016) propose a test of collusion in average-price auctions exploiting cartel members' incentives to coordinate bids. Chassang and Ortner (2019) propose a test of collusion based on changes in behavior around changes in the auction design. Kawai and Nakabayashi (2018) analyze auctions with re-bidding, and exploit correlation in bids across stages to detect collusion.[3]

Our tests are related to revealed preference tests seeking to quantify violations of choice theoretic axioms.[4] Afriat (1967), Varian (1990), and Echenique et al. (2011) propose tests to quantify the extent to which a given consumption dataset violates GARP. More closely related, Carvajal et al. (2013) propose a revealed preference test of the Cournot model. Sullivan (1985) and Ashenfelter and Sullivan (1987) test whether firms in a given industry behave as a perfect cartel.

Our paper makes an indirect contribution to the literature on the internal organization of cartels. Asker (2010) studies stamp auctions, and analyses the effect of a particular collusive scheme on non-cartel bidders and sellers. Pesendorfer (2000) studies the bidding patterns for school milk contracts and compares the collusive schemes used by strong cartels and weak cartels (i.e., cartels that used transfers and cartels that did not). Clark and Houde (2013) document the collusive strategies used by the retail gasoline cartel in Quebec. Clark et al. (2018) study the effect of an investigation on firms' bidding behavior. We add to this literature by documenting a puzzling bidding pattern that is poorly accounted for by

[2]See Harrington (2008) for a recent survey.
[3]Schurter (2017) proposes a test of collusion based on exogenous shifts in the number of bidders.
[4]See Chambers and Echenique (2016) for a recent review of the literature on revealed preferences.

existing theories. We establish that this bidding pattern is non-competitive, and propose some potential explanations.

The paper is structured as follows. Section 2 describes our data and documents missing bids. Section 3 introduces our theoretical framework. Section 4 shows that extreme forms of missing bids cannot be rationalized under any competitive model. Section 5 generalizes this analysis, and develops a test that systematically exploits weak optimality conditions implied by equilibrium. Section 6 delineates the mechanics of our tests in real data, and shows that their implications are consistent with other indicators of collusion. Section 7 concludes with an open ended discussion of why missing bids may arise in the context of collusion. Appendix A generalizes our analysis to auctions with a secret reserve price and re-bidding. Appendix B presents several theoretical extensions, including a characterization of tests exploiting all implications from Bayes Nash equilibrium. Appendix C describes our computational strategy. Proofs are collected in Appendix D unless mentioned otherwise.

# 2 Motivating Facts

Our first dataset consists of roughly 7,100 auctions held between 2004 and 2018 by municipalities located in the Tohoku region and the Ibaraki prefecture of Japan.[5] The median winning bid is about 64,000 USD, and the median number of participants is 7.

For any given firm $i$ participating in auction $a$ with reserve price $r$, we denote by $b_{i,a}$ the bid of firm $i$ in auction $a$, and by $\mathbf{b}_{-i,a}$ the profile of bids by bidders other than $i$. We investigate the distribution of

$$\Delta_{i,a} = \frac{b_{i,a} - \wedge \mathbf{b}_{-i,a}}{r}$$

[5] Our city-level data comes from two datasets. The first dataset contains auctions held by municipalities in the Tohoku region in Japan. For the current analysis, we restrict attention to municipalities using a sealed-bid first-price auction with a public reserve price. The second dataset, studied in Chassang and Ortner (2019), contains auctions held by municipalities in the prefecture of Ibaraki. For the current analysis, we use data from the city of Tsuchiura during 2007-2009, when the city was using sealed-bid first-price auctions with a public reserve price.

aggregated over firms $i$, and auctions $a$, where $\wedge$ denotes the minimum operator. The value $\Delta_{i,a}$ represents the margin by which bidder $i$ wins or loses auction $a$. If $\Delta_{i,a} < 0$ the bidder won, if $\Delta_{i,a} > 0$ she lost. Figure 1(a) plots the distribution of bid differences $\Delta$ aggregating over all firms and auctions in our sample.[6] The mass of missing bids around a difference of 0 is clearly noticeable.[7]

Our second dataset, described in Kawai and Nakabayashi (2018), consists of roughly 78,000 auctions held between 2001 and 2006 by the Ministry of Land, Infrastructure and Transportation in Japan (the Ministry). The auctions are sealed-bid first-price auctions with a *secret* reserve price, and re-bidding in case there is no successful winner. The auctions involve construction projects. The median winning bid is about 1 million USD, and the median number of participants is 10. The bids of all participants are publicly revealed after each auction. Figure 1(b) illustrates the distribution of bid-differences $\Delta$ for national auctions. The missing mass of bids around $\Delta = 0$ is very stark. In addition, this pattern can be traced to individual firms. Figure 2 reports the distribution of bid differences $\Delta$ for a single large firm frequently active in our sample of auctions.

The primary objective of this paper is to assess the extent to which bidding patterns illustrated in Figures 1 and 2 are inconsistent with competitive behavior under any information structure. We note that non competitive behavior need not be collusive (i.e. yield excessively high prices). Still, as we show in our companion paper (Ortner et al., 2019), tests of non-competitive behavior with no false positives provide a safe foundation for antitrust policy: they do not damage competitive industries; they do not create new collusive equilibria; finally, they can reduce profits from operating a cartel. In addition, we note that missing bids are in fact correlated with plausible indicators of collusion.

[6]Note that the distribution of normalized bid-differences is skewed to the right since the most competitive alternative bid is a minimum over other bidders' bids.

[7]Imhof et al. (2018) document a similar bidding pattern in procurement auctions in Switzerland: bidding patterns by several cartels uncovered by the Swiss competition authority presented large differences between the winning bid and the second lowest bid in auctions. See also Tóth et al. (2014).

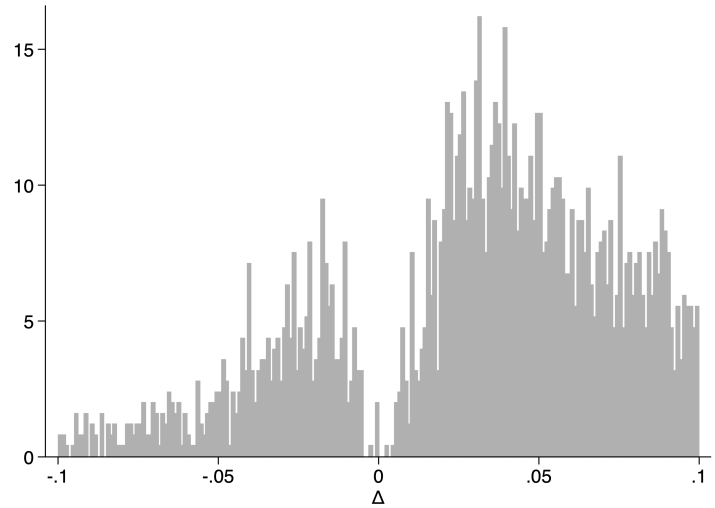


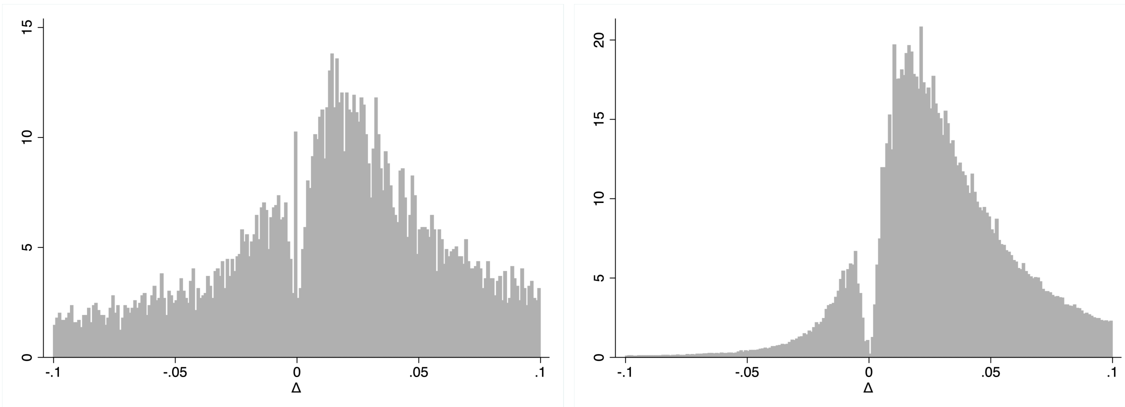

Figure 2: Distribution of normalized bid-differences Δ – single large firm.

**Correlation with indicators of collusion.** Figure 3 breaks down the auctions in Figure 1(b) by bid level: the figure plots the distribution of $\Delta_{i,a} = \frac{b_{i,a} - \wedge \mathbf{b}_{-i,a}}{r}$ for normalized bids $\frac{b_{i,a}}{r}$ below .9 and above .8. The mass of missing bids in Figure 2 is considerably reduced when we look at bids that are low compared to the reserve price.

Figure 4 plots the distribution of $\Delta_{i,a}$ for participants of auctions held by the Ministry that were implicated by the Japanese Fair Trade Commission (JFTC). The JFTC implicated four bidding rings participating in the auctions in our data: (i) firms installing electric traffic signs (Electric); (ii) builders of bridge upper structures (Bridge); (iii) pre-stressed concrete providers (PSC); and (iv) floodgate builders (Flood).[8] The left panels in Figure 4 plot the distribution of Δ for auctions that were run before the JFTC started its investigation, and the right panels plot the distribution in the after period. In all cases except PSC (iii), the pattern of missing bids disappears after the JFTC launched its investigation. Interestingly, firms in case (iii) initially denied the charges against them (unlike firms in the other three cases), and seem to have continued colluding for some time (see Kawai and Nakabayashi

[8]See JFTC Recommendation and Ruling #5-8 (2005) for case (i); JFTC Recommendation and Ruling #12 (2005) for case (ii); JFTC Recommendation #27-28 (2004) and Ruling #26-27 (2010) for case (iii); and JFTC Cease and Desist Order #2-5 (2007) for case (iv).

(2018) for a more detailed account of these collusion cases).




(a) Below 90%     (b) Above 90%

Figure 3: Distribution of bid-difference $\Delta$ for bids below and above 90% of the reserve price – national data.

**What does not explain this pattern.** Although explaining missing bids is not the main goal of this paper, it is useful to clarify what does not explain this pattern.[9] Specifically, we argue that missing bids are not explained by either the granularity of bids, or ex post renegotiation.

Figure 5 plots the distribution of differences $\Delta^2$ between bids after the lowest bid is excluded. Formally, letting $NW(a)$ denote the set of non-winning bidders in auction $a$, we have

$$\forall i \in NW(a), \quad \Delta^2_{i,a} = \frac{b_{i,a} - \min_{j \in NW(a)} b_{j,a}}{r}.$$

The pattern of missing bids essentially disappears in the case of city and national auctions.[10]

What Figures 3, 4 and 5 establish is that the pattern of missing bids of Figure 1 is not a mechanical consequence of the granularity of bids. If this was the case, we should see similar

[9]We propose different plausible explanations for missing bids in Section 7.

[10]We note that Figure 5 is still potentially suspicious – for instance there appears to be a point mass of tied non-winning bids.

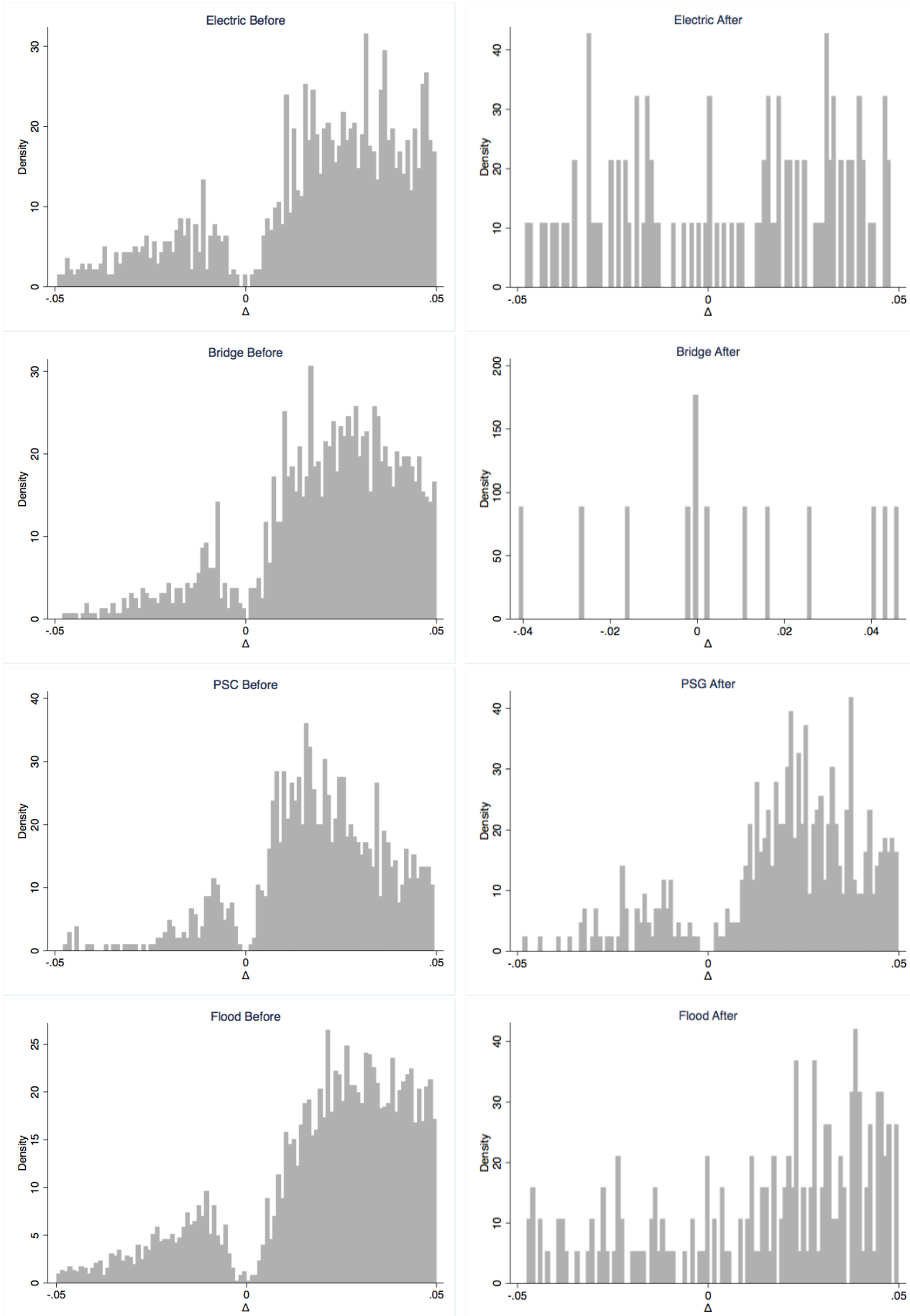
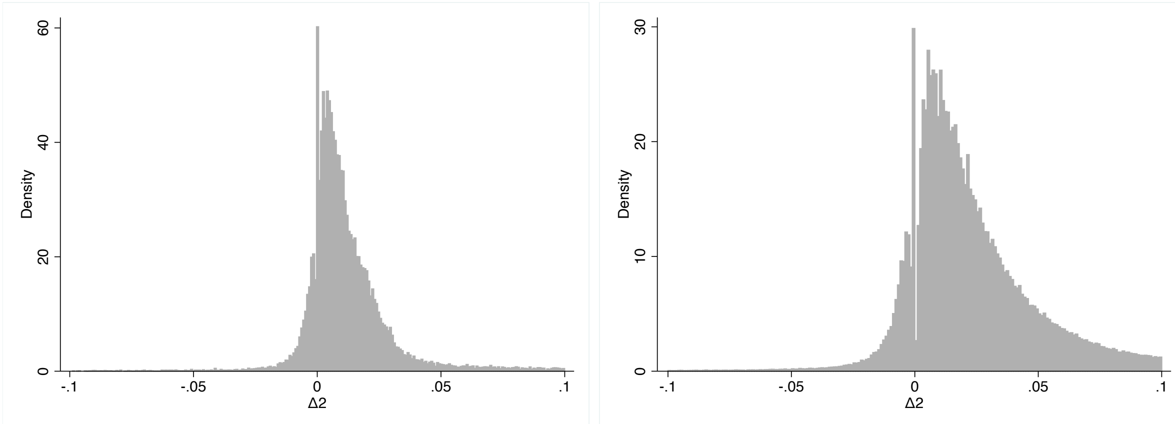


Figure 4: Distribution of bid-difference $\Delta$ – cartel cases in national data, before and after JFTC investigation.

patterns across all bid levels, before and after the JFTC investigations, and when comparing




(a) city auctions (b) national auctions

Figure 5: Distribution of bid-difference $\Delta^2$ excluding winning bids

the second and third lowest bids.

Renegotiation could potentially account for missing bids by making apparent incentive compatibility issues irrelevant. Indeed, some winning firms may seemingly leave money on the table, only to reclaim it through renegotiation ex post. Our national-level data contain information on renegotiated prices, and allow us to rule out this explanation. First, Figure 6 shows that the missing bids pattern persists even if we focus on auctions whose prices were not renegotiated up. Second, in the auctions we study, contracts signed between the awarder and the awardee include renegotiation provisions that greatly reduce firms' incentives to bid aggressively with the hope of renegotiating to a higher price later on. Specifically, the contract stipulates that renegotiated prices should be anchored to the initial bid: if the project is estimated to cost \$$y$ more than initially thought, the renegotiated price is increased by $\frac{\text{initial bid}}{\text{reserve price}} \times y$. This implies that excessively competitive unprofitable bids are likely to remain unprofitable after renegotiation.

Our main objectives in this paper are: (*i*) to formalize why the missing mass of bids around zero is suspicious; (*ii*) to propose a class of systematic tests able to detect failures of competition beyond the specific pattern of missing bids highlighted so far; and (*iii*) to illustrate the mechanics of these tests on real procurement data.

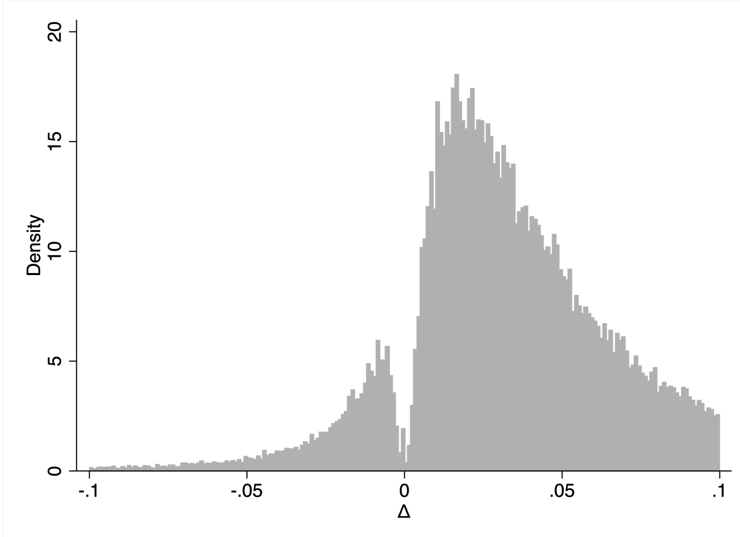



Figure 6: Distribution of bid-difference $\Delta$ – national-level data for auctions whose price was not renegotiated upwards.

# 3 Framework

## 3.1 The Stage Game

We consider a dynamic setting in which, at each period $t \in \mathbb{N}$, a buyer needs to procure a single project. In the main body of the paper, we assume that the auction format is a sealed-bid first-price auction with *public* reserve price $r$, which we normalize to $r = 1$. Appendix A extends the analysis to multistage auctions with *secret* reserve prices and re-bidding (as in the national data). It requires taking into account information revelation across multiple stages of bidding.

In each period $t$, a state $\theta_t \in \Theta$ captures all relevant past information about the environment. State $\theta_t$ may be unknown to the bidders at the time of bidding, but is revealed to bidders at the end of period $t$. We assume that $\theta_t$ is a Markov chain (i.e. given any event $E$ anterior to time $t$, $\theta_t|\{\theta_{t-1}, E\} \sim \theta_t|\{\theta_{t-1}\}$), but do not assume that there are finitely many states, that the chain is irreducible, ergodic, or that $\theta_t$ is observable to the econometrician. The key assumption here is that at the end of each period $t$, bidders observe a sufficient statistic $\theta_t$ of future environments.

In each period $t \in \mathbb{N}$, a set $\widehat{N}_t \subset N$ of bidders is able to participate in the auction, where $N$ is the overall set of bidders. We think of this set of participating firms as those eligible to produce in the current period.[11] The distribution of the set of eligible bidders $\widehat{N}_t$ can vary over time, but depends only on state $\theta_{t-1}$. Participants discount future payoffs using common discount factor $\delta < 1$.

Realized costs of production for eligible bidders $i \in \widehat{N}_t$ are denoted by $\mathbf{c}_t = (c_{i,t})_{i \in \widehat{N}_t}$. Each bidder $i \in \widehat{N}_t$ submits a bid $b_{i,t}$. Profiles of bids are denoted by $\mathbf{b}_t = (b_{i,t})_{i \in \widehat{N}_t}$. We let $\mathbf{b}_{-i,t} \equiv (b_{j,t})_{j \neq i}$ denote bids from firms other than firm $i$, and define $\wedge \mathbf{b}_{-i,t} \equiv \min_{j \neq i} b_{j,t}$ to be the lowest bid among $i$'s competitors at time $t$. The procurement contract is allocated to the bidder submitting the lowest bid at a price equal to her bid. Ties are broken randomly.

**Costs.** The profile of costs $\mathbf{c}_t = (c_{i,t})_{i \in \widehat{N}_t}$ may exhibit correlation across players and over time, but its distribution depends only on state $\theta_t$. All costs are assumed to be positive.

In period $t$, bidder $i \in \widehat{N}_t$ obtains profits

$$\pi_{i,t} = x_{i,t} \times (b_{i,t} - c_{i,t}),$$

where $x_{i,t} \in [0,1]$ is the probability with which $i$ wins the auction at time $t$. Note that costs include both the direct costs of production and the opportunity cost of backlog.

**Information.** In each period $t$, bidder $i$ gets a signal $z_{i,t}$ prior to bidding. The distribution of the profile of signals $(z_{i,t})_{i \in \widehat{N}_t}$ depends only on $(\theta_t, (c_{i,t})_{i \in \widehat{N}_t})$. We stress that signals $(z_{i,t})_{i \in \widehat{N}_t}$ are arbitrary, and may reveal information about current state $\theta_t$, or the realized costs $c_{j,t}$ of other players. This allows our model to nest many informational environments,

[11] For simplicity, we take the set of participating bidders as exogenous. In practice, the set of participants may well be endogenous (see the Online Appendix of Chassang and Ortner (2019) for a treatment of endogenous participation by cartel members). This does not affect our analysis: bids would still have to satisfy the optimality conditions we rely on for inference.

including asymmetric information private value auctions, common value auctions, as well as complete information. Bids $\mathbf{b}_t$ are publicly observable at the end of the auction.[12]

### 3.2 Solution Concepts

A public history $h_t^0$ in period $t$ takes the form $h_t^0 = (\theta_{s-1}, \mathbf{b}_{s-1})_{s \leq t}$. We let $\mathcal{H}^0$ denote the set of all public histories. Our solution concept is perfect public Bayesian equilibrium (Athey and Bagwell, 2008). Because state $\theta_t$ is revealed at the end of each period, past play conveys no information about the private types of other players. As a result we do not need to specify out-of-equilibrium beliefs. A perfect public Bayesian equilibrium consists only of a strategy profile $\sigma = (\sigma_i)_{i \in N}$, such that for all $i \in N$,

$$\sigma_i : h_t^0, z_{i,t} \mapsto b_{i,t}.$$

We emphasize the class of competitive equilibria, which simply corresponds to the class of Markov perfect equilibria (Maskin and Tirole, 2001). In a competitive equilibrium, players condition their play only on payoff relevant parameters.

**Definition 1** (competitive strategy)**.** *We say that $\sigma$ is Markov perfect if and only if $\forall i \in N$ and $\forall h_t^0 \in \mathcal{H}^0$, $\sigma_i(h_t^0, z_{i,t})$ depends only on $(\theta_{t-1}, z_{i,t})$.*

*We say that a strategy profile $\sigma$ is a competitive equilibrium if it is a perfect public Bayesian equilibrium in Markov perfect strategies.*

We note that in a competitive equilibrium, firms must be playing a stage-game Nash equilibrium at every period; that is, firms must play a static best-reply to the actions of their opponents.

[12]Bids are publicly reported in the auctions we study. The assumption that bidders (rather than just the econometrician) observes bids could be relaxed, however, echoing Fershtman and Pakes (2012), this makes it more plausible that real bidding behaviors satisfies the weak optimality conditions used in our identification strategy.

**Competitive histories.** Competitiveness of equilibrium is a fairly coarse notion. Our datasets involve many firms, interacting over an extensive timeframe. Realistically, an equilibrium may include periods in which (a subset of) firms collude and periods in which firms compete. This leads us to define competitiveness at the history level.

**Definition 2** (competitive histories)**.** *Fix a common knowledge profile of play $\sigma$ and a history $h_{i,t} = (h_t^0, z_{i,t})$ of player $i$. We say that player $i$ is competitive at history $h_{i,t}$ if play at $h_{i,t}$ is stage-game optimal for firm $i$ given the behavior of other firms $\sigma_{-i}$.*

# 4 Missing Bids are Inconsistent with Competition

In this section, we show how to use the fact that an oligopolistic competitor must price in the elastic part of its residual demand curve to make inferences about competitiveness under arbitrary incomplete information. First, we identify moments of the subjective residual demand that can be estimated from data, even though the subjective demand of bidders can vary and we do not make any ergodicity assumption about the underlying state $\theta_t$.

## 4.1 Counterfactual demand

Fix a perfect public Bayesian equilibrium $\sigma$. For all histories $h_{i,t}$ and all bids $b' \in [0,1]$, player $i$'s *residual demand* at $h_{i,t}$ is

$$D_i(b'|h_{i,t}) \equiv \mathsf{prob}_\sigma(\wedge \mathbf{b}_{-i,t} > b'|h_{i,t}).$$

In words, bidder $i$'s residual demand at history $h_{i,t}$ represents the probability with which bidder $i$ wins the auction at period $t$, for each possible bid $b'$ she may place.

Take as given a finite set of histories $H$, and a scalar $\rho \in (-1, \infty)$. We define the average residual demand for histories in $H$ as

$$\overline{D}(\rho|H) \equiv \frac{1}{|H|} \sum_{h_{i,t} \in H} D_i((1+\rho)b_{i,t}|h_{i,t}).$$

Its sample equivalent is

$$\widehat{D}(\rho|H) \equiv \frac{1}{|H|} \sum_{h_{i,t} \in H} \mathbf{1}_{\wedge \mathbf{b}_{-i,t} > (1+\rho)b_{i,t}}.$$

**Definition 3.** *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.*

In words, we say that an event is adapted, if it depends only on the information available to individual bidders at the time of bidding. For instance, the entire set of histories for a specific industry or location is adapted – a bidder knows its industry, and its location. In contrast, the set of histories in which a specific bidder wins is not.[13] The ability to legitimately vary the conditioning set $H$ lets us explore the competitiveness of auctions in particular settings of interest, across industries, or time periods. Let $N_{\max}$ denote an upper bound on the number of participants in any auction.[14]

**Lemma 1.** *Consider an adapted set of histories $H$. Under any perfect public Bayesian equilibrium $\sigma$, for any $\nu > 0$,*

$$\textit{prob}(|\widehat{D}(\rho|H) - \overline{D}(\rho|H)| \leq \nu) \geq 1 - 2\exp(-\nu^2|H|/2N_{\max}).$$

*In particular, with probability 1, $\widehat{D}(\rho|H) - \overline{D}(\rho|H) \to 0$ as $|H| \to \infty$.*

In words, in equilibrium, the sample residual demand conditional on an adapted set of histories converges to the true average subjective demand. Indeed, since $H$ is adapted to

[13]Note that the set of competitive histories itself is adapted (a bidder knows whether its bid is subjectively competitive), though unobserved by the econometrician.

[14]It is sufficient for $N_{\max}$ to be a bound on the number of participants with histories in $H$ in each auction.

players' information, by the Law of Iterated Expectations we have that

$$\mathbb{E}_\sigma[\mathsf{prob}_\sigma(\wedge \mathbf{b}_{-i,s} > (1+\rho)b_{i,s}|h_{i,s}) - \mathbf{1}_{\wedge \mathbf{b}_{-i,s}>(1+\rho)b_{i,s}}|h_{i,t} \in H] = 0$$

for all histories $h_{i,t}, h_{i,s} \in H$ with $t < s$. Hence, difference $|H|\left(\widehat{D}(\rho|H) - \overline{D}(\rho|H)\right)$ evolves like a Martingale as the set of histories expands. The Azuma-Hoeffding Inequality yields Lemma 1.

This result is a consequence of the equilibrium assumption, but since bidders observe past bids, it is implied by optimality conditions weaker than equilibrium. For instance, it would hold if participants used data-driven predictors of demand satisfying no-regret (see Hart and Mas-Colell, 2000).

We highlight that Lemma 1 relies on conservative non-asymptotic concentration bounds. In practice, one may be willing to make additional assumptions on the data generating process leading to tighter estimates. Our approach extends to any probabilistic moment conditions of the form

$$\mathsf{prob}\left(\left(\widehat{D}(\rho|H) - \overline{D}(\rho|H)\right)_{\rho \in \mathcal{M}} \in S\right) \geq 1 - \epsilon$$

where $\mathcal{M}$ is a finite set of deviations $\rho \in (-1, +\infty)$, and $S \subset \mathbb{R}^{|\mathcal{M}|}$ is a confidence set for demand estimation errors $\widehat{D}(\rho|H) - \overline{D}(\rho|H)$ at different values of $\rho \in \mathcal{M}$.[15]

## 4.2 A Test of Non-Competitive Behavior

The pattern of bids illustrated in Figures 1 and 2 is striking. Our first main result shows that its more extreme forms are inconsistent with competitive behavior.

[15]For instance, if one is willing to impose that $(\theta_t)_{t\in\mathbb{N}}$ is stationary and ergodic, and that $\sigma$ is a competitive equilibrium, then $\widehat{D}(\rho|H) - \overline{D}(\rho|H)$ is asymptotically Gaussian when set $H$ grows large. One can then derive confidence sets by estimating an asymptotic covariance matrix. Alternatively one may also derive confidence sets using bootstrap. This is the approach we use in our empirical analysis.

**Proposition 1.** *Let $\sigma$ be a competitive equilibrium. Then,*

$$\forall h_i, \quad \frac{\partial \log D_i(b'|h_i)}{\partial \log b'}\Big|_{b'=\sigma_i(h_i)^+} \leq -1, \tag{1}$$

$$\forall H, \quad \frac{\partial \log \overline{D}(\rho|H)}{\partial \rho}\Big|_{\rho=0^+} \leq -1. \tag{2}$$

In words, under any competitive equilibrium, the elasticity of a bidder's *subjective* residual demand must be less than -1 at every history, and the inequality aggregates to sets of histories. Proposition 1 extends to first-price auctions with *secret* reserve prices (see Appendix A).

**Proof.** Consider a competitive equilibrium $\sigma$. Let

$$V(h_{i,t}) \equiv \mathbb{E}_\sigma \left( \sum_{s \geq t} \delta^{s-t}(b_{i,s} - c_{i,s})\mathbf{1}_{b_{i,s} < \wedge \mathbf{b}_{-i,s}} \Big| h_{i,t} \right)$$

denote player $i$'s discounted expected payoff at history $h_{i,t}$. Let $b$ denote the bid that bidder $i$ places at history $h_{i,t}$. Since $b$ is an equilibrium bid, it must be that for all bids $b' > b$,

$$\mathbb{E}_\sigma \left[ (b - c_{i,t})\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b} + \delta V(h_{i,t+1}) \Big| h_{i,t}, b_{i,t} = b \right] \geq \mathbb{E}_\sigma \left[ (b' - c_{i,t})\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b'} + \delta V(h_{i,t+1}) \Big| h_{i,t}, b_{i,t} = b' \right]$$

Since $\sigma$ is competitive, $\mathbb{E}_\sigma[V(h_{i,t+1})|h_{i,t}, b_{i,t} = b] = \mathbb{E}_\sigma[V(h_{i,t+1})|h_{i,t}, b_{i,t} = b']$. Hence,

$$\begin{aligned} bD_i(b|h_{i,t}) - b'D_i(b'|h_{i,t}) &= \mathbb{E}_\sigma \left[ b\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b} - b'\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b'} \Big| h_{i,t} \right] \\ &\geq \mathbb{E}_\sigma \left[ c_{i,t}(\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b} - \mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b'}) \Big| h_{i,t} \right] \geq 0, \end{aligned} \tag{3}$$

where the last inequality uses the assumption that $c_{i,t} \geq 0$. This implies that for all $b' > b$,

$$bD_i(b|h_{i,t}) \geq b'D_i(b'|h_{i,t}) \iff \log b + \log D_i(b|h_{i,t}) \geq \log b' + \log D_i(b'|h_{i,t})$$

$$\iff \frac{\log D_i(b'|h_i) - \log D_i(b|h_i)}{\log b' - \log b} \leq -1.$$

Inequality (1) follows from taking the limit as $b' \to b$. Inequality (2) follows from a similar argument: for all $h_{i,t}$, $b_{i,t}$, we have that

$$b_{i,t}D_i(b|h_{i,t}) \geq (1+\rho)b_{i,t}D_i((1+\rho)b|h_{i,t}) \iff D_i(b|h_{i,t}) \geq (1+\rho)D_i((1+\rho)b|h_{i,t})$$

Averaging over histories $h_{i,t} \in H$, this implies that

$$\overline{D}(0|H) \geq (1+\rho)\overline{D}(\rho|H) \iff \frac{\log \overline{D}(\rho|H) - \log \overline{D}(0|H)}{\log(1+\rho)} \leq -1.$$

Taking $\rho$ to 0 yields inequality (2). ■

Proposition 1 extends the standard result that an oligopolistic competitor must price in the elastic part of her residual demand curve to settings with arbitrary incomplete information. It can be tested by replacing the true average residual demand with its sample average. Extreme forms of missing bids contradict Proposition 1: when the density of $\Delta_i$ at 0 is close to 0, the elasticity of demand is approximately zero.

As the proof highlights, this result exploits the fact that in procurement auctions, zero is a natural lower bound for costs. In contrast, for auctions where bidders are purchasing a good with positive value, there is no corresponding natural upper bound to valuations. One would need to impose an upper bound on values to establish similar results.

Proposition 1 yields a simple test of whether an adapted dataset $H$ can be generated by a competitive equilibrium or not. Importantly, the test is valid under general incomplete information structures, and arbitrary non-stationarity in the underlying environment. This strengthens existing approaches relying on specific information structures such as independent private values (see for instance Bajari and Ye, 2003).

We now refine this test to obtain bounds on the minimum share of non-competitive histories needed to rationalize the data. We begin with a loose bound and then propose a more sophisticated program resulting in tighter bounds.

### 4.3 Estimating the share of competitive histories

It follows from Proposition 1 that missing bids cannot be explained in a model of competitive bidding. We now establish that competitive behavior must fail at a significant number of histories in order to explain isolated winning bids. This implies that bidders have frequent opportunities to learn that their bids are not optimal.

Fix a perfect public Bayesian equilibrium $\sigma$ and a finite set of histories $H$. Let $H^{\mathsf{comp}} \subset H$ be the set of competitive histories in $H$. Define $s_{\mathsf{comp}} \equiv \frac{|H^{\mathsf{comp}}|}{|H|}$ to be the fraction of competitive histories in $H$. For all histories $h_{i,t} = (h_t^0, z_{i,t})$ and all bids $b' \geq 0$, player $i$'s *counterfactual revenue* at $h_{i,t}$ is

$$R_i(b'|h_{i,t}) \equiv b' D_i(b'|h_{i,t}).$$

For any finite set of histories $H$ and scalar $\rho \in (-1, \infty)$, let

$$\overline{R}(\rho|H) \equiv \frac{1}{|H|} \sum_{h_{i,t} \in H} (1+\rho) b_{i,t} D_i((1+\rho) b_{i,t} | h_{i,t})$$

denote the average counterfactual revenue for histories in $H$. An extension of Lemma 1 shows that the sample counterfactual revenue $\widehat{R}(\rho|H) \equiv \frac{1}{|H|} \sum_{h_{i,t} \in H} (1+\rho) b_{i,t} \mathbf{1}_{\wedge \mathbf{b}_{-i,t} > (1+\rho) b_{i,t}}$ is a consistent estimator of $\overline{R}(\rho|H)$, whenever set $H$ is adapted.

Our next result builds on Proposition 1 to derive a bound on $s_{\mathsf{comp}}$.

**Proposition 2.** *The share $s_{\mathit{comp}}$ of competitive histories is such that*

$$s_{\mathit{comp}} \leq 1 - \sup_{\rho > 0} \frac{\overline{R}(\rho|H) - \overline{R}(0|H)}{\rho}.$$

**Proof.** Let $H^{\neg\mathsf{comp}} = H \backslash H^{\mathsf{comp}}$ be the set of non-competitive histories in $H$. For any $\rho > 0$,

$$\frac{1}{\rho}[\overline{R}(\rho|H) - \overline{R}(0|H)] = s_{\mathsf{comp}} \frac{1}{\rho}\left[\overline{R}(\rho|H^{\mathsf{comp}}) - \overline{R}(0|H^{\mathsf{comp}})\right] + (1 - s_{\mathsf{comp}})\frac{1}{\rho}\left[\overline{R}(\rho|H^{\neg\mathsf{comp}}) - \overline{R}(0|H^{\neg\mathsf{comp}})\right].$$

Summing inequality (3) over histories implies that $\overline{R}(\rho|H^{\mathsf{comp}}) - \overline{R}(0|H^{\mathsf{comp}}) \le 0$. In addition, for all histories $h_{i,t}$, we have that $(1+\rho)b_{i,t}D_i((1+\rho)b_{i,t}|h_{i,t}) \le (1+\rho)b_{i,t}D_i(b_{i,t}|h_{i,t})$. Summing up over histories, this implies that

$$\overline{R}(\rho|H^{\neg\mathsf{comp}}) \le (1+\rho)\overline{R}(0|H^{\neg\mathsf{comp}}).$$

Hence, recalling that reserve price $r$ is normalized to 1,

$$\frac{1}{\rho}[\overline{R}(\rho|H^{\neg\mathsf{comp}}) - \overline{R}(0|H^{\neg\mathsf{comp}})] \le \overline{R}(0|H^{\neg\mathsf{comp}}) \le r = 1.$$

Altogether, this implies that $\frac{1}{\rho}[\overline{R}(\rho|H) - \overline{R}(0|H)] \le 1 - s_{\mathsf{comp}}$, which concludes the proof. ■

In words, if total average revenue for histories $H$ increases by more than $\kappa \times \rho$ when bids are multiplied by $(1+\rho)$, the share of competitive histories in $H$ is bounded above by $1-\kappa$. In the extreme case where the density of competing bids is zero just above winning bids, we have that $\overline{R}(\rho|H) - \overline{R}(0|H) \simeq \rho\overline{R}(0|H)$ for $\rho$ small. This implies that $s_{\mathsf{comp}} \le 1 - \overline{R}(0|H)$.

Using estimator $\widehat{R}$, Proposition 2 lets us compute a probabilistic upper-bound to the share of competitive histories. In Section 6, we show how this estimator can provide non-trivial bounds on the share of competitive histories in different datasets.

The next section derives a tighter bound on the share of competitive histories by exploiting a greater set of incentive compatibility constraints. Appendix B shows how one can exploit *all* implications from equilibrium, as well as derive probabilistic bounds on other moments of interest, such as the share of non-competitive auctions, or the sum of deviation temptations.

## 5 A Tighter Bound on Competitive Histories

The bound in Proposition 2 exploits only upward deviations in bids. Our tighter bound exploits the informational content of both upward and downward deviations, as well as subjective restrictions the econometrician or regulator is willing to impose on the environment (here on markups). We emphasize that we exploit optimality conditions that are weaker, and more plausibly satisfied, than equilibrium, consistent with the critique of Fershtman and Pakes (2012). For expositional purposes, we assume private values and address the case of common values in Appendix B.

Take as given an adapted set of histories $H$. Take as given scalars $\rho_n \in (-1, \infty)$ for $n \in \mathcal{M} = \{-\underline{n}, \cdots, \overline{n}\}$, such that $\rho_0 = 0$ and $\rho_n < \rho_{n'}$ for all $n' > n$. For each history $h_{i,t} \in H$ and for each $n \in \mathcal{M}$, let $d_{h_{i,t},n} \equiv D_i((1+\rho_n)b_{h_{i,t}}|h_{i,t})$. That is, $d_{h_{i,t},n}$ is firm $i$'s subjective residual demand at history $h_{i,t}$, when applying a coefficient $1+\rho_n$ to its original bid. For any history $h \in H$, let $\omega_h = ((d_{h,n})_{n\in\mathcal{M}}, c_h)$ be the costs and beliefs of the firm associated with history $h$. Let $\omega_H = (\omega_h)_{h\in H}$ denote the profile of beliefs and costs across all histories in $H$.

For each set of adapted histories $H$, each deviation $n$, and each profile $\omega_H = (\omega_h)_{h\in H}$, let

$$D_n(\omega_H, H) \equiv \frac{1}{|H|} \sum_{h_{i,t}\in H} d_{h_{i,t},n}$$

be the average residual demand when firms' beliefs are given by $\omega_H$. Recall that, for each $n \in \mathcal{M}$

$$\widehat{D}(\rho_n|H) = \frac{1}{|H|} \sum_{h_{i,t}\in H} \mathbf{1}_{(1+\rho_n)b_{h_{i,t}} < \wedge \mathbf{b}_{-i,h_{i,t}}}$$

is the sample counterpart of average residual demand. By Lemma 1, $\widehat{D}(\rho_n|H)$ is a consistent estimator of $D_n(\omega_H, H)$, when $\omega_H$ matches the firms' true equilibrium beliefs.

Under private values, at every competitive history $h \in H$, there must exist costs $c_h$ and subjective demands $d_h = (d_{h,n})_{n \in \mathcal{M}}$ satisfying

$$\text{feasibility:} \quad c_h \geq 0; \quad \forall n, \;\; d_{h,n} \in [0,1]; \quad \forall n, n' > n, \;\; d_{h,n} \geq d_{h,n'} \tag{F}$$

$$\text{incentive compatibility:} \quad \forall n, \quad [(1+\rho_n)b_h - c_h]\, d_{h,n} \leq [(1+\rho_0)b_h - c_h] d_{h,0}. \tag{IC}$$

We allow the analyst or econometrician to include economically plausible constraints on the environment. For simplicity and interpretability, we focus on markup constraints of the form

$$\forall h, \quad \frac{b_h}{c_h} \in [1+m, 1+M] \tag{EP}$$

where $m \geq 0$ and $M \in (m, +\infty]$ are minimum and maximum markups.[16] Constraints (EP) provide what we think is a transparent way for regulators to express minimal subjective beliefs over the environment without making further assumptions of the sort embedded in a Bayesian prior.[17] We provide a detailed discussion of the impact of constraints (EP) on identification in Section 6.

For each profile of costs and beliefs $\omega_H$, define

$$H_{\mathsf{comp}}(\omega_H) \equiv \{h \in H \text{ s.t. } (d_h, c_h) \text{ satisfy (F), (IC) and (EP)}\},$$

the set of histories in $H$ that satisfy plausibility constraint (EP) and are rationalizable as competitive under $\omega_H$.

[16]In Chassang et al. (2019), we discuss plausibility constraints on the informativeness of signals.

[17]A partially sophisticated regulator may express probabilistic constraints without being willing to commit to a full prior. For instance: "90% of the time, margins are greater than 5%". This would add an aggregate constraint to the inference problem.

Let $T > 0$ denote a tolerance margin used to define confidence intervals. We define inference problem (P) and its solution $\widehat{s}$ by

$$\widehat{s} = \max_{\omega_H} \frac{|H_{\mathsf{comp}}(\omega_H)|}{|H|} \tag{P}$$

$$\text{s.t. } \forall n \in \mathcal{M}, \quad D_n(\omega_H, H) \in \left[\widehat{D}(\rho_n|H) - T, \widehat{D}(\rho_n|H) + T\right]. \tag{$\widehat{CR}$}$$

Program (P) finds beliefs and costs $\omega_H$ that maximize the share of competitive histories in $H$, across all profiles $\omega_H$ that are consistent with the data, in the sense that average residual demands $(D_n(\omega_H, H))_{n\in\mathcal{M}}$ are close to their sample counterparts $(\widehat{D}(\rho_n|H))_{n\in\mathcal{M}}$.

As was previously noted, conditions (F), (IC) and ($\widehat{CR}$) exploit some but not all the informational content of equilibrium. We clarify in Appendix B that we would exploit all the empirical content of equilibrium if we imposed demand consistency requirements ($\widehat{CR}$) conditional on all different values of bids and costs $c$ (corresponding to the bidder's private information at the time of bidding). We are concerned that this stretches both the limits of our data, and of bidder sophistication. Relying on a weaker set of optimality conditions makes our estimates more robust to partial failures of optimization.

Our next result shows that estimator $\widehat{s}$ provides a robust upper bound to the share of competitive histories in $H$.

**Proposition 3.** *Assume true beliefs and costs $\omega_H$ satisfy* (EP)*, and denote by $s_{\mathsf{comp}}$ the true share of competitive histories under $\omega_H$. With probability at least* $1-2|\mathcal{M}|\exp(-T^2|H|/2N_{\max})$*, $\hat{s} \geq s_{\mathsf{comp}}$.*

Proposition 3 lets us define robust tests of non-competitive behavior. For any threshold fraction $s_0 \in (0, 1]$ of competitive histories, let $\tau \equiv \mathbf{1}_{\hat{s}\leq s_0}$. Test $\tau$ rejects the null that $s_{\mathsf{comp}} = 1$ whenever $\hat{s}$ is strictly lower than $s_0$. By varying the set $H$ of adapted histories, we can apply test $\tau$ to a single given firm, or a set firms in a given industry. By Proposition 3, for any $s_0$, the probability that test $\tau$ rejects the null when the null is true (i.e., type

1 error) is bounded above by $2|\mathcal{M}|\exp(-T^2|H|/2N_{\max})$, permitting conservative inference. As $|H|$ becomes large, $\tau$ accepts data generated by competing firms with probability 1, and becomes an asymptotically safe test in the sense of Ortner et al. (2019).

Proposition 3 is reassuring because it does not rely on the ergodicity of the underlying state. In practice however, one might impose asymptotic normality assumptions yielding less conservative confidence sets than the ones obtained in Lemma 1. We do so in our empirical analysis, relying on bootstrap to obtain a confidence set for demand profile $(\overline{D}(\rho_n|H))_{n\in\mathcal{M}}$.

# 6 Empirical Evaluation

In this section, we explore the implications of our approach in real data. We take this as an opportunity to clarify the information content of different deviations. In addition, we provide suggestive evidence that our tests tend to accept the null when applied to data from likely competitive auctions, and reject the null when applied to data from likely non-competitive industries. Computational details, as well as sensitivity checks, are collected in Appendix C.

## 6.1 Data

We work with two datasets of first-price auctions held in Japan. Our first dataset consists of 7,109 auctions held between 2004 and 2018 in 14 cities located in the Tohoku region and the Ibaraki prefecture of Japan. The auctions are sealed-bid first-price auctions with a publicly announced reserve price, so that the results of Sections 4 and 5 apply directly.

Our second dataset consists of auctions held by the Ministry of Land, Infrastructure and Transportation in Japan. We have data on 78,272 auctions held between 2001 and 2006. The auctions use a first-price sealed-bid format, with a secret reserve price and re-bidding in case all bids are above the reserve price. In the event the reserve is not met, the lowest bid is typically revealed to the bidders. Our empirical analysis relies on the first round of each auction. The possibility of re-bidding requires specific theoretical adjustments.

Results in Sections 4 and 5 would be unchanged if we assumed that bids in the first round did not influence outcomes in further rounds. We show in Appendix A how to derive bounds on the set of competitive auctions when bids in the first round do affect equilibrium play in later stages. Incentive compatibility constraints for bid increases are essentially unchanged. For bid reductions, the main difficulty is to assess losses in continuation values when the bid reduction changes the lowest bid reported to bidders if the lowest bid is greater than the reserve price. We report bounds on the share of competitive histories computed under the assumption that changing the reported minimum bid reduces a bidder's continuation value by at most 50%.[18]

| | | N | Mean | S.D. |
|---|---|---|---|---|
| **City Auctions** | | | | |
| By Auctions | | | | |
| | reserve price (mil. Yen) | 7,109 | 23.192 | 91.33 |
| | lowest bid (mil. Yen) | 7,109 | 21.539 | 85.17 |
| | lowest bid / reserve | 7,109 | 0.939 | 0.06 |
| | #bidders | 7,109 | 7.425 | 3.77 |
| By Bidders | | | | |
| | participation | 2,267 | 23.29 | 43.58 |
| | number of times lowest bidder | 2,267 | 3.14 | 6.22 |
| | total revenue (mil. Yen) | 2,267 | 67.54 | 219.37 |
| **National Auctions** | | | | |
| By Auctions | | | | |
| | reserve price (mil. Yen) | 78,272 | 105.121 | 259.58 |
| | lowest first round bid (mil. Yen) | 78,272 | 101.909 | 252.30 |
| | lowest bid / reserve | 78,272 | 0.970 | 0.10 |
| | #bidders | 78,272 | 9.883 | 2.27 |
| By Bidders | | | | |
| | participation | 29,670 | 26.40 | 94.61 |
| | number of times lowest bidder | 29,670 | 2.64 | 10.57 |
| | total revenue (mil. Yen) | 29,670 | 264.23 | 1312.77 |

Table 1: Sample Statistics – City and National Level Data

[18]Note that the reported bid is above the fixed reserve price which bidders must beat to win the auction.

Table 1 reports descriptive statistics for both datasets. At the auction level, we report the mean and standard deviation of reserve prices, the lowest bid,[19] the lowest bid as a fraction of the reserve price, and the number of bidders participating in the auction. At the bidder level, we report the mean and standard deviation of the number of auctions a bidder participates in, the number of auctions she wins, and her total revenue.

We note that national-level auctions have higher reserve prices and a greater number of participants than city-level auctions, but the two datasets are broadly comparable. We also note that there is large heterogeneity in reserve prices, and in participation. Some projects are very large, and some bidders participate very often.

## 6.2 A Case Study

We first illustrate the mechanics of inference using data from a specific city – Tsuchiura, located in the Ibaraki prefecture. We select this city for two reasons: first, Chassang and Ortner (2019) provide evidence that there was collusion in auctions held prior to October 2009; second, the data turns out to be well suited to illustrate the information content of different incentive compatibility conditions.

We consider different combinations of deviations $\rho \in \{-.02, 0, .0008\}$. The distribution of $\Delta$ and the deviations we consider (in dashed lines) are illustrated Figure 7. The deviations are selected to deliver crisp illustrative results for this specific dataset. We are specifically interested in illustrating the informative content of individual deviations, as well as complementarities between upward and downward deviations.

We note that tied bids are present in the data, but that their mass is small. While intriguing, tied bids do not play an important part in our results. Observing that tied bids are non-competitive (a very small reduction in bids is a profitable deviation), we label them as non-competitive.

[19] This corresponds to the winning bid when the reserve price is public. If the reserve price is secret, then the lowest bid need not be a winning bid.

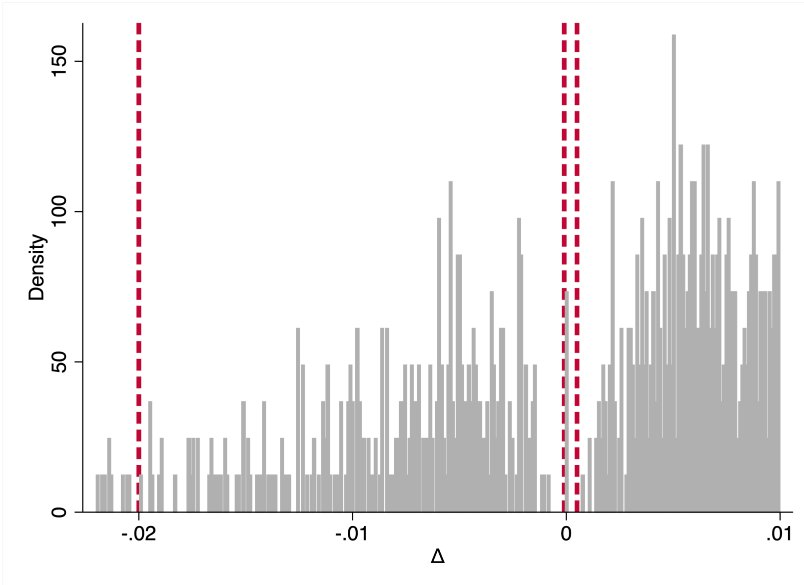



Figure 7: Distribution of $\Delta$ for the city of Tsuchiura, 2007–2009.

Figure 8 presents our estimate of a 95% confidence upper bound on the share of competitive histories as a function of the minimum markup $m$ in constraint (EP). For these estimates and all the estimates that we present below, we use bootstrap to set tolerance $T$ so that demand consistency requirement $(\widehat{CR})$ holds with probability 95%. We set the maximum markup to $M = .5$.[20] We investigate inference for different combinations of deviations.[21]

**A single upward deviation.** We first consider a small upward deviation $\rho = .0008$, corresponding to the analysis of Section 4. At every history $h$ we seek beliefs $d_{h,0} \geq d_{h,1}$ and a cost $c_h$ such that

$$d_{h,0}(b_h - c_h) \geq d_{h,1}((1+\rho_1)b_h - c_h) \iff d_{h,0} - (1+\rho_1)d_{h,1} \geq (d_{h,0} - d_{h,1})\frac{c_h}{b_h}.$$

[20] As Figure C.1 illustrates, our results are not highly sensitive to the choice of $M$.

[21] The ranges of markups that we consider contain prior estimates in the literature. For instance, Krasnokutskaya (2011) estimates markups ranging from 0.1% to 24% for regular bidders participating in Michigan highway procurement auctions. Similarly, Bajari et al. (2014) estimates markups ranging from 2.9% to 26.1% for firms bidding in highway procurement auctions in California.

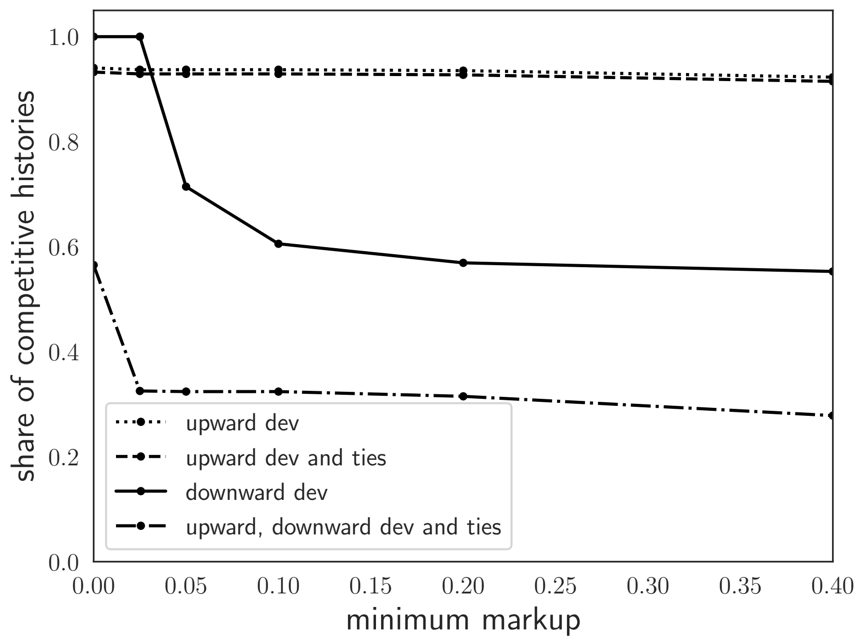



Figure 8: Share of competitive histories, Tsuchiura. Maximum markup $M = .5$.

Since $d_{h,0} \geq d_{h,1}$, this is most easily satisfied by setting $c_h/b_h = 1/(1+M)$. An upward deviation is least profitable (and so the data is best explained) when costs are low. In that case, constraints (IC), (F), and ($\widehat{CR}$) define a convex set. If

$$\widehat{D}(0|H) - (1+\rho_1)\widehat{D}(\rho_1|H) < \frac{1}{1+M}(\widehat{D}(0|H) - \widehat{D}(\rho_1|H)), \tag{4}$$

then for a tolerance $T$ small enough, (IC), (F), and ($\widehat{CR}$) cannot be solved together for all histories $h \in H$. Note that (4) holds even if $M = +\infty$ whenever an increase in bids increases sample revenue.

In the case of auctions from Tsuchiura, a small upward deviation hardly changes a bidder's likelihood of winning an auction: $\widehat{D}(0|H) \simeq \widehat{D}(\rho_1|H)$. As a result a minimum share of histories must be considered non-competitive.

**A single upward deviation and tied bids.** Any mass of tied bids is inherently non-competitive since they create a meaningful benefit from reducing bids by the smallest possible amount. In our data, as Figure 8 shows, this has a very small impact on our estimate of the share of non-competitive histories. Although noticeable, tied bids are not frequent.

**A single downward deviation.** We now consider the implications from a single downward deviation $\rho = -.02$. At every history $h$ we seek beliefs $d_{h,-1} \geq d_{h,0}$ and a cost $c_h$ such that

$$d_{h,0}(b_h - c_h) \geq d_{h,-1}((1+\rho_{-1})b_h - c_h) \iff (1+\rho_{-1})d_{h,-1} - d_{h,0} \leq (d_{h,-1} - d_{h,0})\frac{c_h}{b_h}.$$

Since $d_{h,-1} \geq d_{h,0}$, this IC condition is most easily satisfied if $c_h/b_h = 1/(1+m)$. In that case, constraints (IC), (F), and $(\widehat{CR})$ define a convex set and (setting tolerance $T$ to 0 for simplicity), can be satisfied for all histories $h \in H$ if and only if

$$(1+\rho_{-1})\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) \leq \frac{1}{1+m}\left[\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)\right]. \tag{5}$$

Since $\rho_{-1} < 0$, it follows that condition (5) always hold if $m = 0$. This is intuitive: for a small margin, say $m = 2\%$, then reducing bids by 2% results in net losses for each auction $(.98 \times 1.02 - 1 < 0)$. In contrast, if $m > 1/(1+\rho_{-1}) - 1$ and demand increase $\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)$ is sufficiently large, then (5) does not hold and a minimum share of histories must be considered non competitive.

This is illustrated Figure 8. When considering only downward deviations (and unlike the case of upward deviations), our bound for the share of competitive histories is equal to 1 for low values of minimum markup $m$. However, as minimum markup $m$ increases, downward deviations imply that a significant share of histories must be non-competitive. This is is due to the fact that in the case of Tsuchiura, a 2% drop in prices leads to a 44 percentage-points increase in the probability of winning the auction (roughly doubling demand). For modest minimum markups, such an increase in demand makes reducing one's bid an attractive deviation.

**Complementary upward and downward deviations.** Conditions (4) and (5) highlight that individual upward and downward deviations are best rationalized as competitive by different costs. An upward deviation is least attractive when cost $c_h$ is low. A downward deviation is least attractive when cost $c_h$ is large. This implies that upward and downward deviations are complementary from the perspective of inference. As Figure 8 shows, even for low values of minimum markup $m$ considering both an upward and a downward deviation leads to a lower bound for the share of competitive histories than either upward or downward deviations alone. This is because the high costs needed to ensure that a downward deviation is not attractive also make upward deviations more attractive. Appendix B establishes this complementarity formally in a simple case.

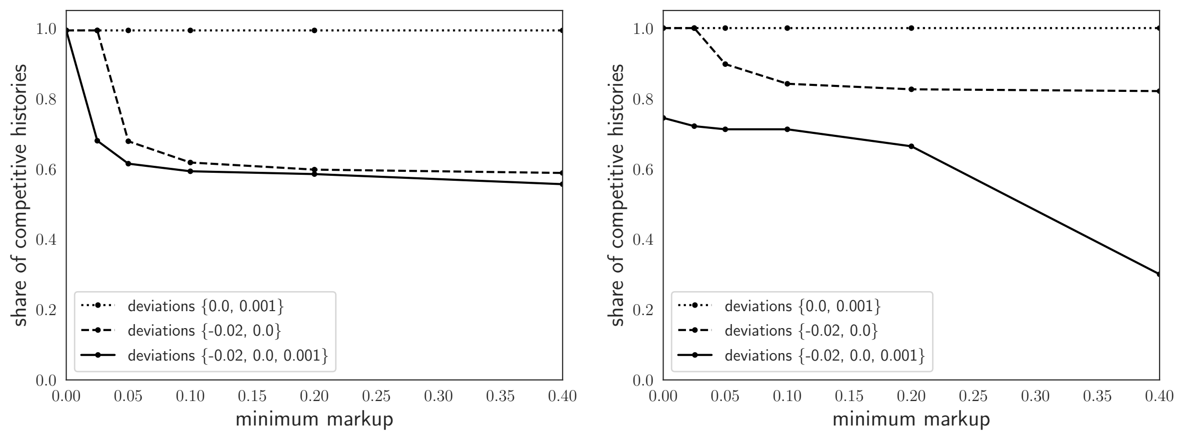
### 6.3 Findings from aggregate data

We now apply our tests to our aggregate datasets. Going forward, we use the fixed set of deviations $\{-.02, 0, .001\}$ for all datasets.

Figure 9 shows our estimates of the share of competitive histories as a function of minimum markup $m$, for city and national auctions.[22] We note that upward deviations alone do not allow us to detect non-competitive histories. One explanation for this is that aggregating causes us to mix competitive and non-competitive histories, thereby weakening our ability to detect non-competitive auctions. Note that this does not mean that upward deviations are not informative: especially in the case of national data, considering both upward and downward deviations yields significantly lower bounds on the share of competitive histories than either deviations alone.[23]

[22] In the case of national auctions we use more conservative incentive compatibility conditions accounting for re-bidding detailed in Appendix A. Appendix C shows that our estimates are not very sensitive to the value of maximum markup $M$, or assumptions about continuation payoffs upon re-bidding.

[23] In particular, when $m$ is small enough, downward deviations alone place no restrictions on the share of competitive histories.




(a) City data (b) National data

Figure 9: Share of competitive histories, city and national level data. Deviations $\{-0.02, 0, 0.001\}$. Maximum markup 0.5.

## 6.4 Zeroing-in on specific firms

As Figure 2 suggests, our tests can be applied to individual firms. As we highlight in Ortner et al. (2019), detecting non-competitive behavior at the firm level helps reduce the potential side-effects of regulatory oversight. Specifically, it ensures that a cartel cannot use the threat of regulatory crackdown to discipline bidders.

For both city and national auction samples, we consider the thirty firms that participate in the most auctions in each data set. Table 2a shows our estimates of a 95% confidence bound on the share of competitive histories for firms active in the city sample. Column 4 shows estimates obtained using the bound from Proposition 2. Column 5 shows tighter estimates obtained using the bound from Proposition 3 using deviations $\{-0.02, 0, 0.001\}$, minimum markup $m = .02$ and maximum markup .5. Table 2b reports the same results for firms active in the national sample.

In the city sample, the bound from Proposition 2 is less than 1 for six out of thirty firms. The bound from Proposition 3 is less than 1 for twenty-three firms. In the national sample, the bound from Proposition 2 is less than 1 for three of the thirty firms, while the bound

| 1<br>Rank | 2<br>Participation | 3<br>Share won | 4<br>Share comp (Prop 2) | 5<br>Share comp (Prop 3) |
|---|---|---|---|---|
| 1 | 347 | 0.19 | **0.85** | **0.19** |
| 2 | 336 | 0.21 | 1.00 | **0.26** |
| 3 | 299 | 0.08 | 1.00 | **0.31** |
| 4 | 293 | 0.05 | 1.00 | 1.00 |
| 5 | 293 | 0.14 | 1.00 | 1.00 |
| 6 | 290 | 0.20 | 1.00 | 1.00 |
| 7 | 287 | 0.14 | **0.90** | **0.30** |
| 8 | 269 | 0.09 | 1.00 | **0.32** |
| 9 | 268 | 0.12 | 1.00 | 1.00 |
| 10 | 262 | 0.18 | 1.00 | **0.24** |
| 11 | 259 | 0.12 | 1.00 | **0.33** |
| 12 | 252 | 0.12 | 1.00 | **0.33** |
| 13 | 241 | 0.16 | 1.00 | **0.22** |
| 14 | 239 | 0.09 | 1.00 | **0.32** |
| 15 | 238 | 0.11 | 1.00 | **0.33** |
| 16 | 227 | 0.12 | 1.00 | **0.32** |
| 17 | 226 | 0.08 | 1.00 | **0.18** |
| 18 | 225 | 0.12 | 1.00 | **0.33** |
| 19 | 223 | 0.07 | **0.97** | **0.23** |
| 20 | 223 | 0.08 | **0.96** | **0.33** |
| 21 | 220 | 0.07 | **0.96** | **0.14** |
| 22 | 218 | 0.14 | **0.91** | **0.33** |
| 23 | 211 | 0.17 | 1.00 | **0.25** |
| 24 | 210 | 0.15 | 1.00 | 1.00 |
| 25 | 209 | 0.11 | 1.00 | **0.13** |
| 26 | 204 | 0.06 | 1.00 | 1.00 |
| 27 | 203 | 0.12 | 1.00 | 1.00 |
| 28 | 199 | 0.06 | 1.00 | **0.32** |
| 29 | 199 | 0.16 | 1.00 | **0.26** |
| 30 | 190 | 0.08 | 1.00 | **0.32** |

(a) City Data

| 1<br>Rank | 2<br>Participation | 3<br>Share won | 4<br>Share comp (Prop 2) | 5<br>Share comp (Prop 3) |
|---|---|---|---|---|
| 1 | 4044 | 0.17 | 1.00 | **0.61** |
| 2 | 3854 | 0.07 | 1.00 | **0.49** |
| 3 | 3621 | 0.12 | 1.00 | **0.56** |
| 4 | 2998 | 0.15 | 1.00 | 1.00 |
| 5 | 2919 | 0.06 | 1.00 | **0.47** |
| 6 | 2547 | 0.08 | 1.00 | **0.69** |
| 7 | 2338 | 0.07 | 1.00 | **0.70** |
| 8 | 2333 | 0.07 | 1.00 | **0.71** |
| 9 | 2328 | 0.04 | 1.00 | **0.58** |
| 10 | 2292 | 0.06 | 1.00 | **0.75** |
| 11 | 2237 | 0.08 | **0.92** | **0.58** |
| 12 | 2211 | 0.03 | 1.00 | **0.54** |
| 13 | 2015 | 0.09 | 1.00 | **0.74** |
| 14 | 1984 | 0.08 | 1.00 | **0.72** |
| 15 | 1727 | 0.07 | 1.00 | 1.00 |
| 16 | 1674 | 0.05 | 1.00 | **0.80** |
| 17 | 1661 | 0.03 | 1.00 | **0.53** |
| 18 | 1660 | 0.08 | 1.00 | **0.75** |
| 19 | 1589 | 0.07 | 1.00 | **0.78** |
| 20 | 1427 | 0.10 | 1.00 | 1.00 |
| 21 | 1393 | 0.06 | 1.00 | **0.79** |
| 22 | 1392 | 0.07 | 1.00 | 1.00 |
| 23 | 1370 | 0.04 | 1.00 | **0.81** |
| 24 | 1368 | 0.14 | 1.00 | 1.00 |
| 25 | 1353 | 0.05 | 1.00 | **0.69** |
| 26 | 1342 | 0.09 | 1.00 | 1.00 |
| 27 | 1337 | 0.04 | 1.00 | **0.78** |
| 28 | 1326 | 0.08 | 1.00 | **0.69** |
| 29 | 1291 | 0.06 | **0.95** | **0.79** |
| 30 | 1260 | 0.06 | **0.95** | **0.65** |

(b) National Data

Table 2: Share of competitive histories, individual firms

95% confidence bound on the share of competitive auctions for top thirty most active firms. The first column corresponds to the ranking of the firms and the second column corresponds to the number of auctions in which each firm participates. Column 3 shows the fraction of auctions that each of these firms wins. Columns 4 and 5 present our 95% confidence bound on the share of competitive histories for each firm based on Proposition 2 and Proposition 3, respectively. For our estimates of column 5, we use deviations $\{-0.02, 0, 0.001\}$, minimum markup $m = 0.02$ and maximum markup $M = 0.5$.

from Proposition 3 is less than 1 for twenty-four firms.

## 6.5 Consistency with proxies for collusion

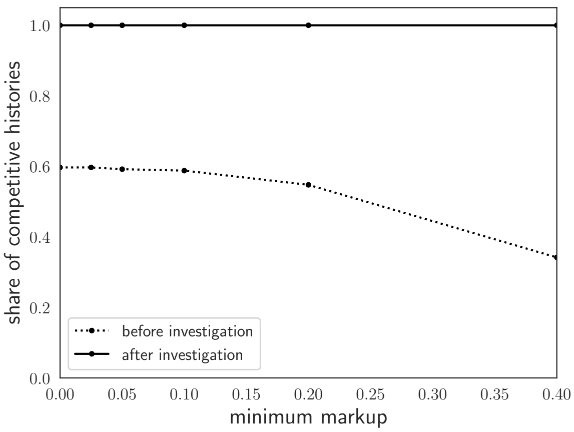
We now show that our bounds on the share of competitive histories are consistent with proxies of collusive behavior.

**Before and after prosecution.** As we noted in Section 2, Figure 4, the JFTC investigated firms bidding in four groups of national auctions during the period for which we have data: auctions labeled Bridges, Electric, Floods, and Pre-Stressed Concrete. We now compute bounds on the set of competitive histories before and after investigation. We exclude the Bridge category because there are too few observations in the post-investigation sample to obtain a reliable confidence set via bootstrap (58 bids, vs more than 560 in the other industries).

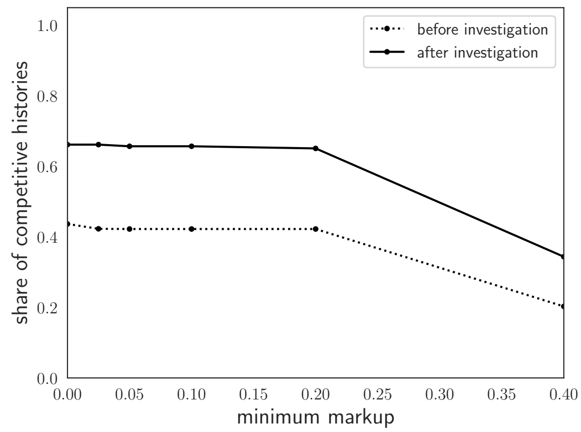
Figure 10 shows our estimates of the 95% confidence bound on the share of competitive histories as a function of minimum markup $m$, for the three remaining groups of firms. For all the figures, we use deviations $\{-0.02, 0, 0.001\}$, and set maximum markup $M = .5$.

For auctions in the Electric category, findings are consistent with the interpretation that collusion took place before but not after the investigation. This is not the case in the case of Pre-Stressed Concrete. Our bounds indicate similar, and if anything higher prevalence of non-competitive behavior. As it turns out, firms in the Pre-Stressed Concrete industry continued to operate as a cartel even after the investigation (see Kawai and Nakabayashi, 2018, for a discussion).

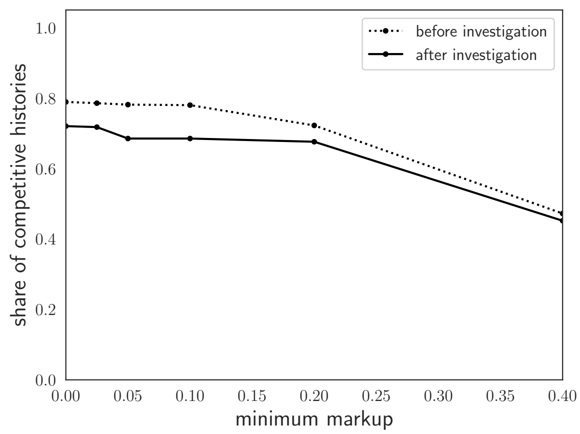
The case of Flood auctions is less clear cut. Our bound on the share of competitive histories is higher after the investigation than before, but is not equal to 1. This follows from the fact that for Flood auctions occurring after investigation, a 2% downward deviation causes a 37.6% increase in demand while a .1% upward deviation causes a 0% decrease in demand. For Electric auctions the changes in demand for downward and upward deviations are respectively 15.5% and -.86%.




(a) Electric




(b) Floods




(c) Pre-Stressed Concrete

Figure 10: Share of competitive histories, before and after JFTC investigation. Deviations $\{-0.02, 0, 0.001\}$, $M = 0.5$.

**High vs. low bids.** Figure 11 plots our estimates of the 95% confidence bound on the share of competitive histories for city-level auctions as a function of minimum markup $m$. We divide the data between histories with high bids relative to the reserve price (i.e., $\frac{b}{r} \geq 0.9$) and histories with low bids relative to the reserve price (i.e., $\frac{b}{r} < 0.9$). Since the reserve price is known to bidders in city level auctions, these two sets of histories are adapted. As the figure shows, the fraction of competitive histories is lower at histories at which bids are high relative to the reserve price, a finding that is consistent with the idea that collusion is more likely at histories at which bidders place higher bids.

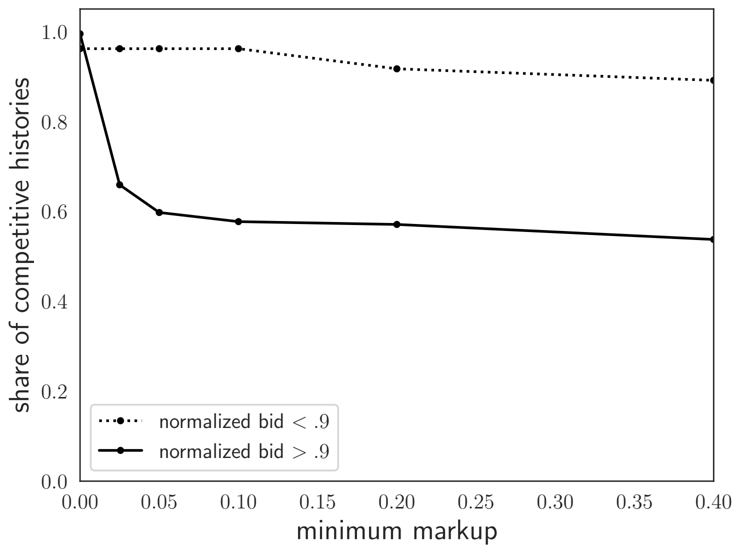



Figure 11: Share of competitive histories by bid level, city data. Deviations $\{-0.02, 0, 0.001\}$. Maximum markup 0.5.

# 7 Discussion

This paper develops tests of non-competitive bidding motivated by the observation that winning bids are isolated in data from procurement auctions held in Japan. The tests we propose establish robust bounds on the minimum share of non-competitive histories needed to rationalize the data. Our tests help identify another suspicious pattern in the data: the increase in demand from a 2% reduction in bids is very large.

While our tests are conservative (they can be passed by any firm that is bidding competitively under some information structure), this concern is echoed by practitioners (Imhof et al., 2018) who emphasize the cost of launching a formal investigation (based on statistical evidence) against non-collusive firms. In a companion paper, Ortner et al. (2019), we identify another important property of such tests: antitrust investigation based on tests with vanishing false positive rate does not generate new collusive equilibria. This addresses the concern that data-driven regulation may inadvertently enhance a cartel's ability to collude (Cyrenne, 1999, Harrington, 2004). For these two reasons, we believe that our tests provide a sensible starting point for data-driven antitrust.

Appendices A and B collect important extensions of our baseline framework: how to deal with secret reserve prices and re-bidding; how to deal with common values; and how to construct money denominated metrics of non-competitive behavior.

We conclude with an open-ended discussion of possible collusive explanations for the bidding behavior observed in the data. We start by clarifying bidding patterns predicted by benchmark models of collusion.

## 7.1 Benchmark cartel behavior

Standard models of collusion (see for instance Rotemberg and Saloner (1986), Athey and Bagwell (2001, 2008)) do not predict the pattern of isolated winning bids we see in the data.

A cartel's main concern is to incentivize losers not to undercut the winning bid (their play is not stage game optimal). In contrast, the behavior of the designated winner is stage game optimal. This is achieved by having a losing bidder bid just above the designated winner, which ensures that the winning bidder is not tempted to increase its bid. This frees up pledgeable future surplus that can be used to dissuade losers from undercutting the winner, and maintain higher equilibrium bids.

As a consequence, standard models of collusion predict a point mass at $\Delta \approx 0$, rather than the pattern of isolated winning bids identified in the data.[24] Isolated winning bids imply that some of the cartel's pledgeable surplus must be wasted on keeping winning bidders from increasing their bid. This is counterintuitive but turns out to be consistent with richer models of collusive behavior.

[24] Note that some datasets do exhibit a spike in density at $\Delta \approx 0$.

## 7.2 Anomalous bids as a side-effect of antitrust oversight

Our data exhibits two surprising and seemingly somewhat contradictory patterns. First, winning bids are locally isolated. Second, there is a large mass of bids within 2% of the winning bid: in our city-level data, a 2% downward deviation increases the probability of winning an auction by 42 percentage points, roughly doubling demand. Both patterns can be explained as an adjustment of the optimal cartel behavior described Section 7.1 in response to scrutiny from antitrust authorities.

The main observation is that antitrust authorities should, and in fact do, scrutinize patterns of nearly tied bids corresponding to the optimal cartel behavior of Section 7.1. Indeed, tied bids cannot be rationalized as competitive under any incomplete information structure, since players would benefit from a small reduction in bids. In addition, as we highlight in Ortner et al. (2019), the corresponding tests of non competitive behavior are very high powered: even with few auctions, it is very improbable that a significant share of winning bids should be approximately tied.[25] In practice, nearly tied bids are flagged out by statistical screens such as variance tests that single out auctions for which the dispersion of bids is small (DOJ, 2005, Imhof et al., 2018).

If antitrust authorities scrutinize auctions with tied bids, a cartel eager to avoid scrutiny will benefit from adjusting its bidding behavior as follows:

(i) ensure that bidders do not submit tied or approximately tied bids;

(ii) keep the winning bidder's temptation to increase its bid small, by having a second bidder bids close to, though not tied with, the winning bid.

Adjustment (*i*) generates the isolated winning bids that we document. Adjustment (*ii*) generates the high sensitivity of demand to a 2% reduction in bids.

[25] In contrast, detecting that winning bids are isolated requires a large amount of data: even under competition, this could occur with positive probability in any finite data set.

## 7.3 Other potential explanations

**Missing bids as robust coordination.** A possible role for isolated winning bids is to facilitate coordination on a specific designated-winner. Being able to guarantee the identity of the winner may be important to a cartel for allocative efficiency, or to reduce the costs of dynamic incentive provision when utility is not transferable.

In this respect keeping winning bids isolated ensures that the designated winner does win the contract, even if bidders cannot precisely agree on exact bids ex ante, or if bids can be perturbed by small trembles (say a fat finger problem).

**Non-collusive explanations.** Finally, it is worth discussing potential non-collusive explanations for missing bids. The first is simply that bidders are committing errors, say playing an $\epsilon$-equilibrium of the game. This explanation is not entirely satisfactory for two reasons. First, the potential gains from downward deviations are not small. Second, natural models of erroneous play do not generate the patterns we see in the data. For instance, by adding noise on bids, quantal response equilibrium (McKelvey and Palfrey, 1995) would smooth out rather than enhance the pattern of missing bids.

A possible direction to explain the anomalous bidding patterns we note in the data is to take into account dynamic payoff consequences of winning an auction, through either capacity constraints, or learning by doing. A rapid analysis suggests that such dynamic considerations are unlikely to explain the data. Capacity constraints essentially correspond to an increase in the bidder's cost reflecting reduced continuation values. This would tend to increase the attractiveness of upward deviations. Similarly, learning-by-doing essentially reduces the cost of accepting a project. This would tend to increase the attractiveness of downward deviations.

# Online Appendix – Not for Publication

# A Multistage Bidding

National level auctions in our data follow a first-price auction format with a secret reserve price. This means that the auction is a multistage game, with stages $k \in \{1, \cdots, K\}$. The auctioneer picks a secret reserve price $r$. At each stage $k$, bidders submit bids $b_{i,k}$. A winner is designated if and only if $\min_i b_{i,k} \leq r$. In this case, the winner is paid her bid. If instead $\min_i b_{i,k} > r$ the game continues to an additional stage. At the end of each stage without a winner the lowest bid is revealed. The reserve price is constant across stages. In this Appendix we extend the revealed preference inequalities of Sections 4 and 5 to multistage first-price auctions.

In a multistage auction, a bidder's continuation strategy after her first bid is a contingent plan dependent on the information revealed at each stage. We denote by $b_{i,1}$ bidder $i$'s first bid, and by $\beta_i$ her continuation play, mapping future information to bids.

Given an equilibrium strategy $\sigma_i = (b_{i,1}, \beta_i)$ by player $i$ we consider first-stage-only deviations $\sigma_i' = (b_{i,1}', \beta_i)$ such that player $i$'s initial bid is different, but her continuation contingent plan, as a function of her own private signals, and the play of others, is unchanged.

Let $\mathsf{win}_{i,k}$ denote the event that bidder $i$ wins in round $k$. Expected profits under $\sigma_i$ and $\sigma_i'$ take the form

$$\mathbb{E}_{\sigma_i}[\pi_i] = (b_{i,1} - c_i)\mathsf{prob}_{\sigma_i,\sigma_{-i}}(\mathsf{win}_{i,1}) + \mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\right]$$

$$\mathbb{E}_{\sigma_i'}[\pi_i] = (b_{i,1}' - c_i)\mathsf{prob}_{\sigma_i',\sigma_{-i}}(\mathsf{win}_{i,1}) + \mathbb{E}_{\sigma_i',\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\right]$$

We now introduce a classification of histories following upward and downward deviations in the first round as a function of how they affect the continuation play. We say that a deviation is marginal for continuation, if it changes whether the auction continues after period 1. When a deviation is marginal for information, it changes the information available to participants in future periods. If a deviation is non-marginal, it does not affect continuation play. This corresponds to the following formal definition.

**Definition A.1.** *Consider an upward deviation $b_{i,1}' > b_{i,1}$. It is* marginal for continuation *(MC)* *if and only if* $b_{i,1} \leq r < \wedge \mathbf{b}_{-i,1}$*, and* $b'_{i,1} > r$*. It is* marginal for information *(MI) if and only if* $r < b_{i,1} < \wedge \mathbf{b}_{-i,1}$*. It is* non-marginal *(NM) otherwise.*

*Consider a downward deviation* $b'_{i,1} < b_{i,1}$*. It is* marginal for continuation *(MC) if and only if* $b'_{i,1} \leq r < \wedge \mathbf{b}_{-i,1}$*, and* $b_{i,1} > r$*. It is* marginal for information *(MI) if and only if* $r < b'_{i,1} < \wedge \mathbf{b}_{-i,1}$*. It is* non-marginal *(NM) otherwise.*

Note that we can assess the marginality of deviations using data, since it only relies on observed period 1 bids. Note also that the probability a given deviation is marginal for continuation or information only depends on the agent's beliefs about bids $b_{-i,1}$,

We have that

$$\begin{aligned}
\mathbb{E}_{\sigma_i}[\pi_i] &= (b_{i,1} - c_i)\mathsf{prob}_{\sigma_i,\sigma_{-i}}(\mathsf{win}_{i,1}) + \mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MC}\right]\mathsf{prob}_{\sigma_{-i}}(\mathsf{MC}) \\
&\quad + \mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right]\mathsf{prob}_{\sigma_{-i}}(\mathsf{MI}) + \mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{NM}\right]\mathsf{prob}_{\sigma_{-i}}(\mathsf{NM}) \\
\mathbb{E}_{\sigma'_i}[\pi_i] &= (b'_{i,1} - c_i)\mathsf{prob}_{\sigma'_i,\sigma_{-i}}(\mathsf{win}_{i,1}) + \mathbb{E}_{\sigma'_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MC}\right]\mathsf{prob}_{\sigma_{-i}}(\mathsf{MC}) \\
&\quad + \mathbb{E}_{\sigma'_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right]\mathsf{prob}_{\sigma_{-i}}(\mathsf{MI}) + \mathbb{E}_{\sigma'_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k} - c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{NM}\right]\mathsf{prob}_{\sigma_{-i}}(\mathsf{NM})
\end{aligned}$$

Equilibrium implies that under player $i$'s beliefs $\mathbb{E}_{\sigma_i}[\pi_i] \geq \mathbb{E}_{\sigma'_i}[\pi_i]$. We now establish implications of this equilibrium condition that can be taken to the data.

For all deviations, the following hold:

- Bids must decrease with the stage of the game: $b_{i,k} > b_{i,k+1}$; indeed, since the reserve price is constant, any bid submitted in period $k$ wins probability 0 in period $k+1$ if the auction continues.

- Continuation payoffs under $\sigma_i$ and $\sigma'_i$ are equal conditional on the deviation being non-marginal.

If the deviation is *an upward deviation* then,

- Player $i$'s continuation value under $\sigma_i$ is equal to zero when the deviation is marginal for continuation.

- If continuation strategies $\beta_i, \beta_{-i}$ are monotonic in observed bids, then

$$\mathbb{E}_{\sigma_i',\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right] \geq \mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right].$$

It follows from this that $\mathbb{E}_{\sigma_i}[\pi_i] \geq \mathbb{E}_{\sigma_i'}[\pi_i]$ implies

$$(b_{i,1}-c_i)\mathsf{prob}_{\sigma_i,\sigma_{-i}}(\mathsf{win}_{i,1}) \geq (b_{i,1}'-c_i)\mathsf{prob}_{\sigma_i',\sigma_{-i}}(\mathsf{win}_{i,1}). \tag{6}$$

This coincides with the IC constraint for upward deviations used in Sections 4 and 5.

If the deviation is *a downward deviation* then player $i$'s continuation value under $\sigma_i'$ is equal to zero when the deviation is marginal for continuation. Furthermore we assume that for some $\alpha \in (0,1)$

$$\mathbb{E}_{\sigma_i',\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right] \geq (1-\alpha)\mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right]. \tag{7}$$

In words, following a downward deviation that's marginal for information (meaning that the bid is in fact above the reserve price, which it would have to beat to win at a later stage) the change in the information provided in the continuation stage does not destroy all the continuation value of the bidder. Note that if at the end of each stage the auctioneer revealed an exogenous signal of the reserve price, rather than the endogenous minimum bid, then condition (7) would hold with $\alpha = 0$. In our empirical investigation, we use $\alpha = .5$.

Finally, we observe that the following bounds hold

$$\mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MI}\right] \leq \mathbb{E}[(r-c_i)^+],$$
$$\mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\middle|\mathsf{MC}\right] \leq \mathbb{E}[(r-c_i)^+].$$

Altogether, with optimality condition $\mathbb{E}_{\sigma_i}[\pi_i] \geq \mathbb{E}_{\sigma_i'}[\pi_i]$ this implies that

$$\begin{aligned}(b_{i,1}-c_i)\mathsf{prob}_{\sigma_i,\sigma_{-i}}(\mathsf{win}_{i,1}) \geq &(b_{i,1}'-c_i)\mathsf{prob}_{\sigma_i',\sigma_{-i}}(\mathsf{win}_{i,1}) \\ &- \left[\mathsf{prob}_{\sigma_{-i}}(\mathsf{MC}) + \alpha\mathsf{prob}_{\sigma_{-i}}(\mathsf{MI})\right]\mathbb{E}[(r-c_i)^+].\end{aligned} \tag{8}$$

Equations (6) and (8) replace (IC) in the inference problem defined Section 5. In addition to disciplining subjective beliefs about residual demand, expanded consistency requirement $(\widehat{CR})$ must ensure that the average subjective belief over events $\mathsf{MI}$ and $\mathsf{MC}$ must also be close to their sample probability for downward deviations. For any $\rho < 0$, denoting by $mi_h$ and $mc_h$ a player's subjective beliefs that downward deviation $\rho$ is marginal for information or continuation at $h$ we must have

$$\frac{1}{|H|}\sum_{h\in H} mi_h \in \left[\frac{1}{|H|}\sum_{h\in H} \mathbf{1}_{r<(1+\rho)b_{i,1}<\wedge \mathbf{b}_{-i,1}} \pm T\right]$$
$$\frac{1}{|H|}\sum_{h\in H} mc_h \in \left[\frac{1}{|H|}\sum_{h\in H} \mathbf{1}_{(1+\rho)b_{i,1}\leq r<\wedge \mathbf{b}_{-i,1}} \pm T\right]$$

where $T$ is a tolerance parameter chosen to ensure adequate coverage. Alternative coverage sets centered around the same sample probabilities of being marginal for information or continuation can be used.

# B Further Theoretical Results

## B.1 Connection with Bayes Correlated Equilibrium

In this section we further extend the estimator introduced Section 5 and clarify what would need to be added so that they exploit all implications from equilibrium. This allows us to connect with the work of Bergemann and Morris (2016).

For simplicity we assume that player identities $i$, bids $b$ and costs $c$ take a fixed finite number of values $(i, b, c) \in I \times B \times C$ that does not grow with sample size $|H|$. Ties between bids are resolved with uniform probability. Deviations $\rho_n \in (-1, \infty)$ correspond to the ratios of different bids on finite grid $B$.

We extend problem (P) as follows. For any environment $\omega_H$ and $(i, b, c) \in I \times B \times C$, let us define $H_{i,b,c}(\omega_H) \equiv \{h \in H | (i_h, b_h, c_h) = (i, b, c)\}$, histories at which bidder $i$ experiences a cost $c$ and bids $b$. Note that $H_{i,b,c}$ is adapted to the information of player $i$. For any tolerance function $T : \mathbb{N} \to \mathbb{R}^+$ such that

$$\lim_{k\to\infty} T(k) = 0 \quad \text{and} \quad \lim_{k\to\infty} \exp\left(-T(k)^2 k/2N_{\max}\right) = 0$$

we consider inference problem (P′)

$$\widehat{s} = \max_{\omega_H} \frac{|H_{\mathsf{comp}}(\omega_H)|}{|H|} \tag{P′}$$
$$\text{s.t. } \forall (i,b,c), \forall n, \quad D_n(\omega_H, H_{i,b,c}(\omega_H)) \in \Big[\widehat{D}(\rho_n | H_{i,b,c}(\omega_H)) - T(|H_{i,b,c}(\omega_H)|), \widehat{D}(\rho_n | H_{i,b,c}(\omega_H)) + T(|H_{i,b,c}(\omega_H)|)\Big].$$

Problem (P′) differs from (P) by imposing demand consistency requirements conditional on all triples $(i,b,c)$. Proposition 3 continues to hold with an identical proof: with probability approaching 1 as $|H|$ goes to $\infty$, $\widehat{s}$ is an upper bound to the share of competitive histories. Imposing consistency requirements conditional on bids and costs lets us establish a converse: if data passes our extended safe tests, then the joint distribution of bids and costs is an $\epsilon$-Bayes correlated equilibrium in the sense of Hart and Mas-Colell (2000).

Consider an $\omega_H$ solving (P′). Let $\widehat{\mu} \in \Delta\left([B \times C]^I\right)$ denote the sample distribution over bids and costs implied by $(H, \omega_H)$.

**Proposition B.1.** *For any $\epsilon > 0$, for $|H|$ large enough, $\widehat{s} = 1$ implies that $\widehat{\mu}$ is an $\epsilon$-Bayes correlated equilibrium.*

**Proof.** Consider beliefs and costs $(d_{n,h}, c_h)_{h \in H}$ solving Problem (P′), and $\widehat{\mu}$ the corresponding sample distribution over profiles of bids $b$ and costs $c$.

In order to deal with ties, we denote by $\wedge \mathbf{b}_{-i} \succ b_i$ the event "$\wedge \mathbf{b}_{-i} > b_i$, or $\wedge \mathbf{b}_{-i} = b_i$ and the tie is broken in favor of bidder $i$."

For $|H|$ large enough, we have that for all $(i,b,c)$ and all $n$,

$$\frac{1}{|H|} \left| \sum_{h \in H_{i,b,c}} d_{n,h} - \mathsf{prob}_{\widehat{\mu}}(\wedge \mathbf{b}_{-i} \succ (1+\rho_n) b_i | i, b, c) \right| \leq \epsilon. \tag{9}$$

In addition, $\widehat{s} = 1$ implies that (IC) holds at all histories: for all $h, n$,

$$d_{n,h}((1+\rho_n) b_h - c_h) \leq d_{h,0}(b_h - c_h).$$

Summing over histories $h \in H_{i,b,c}$ yields

$$\frac{1}{|H|} \sum_{h \in H_{i,b,c}} d_{n,h}((1+\rho_n) b_h - c_h) - d_{h,0}(b_h - c_h) \leq 0.$$

Hence for $|H|$ large enough, for all $(b_i, c_i)$,

$$\sum_{b_{-i},c_{-i}} \widehat{\mu}(b_i, c_i, b_{-i}, c_{-i}) \left( \mathbf{1}_{\wedge \mathbf{b}_{-i} \succ (1+\rho_n) b_i} ((1+\rho_n) b_i - c_i) - \mathbf{1}_{\wedge \mathbf{b}_{-i} \succ b_i} (b_i - c_i) \right) \leq \epsilon.$$

It follows that $\widehat{\mu}$ is an $\epsilon$-Bayes correlated equilibrium in the sense of Hart and Mas-Colell (2000). ■

## B.2 Complementarities between upward and downward deviations

In this appendix we clarify complementarities between downward and upward deviations and establish a possibility result in a stylized setting. Even if neither individual deviation implies that a positive share of auctions is non competitive, the joint restrictions imposed by upward and downward deviations can imply that a positive share of auctions is non competitive. For simplicity we set tolerance $T = 0$.

As we discussed in Section 6, individual upward and downward deviations respectively imply strict bounds on the share of competitive histories if and only if

$$\widehat{D}(0|H) - (1+\rho_1)\widehat{D}(\rho_1|H) < \frac{1}{1+M} \left[ \widehat{D}(0|H) - \widehat{D}(\rho_1|H) \right],$$
$$(1+\rho_{-1})\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) > \frac{1}{1+m} \left[ \widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) \right].[26]$$

To clarify the existence of complementarities between upward and downward deviations, we now consider the special case in which

$$\widehat{D}(0|H) - (1+\rho_1)\widehat{D}(\rho_1|H) = \frac{1}{1+M} \left[ \widehat{D}(0|H) - \widehat{D}(\rho_1|H) \right], \tag{10}$$
$$(1+\rho_{-1})\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) = \frac{1}{1+m} \left[ \widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) \right]. \tag{11}$$

Individual upward and downward deviations imply no restrictions on the set of competitive histories. However, different deviations are potentially rationalized by using different costs at the same history. We show this is indeed the case, and that jointly considering upward and downward deviations can yield strict constraints on the share of competitive histories.

[26] Checking whether these constrains hold can be performed rapidly, and suggests a rough rationale by which one could pick $\rho_{-1}$ and $\rho_1$: obtain a smooth estimate of the true demand, and pick $\rho_{-1}$ and $\rho_1$ so that the conditions above hold with a reliable margin.

The following lemma clarifies that markup constraints will play a role in our argument.

**Lemma B.1.** *Under* (10) *and if* $m = 0$ *and* $M = +\infty$*, then all histories can be rationalized as competitive.*

**Proof.** The following belief and costs rationalize the observed bidding behavior while satisfying consistency requirement $(\widehat{CR})$. At every history $h$ such that the bidder wins, we set $d_{h,0} = 1$, $d_{h,-1} = 1$, $d_{h,1} = \widehat{D}(\rho_1|H)/\widehat{D}(0|H)$ and $c_h = 0$.

At every history $h$ such that the bidder loses, but would win after reducing its bids by $\rho_{-1}$, we set $d_{h,0} = d_{h,1} = 0$, $d_{h,-1} = 1$ and $c_h = b_h$

At every history such that the bidder loses even after deviation $\rho_{-1}$, we set $d_{h,-1} = d_{h,0} = d_{h,1} = 0$, and $c_h = b_h$.

It is immediate that these beliefs and costs satisfy (IC), (F) and $(\widehat{CR})$. ■

We return now to the case where (10) and (11) hold for $m > 0$. Assume that there exists an environment $(\omega_H)$ satisfying (IC), (F), (EP) and $(\widehat{CR})$ with $T = 0$. We establish lower bounds for the number of histories at which $c_h/b_h$ must be equal to $\frac{1}{1+m}$ and $\frac{1}{1+M}$. Whenever these two lower bounds are mutually incompatible, the share of competitive histories is strictly less than one.

**Histories such that** $c_h/b_h = 1/(1+M)$**.** (IC) for upward deviation $\rho_1$ implies that for all histories $h$,

$$d_{h,0} - (1+\rho_1)d_{h,1} \geq (d_{h,0} - d_{h,1})\frac{c_h}{b_h}.$$

Summing over histories, conditions $(\widehat{CR})$ and (10) imply that

$$\begin{aligned}\frac{1}{|H|}\sum_{h\in H}(d_{h,0} - d_{h,1})\frac{c_h}{b_h} \leq \frac{1}{|H|}\sum_{h\in H} d_{h,0} - (1+\rho_1)d_{h,1} &= \frac{1}{1+M}(\widehat{D}(0|H) - \widehat{D}(\rho_1|H)) \\ &= \frac{1}{|H|}\sum_{h\in H}(d_{h,0} - d_{h,1})\frac{1}{1+M}.\end{aligned}$$

Since $d_{h,0} - d_{h,1} \geq 0$ and $c_h/b_h \geq \frac{1}{1+M}$, this implies that whenever $d_{h,0} - d_{h,1} > 0$, $c_h/b_h = 1/(1+M)$.

Note that if $d_{h,0} = d_{h,1} > 0$ then $d_{h,0} - (1+\rho_1)d_{h,1} < 0$ so that (IC) cannot hold. Hence $d_{h,0} - d_{h,1} = 0$ implies $d_{h,0} = 0$. This implies that

$$\frac{1}{|H|}\sum_{h\in H}\mathbf{1}_{d_{0,h}-d_{h,1}>0} \geq \frac{1}{|H|}\sum_{h\in H}\mathbf{1}_{d_{h,0}>0} \geq \frac{1}{|H|}\sum_{h\in H} d_{h,0} = \widehat{D}(0|H).$$

Hence the share of histories such that $c_h/b_h = \frac{1}{1+M}$ is at least equal to $\widehat{D}(0|H)$.

**Histories such that $c_h/b_h = 1/(1+m)$.** (IC) for downward deviation $\rho_{-1}$ implies that for all histories $h$,

$$(1+\rho_{-1})d_{h,-1} - d_{h,0} \leq (d_{h,-1} - d_{h,0})\frac{c_h}{b_h}.$$

Summing over histories, conditions $(\widehat{CR})$ and (11) imply that

$$\begin{aligned}\frac{1}{|H|}\sum_{h\in H}(d_{h,-1} - d_{h,0})\frac{c_h}{b_h} \geq \frac{1}{|H|}\sum_{h\in H}(1+\rho_{-1})d_{h,1} - d_{h,0} &= \frac{1}{1+m}(\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)) \\ &= \frac{1}{|H|}\sum_{h\in H}(d_{h,-1} - d_{h,0})\frac{1}{1+m}.\end{aligned}$$

Since $d_{h,-1} - d_{h,0} \geq 0$ and $c_h/b_h \leq \frac{1}{1+m}$, this implies that whenever $d_{h,-1} - d_{0,h} > 0$, then $c_h/b_h = 1/(1+m)$. In addition, for all $h$, we have that

$$(1+\rho_{-1})d_{h,1} - d_{h,0} = (d_{h,0} - d_{h,1})\frac{1}{1+m} \Rightarrow d_{h,0} = \frac{1+\rho_{-1} - \frac{1}{1+m}}{1 - \frac{1}{1+m}}d_{h,-1} = (1-\nu)d_{h,-1}$$

with $\nu \equiv -\rho_{-1}/(1 - \frac{1}{1+m}) > 0$. Hence, we have that

$$\begin{aligned}\sum_{h\in H} d_{h,-1} - d_{h,0} &\leq \sum_{h\in H}(d_{h,-1} - d_{h,0})\mathbf{1}_{d_{h,-1}-d_{h,0}>0} \\ &\leq \sum_{h\in H}\nu d_{h,-1}\mathbf{1}_{d_{h,-1}-d_{h,0}>0} \leq \sum_{h\in H}\nu\mathbf{1}_{d_{h,-1}-d_{h,0}>0}.\end{aligned}$$

This implies that the share of histories such that $c_h/b_h = 1/(1+m)$ is greater than $\frac{1}{\nu}(\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H))$.

Hence, if $\widehat{D}(0|H) + \frac{1}{\nu}(\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)) > 1$, then joint upward and downward deviations imply strict constraints on the share of competitive histories. For example, if $m = 3\%$, $\rho_{-1} = -1.5\%$, $\widehat{D}(\rho_{-1}|H) = 65\%$ and $\widehat{D}(0|H) = 25\%$, then $\frac{1}{\nu} \simeq 1.94$, and $\widehat{D}(0|H) + \frac{1}{\nu}(\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)) \simeq 1.025$.

## B.3 Common Values

We now show how to extend the analysis in Section 5 to allow for common values. Because expected costs conditional on winning now depend on a bidder's bid, costs and beliefs associated with history $h \in H$ now take the form $\omega_h = (d_{h,n}, c_{h,n})_{n\in\mathcal{M}}$, where for each $n \in \mathcal{M}$, $c_{h,n} = \mathbb{E}[c|h, \wedge\mathbf{b}_{-i,h} > (1+\rho_n)b_h]$ is the bidder's expected cost at history $h$ conditional on winning at bid $(1+\rho_n)b_h$.

We make the following monotonicity assumption.

**Assumption B.1.** *For all histories $h$ and all bids $b, b', b''$ with $b < b' < b''$, $\mathbb{E}[c|h, \wedge\mathbf{b}_{-i,h} \in (b, b')] \leq \mathbb{E}[c|h, \wedge\mathbf{b}_{-i,h} \in (b', b'')]$.*

In words, bidders' expected costs are increasing in opponents' bids. This implies that expected costs $c_{h,n}$ conditional on winning are weakly increasing in the deviation $\rho_n$. This condition on costs follows from affiliation when bidders' signals are one-dimensional and bidders use monotone bidding strategies. We now show that, under these conditions, allowing for common values does not relax the constraints in Program (P).

Note first that, for each deviation $n$, expected costs conditional on winning $(c_{h,n})_{n\in\mathcal{M}}$ satisfy:

$$\forall n \in \mathcal{M}, \quad d_{h,n}c_{h,n} = d_{h,0}c_{h,0} + (d_{h,n} - d_{h,0})\hat{c}_{h,n}, \tag{12}$$

where $\hat{c}_{h,n} = \mathbb{E}[c|h, \wedge\mathbf{b}_{-i,h} \in (b_h, (1+\rho_n)b_h)]$.[27] Our assumptions on costs imply that $\hat{c}_{h,n}$ is weakly increasing in $n$.

Consider first downward deviations $\rho_n < 0$ (i.e., $n < 0$). For such deviations, incentive compatibility constraint (IC) holds if and only if

$$\frac{d_{h,n}(1+\rho_n)b_h - d_{h,0}b_h}{d_{h,n} - d_{h,0}} \leq \hat{c}_{h,n}.$$

Consider next upward deviations $\rho_n > 0$ (i.e., $n > 0$). For any such deviation, constraint (IC) becomes

$$\hat{c}_{h,n} \leq \frac{d_{h,0}b_h - d_{h,n}(1+\rho_n)b_h}{d_{h,0} - d_{h,n}}.$$

Since $\hat{c}_{h,\hat{n}}$ is weakly increasing in $\hat{n}$, $\hat{c}_{h,n} \geq \hat{c}_{h,n'}$ for all $n > 0$ and $n' < 0$. Hence there exist costs $(c_{h,n})_{n\in\mathcal{M}}$ satisfying (IC) if and only if

$$\max_{n<0} \frac{d_{h,n}(1+\rho_n)b_h - d_{h,0}b_h}{d_{h,n} - d_{h,0}} \leq \min_{n>0} \frac{d_{h,0}b_h - d_{h,n}(1+\rho_n)b_h}{d_{h,0} - d_{h,n}}. \tag{13}$$

Condition (13) implies that there also exists a constant profile of costs $c_{h,n} = c_h$ (i.e. a private value cost), that satisfies (IC).

[27] We replace $(b, b')$ by $(b', b)$ in the event that $b' < b$.

## B.4 Bounds on other moments

This appendix shows how to adapt the approach of Section 5 to obtain robust bounds on other moments of interest: (i) the share of competitive auctions, and (ii) the total deviation temptation.

**Maximum share of competitive auctions.** The bound on the share of competitive histories provided by Proposition 3 allows some histories in the same auctions to have different competitive vs. non-competitive status. This may underestimate the prevalence of non-competition in a given dataset. In particular, if one player is non-competitive, she must expect other players to be non-competitive in the future. Otherwise, if all of her opponents played competitively, her stage-game best reply would be a profitable dynamic deviation.

For this reason, one might be interested in providing an upper bound on the share of competitive *auctions*, where an auction is considered to be competitive if and only if every player is competitive at their respective histories.

Take as given an adapted set of histories $H$, corresponding to a set $A$ of auctions. For every beliefs and costs $\omega_H$, let

$$A_{\mathsf{comp}}(\omega_H) \equiv \{A' \subset A \text{ s.t. } \forall a \in A', \forall h \in a, (d_h, c_h) \text{ satisfy (F), (IC) and (EP)}\}.$$

be the set of competitive auctions under $\omega_H$. Program (P) then becomes

$$\begin{aligned} \widehat{s}_{\mathsf{auc}} &= \max_{\omega_H} \frac{|A_{\mathsf{comp}}(\omega_H)|}{|A|} \\ \text{s.t. } \forall n, \quad D_n(\omega_H, H) &\in \left[\widehat{D}(\rho_n|H) - T, \widehat{D}(\rho_n|H) + T\right]. \end{aligned}$$

$\widehat{s}_{\mathsf{auc}}$ provides an upper bound to the fraction of competitive auctions.

**Total deviation temptation.** Regulators may want to investigate an industry only if firms fail to optimize in a significant way. Our methods can be used to derive a lower bound on the bidders' deviation temptation.

Given beliefs and costs $\omega_H$, define

$$U(\omega_H) \equiv \frac{1}{|H|} \sum_{h \in H} \left[(b_h - c_h)d_{h,0} - \max_{n \in \{-\underline{n}, \cdots, \overline{n}\}} [(1+\rho_n)b_h - c_h] d_{h,n}\right].$$

Our inference problem now becomes:

$$\widehat{DT} = \max_{\omega_H} U(\omega_H)$$
$$\text{s.t. } \forall n, \quad D_n(\omega_H, H) \in \left[\widehat{D}(\rho_n|H) - T, \widehat{D}(\rho_n|H) + T\right].$$

In this case, with probability approaching 1 as $|H|$ gets large, $-\widehat{DT}$ is a lower bound for the average total deviation-temptation per auction. This lets a regulator assess the extent of firms' failure to optimize before launching a costly audit. In addition, since the sum of deviation temptations must be compensated by a share of the cartel's future excess profits (along the lines of Levin (2003)), $\widehat{DT}$ provides an indirect measure of the excess profits generated by the cartel.

Figure B.1 reports estimates for firms in the city of Tsuchiura, as a function of minimum markup $m$.

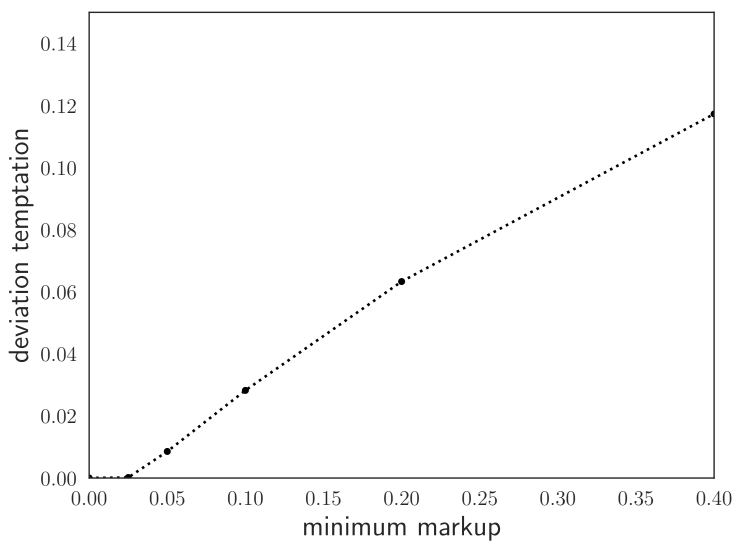



Figure B.1: Total deviation temptation as a fraction of profits, Tsuchiura. Deviations $\{-0.02, 0.0, 0.001\}$. Maximum markup 0.5.

# C Computational Strategy

In this appendix, we discuss computational implementations of the estimates of competitiveness derived in Sections 4 and 5.

## C.1 Bounds of Section 4

Column 3 of Table 2 reports 95% confidence upper bounds on the share of competitive histories for individual firms using Proposition 2:

$$s_{\mathsf{comp}} \leq 1 - \sup_{\rho > 0} \frac{\overline{R}(\rho|H) - \overline{R}(0|H)}{\rho}.$$

We take $\rho$ equal to $\{0.01\%, 0.02\%, \cdots, 0.3\%\}$ of the reserve price and fix this set in our asymptotics. We estimate the counterfactual revenue using a triangular kernel with bandwidth equal to 0.01% of the reserve price. We use the central limit theorem for Martingale difference sequences to obtain the 95% confidence bound, see e.g., Liu and Yang (2008).

## C.2 Bounds of Section 5

Problem (P) is not naturally suited for computational implementation. Recalling that $\mathcal{M}$ denotes the set of possible deviations, a solution $\omega_H$ of (P) is a real vector of dimension $|H| \times (|\mathcal{M}| + 1)$, and the objective function is not concave. For this reason we study a natural relaxation of (P). Observe that $\frac{|H_{\mathsf{comp}}|}{|H|} = \mathbb{E}_{\mu(\omega_H)}[v(\omega_h)]$ where $\mu_{\omega_H} \in \Delta([0,1]^{|\mathcal{M}|+1})$ is the sample distribution of environments $\omega_h$ induced by profile $\omega_H$, and $v(\omega_h) = \mathbf{1}_{\omega_h \text{ satisfies (IC),(F),(EP)}}$.[28] Hence, it follows that

$$\begin{aligned} \widehat{s} \leq & \max_{\mu \in \Delta([0,1]^{|\mathcal{M}|+1})} \mathbb{E}_{\mu} v(\omega_h) \\ & \text{under the constraint that} \\ & \forall n,\ \mathbb{E}_{\mu} d_{n,h} \in [\widehat{D}(\rho_n|H) - T, \widehat{D}(\rho_n|H) + T]. \end{aligned}$$

The convexified right-hand side problem is linear, and the dimensionality of a solution is no longer related to sample size $|H|$. Note that the constraint on expected demand could be replace by any convex coverage set. The difficulty is that, $\Delta([0,1]^{|\mathcal{M}|+1})$ is infinite dimensional. Still, because $[0,1]^{|\mathcal{M}|+1}$ is compact and finite dimensional it is covered by finitely many balls of radius $r$ for any $r > 0$. Hence for all $\epsilon > 0$, there exists a finite set $O \subset [0,1]^{|\mathcal{M}|+1}$ such

[28] We rexpress cost $c_h$ as a share $c_h/b_h$ of bids.

that

$$\widehat{s} - \epsilon \leq \max_{\mu \in \Delta(O)} \mathbb{E}_\mu v(\omega_h) \tag{CVX-P}$$

$$\text{under the constraint that}$$

$$\forall n,\ \mathbb{E}_\mu d_{n,h} \in [\widehat{D}(\rho_n|H) - T, \widehat{D}(\rho_n|H) + T].$$

The right-hand-side problem is now a well behaved finite dimensional linear problem. Practically we use the following parallelized algorithm.

1. Draw $K$ samples of $L$ points in $[0,1]^{|\mathcal{M}|+1}$ using a full-support distribution.
2. For each sample $k \leq K$ of $L$ points (generating a set $O_k$) compute the solution $\mu_k^*$ to the associated (CVX-P) problem. Let $\underline{O}_k$ denote the support of $\mu_k^*$, truncated to cover $1 - \nu$ of the mass under $\mu^* k$
3. Set $O = \cup_k \in \{1, \cdots, K\} \underline{O}_k$, and solve the associated (CVX-P) problem.
4. Assess convergence by comparing solution to that obtained starting from different random seeds.

Practically, we set $K = 400$, $L = 5000$ and $\nu = 1\%$.

**Sensitivity to economic plausibility constraints.** Figure C.1 shows that, for our city-level data, our estimates on the share of competitive histories are insensitive to changes in maximum markup $M$. Figure C.2 illustrates the sensitivity of our estimates to parameter $\alpha \in [0,1]$ in downward deviation IC constraint (8) for auctions with re-bidding. Recall that parameter $\alpha$ measures the extent to which a deviation by a firm in round 1 affects her continuation profits in the following rounds when the deviation changes the information bidders have in the following rounds.

# D Proofs

**Proof of Lemma 1.** Let $H$ be an adapted set of histories, and fix $\rho \in (-1, \infty)$. Bidding data from each auction is progressively revealed over time. Assume for simplicity that one

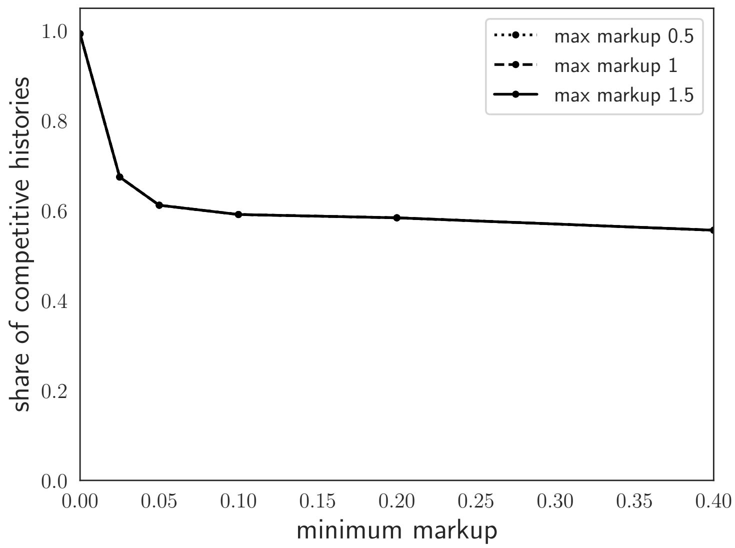



Figure C.1: Share of competitive histories for different maximum markups, city data, deviations $\{-0.02, 0, 0.001\}$.

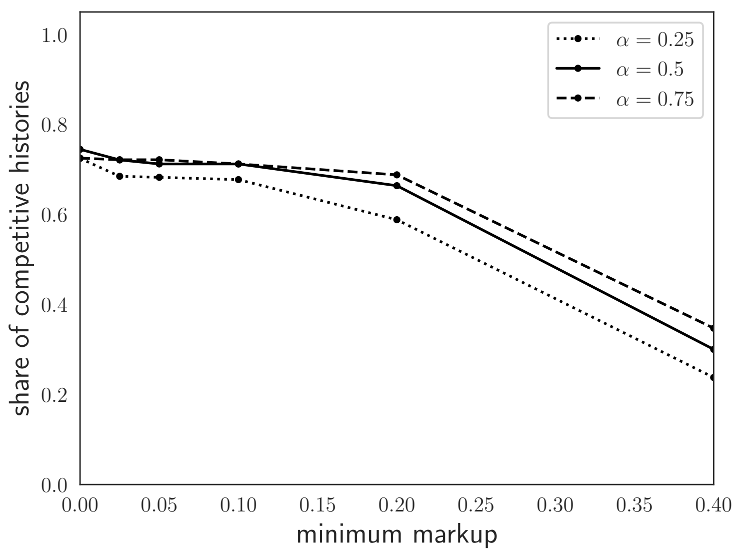



Figure C.2: Share of competitive histories, national-level data. Deviations $\{-0.02, 0, 0.001\}$, $M = 0.5$.

auction happens at each time $t \in \{1, \cdots, |A|\}$ so that we can associate each auction with a single time $t$. For each time $t$, define

$$\varepsilon_t \equiv \sum_{h_{i,t} \in H} \mathbb{E}_\sigma[\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b_{i,t}(1+\rho)} | h_{i,t}] - \mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b_{i,t}(1+\rho)}.$$

Note that $\widehat{D}(\rho|H) - \overline{D}(\rho|H) = \frac{1}{|H|}\sum_{t=1}^{|A|}\varepsilon_t$.

Note further that, by the law of iterated expectations, for all public histories $h^0_{t-s} \in H$ with $s \geq 0$, $\mathbb{E}_\sigma[\varepsilon_t|h^0_{t-s}] = 0$.[29] Hence, $S_T \equiv \sum_{t=1}^{T}\varepsilon_t$ is a Martingale, with increments $\varepsilon_{\hat{t}}$ whose absolute value is bounded above by $\overline{N}_t$, the number of bidders participating at time $t$ (with histories in $H$). By the Azuma-Hoeffding Inequality, for every $\nu > 0$,

$$\mathsf{prob}(|S_{|H|}| \geq \nu|H|) \leq 2\exp\left(\frac{-\nu^2|H|^2}{2\sum_{t=1}^{|A|}\overline{N}_t^2}\right).$$

Observing that $\sum_{t=1}^{|A|}\overline{N}_t^2 \leq \sum_{t=1}^{|A|}\overline{N}_t N_{\max} = N_{\max}|H|$, this implies that

$$\mathsf{prob}(|S_{|H|}| \geq |H|x) \leq 2\exp\left(-\nu^2|H|/2N_{\max}\right).$$

This concludes the proof. ■

**Proof of Proposition 3.** By Lemma 1, we have that

$$\mathsf{prob}\left(|\widehat{D}(\rho_n|H) - D_n(\omega_H, H)| \geq T\right) \leq 2\exp(-T^2|H|/2N_{\max})$$

for each deviation $n$, with $\omega_H$ denoting the true environment. It then follows that

$$\mathsf{prob}\left(\forall n, |\widehat{D}(\rho_n|H) - D_n(\omega_H, H)| \geq T\right) \leq 2|\mathcal{M}|\exp(-T^2|H|/2N_{\max}).$$

This implies that, with probability at least $1 - 2|\mathcal{M}|\exp(-T^2|H|/2N_{\max})$, the constraints in Program (P) are satisfied when we set the environment equal to the true environment $\omega_H$. Hence, with probability at least $1 - 2|\mathcal{M}|\exp(-T^2|H|/2N_{\max})$, $\hat{s}$ is weakly larger than the share of competitive histories $s_{\mathsf{comp}}$ under the true environment $\omega_H$. ■

# References

Afriat, S. N. (1967): "The construction of utility functions from expenditure data," *International economic review*, 8, 67–77.

[29] This holds since, in a perfect public Bayesian equilibrium, bidders' strategies at any time $t$ depend solely on the public history and on their private information at time $t$.

Ashenfelter, O. and D. Sullivan (1987): "Nonparametric tests of market structure: an application to the cigarette industry," *The Journal of Industrial Economics*, 483–498.

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

Athey, S. and K. Bagwell (2001): "Optimal collusion with private information," *RAND Journal of Economics*, 32, 428–465.

——— (2008): "Collusion with persistent cost shocks," *Econometrica*, 76, 493–540.

Bajari, P., S. Houghton, and S. Tadelis (2014): "Bidding for incomplete contracts: An empirical analysis of adaptation costs," *American Economic Review*, 104, 1288–1319.

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

Bergemann, D. and S. Morris (2016): "Bayes correlated equilibrium and the comparison of information structures in games," *Theoretical Economics*, 11, 487–522.

Carvajal, A., R. Deb, J. Fenske, and J. K.-H. Quah (2013): "Revealed preference tests of the Cournot model," *Econometrica*, 81, 2351–2379.

Chambers, C. P. and F. Echenique (2016): *Revealed preference theory*, vol. 56, Cambridge University Press.

Chassang, S., K. Kawai, J. Nakabayashi, and J. M. Ortner (2019): "Data Driven Regulation: Theory and Application to Missing Bids," Tech. rep., National Bureau of Economic Research.

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.

Clark, R. and J.-F. Houde (2013): "Collusion with asymmetric retailers: Evidence from a gasoline price-fixing case," *American Economic Journal: Microeconomics*, 5, 97–123.

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

Cyrenne, P. (1999): "On antitrust enforcement and the deterrence of collusive behaviour," *Review of Industrial Organization*, 14, 257–272.

DOJ, U. (2005): "Price fixing, bid rigging, and market allocation schemes: What they are and what to look for. An Antitrust Primer." 2005.

Echenique, F., S. Lee, and M. Shum (2011): "The money pump as a measure of revealed preference violations," *Journal of Political Economy*, 119, 1201–1223.

Ellison, G. (1994): "Theories of cartel stability and the joint executive committee," *The Rand journal of economics*, 37–57.

Fershtman, C. and A. Pakes (2012): "Dynamic games with asymmetric information: A framework for empirical work," *The Quarterly Journal of Economics*, 127, 1611–1661.

Green, E. J. and R. H. Porter (1984): "Noncooperative collusion under imperfect price information," *Econometrica: Journal of the Econometric Society*, 87–100.

Harrington, J. E. (2004): "Cartel pricing dynamics in the presence of an antitrust authority," *RAND Journal of Economics*, 651–673.

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

Hart, S. and A. Mas-Colell (2000): "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, 68, 1127–1150.

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

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

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

Levin, J. (2003): "Relational incentive contracts," *The American Economic Review*, 93, 835–857.

Liu, R. and L. Yang (2008): "Kernel estimation of multivariate cumulative distribution function," *Journal of Nonparametric Statistics*, 20, 661–677.

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

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

Ortner, J., S. Chassang, K. Kawai, and J. Nakabayashi (2019): "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. H. (1983): "A study of cartel stability: the Joint Executive Committee, 1880-1886," *The Bell Journal of Economics*, 301–314.

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.

Rotemberg, J. J. and G. Saloner (1986): "A supergame-theoretic model of price wars during booms," *The American Economic Review*, 390–407.

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

Sullivan, D. (1985): "Testing hypotheses about firm behavior in the cigarette industry," *Journal of Political Economy*, 93, 586–598.

Tóth, B., M. Fazekas, Á. Czibik, and I. J. Tóth (2014): "Toolkit for detecting collusive bidding in public procurement, with examples from Hungary," .

Varian, H. R. (1990): "Goodness-of-fit in optimizing models," *Journal of Econometrics*, 46, 125–140.