Robust Screens for Non-Competitive Bidding in Procurement Auctions

Sylvain Chassang, Princeton University
Kei Kawai, U.C. Berkeley
Jun Nakabayashi, Kindai University
Juan Ortner*†, Boston University

June 22, 2021

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: its extreme forms are inconsistent with competitive behavior under arbitrary information structures. Building on this observation, we develop systematic tests of competitive behavior in procurement auctions that allow for general information structures as well as non-stationary unobserved heterogeneity. We provide an empirical exploration of our tests, and show they can help identify other suspicious patterns in the data.

KEYWORDS: procurement, competition, antitrust, missing bids, robustness.

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

†We are especially indebted to Hiroaki Kaido and Steve Tadelis for detailed feedback. The paper benefited from discussions with Pierpaolo Battigali, Eric Budish, Yeon-Koo Che, Francesco Decarolis, Rieko Ishii, Emir Kamenica, Ulrich Müller, Roger Myerson, Ariel Pakes, Wolfgang Pesendorfer, Andrea Prat, Michael Riordan, Jozsef Sakovics, Larry Samuelson, Andy Skrzypacz, Paulo Somaini, as well as comments from participants at several seminars and conferences. The study has been supported by JSPS KAKENHI (Grant Number JP18H03643)

1 Introduction

One of the key functions of antitrust authorities is to detect and punish collusive agreements. Although concrete evidence is required for successful prosecution, screening devices that flag suspicious firms help regulators identify such collusive agreements, and encourage members of existing cartels to apply for leniency programs.[1] Correspondingly, an active research agenda has sought to build methods to detect collusion 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 competitive behavior in procurement auctions under weak assumptions on the environment.

We begin by documenting a suspicious bidding pattern observed in first-price sealed-bid 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), Switzerland (Imhof et al., 2018), and Canada (Clark et al., 2020). We establish that extreme forms of this pattern are inconsistent with competitive behavior under a general class of asymmetric information structures. Indeed, when winning bids are isolated, bidders can profitably deviate by increasing their bids. Expanding on this observation, we propose general tests of competitive behavior in procurement auctions that are robust, in the sense of holding under weak assumptions on the information structure and arbitrary unobserved heterogeneity.

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 Ibaraki prefecture and the Tohoku region of Japan.

[1]A growing number of agencies are adopting algorithm-based screens that analyze bidding data from public procurement auctions to flag suspicious behavior including those in Brazil, South Korea, Switzerland, the United Kingdom, and most recently, the United States. For example, the U.S. Department of Justice announced the formation of a procurement collusion strike force whose goals include bolstering "data analytics employment to identify signs of potential anticompetitive, criminal collusion." (Announcement of the Antitrust Division's Procurement Collusion Strike Force, November 22, 2019.)

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' margin of victory and defeat. For every (bidder, auction) pair, we compute $\Delta \equiv \frac{\text{own bid} - \text{min(other bids)}}{\text{reserve}}$, 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. For both the municipal and national datasets, we document a missing mass in the distribution of Δ around $\Delta = 0$. Our results clarify the sense in which this missing mass of close losing bids is suspicious, and help us identify other patterns in the data that are inconsistent with competition.

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 correlated within and across periods. Importantly, we rule out dynamic considerations such as capacity constraints or learning by doing: we assume that current auction outcomes don't affect firms' future costs. 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 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 portions of our data: because winning bids are isolated, the elasticity of sample residual demand is close to zero.

Our second set of results generalizes this test. In particular, we show how to exploit equilibrium conditions to derive bounds on the extent of non-competitive behavior in our data. The bounds that we propose allow for very general information structures, contrasting with existing approaches that rely on specific assumptions such as independent private values (e.g. Bajari and Ye, 2003). As we show in our companion paper Ortner et al. (2020), antitrust policy based on tests that are robust to information structure cannot be exploited by cartels to enhance collusion. This addresses the concern articulated by Cyrenne (1999) and Harrington (2004) that data driven antitrust policies may end up facilitating collusion by providing cartel members a more effective threat point.

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 moderate drop in bids (on the order of 2%) may also be attractive to bidders: it yields large increases in demand. In our data, downward deviations tend to be more informative about the competitiveness of auctions than upward deviations. In addition, upward and downward deviations are more informative together than separately. Finally, although failing our tests does not necessarily imply bidder collusion, we show that the outcomes of our tests are consistent with other proxy evidence for competitiveness and collusion. Bids that are high relative to the reserve price are more likely to fail our tests than bids that are low. Bids placed before an industry is investigated for collusion are more likely to fail our tests than bids placed after it is investigated for collusion. Altogether this suggests that, although our tests are conservative, they still have bite in practice.

Our paper relates primarily to the literature on cartel detection in auctions.[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. 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) focus on auctions with re-bidding, and exploit correlation

[2]See Harrington (2008) for a recent survey.

in bids across stages to detect collusion.[3] Marmer et al. (2016) and Schurter (2017) design tests of collusion for English auctions and for first-price sealed bid auctions focusing on partial cartels. The tests that we propose relax assumptions imposed in previous work such as symmetry, independence and private values (at the cost of reduced power), and can be used to detect both all-inclusive cartels and partial cartels.

More broadly, our paper relates to prior work that seeks to test for competitive behavior in other (non auction) markets. Sullivan (1985) and Ashenfelter and Sullivan (1987) propose tests of whether firms behave as a perfect cartel, and apply these tests to the cigarette industry. Bresnahan (1987) and Nevo (2001) test for competition in the automobile and ready-to-eat cereal industries.[4]

Finally, our tests are also related to revealed preference tests seeking to quantify violations of choice theoretic axioms.[5] 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.

2 Motivating Facts

Our first dataset consists of roughly 7,100 auctions for public works contracts held between 2004 and 2018 by municipalities located in the Tohoku region and Ibaraki prefecture of Japan. The auctions are sealed-bid first-price auctions with a publicly announced reserve price.[6] The top panel of Table 1 reports summary statistics. The mean reserve price is 23.2

[3] Kawai and Nakabayashi (2018) study the subset of auctions with rebidding while the current analysis applies to all auctions.

[4] Also related are Porter (1983) and Ellison (1994), who exploit dynamic patterns of play predicted by the theory of repeated games (Green and Porter, 1984, Rotemberg and Saloner, 1986) to test for tacit collusion.

[5] See Chambers and Echenique (2016) for a recent review of the literature on revealed preferences.

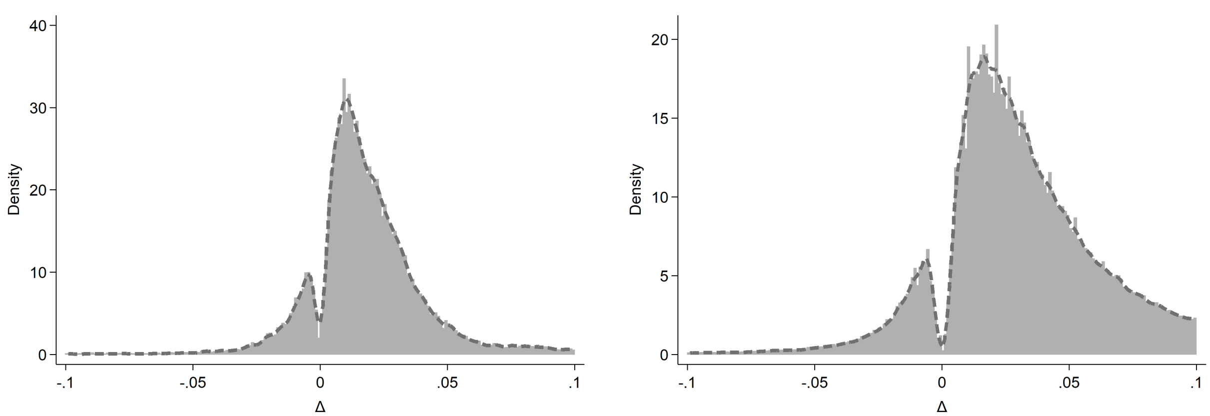
[6] Our city-level dataset combines two datasets. The first dataset contains auctions held by municipalities in the Tohoku region in Japan. For the current analysis, we restrict attention to municipalities using a sealed-bid first-price auction with a public reserve price. The second dataset, studied 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

million yen, or about 230,000 USD, and the mean winning bid is 21.5 million yen. The mean number of bidders is 7.4. On average, a bidder in the dataset participates in 23.3 auctions and wins 3.1 auctions.

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

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

aggregated over firms i, and auctions a, where $\wedge \mathbf{b}_{-i,a}$ denotes the minimum bid from bidders other than i. 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 histogram and the density estimate of bid differences Δ aggregating over all firms and auctions in the sample. The mass of missing bids around 0 is clearly noticeable.



(a) city auctions

(b) national auctions

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

The dotted curves correspond to local (6th order) polynomial density estimates with bandwidth set to 0.0075.

Our second dataset, studied in Kawai and Nakabayashi (2018), consists of roughly 78,000 auctions for construction projects held between 2001 and 2006 by the Ministry of Land,

auctions with a public reserve price.

Infrastructure and Transportation in Japan (the Ministry). The auctions are sealed-bid first-price auctions with a *secret* reserve price. The average reserve price is 105.1 million yen, or about 1 million USD and the mean lowest bid is 101.9 million yen, which is 97.0% of the reserve price. Because the reserve price is secret, the lowest bid may be higher than the reserve price in which case there is rebidding. In that event, the reserve price remains secret to the bidders, but the lowest bid from the initial round is announced. There are at most two rounds of rebidding. If none of the bids are below the reserve price at the end of the second round of rebidding, the lowest bidder from the last round enters into a bilateral negotiation with the buyer. The auction concludes in the initial round of bidding about 75% of the time. The auction concludes after one round of rebidding in more than 97% of auctions and concludes after two rounds of rebidding in more than 99% of auctions. The mean number of participants is 9.9. For both datasets, all bids become public information after the auction. Figure 1(b) illustrates the distribution of bid-differences Δ for national auctions, where Δ is defined using first-round bids. The missing mass of bids around $\Delta = 0$ is stark.

We point out another important (though less visually striking) feature of the densities plotted in Figure 1: the tails of the distribution taper off rapidly. This implies that much of the mass of Δ is concentrated within a relatively small interval around 0. For example, Δ lies between 0 and 0.02 for 50.0% of the losing bids in the city auctions and 25.6% of the losing bids in the national auctions. This implies that a drop in bids of 2% increases demand considerably (by 349% and 307%, respectively).[7] Hence, while the missing bid pattern in Figure 1 suggests small increases in bids are profitable, the relatively large concentration of mass around $\Delta = 0$ suggests that a small reduction in bids is also attractive (unless firms' profit margins are very small).

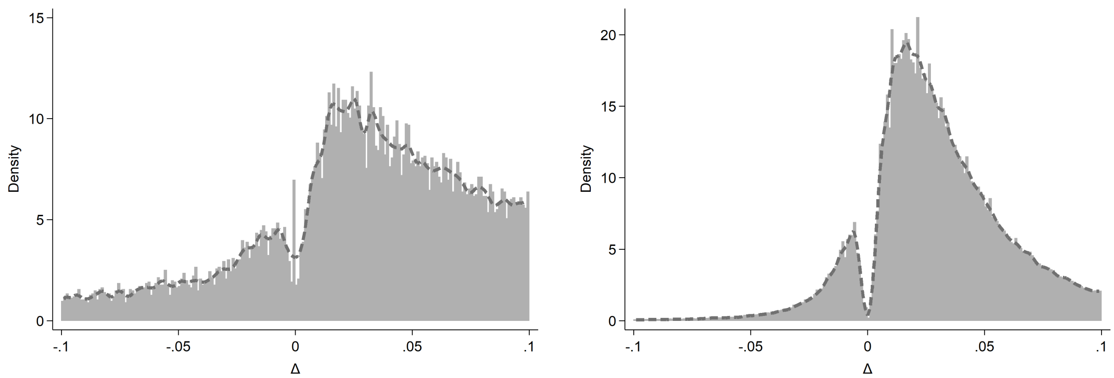
[7]Figure OC.1 (Online Appendix OC.1) illustrates this point by plotting firms' sample residual demand for our two datasets.

		Mean	S.D.	N
City Auctions				
By Auctions				
	reserve price (mil. Yen)	23.189	91.32	7,111
	lowest bid (mil. Yen)	21.500	85.10	7,111
	lowest bid / reserve	0.938	0.06	7,111
	#bidders	7.425	3.77	7,111
By Bidders				
	participation	23.29	43.58	2,267
	number of times lowest bidder	3.14	6.22	2,267
National Auctions				
By Auctions				
	reserve price (mil. Yen)	105.121	259.58	78,272
	lowest initial bid (mil. Yen)	101.909	252.30	78,272
	winning bid (mil. Yen)	100.338	252.30	78,272
	lowest bid / reserve	0.970	0.10	78,272
	winning bid / reserve	0.946	0.10	78,272
	ends in one round of bidding	0.752	0.43	78,272
	ends after one rebidding	0.971	0.17	78,272
	ends after two rebidding	0.996	0.06	78,272
	#bidders	9.883	2.27	78,272
By Bidders				
	participation	26.40	94.61	29,670
	number of times lowest bidder	2.64	10.57	29,670

Table 1: Sample Statistics – City and National Level Data

Correlation with indicators of collusion. Our analysis studies the extent to which the bidding patterns in Figure 1 are inconsistent with competitive behavior under arbitrary information structures. While non-competitive behavior need not be collusive, we note that missing bids are in fact correlated with plausible indicators of collusion.

Since the goal of collusion is to elevate prices, we would expect to see suspicious bidding patterns in auctions with high bids. Figure 2 breaks down the auctions in Figure 1(b) by bid level: the figure plots the distribution of $\Delta_{i,a}$ for normalized bids $\frac{b_{i,a}}{r}$ below .8 (Panel (a)) and above .9 (Panel (b)). The mass of missing bids is considerably reduced in Panel (a). The tails of the distribution taper off more gradually in Panel (a).



(a) Bids below 80% of reserve price

(b) Bids above 90% of reserve price

Figure 2: Distribution of bid-difference Δ – national data.

The dotted curves correspond to local (6th order) polynomial density estimates with bandwidth set to 0.0075.

Figure 3 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 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). The left panels in Figure 3 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 case (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 (2018) for a more detailed account of these collusion cases).

What does not explain this pattern. We end this section by arguing that missing bids are not explained by either the granularity of bids, or ex post renegotiation.

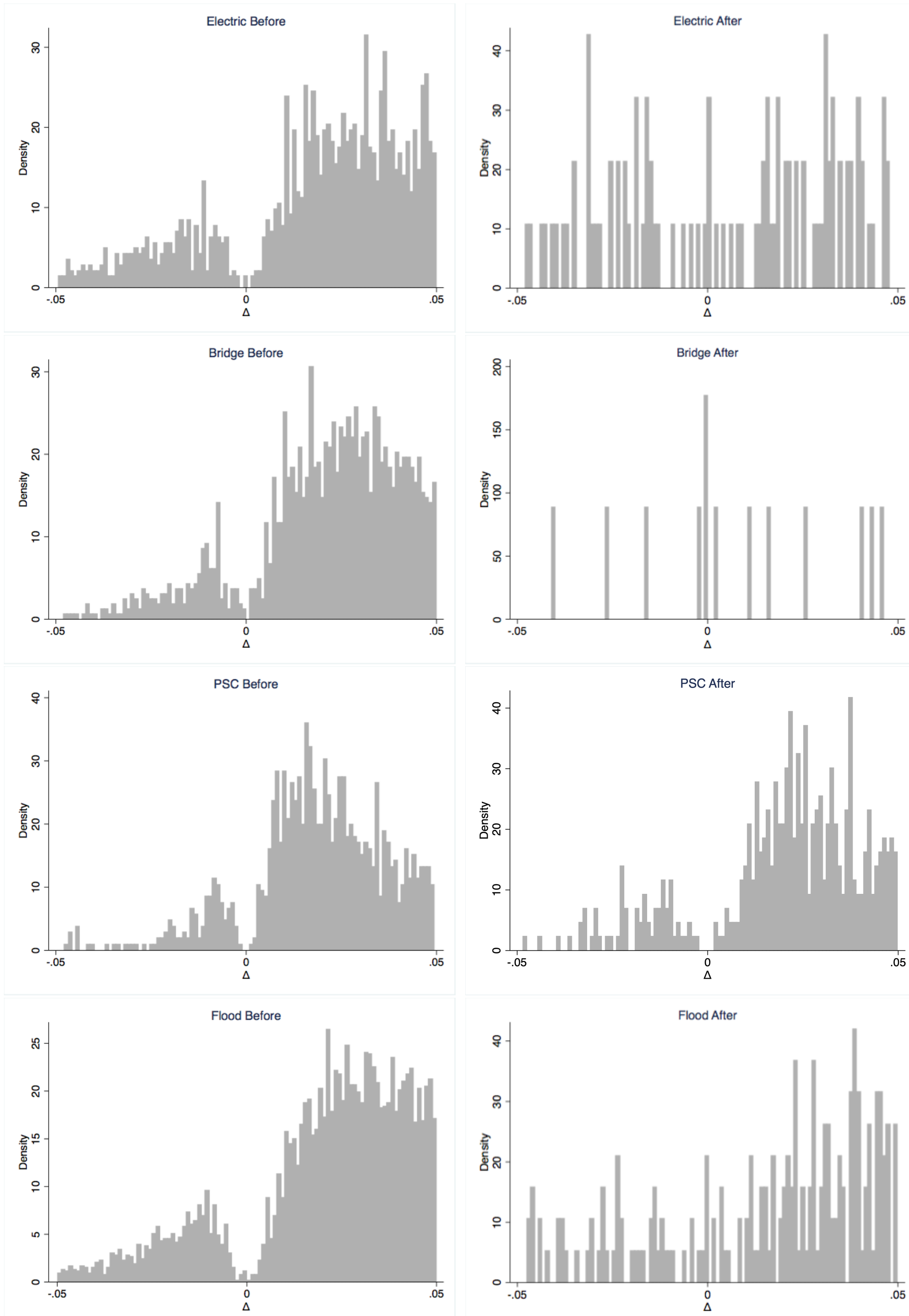


Figure 3: Distribution of bid-difference Δ, before and after JFTC investigation.

Figures 2 and 3 show that the pattern of missing bids in Figure 1 is not a mechanical consequence of the granularity of bids. If this was the case, we should see similar patterns across all bid levels, or before and after the JFTC investigations. In addition, Figure OC.2 in Online Appendix OC.1 plots the distribution of placebo statistic Δ^2, defined as the difference between bids and the most competitive other bid in bidding data from which each auction's lowest bid is excluded. The figure shows that the distribution of Δ^2 has no corresponding missing mass at 0.

Renegotiation could potentially account for missing bids by making apparent incentive compatibility issues irrelevant. However, in the auctions we study, contracts signed between the awarder and the awardee include a renegotiation provision which stipulates that renegotiated prices should be anchored to the initial bid. Specifically, if the project is deemed to require additional work, the government engineers estimate the costs associated with extra work. The firm is then paid $\frac{\text{initial bid}}{\text{reserve price}} \times$ costs of extra work. This implies that a missing mass of Δ makes a small increase in bids profitable even taking into account the possibility of renegotiation. We provide further evidence on this point in Online Appendix OC.1: the missing bid pattern is present even in a subset of auctions for which we can ascertain that no price renegotiation took place.

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 a *public* reserve price r, which we normalize to $r = 1$. Online Appendix OA extends the analysis to auctions with *secret* reserve prices and re-bidding (as in the national data).

In each period t, a state $\theta_t \in \Theta$ captures all relevant past information about the environment. Some elements of θ_t may be observed by the bidders at the time of bidding, but other elements may not be.[8] All elements of θ_t are revealed to the bidders by the end of period t. Importantly, θ_t need not be observed by the econometrician. We assume that θ_t is an exogenous 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, or ergodic. The important assumption here is that by the end of each period t, bidders observe a sufficient statistic θ_t of future environments. Since θ_t evolves as an exogenous Markov process, we rule out intertemporal linkages between actions and payoffs.[9] In practice, state θ_t may include the vector of distances between the project site and each of the firms, the vector of inputs specified in the construction plan, or the vector of current input prices. It may also include the current calendar date if there is seasonality in firms' costs. Lastly, θ_t may also include variables that are unknown to the bidders at the time of bidding, such as underground soil conditions. We do not assume that θ_t is observed by the econometrician.

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.[10] The distribution of the set of eligible bidders $\widehat{N}_t$ can vary over time, but depends only on state θ_{t-1}. Participants discount future payoffs using a common discount factor $\delta < 1$.

Costs. Costs of production for eligible bidders $i \in \widehat{N}_t$ are denoted by $\mathbf{c}_t = (c_{i,t})_{i \in \widehat{N}_t}$. The bidder may or may not know its own costs at the time of bidding. 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 θ_t. Because θ_t need not be finite-valued or ergodic, the distribution of

[8] As we discuss below, we assume that each bidder receives a (possibly multi-dimensional) signal $z_{i,t}$ prior to bidding. Components of θ_t that are observed to the bidder will be included in $z_{i,t}$. For components of θ_t that are not observed, the bidder will form beliefs over their values based on the signal.

[9] This rules out dynamic considerations such as capacity constraints or learning by doing. In our companion paper, Kawai et al. (2020), we propose tests of competition that hold in the presence of intertemporal links between bids and costs.

[10] 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.

costs can be arbitrarily different at every t. For example, if time t is included as part of state θ_t, then the distribution of costs can be different for every auction. All costs are assumed to be positive.

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 $(\theta_t, (c_{i,t})_{i\in\widehat{N}_t})$. Signals $z_{i,t}$ can take arbitrary values, including vectors in $\mathbb{R}^n$. Signals $z_{i,t}$ may reveal information about current state θ_t, bidder i's own costs $c_{i,t}$, or the costs $c_{j,t}$ of other players. This allows our model to nest many informational environments, including private and common values, correlated values, asymmetric bidders and asymmetric information.[11] Since θ_t may not be finite valued or ergodic, our framework allows the distribution of signals $(z_{i,t})_{i\in\widehat{N}_t}$ to be different in every period t.

Bids and payoffs. 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. Bids $\mathbf{b}_t$ are publicly observed at the end of the auction.[12]

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.

[11] For instance, our model encompasses the affiliated values model, with the distribution of costs and signals allowed to depend on state θ_t.

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

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 θ_t is revealed by 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 $i \in N$, σ_i maps public histories and payoff-relevant private signals to bids

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

Because our framework doesn't allow for intertemporal linkages between past actions and future payoffs, we can identify the class of competitive equilibria with the class of Markov perfect equilibria (Maskin and Tirole, 2001).

Definition 1 (competitive strategy)**.** *We say that σ 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 σ is a competitive equilibrium if it is a perfect public Bayesian equilibrium in Markov perfect strategies.

In a competitive equilibrium, firms must be playing a stage-game Nash equilibrium in every period; i.e. firms play a static best-reply to the actions of their opponents.

Competitive histories. 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: we allow for both full and partial cartels. This leads us to define competitiveness at the history level.

Definition 2 (competitive histories)**.** *Fix a common knowledge profile of strategies σ 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 σ_{-i}.*

We build up to our main inference problem, described in Section 6, in two steps. First, we show in Section 4 that even under general information structures, it is possible to use data to place restrictions on bidders' beliefs at the time of bidding. Second, we show in Section 5 that competitive behavior has testable implications, even under general incomplete information: the standard result that firms bid in the elastic part of the demand curve continues to hold for averages of demand. The results from Section 4 and Section 5 can be combined to construct a simple test of competition based on the idea that missing bids patterns are inconsistent with competition. Section 6 expands on these insights to obtain a probabilistic upper bound on the maximum share of histories consistent with competitive behavior.

4 Data-Driven Restrictions on Beliefs

Because we make few assumptions on the environment, it is not obvious that bidding data can be used to test for competition. This section shows that as long as bidders play a perfect public Bayesian equilibrium, we can still place probabilistic constraints on the players' beliefs about their probability of winning.[13] We show that in equilibrium the difference between realized demand (i.e., $\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b}$, for different values of b) and bidders' beliefs regarding demand is a martingale. Versions of the central limit theorem applying to sums of martingale increments imply that as the sample size grows, sample averages of demand must be close to the historical average of bidders' beliefs, even if those beliefs vary in a non-stationary way across histories.

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

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

[13] In this section, we do not assume that bidders play Markov perfect strategies. We only assume perfect public Bayesian equilibrium. Weaker solution concepts that do not require common knowledge of equilibrium, such as no-regret play (Hart and Mas-Colell, 2000) would lead to similar results.

In words, bidder i's residual demand at history $h_{i,t}$ represents the bidders' beliefs regarding the probability of winning auction t for each possible bid b' she may place. Note that the conditioning variable $h_{i,t}$ in the expression includes the bidder's signal $z_{i,t}$ on which we have made very few assumptions. In particular, the distribution from which the signals $z_{i,t}$ are drawn can be different for each period t unlike in the conditional i.i.d. setting used in previous work.[14] Indeed, bidders' beliefs can depend on public information and private signals that are unobserved to the econometrician. Still, it is possible to consistently estimate appropriately constructed averages of beliefs.

Take as given a finite set of histories H, and a scalar $\rho \in (-1, \infty)$. We denote by $\overline{D}(\rho|H)$ the average residual demand for histories in H and by $\widehat{D}(\rho|H)$ its sample equivalent:

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

$$\widehat{D}(\rho|H) \equiv \frac{1}{|H|} \sum_{h_{i,t} \in H} \mathbf{1}_{\wedge \mathbf{b}_{-i,t} > (1+\rho)b_{i,t}} \tag{2}$$

where $|H|$ denotes the cardinality of set H.

We now provide conditions on the set of histories H under which expression (2) consistently estimates (1).

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.*

A subset H can be thought of as a selection of histories that satisfy certain criteria defined by the analyst. Definition 3 states that H is adapted if it is possible to check whether $h_{i,t}$ satisfies the criteria needed for inclusion in H using only information available to bidder i at time t, prior to bidding.

[14]Note that in conditional i.i.d. environments with no unobserved heterogeneity, $D_i(b'|h_{i,t})$ is consistently estimated by the empirical distribution of $\wedge \mathbf{b}_{-i,t}$ conditional on observables. In our setting, we cannot consistently estimate demand $D_i(b'|h_{i,t})$ at each $h_{i,t}$. We will only be able to obtain consistent estimates for averages of demand.

Consider, for example, taking H to be the entire set of histories for a specific industry or location. The criteria for inclusion is that a particular history is for a specific industry or location. In this case, H is adapted because a bidder knows at the time of auction t that auction t is for a given industry or location. Hence, we do not need any information that the bidder does not know at the time of bidding to determine whether or not to include any such history in H.

Similarly, the set of histories in which a bidder bids a particular value is adapted since the bidder knows how she bids. In contrast, the set of histories in which a specific bidder wins the auction, or the set of histories in which the winning bid is equal to some value are not adapted.[15]

It is necessary for us to focus on adapted sets of histories to link realized demand and beliefs regarding demand (and payoffs). When we select a subset of histories H for data analysis, we are effectively evaluating outcomes under the conditioning event that $h_{i,t} \in H$. When this event is in the information set of bidders at the time of bidding, then even conditional on this event, the differences between the bidders' beliefs and realizations of demand are zero, in expectation. This allows us to link realized outcomes to bidders' expectations, and obtain consistent estimates of the bidders' beliefs regarding demand and payoffs using realized data. This link disappears if we focus on a set of histories that is not adapted: the realized outcomes may be systematically different from the bidder's expectation at the time of bidding. For instance, if we focus on the set of histories such that a given bidder wins, then the bidder's realized demand under this conditioning event is 1, whereas the bidder's expected demand at the time of bidding will likely have been strictly below 1.

The notion of adpatedness formalizes the conditions on the sample selection rule such that sample averages consistently estimate bidder beliefs without introducing selection bias. In our empirical application, we take the set H to be histories in which the bids are above or below a particular value, histories in which a given firm bids, etc. In each of these

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

applications, the fact that we select the set of histories to be adapted guarantees that sample demand consistently estimates averages of bidders' beliefs.

Let $N_{\max}$ denote an upper bound on the number of participants in any auction.[16]

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

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

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

In equilibrium, the sample residual demand conditional on an adapted set of histories converges to the historical average of bidders' beliefs about demand. This implies a confidence set $[\widehat{D}(\rho|H) - \nu, \widehat{D}(\rho|H) + \nu]$ for the unobserved historical average of beliefs $\overline{D}(\rho|H)$. We note that for simplicity, Proposition 1 uses non-asymptotic concentration results, and symmetric two-sided confidence sets. We revisit these choices with power in mind in Section 6, after introducing our main inference problem.[17]

What drives Proposition 1 and when can it fail? The key step of the proof of Proposition 1 is the observation that for all periods t, coefficients ρ, and bidding histories $(\mathbf{b}_s)_{s<t}$

$$\mathbb{E}_\sigma[\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > (1+\rho)b_{i,t}} - \mathsf{prob}_\sigma(\wedge \mathbf{b}_{-i,t} > (1+\rho)b_{i,t}|h_{i,t}) \mid (\mathbf{b}_s)_{s<t},\ h_{i,t} \in H] = 0. \tag{3}$$

In words, bidders' conditional beliefs about demand are correct in expectation, conditional on past bidding information $(\mathbf{b}_s)_{s<t}$ and $h_{i,t} \in H$, i.e., the fact that history $h_{i,t}$ satisfies the criteria for inclusion in H. This implies that the difference $|H| \times \left(\widehat{D}(\rho|H) - \overline{D}(\rho|H)\right)$

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

[17] Among other things, we use tighter asymptotic bounds relying on the central limit theorem for sums of martingale increments (Billingsley, 1995).

evolves like a martingale as $|H|$ grows. The Azuma-Hoeffding Inequality yields Proposition 1.

Condition (3) clarifies the role that public information about past bids, and adapted histories play in the proof of Proposition 1. Note that bidders are required to have correct beliefs in equilibrium, which implies

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

If the history of bids $(\mathbf{b}_s)_{s<t}$ and the event $\{h_{i,t} \in H\}$ are both known to the bidder at the time of bidding, the law of iterated expectations implies (3). Condition (3) would fail if set H was not adapted (in that case, being included in our analysis carries information unavailable to bidders at the time of bidding), or if bidders did not obtain feedback about past outcomes (in that case, the history of bids $(\mathbf{b}_s)_{s<t}$ carries information unavailable to bidders at the time of bidding). For instance, imagine that bidders other than i experience a persistent drop in costs, so that they start bidding more aggressively than what firm i expected in period $t = 0$. If firm i gets no feedback about either bids or auction outcomes, it cannot update its beliefs and will systematically overestimate its demand at periods $t > 0$. Similarly, if we used a non-adapted set H, such as the set of bidding histories where bidder i wins, then (3) need not hold. Conditional on $h_{i,t} \in H$, the bidder wins the auction with probability 1, while bidder i's belief at the time of bidding can be strictly less than 1.

Condition (3) also clarifies the role of rational expectations. Outside expectation $\mathbb{E}_\sigma$ and inside probability prob_σ are indexed on the same stochastic process for bids generated by equilibrium σ. If bidders expected others' bids to be generated according to a strategy profile σ' while actual bids are generated by strategy profile $\sigma \neq \sigma'$, then (3) need not hold. We note however that because past bids are observable, there exist prior-free solution concepts based on no-regret learning rules (Hart and Mas-Colell, 2000) which guarantee that a version of Proposition 1 would hold even without the rational expectations assumption.

5 Missing Bids are Inconsistent with Competition

Our first main result shows that *extreme forms* of the pattern of bids illustrated in Figure 1 are inconsistent with competitive behavior.

Proposition 2. *Let σ be a competitive equilibrium. Then,*

$$\forall h_i, \forall b' > \sigma_i(h_i) = b, \quad \frac{\log D_i(b'|h_i) - \log D_i(b|h_i)}{\log b' - \log b} \leq -1, \tag{4}$$

$$\forall H, \forall \rho > 0, \quad \frac{\log \overline{D}(\rho|H) - \log \overline{D}(0|H)}{\log(1+\rho)} \leq -1. \tag{5}$$

Under any competitive equilibrium, the elasticity of a bidder's residual demand must be less than -1 at every history, and the inequality aggregates to sets of histories. Proposition 2 extends to first-price auctions with *secret* reserve prices (see Online Appendix OA).

Proof. Consider a competitive equilibrium σ. Let

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

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 $b' > b$,

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

Since σ 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,

$$\begin{aligned} bD_i(b|h_{i,t}) - b'D_i(b'|h_{i,t}) &= \mathbb{E}_\sigma \left[b\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b} - b'\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b'} \middle| h_{i,t} \right] \\ &\geq \mathbb{E}_\sigma \left[c_{i,t} (\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b} - \mathbf{1}_{\wedge \mathbf{b}_{-i,t} > b'}) \middle| h_{i,t} \right] \geq 0, \end{aligned} \tag{6}$$

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 (5) follows from a similar argument: for all $h_{i,t}$, $b_{i,t}$ and $\rho > 0$, we have that

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

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

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

■

Proposition 2 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. Extreme forms of missing bids contradict Proposition 2: when the density of Δ 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.[18] 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.

Because Proposition 1 allows us to obtain estimates (and confidence sets) of $\overline{D}(0|H)$ and $\overline{D}(\rho|H)$, Propositions 1 and 2 together yield a simple test of whether or not an adapted set of histories H can be generated by a competitive equilibrium. This test holds under weak restrictions on the environment, and hence strengthens existing approaches that make specific assumptions such as symmetry, independent values and private values (see for instance Bajari and Ye, 2003).

[18]Dynamic considerations such as learning by doing might imply that firms incur a negative net cost of winning an auction. As we argue in Section 8, such dynamic considerations are unlikely to explain the bidding patterns in our data.

6 Bounding the Share of Competitive Histories

Proposition 2 derives testable implications of competition by using only the restriction that costs are non-negative, and incentive compatibility conditions with respect to a single deviation: an increase in bids. In this section we show how to obtain an upper bound on the share of competitive histories (or equivalently a lower bound on the share of non-competitive histories) consistent with observed data. For this we exploit the information content of both upward and downward deviations, as well as possible restrictions on costs taking the form of markup constraints. Proposition 2 can be viewed as a special case of the results we present below. We also present asymptotic confidence sets that offer better power than Proposition 1. For simplicity, we assume that costs are private, and treat the case of common-value costs in Online Appendix OB.

We believe that estimating the share of non-competitive histories, rather than just offering a binary test of competition, is practically important. Cartels are often partial, and regulators may want to prioritize more egregious cases. Measuring the prevalence of non-competitive behavior can help regulators gauge the magnitude of potential cartels and target investigations efficiently. This finer measure can also be used to track changes in cartel behavior over time. Finally, establishing that failures of non-competitive behavior are not rare clarifies that bidders have plenty of opportunities to learn how to improve their bids. This suggests that failures to optimize stage-game profits are not merely errors.

6.1 Deviations, Beliefs and Constraints

We begin by describing bidders' beliefs and the constraints they must satisfy: incentive compatibility constraints for competitive histories, markup constraints imposed by the analyst, and consistency with empirical demand (along the lines of Proposition 1).

Deviations and beliefs. Take as given scalars $\rho_n \in (-1, \infty)$ indexed by $n \in \mathcal{M} = \{-\underline{n}, \cdots, \overline{n}\} \subset \mathbb{Z}$, such that $\rho_0 = 0$ and $\rho_n < \rho_{n'}$ for all $n' > n$. Each scalar ρ_n parameterizes deviations $(1+\rho_n)b_{i,t}$ from equilibrium bids $b_{i,t}$.

For a history $h_{i,t}$ with equilibrium bid $b_{i,t}$, deviation $n \in \mathcal{M}$ is associated with an expected demand $d_{h_{i,t},n}$ corresponding to bidder i's belief that she would win the auction conditional on history $h_{i,t}$ and bid $(1+\rho_n)b_{i,t}$:

$$d_{h_{i,t},n} \equiv D_i((1+\rho_n)b_{h_{i,t}}|h_{i,t}).$$

The profile of beliefs $\mathbf{d}_{h_{i,t}} \equiv (d_{h_{i,t},n})_{n\in\mathcal{M}}$ associated with deviations $\mathcal{M}$ is a vector in $[0,1]^{\mathcal{M}}$. The set of *feasible* beliefs $\mathcal{F}$ is defined by $\mathcal{F} \equiv \{\mathbf{d} \in [0,1]^{\mathcal{M}} \text{ s.t. } \forall n < n',\ d_n \geq d_{n'}\}$.

Given a set of histories H, consider the associated collection of bidders' beliefs $(\mathbf{d}_h)_{h\in H}$ at those histories. It will be useful to consider the historical distribution μ^* of feasible beliefs $\mathbf{d} \in \mathcal{F}$ induced by $(\mathbf{d}_h)_{h\in H}$:

$$\forall \mathbf{d} \in \mathcal{F}, \quad \mu^*(\mathbf{d}) = \frac{1}{|H|}\sum_{h\in H} \mathbf{1}_{\mathbf{d}_h=\mathbf{d}}.$$

Note that μ^* is simply the (discrete) empirical measure defined on $\mathcal{F}$ which puts mass $1/|H|$ on realized beliefs and mass 0 on all other points. Distribution μ^* will be used to express averages of the history of beliefs $(\mathbf{d}_h)_{h\in H}$ as expectations under distribution μ^*. For instance, the average expected demand $(\overline{D}(\rho_n|H))_{n\in\mathcal{M}}$ at different deviations can be expressed as $\mathbb{E}_{\mu^*}[\mathbf{d}]$. Indeed, for each $n \in \mathcal{M}$,

$$\mathbb{E}_{\mu^*}[d_n] \equiv \frac{1}{|H|}\sum_{h_{i,t}\in H} d_{h_{i,t},n} = \overline{D}(\rho_n|H).$$

We note that distribution μ^* is itself a random variable, depending on the realization of bidders' beliefs at the time of bidding. It is not observed by the econometrician.

Markup constraints. We allow the analyst or econometrician to impose constraints on costs c_h taking the form of markup constraints:

$$\frac{b_h}{c_h} \in [1+m, 1+M] \tag{MKP}$$

where $m \geq 0$ and $M \in (m, +\infty]$ are minimum and maximum markups.[19] Constraint (MKP) provides what we think is a transparent and convenient way for regulators to express intuitive subjective restrictions over the environment. Markups are simple familiar objects that analysts may have direct intuition or information about. In contrast, Bayesian priors over the underlying environment are high-dimensional objects for which little direct information is available.[20,21] We discuss the impact of (MKP) on empirical inference in Section 7.

Incentive compatibility constraints. When a history is competitive, the associated beliefs must be consistent with stage-game best response (Definition 2). Hence, every competitive history $h \in H$ is associated with feasible beliefs $(d_{h,n})_{n \in \mathcal{M}} \in \mathcal{F}$ satisfying

$$\exists c_h \geq 0 \text{ s.t. } \forall n \in \mathcal{M}, \quad [(1+\rho_n)b_h - c_h]\, d_{h,n} \leq [(1+\rho_0)b_h - c_h] d_{h,0}. \tag{IC}$$

In words, beliefs can be rationalized as competitive if there exists a cost for which all deviations are suboptimal given beliefs. Condition (IC) can be rexpressed more conveniently as

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

[20] 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%". Our existing framework lets us test whether at least 90% of histories can be deemed competitive when markups are greater than 5%. Alternatively, we could accommodate probabilistic restrictions by specifying that (MKP) must hold for a minimum share of histories.

[21] Constraint (MKP) places restrictions on the set of environments that we allow for. For instance, it rules out settings in which all bidders share the same publicly observed cost. We note, however, that the minimum markup constraint in (MKP) can be microfounded with a model in which bidders incur a small bid preparation cost $\kappa > 0$ if they participate in an auction. In such a model, bidder i will only participate at time t and bid $b_{i,t}$ if $D_i(b_{i,t}|h_{i,t})(b_{i,t} - c_{i,t}) \geq \kappa$. Since $D_i(b_{i,t}|h_{i,t}) \leq 1$, if bidder i chooses to participate, it must be that $\frac{b_{i,t}}{c_{i,t}} \geq 1 + \frac{\kappa}{c_{i,t}} \geq 1 + \kappa$. Previous papers (e.g., Krasnokutskaya and Seim, 2011) have estimated that bid preparation costs can be substantial, between 2.2% and 3.9% of the engineer's cost estimate.

follows: for all $n \in \mathcal{M}$,

$$\begin{aligned}
&[(1+\rho_n)b_h - c_h]\, d_{h,n} \leq [(1+\rho_0)b_h - c_h]d_{h,0} \\
&\qquad \iff (1+\rho_n)d_{h,n} - (1+\rho_0)d_{h,0} \leq \frac{c_h}{b_h}(d_{h,n} - d_{h,0}). \qquad (7)
\end{aligned}$$

Since $d_{h,n} \geq d_{h,n'}$ whenever $n < n'$, it follows that (7) is equivalent to[22]

$$\frac{(1+\rho_n)d_{h,n} - (1+\rho_0)d_{h,0}}{d_{h,n} - d_{h,0}} \leq \frac{c_h}{b_h} \quad \text{if } n < 0 \qquad (8)$$

$$\frac{(1+\rho_n)d_{h,n} - (1+\rho_0)d_{h,0}}{d_{h,n} - d_{h,0}} \geq \frac{c_h}{b_h} \quad \text{if } n > 0. \qquad (9)$$

Conditions (8) and (9) imply that there exists a cost c_h satisfying (MKP) and (IC) if and only if beliefs $\mathbf{d}_h$ satisfy

$$\begin{aligned}
&\max\left\{\frac{1}{1+M}, \frac{(1+\rho_n)d_{h,n} - (1+\rho_0)d_{h,0}}{d_{h,n} - d_{h,0}}\Big| n < 0\right\} \qquad \text{(IC-MKP)} \\
&\qquad \leq \min\left\{\frac{1}{1+m}, \frac{(1+\rho_n)d_{h,n} - (1+\rho_0)d_{h,0}}{d_{h,n} - d_{h,0}}\Big| n > 0\right\}.
\end{aligned}$$

Empirical consistency constraints. The third constraint that beliefs must satisfy is that they must be consistent with data. Although we cannot pin down bidders' beliefs at individual histories, Proposition 1 implies probabilistic restrictions on average beliefs. With probability close to 1 as the number of histories gets large, the average expected demand at different deviations $\mathbb{E}_{\mu^*}[\mathbf{d}] = (\overline{D}(\rho_n|H))_{n\in\mathcal{M}}$ must be close to its empirical counterpart $\widehat{\mathbf{D}} \equiv (\widehat{D}(\rho_n|H))_{n\in\mathcal{M}}$, if H is adapted.

Formally, for any $\alpha \in [0,1]$ we take as given a confidence set $\mathcal{D}_\alpha$ that covers the realized average expected demand $\mathbb{E}_{\mu^*}[\mathbf{d}]$ with probability α. Set $\mathcal{D}_\alpha$ depends on realized bidding data alone, and satisfies

$$\mathsf{prob}_\sigma(\mathbb{E}_{\mu^*}[\mathbf{d}] \in \mathcal{D}_\alpha) \geq \alpha.$$

[22]In the event that $d_{h,n} - d_{h,0} = 0$, we set the left-hand side of (8) equal to $+\infty$, $-\infty$ or 0, depending on whether the numerator $(1+\rho_n)d_{h,n} - (1+\rho_0)d_{h,0}$ is positive, negative or equal to zero.

Proposition 1 implies that for any $\alpha > 0$, there exist confidence sets $\mathcal{D}_\alpha$ converging to the singleton set given by the historical mean beliefs $\{\mathbb{E}_{\mu^*}[\mathbf{d}]\}$ when the sample size gets sufficiently large. For finite data however, the choice of confidence set $\mathcal{D}_\alpha$ will impact inference: some restrictions matter more than others to obtain upper bounds on the share of competitive histories. We describe the confidence sets $\mathcal{D}_\alpha$ used in our empirical implementation in Section 6.3 after presenting our main inference result.

6.2 Inferring a Bound on Competitive Histories

Given a distribution of historical beliefs μ^*, the share of histories s_{comp} in H that can be rationalized as competitive can be expressed as follows:

$$s_{\mathsf{comp}} = \mathbb{E}_{\mu^*}[\text{IsComp}(\mathbf{d})] \quad \text{with} \quad \text{IsComp}(\mathbf{d}) = \begin{cases} 1 & \text{if } \mathbf{d} \text{ satisfies (IC-MKP)} \\ 0 & \text{otherwise} \end{cases}.$$

While s_{comp} is not observable, we can define a probabilistic upper bound $\widehat{s}_{\mathsf{comp}}$:

$$\widehat{s}_{\mathsf{comp}} = \max_{\mu \in \Delta(\mathcal{F})} \mathbb{E}_\mu[\text{IsComp}(\mathbf{d})] \tag{P}$$

$$\text{s.t. } \mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha. \tag{$\widehat{CR}$}$$

Program (P) maximizes the share of histories that can be rationalized as competitive over distributions of historical beliefs $\mu \in \Delta(\mathcal{F})$ consistent with empirical demand, i.e. distributions μ such that $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$.

Proposition 3. *With probability greater than α, $\widehat{s}_{\mathsf{comp}} \geq s_{\mathsf{comp}}$.*

For any threshold $s_0 \in [0, 1]$, consider the test $\tau \equiv \mathbf{1}_{\widehat{s}_{\mathsf{comp}} < s_0}$. Under the null that $s_{\mathsf{comp}} \geq s_0$, test τ rejects the null with probability less than $1 - \alpha$.

We remark that when we consider a single upward deviation and take $M = +\infty$, the test $\tau \equiv \mathbf{1}_{\widehat{s}_{\mathsf{comp}} < 1}$ reduces to a test of whether or not the elasticity of sample demand is higher than -1. Indeed, with a single upward deviation $\rho_1 > 0$ and $M = +\infty$, (IC-MKP) reduces to $(1+\rho_1)d_{h,1} \leq d_{h,0}$. Hence, if $(1+\rho_1)\widehat{D}(\rho_1|H) > \widehat{D}(0|H)$, (IC-MKP) and $(\widehat{CR})$ cannot hold simultaneously for all histories $h \in H$ whenever confidence set $\mathcal{D}_\alpha$ is a sufficiently small interval containing $\widehat{\mathbf{D}} = (\widehat{D}(\rho_n|H))_{n\in\mathcal{M}}$. In this sense, Proposition 2 can be viewed a special case of Proposition 3.[23]

The proof of Proposition 3 follows the logic of calibrated projection (Kaido et al., 2019): a confidence set for an underlying parameter — here the true average expected demand $\mathbb{E}_{\mu^*}[\mathbf{d}]$ — implies a confidence set for a function of the parameter — here, the share of competitive histories. Statistic $\widehat{s}_{\mathsf{comp}}$ is the upper-bound of a confidence interval for underlying true parameter s_{comp}.

The space of distributions over feasible beliefs $\mu \in \Delta(\mathcal{F})$ is infinite dimensional, which makes program (P) computationally intractable. However, for each finite set $\mathcal{F}_0 \subset \mathcal{F}$, we can consider the following program:

$$\max_{\mu\in\Delta(\mathcal{F}_0)} \mathbb{E}_\mu[\text{IsComp}(\mathbf{d})] \qquad \text{(Approx-P)}$$
$$\text{s.t. } \mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$$

Program (Approx-P) is now a well behaved finite dimensional linear program in $\mu \in \Delta(\mathcal{F}_0)$.[24] Our next result shows that, provided set $\mathcal{F}_0$ is sufficiently dense within $\mathcal{F}$, the solution to (Approx-P) is approximately equal to $\widehat{s}_{\mathsf{comp}}$.

Consider a sequence $(\mathcal{F}_0^n)_{n\in\mathbb{N}}$ of finite subsets of $\mathcal{F}$, such that $\mathcal{F}_0^n$ becomes dense in $\mathcal{F}$ as n grows large. Denote by $\widehat{s}^n_{\mathsf{comp}}$ the solution to the approximate problem (Approx-P) associated with $\mathcal{F}_0^n$. The following result holds:

Lemma 1. $\lim_{n\to\infty} \widehat{s}^n_{comp} = \widehat{s}_{comp}$.

[23]For further details, see Appendix OB.2, where we solve problem (P) in closed form for the case of a single downward or upward deviation.

[24]If $\mathcal{F}_0 = \{\mathbf{d}_1, \cdots \mathbf{d}_K\} \subset [0,1]^{\mathcal{M}}$, then $\mu = (\mu_k)_{k=1}^K \in \Delta(\mathcal{F}_0)$ is simply a $K \times 1$ vector of non-negative numbers that sum to 1. The objective function is $\sum_{k=1}^K \mu_k \text{IsComp}(\mathbf{d}_k)$ and the constraint is $\sum_{k=1}^K \mu_k \mathbf{d}_k \in \mathcal{D}_\alpha$. Note that both the objective function and the constraint are linear in μ.

The proof of Lemma 1 is provided in Online Appendix OD.

A natural question is whether $\widehat{s}_{\mathsf{comp}}$ is the tightest possible bound on the share of competitive histories. It isn't. Program (P) exploits restrictions on beliefs imposed by (IC-MKP) and ($\widehat{CR}$). We show in Appendix OB that when sample size is arbitrarily large, we would be exploiting all of 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). In practice, we find that this stretches both the limits of our data (in finite samples confidence levels drop as we seek to cover many conditional demands), and of bidder sophistication. Relying on a weaker set of optimality conditions makes our estimates more robust to partial failures of optimization, consistent with the critique of Fershtman and Pakes (2012).

6.3 Confidence Sets

In our empirical application, we consider confidence sets $\mathcal{D}_\alpha$ of feasible beliefs $\mathbf{d} \in \mathcal{F}$, taking the form of convex sets centered around empirical demand vector $\widehat{\mathbf{D}} \equiv (\widehat{D}(\rho_n|H))_{n\in\mathcal{M}}$:

$$\mathcal{D}_\alpha = \left\{\mathbf{d} \in \mathcal{F} \text{ s.t. } \forall \lambda \in \Lambda, \left\langle \lambda, \mathbf{d} - \widehat{\mathbf{D}} \right\rangle \leq x_\lambda \right\} \tag{10}$$

where: $\Lambda \subset \mathbb{R}^{\mathcal{M}}$ is a finite set of real vectors λ; parameters $x_\lambda \geq 0$ are given positive thresholds; and $\langle \cdot, \cdot \rangle$ denotes the usual scalar product between vectors. Note that λ is a vector that is chosen by the researcher. For example, if we take λ to be a standard basis vector of the form $(0, \ldots, 0, 1, 0, \ldots, 0)$, with the n^{th} coordinate equal to 1, the corresponding inequality defines an upper bound on the n-th element of $\mathbf{d}$ as $d_n \leq \widehat{D}(\rho_n|H) + x_\lambda$.

The Boole-Frechet inequality implies that

$$\begin{aligned} \mathsf{prob}(\mathbb{E}_{\mu^*}[\mathbf{d}] \in \mathcal{D}_\alpha) &= \mathsf{prob}\left(\forall \lambda \in \Lambda,\ \left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle \leq x_\lambda \right) \\ &\geq 1 - \sum_{\lambda \in \Lambda} \mathsf{prob}\left(\left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle > x_\lambda \right). \end{aligned} \tag{11}$$

This implies that to obtain coverage rates for $\mathcal{D}_\alpha$, it is sufficient to provide upper bounds for $\mathsf{prob}\left(\left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle > x_\lambda\right)$. As in Proposition 1, a non-asymptotic bound holds.

Lemma 2 (non-asymptotic coverage)**.** *For any adapted set of histories H, any vector $\lambda = (\lambda_n)_{n\in\mathcal{M}} \in \mathbb{R}^{\mathcal{M}}$, and any threshold $x_\lambda > 0$,*

$$\mathsf{prob}\left(\left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle > x_\lambda\right) \leq \exp\left(-\frac{x_\lambda^2 |H|}{2||\lambda||_1^2 N_{\max}}\right), \tag{12}$$

where $||\lambda||_1 \equiv \sum_{n\in\mathcal{M}} |\lambda_n|$.

Together with (11), Lemma 2 allows us to construct confidence sets $\mathcal{D}_\alpha$ that have appropriate coverage for average expected demand $\mathbb{E}_{\mu^*}[\mathbf{d}]$. Note that we can set $|\mathcal{M}| = 1$ and $\lambda = -1, 1$ in expression (12) to obtain Proposition 1. The proof of Lemma 2 is very similar to that of Proposition 1. As in the case of Proposition 1, taking the set H to be adapted plays a crucial role.[25]

The attraction of bound (12) is that it is non-asymptotic. Unfortunately it is also very conservative. For this reason we provide less conservative asymptotic bounds relying on a central limit theorem for renormalized sums of martingale increments (see Billingsley (1995), Theorem 35.11).

Assume H includes histories corresponding to auctions taking place at $\{0, 1, ..., T\}$, so that $|H|$ grows with T. For any history $h_{i,t} \in H$, let $\widehat{\mathbf{d}}_{h_{i,t}} \equiv (\mathbf{1}_{(1+\rho_n)b_{i,t} \leq \wedge \mathbf{b}_{-i,t}})_{n\in\mathcal{M}}$ denote the empirical demand at history $h_{i,t}$. Recall that $\mathbf{d}_{h_{i,t}} \equiv (d_{h_{i,t},n})_{n\in\mathcal{M}}$ denotes bidder i's actual expected demand at history $h_{i,t}$. We assume that $\lambda \neq 0$ and

$$\lim_{T\to+\infty} \sum_{t=0}^{T} \mathsf{var}\left(\sum_{h_{i,t}\in H} \left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \middle| h_t^0 \right) = +\infty.$$

[25] Vectors $\lambda \in \Lambda$ allow us to obtain bounds for demand (and for payoffs) following different deviations. For instance, when we consider a single downward deviation, we use $\Lambda = \{(-1, 0), (-1, 1), (0, 1)\}$. Hence, set $\mathcal{D}_\alpha$ includes beliefs $\mathbf{d} = (d_{-1}, d_0)$ satisfying $d_{-1} \geq \widehat{D}(\rho_{-1}|H) - x_{\lambda_1}$, $d_{-1} - d_0 \geq D(\rho_{-1}|H) - D(\rho_0|H) - x_{\lambda_2}$ and $d_0 \leq D(\rho_{-1}|H) + x_{\lambda_3}$. The key observation is that we apply negative coefficients to demand following a deviation, and positive coefficients to equilibrium demand. This implies a lower bound for demand following a deviation, and an upper bound for equilibrium demand.

In other words, bidders remain uncertain about their demand in the long run. This ensures that a central limit theorem for renormalized sums of martingale increments holds.

Let $H_t \equiv \{h_s \in H, s < t\}$ denote the set of bidding histories in H occurring prior to period t. Let $\widehat{\mathbf{D}}_t \equiv \frac{1}{|H_t|}\sum_{h \in H_t} \widehat{\mathbf{d}}_h$ denote the average empirical demand given bidding data available at time t. For any λ, we define

$$\widehat{\sigma}_\lambda \equiv \sqrt{\frac{1}{T+1}\sum_{t=0}^{T} |\{h_{i,t} \in H\}| \sum_{h_{i,t} \in H} \left\langle \lambda, \widehat{\mathbf{D}}_t - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle^2}.$$

Lemma 3 (asymptotic coverage)**.** *For any adapted set of histories H, any vector $\lambda = (\lambda_n)_{n \in \mathcal{M}} \in \mathbb{R}^{\mathcal{M}}$, and any threshold $x_\lambda > 0$,*

$$\limsup_{T \to +\infty} \textit{prob}\left(\frac{|H|}{\widehat{\sigma}_\lambda \sqrt{T+1}} \left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle > x_\lambda \right) \leq \Phi(-x_\lambda)$$

where Φ is the c.d.f. of the standard normal $\mathcal{N}(0,1)$.

Along with condition (11), this yields the coverage rates for sets $\mathcal{D}_\alpha$ that we use in our empirical investigation. Proofs for Lemmas 2 and 3 are provided in Online Appendix OD.

7 Empirical Findings

In this section, we estimate upper bounds on the share of competitive histories for the city-level and national-level auctions described in Section 2. Before reporting our findings we discuss a few points regarding implementation and computation. We also give a brief discussion of how we address rebidding.

Implementation and computation. We use confidence sets $\mathcal{D}_\alpha$ of the form described by (10) and compute confidence levels using Lemma 3. In our application, we consider the bound $\widehat{s}_{\mathsf{comp}}$ corresponding to several different choices of deviations $\mathcal{M}$. When we consider a single downward, or a single upward deviation, we use $\Lambda = \{(-1,0),(-1,1),(0,1)\}$ and

$\Lambda = \{(1,0),(1,-1),(0,-1)\}$, respectively. The key observation is that we apply negative coefficients to demand following deviations and positive coefficients to equilibrium demand. This implies lower bounds for demand (and therefore payoffs) after deviations, and upper bounds for demand (and therefore payoffs) in equilibrium.[26] We pick thresholds x_λ to ensure a 98.33% confidence interval for each dot product $\left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle$, resulting in a 95% confidence level for $\mathcal{D}_\alpha$.

When we consider both a downward, and an upward deviation, we set

$$\Lambda = \{(-1,0,0),(-1,1,0),(0,1,0),(0,1,-1),(0,0,-1)\}.$$

For this case, we pick thresholds x_λ to ensure a 99% confidence interval for each dot product to obtain a 95% confidence level for $\mathcal{D}_\alpha$.

In practice, we solve problem (Approx-P) using the following parallelized algorithm for the case of $|\mathcal{M}| = 3$.

1. Draw 200 samples of 1000 tuples in $[0,1]^3$ using a seeded uniform distribution and sort each tuple in decreasing order.[27] Each sample of 1000 points in $[0,1]^3$ corresponds to a finite subset $\mathcal{F}_j \in \mathcal{F}$ $(j = 1, \cdots 200)$ of feasible beliefs.

2. For each sample $\mathcal{F}_j$, compute the solution μ_j^* to the associated problem (Approx-P). The solution μ_j^* is a 1000×1 vector of non-negative numbers that sum to 1. Let $\underline{\mathcal{F}}_j$ denote the support of μ_j^*, truncated to cover 99% of the mass under μ_j^* by dropping belief vectors $\mathbf{d} \in \mathcal{F}_j$ in order of ascending probability $\mu_j^*(\mathbf{d})$.[28]

3. Set $\mathcal{F}_0 = \cup_{j\in\{1,\cdots,200\}} \underline{\mathcal{F}}_j$, and solve the associated (Approx-P) problem.

[26] Bounds on differences between demands help exploit correlation between demand at different bids. The set of directions Λ was chosen to exhaust the set of bounds on individual demands, and differences between demand at deviations and in equilibrium, but does not seek to optimize weights.

[27] The uniform distribution here does not matter except for ensuring full support. The reason for sorting each tuple in decreasing order is because demand is decreasing, i.e., $d_n \geq d_{n'}$ if $n < n'$

[28] The support of optimal distributions tends to include much less than 1000 points, i.e., most of the elements in μ_j^* are 0. This approach allows us to collect many good guesses using parallelized computations on relatively small samples, and then compute the solution associated with the union of these goods guesses.

4. Assess convergence by comparing the solution to that obtained starting from different random seeds up to 10^{-4}.

Rebidding. City-level auctions use public reservation prices, and the results of Section 5 and 6 apply directly. National-level auctions use secret reserve prices, and dealing with rebidding requires theoretical adjustments. As we explain in Online Appendix OA, incentive compatibility constraints for bid increases are essentially unchanged. For bid reductions, we need to assess losses in continuation values when (1) there is rebidding, (2) the bid reduction changes the lowest bid reported to bidders, thereby affecting their continuation information.[29] 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%.[30]

7.1 A Case Study

We first illustrate the mechanics of inference using bidding data from the city of Tsuchiura, located in Ibaraki prefecture. We do so by pooling bidding histories associated with auctions held prior to October 2009. Note that this set of histories is adapted. 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 $\rho \in \{-.02, 0, .0008\}$, where by convention, $\rho = 0$ is an infinitesimal downward deviation that amounts to breaking ties. The distribution of Δ and the deviations we consider (in dashed lines) are illustrated in Figure 4. The deviations are selected to deliver crisp illustrative results. We are specifically interested

[29] Recall that the lowest bid in each round of bidding is announced to all of the bidders before the next round of bidding.

[30] Note that the reported bid is above the fixed reserve price which bidders must beat to win the auction. If the reported bid was a function of r alone, then there would be no loss in continuation value associated with a change in bidders information conditional on rebidding.

in illustrating the empirical content of individual deviations, as well as complementarities between upward and downward deviations.

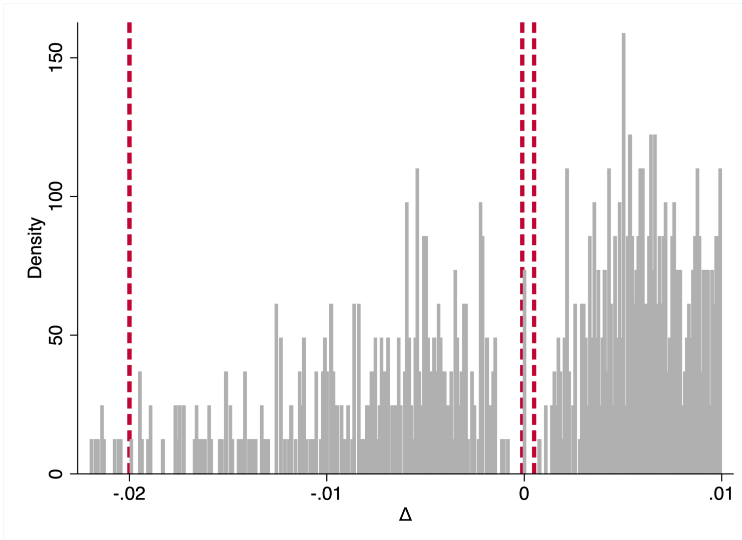


Figure 4: Distribution of Δ for the city of Tsuchiura, 2007–2009.

A single upward deviation. We first consider a single small upward deviation $\rho_1 = .0008$. Under inference problem (P), we seek to maximize the share of histories such that demand $(\mathbf{d}_h)_{h\in H} = ((d_{h,0}, d_{h,1}))_{h\in H}$ satisfies (IC-MKP) subject to $(\widehat{CR})$. For the case of a single upward deviation, (IC-MKP) reduces to

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

or equivalently,

$$(1+\rho_1)d_{h,1} - d_{h,0} \leq \frac{1}{1+M}(d_{h,1} - d_{h,0}).$$

Note that this condition corresponds to the IC constraint (7) in which c_h/b_h is set to $1/(1+M)$. An upward deviation is least profitable (and so the data is best explained) when costs are low. Aggregating across all histories, the IC constraints imply that:

$$(1+\rho_1)\overline{D}(\rho_1|H) - \overline{D}(0|H) \le \frac{1}{1+M}(\overline{D}(\rho_1|H) - \overline{D}(0|H)). \tag{13}$$

Hence, if

$$(1+\rho_1)\widehat{D}(\rho_1|H) - \widehat{D}(0|H) > \frac{1}{1+M}(\widehat{D}(\rho_1|H) - \widehat{D}(0|H)), \tag{14}$$

then for $(x_\lambda)_{\lambda\in\Lambda}$ small enough, (IC-MKP) and ($\widehat{CR}$) cannot be satisfied together for all histories $h \in H$. In auctions from Tsuchiura, a small upward deviation does not change a bidder's demand: $\widehat{D}(0|H) \simeq \widehat{D}(\rho_1|H)$. This implies that (14) holds regardless of the value of M, and at least some fraction of histories in H must be deemed non-competitive. Note that when $M = +\infty$, expression (13) reduces to a statement about the elasticity of demand, i.e., Proposition 2.

The dotted line in Figure 5 corresponds to our estimate of the upper bound on the share of competitive histories based on Proposition 3 as a function of minimum markup m, using single upward deviation $\rho_1 = .0008$. For these estimates and all the estimates we present below, we use Lemma 3 to construct confidence bounds and set tolerance $((x_\lambda))_{\lambda\in\Lambda}$ so that our estimate is an upper bound for the true share of competitive histories with 95% confidence. We set $M = 0.5$.[31] Since bounds based only on upward deviations do not depend on m, the dotted line in Figure 5 is constant at 0.87.

A single upward deviation and tied bids. We now consider combining the upward deviation with an an infinitesimal downward deviation ($\rho \in \{0, .0008\}$). Any mass of tied bids is inherently non-competitive since they create a meaningful benefit from reducing bids by the smallest possible amount. We note that tied bids are present in the data, but that their mass is small. Combining the upward deviation with an infinitesimal downward deviation, we estimate a 95% confidence bound on the share of competitive histories to be 0.86. This

[31] As Figure OC.7 in the Online Appendix shows, our results are not highly sensitive to the choice of M.

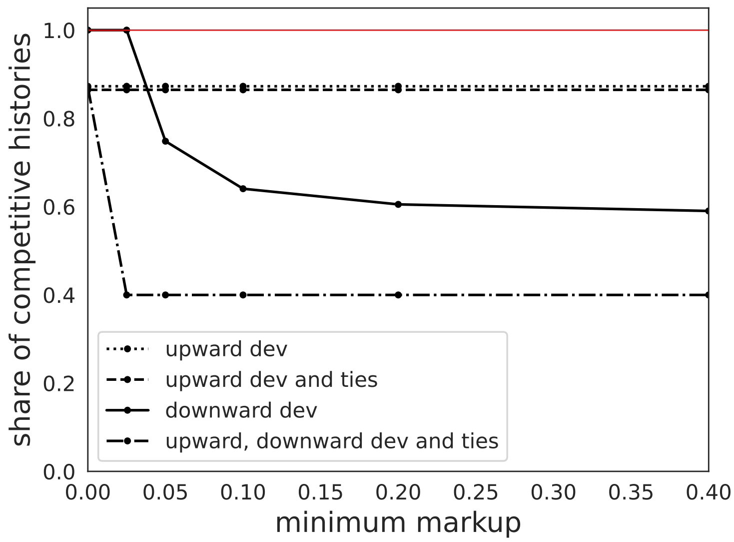


Figure 5: Share of competitive histories, Tsuchiura.

Deviations $\{-0.02, 0, 0.0008\}$; maximum markup 0.5.

is illustrated as the horizontal dashed line in Figure 5. As the figure shows, the presence of ties has a very small impact on our estimate of the share of non-competitive histories.

A single downward deviation. We now consider the implications from a single downward deviation $\rho = -.02$. For a single downward deviation, (IC-MKP) reduces to

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

or equivalently,

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

This condition is equivalent to setting $c_h/b_h = 1/(1+m)$ in the IC constraint (7), which means that the data patterns are best explained when costs are high.

Constraints (IC-MKP) and $(\widehat{CR})$ define a convex set. For $(x_\lambda)_{\lambda\in\Lambda}$ small enough, (IC-MKP) and $(\widehat{CR})$ can be satisfied for all histories $h \in H$ if and only if the mean empirical demand satisfies (IC-MKP):

$$(1+\rho_{-1})\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) \leq \frac{1}{1+m}\left[\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)\right]. \tag{15}$$

Since $\rho_{-1} < 0$, it follows that condition (15) always holds if m is close to zero. This is intuitive: for a sufficiently small margin, say $m < 2\%$, reducing bids by 2% results in losses for each auction ($.98 \times 1.02 - 1 < 0$). In contrast, if $1+\rho_{-1} > 1/(1+m)$ and demand increase $\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)$ is sufficiently large, then (15) does not hold and some share of histories must be considered non competitive.

The solid line in Figure 5 plots our estimate of the 95% confidence bound as a function of m. In this data, a 2% drop in prices leads to a 44 percentage-point increase in the probability of winning the auction, from $\widehat{D}(0|H) = 22.6\%$ to $\widehat{D}(\rho_{-1}|H) = 66.2\%$, almost tripling demand. Hence, as minimum markup m increases from 0, inequality (15) fails, implying that (IC) and ($\widehat{CR}$) cannot be solved together for all histories $h \in H$. Correspondingly, the bound in Figure 5 is equal to 1 for low values of m, and becomes less than 1 as m increases.[32]

Complementary upward and downward deviations. Conditions (14) and (15) highlight that individual upward and downward deviations are 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. Hence, upward and downward deviations are complementary from the perspective of inference. The dashed-dotted line in Figure 5 plots our bound on the share of competitive histories using all three deviations ($\rho \in \{-.02, 0, .0008\}$). For all values of m, considering both upward and downward deviations leads to a tighter bound for the share of competitive histories than either upward or downward deviations alone ($\widehat{s}_{\mathsf{comp}}$ is 0.40 when $m = 0.025$). The high costs needed to ensure that a downward deviation is not attractive also make upward deviations more attractive.

[32]The markups we consider contain prior estimates in the literature. Krasnokutskaya (2011) estimates markups ranging from 0.1% to 24% in Michigan highway procurement auctions. Bajari et al. (2014) estimate markups from 2.9% to 26.1% in California highway procurement auctions.

Online Appendix OB.3 establishes this complementarity formally in a simple case.

7.2 Findings from Aggregate Data

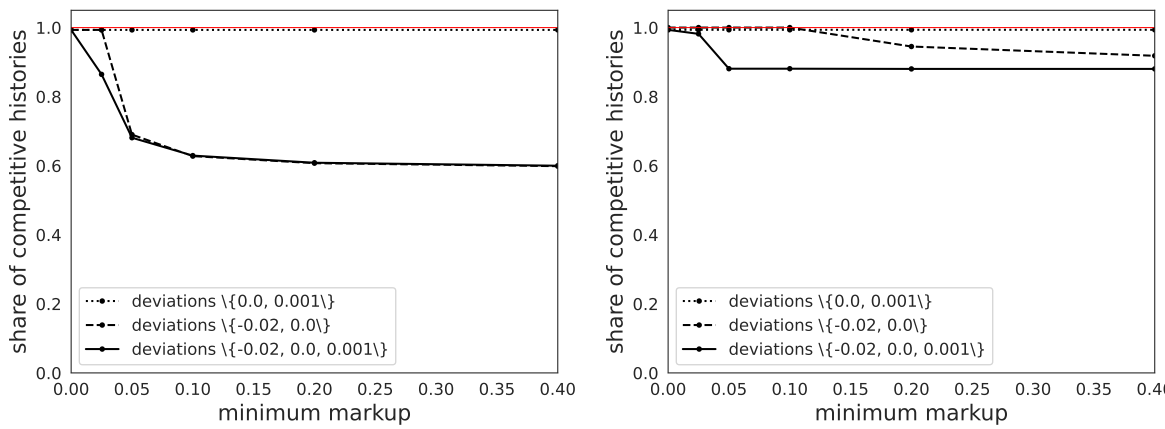
We now apply our tests to the full set of auctions in each of our datasets, taking H as the set of histories corresponding to all municipal or national auctions. Clearly, H is adapted. Going forward, when applying the results of Section 6, we set $M = .5$ and use the fixed set of deviations $\{-.02, 0, .001\}$ for all datasets. Using a fixed set of deviations for all datasets is likely suboptimal for statistical power, but offers more transparency.[33]

Figure 6 shows our estimates of the 95% confidence bound on the share of competitive histories as a function of minimum markup m, for city and national auctions. The dotted line corresponds to the estimated bound when we set deviations $\rho \in \{0, .001\}$. The dashed and the solid lines correspond to estimates for deviations $\rho \in \{-.02, 0\}$ and $\rho \in \{-.02, 0, .001\}$, respectively. In the case of national auctions we use penalized incentive compatibility conditions accounting for rebidding detailed in Online Appendix OA.[34] We note that upward deviations alone allow us to detect only a very small number of non-competitive histories both in city-level auctions and in national-level auctions. One explanation for this is that looking at the full set of auctions causes us to mix competitive and non-competitive histories, thereby weakening our ability to detect non-competitive behavior.[35] Correspondingly, our bound does not have much bite when we consider a single upward deviation and we take H to be the entire sample of auctions. Note that this does not mean that upward deviations are uninformative: especially in the case of national data, considering both upward and downward deviations can yield significantly tighter bounds on the share of competitive histories than either deviation alone.

[33] The magnitude of deviations was set using residual demand plotted in Figures OC.1(a) and OC.1(b) to achieve large gains in demand for downward deviations, and small drops in demand for upward deviations.

[34] Online Appendix OC shows that our estimates are not very sensitive to the value of maximum markup M, or assumptions about continuation payoffs upon re-bidding.

[35] Even though the missing mass at $\Delta = 0$ is very noticeable in Figure 1, the density at $\Delta = 0$ is not quite low enough at the aggregate level for upward deviations alone to have statistical power.



(a) city data (b) national data

Figure 6: Share of competitive histories, city and national level data.

Deviations $\{-0.02, 0, 0.001\}$; maximum markup 0.5.

7.3 Zeroing-in on Specific Firms

We now consider applying our tests to individual firms. As we highlight in Ortner et al. (2020), 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 samples, we consider the thirty firms that participate in the most auctions in each dataset. For each firm, we estimate a bound on the share of competitive histories taking H to be the set of histories corresponding to all instances in which the firm participates (set H is clearly adapted to the firm's information).

Panel (a) of Table 2 reports the results for firms active in the city sample. We order firms according to the number of auctions in which they participate. We report this number in column 2. Column 3 reports the share of auctions each firm wins as a fraction of the number of auctions in which the firm participates. Column 4 reports our estimates for the bound on the share of competitive histories using deviations $\{-0.02, 0, 0.001\}$, minimum markup $m = .025$ and maximum markup $M = .5$. Panel (b) of Table 2 reports the corresponding results for firms in the national sample. Our bound is less than 1 for eighteen firms for the city sample and twenty-four for the national sample.

7.4 Consistency with Proxies for Collusion

We now show that our bounds on the share of competitive histories are consistent with proxies of collusive behavior. We consider different adapted subsets H associated with markers of collusion: whether or not the industry has been prosecuted for collusive practices, and whether bids are high compared to reserve prices.[36] We then compute the associated estimates for the share of competitive histories.

Before and after prosecution. As we noted in Section 2, 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 share 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 (58 bids, vs more than 560 in the other industries).

Figure 7 shows our estimates of the 95% confidence bound on the share of competitive histories as a function of m, for the three remaining groups of firms. As above, we consider deviations $\{-0.02, 0, .001\}$ and set $M = .5$. For all three industries, we find that the share of competitive histories is higher after the investigation than before, consistent with the interpretation that collusion was more rampant before the investigation and less severe afterwards.[37]

High vs. low bids. We now compare the estimates that we obtain when focusing on histories with low bids relative to high bids. Because collusion typically elevates prices, we

[36] In Online Appendix OC.3 we correlate our firm-level bounds on competitiveness with an indicator of collusive behavior introduced by Imhof et al. (2018) in the context of Swiss antitrust policy.

[37] In Online Appendix OC.2 we perform sensitivity analysis of the share of competitive histories in these three industries to changes in the confidence level used for inference, and the set of deviations considered.

(1) Rank	(2) Participation	(3) Share won	(4) Share comp
1	347	0.19	**0.88**
2	336	0.21	**0.86**
3	299	0.08	**0.98**
4	293	0.05	1.00
5	290	0.14	1.00
6	287	0.20	1.00
7	269	0.14	**0.94**
8	268	0.09	**0.97**
9	262	0.12	1.00
10	259	0.18	**0.90**
11	252	0.12	**0.97**
12	241	0.12	**0.95**
13	239	0.16	**0.93**
14	238	0.09	**0.99**
15	227	0.11	**0.97**
16	226	0.12	**0.99**
17	225	0.08	**0.96**
18	223	0.12	**0.98**
19	220	0.07	1.00
20	218	0.08	1.00
21	211	0.07	1.00
22	210	0.14	**0.95**
23	209	0.17	**0.93**
24	204	0.15	1.00
25	203	0.11	**0.98**
26	199	0.06	1.00
27	190	0.12	1.00
28	189	0.06	1.00
29	188	0.16	**0.94**
30	187	0.08	1.00

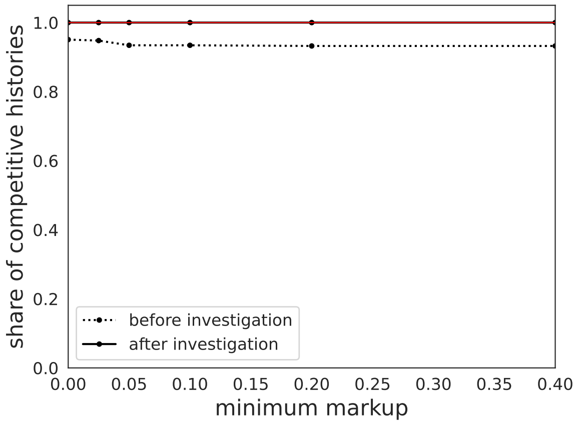
(a) City Data

(1) Rank	(2) Participation	(3) Share won	(4) Share comp
1	4,044	0.17	**0.84**
2	3,854	0.07	**0.91**
3	3,621	0.12	**0.85**
4	2,998	0.15	1.00
5	2,919	0.06	**0.92**
6	2,547	0.08	**0.71**
7	2,338	0.07	**0.74**
8	2,333	0.07	**0.74**
9	2,328	0.04	**0.95**
10	2,292	0.06	**0.75**
11	2,237	0.08	**0.90**
12	2,211	0.03	**0.96**
13	2,015	0.09	**0.76**
14	1,984	0.08	**0.75**
15	1,727	0.07	1.00
16	1,674	0.05	**0.84**
17	1,661	0.03	**0.94**
18	1,660	0.08	**0.75**
19	1,589	0.07	**0.79**
20	1,427	0.10	1.00
21	1,393	0.06	**0.86**
22	1,392	0.07	1.00
23	1,370	0.04	**0.92**
24	1,368	0.14	1.00
25	1,353	0.05	**0.80**
26	1,342	0.09	1.00
27	1,337	0.04	**0.87**
28	1,326	0.08	**0.92**
29	1,291	0.06	**0.86**
30	1,260	0.06	**0.93**

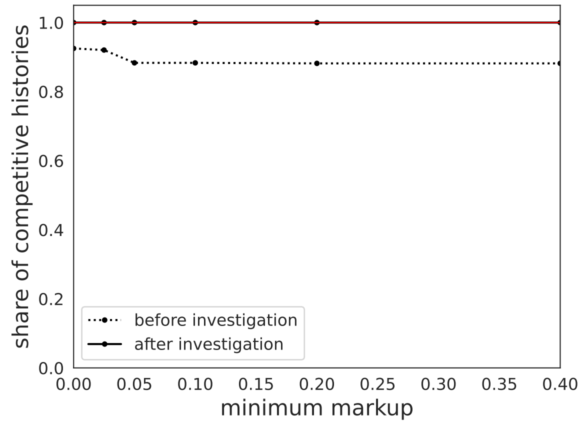
(b) National Data

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. Column 4 present our 95% confidence bound on the share of competitive histories for each firm based on Proposition 3. For our estimates of column 5, we use deviations $\{-0.02, 0, 0.001\}$, minimum markup $m = 0.025$ and maximum markup $M = 0.5$.

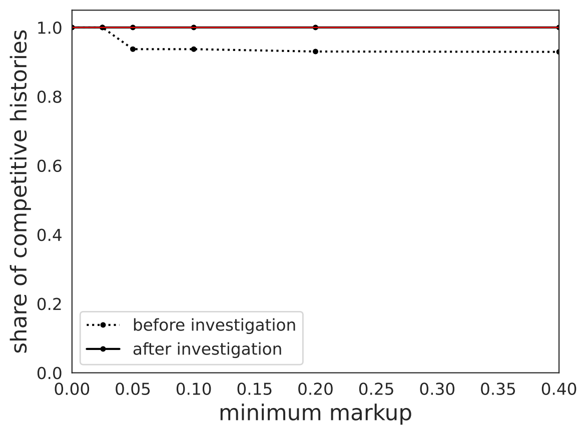
Table 2: Share of competitive histories, individual firms



(a) Electric



(b) Floods



(c) Pre-Stressed Concrete

Figure 7: Share of competitive histories, before and after JFTC investigation.

Deviations $\{-0.02, 0, 0.001\}$; maximum markup 0.5.

expect the share of competitive histories to be higher for histories with low bids and vice versa. Specifically, we divide the city-level data into histories with high bids relative to the reserve price (i.e., $\frac{b}{r} \geq 0.9$) and those with low bids (i.e., $\frac{b}{r} < 0.9$). Since the reserve price is known to bidders in the city auctions, these two sets of histories are adapted.

Figure 8 plots our estimates of the bound on the share of competitive histories for histories with high bids (solid line) and low bids (dotted line) for the city auctions. As before, we use deviations $\{-0.02, 0, .001\}$ and set $M = .5$. As the figure shows, the fraction of competitive histories is lower for histories with high bids. This is consistent with the fact that collusion increases bids.

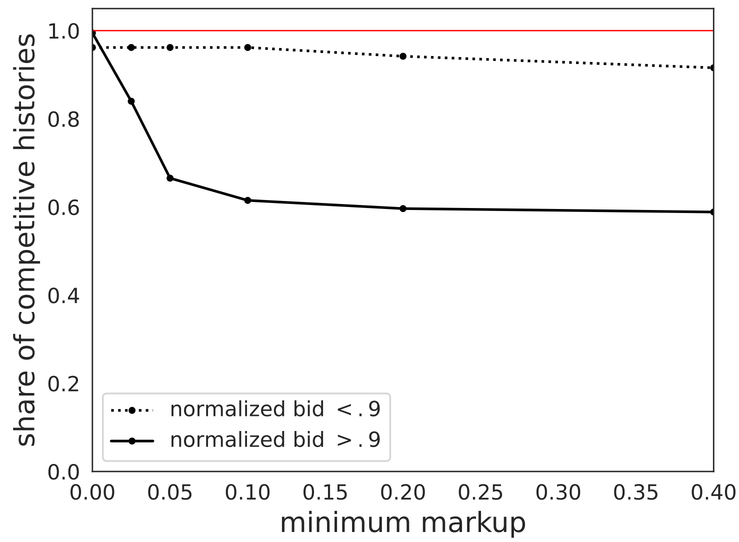


Figure 8: Share of competitive histories by bid level, city data.

Deviations $\{-0.02, 0, 0.001\}$; maximum markup 0.5.

8 Discussion

This paper develops tests of competitive bidding valid under general assumptions allowing for non-stationary costs and signals. In addition to the motivating observation that winning bids are isolated, we identify another suspicious pattern in the data: a 2% reduction in bids causes a large increase in demand.

Our tests are conservative: they can be passed by any firm that is bidding competitively under some information structure. Such caution is justified by the high direct and collateral costs of launching a formal investigation against non-collusive firms (Imhof et al., 2018). In a companion paper building on the same framework (Ortner et al., 2020), we identify another important property of such tests: antitrust investigation based on tests that are passed by any competitive firm 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 in public procurement.

The Online Appendices 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 a discussion of practical aspects of our tests: (i) firms' responses to antitrust oversight, (ii) the relation between the rejection of the test and collusion, (iii) non-collusive explanations for missing bids, and (iv) issues related to the implementation of our tests.

What if firms adapt? Non-competitive bidders may adapt to the screens used by antitrust authorities, reducing their efficacy. We show in a companion paper (Ortner et al., 2020) that antitrust oversight based on tests that are robust to the information structure always reduces the set of enforceable collusive schemes available to cartels. That is, screens based on robust tests always make cartels worse off, even if firms know they are being monitored and adapt their play accordingly.

Moreover, we note that simple adaptive responses to our tests may themselves lead to suspicious patterns. For instance, collusive firms may adapt their play to our tests by "filling the gap" in the distribution of Δ to avoid generating the suspicious patterns in Figure 1.[38] While such adaptive response would reduce the profitability of upward deviations by winning bidders, it would also make downward deviations by losing bidders more profitable, potentially also leading to patterns that would fail our tests.[39]

[38] In Kawai et al. (2020), we propose screens of non-competitive behavior that have statistical power precisely when the distribution of bid differences Δ has sufficient mass around $\Delta = 0$.

[39] In fact, as we argue in Ortner et al. (2020), the bidding patterns we highlight as suspicious can be interpreted as by-products of cartel responses to existing antitrust oversight flagging-out tied and nearly tied bids. This view connects our findings to a body of work on the internal organization of cartels. Asker (2010) describes the internal allocation procedure used in a stamp auction cartel. Pesendorfer (2000) studies the bidding patterns for school milk contracts and compares the collusive schemes used by cartels that used transfers and those that did not. Clark and Houde (2013) document collusive strategies used by the retail gasoline cartel in Quebec. Clark et al. (2018, 2020) study the effect of an investigation on bidding behavior.

Rejection of the test and collusion. Our tests, which are aimed at detecting failures of competitive behavior, do not distinguish among the various reasons why a given dataset may be inconsistent with competition. Failure to pass our test does not necessarily imply bidder collusion.[40] However, findings from Section 7 show that rejection of our tests is in fact correlated with different markers of collusion. Indeed, in industries that were investigated for bid rigging, the fraction of competitive histories is lower before the investigation than after. The fraction of competitive histories is lower at histories at which bids are high relative to the reserve price. Altogether, this suggests that our test are sufficiently powered to flag cartels in practice.

Non-collusive explanations. It is instructive to evaluate potential non-collusive explanations for the bidding patterns in our data. One possibility is that bidders are committing errors, say playing an ϵ-equilibrium of the game. This explanation is not entirely satisfactory for two reasons. First, the potential gains from downward deviations are not small, and bidders have many opportunities to learn. 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.

Another possible explanation for 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 increases the attractiveness of upward deviations. Similarly, learning-by-doing reduces the cost of accepting a project. This increases the attractiveness of downward deviations.

[40]For instance, firms that hold systematically incorrect beliefs about their opponents' bids may fail our test, even if they are bidding to maximize their stage-game subjective expected payoffs. In addition, our tests cannot distinguish between tacit and explicit collusion.

Practical implementation. An important question for practical implementation concerns the choice of what set of histories H to apply our methods to. Besides the requirement that set H should be adapted, our analysis assumes that bidders take into account all the relevant information provided by past bidding data.[41] Pooling data from competitive and non-competitive environments may obscure non-competitive bidding patterns, making it harder to reject the null of competition. The value of pooling data from different settings is that it potentially affords better power if non-competitive behavior is broadly prevalent. In addition, separately analyzing sub-industries raises multiple hypothesis testing concerns.

A second important aspect of practical implementation is the choice of deviations $\rho \in (-1, +\infty)$ being considered. As we have already noted, downward deviations are most binding when small drops in price lead to a large increase in demand. Upward deviations are most binding when relatively large increases in price cause relatively small drops in demand. Note that it may be difficult to estimate drops in demand for very small upward deviations (even if the empirical standard deviation of estimates is an exact 0). For that reason, we suggest that ρ should not be smaller (in absolute value) than the granularity of bids; i.e., ρ should lie in the support of $b_{i,k}/b_{i,k'} - 1$ where $b_{i,k}$ and $b_{i,k'}$ are adjacent (not necessarily winning) bids.

Appendix

A Proofs

Proof of Proposition 1. Let H be an adapted set of histories, and fix $\rho \in (-1, \infty)$. Recall that one auction happens at each time $t \in \{0, 1, ..., T\}$ and that bidding outcomes are

[41] Note that our analysis extends even if bidders do not keep track of bids from other regions, provided that realized demand is independent across regions, conditional on the information bidders do keep track of. If bidders do not keep track of external information even though it is relevant, then our estimator $\widehat{D}(\rho|H)$ of demand $\overline{D}(\rho|H)$ is unbiased (and consistent provided the sample size of the data bidders do track becomes arbitrarily large), but our analysis underestimates its true standard error.

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

$$\mathsf{prob}(|S_T| \geq \nu|H|) \leq 2\exp\left(\frac{-\nu^2|H|^2}{2\sum_{t=0}^{T}\overline{N}_t^2}\right).$$

Observing that $\sum_{t=0}^{T}\overline{N}_t^2 \leq \sum_{t=0}^{T}\overline{N}_t N_{\max} = N_{\max}|H|$, this implies that

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

This concludes the proof. ■

Proof of Proposition 3. By assumption, constraint $(\widehat{CR})$ is satisfied by the distribution of historical beliefs μ^* with probability greater than α. Whenever μ^* satisfies $(\widehat{CR})$, $\widehat{s}_{\mathsf{comp}}$ is weakly larger than s_{comp} by construction. This implies that $\widehat{s}_{\mathsf{comp}} \geq s_{\mathsf{comp}}$ with probability greater than α.

Hence, if $s_{\mathsf{comp}} \geq s_0$, then with probability greater than α, $\widehat{s}_{\mathsf{comp}} \geq s_{\mathsf{comp}} \geq s_0$, so that test $\tau = \mathbf{1}_{\widehat{s}_{\mathsf{comp}}<s_0}$ rejects the null that $s_{\mathsf{comp}} \geq s_0$ with probability less than $1-\alpha$. ■

References

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

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 (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.

Billingsley, P. (1995): *Probability and measure*, John Wiley & Sons.

Bresnahan, T. F. (1987): "Competition and collusion in the American automobile industry: The 1955 price war," *The Journal of Industrial Economics*, 457–482.

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, and A. De Leveranoc (2020): "Complementary Bidding and the Collusive Arrangement: Evidence from an Antitrust Investigation," .

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.

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.

Kaido, H., F. Molinari, and J. Stoye (2019): "Confidence intervals for projections of partially identified parameters," *Econometrica*, 87, 1397–1432.

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

Kawai, K., J. Nakabayashi, J. Ortner, and S. Chassang (2020): "Using Bid Rotation and Incumbency to Detect Collusion: A Regression Discontinuity Approach," .

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

Krasnokutskaya, E. and K. Seim (2011): "Bid preference programs and participation in highway procurement auctions," *The American Economic Review*, 101, 2653–2686.

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

Marmer, V., A. A. Shneyerov, and U. Kaplan (2016): "Identifying collusion in english auctions," *Available at SSRN 2738789*.

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.

Nevo, A. (2001): "Measuring market power in the ready-to-eat cereal industry," *Econometrica*, 69, 307–342.

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

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

Porter, R. 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.

Online Appendix – Not for Publication

OA 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 $k \in \{1, \cdots, \overline{k}\}$. The auctioneer picks a secret reserve price r. At each stage k, bidders submit bids $b_{i,k}$. A winner is declared 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 Section 6 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 β_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 σ_i and σ_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 OA.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 bidder i's belief that a given deviation is marginal for continuation or information only depends on the bidder i's beliefs about bids $b_{-i,1}$.

For bids $b_{i,1}$ and $b'_{i,1}$, 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 with probability 0 in period $k+1$ if the auction continues.

- Continuation payoffs under σ_i and σ'_i are equal conditional on the deviation being non-marginal.

If the deviation is *an upward deviation* then,

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

- If continuation strategies β_i, β_{-i} are monotonic in observed bids, then

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

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

$$(b_{i,1}-c_i)\mathsf{prob}_{\sigma_i,\sigma_{-i}}(\mathsf{win}_{i,1}) \geq (b_{i,1}'-c_i)\mathsf{prob}_{\sigma_i',\sigma_{-i}}(\mathsf{win}_{i,1}). \tag{O1}$$

This coincides with the IC constraint for upward deviations used in Sections 5 and 6.

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

$$\mathbb{E}_{\sigma_i',\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\Big|\mathsf{MI}\right] \geq (1-\alpha)\mathbb{E}_{\sigma_i,\sigma_{-i}}\left[\sum_{k>1}(b_{i,k}-c_i)\mathbf{1}_{\mathsf{win}_{i,k}}\Big|\mathsf{MI}\right]. \tag{O2}$$

In words, following a downward deviation that is 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 (O2) would hold with $\alpha = 0$. In our empirical investigation, we use $\alpha = .5$.

Finally, we observe that the following bounds hold

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

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

$$\begin{aligned}(b_{i,1}-c_i)\mathsf{prob}_{\sigma_i,\sigma_{-i}}(\mathsf{win}_{i,1}) \geq & (b_{i,1}'-c_i)\mathsf{prob}_{\sigma_i',\sigma_{-i}}(\mathsf{win}_{i,1}) \\ & - \left[\mathsf{prob}_{\sigma_{-i}}(\mathsf{MC}) + \alpha\mathsf{prob}_{\sigma_{-i}}(\mathsf{MI})\right]\mathbb{E}[(r-c_i)^+].\end{aligned} \tag{O3}$$

Equations (O1) and (O3) replace (IC) in the inference problem defined Section 6. In addition to disciplining residual demand under μ, expanded consistency requirement $(\widehat{CR})$ must ensure that the probability of events MI and MC under the true historical average distribution of beliefs μ^* must also be close to their sample probability. Denote by mi_h and mc_h the probability that downward deviation ρ is marginal for information or continuation at h. An extension of Proposition 1 implies that for any $\rho < 0$, with large probability as $|H|$ becomes large,

$$\mathbb{E}_{\mu^*}[mi] \equiv \frac{1}{|H|}\sum_{h \in H} mi_h \in \left[\frac{1}{|H|}\sum_{h \in H} \mathbf{1}_{r<(1+\rho)b_{i,1}<\wedge \mathbf{b}_{-i,1}} \pm K\right]$$
$$\mathbb{E}_{\mu^*}[mc] \equiv \frac{1}{|H|}\sum_{h \in H} mc_h \in \left[\frac{1}{|H|}\sum_{h \in H} \mathbf{1}_{(1+\rho)b_{i,1}\leq r<\wedge \mathbf{b}_{-i,1}} \pm K\right]$$

where K is an arbitrary fixed tolerance parameter. This implies that we can expand coverage sets $\mathcal{D}_\alpha$ introduced in Section 6.3 to cover not only the true expected vector of demands $\mathbb{E}_{\mu^*}[\mathbf{d}]$, but also the true expected probability that deviations are marginal: $\mathbb{E}_{\mu^*}[mi]$ and $\mathbb{E}_{\mu^*}[mc]$.

OB Further Theoretical Results

OB.1 Connection with Bayes Correlated Equilibrium

In this section we further extend the estimator introduced in Section 6 and clarify what would need to be added so that asymptotically, it exploits 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.

In Section 6 we were able to express inference problem (P) as a function of beliefs alone. Because both costs and bids are part of bidders' information set (so that selecting histories on the basis of both beliefs and costs is adapted), to exploit all the information content of equilibrium, we must impose constraints on demand conditional on both bids and costs. For this reason, instead of expressing our inference problem using the distribution of beliefs alone, we express this new inference problem using the historical profile of beliefs, costs and bids.[42]

For any history $h \in H$, let $\omega_h = ((d_{h,n})_{n\in\mathcal{M}}, c_h)$ be the demand and cost of the firm associated with history h. Let $\omega_H = (\omega_h)_{h\in H}$ denote the profile of demands and costs across all histories in H. Let $\Omega \equiv \{\omega_H : \forall h \in H, (d_{h,n})_{n\in\mathcal{M}} \in \mathcal{F}\}$ be the set of environments ω_H with feasible demands.

For each profile $\omega_H \in \Omega$, define

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

to be the set of histories in H that satisfy markup constraint (MKP) and are rationalizable as competitive under ω_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' demands and costs are given by ω_H.

We extend problem (P) as follows. For any environment ω_H and $(i, b, c) \in I \times B \times C$, let us define $H_{i,b,c}(\omega_H) \equiv \{h \in H | (i_h, b_h, c_h) = (i, b, c)\}$ to be the 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 $K : \mathbb{N} \to \mathbb{R}^+$ such that

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

we consider inference problem (P′)

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

$$\text{s.t. } \forall(i,b,c), \forall n, \quad D_n(\omega_H, H_{i,b,c}(\omega_H)) \in \Big[\widehat{D}(\rho_n|H_{i,b,c}(\omega_H)) - K(|H_{i,b,c}(\omega_H)|),$$
$$\widehat{D}(\rho_n|H_{i,b,c}(\omega_H)) + K(|H_{i,b,c}(\omega_H)|)\Big].$$

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 ∞, $\widehat{s}_{\mathsf{comp}}$ is an upper bound to the share of competitive histories. Imposing consistency requirements conditional on bids and costs lets us establish a converse:

[42]As in Section 6, we could directly consider the joint distribution of beliefs, costs and bids.

if data passes our extended tests, then the joint distribution of bids and costs is an ϵ-Bayes correlated equilibrium in the sense of Hart and Mas-Colell (2000).

Consider an ω_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, ω_H).

Proposition OB.1. *For any $\epsilon > 0$, for $|H|$ large enough, $\widehat{s}_{\mathsf{comp}} = 1$ implies that $\widehat{\mu}$ is an ϵ-Bayes correlated equilibrium of the stage-game first-price auction.*

Proof. Consider demand and costs $(d_{h,n}, c_h)_{h\in H}$ 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,

$$\frac{1}{|H|}\left|\sum_{h\in H_{i,b,c}} d_{n,h} - \mathsf{prob}_{\widehat{\mu}}(\wedge\mathbf{b}_{-i} \succ (1+\rho_n)b_i | i, b, c)\right| \leq \epsilon. \tag{O1}$$

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

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

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

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

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 ϵ-Bayes correlated equilibrium in the sense of Hart and Mas-Colell (2000). ■

OB.2 Bounding the Share Competitive Histories in Simple Cases

This section provides an explicit characterization of bound $\widehat{s}_{\mathsf{comp}}$ when we consider either a single upward deviation, or a single downward deviation.

OB.2.1 Inference from an upward deviation

Consider first the case of a single upward deviation $\rho_1 > 0$. Let $\Lambda = \{(1,0),(0,-1)\}$ and $x_\lambda = x > 0$ for all $\lambda \in \Lambda$. Set $\mathcal{D}_\alpha$ then takes the form:

$$\mathcal{D}_\alpha = \left\{\mathbf{d} = (d_0, d_1) \in \mathcal{F} \text{ s.t. } d_0 \leq \widehat{D}(\rho_0|H) + x \text{ and } d_1 \geq \widehat{D}(\rho_1|H) - x\right\}.$$

Bound $\widehat{s}_{\mathsf{comp}}$ takes the form

$$\widehat{s}_{\mathsf{comp}} = \min\left\{1, 1 - \frac{(\widehat{D}(\rho_1|H) - x)\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) - (\widehat{D}(\rho_0|H) + x)}{\rho_1\left(1 + \frac{1}{M}\right)}\right\}.$$

Note that when $M = +\infty$, $\widehat{s}_{\mathsf{comp}} < 1$ is equivalent to the condition that elasticity of demand is larger than -1:

$$\widehat{s}_{\mathsf{comp}} < 1 \iff \frac{\log(\widehat{D}(\rho_1|H) - x) - \log(\widehat{D}(\rho_0|H) + x)}{\log(1 + \rho_1)} > -1.$$

When bidder's demand is unchanged following an upward deviation (i.e., when $\widehat{D}(\rho_1|H) \approx \widehat{D}(\rho_0|H)$), for $x > 0$ small enough we have $\widehat{s}_{\mathsf{comp}} < 1$ even as M goes to $+\infty$.

Proof. For any competitive history $h \in H$, beliefs $d_{h,1}$ and $d_{h,0}$ must be such that

$$\left(1 + \rho_1 - \frac{1}{1+M}\right) d_{h,1} \leq \left(1 - \frac{1}{1+M}\right) d_{h,0}$$
$$\iff d_{h,0} \geq d_{h,1}\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right), \tag{O2}$$

where the first inequality uses the mark-up constraint $\frac{c_h}{b_h} \geq \frac{1}{1+M}$. Suppose that

$$\widehat{D}(\rho_0|H) + x \geq (\widehat{D}(\rho_1|H) - x)\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) \tag{O3}$$

Note that in this case, $\widehat{s}_{\mathsf{comp}} = 1$. Indeed, let $\mu \in \Delta(\mathcal{F})$ be a distribution that puts all its mass at (d_0, d_1), with $d_0 = \widehat{D}(\rho_0|H) + x$, and $d_1 = \widehat{D}(\rho_1|H) - x$. Note that $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$, and

$\mathbb{E}_\mu[\text{IsComp}(\mathbf{d})] = 1$.

Suppose next that (O3) does not hold. Let $\mu \in \Delta(\mathcal{F})$ be a distribution satisfying the constraint $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$, and let $\widehat{s}(\mu) = \mathbb{E}_\mu[\text{IsComp}(\mathbf{d})]$ be the share of competitive histories under μ. Note that

$$\widehat{D}(\rho_0|H) + x \geq \mathbb{E}_\mu[d_0] = \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 1] + (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 0] \tag{O4}$$

$$\widehat{D}(\rho_1|H) - x \leq \mathbb{E}_\mu[d_1] = \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) = 1] + (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) = 0] \tag{O5}$$

Since equation (O2) holds for $\mathbf{d}$ with $\text{IsComp}(\mathbf{d}) = 1$, we have that

$$\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 1] \geq \mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) = 1]\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) \tag{O6}$$

Using (O4)-(O6), we get

$$\begin{aligned}
&\widehat{D}(\rho_0|H) + x - (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 0] \\
\geq& \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 1] \\
\geq& \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) = 1]\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) \\
\geq& \left(\widehat{D}(\rho_1|H) - x - (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) = 0]\right)\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right).
\end{aligned} \tag{O7}$$

From equation (O7), we get

$$\begin{aligned}
&(\widehat{D}(\rho_1|H) - x)\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) - (\widehat{D}(\rho_0|H) + x) \\
\leq& (1 - \widehat{s}_{\mathsf{comp}}(\mu))\left(\mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) = 0]\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) - \mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 0]\right) \\
\leq& (1 - \widehat{s}_{\mathsf{comp}}(\mu))\rho_1\left(1 + \frac{1}{M}\right) \\
\Longrightarrow& \widehat{s}_{\mathsf{comp}}(\mu) \leq 1 - \frac{(\widehat{D}(\rho_1|H) - x)\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) - (\widehat{D}(\rho_0|H) + x)}{\rho_1\left(1 + \frac{1}{M}\right)}.
\end{aligned}$$

where the second inequality follows since $1 \geq \mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 0] \geq \mathbb{E}_\mu[d_1|\text{IsComp}(\mathbf{d}) =$

$0] \geq 0$. Since this inequality holds for all $\mu \in \Delta(\mathcal{F})$ satisfying the constraint $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$,

$$\widehat{s}_{\mathsf{comp}} \leq \overline{s}_1 \equiv 1 - \frac{(\widehat{D}(\rho_1|H) - x)\left(1 + \rho_1\left(1 + \frac{1}{M}\right)\right) - (\widehat{D}(\rho_0|H) + x)}{\rho_1\left(1 + \frac{1}{M}\right)}.$$

Finally, to see that $\widehat{s}_{\mathsf{comp}} = \overline{s}_1$ when (O3) does not hold, let $\overline{\mu} \in \Delta(\mathcal{F})$ be a distribution that puts weight $1 - \overline{s}_1$ on beliefs $d^{nc} = (d_0^{nc}, d_1^{nc}) = (1, 1)$ and puts weight $\overline{s}_1$ on beliefs $d^c = (d_0^c, d_1^c)$ such that $\overline{s}_1 d_0^c + (1 - \overline{s}_1) = \widehat{D}(\rho_0|H) + x$ and $\overline{s}_1 d_1^c + (1 - \overline{s}_1) = \widehat{D}(\rho_1|H) - x$. One can check that $\mathbb{E}_{\overline{\mu}}[\mathbf{d}] \in \mathcal{D}_\alpha$ and that $\mathbb{E}_{\overline{\mu}}[\text{IsComp}(\mathbf{d})] = \overline{s}_1$. Hence, when (O3) does not hold, $\widehat{s}_{\mathsf{comp}} = \overline{s}_1$. ■

OB.2.2 Inference from a downward deviation

Consider next the case of a single downward deviation $\rho_{-1} < 0$. Let $\Lambda = ((-1, 0), (0, 1))$ and $x_\lambda = x > 0$ for all $\lambda \in \Lambda$. Set $\mathcal{D}_\alpha$ then takes the form:

$$\mathcal{D}_\alpha = \left\{\mathbf{d} = (d_{-1}, d_0) \in \mathcal{F} \text{ s.t. } d_{-1} \geq \widehat{D}(\rho_{-1}|H) - x \text{ and } d_0 \leq \widehat{D}(\rho_0|H) + x\right\}.$$

With a single deviation $\rho_{-1} < 0$ such that $m \leq 1/(1+\rho_{-1}) - 1$, all histories can be rationalized as competitive. Indeed, reducing bids by ρ_{-1} would result in negative profits if costs are such that $b_h/c_h = 1 + m$. Instead, if $m > 1/(1 + \rho_{-1}) - 1$, the solution to program (P) is

$$\widehat{s}_{\mathsf{comp}} = \min\left\{1, 1 - \frac{(\widehat{D}(\rho_{-1}|H) - x)\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) - (\widehat{D}(\rho_0|H) + x)}{1 + \rho_{-1}\left(1 + \frac{1}{m}\right)}\right\}.$$

Proof. For any competitive history h, beliefs $d_{h,-1}$ and $d_{h,0}$ must be such that

$$\left(1 + \rho_{-1} - \frac{1}{1+m}\right) d_{h,-1} \leq \left(1 - \frac{1}{1+m}\right) d_{h,0}$$
$$\iff d_{h,0} \geq d_{h,-1}\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right), \tag{O8}$$

where the first inequality uses the mark-up constraint $\frac{c_h}{b_h} \leq \frac{1}{1+m}$. Suppose that

$$\widehat{D}(\rho_0|H) + x \geq (\widehat{D}(\rho_{-1}|H) - x)\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) \tag{O9}$$

Note that in this case, $\widehat{s}_{\mathsf{comp}} = 1$. Indeed, let $\mu \in \Delta(\mathcal{F})$ be a distribution that puts all its mass at (d_{-1}, d_0), with $d_{-1} = \widehat{D}(\rho_{-1}|H) - x$ and $d_0 = \min\{\widehat{D}(\rho_0|H) + x, \widehat{D}(\rho_{-1}|H) - x\}$. Note that $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$, and $\mathbb{E}_\mu[\text{IsComp}(\mathbf{d})] = 1$. Note that (O9) always holds when $m \leq 1/(1+\rho_{-1}) - 1$.

Suppose next that (O9) does not hold. Let $\mu \in \Delta(\mathcal{F})$ be a distribution satisfying the constraint $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$, and let $\widehat{s}_{\mathsf{comp}}(\mu) = \mathbb{E}_\mu[\text{IsComp}(\mathbf{d})]$ be the share of competitive histories under μ. Note that equation (O4) must hold. In addition,

$$\widehat{D}(\rho_{-1}|H) - x \leq \mathbb{E}_\mu[d_{-1}] = \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_{-1}|\text{IsComp}(\mathbf{d}) = 1] + (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_{-1}|\text{IsComp}(\mathbf{d}) = 0] \tag{O10}$$

Since equation (O8) holds for $\mathbf{d}$ with $\text{IsComp}(\mathbf{d}) = 1$,

$$\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 1] \geq \mathbb{E}_\mu[d_{-1}|\text{IsComp}(\mathbf{d}) = 1]\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) \tag{O11}$$

Using (O4), (O10) and (O11), we get

$$\begin{aligned}
\widehat{D}(\rho_0|H) + x \geq & \widehat{D}(\rho_0|H) + x - (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 0] \\
\geq & \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 1] \\
\geq & \widehat{s}_{\mathsf{comp}}(\mu)\mathbb{E}_\mu[d_{-1}|\text{IsComp}(\mathbf{d}) = 1]\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) \\
\geq & \left(\widehat{D}(\rho_{-1}|H) - x - (1 - \widehat{s}_{\mathsf{comp}}(\mu))\mathbb{E}_\mu[d_{-1}|\text{IsComp}(\mathbf{d}) = 0]\right)\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) \\
\geq & \left(\widehat{D}(\rho_{-1}|H) - x - (1 - \widehat{s}_{\mathsf{comp}}(\mu))\right)\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right),
\end{aligned} \tag{O12}$$

where the first inequality uses $0 \leq \mathbb{E}_\mu[d_0|\text{IsComp}(\mathbf{d}) = 0]$ and the last inequality uses $\mathbb{E}_\mu[d_{-1}|\text{IsComp}(\mathbf{d}) = 0] \leq 1$. It follows from (O12) that

$$\widehat{s}_{\mathsf{comp}}(\mu) \leq 1 - \frac{(\widehat{D}(\rho_{-1}|H) - x)\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) - (\widehat{D}(\rho_0|H) + x)}{1 + \rho_{-1}\left(1 + \frac{1}{m}\right)}.$$

Since this holds for all $\mu \in \Delta(\mathcal{F})$ satisfying the constraint $\mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$,

$$\widehat{s}_{\mathsf{comp}} \leq \overline{s}_{-1} \equiv 1 - \frac{(\widehat{D}(\rho_{-1}|H) - x)\left(1 + \rho_{-1}\left(1 + \frac{1}{m}\right)\right) - (\widehat{D}(\rho_0|H) + x)}{1 + \rho_{-1}\left(1 + \frac{1}{m}\right)}.$$

Finally, to see that $\widehat{s}_{\mathsf{comp}} = \overline{s}_{-1}$ when (O9) does not hold, let $\overline{\mu} \in \Delta(\mathcal{F})$ be a distribution that puts weight $1 - \overline{s}_{-1}$ on beliefs $d^{nc} = (d^{nc}_{-1}, d^{nc}_0) = (1, 0)$ and puts weight $\overline{s}_{-1}$ on beliefs $d^c = (d^c_{-1}, d^c_0)$ such that $\overline{s}_{-1}d^c_0 + (1 - \overline{s}_{-1})d^{nc}_0 = \overline{s}_{-1}d^c_0 = \widehat{D}(\rho_0|H) + x$ and

$\overline{s}_{-1}d^c_{-1}+(1-\overline{s}_{-1})d^{nc}_{-1} = \overline{s}_{-1}d^c_{-1}+(1-\overline{s}_{-1}) = \widehat{D}(\rho_{-1}|H)-x$. One can check that $\mathbb{E}_{\overline{\mu}}[\mathbf{d}] \in \mathcal{D}_\alpha$ and that $\mathbb{E}_{\overline{\mu}}[\text{IsComp}(\mathbf{d})] = \overline{s}_{-1}$. Hence, when (O9) does not hold, $\widehat{s}_{\mathsf{comp}} = \overline{s}_{-1}$. ■

OB.3 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 focus on the case of arbitrarily large data so that we can use limit confidence set $\mathcal{D}_\alpha = \{\widehat{\mathbf{D}}\}$, where $\widehat{\mathbf{D}} = (\widehat{D}(\rho_n|H))_{n\in\mathcal{M}}$.

As we discussed in Section 7, individual upward and downward deviations respectively imply strict bounds on the share of competitive histories if and only if

$$\widehat{D}(0|H) - (1+\rho_1)\widehat{D}(\rho_1|H) < \frac{1}{1+M}\left[\widehat{D}(0|H) - \widehat{D}(\rho_1|H)\right],$$
$$(1+\rho_{-1})\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) > \frac{1}{1+m}\left[\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)\right].$$ [43]

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

$$\widehat{D}(0|H) - (1+\rho_1)\widehat{D}(\rho_1|H) = \frac{1}{1+M}\left[\widehat{D}(0|H) - \widehat{D}(\rho_1|H)\right], \tag{O13}$$
$$(1+\rho_{-1})\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H) = \frac{1}{1+m}\left[\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)\right]. \tag{O14}$$

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

Lemma OB.1. *Under* (O13) *and if* $m = 0$ *and* $M = +\infty$*, then all histories can be ratio-*

[43] Checking whether these constrains hold can be performed rapidly, and suggests a rough rationale by which one could pick ρ_{-1} and ρ_1: obtain a smooth estimate of the true demand, and pick ρ_{-1} and ρ_1 so that the conditions above hold with a reliable margin.

nalized as competitive.

Proof. The following demand 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 $c_h = 0$. Since $\rho_1 > 0$, $d_{h,1} \leq 1$.

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

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

We return now to the case where (O13) and (O14) hold for $m > 0$. 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 ρ_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 all histories, conditions $(\widehat{CR})$ and (O13) 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 $d_{h,0} = d_{h,1} = 0$. This implies that

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

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 ρ_{-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 all histories in H, conditions ($\widehat{CR}$) and (O14) 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

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

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 $\widehat{D}(0|H) + \frac{1}{\nu}(\widehat{D}(\rho_{-1}|H) - \widehat{D}(0|H)) \simeq 1.027$.

OB.4 Common Values

We now show how to extend the analysis in Section 6 to allow for common values. Because expected costs conditional on winning now depend on a bidder's bid, costs and demand associated with history $h \in H$ now take the form $(d_{h,n}, c_{h,n})_{n\in\mathcal{M}}$, 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 ρ_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:

$$\forall n \in \mathcal{M}, \quad d_{h,n}c_{h,n} = d_{h,0}c_{h,0} + (d_{h,n} - d_{h,0})\hat{c}_{h,n}, \tag{O15}$$

where $\hat{c}_{h,n} = \mathbb{E}[c|h, \wedge\mathbf{b}_{-i,h} \in (b_h, (1+\rho_n)b_h)]$.[44] 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 constraints hold 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, incentive compatibility constraints become

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

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

$$\max_{n<0} \frac{d_{h,n}(1+\rho_n)b_h - d_{h,0}b_h}{d_{h,n} - d_{h,0}} \leq \min_{n>0} \frac{d_{h,0}b_h - d_{h,n}(1+\rho_n)b_h}{d_{h,0} - d_{h,n}}. \tag{O16}$$

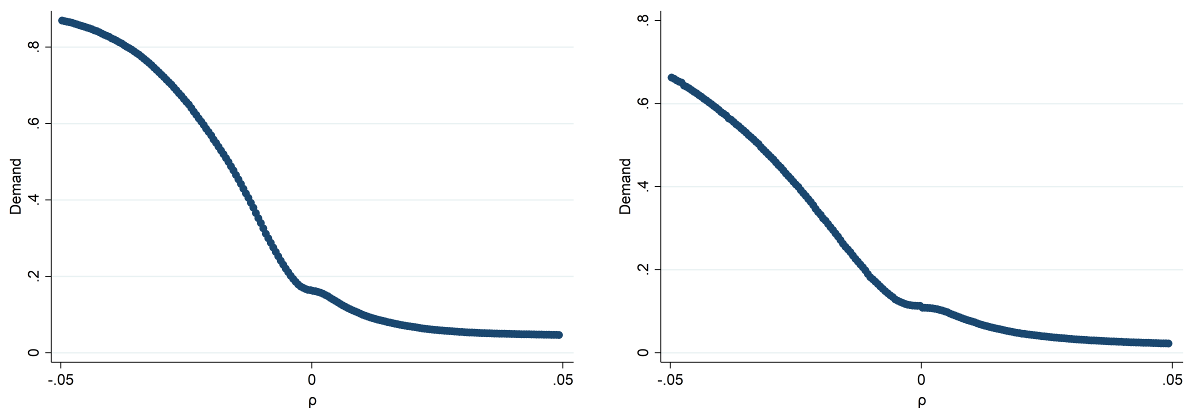
Condition (O16) implies that there also exists a constant profile of costs $c_{h,n} = c_h$ (i.e. a private value cost), that satisfies (IC).

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

OC Further Empirical Findings

OC.1 Additional Figures for Section 2

Figure OC.1 illustrates the clustering of bids we highlight in Section 2 more directly. The two panels of Figure OC.1 plot the sample demand function, $\widehat{D}(\cdot)$, for city-level and national-level auctions. We define the sample demand as follows: $\widehat{D}(\rho) = \frac{1}{|H|}\sum_{i,a} \mathbf{1}_{b_{i,a}(1+\rho)<\wedge \mathbf{b}_{-i,a}}$, where $|H|$ denotes the number of all bids in the dataset. In words, $\widehat{D}(\rho)$ is the sample probability with which bidders win an auction if the bids are changed by a factor of $(1+\rho)$. For panel (a), we find that a drop in bids of 2% increases the likelihood of winning by more than 3-fold from 16.3% to 56.9%. For panel (b), we find that a 2% drop in bids also increases the likelihood of winning by about 3-fold from 10.8% to 33.2%.



(a) city auctions

(b) national auctions

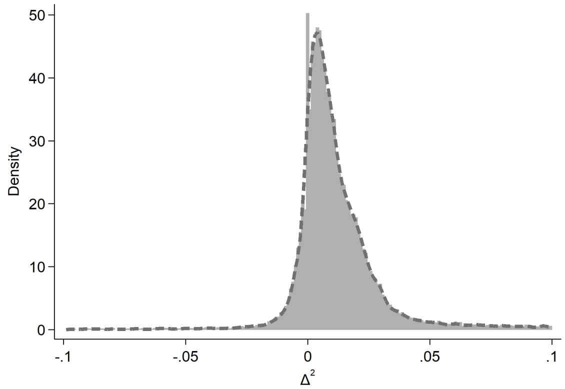
Figure OC.1: Sample demand.

Figure OC.2 plots the distribution of differences Δ^2 between bids after the lowest bid is excluded. Formally, let $NW(a)$ denote the set of non-winning bidders in auction a. Then,

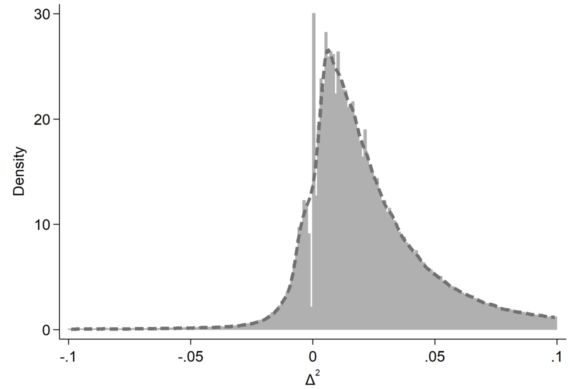
$$\forall i \in NW(a), \quad \Delta^2_{i,a} = \frac{b_{i,a} - \min_{j\in NW(a)} b_{j,a}}{r}.$$

Although there is some bunching at exactly zero making the distributions of Δ^2 somewhat irregular, the kernel density estimates show that there is no corresponding missing mass.

Figure OC.3 plots the distribution of bid differences Δ for the set of auctions whose prices were not renegotiated up (about 15.3% of the sample). As the figure shows, the



(a) city auctions



(b) national auctions

Figure OC.2: Distribution of bid-difference Δ^2 excluding winning bids.

The dotted curves correspond to local (6th order) polynomial density estimates with bandwidth set to 0.0075.

missing mass at $\Delta = 0$ is just as visible when we focus on this set of auctions. This suggests that renegotiation does not drive the patterns we document.

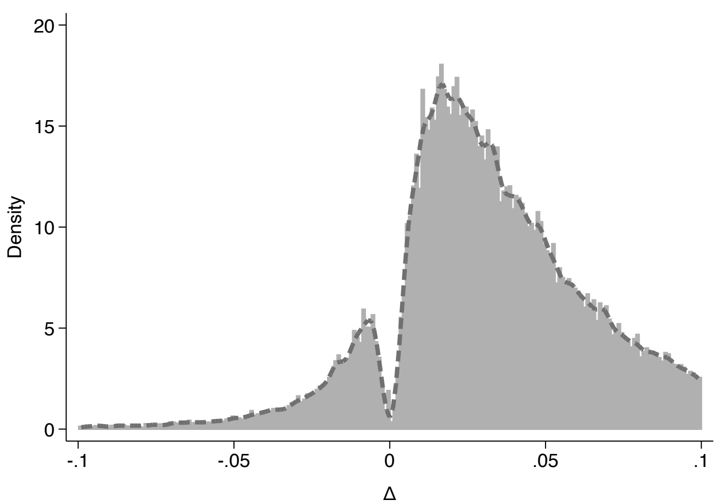


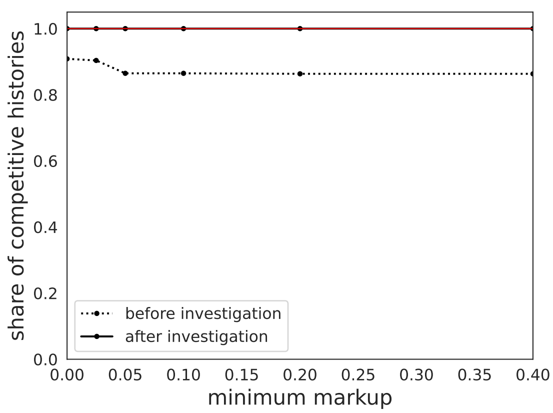
Figure OC.3: Distribution of bid-difference Δ.

National data, auctions with no renegotiation. The dotted curves correspond to local polynomial density estimates with bandwidth set to 0.0075.

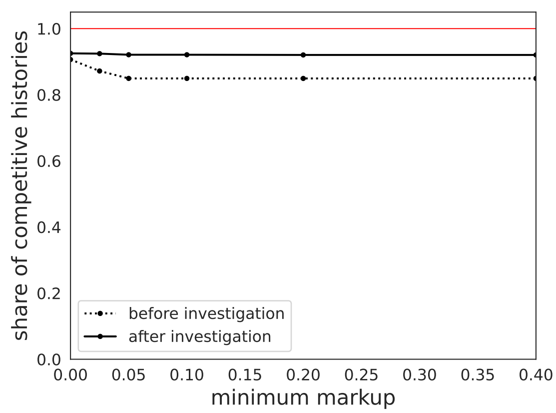
OC.2 Additional Findings Related to Section 7.4

This appendix provides sensitivity analysis of the share of competitive histories for the industries that were investigated for bid rigging by the JFTC to changes in (a) the choice of confidence level, and (b) the set of deviations considered.

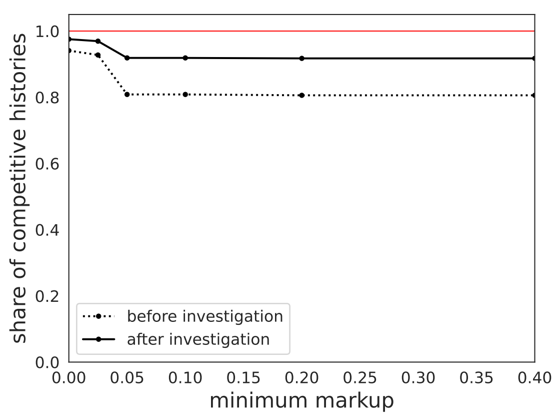
Figure OC.4 presents estimates of the 90% confidence bound on the share of competitive histories for the three industries we studied in Section 7.4. Relative to the bounds in Figure 7 in Section 7.4, these bounds are obtained using smaller tolerance parameters $(x_\lambda)_{\lambda \in \Lambda}$ when computing set $\mathcal{D}_\alpha$. Under this less conservative bound, firms labeled Floods and Pre-Stressed Concrete appear to have continued colluding in the period after the investigation.



(a) Electric



(b) Floods



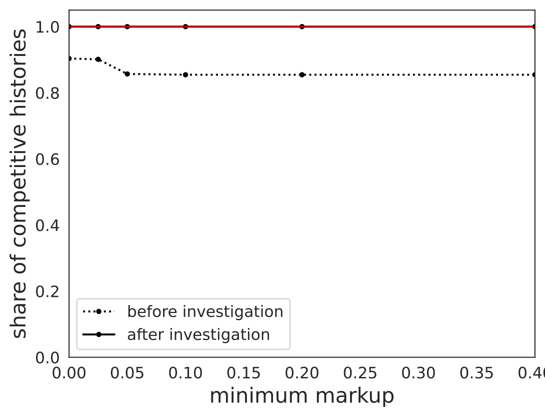
(c) Pre-Stressed Concrete

Figure OC.4: 90% confidence bound on share of competitive histories.

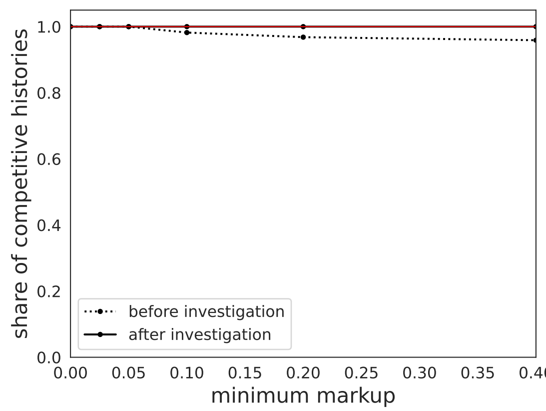
Deviations $\{-0.02, 0, 0.001\}$; maximum markup 0.5.

Figure OC.5 again presents estimates of the 90% confidence bound on the share of competitive histories for the three industries we studied in Section 7.4, but with the set of deviations $\{-0.02, 0, 0.002\}$. Relative to the bounds in Figure OC.4, we now only find evidence of non-competitive behavior for firms labeled Pre-Stressed Concrete in the after period.

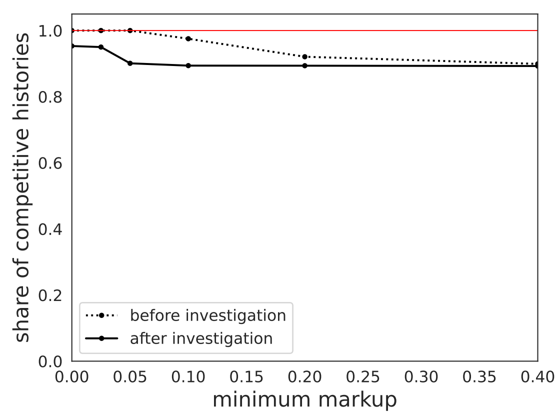
The difference between our estimates in Figure OC.4 and Figure OC.5 can be explained as follows. For Flood auctions occurring after investigation, a .1% upward deviation causes no change in demand. As a result, our estimate on the share of competitive histories in the after period are strictly less than 1 when we consider such a small upward deviation. However, we are worried that this insensitivity of demand to a small upward deviation might be a mechanical consequence of the small number of observations we have in the after period for this industry. Indeed, a .2% upward deviation causes a small drop in demand, and our estimate on the share of competitive histories in the after period is exactly 1 in Figure OC.5.



(a) Electric



(b) Floods



(c) Pre-Stressed Concrete

Figure OC.5: 90% confidence bound on share of competitive histories.

Deviations $\{-0.02, 0, 0.002\}$; maximum markup 0.5.

OC.3 Relation to the Cover Bidding Screen of Imhof et al. (2018)

One screen of collusion that is closely related to ours is the cover bidding screen proposed by Imhof et al. (2018). The authors study road construction projects in Switzerland and document a suspicious bidding pattern for a subset of the auctions in which the difference between the two lowest bids is substantially larger than the gap between any two losing bids. Based on this pattern, the authors propose a screen based on a comparison of the gap between the two lowest bids and the standard error of the losing bids. In particular, they consider computing the following statistic for each auction,

$$cover_t = \frac{\mathbf{b}_t^{(2)} - \mathbf{b}_t^{(1)}}{\sigma_t},$$

where $\mathbf{b}_t^{(1)}$ and $\mathbf{b}_t^{(2)}$ are the lowest and second lowest bids in the auction taking place in period t; and σ_t is the standard error of the losing bids in auction t. Imhof et al. (2018) identifies a subset of auctions in which $cover_t$ is consistently above 1 and flag them as potentially uncompetitive. Interestingly, the Swiss competition commission (COMCO) launched an investigation based partly on the results of this statistical test which led to sanctions against eight firms.

The distribution of the gap between the two lowest bids determine the elasticity of residual demand, and hence, the cover bidding screen of Imhof et al. (2018) is related to Proposition 2. Moreover, the distribution of the gap between the two lowest bids determine the upper bound on costs that must hold under competitive bidding. Together with our upper bound on markups, the gap between the lowest bids determine how much power upward deviations have in problem (P) in Section 6. This relates the cover bidding screen to the metric of competitive behavior introduced in Section 6.

Table OC.1 reports the firm-level bounds on the share of competitive auctions the test statistic of Imhof et al. (2018). The left panel of the table corresponds to the top 30 firms in the municipal sample and the right panel corresponds to the top 30 firms in the national sample. Columns (1)-(4) are the same as in Table 2. In column (5), we report for each firm the fraction of auctions it participates in, such that $cover_t$ is above 1. We find that $cover_t$ is above 0.5 for all of the firms except one (firm 4 in the sample of national auction). In fact, except this one firm, we can reject with 95% confidence that $cover_t$ is equal to 0.5. Hence, almost all of the firms in our dataset would be considered as somewhat suspicious by the cover bidding screen.

The overall correlation between the share of competitive histories that we report in column (4) and column (5) is essentially zero across both municipal and national data. If we instead take the correlation between a dummy of whether or not the estimated share of competitive histories is less than 1 and the statistic in column (5), the two statistics become more negatively correlated, with a correlation coefficient of -0.18.

A possible reason why the association between the two measures is somewhat weak is because our measure exploits information from downward deviations as well as from upward deviations. The metric proposed in Imhof et al. (2018) only captures information from upward deviations. This may explain why the correlation between the two is not very strong.[45]

OC.4 Bounds on Other Moments

This appendix shows how to adapt the approach of Section 6 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. Recall from Appendix OB that $\omega_H = (\omega_h)_{h \in H}$ denotes an environment, with $\omega_h = ((d_{h,n})_{n \in \mathcal{M}}, c_h)$. Recall further that $\Omega = \{\omega_H : \forall h \in H, (d_{h,n})_{n \in \mathcal{M}} \in \mathcal{F}\}$ is the set of environments ω_H with feasible demands.

For every environment $\omega_H \in \Omega$, let

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

[45] Another potential reason is that, for the municipal data, the presence of the public reserve price compresses the distribution of the losing bids.

(1) Rank	(2) Participation	(3) Shr won	(4) Shr comp	(5) Shr cover ≥ 1
1	347	0.19	**0.88**	0.69
2	336	0.21	**0.86**	0.74
3	299	0.08	**0.98**	0.79
4	293	0.05	1.00	0.94
5	290	0.14	1.00	0.81
6	287	0.20	1.00	0.70
7	269	0.14	**0.94**	0.72
8	268	0.09	**0.97**	0.75
9	262	0.12	1.00	0.69
10	259	0.18	**0.90**	0.74
11	252	0.12	**0.97**	0.85
12	241	0.12	**0.95**	0.83
13	239	0.16	**0.93**	0.82
14	238	0.09	**0.99**	0.76
15	227	0.11	**0.97**	0.82
16	226	0.12	**0.99**	0.84
17	225	0.08	**0.96**	0.81
18	223	0.12	**0.98**	0.84
19	220	0.07	1.00	0.91
20	218	0.08	1.00	0.93
21	211	0.07	1.00	0.77
22	210	0.14	**0.95**	0.84
23	209	0.17	**0.93**	0.78
24	204	0.15	1.00	0.74
25	203	0.11	**0.98**	0.72
26	199	0.06	1.00	0.75
27	190	0.12	1.00	0.82
28	189	0.06	1.00	0.75
29	188	0.16	**0.94**	0.82
30	187	0.08	1.00	0.85

(a) Municipal Data

(1) Rank	(2) Participation	(3) Shr won	(4) Shr comp	(5) Shr cover ≥ 1
1	4,044	0.17	**0.84**	0.56
2	3,854	0.07	**0.91**	0.82
3	3,621	0.12	**0.85**	0.58
4	2,998	0.15	1.00	0.49
5	2,919	0.06	**0.92**	0.89
6	2,547	0.08	**0.71**	0.76
7	2,338	0.07	**0.74**	0.79
8	2,333	0.07	**0.74**	0.77
9	2,328	0.04	**0.95**	0.60
10	2,292	0.06	**0.75**	0.77
11	2,237	0.08	**0.90**	0.59
12	2,211	0.03	**0.96**	0.62
13	2,015	0.09	**0.76**	0.78
14	1,984	0.08	**0.75**	0.75
15	1,727	0.07	1.00	0.81
16	1,674	0.05	**0.84**	0.80
17	1,661	0.03	**0.94**	0.87
18	1,660	0.08	**0.75**	0.78
19	1,589	0.07	**0.79**	0.77
20	1,427	0.10	1.00	0.58
21	1,393	0.06	**0.86**	0.83
22	1,392	0.07	1.00	0.76
23	1,370	0.04	**0.92**	0.86
24	1,368	0.14	1.00	0.54
25	1,353	0.05	**0.80**	0.80
26	1,342	0.09	1.00	0.72
27	1,337	0.04	**0.87**	0.85
28	1,326	0.08	**0.92**	0.75
29	1,291	0.06	**0.86**	0.83
30	1,260	0.06	**0.93**	0.74

(b) National Data

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. Column 4 present our 95% confidence bound on the share of competitive histories for each firm based on Proposition 3. For our estimates of column 5, we use deviations $\{-0.02, 0, 0.001\}$, minimum markup $m = 0.025$ and maximum markup $M = 0.5$.

Table OC.1: Share of competitive histories, individual firms be the set of competitive auctions under ω_H. Consider the following program:

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

where, for each n and each ω_H, $D_n(\omega_H, H) = \frac{1}{|H|}\sum_{h\in H} d_{h,n}$. $\widehat{s}_{\mathsf{auc}}$ 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 demand and costs ω_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:

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

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 OC.6 reports estimates for firms in the city of Tsuchiura, as a function of minimum markup m.

OC.5 Sensitivity to Economic Plausibility Constraints

Figure OC.7 shows that, for our city-level data, our estimates on the share of competitive histories are insensitive to changes in maximum markup M. Figure OC.8 illustrates the

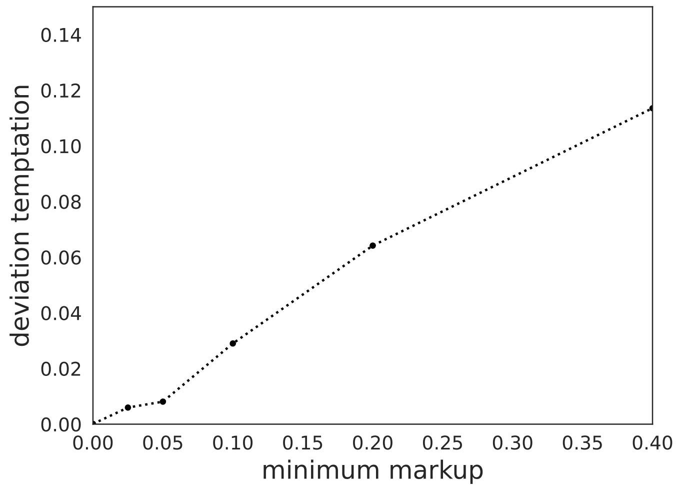


Figure OC.6: Total deviation temptation as a fraction of profits, Tsuchiura. Deviations $\{-.02, 0, .001\}$. Maximum markup 0.5.

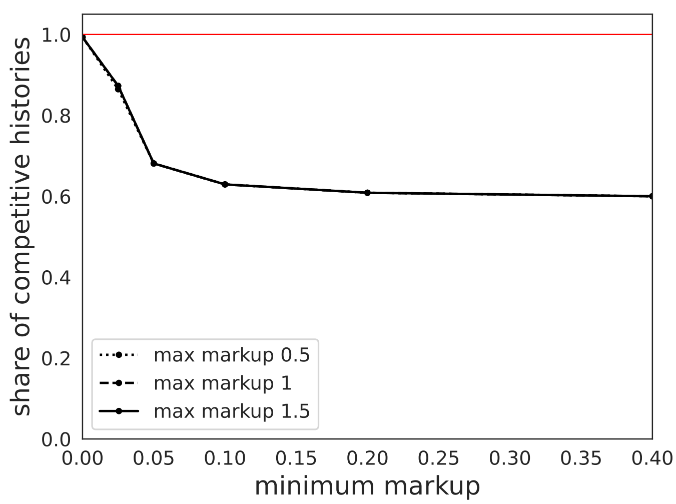


Figure OC.7: Share of competitive histories for different maximum markups, city data, deviations $\{-0.02, 0, 0.001\}$.

sensitivity of our estimates to parameter $\alpha \in [0, 1]$ in downward deviation IC constraint (O3) for auctions with re-bidding. Recall that parameter α 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.

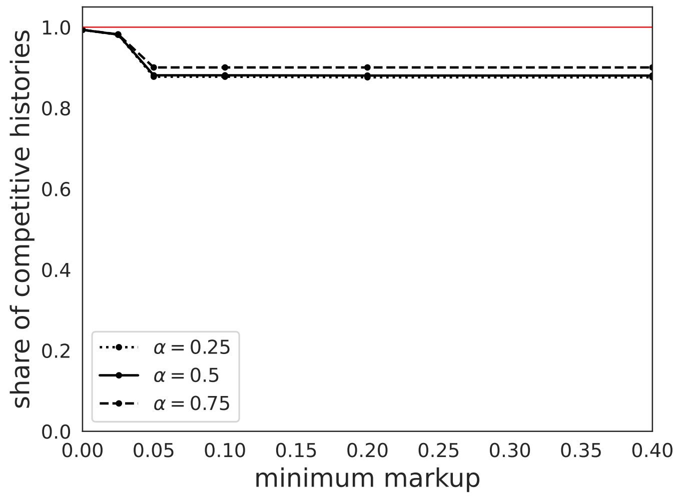


Figure OC.8: Share of competitive histories, national-level data. Deviations $\{-0.02, 0, 0.001\}$, $M = 0.5$.

OD Proofs for Lemmas 1, 2 and 3

Proof of Lemma 1. Let us first establish that problem (P) does admit a solution. Since $\mathcal{F}$ is a compact subset of $\mathbb{R}^{\mathcal{M}}$, Prokhorov's theorem implies the set $\Delta(\mathcal{F})$ of distributions over $\mathcal{F}$ is compact under the weak topology. Since $\mathcal{D}_\alpha$ is compact and IsComp is upper semi-continuous, it follows that

$$\sup_{\mu \in \Delta(\mathcal{F})} \mathbb{E}_\mu[\text{IsComp}(\mathbf{d})]$$
$$\text{s.t. } \mathbb{E}_\mu[\mathbf{d}] \in \mathcal{D}_\alpha$$

does admit a solution, and the supremum is in fact a maximum. Let us denote by $\widehat{\mu}$ this solution.

Let us denote by C the set of competitive belief profiles $\mathbf{d}$ satisfying (IC-MKP), and by C^{o} the interior of set C, i.e. the set of beliefs $\mathbf{d}$ satisfying (IC-MKP) with strict inequalities.

Since $\mathcal{F}_0^n$ becomes dense in $\mathcal{F}$ as n grows, it follows that $\mathcal{F}_0^n \cap C^{\text{o}}$ becomes dense in C^{o}. Since C is equal to the closure of C^{o}, it follows that $\mathcal{F}_0^n \cap C^{\text{o}}$ becomes dense in C.

Pick $\epsilon > 0$. Since C is compact, it is covered with finitely many balls of radius ϵ, and since $\mathcal{F}_0^n \cap C^{\text{o}}$ becomes dense in C, for n large enough and for all $\mathbf{d} \in C$, there exists $\mathbf{d}' \in \mathcal{F}_0^n \cap C^{\text{o}}$ such that $||\mathbf{d} - \mathbf{d}'|| \leq \epsilon$, where $|| \cdot ||$ is the Euclidean distance on $\mathbb{R}^{\mathcal{M}}$.

Hence, for n large enough and for every $\mathbf{d} \in \mathsf{supp}\, \widehat{\mu} \cap C$ we can associate $f(\mathbf{d}) = \mathbf{d}' \in$

$\mathcal{F}_0^n \cap C^{\mathrm{o}}$ such that $||\mathbf{d} - f(\mathbf{d})|| \leq \epsilon$.

Let $\mathbf{d}_{\neg C} \equiv \mathbb{E}_{\widehat{\mu}}[\mathbf{d} \,|\, \mathbf{d} \notin C]$ denote the weighted average of $\mathbf{d}$ under $\widehat{\mu}$ conditional on $\mathbf{d} \notin C$ (i.e. $\mathbf{d}$ not competitive). For n large enough, there exists $f(\mathbf{d}_{\neg C}) \in \mathcal{F}_0^n$ such that $||f(\mathbf{d}_{\neg C}) - \mathbf{d}_{\neg C}|| \leq \epsilon$.

Similarly, consider the sample demands $\widehat{\mathbf{D}}$. For n large enough, we can find $f(\widehat{\mathbf{D}}) \in \mathcal{F}_0^n$ such that $||f(\widehat{\mathbf{D}}) - \widehat{\mathbf{D}}|| \leq \epsilon$. We assume for simplicity that $\widehat{\mathbf{D}} \neq \mathbf{d}_{\neg C}$.

Finally consider distributions $\mu_n \in \Delta(\mathcal{F}_0^n)$ such that

$$\mu_n \circ f(\widehat{\mathbf{D}}) = \nu$$
$$\mu_n \circ f(\mathbf{d}_{\neg C}) = (1-\nu) \times \widehat{\mu}(\mathcal{F} - C)$$
$$\forall \mathbf{d} \in C, \mu_n \circ f(\mathbf{d}) = (1-\nu) \times \widehat{\mu}(\mathbf{d})$$

where $\nu > 0$.[46]

For each $\nu > 0$, we can find $\epsilon > 0$ small and $\overline{N}$ large such that, for all $n \geq \overline{N}$, $\mathbb{E}_{\mu_n}[\mathbf{d}] \in \mathcal{D}_\alpha$. In addition, by construction $\mu_n \in \Delta(\mathcal{F}_0^n)$ and $\mathbb{E}_{\mu_n}[\text{IsComp}(\mathbf{d})] \geq (1-\nu) \times \mathbb{E}_{\widehat{\mu}}[\text{IsComp}(\mathbf{d})] = (1-\nu)\widehat{s}_{\mathsf{comp}}$. Since $\widehat{s}^n_{\mathsf{comp}} \leq \widehat{s}_{\mathsf{comp}}$ by construction, this implies that $\lim_{n\to\infty} \widehat{s}^n_{\mathsf{comp}} = \widehat{s}_{\mathsf{comp}}$. ■

Proof of Lemma 2. Along the lines of the proof of Proposition 1, we define

$$\varepsilon_t \equiv \sum_{h_{i,t}} \left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \quad \text{and} \quad S_T \equiv \sum_{t=0}^{T} \varepsilon_t,$$

where for each history $h_{i,t}$, $\widehat{\mathbf{d}}_{h_{i,t}} = (\mathbf{1}_{\wedge \mathbf{b}_{-i,t} > (1+\rho_n) b_{i,t}})_{n \in \mathcal{M}}$. S_T is a sum of martingale increments ε_t whose absolute value is bounded by $||\lambda||_1 \overline{N}_t$, where $||\lambda||_1 \equiv \sum_{n\in\mathcal{M}} |\lambda_n|$ and $\overline{N}_t$ denotes the number of bidders participating at time t with histories in H.

The Azuma-Hoeffding inequality implies that

$$\mathsf{prob}(S_T > x_\lambda |H|) \leq \exp\left(- \frac{x_\lambda^2 |H|^2}{2||\lambda||_1^2 \sum_{t=0}^T \overline{N}_t^2} \right).$$

Observing that $\sum_{t=0}^T \overline{N}_t^2 \leq \sum_{t=0}^T \overline{N}_t N_{\max} = N_{\max}|H|$, this implies that

$$\mathsf{prob}(|S_T| > \nu |H|) \leq \exp\left(- \frac{x_\lambda^2 |H|}{2||\lambda||_1^2 N_{\max}} \right).$$

[46]If $\widehat{\mathbf{D}} = \mathbf{d}_{\neg C}$, we can define μ_n to be such that: $\mu_n \circ f(\mathbf{d}_{\neg C}) = \nu + (1-\nu) \times \widehat{\mu}(\mathcal{F} - C)$, and such that $\forall \mathbf{d} \in C, \mu_n \circ f(\mathbf{d}) = (1-\nu) \times \widehat{\mu}(\mathbf{d})$.

This concludes the proof. ■

Proof of Lemma 3. As in the proof of Lemma 2, we define

$$\varepsilon_t \equiv \sum_{h_{i,t}} \left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \quad \text{and} \quad S_T \equiv \sum_{t=0}^{T} \varepsilon_t = |H| \left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle .$$

S_T is a sum of martingale increments ε_t whose absolute value is bounded by $||\lambda||_1 N_{\max}$. This implies that the central limit theorem for sums of martingale increments holds (Billingsley (1995), Theorem 35.11):

$$\lim_{T\to\infty} \mathsf{prob}\left(\frac{1}{\sigma_\lambda \sqrt{T+1}} S_T \geq x \right) = 1 - \Phi(x) = \Phi(-x)$$

with

$$\sigma_\lambda \equiv \sqrt{\frac{1}{T+1} \sum_{t=0}^{T} \mathsf{var}\left(\sum_{h_{i,t}\in H} \left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \Big| h_t^0 \right)}.$$

We cannot directly exploit this result to get explicit bounds on the distribution of $\left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle$ because $\mathbf{d}_{h_{i,t}}$ is not directly observed, so that we can't form a consistent estimate of σ_λ. Instead, we show that

$$\widehat{\sigma}_\lambda = \sqrt{\frac{1}{T+1} \sum_{t=0}^{T} |\{h_{i,t} \in H\}| \sum_{h_{i,t}\in H} \left\langle \lambda, \widehat{\mathbf{D}}_t - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle^2}.$$

can be used as an asymptotic upper bound to σ_λ.

Indeed, for any period t, Jensen's inequality implies that

$$\begin{aligned} \mathsf{var}\left(\sum_{h_{i,t}\in H} \left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \Big| h_t^0 \right) &= |\{h_{i,t} \in H\}|^2 \mathsf{var}\left(\frac{1}{|\{h_{i,t} \in H\}|} \sum_{h_{i,t}\in H} \left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \Big| h_t^0 \right) \\ &\leq |\{h_{i,t} \in H\}| \sum_{h_{i,t}\in H} \mathsf{var}\left(\left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \Big| h_t^0 \right). \end{aligned}$$

Furthermore, since $\mathbf{d}_{h_{i,t}} = \mathbb{E}[\widehat{\mathbf{d}}_{h_{i,t}} | h_{i,t}]$, and since $h_{i,t}$ includes all the information provided in history h_t^0, it follows that for any h_t^0-measurable random variable $Z_t \in \mathbb{R}^{\mathcal{M}}$,

$$\mathsf{var}\left(\left\langle \lambda, \mathbf{d}_{h_{i,t}} - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle \Big| h_t^0\right) \leq \mathbb{E}\left[\left\langle \lambda, Z_t - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle^2 \Big| h_t^0\right].$$

The Law of Large Number for martingale increments implies that almost surely,

$$\lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \sum_{h_{i,t}\in H} \mathbb{E}\left[\left\langle \lambda, Z_t - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle^2 \Big| h_t^0\right] - \left\langle \lambda, Z_t - \widehat{\mathbf{d}}_{h_{i,t}} \right\rangle^2 = 0.$$

Hence, setting $Z_t = \widehat{\mathbf{D}}_t$, it follows that for any $\epsilon > 0$, almost surely as T becomes large, $(1+\epsilon)\widehat{\sigma}_\lambda \geq \sigma_\lambda$. Since $x > 0$, this implies that

$$\begin{aligned} \limsup \mathsf{prob}\left(\frac{1}{\widehat{\sigma}_\lambda \sqrt{T+1}} S_T \geq x\right) &\leq \limsup \mathsf{prob}\left(\frac{1}{\sigma_\lambda \sqrt{T+1}} S_T \geq x(1+\epsilon)\right) \\ &= \Phi(-x(1+\epsilon)) \end{aligned}$$

Observing that $\frac{1}{\widehat{\sigma}_\lambda\sqrt{T+1}} S_T = \frac{|H|}{\widehat{\sigma}_\lambda\sqrt{T+1}} \left\langle \lambda, \mathbb{E}_{\mu^*}[\mathbf{d}] - \widehat{\mathbf{D}} \right\rangle$, Lemma 3 follows from the fact that the result holds for any $\epsilon > 0$, and Φ is continuous. ■