

# An Equilibrium Model of the First-Price Auction with Strategic Uncertainty: Theory and Empirics

Bernhard Kasberger\*

March 28, 2022

## Abstract

In many first-price auctions, bidders face considerable strategic uncertainty: they cannot perfectly anticipate the other bidders' bidding behavior. We propose a model in which bidders do not know the entire distribution of opponent bids but only the expected winning bid and lower and upper bounds on the opponents' bids. We characterize the optimal bidding strategies and prove the existence of a novel equilibrium. Finally, we apply the model to estimate the cost distribution in highway procurement auctions. Our model fits the data better than existing approaches in which the entire distribution of opponent bidders is assumed to be known.

*JEL Classification:* C57, C72, D44, D81.

*Keywords:* First-price auctions, independent private value, non-parametric identification, moment constraints, strategic uncertainty, ambiguity, minimax regret, minimax loss.

---

\*Düsseldorf Institute for Competition Economics (DICE), Heinrich Heine University Düsseldorf, Germany

Email: [bernhard.kasberger@hhu.de](mailto:bernhard.kasberger@hhu.de)

I thank David Delacrétaz, Matteo Escudé (discussant), Simon Finster, Paul Heidhues, Stefan Hubner, Maarten Janssen, Willemien Kets, Paul Klemperer, Simon Martin, Martin Pesendorfer, Karl Schlag and audiences at Berlin, Bonn, Düsseldorf, Manchester, Mannheim, NYU Abu Dhabi, Oxford, Vienna, the 2020 Conference on Mechanism and Institution Design, the 2020 INFORMS Annual Meeting, the 2020 European Winter Meeting of the Econometric Society, the 2021 Spring Meeting of Young Economists, and the 2021 Annual Meeting of the Austrian Economic Association for their helpful comments and suggestions. All errors are my own.

# 1 Introduction

This paper proposes a novel game-theoretic equilibrium model of the first-price auction in which observable statistics characterize the players' beliefs. In particular, bidders have a belief about the expected winning bid, a statistic that is observable in many auction markets. Indeed, households interested in bidding for a house can look up the average sales price in a neighborhood but typically not the history of transaction prices. In procurement, bidders may have a small sample of comparable transaction prices and corresponding covariates but do not observe losing offers. Bidders can then estimate the expected winning bid but not much more. Finally, while not necessarily a first-price auction, Google informs bidders about expected prices in ad auctions but keeps the bid history secret.<sup>1</sup> In addition to the belief about the expected winning bid, bidders also identify lower and upper bounds on the opponents' bids in our model. These can be, for example, prices in the next worst or next best neighborhoods. These three statistics, the lowest opponent bid, the expected winning bid, and the highest opponent bid, independence, and the number of bidders are all a bidder knows about the opponent bid distribution in our model.

A key characteristic of the model is strategic uncertainty, as many bid distributions are consistent with the observed statistics. Due to this uncertainty about the opponent bid distribution, bidders cannot maximize the (subjective) expected payoffs. Instead, any bid will likely lead to a lower payoff than the optimal bid had the bidder known the opponent bid distribution. We quantify this (interim) suboptimality with a loss function; the loss of a bid given the opponents' bid distribution is the difference between the highest possible payoff and the payoff from the bid. It is then natural to desire a low loss. After all, if the loss of a bid is small, the payoff is close to optimal. Therefore, we assume that bidders choose bids that minimize the maximal loss, leading to a close-to-optimal payoff for any possible bid distribution (Savage, 1951).

The difference to traditional game-theoretic equilibrium models is that the bidders know much less about their opponents, and what they know is, in principle, observable. In a Bayes-Nash equilibrium, the predominant game-theoretic tool for analyzing the first-price auction, the bidders' value distribution is assumed to be common knowledge, and bidders know the opponents' bidding strategies in equilibrium. Neither is typically observable. Usually, only (statistics of) the bids are

---

<sup>1</sup>There are also other settings in which averages may characterize beliefs. For example, central banks communicate that they want an average inflation rate of 2%. In a different context, there is a (known) target for the average block size on the Ethereum blockchain (Roughgarden, 2020).

observable. In our model, there is no common prior as bidders are ignorant of the value distribution and opponents' strategies; they reason directly about the bid distribution. Their beliefs take the form of statistics of the bid distribution. In contrast, bidders know the entire bid distribution in a Bayes-Nash equilibrium; that is, they can compute for each bid the probability that it wins the auction.<sup>2</sup> We maintain correct beliefs but dispense with the assumption that the entire distribution is known.

We introduce with *aggregate moment equilibrium* a novel equilibrium notion. In an aggregate moment equilibrium, the players play optimally given their beliefs, and the beliefs are correct given the optimal strategies and the latent (independent) value distribution. We prove that an aggregate moment equilibrium exists and that the beliefs induce bidding functions that are strictly increasing in the bidder's private value (Theorem 1). The equilibrium beliefs naturally depend on the latent value distribution and are such that the lowest bid by any type is the belief about the lowest bid, the highest bid by any type is the belief about the highest bid, and the expected winning bid equals the belief about the expected winning bid. The features of optimal behavior given beliefs and consistent beliefs given optimal behavior and the latent value distribution are shared with Bayes-Nash equilibrium.

Instead of observing the average winning bid, bidders may observe the average bid. Hence, we also present a closely related but alternative model in which bidders have a belief not about the expected winning bid but rather the expected opponent bid (in addition to the lowest and highest possible bids). Such a belief might come from an observed list of past bids and corresponding auction covariates. Suppose bidders use the sample to estimate a (possibly linear) relationship between the bids and the covariates.<sup>3</sup> Plugging in the covariates of the current auction gives an estimate of the expected opponent bid.<sup>4</sup> Alternatively, bidders might directly observe the average bid. For concreteness, we assume that each bidder believes that the expected bids of the other bidders are the same. We again define an equilibrium notion (*moment equilibrium*) in which beliefs are consistent with optimal behavior and the latent value distribution. We characterize the minimax-loss bidding func-

---

<sup>2</sup>A survey among academic and professional auction consultants shows that many bidders in high-stake auctions struggle to assign probabilities to opponent bids and values (Kasberger and Schlag, 2022).

<sup>3</sup>Due to the auction heterogeneity, it might not be best to naively best respond to the sample as players do in a sampling equilibrium (Osborne and Rubinstein, 2003). Salant and Cherry (2020) consider players who use a sample estimate the distribution of actions. In our model, players would only estimate the first moment due to the covariates. We implicitly assume that the sample is small in relation to the number covariates.

<sup>4</sup>The survey of Kasberger and Schlag (2022) reveals that some real-world bidders estimate the expected opponent bids.

tion and prove that a moment equilibrium exists in which the minimax loss bidding function is strictly increasing in the bidder’s private value (Theorem 2). To emphasize that beliefs are now directly about individual behavior, we sometimes refer to this moment equilibrium as the individual moment equilibrium.

The difference between the two types of belief is that with individual moment beliefs, bidders do not face the inference problem that they face with aggregate moment beliefs. With aggregate moment beliefs, bidders have to translate their belief about the expected winning bid to a belief about the distribution of the highest opponent bid. To do so, we assume that bidders believe that their belief about the expected winning bid comes from auctions with  $n$  bidders whose bids are i.i.d. draws from some individual bid distribution. In the current auction are also  $n$  bidders, so a bidder’s payoff depends on the distribution of the maximum of the  $n - 1$  highest bid. Bidders then consider all individual bid distributions as possible for which the expected value of the highest of  $n$  independent draws is their belief about the expected winning bid. They then use independence to identify the worst-case distribution of the highest of  $n - 1$  independent bids. With moment beliefs directly about the individual bid distributions, the first inference is unnecessary.

Next, we show how we can use our equilibrium models with beliefs in the form of observable statistics for the structural estimation of the latent value distribution. We prove analogous results for both types of belief but now discuss the estimation for the model with a belief about the expected winning bid. The backbone of the structural estimation is a non-parametric identification result: the bid distribution identifies the value distribution if bidders play the strictly increasing aggregate moment equilibrium strategies. The first step in the estimation estimates the players’ beliefs non-parametrically by computing the sample minimum and maximum bids as well as the average winning bid. The second step uses the estimated beliefs and the characterized minimax-loss bidding function to translate bids into “pseudo values.” In contrast, the first step in the leading non-parametric estimator based on the Bayes-Nash equilibrium as proposed by Guerre, Perrigne, and Vuong (2000) [GPV] requires specifying a kernel and a bandwidth to estimate the probability density function of the bid distribution. The second step in GPV uses the first-order condition of expected utility maximization to translate bids into pseudo values. A notable difference is how easy it is to estimate the bidders’ beliefs in the moment-equilibrium-based approach as one only needs to compute some summary statistics. While both estimators are consistent (given that the respective theoretical model generates the data), we expect the moment-equilibrium-based estimator to perform well in small samples as the estimation of the beliefs is super-consistent (Donald

and Paarsch, 2002).

Finally, using the data of Bajari, Houghton, and Tadelis (2014), we estimate the latent cost distribution in Californian highway procurement auctions non-parametrically. Out-of-sample predictions reveal that our equilibrium approaches perform empirically at least as well as Bayes-Nash equilibrium. In these out-of-sample predictions, we compare five estimators for estimating the latent cost distribution. On the one hand, we use the estimators based on aggregate and individual moment beliefs, respectively, each time with and without removing (high losing) outliers based on the interquartile range. On the other hand, we use the Bayes-Nash-equilibrium-based estimator of GPV as a benchmark. For each number of bidders per auction  $n$ , we use the auctions with a different number of bidders to estimate the cost distribution and then compute the equilibrium bid distribution had there been  $n$  bidders. Finally, we assess the fit by computing the distance between the sample and the predicted bid distribution. Across the specifications, we find that robust models that do not make strong assumptions about bidders' knowledge have higher explanatory power of the observed data than BNE. When all bidders are large corporations, the aggregate moment equilibrium performs best. When all bidders are small, the individual moment equilibrium performs best. In that case, removing outliers for estimating the highest possible bid turns out to be particularly useful because there are some unduly high bids made by fringe firms in the sample.

## 1.1 Related Literature

The paper stands in the tradition of the so-called Wilson (1987) doctrine that seeks to weaken the common knowledge assumptions made in analyzing practical problems with game-theoretic methods. The extensive literature in this tradition has mostly focused on robust mechanism design. Ex-post implementation weakens the common prior but maintains the assumption that players know the other players' strategies (Bergemann and Morris, 2005). Note that the first-price auction does not admit an ex-post equilibrium. Other approaches consider non-Bayesian designers dealing with uncertainty (ambiguity) about aspects of the environment by minimizing the maximal regret (Bergemann and Schlag, 2011; Guo and Shmaya, 2019, 2021) or maximizing worst-case expected utility (Carroll, 2015; Carroll and Segal, 2019). We consider non-Bayesian players and define an equilibrium notion when these players do not have a common prior and do not know other players' strategies.

Other papers that define equilibrium concepts for games with non-Bayesian players facing ambiguity broadly fall into two strands. A first strand considers games of complete information and players being uncertain about the strategies of

the other players. A second strand considers games of incomplete information with players being uncertain about the type (i.e., value) distribution. The first strand is more related as the second assumes that players know the other players' strategies in equilibrium.<sup>5</sup> Within the first strand, Lehrer (2012) defines with *partially specified equilibrium* a related equilibrium notion for games of complete information. In his model, similar to the moment and range, players' beliefs are characterized by expectations of random variables that depend on the mixed strategies of the other players. However, in a first-price auction, forming beliefs about strategies is only useful if one also has a belief about the value distribution. Such knowledge conflicts with our assumption that bidders' beliefs are characterized by observable statistics; the value distribution is typically unobserved.<sup>6</sup>

There are also other solution concepts for games of complete information with equilibrium strategic uncertainty in the sense of multiple priors. These concepts differ in how they determine the belief—in the form of a set of other players' possible strategies—in equilibrium. Lo (1996) has the restrictive requirement that only optimal mixed strategies are deemed possible. Klibanoff (1996) and Renou and Schlag (2010) demand that the optimal strategy is just one of the strategies deemed possible, which is quite permissive as they do not determine the entire belief set in equilibrium.<sup>7</sup> In our model, the belief set does not contain strategies (actions) but rather distributions of actions (bids). Moreover, we parameterize the belief set with the moment and range beliefs. Similar to Lo (1996), all possible bid distributions reflect the truth in equilibrium by having the same mean as the true bid distribution in our model. At the same time, echoing Klibanoff (1996) and Renou and Schlag (2010), the equilibrium bid distribution is just one of the possible.

There are other papers in economics (Ollár and Penta, 2017; Carrasco, Luz, Kos, Messner, Monteiro, and Moreira, 2018) and in operations research (Smith, 1995; Popescu, 2005; Perakis and Roels, 2008) that use moment constraints. These papers do not determine them in equilibrium.

Kasberger and Schlag (2022) also analyze the first-price auction under strategic

---

<sup>5</sup>Lo (1998), Chen, Katuščák, and Ozdenoren (2007), and Gretschko and Mass (2018) analyze the first-price auction and belong to the second category as the players are uncertain about the value distribution, but there is no strategic uncertainty in equilibrium. Hyafil and Boutilier (2004) and Schlag and Zapechelnyuk (2020) do not analyze first-price auctions but use minimax loss to deal with the unknown type distribution.

<sup>6</sup>Another difference is the decision criterion. Lehrer (2012) uses maximin expected utility, whereas we use minimax loss. Maximin expected utility does not lead to sharp predictions in the first-price auction with moment beliefs. Using smooth ambiguity, Battigalli, Cerreia-Vioglio, Maccheroni, and Marinacci (2015) propose a related solution concept (cf. Battigalli, Cerreia-Vioglio, Maccheroni, and Marinacci, 2016).

<sup>7</sup>Other solution concepts in this vein are by Dow and Werlang (1994), Eichberger and Kelsey (2000), and Marinacci (2000) for players maximizing Choquet expected utility (Schmeidler, 1989).

uncertainty with minimax loss as the decision criterion. Their objective is not to explain behavior but to find bidding rules that work well for a single bidder. Without using moment beliefs, they construct the set of possible bid distributions for a single bidder; we determine the set in equilibrium. Empirically, they find that their bidding rules would have led to a higher expected utility, on average, which implies that their approach is not well suited for explaining observed behavior. Note that one can use the minimax loss bidding functions of the current paper in such a normative way. Linhart and Radner (1989) is an early paper that uses minimax loss for the analysis of a strategic decision problem (bilateral bargaining). However, they do not define an equilibrium.

## 2 The Model

In a nutshell, the  $n$  bidders in a first-price auction have independent private values and a belief about the lowest possible bid, the highest possible bid, and the expected winning bid. As there are many bid distributions that are consistent with these three statistics, the bidders face (strategic) uncertainty. They deal with the uncertainty by minimizing the maximal loss. The beliefs are correct in equilibrium.

### 2.1 Basics

A single good is for sale in a first-price sealed-bid auction. The auction rules are such that all bidders simultaneously submit sealed bids. The bidder with the highest bid wins and pays her bid. In the case of ties, the good is allocated uniformly between the bidders with the highest bid. There are  $n$  bidders in the auction,  $n \geq 2$ . Let  $N$  denote the set of players.

The bidders are risk neutral. Before the auction, each bidder  $i$  learns her willingness-to-pay  $v_i$  for the auctioned item. These private values are independent, identically distributed, and drawn from a distribution with the cumulative distribution function  $F$ . The lowest and the highest value in the support are denoted by  $\underline{v}$  and  $\bar{v}$ , respectively, with  $0 \leq \underline{v} < \bar{v} < \infty$ . Let  $\text{supp}(F)$  denote the support of  $F$ . The payoff from losing is normalized to 0. The payoff from winning with bid  $b_i$  equals  $v_i - b_i$ . Hence, conditional on a profile of bids  $(b_1, \dots, b_n)$ , the expected utility is equal to

$$u_i(b_1, \dots, b_n) = \begin{cases} \frac{1}{k}(v_i - b_i) & \text{if } b_i = \max_{j \in N} b_j \text{ and } k = |\{j \in N : b_j = b_i\}| \\ 0 & \text{if } b_i < \max_{j \in N} b_j. \end{cases}$$

Bidders choose a bid conditional on their willingness-to-pay. From an interim perspective, bidder  $i$ 's behavior takes the form of choosing a strategy  $\beta_i : \text{supp}(F) \rightarrow \mathbb{R}_+$  that maps values into bids. The strategy is measurable with respect to the underlying Borel  $\sigma$ -algebra. Importantly, the value distribution  $F$  and the bidding strategy  $\beta_i$  generate a bid distribution  $B_i \in \mathcal{P}$ , where  $\mathcal{P}$  is the set of measures on  $\mathbb{R}_+$  (defined on the Borel  $\sigma$ -algebra). Bidders bid independently so that the distributions  $B_i$  and  $B_j$  are independent for  $i \neq j$ . The set of probability distributions on set  $X$  is denoted by  $\Delta X$ .

The bid distributions of the other bidders determine the expected utility of a bid  $b_i$ . Formally, for joint bid distributions  $B_{-i} \in \Delta \mathbb{R}_+^{n-1}$ , the expected utility  $U_i(b_i, B_{-i})$  of bidder  $i$  while bidding  $b_i$  is given by

$$U_i(b_i, B_{-i}) = \int u_i(b_i, b_{-i}) dB_{-i}(b_{-i}).$$

In the absence of ties and with independent bid distributions,  $\prod_{j \neq i} B_j(b_i)$  is the probability that  $b_i$  is higher than all of the other bidders' bids. Then the expected utility is  $U_i(b_i, B_{-i}) = \prod_{j \neq i} B_j(b_i)(v_i - b_i)$ .

Bidders know that there are  $n$  bidders in total, that values are independent draws, and that everyone bids independently. Each bidder believes that the other players are ex-ante identical, i.e., that their bids come from the same bid distribution.

The bidders do not know the value distribution (not even the support) and do not know the strategies of the other players.<sup>8</sup> Instead of forming beliefs about the value distribution and the strategies, they form beliefs directly about the bid distribution, the object that determines the expected utility. From bidder  $i$ 's perspective, the bid of bidder  $j$  is a random variable with distribution  $B \in \mathcal{P}$ .

Bidder  $i$ 's belief about the other bidders' behavior takes the form of the set of all joint bid distributions  $B_{-i}$  that she thinks are possible. Let  $\mathcal{B}_i \subseteq \Delta \mathbb{R}_+^{n-1}$  denote this set of possible bid distributions (the belief set). Thus, the belief is a set of multiple priors. Note that the agents are not Bayesians (Savage, 1954), i.e., they do not have a probabilistic prior over the set of possible bid distributions. We put structure on the players' belief sets after defining how players choose their bids for a given belief set.

Bidder  $i$  deals with the uncertainty about the opponent bid distribution by minimizing the maximal loss (Savage, 1951). Given a bid distribution, the loss of

---

<sup>8</sup>Bidders using their belief about the value distribution would impose additional constraints on the set of possible bid distributions. It would be interesting to study the interaction of beliefs about the value distribution and the bid distribution in future work.



a bid is the difference between the highest possible expected utility and the actual payoff. Formally, the loss of bidding  $b_i$  is

$$\lambda_i(b_i, B_{-i}|v_i) = \sup_{\tilde{b}_i \in \mathbb{R}_+} U_i(\tilde{b}_i, B_{-i}) - U_i(b_i, B_{-i}).$$

Player  $i$  chooses the bid  $b_i$  to minimize the maximal loss, i.e.,

$$b_i \in \arg \inf_{\tilde{b}_i \in \mathbb{R}_+} \sup_{B_{-i} \in \mathcal{B}_i} \lambda_i(\tilde{b}_i, B_{-i}|v_i).$$

Minimax loss seeks good performance for all possible bid distributions. As such, it qualifies as *robust*. If the loss of a bid is small, then the bid is close to optimal for the given bid distribution. If maximal loss is small, then one loses little payoff due to not knowing the bid distribution. According to Linhart and Radner (1989), minimax loss can be used to find compromises in group decisions with diverse subjective probability judgments. Minimax loss bids are easy to justify as one can lay down the objective facts that characterize the belief set and then choose the bid that performs well for all possible bid distributions. Note that we adopt an interim perspective and compute loss conditional on the opponent bid distribution, as in Kasberger and Schlag (2022).<sup>9</sup>

An alternative to minimax loss is maximin expected utility (Gilboa and Schmeidler, 1989). In contrast to minimax loss, it does not lead to sharp predictions as all bids are optimal for low values. To see this, consider a bidder with a willingness-to-pay below the expected winning bid. The worst case occurs when one loses the auction. A corresponding worst-case bid distribution puts all the mass on the expected winning bid. Worst-case expected utility is then maximized by any bid below the willingness-to-pay. Another alternative is to assume that bidders have a probabilistic belief about which bid distribution materializes, in which case the beliefs are no longer characterized by statistics.

## 2.2 Aggregate Moment Belief

The belief (set)  $\mathcal{B}_i$  of bidder  $i$  about the other bidders' bidding behavior is characterized by two types of belief. First, bidder  $i$  believes that other bidders only choose bids in  $[l_i, u_i]$ , where  $0 \leq l_i < u_i$ . Note that bidder  $i$  forms identical beliefs about the other bidders, i.e.,  $(l_i, u_i)$  is independent of  $j$ . Second, bidder  $i$  believes that in

---

<sup>9</sup>An alternative would be to compute loss ex-post conditional on the opponents' realized bids. To distinguish these two perspectives, Kasberger and Schlag (2022) refer to the interim concept as loss and to the ex-post concept as regret.

auctions like the current one the average winning bid is  $m_i$ . Thus, the other bidders bid in such a way that the expected value of the maximum of  $n$  independent draws from their bid distributions is  $m_i$ . We refer to the first type of belief as *range belief* and to the second type as (aggregate) *moment belief*.

The range and moment beliefs parameterize the belief set, which is from now on denoted by  $\mathcal{B}_i(l_i, m_i, u_i)$ . Let  $G$  be the distribution function of a real-valued random variable and let  $G^n$  denote the distribution of the maximum of  $n$  independent draws from  $G$ . The set of possible bid distributions is

$$\mathcal{B}_i(l_i, m_i, u_i) = \{B_{-i} \in \Delta \mathbb{R}_+^{n-1} : \exists B \in \mathcal{P} \text{ with } B_{-i}(b_{-i}) = \prod_{j \neq i} B_j(b_j), \\ \int_{l_i}^{u_i} dB = 1 \text{ and } \int x dB^n(x) = m_i\}.$$

Only bid distributions that put mass one on  $[l_i, u_i]$  are possible. Moreover, the highest bid of  $n$  independent bids must be  $m_i$  in expectation.

In first-price auctions, the expected utility depends on the highest competing bid, that is, the highest bid of  $n - 1$  bids. In the current model, bidders use the belief about the maximum of  $n$  bids to form beliefs about the distribution of individual bids and then use the distribution of individual bids to form beliefs about the distribution of the highest bid of  $n - 1$  bids. The key assumptions are that the number of bidders  $n$  is known (that “generated”  $m_i$  and that participate in the current auction), and that the competing bids are i.i.d. There are clearly interesting alternative models.<sup>10</sup> The current paper focuses on the case of known  $n$  and i.i.d. bids to keep the model simple, close to the standard textbook case, and applicable for the structural estimation.

### 2.3 Aggregate Moment Equilibrium

The model is closed by assuming that the beliefs are not systematically wrong. The beliefs are consistent or in equilibrium if the beliefs match the statistics of the bid distribution generated by the latent value distribution and the optimal strategies.

Let  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$ ,  $l = (l_i)_{i \in N}$ ,  $m = (m_i)_{i \in N}$ , and  $u = (u_i)_{i \in N}$ .

**Definition 1.** The tuple  $(\beta, l, m, u)$  is an *aggregate moment equilibrium* if

---

<sup>10</sup>In our model, bidders are sophisticated as they translate their belief about the expected winning to a (worst-case) belief about the distribution of the highest bid of  $n - 1$  bidders. Alternatively, bidders could be naive in that they ignore the necessary inference and act as if the highest opponent bid is  $m$  on average. As the number of bidders grows, the difference between naive and sophisticated bidders disappears.

1.  $\beta_i$  minimizes the maximal loss for player  $i \in N$  and value  $v_i \in \text{supp}(F)$ , i.e.,

$$\beta_i(v_i|l_i, m_i, u_i) \in \arg \inf_{b_i \in \mathbb{R}_+} \sup_{B_{-i} \in \mathcal{B}_i(l_i, m_i, u_i)} \lambda_i(b_i, B_{-i}|v_i); \quad (1)$$

2. The aggregate moment belief is consistent across the players and with  $\beta$  and  $F$ , that is,

$$\int \int \cdots \int \max_{i \in N} \beta_i(v_i|l_i, m_i, u_i) dF(v_1) dF(v_2) \cdots dF(v_n) = m_i \quad (2)$$

for all  $i \in N$ ;

3. The range beliefs are consistent, i.e., for all  $i \in N$  and  $j \neq i$

$$\begin{aligned} \inf_{v \in \text{supp}(F)} \beta_i(v|l_i, m_i, u_i) &= l_j, \text{ and} \\ \sup_{v \in \text{supp}(F)} \beta_i(v|l_i, m_i, u_i) &= u_j. \end{aligned}$$

With symmetric strategies and beliefs (dropping the indices), Equation (2) becomes

$$\int_{\underline{v}}^{\bar{v}} \beta(v|l, m, u) dF^n(v) = m,$$

that is, we can compute the expected winning minimax bid with respect to the distribution of the highest value. Moreover, the requirement of consistent range beliefs simplifies to  $\inf_{v \in \text{supp}(F)} \beta(v|l, m, u) = l$  and  $\sup_{v \in \text{supp}(F)} \beta(v|l, m, u) = u$ .

The definition requires that (i) behavior is optimal given the beliefs, and that (ii) the beliefs are consistent with behavior and the latent value distribution. The same two elements are present in a Bayes-Nash equilibrium. The difference is that in a BNE the belief is about the entire bid distribution and not just some (aggregate) statistics. Thus, one can view moment equilibrium as an analogue to Bayes-Nash equilibrium in which bidders know less than the full distribution in equilibrium.

### 3 Existence of an Aggregate Moment Equilibrium

The section contains the main theoretical result of the paper, namely that an aggregate moment equilibrium exists for all value distributions. The proof of the theorem relies on the characterization of the minimax bidding strategy, with which we begin the analysis.

### 3.1 Optimal Bidding Function

Bidder  $i$  faces the problem of not being able to know the bid that maximizes the expected utility (the ‘best response’ bid) as she does not know the competing bid distribution. Thus, any bid is likely to be suboptimal in the sense of not maximizing the expected utility. The bid can be suboptimal for two reasons. On the one hand, the bid can be too high relative to the best response or it can be too low. Both types of mistake lead to a loss in payoff relative to the highest possible.

The bidder’s solution to the problem is to choose the bid that minimizes the maximal loss in payoff due to not knowing the bid distribution. Such a bid performs uniformly relatively well as it leads to a payoff that is close to the highest possible for any feasible bid distribution.

Intuitively, loss associated with bidding too high increases in the bid: the higher the bid, the higher the loss in utility if the best response is a very low bid. If the bid is as low as the lowest possible best response, i.e.,  $b = l$ , then the bid cannot be too high. So the loss of bidding too high is 0 at  $b = l$ .

Conversely, the loss of bidding too low decreases in the bid: the lower the bid  $b$ , the higher the loss in utility if the best response is above  $b$ . If the bid is the highest undominated bid, i.e.,  $b = v$ , then the bid cannot be too low. The loss of bidding too low is 0 for  $b = v$ .

Maximal loss is the maximum of the maximal loss of bidding too high and bidding too low. The maximal loss is minimized by equalizing the two maximal losses. This explains the form of the minimax bidding function in the following proposition, which characterizes the minimax bidding function.

**Proposition 1.** *Suppose bidder  $i$  believes that the other bidders’ bids are i.i.d. draws from some distribution with the lowest bid  $l$  and the highest bid  $u$ ,  $l \leq v$ . Moreover, the distribution is such that the expected winning bid is  $m$  in auctions with  $n$  bidders. If  $v_i \leq m + (m - l) \left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}}$ , let  $b_i$  be the unique solution of*

$$\left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}} (b_i - l) = \max\left\{\left(\frac{u-m}{u-b_i}\right)^{\frac{n-1}{n}} (v_i - b_i), v_i - m\right\}. \quad (3)$$

*For higher values, let  $b_i$  be defined as the unique solution of*

$$\left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}} (b_i - l) = (v_i - b_i) \left(1 - \left(\frac{b_i - m}{b_i - l}\right)^{\frac{n-1}{n}}\right) \quad (4)$$

*The optimal bid of bidder  $i$  is  $\beta^{AGG}(v_i|l, m, u) := \min\{b_i, u\}$ . The optimal bidding function  $\beta^{AGG}$  is continuous in  $v$ ,  $l$ ,  $m$ , and  $u$ .*

The implicit equations in (3) and (4) equate the maximal loss from bidding too high and too low. The respective left-hand side of the equations give the worst-case loss of bidding too high (which is increasing in  $b_i$ ), and the right-hand side state the maximal loss when the best response is higher than  $b_i$  (which decreases in  $b_i$ ).

The minimax bidding function takes a different form below  $m$  and above  $m$ . This kink is due the constraints of the moment belief on the worst-case distributions. In particular, the implicit equations (3) and (4) show that the worst case when bidding too low depends on the bid being below or above the moment belief  $m$ .

We now explain the reason for the kink in the bidding function and the shape of worst-case loss. The proof of the proposition shows that the worst-case bid distribution has at most two elements in the support. Two elements are enough to satisfy the moment constraint and accentuate what bidding too high and bidding too low means. The moment constraint requires one mass point below  $m$  and one mass point above  $m$ . The best response to such a discrete bid distribution is to bid marginally above one of these mass points. The worst case never places all the mass above the value so that any bid is optimal.

In the case of bidding too high, one leaves most money on the table if lowering the bid does not change the probability of winning (while raising the surplus conditional on winning). This is pushed to the extreme if one mass point of the bid distribution is at the lowest possible bid, i.e.,  $x_1 = l$ . The worst case then maximizes the probability with which others bid  $l$ . The second mass point that maximizes this probability is  $x_2 = u$ , implying the probability that satisfies the moment constraint. This maximizes the loss of bidding too high.

In the case of bidding too low, the worst that can happen is if raising the bid marginally leads to a discrete increase in the winning probability. Note that this decreases the surplus conditional on winning only marginally. If one bids below  $m$ , then one mass point of the bid distribution is just marginally above  $b$ . To maximize the probability with which one could win by raising the bid marginally, the second mass point is at  $u$ . However, it can be the case for high values bidding below  $m$  that loss is higher if all the mass is placed on  $m$ . This follows from the trade-off between the probability of winning and the surplus conditional on winning. For bids above  $m$ , the mass point is again just above the bid. But now some mass below  $m$  is needed and this mass is to be minimized as it positively impacts the payoff of the bid  $b$ . The worst that can happen is a mass point at  $l$ .

Note that the equalization of the two formulas for loss may lead to a bid above  $u$ . It is never optimal to bid above  $u$  as such a bid is believed to win with certainty; a marginally lower bid would increase the payoff. Thus, the bidding function is

capped at  $u$ .

Probabilities take the form of  $q^{\frac{n-1}{n}}$  in the minimax bidding function. This form results from the inference from the aggregate moment belief. If the bid distribution puts mass on two bids  $x_1$  and  $x_2$ , then the moment constraint  $q \cdot x_1 + (1 - q) \cdot x_2 = m$  pins down  $q$ . The probability that a single bidder bids  $x_1$  is  $q^{\frac{1}{n}}$ , and the probability that the maximum of  $n - 1$  independent draws is  $x_1$  is  $q^{\frac{n-1}{n}}$ .

The following proposition discusses how the minimax bid reacts to changes in the parameters.

**Proposition 2.** *Let  $v \in [\underline{v}, \bar{v}]$  be such that  $\beta^{AGG}(v|l, m, u) \in (l, u)$ . The minimax bidding function increases in  $v$ ,  $m$ , and  $n$ . The minimax bidding function decreases in  $u$ . For  $v < m + (m - l) \left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}}$ , the minimax bid increases in  $l$ . The minimax bid reacts ambiguously to changes in  $l$  for higher types.*

Minimax loss bids increase in value. The other comparative statics are also as expected: More competing bidders and a higher expected winning bid increase the bids. An increase in  $u$  increases the probability with which other bidders bid  $l$  in the worst case when bidding too high; the optimal reaction is to lower the bid. Similarly, low types can bid higher as  $l$  increases. For types bidding above  $m$  a higher  $l$  decreases the maximal loss of bidding too high (which has a positive effect on the minimax bid) and decreases the maximal loss of bidding too low (which has a negative effect on the minimax bid). Depending on the type, both effects can dominate.

## 3.2 Equilibrium Existence

The minimax bidding function of Proposition 1 takes arbitrary, possibly misspecified, beliefs as input. We now show that there are consistent (equilibrium) beliefs.

The following theorem establishes the existence of an aggregate moment equilibrium. The beliefs, the bidding functions, and the value distribution induce a bid distribution such that the range and the expected winning bid match the players' beliefs. Moreover, it shows that there is a *separating* moment equilibrium, i.e., the beliefs  $(l, m, u)$  are such that the bidding functions are strictly increasing in  $v$ .

**Theorem 1.** *There is an aggregate moment equilibrium  $(\beta^{AGG}, l^*, m^*, u^*)$  for any value distribution  $F$  such that the bidding function  $\beta^{AGG}(v|l^*, m^*, u^*)$  is strictly increasing in  $v$ .*

We first discuss the result and then the main points of the proof. The theorem proves that a pure-strategy separating moment equilibrium exists for any value

distribution. This stands in contrast to the Bayes-Nash equilibrium, where pure equilibria tend to exist in first-price auctions only for continuous distributions. The difference is due to the nature of beliefs. In a BNE, bidders have a (correct) belief about the other bidders' strategies. If a bidder knows that the other bidders make a certain bid with positive probability, then it is optimal for some types to bid marginally higher. This leads to a mixed strategy in equilibrium. No such detailed information is available to the bidders in a moment equilibrium.

The moment equilibrium is similar to the Bayes-Nash equilibrium as the beliefs are correct in equilibrium. The difference is the level of detail. In an aggregate moment equilibrium the beliefs are statistics: the lowest and highest bids and the expected winning bid. In a BNE, the beliefs are about the entire bid distribution and not just statistics. A monotone (separating) equilibrium can exist in both types of equilibrium.

Coming to the main points of the proof, the second point of the definition of an aggregate moment equilibrium (Def. 1) implies that the moment beliefs of the players must be the same in equilibrium, i.e.,  $m_1 = m_2 = \dots = m_n$ . The third point of the definition implies that the bidders must have identical range beliefs.

The first step in the proof shows that  $l^* = \underline{v}$  is the unique consistent lower bound belief. This can be inferred from the minimax bidding function. As the minimax bidding function increases in  $v$ , the minimum is attained at  $\underline{v}$ . Suppose  $l < \underline{v}$  and  $\beta^{\text{AGG}}(\underline{v}|l, m, u) = l$ . The equality

$$\left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}}(b_i-l) = \max\left\{\left(\frac{u-m}{u-b_i}\right)^{\frac{n-1}{n}}(v_i-b_i), v_i-m\right\}$$

does not hold for  $b_i = l$  as the left-hand side is 0 and the right-hand side is strictly positive.

Requiring separation pins down a unique consistent upper bound belief  $u$  for each moment belief  $m$ . The requirement is that the highest type  $\bar{v}$  is the unique type that bids  $u$ . This is the case if  $u$  solves

$$\left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}}(u-l) = (\bar{v}-u)\left(1 - \left(\frac{u-m}{u-l}\right)^{\frac{n-1}{n}}\right).$$

There is a unique  $u$  that solves the equation for each  $m$ . Let  $u(m)$  denote the solution as a function of  $m$ . This function increases in  $m$  (as shown in the appendix).

The final step of the proof studies the function

$$\varphi^{\text{AGG}}(m) = \int_{\underline{v}}^{\bar{v}} \beta^{\text{AGG}}(v|l = \underline{v}, m, u = u(m)) dF^n(v)$$

for  $m \in [\underline{v}, \bar{v}]$ . The function  $\varphi^{\text{AGG}}$  returns for each aggregate moment belief  $m$  the corresponding expected winning bid. A fixed point of this function is thus an aggregate moment equilibrium belief. The domain of the function is the set of feasible moment beliefs. This set is  $[\underline{v}, \bar{v}]$  as all types bid above  $l$ , which is equal to  $\underline{v}$ , and no type bids above value, so all bidders bid below  $\bar{v}$ . The same arguments establish that the interval  $[\underline{v}, \bar{v}]$  is also the codomain. The existence of a fixed point is shown by arguing that  $\varphi^{\text{AGG}}(m) > m$  for  $m$  in a neighborhood of  $\underline{v}$  and  $\varphi^{\text{AGG}}(m) < m$  for  $m$  in a neighborhood of  $\bar{v}$ . For  $m$  close to  $\underline{v}$ , the vast majority of types bid above  $m$ , so the expected winning bid is higher than the moment belief. Similarly, for very high moment beliefs the vast majority bid below  $m$  so that the expected winning bid is below  $m$ . As the function is continuous, it must intersect the 45-degree line at least once. The aggregate moment belief at the intersection is consistent by construction and it induces separating range beliefs.

Note that if there is only one possible value ( $\underline{v} = \bar{v}$ ), then a separating equilibrium exists trivially ( $l^* = m^* = u^* = \underline{v}$ ). If there are at least two types, then the previous limit arguments apply.

The following example illustrates the equilibrium construction and behavior.

**Example 1.** *We consider two bidders,  $n = 2$ , and the uniform value distribution, i.e.,  $F(v) = v$  for  $v \in [0, 1]$ . Figure 1a shows the function  $\varphi^{\text{AGG}}(m)$  that returns for the aggregate moment belief  $m$  the corresponding expected winning bid if the range beliefs induce separating bidding functions. The aggregate moment equilibrium is at the intersection of  $\varphi^{\text{AGG}}$  and the identity function  $\text{id}(m) = m$ . The equilibrium beliefs are  $l^* = 0$ ,  $m^* = 0.37$ , and  $u^* = 0.5$ . The average bid is 0.28, which is higher than the corresponding average bid of 0.25 in the Bayes-Nash equilibrium. Figure 1b displays the moment equilibrium bidding function. The kink at  $m^*$  is clearly visible. The figure also shows the BNE bidding function for comparison. BNE bids are lower than moment equilibrium bids.*

Figure 1a reveals that there is also a pooling equilibrium at  $l = m = u = \underline{v}$ . This pooling equilibrium exists for all value distributions. It exists for technical reasons. Suppose a bidder thinks that all other bidders bid  $l$  with certainty. There is no uncertainty and the (supremum) best response is  $l$ , confirming the belief that everyone bids  $l$ . The equilibrium is unstable in the sense that marginally different



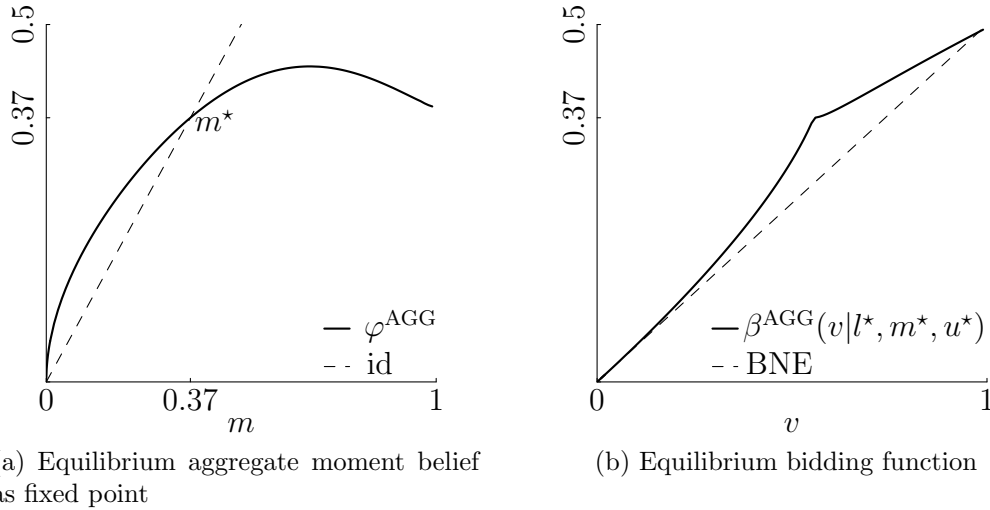


Figure 1: Aggregate moment equilibrium with uniform values and two bidders

*Notes: The equilibrium aggregate moment belief  $m^*$  is at the intersection of  $\varphi(m)$ , which gives the expected winning bid when players have the belief  $m$ , and the identity function  $id(m)$ .*

beliefs ( $l < m$ ) lead to bids above  $l$ . We do not consider the pooling equilibrium to be of economic interest as, typically, there is variation in the bids.

## 4 Individual Moment Beliefs

The model of the previous sections was slightly unconventional from a modeling perspective as players formed beliefs about an aggregate statistic: the expected winning bid. This section presents an alternative model that follows the game-theoretic convention of players forming beliefs about the other players' individual behavior. The beliefs take the form of statistics again. Note that both types of belief may come from what is observable, or, as mentioned in the introduction, what can be learned from a small sample of past (winning) bids.<sup>11</sup>

### 4.1 Individual Moment Beliefs and Equilibrium

Bidder  $i$  now forms beliefs directly about the behavior of bidder  $j$ ; there is no detour via the distribution of the winning bid. Bidder  $i$ 's belief about bidder  $j$ 's behavior is that the bid distribution of bidder  $j$  belongs to the set  $\mathcal{B}_{ij}$ , where  $\mathcal{B}_{ij} \subseteq \mathcal{P}$ . The set  $\mathcal{B}_i = \times_{j \neq i} \mathcal{B}_{ij}$  is the set of possible competing bid distributions or the belief set.

<sup>11</sup>For simplicity, we confine ourselves to bidders learning either the average winning bid or the average opponent bid. In a sample of losing and winning bids, one can estimate both statistics. We leave this case and the one in which bidders also learn higher moments for future research.

Bidder  $i$  believes that bidder  $j$ 's bid is  $\mu_i$  on average. We refer to  $\mu_i$  as an (individual) moment belief. As in the previous section, bidder  $i$  believes that bidder  $j$  does not bid below  $l_i$  and does not bid above  $u_i$ . Thus, bidder  $i$  believes that the bid of bidder  $j$  is drawn from a distribution in the set  $\mathcal{B}_{ij}(l_i, \mu_i, u_i)$ , where this set is

$$\mathcal{B}_{ij}(l_i, \mu_i, u_i) = \{B \in \mathcal{P} : \int_{l_i}^{u_i} dB = 1 \text{ and } \int x dB(x) = \mu_i\}.$$

Note that the expected bids of the other bidders are identical. Their bids do not necessarily come from the same distribution.

The new model requires adapting the definition of a moment equilibrium.

**Definition 2.** The tuple  $(\beta, l, \mu, u)$  is a *moment equilibrium* if

1.  $\beta_i$  minimizes the maximal loss for player  $i \in N$  and value  $v_i \in \text{supp}(F)$ , i.e.,

$$\beta_i(v_i | l_i, \mu_i, u_i) \in \arg \inf_{b_i \in \mathbb{R}_+} \sup_{B_{-i} \in \mathcal{B}_i(l_i, \mu_i, u_i)} \lambda_i(b_i, B_{-i} | v_i); \quad (5)$$

2. The moment beliefs are consistent with  $\beta$  and  $F$ , that is,

$$\int \beta_i(v_i | l_i, \mu_i, u_i) dF(v_i) = \mu_j \quad (6)$$

for all  $i \in N$  and  $j \neq i$ ;

3. The range beliefs are consistent, i.e., for all  $i \in N$  and  $j \neq i$

$$\begin{aligned} \inf_{v \in \text{supp}(F)} \beta_i(v | l_i, \mu_i, u_i) &= l_j, \text{ and} \\ \sup_{v \in \text{supp}(F)} \beta_i(v | l_i, \mu_i, u_i) &= u_j. \end{aligned}$$

In an (individual) moment equilibrium, the same kind of consistency requirements as in an aggregate moment equilibrium are met. The range beliefs are consistent with optimal behavior and the value distribution. Moreover, the expected optimal bid of a player equals the other bidders' beliefs about the expected bid.

Bayesian priors can be thought of as an infinite sequence of moment beliefs that pin down a unique distribution. A Bayes-Nash equilibrium then requires an infinite number of moment beliefs to be consistent. In a moment equilibrium we only require a finite number of moment beliefs to be consistent.

## 4.2 Optimal Bidding Function

We again start the analysis by characterizing the minimax bidding function. The bidders face the same type of decision problem as above. They engage in worst-case reasoning to find the bid that performs uniformly well in the sense of worst-case loss for all possible bid distributions. Loss again comes from bidding too high or low relative to the best response bid. The bid that minimizes maximal loss equates the maximal loss when bidding too high with the maximal loss when bidding too low.

The difference between individual and aggregate moment beliefs is that bidders now do not need to infer something from the distribution of the highest bid of  $n$  bidders about the highest bid of  $n - 1$  bidders. Bidders now form beliefs directly about the distribution of each other bidder's bid distribution and can use these beliefs to form a belief about the distribution of the highest bid of  $n - 1$  opponent bidders.

It is possible to state the minimax bidding function in closed form when  $n = 2$ . We define the following two cutoff values

$$\hat{v}_1 = \frac{u(\mu - l) + \mu(u - \mu)}{u - l} \quad \text{and} \quad \hat{v}_2 = \frac{u(u - l) - l(u - \mu)}{\mu - l}$$

for the two-bidder case.

**Proposition 3.** *Suppose bidder  $i$  believes that the other bidders' bids are independent draws from bid distributions with lowest bid  $l$  and highest bid  $u$ ,  $l \leq \underline{v}$ . Moreover, each bidder  $j$ 's bid distribution is such that the expected bid is  $\mu$  in auctions with  $n$  bidders.*

*Let  $n = 2$ . The minimax bidding function is*

$$\beta^{IND}(v_i | l, \mu, u) := \begin{cases} u - \sqrt{(u - l)(u - v_i)} & \text{if } v_i < \hat{v}_1 \\ l + \frac{-(\mu - l)(u - l) + \sqrt{(\mu - l)(u - l)((4v - 3l)(u - \mu) + u(\mu - l) - l(u - l))}}{2(u - \mu)} & \text{if } \hat{v}_1 \leq v_i < \hat{v}_2 \\ u & \text{if } v_i \geq \hat{v}_2. \end{cases} \quad (7)$$

*Let  $n \geq 3$ . Let  $\tilde{x}_1(b) = \max\{l, \min\{\mu, \frac{(n-1)b-u}{n-2}\}\}$ . If  $v_i \leq \mu + (\mu - \tilde{x}_1(\mu)) \left(\frac{u - \mu}{u - \tilde{x}_1(\mu)}\right)^{n-1}$ , let  $b_i$  be the unique solution of*

$$\left(\frac{u - \mu}{u - \tilde{x}_1(b_i)}\right)^{n-1} (b_i - \tilde{x}_1(b_i)) = \left(\frac{u - \mu}{u - b_i}\right)^{n-1} (v_i - b_i). \quad (8)$$

For higher values, let  $b_i$  be defined as the unique solution of

$$\left(\frac{u - \mu}{u - \tilde{x}_1(b_i)}\right)^{n-1} (b_i - \tilde{x}_1(b_i)) = (v_i - b_i) \left(1 - \left(\frac{b_i - \mu}{b_i - l}\right)^{n-1}\right) \quad (9)$$

The optimal bid of bidder  $i$  is  $\beta^{IND}(v_i|l, \mu, u) := \min\{b_i, u\}$ . The optimal bidding function  $\beta^{IND}$  is continuous in  $v$ ,  $l$ ,  $\mu$ , and  $u$ .

The first step of the proof shows that the worst-case bid distributions are identical across the competing bidders. This was imposed in the previous section to infer something from the expected winning bid about the individual bid distributions. Now the bid distributions could be different (up to the range and the first moment), but they are identical in the worst case. Loss is maximized by certain extreme cases, making possible asymmetries irrelevant.

The second step shows that the worst-case bid distributions have at most two elements in the support. The result holds more generally. A moment belief is the expectation of a measurable function. In the current model, the measurable function is  $g(b_j) = b_j$  and the moment belief is  $\mu_i$ . Let  $M_{ij}$  be the number of moment beliefs that bidder  $i$  has about bidder  $j$ 's bid distribution. The worst-case bid distribution has at most  $M_{ij} + 1$  elements in the support (Winkler, 1988).

Let  $x_1$  and  $x_2$  denote the two bids in the support, where  $l \leq x_1 \leq \mu \leq x_2 \leq u$ . The moment constraint then implies that  $\pi \cdot x_1 + (1 - \pi)x_2 = \mu$ , so the moment constraint pins down the mass  $\pi$  on  $x_1$ . Bidder  $i$ 's payoff depends on the distribution of the highest bid of  $n - 1$  independent bids. The probability that the highest bid of the competing bidders is  $x_1$  is then  $\pi^{n-1}$ . Worst-case loss is now convex in the probability, whereas it was concave in the previous section.

The minimax bid again equalizes the maximal loss from bidding too high with the maximal loss from bidding too low. The implicit equations (8) and (9) state the maximal loss of bidding too high on the left-hand side and the maximal loss of bidding too low on the right-hand side.

The maximal loss of bidding too high is more elaborate than in the previous section. Depending on  $b$ , it can be maximized by any  $x_1 \in [l, \mu]$ . Why is the worst case of bidding too high more involved with individual moment beliefs than with aggregate moment beliefs? In both cases, maximal loss takes the form  $\lambda^H = \pi^\alpha(b - x_1)$ , where  $\pi = (x_2 - m)/(x_2 - x_1)$  or  $\pi = (x_2 - \mu)/(x_2 - x_1)$  and  $\alpha \in \{\frac{n-1}{n}, n-1\}$ . Loss trades off the probability with which  $b$  becomes winning and the ex-post loss  $b - x_1$  conditional on winning. A higher  $x_1$  raises the probability while it decreases the difference. If  $\alpha = (n-1)/n$ , then loss is concave in the probability. This means

that raising  $\pi$  to the power of  $\alpha$  “boosts” the probability as  $\pi^\alpha > \pi$  for  $\alpha < 1$ . This is not the case when  $\alpha = n - 1 \geq 1$ . Then raising  $\pi$  to the power of  $\alpha$  depresses the probability, which can then be increased by raising  $x_1$ . In the latter case, there is a non-trivial interaction between  $\pi$  and  $b - x_1$ . In the former case, loss is always maximized by putting as much mass as possible on  $l$ , as it anyway leads to a high probability.

The maximal loss of bidding too low is simpler as the worst case always puts a mass point slightly above  $b$ . In contrast to the case of aggregate moment beliefs, the worst case when bidding the minimax bid never puts all the mass on  $\mu$ . The proof shows that this can be the worst case for bids that are not the minimax bids.

**Proposition 4.** *Let  $v \in [\underline{v}, \bar{v}]$  such that the minimax bid  $\beta^{IND}(v|l, \mu, u) \in (l, u)$ . The minimax bidding function increases in  $v$ ,  $\mu$ , and  $n$ .*

The comparative statics are again as one would expect. The minimax bid is strictly increasing in  $v$  and  $n$ .

### 4.3 Equilibrium Existence

A separating moment equilibrium exists also in the case of individual moment beliefs for any value distribution.

**Theorem 2.** *There is a moment equilibrium  $(\beta^{IND}, l^*, \mu^*, u^*)$  for any value distribution  $F$  such that the bidding function  $\beta^{IND}(v|l^*, \mu^*, u^*)$  is strictly increasing in  $v$ .*

The unique consistent lower bound belief is again  $l^* = \underline{v}$ . In terms of the upper bound belief, separation and consistency requires  $\bar{v} = \hat{v}_2$  if  $n = 2$  and

$$\bar{v} = u + \frac{(u - \mu)(u - l)^{n-1}}{(u - l)^{n-1} - (u - \mu)^{n-1}}$$

if  $n \geq 3$ . In either case, there is a unique  $u^*(\mu)$  for each  $l$  and  $\mu$  that solves the respective equation. Uniqueness follows from the right-hand side being strictly increasing in  $u$ .

The function  $\psi^{IND}$  returns for each moment belief  $\mu$  the expected bid if the range beliefs are consistent and separating, i.e.,

$$\psi^{IND}(\mu) = \int_{\underline{v}}^{\bar{v}} \beta^{IND}(v|l^*, \mu, u^*(\mu)) dF(v). \quad (10)$$

The consistent moment belief  $\mu^*$  is a fixed point of  $\psi^{IND}$ .

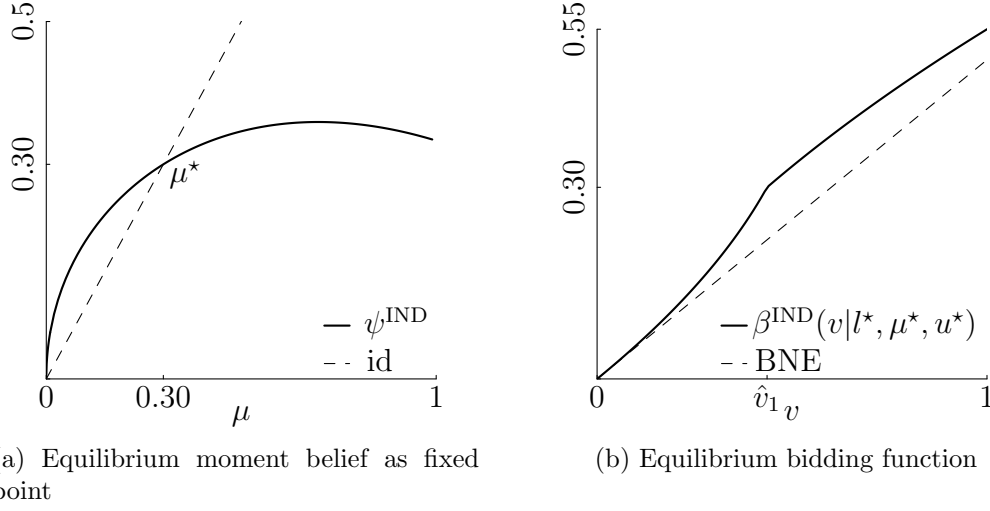


Figure 2: Moment equilibrium with uniform values and two bidders  
Notes: The equilibrium moment belief  $\mu^*$  is at the intersection of  $\varphi(\mu)$ , which gives the expected bid when players have the belief  $\mu$ , and the identity function  $id(\mu)$ .

The domain of  $\psi^{\text{IND}}$  is  $[\underline{v}, \bar{v}]$ . If  $\mu = \underline{v}$ , then all the mass is on  $l^*$ . Likewise, if  $\mu = \bar{v}$ , then  $u^* = \mu = \bar{v}$  and all the mass is on  $u^*$ . The codomain of  $\psi^{\text{IND}}$  is also  $[\underline{v}, \bar{v}]$  as no type bids below  $l^*$  and no type bids above value. The function  $\psi^{\text{IND}}$  is continuous by Lebesgue's dominated convergence theorem (Elstrodt, 2018). A fixed point exists according to Brouwer's fixed point theorem. We argue that the fixed point is in the interior so that the moment equilibrium is separating. If  $\mu = \underline{v}$ , then the minimax bid of all types is  $\underline{v}$ . If  $\mu$  is marginally higher, then almost all values bid strictly higher than  $\underline{v}$ . Similarly, if  $\mu = \bar{v}$ , then almost all values  $v$  bid less than  $\mu$ . The function  $\psi^{\text{IND}}$  must intersect the identity function at least once in the interior of  $[\underline{v}, \bar{v}]$ .

**Example 2.** The value distribution is uniform with support  $[0, 1]$  and there are two players, i.e.,  $n = 2$ . The upper belief is  $u(\mu) = \sqrt{\mu}$  in a separating equilibrium. Figure 2a shows the function  $\psi(\mu) = \int_0^1 \beta^{\text{IND}}(v|l = 0, \mu, u(\mu)) dv$ . The concave shape of  $\psi$  implies that there is a unique separating moment equilibrium with consistent range beliefs. The equilibrium beliefs are  $l^* = 0$ ,  $m^* = 0.3$ , and  $u^* = 0.55$ . For comparison, the expected bid in the unique BNE is 0.25 and the highest bid is 0.5. Figure 2 illustrates the respective equilibrium bidding functions.

A final note is that there is a multiplicity of equilibria with inconsistent range beliefs. For any  $l \in [0, \underline{v}]$  there is a consistent moment belief  $\mu^*(l)$  and an upper bound belief such that the bidding function is strictly increasing.

## 5 Identification and Estimation

This second part of the paper is concerned with the estimation of the latent value distribution  $F$  from observed bids. We first provide non-parametric identification results and then non-parametric estimators of the value distribution. The next section applies these to data from highway procurement auctions and assesses the fit in out-of-sample predictions.

### 5.1 Non-parametric Identification

Suppose we observe a distribution  $G$  of bids and know that these bids come from auctions with  $n$  bidders. Our objective is to infer the value distribution  $F$  from the bid distribution  $G$ . To do so, we need a (game-theoretic) model that informs us about the relation between  $F$  and  $G$ . Guerre et al. (2000) [GPV], the landmark paper on the structural analysis of first-price auctions, assumes that the data is generated by bidders playing a Bayes-Nash equilibrium. We assume in contrast that the bidders play an aggregate moment equilibrium or a moment equilibrium. In any case, the value distribution  $F$  is identified if it is uniquely determined by the bid distribution and the assumption that the data is generated by a certain model.

Theorems 1 and 2 show that there are equilibrium beliefs that induce strictly increasing bidding functions. The bidding functions depend on the bidder's value and on statistics that can be inferred from the bid distribution. The inverse minimax bidding functions then translate the bid distribution into a unique value distribution. Let  $v^{\text{AGG}}$  denote the inverse minimax bidding function when bidders have aggregate moment beliefs, i.e.,  $v^{\text{AGG}}(\beta^{\text{AGG}}(v|l^*, m^*, u^*)) = v$ . Let  $v^{\text{IND}}$  denote the inverse minimax bidding function when bidders have individual moment beliefs.

The next theorem shows that the value distribution  $F$  is non-parametrically identified by the bid distribution if the bid distribution is generated by the bidders playing an aggregate moment equilibrium. Guerre et al. (2000) prove an analogous theorem when BNE is the underlying behavioral model.

**Theorem 3.** *Let  $G$  be a distribution with support  $[l, u]$  and expected winning bid  $m$ , i.e.,  $m = \int b dG^n(b)$ . There exists a distribution of bidders' private values  $F$  such that  $G$  is the bid distribution of a separating aggregate moment equilibrium if and only if  $l < m < u$ . Moreover, when  $F$  exists, it is unique with support  $[\underline{v}, \bar{v}]$  and satisfies  $F(v) = G(\beta^{\text{AGG}}(v|l, m, u))$  for all  $v \in [\underline{v}, \bar{v}]$ .*

The theorem relies on the minimax bidding function  $\beta^{\text{AGG}}$  being strictly increasing in  $v$  in a separating moment equilibrium (Proposition 2) because this allows

transforming the value distribution to a bid distribution and vice versa. In a separating equilibrium, the unique type bidding  $u$  is  $\bar{v}$ . The necessary condition for the bidding function to be strictly increasing is then  $l < m < u$ .

The value distribution is non-parametrically identified from the observed bid distribution whenever  $l < m < u$ . This is a minimal prerequisite as it only requires that there is some dispersion in the bids. If all the mass were on one point, then one could not tell whether the data came from a point mass value distribution or a pooling equilibrium.

We highlight a difference to the Bayes-Nash equilibrium-based approach by GPV. In contrast to GPV, we do not require  $G$  to be absolutely continuous. This is because a pure moment equilibrium exists for all value distributions whereas there is no guarantee that a pure Bayes-Nash equilibrium exists if the value distribution has mass points. We also do not need  $G$  to admit a density as we do not consider first-order conditions.

The identification is non-parametric, which implies that the estimates are robust to possible misspecification of the underlying parametric class of distributions. Clearly, the identification also holds if one makes parametric assumptions.

The following theorem shows that the value distribution  $F$  is non-parametrically identified by the bid distribution if one assumes that the bidders play a separating moment equilibrium as in Section 4.

**Theorem 4.** *Let  $G$  be a distribution with support  $[l, u]$  and expected value  $\mu$ , i.e.,  $\mu = \int bdG(b)$ . There exists a distribution of bidders' private values  $F$  such that  $G$  is the bid distribution of a separating moment equilibrium if and only if  $l < \mu < u$ . Moreover, when  $F$  exists, it is unique with support  $[\underline{v}, \bar{v}]$  and satisfies  $F(v) = G(\beta^{IND}(v|l, m, u))$  for all  $v \in [\underline{v}, \bar{v}]$ .*

We omit the statement of the proof as it is analogous to the proof of Theorem 3.

## 5.2 Non-parametric Estimation

In practice, we observe a sample of the underlying bid distribution  $G$ . The sample can be structured in the sense of identifying the bids that come from the same auction. A structured sample  $H$  takes the form of a  $T \times n$  matrix, where  $T$  is the number of observed auctions and  $n$  is the number of bidders. A typical element (bid) of  $H$  is denoted by  $h_{tj}$ ,  $1 \leq t \leq T$  and  $1 \leq j \leq n$ . An unstructured sample  $h$  takes the form of a vector of bids. Abusing notation, the sample has length  $T$  and we use  $h_t$  to denote a typical element,  $1 \leq t \leq T$ . We do not know which bids were winning or losing, we only assume that the bids in  $h$  are i.i.d. draws from  $G$ .



We assume that we see the number of bidders  $n$  that generated sample  $h$ . One can immediately go from a structured sample  $H$  to an unstructured sample  $h$  by listing all rows of  $H$  in one vector. One cannot translate an unstructured sample into a structured sample.

We use the inverse minimax bidding function to estimate  $F$  from the sample. All we need to do is estimate the players' beliefs. An immediate estimator for the lower bound belief  $l$  is the minimum bid in the sample, i.e.,  $\hat{l} = \min_t h_t$ . Similarly, an estimator for  $u$  is the highest bid in the sample, that is,  $\hat{u} = \max_t h_t$ . Donald and Paarsch (2002) show that the sample extreme values are super-consistent estimators. In a structured sample, the estimator for the expected winning bid is the average winning bid, i.e.,  $\hat{m} = \frac{1}{T} \sum_{t=1}^T \max_j h_{tj}$ . If the data does not identify the winning bids, then one can estimate  $m$  by computing the expected value of the highest of  $n$  independent draws from  $h$ . To do so, we compute the probability mass function implied by the sample  $h$  and then compute the expectation of the distribution of the highest bid that only takes values in  $h$ . In the moment equilibrium with individual moment beliefs, one estimates  $\mu$  by taking the sample mean, i.e.,  $\hat{\mu} = \frac{1}{T} \sum_t h_t$ . For the bidders' belief about the number of player  $n$ , we use the observed number of bidders.

The two estimators are consistent given the data is generated by the respective equilibrium notion as the beliefs  $(l, m, u)$  and  $(l, \mu, u)$  are estimated consistently.

The estimation is fully non-parametric. At a high level, there are two differences to BNE-based approaches. The first difference is that the beliefs take the form of statistics in the moment equilibrium approach, whereas they are about entire continuous distributions in the BNE approach. No kernel estimation is necessary in contrast to the BNE-based approach of Guerre et al. (2000). A second difference is that the beliefs are about the bid distribution in the moment equilibrium approach, so about something that is observed. BNE bases primitive beliefs on the unobserved value distribution.

In the empirical application, we use a total of four estimators (in addition to GPV). The first two estimators are those already discussed. They use the sample minimum and maximum as estimates for  $l$  and  $u$ , respectively, but differ in the underlying game-theoretic model. While the first set of estimators is super-consistent, it may be sensitive to outliers.

The two other estimators follow a different approach to estimating  $l$ .<sup>12</sup> In a parametric setting, Donald and Paarsch (2002) use an estimator for the lower bound

---

<sup>12</sup>As the application considers procurement auctions, it is actually the highest bid  $u$  that receives different treatment. Note that in procurement auctions the highest bid  $u$  takes the role of the lowest bid  $l$  in standard bidders-as-buyers auctions.

of the support that is higher than the sample minimum. As we estimate the value distribution non-parametrically, we cannot use their estimator. Instead, we estimate the minimum bid after removing outliers. Specifically, we use the interquartile range so that outliers are bids below  $Q_1 - k(Q_3 - Q_1)$ , where  $Q_1$  and  $Q_3$  are the lower and upper quartiles, respectively, and  $k = 1.5$  (Tukey, 1977). The estimator for  $l$  is then the minimum sample bid that is higher than  $Q_1 - k(Q_3 - Q_1)$ . To minimize the effect of removing outliers, we estimate the belief  $\mu$  before removing the outliers. Note that we only change the estimator for  $l$ . Our rationale is that (potentially) winning bids are probably more meaningful than losing bids. Sample extrema are not sensitive to outliers, which may be obtained by mistakes in the data collection, the treatment of covariates, or just noisy bidding. By focusing on low bids we implicitly assume that these issues matter more at the bottom of the distribution. In particular, noisy bidding arguably matters more for losing than for winning bids.

## 6 Structural Analysis of Highway Procurement Auctions

We now use the moment equilibrium approach to estimate the distribution of private cost in highway procurement auctions. The section first introduces the data and then discusses how we deal with the observed and unobserved auction heterogeneity. We then estimate the cost distribution and assess the estimates in out-of-sample predictions. Our benchmark is the estimator based on the BNE-based approach of Guerre et al. (2000). Appendix B shows how the theory has to be adapted for procurement auctions.

### 6.1 Data

We analyze procurement auctions of California’s Department of Transportation for highway paving contracts from 1999 to 2005. The data set was originally prepared and analyzed by Bajari, Houghton, and Tadelis (2014) (BHT).<sup>13,14</sup>

The timing of these auctions is that government engineers first prepare information on the project and a cost estimate. The projects typically involve different work and cost items. The government engineers specify the expected quantities for each item, e.g., the tons of asphalt needed. Interested contractors can submit bids

---

<sup>13</sup>I thank Bajari, Houghton, and Tadelis for making their data publicly available.

<sup>14</sup>The only minor difference is that we correct the number of bidders when the variable giving the number of bidders (NBIDDERS) does not match the number of recorded bids. We adjust the number of bidders to match the number of recorded bids. There are two such auctions.

for these individual items until a certain set date. The bidder with the lowest total bid wins the contract. We assume that bidders directly submit the total bid. The total cost of undertaking the project is assumed to be privately known to the bidder at the time of bidding.

In the data, 347 firms submitted a total of 3,661 bids in 819 auctions. There is substantial heterogeneity in these firms. BHT report that the 20 largest firms won 73.4% of the total contract dollars awarded. There is a sizable difference even within the top 20 contractors, as only two companies participated in more than 10% of the auctions. More than half of the firms never won a paving contract in the data.

BHT deal with the bidder heterogeneity by classifying bidders into two categories: fringe and non-fringe bidders. A fringe firm is a contractor that won less than 1% of the value of contracts awarded. Using this definition, 20 firms are non-fringe, while 327 firms are classified as fringe. Similar to BHT, we assume that fringe and non-fringe firms have different cost distributions.

The data contains the bids and many covariates, of which we use the following. The engineer’s estimate (ENG) is known to the bidders before bidding and reflects an estimate of a “fair and reasonable price.” The variable DIST is the firm’s distance to the project’s location. Distances are measured in 100 miles. As firms probably have capacity constraints, BHT provide with UTIL a measure of the estimated capacity utilization rate. The capacity utilization rate is the ratio of backlog and capacity, where capacity is the maximum backlog in the sample. Assuming a linear depreciation, backlog is the remaining dollar value of previously won projects in the data. BHT also use RDIST and RUTIL to provide the minimum of the rivals’ distance and capacity utilization rate, respectively.

The level of the bids is affected by factors that are unique to the auction, such as the engineer’s estimate. We control for the observed and unobserved auction heterogeneity by homogenizing the bids with a random effects regression. A similar homogenization procedure was used, for example, by Hong, Haile, and Shum (2003), Krasnokutskaya (2011), and Bajari et al. (2014).

The theoretical underpinning of the homogenization regression is the following proposition that shows the moment equilibrium scales analogously with affine transformations of the values. Athey and Haile (2007) prove an analogous result for BNE. More formally, and in the framing of bidders-as-buyers auctions, we assume that the bidders’ values are an affine function of their private value (signal), i.e., their willingness-to-pay is  $\alpha(v; \mathbf{y}) = p(\mathbf{y})v + q(\mathbf{y})$ , where  $\mathbf{y}$  is the vector of observables and  $p$  and  $q$  are functions.

**Proposition 5.** *If  $\beta^{IND}$  is the minimax bidding function for  $l_0, \mu_0, u_0$ , and values  $v$ , then  $\beta_\alpha^{IND}$ ,*

$$\beta_\alpha^{IND}(\alpha(v; \mathbf{y}) | \alpha(l_0; \mathbf{y}), \alpha(\mu_0; \mathbf{y}), \alpha(u_0; \mathbf{y})) = p(\mathbf{y})\beta^{IND}(v | l_0, \mu_0, u_0) + q(\mathbf{y}),$$

*is the minimax bidding function for  $\alpha(l_0; \mathbf{y}), \alpha(\mu_0; \mathbf{y}), \alpha(u_0; \mathbf{y})$ , and values  $\alpha(v; \mathbf{y})$ .*

*Let values and covariates  $\mathbf{y}$  be independent. If  $(\beta^{IND}, l_0, \mu_0, u_0)$  is a moment equilibrium for valuations  $v$  distributed according to distribution  $F$ , then  $(\beta_\alpha^{IND}, \alpha(l_0; \mathbf{y}), \alpha(\mu_0; \mathbf{y}), \alpha(u_0; \mathbf{y}))$  is a moment equilibrium for valuations  $\alpha(v; \mathbf{y})$  distributed according to distribution  $F_\alpha$ ,  $F_\alpha(\alpha(v; \mathbf{y})) = F(v)$  for all  $v \in [\underline{v}, \bar{v}]$ .*

*Analogous statements hold for minimax bidding functions with aggregate moment beliefs and aggregate moment equilibria.*

Relying on Proposition 5, we homogenize the data by first running the random effects regression

$$\begin{aligned} \frac{b_{ij}}{\text{ENG}_{ij}} = & \alpha_1 + \beta_i + \sum_{k=2}^9 \alpha_k \chi_{\{n_i=k\}} + \alpha_{10} \text{FRINGE}_{ij} + \alpha_{11} \text{DIST}_{ij} + \alpha_{12} \text{UTIL}_{ij} \\ & + \alpha_{13} \text{RUTIL}_{ij} + \alpha_{14} \text{RDIST}_{ij} + \varepsilon_{ij}, \end{aligned} \quad (11)$$

where  $i$  identifies the auction and  $j = 1, 2, \dots, n_i$  the identity of the bidder in auction  $i$ . The variable  $\chi_{\{n_i=k\}}$  is a dummy variable for the number of bidders. The variable is 0 if  $n_i \geq 10$ . Note that there are no auctions with one bidder. Intuitively, the coefficient estimates should be positive and decreasing as one would expect that more competition in a procurement auction in the form of more participating firms brings bids down. The variable FRINGE is a dummy variable that is 1 if firm  $j$  is classified as a fringe firm and 0 otherwise. Following Athey and Haile (2007), the homogenized bid  $\hat{b}_{ij}$  is then  $\hat{\varepsilon}_{ij} + \hat{\alpha}_1 + \hat{\alpha}_{n_i}$ , where  $\hat{\alpha}_{n_i} = 0$  for  $n_i \geq 10$ .

Table 1 presents the estimated coefficients. One can see that the effect of the number of bidders on the mean bid is not always as theory would predict. Bids increase when moving from four to five bidders, which is counterintuitive. The structural estimation only uses data with at most seven bidders per auction.

We assume that the costs of fringe and non-fringe bidders are drawn from different distributions. We focus on symmetric auctions for simplicity, and thus divide the data into two data sets. In one data set all bidders in all auctions are fringe firms, while in the other set in all auctions all bidders are non-fringe firms. Table 2 summarizes the homogenized data. The number of observed bids lies between 10 and 108 for the various categories. For the non-fringe data, there are auctions with

Table 1: Homogenization random effects regression (11): coefficient estimates

| Intercept | FRINGE  | DIST    | UTIL    | RDIST   | RUTIL   |         |         |
|-----------|---------|---------|---------|---------|---------|---------|---------|
| 0.967     | 0.040   | 0.000   | 0.007   | 0.000   | -0.042  |         |         |
| $n = 2$   | $n = 3$ | $n = 4$ | $n = 5$ | $n = 6$ | $n = 7$ | $n = 8$ | $n = 9$ |
| 0.117     | 0.071   | 0.058   | 0.081   | 0.009   | 0.012   | -0.066  | 0.019   |
| # of bids | 3661    |         |         |         |         |         |         |

two, three, four, and five bidders. For the fringe data, we only use the auctions with at most seven bidders. With the exception of five bidders, the mean homogenized bid decreases in the number of bidders. Next, consider the lowest bids  $\hat{l}$  in the sample. Theory predicts that the lowest bid  $l$  decreases as the auction becomes more competitive (higher  $n$ ). The sample does not reflect this property; we take this into account in the estimation. Theory also predicts that the highest bid  $u$  equals the highest cost  $\bar{c}$  for any number of bidders ( $n \geq 2$ ). The sample also exhibits variation in the highest observed bids. Moreover, in auctions with fringe bidders there are, for  $n = 3$  and  $n = 5$ , bids that are significantly higher than the other maximal bids. As discussed in Section 5.2, we take these outliers into account in the estimation.

Table 2: Summary of the homogenized bids

|   | $\hat{l}$ | $\hat{m}$ | $\hat{\mu}$ | $\hat{u}$ | # of bids |
|---|-----------|-----------|-------------|-----------|-----------|
| <b>Auctions with exclusively non-fringe bidders</b> |           |           |             |           |           |
| $n = 2$   | 0.93      | 1.02      | 1.07        | 1.29      | 108       |
| $n = 3$   | 0.85      | 0.96      | 1.05        | 1.38      | 90        |
| $n = 4$   | 0.90      | 0.93      | 1.02        | 1.26      | 20        |
| $n = 5$   | 0.94      | 0.94      | 1.04        | 1.19      | 10        |
| <b>Auctions with exclusively fringe bidders</b>     |           |           |             |           |           |
| $n = 2$   | 0.91      | 1.03      | 1.10        | 1.36      | 48        |
| $n = 3$   | 0.64      | 0.91      | 1.05        | 2.01      | 54        |
| $n = 4$   | 0.77      | 0.90      | 1.03        | 1.53      | 92        |
| $n = 5$   | 0.80      | 0.90      | 1.05        | 1.97      | 70        |
| $n = 6$   | 0.62      | 0.81      | 0.97        | 1.28      | 48        |
| $n = 7$   | 0.75      | 0.82      | 0.97        | 1.46      | 77        |

Notes: The minimum (maximum) bid is denoted by  $\hat{l}$  ( $\hat{u}$ ). The average winning bid is  $\hat{m}$ . The average bid is  $\hat{\mu}$ .

## 6.2 Estimation and Out-of-sample Predictions

We now come to the estimation of the cost distribution and the out-of-sample assessment of the estimates. We take the following approach. Suppose we want to predict the bid distribution of auctions with  $n$  bidders. We take the data from auctions with  $n'$  bidders,  $n' \neq n$ . We estimate the cost distribution for each  $n'$ . We pool the cost estimates across  $n'$  to estimate the cost distribution and compute the counterfactual equilibrium had there been  $n$  bidders. We compare the counterfactual equilibrium distribution to the sample with  $n$  bidders (without removing outliers).

We first estimate the cost distribution with each of the four non-parametric estimators based on the moment equilibrium approach. As discussed in Section 5.2, these estimators differ in the underlying game-theoretic model (aggregate or individual moment equilibrium) and the treatment of the range beliefs (with or without removing outliers). Let  $(b_t)_{t=1}^{T_n}$  denote the sample of bids from auctions with  $n$  bidders, where  $T_n$  denotes the sample size. To be clear, the  $n$  bidders are all either fringe or non-fringe firms and the sample size differs when removing and not removing outliers. Let  $\hat{l}(n) = \min_{t:1 \leq t \leq T_n} b_t$  denote the sample minimum and  $\hat{u}(n)$  the sample maximum.

We estimate the moment and range beliefs when we want to predict the bid distribution with  $n$  bidders as follows. In the treatment of the auctions with  $n'$  bidders,  $n' \neq n$ , we take the estimates for  $m$  and  $\mu$ , denoted by  $\hat{m}_{n'}$  and  $\hat{\mu}_{n'}$ , from Table 2. The estimate for  $l$  is the lowest  $\hat{l}$  for at most  $n'$  bidders. Formally, we estimate  $l$  by  $\hat{l}_{n'} = \min_{k \in \{2,3,\dots,n'\} \setminus \{n\}} \hat{l}(k)$ . The idea behind this approach is that this estimate for  $l$  is consistent with theory as it decreases in  $n'$ . The estimate for  $u$  is the pooled maximum of the auctions without  $n$  bidders, i.e.,  $\hat{u}_{n'} = \min_{k \neq n} \hat{u}(k)$ . Here, we do not restrict attention to auctions with at most  $n'$  bidders as the theoretical upper bound is the same for any  $n$ . By pooling across the different auctions we can estimate the support more reliably. In the case of estimating  $u$ , we distinguish the maximum before and after removing outliers as based on the interquartile range. We compute the interquartile range with the pooled bids, where we discard bids above the cutoff  $Q_3 + 1.5(Q_3 - Q_1)$ .

We plug the estimated beliefs into the inverse minimax bidding functions  $c^{\text{AGG}}$  and  $c^{\text{IND}}$  as described in Appendix B. The appropriate inverse minimax bidding function maps the sample  $(b_t)_{t=1}^{T_{n'}}$  into a sample of pseudo costs  $(c_t)_{t=1}^{T_{n'}}$ , where  $c_t = c^{\text{AGG}}(b_t | \hat{l}_{n'}, \hat{m}_{n'}, \hat{u}_{n'})$  or  $c_t = c^{\text{IND}}(b_t | \hat{l}_{n'}, \hat{m}_{n'}, \hat{u}_{n'})$ .

We use auctions with at most four bidders for auctions in which all bidders are non-fringe firms and auctions with at most seven bidders in auctions in which all

bidders are fringe firms. We limit the number of bidders per auction due to the sample size. For example, we only observe two auctions in which there are five non-fringe bidders and no fringe bidders. Let  $\bar{n}^{NF} = 4$  and  $\bar{n}^F = 7$  be the respective maxima of the number of bidders for auctions with non-fringe and fringe bidders. We estimate the pseudo costs for each  $n \in \{2, 3, \dots, \bar{n}^\theta\}$  and  $\theta \in \{NF, F\}$ .

We assess the estimates out-of-sample. Let  $\theta \in \{NF, F\}$  and  $n \in \{2, 3, \dots, \bar{n}^\theta\}$ . Suppose we want to predict the bid distribution for auctions with  $n$  bidders. We pool the pseudo costs for all  $n' \in \{2, 3, \dots, \bar{n}^\theta\} \setminus \{n\}$  to get an estimate of the cost distribution. We then compute the bid distributions predicted by the respective moment equilibrium had there been  $n$  bidders. Let  $T$  denote the number of pseudo costs in the estimated cost distribution, i.e.,  $T = \sum_{n' \neq n, 2 \leq n' \leq \bar{n}^\theta} T_{n'}$ . Let  $(c_t)_{t=1}^T$  denote the pooled pseudo costs. The theory implies that the unique consistent upper bound belief  $u^*$  is the highest possible cost, i.e.,  $u_{n\theta}^* = \max c_t$ . For the moment equilibrium with beliefs about individual behavior, we find a fixed point of the function

$$\psi^{\text{IND}}(l, \mu) = \left( \beta^{\text{IND}}(\min c_t | l, \mu, u_{n\theta}^*), \frac{1}{T} \sum_{t=1}^T \beta^{\text{IND}}(c_t | l, \mu, u_{n\theta}^*) \right).$$

The function returns for each  $(l, \mu)$  the corresponding minimum bid and expected bid given the sample of pseudo costs. Note that the computation of the counterfactual is fully non-parametric; no kernel estimation is used. The fixed point of the function  $(l_{n\theta}^*, \mu_{n\theta}^*)$  forms the basis of the out-of-sample prediction because we compute  $\hat{b}_t^{\text{IND}} = \beta^{\text{IND}}(c_t | l_{n\theta}^*, \mu_{n\theta}^*, u_{n\theta}^*)$  for all  $c_t$ ,  $1 \leq t \leq T$ .

In the case of aggregate moment equilibrium, we also use  $u_{n\theta}^* = \max c_t$ . To compute the fixed point of the sample analogue of  $\varphi$ , we need to estimate the distribution of the lowest bid among  $n$  bidders. We could use the pseudo cost sample to estimate the density and use it, together with the empirical c.d.f. of the cost distribution, to compute the estimate of the density of the distribution of the lowest of  $n$  independent cost draws. Instead, we implemented a fully non-parametric and computationally less demanding approach. We compute the probability mass function of the distribution of the minimum of  $n$  independent draws from the discrete distribution of  $(c_t)_{t=1}^T$ . Let  $p_1 = \mathbb{P}(C < x)$ ,  $p_2 = \mathbb{P}(C = x)$ , and  $p_3 = \mathbb{P}(C > x)$ . Then, the probability that the minimum of  $n$  independent draws is  $x$  is

$$\mathbb{P}(\min\{C_1, C_2, \dots, C_n\} = x) = \sum_{j=0}^{n-1} \binom{n}{j} (p_3^j (p_1 + p_2)^{n-j} - (p_2 + p_3)^j p_1^{n-j}).$$

Let  $p^{\min}(x) = \mathbb{P}(\min\{C_1, C_2, \dots, C_n\} = x)$ . We then find a fixed point  $(l_{n\theta}^*, m_{n\theta}^*)$  of the function

$$\varphi^{\text{AGG}}(l, m) = \left( \beta^{\text{AGG}}(\min c_t | l, m, u_{n\theta}^*), \frac{1}{T} \sum_{t=1}^T p^{\min}(c_t) \beta^{\text{AGG}}(c_t | l, m, u_{n\theta}^*) \right).$$

The predicted bid distribution is then  $(\hat{b}_t^{\text{AGG}})_{t=1}^T$ , where  $\hat{b}_t^{\text{AGG}} = \beta^{\text{AGG}}(c_t | l_{n\theta}^*, m_{n\theta}^*, u_{n\theta}^*)$  for all pseudo costs  $c_t$ .

The benchmark for the estimation of the cost distribution is the estimator proposed by Guerre et al. (2000). They assume that bidders play the BNE. Each bid is understood to be the best response to the empirical bid distribution. The estimator depends on the number of bidders, as well as the density and cumulative distribution function of the bid distribution. The estimator for the c.d.f. is the empirical c.d.f. GPV propose a density estimation for the density of the bid distribution. We use the triweight kernel and parameters that GPV use in their Section 2.4. Bidders best respond to the empirical bid distribution  $G$ , with density  $g$ , which clearly depends on whether the auction consists of fringe or non-fringe bidders. The first-order condition of the maximization of expected utility leads to

$$c_{ij} = b_{ij} - \frac{1}{n-1} \frac{1 - G(b_{ij})}{g(b_{ij})}. \quad (12)$$

One can use the first-order condition to transform a sample of bids into a sample of pseudo costs. For a set of pseudo costs, we estimate the density of the cost distribution with a kernel density estimation. We only apply the BNE estimator to the entire sample. Note that the kernel density estimation removes extreme observations to correct for the bias on the boundaries of the support.

For the BNE, we plug in the empirical c.d.f.  $\hat{F}_{(n)}$  of all estimated pseudo costs from all bids from auctions with  $n' \neq n$  bidders. The closed-form solution for a BNE of symmetric procurement auctions is

$$\beta^{\text{BNE}}(c_{ij}) = c_{ij} + \frac{\int_c^{\bar{c}} (1 - F(x))^{n-1} dx}{(1 - F(c_{ij}))^{n-1}}. \quad (13)$$

Table 3 reports the mean bid and the standard deviation (SD) of the sample bid distribution and the predictions of the various theories and estimators. The table suggests that aggregate moment equilibrium predicts the bid distributions rather accurately in auctions in which the bidders are large firms. In auctions in which the bidders are small firms, it performs less well.

To better interpret Table 3, we compute the difference between the mean and



Table 3: Summary statistics in the out-of-sample predictions

|   | Mean    | SD    | Mean    | SD    | Mean    | SD    |
|---|---------|-------|---------|-------|---------|-------|
| <b>Auctions with non-fringe bidders</b> |         |       |         |       |         |       |
|   | $n = 2$ |       | $n = 3$ |       | $n = 4$ |       |
| Sample                                  | 1.071   | 0.079 | 1.046   | 0.090 | 1.018   | 0.096 |
| AGG                                     | 1.062   | 0.073 | 1.059   | 0.088 | 1.046   | 0.097 |
| IND                                     | 1.106   | 0.074 | 1.027   | 0.096 | 1.012   | 0.087 |
| BNE                                     | 1.082   | 0.062 | 1.027   | 0.066 | 1.019   | 0.080 |
| AGG Out.                                | 1.053   | 0.065 | 1.050   | 0.079 | 1.043   | 0.087 |
| IND Out.                                | 1.078   | 0.065 | 1.021   | 0.090 | 1.016   | 0.087 |
| <b>Auctions with fringe bidders</b>     |         |       |         |       |         |       |
|   | $n = 2$ |       | $n = 3$ |       | $n = 4$ |       |
| Sample                                  | 1.098   | 0.100 | 1.049   | 0.219 | 1.032   | 0.144 |
| AGG                                     | 1.111   | 0.121 | 1.066   | 0.128 | 1.066   | 0.145 |
| IND                                     | 1.241   | 0.143 | 1.067   | 0.156 | 1.018   | 0.158 |
| BNE                                     | 1.104   | 0.109 | 1.045   | 0.101 | 1.023   | 0.116 |
| AGG Out.                                | 1.035   | 0.089 | 1.028   | 0.111 | 1.022   | 0.116 |
| IND Out.                                | 1.086   | 0.089 | 1.043   | 0.123 | 1.013   | 0.129 |
|   | $n = 5$ |       | $n = 6$ |       | $n = 7$ |       |
| Sample                                  | 1.050   | 0.165 | 0.969   | 0.120 | 0.973   | 0.133 |
| AGG                                     | 1.050   | 0.145 | 1.058   | 0.155 | 1.054   | 0.161 |
| IND                                     | 0.993   | 0.145 | 0.998   | 0.150 | 0.993   | 0.148 |
| BNE                                     | 1.003   | 0.122 | 1.011   | 0.126 | 1.004   | 0.129 |
| AGG Out.                                | 1.014   | 0.120 | 1.024   | 0.125 | 1.026   | 0.129 |
| IND Out.                                | 0.991   | 0.127 | 0.994   | 0.128 | 0.990   | 0.127 |

Notes: AGG Out. and IND Out. refer to the predictions based on aggregate and individual moment equilibrium after removing outliers in the estimation.

the standard deviation. Specifically, the “moment distance” (MD) is the distance between two vectors of the mean and the standard deviation. Let  $EQ \in \{\text{AGG}, \text{IND}, \text{BNE}\}$  denote the type of equilibrium. Let  $(\hat{m}_n^{EQ}, \hat{s}_n^{EQ})$  be the estimated mean and standard deviation of the predicted equilibrium bid distribution and  $(\hat{m}_n, \hat{s}_n)$  the sample mean and standard deviation of the bid distribution with  $n$  bidders. The statistic  $MD$  is formally defined as

$$MD = \sqrt{(\hat{m}_n^{EQ} - \hat{m}_n)^2 + (\hat{s}_n^{EQ} - \hat{s}_n)^2}.$$

We also report a second criterion that compares not only the first two moments but the entire cumulative distribution functions. Specifically, let  $L^1$  denote the distance between the respective estimated equilibrium bid distribution  $\hat{G}_n^{EQ}$  and the empirical bid distribution  $\hat{G}_n$  as given by

$$L^1 = \int |\hat{G}_n^{EQ}(x) - \hat{G}_n(x)| dx.$$

We also aggregate the two distances across the different  $n$ . We do so by computing the weighted average, where the weights are proportional to the number of observations in Table 2.

Table 4 reports the two notions of distance: the distance in the first two moments (MD) and the  $L^1$  norm. In auctions with non-fringe firms, aggregate moment equilibrium without removing outliers leads to the smallest distances overall. In terms of the (weighted average of the) MD metric, aggregate moment equilibrium performs clearly better than the other theories. There is not much difference between the theories in terms of the  $L^1$  norm. Individual moment equilibrium benefits from removing the outliers. In auctions with small firms, individual moment equilibrium after removing the outliers performs best overall. Note that this is not the first empirical study that finds that the behavior of small and large firms can be rationalized by different models of behavior (Hortaçsu, Luco, Puller, and Zhu, 2019). With small firms, both types of moment equilibrium benefit from removing outliers. This suggests that it is the small and not the large firms that submit more noisy bids at the top of the distribution. Overall, moment equilibrium performs well in the data and is not outperformed by Bayes-Nash equilibrium.

Table 4: Distances between the sample and predicted bid distribution

|   | MD      | $L^1$ | MD      | $L^1$ | MD      | $L^1$ | MD           | $L^1$ |
|---|---------|-------|---------|-------|---------|-------|--------------|-------|
| <b>Auctions with non-fringe bidders</b> |         |       |         |       |         |       |              |       |
|   | $n = 2$ |       | $n = 3$ |       | $n = 4$ |       | Weighted AVG |       |
| AGG                                     | 0.011   | 0.019 | 0.013   | 0.016 | 0.028   | 0.035 | 0.013        | 0.019 |
| IND                                     | 0.035   | 0.035 | 0.020   | 0.024 | 0.011   | 0.018 | 0.027        | 0.029 |
| BNE                                     | 0.020   | 0.018 | 0.030   | 0.022 | 0.017   | 0.016 | 0.024        | 0.019 |
| AGG Out.                                | 0.023   | 0.022 | 0.012   | 0.013 | 0.026   | 0.034 | 0.019        | 0.019 |
| IND Out.                                | 0.016   | 0.017 | 0.024   | 0.026 | 0.009   | 0.015 | 0.019        | 0.021 |
| <b>Auctions with fringe bidders</b>     |         |       |         |       |         |       |              |       |
|   | $n = 2$ |       | $n = 3$ |       | $n = 4$ |       |              |       |
| AGG                                     | 0.025   | 0.033 | 0.093   | 0.064 | 0.034   | 0.045 |              |       |
| IND                                     | 0.149   | 0.144 | 0.066   | 0.053 | 0.020   | 0.022 |              |       |
| BNE                                     | 0.011   | 0.016 | 0.118   | 0.059 | 0.029   | 0.021 |              |       |
| AGG Out.                                | 0.064   | 0.064 | 0.110   | 0.048 | 0.029   | 0.025 |              |       |
| IND Out.                                | 0.017   | 0.015 | 0.096   | 0.054 | 0.024   | 0.026 |              |       |
|   | $n = 5$ |       | $n = 6$ |       | $n = 7$ |       | Weighted AVG |       |
| AGG                                     | 0.020   | 0.019 | 0.096   | 0.090 | 0.085   | 0.084 | 0.056        | 0.055 |
| IND                                     | 0.061   | 0.059 | 0.042   | 0.031 | 0.024   | 0.031 | 0.053        | 0.051 |
| BNE                                     | 0.064   | 0.047 | 0.042   | 0.042 | 0.030   | 0.036 | 0.047        | 0.036 |
| AGG Out.                                | 0.058   | 0.036 | 0.056   | 0.056 | 0.052   | 0.061 | 0.058        | 0.046 |
| IND Out.                                | 0.071   | 0.060 | 0.026   | 0.026 | 0.017   | 0.030 | 0.040        | 0.035 |

Notes: AGG Out. and IND Out. refer to the predictions based on aggregate and individual moment equilibrium after removing outliers in the estimation. MD is the Euclidean distance between the sample mean bid and standard deviation and the predicted mean and standard deviation.  $L^1$  is the  $L^1$  norm between the sample and the predicted bid distribution. Weighted AVG is the average distance weighted by the number of observations in the sample.

## 7 Conclusion

The paper proposes a new way of modeling bidders' beliefs in first-price auctions. The idea is that the bidders' beliefs are characterized by statistics of the bid distribution that are typically observable, such as the expected winning bid or the expected bid. The idea that beliefs are characterized by observables can be taken to other settings. For example, firms might not know the distribution of competitors' costs and inventories but might only observe their prices. In a job search, a worker might know the average wage but not the entire wage distribution.

In our setting, bidders have a belief about the first moment of the distribution of opponents' (aggregate) actions. One can generalize the single belief to a finite number of moments. Intuitively, the model would then nest Bayes-Nash equilibrium, which is as if bidders have an infinite sequence of moment beliefs, as a limiting case. It would be interesting to formalize the relation between moment equilibrium

and Bayes-Nash equilibrium for a general class of games. Moreover, assuming that beliefs are characterized by data from past auctions, which moments of the bid distribution would an auctioneer want to reveal?

A characteristic of moment equilibrium is the presence of strategic uncertainty in equilibrium: Bidders cannot perfectly anticipate the distribution of opponents' actions. Bidders consider many bid distributions to be possible in a moment equilibrium and all these distributions reflect the truth as captured by the moment and range beliefs. However, bidders do not know the (unobservable) value distributions and strategies as in traditional Bayesian treatments of the first-price auction. Players directly forming beliefs about the distribution of other players' actions makes the model more parsimonious, based on observables, and eliminates the need to invert strategies. Inverting the bidding function is often mathematically challenging, particularly in models with multidimensional types or multiple goods. Moment equilibrium might pose a tractable alternative.

We illustrated that (aggregate) moment equilibrium can form the basis of the structural estimation of the latent value distribution. The estimates performed well in out-of-sample predictions in comparison to Bayes-Nash equilibrium. The estimation is simple as only some statistics of observables need to be computed. Hence, moment equilibrium might perform better than Bayes-Nash equilibrium in small samples.

## A Omitted Proofs

### A.1 Proof of Proposition 1

Equipped with a willingness-to-pay  $v$ , bidder  $i$ 's loss of not knowing the distribution  $B$  of the highest bid among  $n$  independent bidders while facing  $n - 1$  of these is

$$\begin{aligned} \sup_{\tilde{b} \in \mathbb{R}_+} \sup_{B \in \Delta \mathbb{R}_+} B(\tilde{b})^{\frac{n-1}{n}} (v - \tilde{b}) - B(b)^{\frac{n-1}{n}} (v - b) \quad (14) \\ \text{s.t. } \int x dB(x) = m \\ \int_l^u dB(x) = 1. \end{aligned}$$

Note that we have swapped the order of taking the supremum, that is, the “best response”  $\tilde{b}$  is chosen and then the bid distribution. The supremum is not affected by the order.

A first observation is that it is never optimal to bid strictly above  $u$  or strictly

lower than  $l$ . To see this, note that bidding below  $l$  means that one loses with certainty. As the lowest value  $\underline{v}$  is at least  $l$ , it is weakly better to bid  $l$ . Similarly, bidding above  $u$  means that one bids strictly higher than all bids of the competitors. One can lower the bid to (marginally above)  $u$  without decreasing the likelihood of winning. Formally, for all possible bid distributions  $B_{-i}$  we have  $\lambda(u, B_{-i}|v_i) < \lambda(b, B_{-i}|v_i)$  for  $u < b$  as  $v_i - u = U_i(u, B_{-i}) > U_i(b, B_{-i}) = v_i - b$ .

The following lemma shows that the worst-case bid distribution has at most two elements in the support.

**Lemma 1.** *For any fixed pair  $(b, \tilde{b})$  in equation (14), the bid distribution that maximizes loss has at most two elements in the support.*

*Proof.* Let  $\tilde{b} < b$ . We observe that loss is higher when there is no mass in  $(\tilde{b}, b)$ . Loss then equals  $B(\tilde{b})^{\frac{n-1}{n}}(b - \tilde{b})$ . We distinguish  $m \leq \tilde{b}$  and  $\tilde{b} < m$ .

If  $m \leq \tilde{b}$ , then the worst case puts all the mass on  $m$ . Loss then equals  $b - \tilde{b}$ . The support is a singleton.

Let  $\tilde{b} < m$  and let  $B$  denote the worst-case bid distribution. Let  $\text{supp}(B)$  denote the support of  $B$ . We show that  $\text{supp}(B) = \{\tilde{b}, u\}$ .

We first consider the mass above  $\tilde{b}$ . Let  $y_1$  be the mass  $B$  places on  $(\tilde{b}, u)$ , and let  $m_1$  be the contribution of that interval to the expected value. Formally, let

$$y_1 = \lim_{\epsilon \searrow 0} \int_{\tilde{b}+\epsilon}^{u-\epsilon} dB = \lim_{\epsilon \searrow 0} B(u - \epsilon) - B(\tilde{b} + \epsilon)$$

and

$$m_1 = \lim_{\epsilon \searrow 0} \int_{\tilde{b}+\epsilon}^{u-\epsilon} x dB(x).$$

By way of contradiction, suppose  $y_1 > 0$ . Then, the mass  $y_1$  can be shifted from  $(\tilde{b}, u)$  to  $\{\tilde{b}, u\}$  without changing the expected value while (strictly) increasing  $B(\tilde{b})$  and (weakly) decreasing  $B(b)$ . Such a shift would increase loss. To see that such a reallocation of mass is feasible, let  $p_1$  denote the mass that is additionally put on  $\tilde{b}$ . We choose  $p_1$  such that  $p_1 \cdot \tilde{b} + (y_1 - p_1) \cdot u = m_1$ , i.e.,  $p_1 = (y_1 u - m_1)/(u - \tilde{b})$ . Note that  $0 < p_1$  if and only if  $m_1 < y_1 u$ , which is true whenever  $y_1 > 0$ . Moreover,  $p_1 \leq y_1$  if and only if  $y_1 \tilde{b} \leq m_1$ , which is also true. Hence,  $B$  cannot be the worst-case distribution. The worst-case bid distribution cannot put mass on  $(\tilde{b}, u)$ . The argument also establishes that  $u$  is the only element in the support of the worst-case bid distribution  $B$  above the average  $m$ , i.e.,  $\text{supp}(B) \cap [m, \infty) = \{u\}$ .

We now show that the worst-case bid distribution also cannot put positive mass on  $[l, \tilde{b})$ . We show that if  $\text{supp}(B) \cap [l, \tilde{b}) \neq \emptyset$ , then one can decrease the mass on  $u$  and increase the mass on  $\tilde{b}$  by removing the mass on  $[l, \tilde{b})$ . There is such a shift in

mass that it does not alter the expected value but instead increases loss. Formally, let  $y_2$  denote the mass below  $\tilde{b}$ , i.e.,  $y_2 = \lim_{\epsilon \searrow 0} \int_l^{\tilde{b}-\epsilon} dB$ , and let  $m_2$  denote the contribution of  $m_2$  to the expected value, i.e.,  $m_2 = \lim_{\epsilon \searrow 0} \int_l^{\tilde{b}-\epsilon} x dB(x)$ . Let  $p_u$  denote the mass on  $u$ . In accordance with the previous paragraph we have

$$m = m_2 + p_u u + (1 - y_2 - p_u)\tilde{b}.$$

If there was only mass on  $\{\tilde{b}, u\}$ , then the distribution puts mass  $p$  on  $\tilde{b}$ , where  $p = (u - m)/(u - \tilde{b})$ . By way of contradiction, assume that  $y_2 > 0$ . The claim is that reducing the mass below  $\tilde{b}$  decreases the mass on  $u$ , that is,  $1 - p \leq p_u$ . The inequality is equivalent to  $-\tilde{b} \leq -m_2 - (1 - y_2)\tilde{b}$ , which is true as  $m_2 \leq y_2\tilde{b}$ . Hence,  $y_2 > 0$  cannot be the case in the worst-case bid distribution. The worst-case bid distribution has only the two elements  $\{\tilde{b}, u\}$  in its support.

Let  $b < \tilde{b}$ . We distinguish three sub-cases:  $b < \tilde{b} < m$ ,  $b < m < \tilde{b}$ , and  $m \leq b$ . We first consider  $b < \tilde{b} < m$ . As the proof is similar to the above arguments, we keep it short. The first insight is that if the distribution puts mass on  $(\tilde{b}, u)$ , loss can be made higher while maintaining the moment constraint, and shift the mass from  $(\tilde{b}, u)$  to  $\{\tilde{b}, u\}$ . This raises the mass on  $\tilde{b}$  and, therefore, loss. The second insight is that one can then shift the mass below  $\tilde{b}$  to  $\{\tilde{b}, u\}$ . This increases the mass on  $\tilde{b}$  and decreases the mass on  $u$ . Moreover, it decreases the mass below  $b$ . The only elements in the support of the worst-case bid distribution can be  $\{\tilde{b}, u\}$ .

The second sub-case is characterized by  $b < m < \tilde{b}$ . All the mass can be put on  $m$ . This maximizes loss as it keeps  $B(b)$  down to 0 and raises  $B(\tilde{b})$  to 1.

The third sub-case has  $m \leq b < \tilde{b}$ . Now there must be mass below  $b$  as there must be mass below  $m$ . The mass that is required below  $m$  to satisfy the moment constraint is lower the lower the elements below  $m$  in the support. Hence, we can increase  $B(\tilde{b})$  and lower  $B(b)$  by reallocating the mass from  $(l, \tilde{b})$  to  $\{l, \tilde{b}\}$ . This can be done without changing the expected value. What is left to show is that there is no mass above  $\tilde{b}$  in the worst case. Such a mass could be reallocated to  $\{l, \tilde{b}\}$ . This would decrease the mass required on  $l$  to satisfy the moment constraint, and therefore decrease  $B(b)$ . Moreover, it would increase  $B(\tilde{b})$ .  $\square$

Now that we know that the worst-case bid distribution only has two bids  $x_1$  and  $x_2$  in its support, we can maximize loss for a given bid  $b$ . The first thing to note is that the “best response”  $\tilde{b}$  is either  $x_1$  or  $x_2$  in the worst case. The nature of the first-price auction ensures that the optimal response to a discrete distribution is always to bid just slightly above a mass point. Note that, strictly speaking, there is no best response. This poses no problem as we operate in the payoff space and

consider suprema. The best response  $\tilde{b}$  is, therefore, either  $x_1$  or  $x_2$ . If it was neither of them, then loss would be zero as one does not wish to win the auction.

With two bids in its support, the distribution of the maximum of  $n$  bids is  $q \cdot [x_1] + (1 - q) \cdot [x_2]$  with  $l \leq x_1 \leq m < x_2 \leq u$ . The moment constraint  $q \cdot x_1 + (1 - q) \cdot x_2 = m$  implies

$$q = \frac{x_2 - m}{x_2 - x_1}.$$

A single competing bidder bids  $x_1$  with probability  $q^{\frac{1}{n}}$  and  $x_2$  with probability  $1 - q^{\frac{1}{n}}$ . Bidder  $i$  wins if she bids higher than the maximum of  $n - 1$  draws. The maximum of  $n - 1$  draws is  $x_1$  with probability  $q^{\frac{n-1}{n}}$  and  $x_2$  with probability  $1 - q^{\frac{n-1}{n}}$ .

To simplify notation, we define the probability  $p$  as

$$p(x_1, x_2) = q^{\frac{n-1}{n}} = \left( \frac{x_2 - m}{x_2 - x_1} \right)^{\frac{n-1}{n}}.$$

We will often drop the arguments. The first derivatives are

$$\begin{aligned} \frac{\partial p(x_1, x_2)}{\partial x_1} &= \frac{n-1}{n} \frac{p}{x_2 - x_1}, \text{ and} \\ \frac{\partial p(x_1, x_2)}{\partial x_2} &= \frac{n-1}{n} p \frac{m - x_1}{(x_2 - x_1)(x_2 - m)}. \end{aligned}$$

As both derivatives are positive for the non-degenerate case  $x_1 < m < x_2$ , the probability  $p$  is increasing in  $x_1$  and  $x_2$ , respectively.

We distinguish two primary cases. The first case features  $\tilde{b} < b$  and the second case is characterized by  $\tilde{b} > b$ . Note that  $\tilde{b} = b$  does not maximize loss as it leads to a loss of 0.

The bid  $b$  can be too high, i.e., the best response  $\tilde{b}$  is less than  $b$ . There are the following three exhaustive cases: (i)  $x_1 < b < x_2$  and  $x_1$  is the best response, (ii)  $x_2 < b$  and  $x_1$  is the best response, and (iii)  $x_2 < b$  and  $x_2$  is the best response.

In the first case ( $x_1 < b < x_2$  and  $x_1$  the best response), loss equals

$$\lambda_1^H(x_1, x_2) = p \cdot (v - x_1) - p \cdot (v - b) = p \cdot (b - x_1).$$

Loss is increasing in  $x_2$  as the first derivative of loss with respect to  $x_2$  is positive. The first derivative of loss with respect to  $x_1$  equals

$$\frac{\partial \lambda_1^H}{\partial x_1} = \frac{\partial p}{\partial x_1} (b - x_1) - p = \frac{n-1}{n} p \frac{b - x_1}{x_2 - x_1} - p.$$

Assuming that  $p > 0$ , this expression is negative if and only if  $(n - 1)(b - x_1) < n(x_2 - x_1)$ , which is true as  $b < x_2$  by assumption. Note that  $p = 0$  does not maximize loss. Hence, the worst-case bid distribution has  $x_1 = l$  and  $x_2 = u$ . Conditional worst-case loss equals

$$\bar{\lambda}_1^H(b) = \left( \frac{u - m}{u - l} \right)^{\frac{n-1}{n}} (b - l). \quad (15)$$

In the second case ( $x_2 < b$  and  $x_1$  the best response), loss equals

$$\lambda_2^H(x_1, x_2) = p \cdot (v - x_1) - (v - b). \quad (16)$$

This case does not maximize loss. By placing  $x_2$  above  $b$  loss is made higher and changed to  $\lambda_1^H$ . Note that loss  $\lambda_2^H$  increases in  $x_2$ , so  $x_2 < b$  poses a binding constraint. The constraint can be avoided by placing  $x_2$  above  $b$ , which leads to the first case.

In the third case ( $x_2 < b$  and  $x_2$  the best response), loss equals

$$\lambda_3^H(x_1, x_2) = b - x_2.$$

Loss is maximized by  $x_2 = m$ . We now show that worst-case loss of the first case is higher than  $b - m$ , i.e.,  $\bar{\lambda}_1^H \geq b - m$ . We show that

$$b - m \leq \frac{u - m}{u - l} (b - l) \leq \left( \frac{u - m}{u - l} \right)^{\frac{n-1}{n}} (b - l) = \bar{\lambda}_1^H.$$

The second inequality is true as  $q \leq q^{\frac{n-1}{n}}$  for  $q \in (0, 1)$ . The first inequality is true for  $b \leq u$ . To summarize, the worst-case loss conditional on  $b$  being too high is  $\bar{\lambda}_1^H$ .

The bid  $b$  can also be too low, i.e., the best response is higher than  $b$ . There are again three exhaustive cases to consider: (i)  $b < x_1$  and  $x_1$  the best response, (ii)  $b < x_1$  and  $x_2$  the best response, and (iii)  $x_1 < b < x_2$  and  $x_2$  the best response. Note that the first two cases require  $b < m$ , so only the third applies to the case in which one bids above the average  $m$ .

In the first case of bidding too low ( $b < x_1$  and  $x_1$  the best response), loss equals

$$\lambda_1^L(x_1, x_2) = p \cdot (v - x_1).$$

Loss increases in  $x_2$ , so  $x_2 = u$  and  $p > 0$  in the worst case. The first derivative



with respect to  $x_1$  equals

$$\frac{\partial \lambda_1^L}{\partial x_1} = \frac{n-1}{n} \frac{p}{x_2 - x_1} (v - x_1) - p.$$

The first derivative is negative if  $x_1 < nx_2 - (n-1)v$  and positive if  $x_1 > nx_2 - (n-1)v$ . Hence, loss is convex in  $x_1$ . Recall that we have  $x_2 = u$  in the worst case. Loss is decreasing in  $x_1$  if  $x_1 \leq m \leq nu - (n-1)v$  (low value  $v$ ). Loss is then maximized by  $x_1$  as low as possible, that is, by  $x_1 = b$ . For higher values, the convexity of loss matters. The maximum of the loss is attained on the boundary, so  $x_1$  is either  $b$  or  $m$ .

In the second case ( $b < x_1$  and  $x_2$  the best response), loss equals  $v - x_2$ . It is maximized by  $x_2 = m$ . Note that the same loss can be achieved in the first case by  $x_1 = m$ . This is, however, not always optimal. Thus, the second case is subsumed by the first.

In the third case ( $x_1 < b < x_2$  and  $x_2$  the best response), loss equals

$$\lambda_3^L(x_1, x_2) = v - x_2 - p \cdot (v - b).$$

Loss decreases in  $x_1$  and  $x_2$ , respectively. Hence, the worst-case bid distribution has the two mass points  $x_1 = l$  and  $x_2 = \max\{b, m\}$ .

To summarize, the worst-case loss conditional on bidding too low is

$$\bar{\lambda}^L(b) = \begin{cases} \max\left\{\left(\frac{u-m}{u-b}\right)^{\frac{n-1}{n}} (v-b), v-m\right\} & \text{if } b < m \\ (v-b) \left(1 - \left(\frac{b-m}{b-l}\right)^{\frac{n-1}{n}}\right) & \text{if } m \leq b. \end{cases}$$

The maximal loss when bidding  $b$  is then  $\max\{\bar{\lambda}_1^H(b), \bar{\lambda}^L(b)\}$ . The maximal loss of bidding too high  $\bar{\lambda}_1^H$  is 0 when bidding  $l$  and increases in  $b$ . The maximal loss of bidding too low is strictly positive at  $b = l$  and is weakly decreasing in  $b$ . Thus, maximal loss is minimized by equalizing the two conditional maximal losses. The monotonicity implies that there is a unique solution.

## A.2 Proof of Proposition 2

To study the comparative statics, we introduce the three implicit equations

$$\begin{aligned} F_1(v, b, l, m, u, n) &:= p(l, u) \cdot (b - l) - p(b, u) \cdot (v - b) = 0, \\ F_2(v, b, l, m, u, n) &:= p(l, u) \cdot (b - l) - (v - m) = 0, \text{ and} \\ F_3(v, b, l, m, u, n) &:= p(l, u) \cdot (b - l) - (v - b) \cdot (1 - p(l, b)) = 0, \end{aligned}$$

all of which are  $\lambda^H - \lambda^L$  for different values  $v$ . The first two apply if  $b < m$  and the last applies if  $m \leq b$ . The roots of these equations are the minimax bids. The implicit function theorem states that

$$\frac{\partial \beta}{\partial x} = -\frac{\partial F_k}{\partial x} \cdot \left( \frac{\partial F_k}{\partial b} \right)^{-1},$$

where  $k = 1, 2, 3$  defines the relevant region for  $v$  and  $x \in \{v, l, m, u, n\}$ . Note that  $\frac{\partial p}{\partial m} < 0$ .

We have

$$\begin{aligned} \frac{\partial F_1}{\partial v} &= -p(b, u) &< 0 \\ \frac{\partial F_1}{\partial b} &= p(l, u) - \frac{\partial p(b, u)}{\partial b}(v - b) + p(b, u) &> 0 \\ \frac{\partial F_1}{\partial l} &= \frac{\partial p(l, u)}{\partial l}(b - l) - p(l, u) &< 0 \\ \frac{\partial F_1}{\partial m} &= \frac{\partial p(l, u)}{\partial m}(b - l) - \frac{\partial p(b, u)}{\partial m}(v - b) &= 0 \\ \frac{\partial F_1}{\partial u} &= \frac{\partial p(l, u)}{\partial u}(b - l) - \frac{\partial p(b, u)}{\partial u}(v - b) &> 0. \end{aligned}$$

To see why the derivative with respect to  $b$  is positive, recall that loss associated with bidding too low is decreasing in  $x_1$ . Hence, the partial derivative of the loss of bidding too low with respect to  $b$  ( $x_1$ ) is negative.

We show that  $\frac{\partial F_1}{\partial m} = 0$ . We consider the case in which the minimax bid is found by solving

$$\left( \frac{u - m}{u - l} \right)^{\frac{n-1}{n}} (b - l) - \left( \frac{u - m}{u - b} \right)^{\frac{n-1}{n}} (v - b) = 0.$$

We divide by  $(u - m)^{\frac{n-1}{n}}$  and observe that the derivative with respect to  $m$  is then 0.

We now show that  $\frac{\partial F_1}{\partial u} > 0$ . Note that for these relatively low values the inverse bid equals

$$v = b + \left( \frac{u - b}{u - l} \right)^{\frac{n-1}{n}} (b - l)$$

We plug the inverse minimax bid into the first partial derivative and obtain

$$\begin{aligned}
\frac{\partial F_1}{\partial u} &= \frac{\partial p(l, u)}{\partial u}(b-l) - \frac{\partial p(b, u)}{\partial u}(v-b) \\
&= \frac{n-1}{n} \frac{b-l}{u-m} \cdot \left( p(l, u) \frac{m-l}{u-l} - p(b, u) \frac{m-b}{u-b} \left( \frac{u-b}{u-l} \right)^{\frac{n-1}{n}} \right) \\
&= \frac{n-1}{n} \frac{b-l}{u-m} \cdot \left( p(l, u) \frac{m-l}{u-l} \left( \frac{u-l}{u-l} \right)^{\frac{n-1}{n}} - p(b, u) \frac{m-b}{u-b} \left( \frac{u-b}{u-l} \right)^{\frac{n-1}{n}} \right).
\end{aligned}$$

The sign of the partial derivative is determined by the sign of  $h(l) - h(b)$ , where  $h$  is defined as

$$h(x) := p(x, u) \left( \frac{u-x}{u-l} \right)^{\frac{n-1}{n}} \frac{m-x}{u-x} = \left( \frac{u-m}{u-l} \right)^{\frac{n-1}{n}} \frac{m-x}{u-x}.$$

The first derivative of  $h$  with respect to  $x$  is  $h'(x) = \left( \frac{u-m}{u-l} \right)^{\frac{n-1}{n}} \frac{-(u-m)}{(u-x)^2} < 0$ . We conclude that  $h(l) - h(b) > 0$ , so  $\frac{\partial F_1}{\partial u} > 0$ .

Note that  $\frac{\partial F_1}{\partial l} < 0$  as  $\lambda_1^H$  decreases in  $x_1$ .

We recall that  $p = q^{\frac{n-1}{n}}$  for taking the partial derivative with respect to  $n$ , i.e.,

$$\begin{aligned}
\frac{\partial F_1}{\partial n} &= \frac{1}{n^2} \cdot p(l, u) \cdot \log(q(l, u)) \cdot (b-l) - \frac{1}{n^2} \cdot p(b, u) \cdot \log(q(b, u)) \cdot (v-b) \\
&= \frac{1}{n^2} \cdot (p(l, u) \cdot \log(q(l, u)) \cdot (b-l) - p(b, u) \cdot \log(q(b, u)) \cdot (v-b)) \\
&= \frac{1}{n^2} \cdot \left( p(l, u) \cdot \log(q(l, u)) \cdot (b-l) - p(b, u) \cdot \log(q(b, u)) \cdot (b-l) \cdot \left( \frac{u-b}{u-l} \right)^{\frac{n-1}{n}} \right) \\
&= \frac{1}{n^2} \cdot (b-l) \cdot p(l, u) \cdot (\log(q(l, u)) - \log(q(b, u)))
\end{aligned}$$

Note that we plug in the inverse minimax bid. As  $\log(q(x, u))$  increases in  $x$ , the first partial derivative is negative.

For intermediate values  $v$ , the partial derivatives are

$$\begin{aligned}
\frac{\partial F_2}{\partial v} &:= -1 &< 0 \\
\frac{\partial F_2}{\partial b} &:= p(l, u) &> 0 \\
\frac{\partial F_2}{\partial l} &:= \frac{\partial p(l, u)}{\partial l} (b - l) - p(l, u) &< 0 \\
\frac{\partial F_2}{\partial m} &:= \frac{\partial p(l, u)}{\partial m} (b - l) + 1 &< 0 \\
\frac{\partial F_2}{\partial u} &:= \frac{\partial p(l, u)}{\partial u} (b - l) &> 0 \\
\frac{\partial F_2}{\partial n} &:= \frac{1}{n^2} \cdot (b - l) \cdot p(l, u) \cdot \log(q(l, u)) &< 0.
\end{aligned}$$

The partial derivative with respect to  $l$  is negative as the loss of bidding too high decreases in  $x_1 = l$ . We show that  $\frac{\partial F_2}{\partial m} < 0$ . Recall that this case applies for the intermediate  $v$  that bid below the mean. In particular, loss of bidding too low is decreasing and increasing in  $x_1$  for  $x_1 \in [b, m]$  only if  $v \geq (n \cdot u - m)/(n - 1)$ . Moreover, note that choosing  $b$  such that  $p(l, u)(b - l) = v - m$  implies that  $b - l = (v - m)/p(l, u)$ . The first partial derivative with respect to  $m$  simplifies to

$$\begin{aligned}
\frac{\partial F_2}{\partial m} &= \frac{\partial p(l, u)}{\partial m} (b - l) + 1 \\
&= \frac{n - 1}{n} \left( \frac{u - m}{u - l} \right)^{\frac{n-1}{n} - 1} \frac{-1}{u - l} (b - l) + 1 \\
&= \frac{n - 1}{n} \left( \frac{u - m}{u - l} \right)^{\frac{n-1}{n} - 1} \frac{-1}{u - l} (v - m) \left( \frac{u - l}{u - m} \right)^{\frac{n-1}{n}} + 1 \\
&= -\frac{n - 1}{n} \frac{v - m}{u - m} + 1 \\
&\leq -\frac{n - 1}{n} \frac{1}{u - m} \left( \frac{n \cdot u - m}{n - 1} - m \right) + 1 \\
&= -\frac{n - 1}{n} \frac{1}{u - m} \frac{n(u - m)}{n - 1} + 1 = 0,
\end{aligned}$$

where we use the lower bound on  $v$  for the inequality.

The partial derivatives for high types that bid above the expected winning bid

$m$  are

$$\begin{aligned}
\frac{\partial F_3}{\partial v} &:= -(1 - p(l, b)) &< 0, \\
\frac{\partial F_3}{\partial b} &:= p(l, u) + (1 - p(l, b)) + (v - b) \frac{\partial p(l, b)}{\partial b} &> 0 \\
\frac{\partial F_3}{\partial m} &:= \frac{\partial p(l, u)}{\partial m} (b - l) + (v - b) \frac{\partial p(l, b)}{\partial m} &< 0 \\
\frac{\partial F_3}{\partial u} &:= \frac{\partial p(l, u)}{\partial u} (b - l) &> 0.
\end{aligned}$$

The partial derivative with respect to  $n$  is negative,  $\frac{\partial F_3}{\partial n} < 0$ , as  $p$  decreases in  $n$ .

The following numerical example demonstrates the ambiguous impact of  $l$  on the minimax bid for high types. Let  $n = u = 2$  and  $m = 1$ . We consider two values:  $v = 2$  and  $v' = 5$ . We have  $\beta(v = 2|l = 0) = 1.05$  but  $\beta(v = 2|l = 0.1) = 1.06$ . The minimax bid increases in  $l$ . On the other hand, we have  $\beta(v = 5|l = 0) = 1.69$  but  $\beta(v = 5|l = 0.1) = 1.68$ . The minimax bid decreases in  $l$ .

To summarize, the minimax bid increases in  $v$ ,  $m$ , and  $n$ . It decreases in  $u$ . For types that bid below the average winning bid  $m$ , the minimax bid increases in  $l$ . The impact of  $l$  on high values' bids is ambiguous.

### A.3 Proof of Theorem 1

The only consistent lower bound belief is  $l^* = \underline{v}$ . To see this, we first look for the value  $v$  whose minimax bid is  $l$  given arbitrary beliefs  $(l, m, u)$ ,  $l < m < u$ . The worst-case loss associated with bidding too high (Eq. (15)) is 0 if  $b = l$ . The maximal loss associated with bidding too low (Eq. (A.1)) is  $\max\{v - m, (v - l) \cdot p(l, u)\}$ . The bid  $b = l$  minimizes worst-case loss if the maximal loss associated with bidding too low is also 0. This cannot be the case for  $v > m$ . The only way that the worst-case loss of bidding too low is 0 is if  $v = l$ . The only way to satisfy this equality in a monotone equilibrium is  $l^* = \underline{v}$ .

For  $l^* = \underline{v}$  and any  $m \leq \bar{v}$ , there exists a unique upper bound belief  $u(m)$  such that the bidding function  $\beta(v|l^*, m, u(m))$  is strictly increasing in  $v$ . The bidding function is strictly increasing if the bid  $u$  equalizes the maximal loss of bidding too low and the maximal loss of bidding too high, i.e., if  $\bar{\lambda}^L(u) = \bar{\lambda}^H(u)$  for the highest type  $\bar{v}$ . We solve for a root of the function

$$\bar{\lambda}^L(u) - \bar{\lambda}^H(u) = (\bar{v} - u)(1 - p(l, u)) - p(l, u)(u - l) = \bar{v} - u - p(l, u)(\bar{v} - l).$$

in  $u$ . The function is clearly strictly decreasing in  $u$ . The function equals  $\bar{v} - m > 0$

for  $u$  as small as possible ( $u = m$ ) and is negative for  $u$  as large and plausible values (e.g.,  $u = \bar{v}$ ). Hence, for any  $m < \bar{v}$  there is exactly one consistent upper bound belief  $u(m)$ . The implicit function theorem allows us to show that

$$\begin{aligned} u'(m) &= - \left( \frac{\partial \bar{v} - u - p(l, u)(\bar{v} - l)}{\partial m} \right) \cdot \left( \frac{\partial \bar{v} - u - p(l, u)(\bar{v} - l)}{\partial u} \right)^{-1} \\ &= - \frac{\partial p(l, u)}{\partial m} (\bar{v} - l) \left( 1 + \frac{\partial p(l, u)}{\partial u} (\bar{v} - l) \right)^{-1} > 0. \end{aligned}$$

The consistent upper bound belief  $u(m)$  is increasing in  $m$ .

Suppose there is a single type, i.e.,  $\underline{v} = \bar{v}$ . Then the equalities  $\underline{v} = l^* = m^* = u^*$  must hold in an aggregate moment equilibrium.

Let  $\underline{v} < \bar{v}$ . We now show that  $\varphi^{\text{AGG}}$  must have an interior fixed point, that is,  $m^*$  is such that  $\underline{v} < m^* < \bar{v}$ . Recall from Proposition 1 that there is a cutoff type, now denoted by  $\hat{v}(m)$ ,

$$\hat{v}(m) = m + (m - l) \left( \frac{u - m}{u - l} \right)^{\frac{n-1}{n}},$$

such that types below the cutoff bid below  $m$  and types above the cutoff bid above  $m$ . Observe that as  $m$  converges to  $l$ , which equals  $\underline{v}$  in equilibrium, the cutoff converges to  $\underline{v}$ . Hence, as  $m$  converges to  $\underline{v}$ , an increasing share of types bid above  $m$ . As only types close to  $\underline{v}$  bid close to  $m$ , the expected bid must be higher than  $m$ .

## A.4 Proof of Proposition 3

The proof first shows that the worst-case bid distributions only put mass on at most two bids (Lemma 2). These bids are the same for all other bidders in the worst case (Lemma 3). We then maximize loss with respect to these two bids. Loss may again come from bidding too high and too low. The minimax bid equalizes the highest loss from bidding too high and the highest loss from bidding too low.

**Lemma 2.** *The bid distribution of bidder  $j$  that maximizes bidder  $i$ 's loss has at most two elements in the support.*

*Proof.* Recall that we maximize loss with respect to  $\tilde{b}_i$  and the bid distributions  $(B_k)_{k \neq i}$ . We first change the order of taking suprema, i.e., we fix  $(b_i, \tilde{b}_i)$ . Then we consider any bid distributions of bidders  $k \neq i, j$  and obtain the maximization

problem

$$\sup_{B_j \in \mathcal{B}_{ij}(l_i, \mu_i, u_i)} \int \mathbb{E}_{B_{-i,j}} u_i(\tilde{b}_i, b_j, b_{-i,j}) - u_i(b_i, b_j, b_{-i,j}) dB_j(b_j) \quad (17)$$

The extreme points of the constraint set  $\mathcal{B}_{ij}(l_i, \mu_i, u_i)$  are distributions with at most two elements in the support (Winkler, 1988). The objective function is linear in  $B_j$ . The maximum of a linear function is attained at an extreme point.  $\square$

The lemma states that the worst-case distribution  $B_j$  has at most two elements in its support.<sup>15</sup> Thus, it suffices to consider bid distributions of the form  $\pi \cdot [x_1] + (1 - \pi) \cdot [x_2]$ , where  $\pi \in [0, 1]$  and  $l \leq x_1 \leq x_2 \leq u$ . The first moment constraint implies  $\pi \cdot x_1 + (1 - \pi) \cdot x_2 = \mu$ , which clearly requires  $x_1 \leq \mu \leq x_2$ . The moment constraint implies

$$\pi(x_1, x_2) = \frac{x_2 - \mu}{x_2 - x_1}. \quad (18)$$

We sometimes drop the arguments.

Let  $\pi_1$  and  $\pi_2$  denote the partial derivatives of  $\pi$  with respect to  $x_1$  and  $x_2$ , respectively. These derivatives are equal to

$$\pi_1 = \frac{\partial \pi}{\partial x_1} = \frac{x_2 - \mu}{(x_2 - x_1)^2} \text{ and } \pi_2 = \frac{\partial \pi}{\partial x_2} = \frac{\mu - x_1}{(x_2 - x_1)^2}.$$

Observe that  $\pi_1 > 0$  and  $\pi_2 > 0$  for  $x_1 < \mu < x_2$ .

**Lemma 3.** *The worst-case bid distributions of the other bidders are identical.*

*Proof.* The worst-case bid distribution for each other bidder  $j$  has at most two elements,  $x_{j1}$  and  $x_{j2}$ , in the support. The best response to a discrete bid distribution is to bid (slightly above) one of the mass points. Tie-breaking does not play a role as suprema are taken. With the current moment constraints and for a given pair  $(\tilde{b}, b)$  as in (17), the maximization of loss equals

$$\begin{aligned} \sup_{x_{j1}, x_{j2}: l \leq x_{j1} \leq \mu \leq x_{j2} \leq u} & (v - \tilde{b}) \prod_{j \neq i} (\pi(x_{j1}, x_{j2}) \chi_{x_{j1} \leq \tilde{b}} + (1 - \pi(x_{j1}, x_{j2})) \chi_{x_{j2} \leq \tilde{b}}) \\ & - (v - b) \prod_{j \neq i} (\pi(x_{j1}, x_{j2}) \chi_{x_{j1} < b} + (1 - \pi(x_{j1}, x_{j2})) \chi_{x_{j2} < b}), \end{aligned}$$

where we break ties in favor of higher loss and  $\chi_A$  denotes the indicator function, i.e., the function that takes the value 1 if  $A$  is true and 0 otherwise.

<sup>15</sup>One can generalize the lemma. Let  $M$  be the number of moment beliefs, where a moment belief is the expectation of a measurable function. The worst-case bid distribution has at most  $M + 1$  elements in the support.

If  $\tilde{b} < \mu$ , then loss is maximized by  $x_{j_1} = \tilde{b}$  as this maximizes the mass on  $\tilde{b}$ . The second mass point  $x_{j_2}$  is then chosen equal to  $u$  as this maximizes the mass on  $\tilde{b}$  and is higher than  $b$ .

Let  $\mu \leq \tilde{b}$ . If  $\tilde{b} < b$ , then loss is maximized by  $x_{j_2} = \tilde{b}$  and  $x_{j_1} = l$  as this maximizes the mass on  $\tilde{b}$ . If  $\mu < b < \tilde{b}$ , then the same construction maximizes loss. If  $b < \mu < \tilde{b}$ , then loss is again maximized by  $x_{j_2} = \tilde{b}$ , but now  $x_{j_1} = b$ . The bid  $b$  always loses, while the bid  $\tilde{b}$  wins with certainty. Thus, loss is always highest with symmetric worst-case distributions.  $\square$

Due to symmetric worst-case beliefs, the maximization of loss takes the form

$$\begin{aligned} \sup_{x_1, x_2} \max\{\pi(x_1, x_2)^{n-1}(v - x_1), v - x_2, 0\} - (v - b)(\chi_{x_1 < b < x_2} \pi(x_1, x_2)^{n-1} + \chi_{x_2 < b}) \\ \text{s.t. } l \leq x_1 \leq \mu \leq x_2 \leq u, \end{aligned}$$

for a given  $b \in [l, u]$ . Bidding above value is dominated, so we consider  $b \leq v$ . The case in which the highest possible utility is 0 is then irrelevant as it never maximizes loss.

The maximization of loss requires the discussion of six exhaustive cases. There are three cases in which the bid  $b$  is too high as the best response is lower. There are also six cases in which the bid  $b$  is too low as the best response is higher. In what follows, we maximize loss for each of these six cases with respect to  $x_1$  and  $x_2$ . We then find the minimax bids.

**Loss of bidding too high.** There are three cases in which one bids too high, that is, when the best response to the bid distribution is lower than  $b$ . The first case has  $x_1 < b < x_2$  and  $x_1$  is the best response. The second case has  $x_2 < b$  with  $x_1$  as the best response. The third case has  $x_2 < b$  with  $x_2$  being the best response.

Consider the first case of bidding too high, i.e.,  $x_1 < b < x_2$  and  $x_1$  the best response. Loss equals  $\lambda^{H1}(x_1, x_2) = \pi^{n-1}(b - x_1)$ , which is to be maximized with respect to  $x_1$  and  $x_2$  under the constraints  $l \leq x_1 \leq \mu < x_2 \leq u$  and  $x_1 < b < x_2$ .

We show that  $\tilde{x}_1 = \max\{l, \min\{\hat{x}_1(b, u), \mu\}\}$  and  $x_2 = u$  maximize loss, where

$$\hat{x}_1(b, x_2) = \begin{cases} \frac{(n-1)b-x_2}{n-2} & \text{for } n \geq 3 \\ l & \text{for } n = 2. \end{cases}$$

Note that  $\hat{x}_1(b, x_2) < b$  for  $b < x_2$ . The partial derivatives of  $\lambda^{H1}$  with respect to  $x_1$



and  $x_2$  equal

$$\lambda_1^{H1}(x_1, x_2) = \pi^{n-1} \left( (n-1) \frac{b-x_1}{x_2-x_1} - 1 \right) \text{ and}$$

$$\lambda_2^{H1}(x_1, x_2) = (n-1)\pi^{n-2}\pi_2(b-x_1),$$

respectively. Loss increases in  $x_2$  as  $\lambda_2^{H1} > 0$ . Thus, loss is maximized by  $x_2 = u$ . For  $n = 2$ , loss decreases in  $x_1$  as  $\lambda_1^{H1} < 0$ . It is, therefore, maximized by  $\tilde{x}_1 = l = \max\{l, \min\{\hat{x}_1(b, u), \mu\}\}$ . For  $n \geq 3$ , the root of  $\lambda_1^{H1}$  is  $\hat{x}_1(b, x_2)$ . The second derivative of  $\lambda^{H1}$  with respect to  $x_1$  equals

$$\frac{\partial^2 \lambda^{H1}}{\partial x_1^2} = (n-1)\pi^{n-1} \frac{bn - (n-2)x_1 - 2x_2}{(x_2 - x_1)^2},$$

which is negative at  $x_1 = \hat{x}_1(b, x_2)$ . To see this, note that  $bn - (n-2)\hat{x}_1 - 2x_2 < 0$  simplifies to  $b < x_2$ , which is true. Hence, for  $\hat{x}_1(b, u) < l$ , loss decreases in  $x_1$  and is maximized by  $x_1 = l$ . For  $l \leq \hat{x}_1 < \mu$ , loss is maximized by  $x_1 = \hat{x}_1(b, u)$ . For  $\mu \leq \hat{x}_1(b, u)$ , loss increases in  $x_1$  and is maximized by  $x_1 = \mu$ . To summarize, loss is maximized by  $\tilde{x}_1 = \max\{l, \min\{\hat{x}_1(b, u), \mu\}\}$  and  $x_2 = u$ .

The second and the third case of bidding too high both lead to a lower loss than in the first case. To see this, note that loss in the second case equals  $\pi^{n-1}(v-x_1) - (v-b) \leq \pi^{n-1}(b-x_1) = \lambda^{H1}(x_1, x_2)$ . The worst-case distribution never puts a mass point below  $b$  when  $x_1$  is the best response. In the third case, loss equals  $b-x_2$ , which is maximized by  $x_2 = \mu$ . Note that this loss equals the loss of the first case when choosing  $x_1 = \mu$ . This choice is optimal only if  $\hat{x}_1(b, u) > \mu$ , in which case the two cases yield the same maximal loss.

To summarize, the worst-case loss when bidding too high  $\bar{\lambda}^H(b)$  is given by

$$\bar{\lambda}^H(b) = \left( \frac{u-\mu}{u - \max\{l, \min\{\hat{x}(b, u), \mu\}\}} \right)^{n-1} (b - \max\{l, \min\{\hat{x}(b, u), \mu\}\}).$$

Observe that  $\bar{\lambda}^H$  is continuous in  $b$ .

The range and slope of  $\bar{\lambda}^H$  are as follows. First, the loss of bidding too high when bidding the lowest bid is 0, that is,  $\bar{\lambda}^H(l) = 0$  as  $\hat{x}(l, u) \leq l$ . Second,  $\bar{\lambda}^H$  strictly increases in  $b$ . The statement is obvious unless  $\hat{x}(b, u) = \max\{l, \min\{\hat{x}(b, u), \mu\}\}$ . In this case, the derivative is

$$\frac{d\bar{\lambda}^H(b)}{db} = \frac{d}{db} \max_{x_1} \lambda^{H1}(x_1, u) = \pi(\hat{x}(b, u), u)^{n-1}$$

by the envelop theorem.

**Loss of bidding too low.** There are also three cases when one bids too low, i.e., when the best response to the bid distribution is higher than  $b$ . The first case has  $b < x_1$  with  $x_1$  as the best response. The second case has  $b < x_1$  and  $x_2$  is the best response. The third case has  $x_1 < b < x_2$  with  $x_2$  being the best response.

In the first case of bidding too low, loss equals  $\lambda^{L1}(x_1, x_2) = \pi^{n-1}(v - x_1)$  as  $b < x_1$  and  $x_1$  is the best response. Clearly, the case applies only if  $b < \mu$ . Observe that we can draw from the analysis of  $\lambda^{H1}$  as  $\lambda^{L1}$  is  $\lambda^{H1}$  when substituting  $v$  for  $b$ . It follows that the partial derivative  $\lambda_2^{L1}$  is positive, so that loss is maximized by  $x_2 = u$ . For  $n = 2$  and  $v \leq u$ , loss  $\lambda^{L1}$  is decreasing in  $x_1$  as  $\lambda_1^{L1} \leq 0$  so that loss is maximized by  $x_1 = b$  and  $x_2 = u$ . For  $u < v$ , loss is increasing in  $x_1$  so that it is maximized by  $x_1 = \mu$ . Maximal loss then equals  $v - \mu$ . For  $n \geq 3$ , the root of  $\lambda_1^{L1}$  is  $\hat{x}_1(v, x_2)$ . The calculations above show that the second derivative  $\lambda_{11}^{L1}$  is negative at  $\hat{x}_1(v, x_2)$ . Note that  $\hat{x}_1(v, u) \leq v$  if and only if  $v \leq u$ . Hence, for  $v \leq u$ , loss  $\lambda^{L1}$  is maximized by  $x_1 = \max\{b, \min\{\hat{x}_1(v, u), \mu\}\}$  and  $x_2 = u$ . Observe that maximal loss is weakly larger than  $v - \mu = \lambda^{L1}(\mu, u)$ . For  $u < v$ , loss increases in  $x_1$  and is maximized by  $x_1 = \mu$  and  $x_2 = u$ . Note that then  $\mu < u < v \leq \hat{x}_1(v, u)$  so that  $\max\{b, \min\{\hat{x}_1(v, u), \mu\}\} = \mu$  as  $b < \mu$ .

In the second case of bidding too low, loss equals  $v - x_2$  as  $b < x_1$  and  $x_2$  is the best response. Maximal loss equals  $v - \mu$ , which is weakly less than the worst case of the previous case.

In the third case associated with bidding too low, loss equals  $\lambda^{L3}(x_1, x_2) = v - x_2 - \pi^{n-1}(v - b)$  as  $x_1 < b < x_2$  with  $x_2$  as the best response. We show that loss decreases in  $x_1$  and  $x_2$ . The first partial derivatives of  $\lambda^{L3}$  are  $\lambda_1^{L3}(x_1, x_2) = -(n-1)(v-b)\pi^{n-2}\pi_1$  and  $\lambda_2^{L3}(x_1, x_2) = -1 - (n-1)(v-b)\pi^{n-2}\pi_2$ . Both derivatives are negative. As a result, if  $\mu \leq b$ , then loss is maximized by  $x_1 = l$  and  $x_2 = b$ . If  $b < \mu$ , then the worst-case has  $x_2 = \mu$ , so that loss then equals  $v - \mu$ . This loss is less than the worst-case loss  $\lambda^{L1}(\max\{b, \min\{\mu, \hat{x}_1(v, u)\}\}, u)$ .

To summarize, let  $\bar{\lambda}^L$  denote the worst-case loss when bidding too low. The maximal loss  $\bar{\lambda}^L$  equals

$$\bar{\lambda}^L(b) = \begin{cases} \left( \frac{u-\mu}{u-\max\{b, \min\{\hat{x}_1(v, u), \mu\}\}} \right)^{n-1} (v - \max\{b, \min\{\hat{x}_1(v, u), \mu\}\}) & \text{for } b < \mu \\ (v-b) \left( 1 - \left( \frac{b-\mu}{b-l} \right)^{n-1} \right) & \text{for } \mu \leq b. \end{cases}$$

Note that  $\bar{\lambda}^L$  is continuous in  $b$ .

The range and slope of  $\bar{\lambda}^L$  are as follows. First, the maximal loss of bidding too low is positive when bidding the lowest bid, i.e.,  $\bar{\lambda}^L(l) > 0$ . Second, loss  $\bar{\lambda}^L$  is constant in  $b$  if  $b < \min\{\hat{x}(v, u), \mu\}$ . Third, we show that maximal loss

$\bar{\lambda}^L$  is strictly decreasing in  $b$  when  $\min\{\hat{x}(v, u), \mu\} \leq b$ . When  $b < \mu$ , then the maximization of loss  $\lambda^{L1}$  is a constrained maximization problem  $\max_{x_1} \lambda^{L1}(x_1, u)$  subject to  $b \leq x_1 \leq \mu$ . The envelope theorem implies

$$\left. \frac{d\bar{\lambda}^L(b)}{db} \right|_{b < \mu} = \frac{\partial}{\partial b} \lambda^{L1}(x_1, u) - \nu_1(-x_1 + b) - \nu_2(x_1 - \mu) = -\nu_1,$$

where  $\nu_1$  and  $\nu_2$  are the respective Lagrange multipliers in the optimal solution. The derivative is negative as the constraint  $x_1 \geq b$  is binding so that the multiplier  $\nu_1$  must be strictly positive by complementary slackness (Simon and Blume, 1994). For  $\mu \leq b$ , the derivative of  $\bar{\lambda}^L(b) = (v - b)(1 - \pi(l, b)^{n-1})$  with respect to  $b$  is

$$\left. \frac{d\bar{\lambda}^L(b)}{db} \right|_{\mu \leq b} = -(1 - \pi(l, b)^{n-1}) - (n - 1)(v - b)\pi(l, b)^{n-2}\pi_2(l, b) < 0.$$

Finally, the loss of bidding too low is 0 when bidding value, that is,  $\bar{\lambda}^L(v) = 0$ .

**Minimization of maximal loss.** Maximal loss is the upper envelope of  $\bar{\lambda}^L(b)$  and  $\bar{\lambda}^H(b)$ , i.e.,  $\max\{\bar{\lambda}^L(b), \bar{\lambda}^H(b)\}$ . Recall that  $\bar{\lambda}^L(l) > 0$  and  $\bar{\lambda}^L(b)$  decreases in  $b$ . Moreover,  $\bar{\lambda}^H(l) = 0$  and  $\bar{\lambda}^H(b)$  increases in  $b$ . Hence, maximal loss is minimized by the bid that equalizes the two expressions. Such a bid exists in  $[l, v]$  as maximal loss is continuous in  $b$  and  $\bar{\lambda}^L(v) = 0$ .

We now show that certain cases related to the loss of bidding too low are irrelevant at the minimax bid. First, we argue that the minimax bid  $b^*$  is never such that  $b^* < \hat{x}(v, u) \leq \mu$ . Let  $v \leq u$  so that  $\hat{x}(v, u) \leq v$ . Suppose  $l < \hat{x}(v, u) \leq \mu$  so that  $\bar{\lambda}^L(b)$  is constant in  $b$  for  $l \leq b \leq \hat{x}(v, u)$ . We show  $\bar{\lambda}^L(b) > \bar{\lambda}^H(b)$  for all  $b \in [l, \hat{x}(v, u)]$ . Hence, the two functions can only intersect at  $b > \hat{x}(v, u)$ . Recall that for relatively low bids  $\bar{\lambda}^L(b) > \bar{\lambda}^H(b)$ . Let  $b \in [l, \hat{x}(v, u)]$ . Recall that  $x_1 = \hat{x}(v, u)$  maximizes  $\pi(x_1, u)^{n-1}(v - x_1)$ . In particular, for  $x'_1 = \max\{l, \hat{x}(b, u)\}$  we have  $\bar{\lambda}^L(b) = \bar{\lambda}^L(\hat{x}(v, u)) \geq \pi(x'_1, u)^{n-1}(v - x'_1)$ . Hence, we have

$$\bar{\lambda}^L(b) \geq \pi(x'_1, u)^{n-1}(v - x'_1) > \pi(x'_1, u)^{n-1}(b - x'_1) = \bar{\lambda}^H(b),$$

where the second inequality follows from  $v > \hat{x}(v, u) \geq b$ . The equality holds as  $x'_1$  maximizes  $\pi(x_1, u)^{n-1}(b - x_1)$ .

We show that the minimax bid  $b^*$  is never such that  $b^* < \mu$  and  $\bar{\lambda}(b^*) = v - \mu$ . By way of contradiction, suppose this was true. Then we have either  $u < v$  or  $v \leq u$  and  $\hat{x}(v, u) > \mu$ . Let  $b < \mu$  and observe that  $\lambda^{L1}(x_1, u)$  increases in  $x_1$  so that the highest loss of bidding too low is  $v - \mu$ . Note that  $\max\{l, \min\{\hat{x}(\mu, u), \mu\}\} < \mu$ .

We note that  $\bar{\lambda}^H(b) \leq \bar{\lambda}^H(\mu)$  and

$$\begin{aligned}\bar{\lambda}^H(\mu) &= \frac{(u - \mu)^{n-1}}{(u - \max\{l, \min\{\hat{x}(\mu, u), \mu\}\})^{n-1}} (\mu - \max\{l, \min\{\hat{x}(\mu, u), \mu\}\}) \\ &< \frac{(u - \mu)^{n-1}}{(u - \max\{l, \min\{\hat{x}(\mu, u), \mu\}\})^{n-1}} (v - \max\{l, \min\{\hat{x}(\mu, u), \mu\}\}) \\ &= \lambda^{L1}(\max\{l, \min\{\hat{x}(\mu, u), \mu\}\}, u) \leq \bar{\lambda}^L(b) = v - \mu.\end{aligned}$$

The first line states the formula for the maximal loss when the bid  $\mu$  is too high. The second line uses the fact that  $\mu < v$  must hold. The third line observes that the second equals  $\lambda^{L1}$  and that  $\lambda^{L1}$  increases in  $x_1$ . We conclude that  $\bar{\lambda}^H(b) = \bar{\lambda}^L(b)$  can be true only for  $b \geq \mu$  if  $\lambda^{L1}$  increases in  $x_1$  for  $b \leq \mu$ .

The previous two paragraphs show that  $\bar{\lambda}^H$  and  $\bar{\lambda}^L$  never intersect in the interior of the region on which  $\bar{\lambda}^L(b)$  is constant in  $b$ . The worst-case loss of bidding too low  $\bar{\lambda}^L$  simplifies to the relevant expression

$$\bar{\lambda}^L(b) = \begin{cases} \left(\frac{u-\mu}{u-b}\right)^{n-1} (v-b) & \text{for } b < \mu \\ (v-b) \left(1 - \left(\frac{b-\mu}{b-l}\right)^{n-1}\right) & \text{for } \mu \leq b \leq v. \end{cases}$$

Equating the worst-case losses  $\bar{\lambda}^H$  and  $\bar{\lambda}^L$  gives the minimax bid in Proposition 3.

Let  $n = 2$ . The first cutoff value  $\hat{v}_1$  is such that  $\beta^*(\hat{v}_1) = m$ , and the second cutoff value is such that  $\beta^*(\hat{v}_2) = u$ .

## A.5 Proof of Proposition 4

The proof is analogous to the proof of Proposition 2. The minimization of maximal loss takes the form of finding  $b$  such that  $F(v, \mu, b) = \bar{\lambda}^L(b) - \bar{\lambda}^H(b) = 0$ . Note that there is a unique  $b$  that solves the equation for each  $v$  and  $\mu$ .

The proof of Proposition 3 implies that  $\partial F / \partial b < 0$  as  $\bar{\lambda}^L$  decreases in  $b$  and  $\bar{\lambda}^H$  increases in  $b$ .

The IFT implies that the minimax bid  $\beta$  increases in  $v$  as  $\frac{\partial \beta}{\partial v} = -\frac{\partial F}{\partial v} \left(\frac{\partial F}{\partial b}\right)^{-1}$  and  $\frac{\partial F}{\partial v} = \frac{\partial \bar{\lambda}^L(b)}{\partial v} > 0$ .

We show that the minimax bid increases in  $\mu$ . The minimax bid is independent of the mean  $\mu$  for values  $v$  such that  $\beta(v|l, \mu, u) \leq \mu$ . In general, the minimax bid increases in  $\mu$  if  $\partial F / \partial \mu \geq 0$ . For bids above the mean, the maximal loss associated with bidding too low increases in  $\mu$  as

$$\left. \frac{\partial \bar{\lambda}^L}{\partial \mu} \right|_{\mu \leq b} = (n-1) \frac{v-b}{b-l} \left(\frac{b-\mu}{b-l}\right)^{n-2} > 0.$$

We now show that the maximal loss from bidding too high  $\bar{\lambda}^H$  decreases in  $\mu$ . If  $\max\{l, \min\{\hat{x}_1(b, u), \mu\}\} \neq \mu$ , then the maximal loss of bidding too high decreases in  $\mu$  as

$$\left. \frac{\partial \bar{\lambda}^H}{\partial \mu} \right|_{\hat{x}_1(b, u) \leq \mu} = -(n-1)\pi(\max\{l, \hat{x}_1(b, u)\}, u)^{n-2} \frac{b - \max\{l, \hat{x}_1(b, u)\}}{u - \max\{l, \hat{x}_1(b, u)\}} < 0.$$

If  $\mu = \max\{l, \min\{\hat{x}_1(b, u), \mu\}\}$ , then  $\bar{\lambda}^H(b) = b - \mu$  so that the derivative with respect to  $\mu$  is  $-1$ . As a result,  $\partial F / \partial \mu > 0$  and  $\partial \beta / \partial \mu \geq 0$ .

The first derivative of  $\bar{\lambda}^H$  with respect to  $n$  is

$$\frac{\partial \bar{\lambda}^H}{\partial n} = \begin{cases} (b-l) \left(\frac{u-\mu}{u-l}\right)^{n-1} \log\left(\frac{u-\mu}{u-l}\right) & \text{if } \tilde{x}_1 = l \\ \frac{(n-1)(u-b)^2}{(n-2)^2(u-\mu)} \left(\frac{(n-2)(u-\mu)}{(n-1)(u-b)}\right)^n \log\left(\frac{(n-2)(u-\mu)}{(n-1)(u-b)}\right) & \text{if } \tilde{x}_1 = \hat{x}(b, u) \\ 0 & \text{if } \tilde{x}_1 = \mu. \end{cases}$$

The partial derivative is negative.

For bids above the mean, the partial derivative of  $\bar{\lambda}^L$  with respect to  $n$  is

$$\left. \frac{\partial \bar{\lambda}^L}{\partial n} \right|_{\mu \leq b} = -(v-b) \left(\frac{b-\mu}{b-l}\right)^{n-1} \log\left(\frac{b-\mu}{b-l}\right) > 0.$$

Hence, for bids above the mean  $\partial(\bar{\lambda}^L - \bar{\lambda}^H) / \partial n > 0$ .

For bids below the mean  $\mu$ , we consider the inverse minimax bid  $\beta^{-1}$ . The partial derivative of  $\beta^{-1}$  with respect to  $n$  is

$$\left. \frac{\partial \beta^{-1}}{\partial n} \right|_{b < \mu} = \begin{cases} (b-l) \left(\frac{u-b}{u-l}\right)^{n-1} \log\left(\frac{u-b}{u-l}\right) & \text{if } l \geq \hat{x}(b, u) \\ \frac{u-b}{n-2} \left(\frac{n-2}{n-1}\right)^{n-1} \log\left(\frac{n-2}{n-1}\right) & \text{else.} \end{cases}$$

The derivative is negative in both cases. Note that taking the total derivative of  $\beta^{-1}(\beta(v|l, \mu, u)|l, \mu, u) = v$  with respect to  $\mu$  on both sides and reformulating leads to

$$\frac{\partial \beta(v|l, \mu, u)}{\partial \mu} = - \frac{\partial \beta^{-1}(\beta(v|l, \mu, u)|l, \mu, u)}{\partial \mu} \left( \frac{\partial \beta^{-1}(\beta(v|l, \mu, u)|l, \mu, u)}{\partial b} \right)^{-1}.$$

The inverse bid increases in  $b$  and the inverse bid decreases in  $n$ . Thus, the minimax bid increases in  $n$ .

## A.6 Proof of Theorem 3

Let  $l < m < u$ . The inverse bidding function is strictly increasing on  $[l, u]$  as  $l < m < u$ . We use the strictly increasing inverse bidding function  $v^{\text{AGG}}(\cdot|l, m, u)$  to construct the value distribution  $F$  by defining  $v = v^{\text{AGG}}(b|l, m, u)$  for all  $b \in [l, u]$ . Clearly,  $F$  is a c.d.f. with support  $[\underline{v}, \bar{v}]$ , where  $\underline{v} = v^{\text{AGG}}(l|l, m, u) = l$  and  $\bar{v} = v^{\text{AGG}}(u|l, m, u)$ . Note that as  $v^{\text{AGG}}$  is strictly increasing in  $b$ , so is its inverse  $\beta^{\text{AGG}}$  in  $v$ , and  $\beta^{\text{AGG}}(v^{\text{AGG}}(b)) = b$ . Hence,  $F(v) = G(\beta^{\text{AGG}}(v|l, m, u))$  for all  $v \in [\underline{v}, \bar{v}]$ . The equilibrium is separating by construction and has consistent range beliefs by construction. To see that  $G$  is the bid distribution of a moment equilibrium, observe that

$$\int_{\underline{v}}^{\bar{v}} \beta^{\text{AGG}}(v|l, m, u) dF^n(v) = \int_l^u \beta^{\text{AGG}}(v^{\text{AGG}}(b|l, m, u)|l, m, u) dG^n(b) = \int_l^u b dG^n(b) = m.$$

Hence, the moment condition holds, and  $(\beta^{\text{AGG}}, l, m, u)$  is a separating aggregate moment equilibrium bid distribution  $G$  when the value distribution is  $F$  as constructed.

Let there be a value distribution  $F$  and  $G$  a corresponding bid distribution of a separating aggregate moment equilibrium. Separation and  $\underline{v} < \bar{v}$  implies  $l < m < u$ .

## A.7 Proof of Proposition 5

To prove that if  $b$  is a minimax bid for  $l_0, \mu_0, u_0$ , and  $v$ , then  $p \cdot b + q$  is a minimax bid for  $p \cdot l_0 + q, p \cdot \mu_0 + q, p \cdot u_0 + q$ , and  $p \cdot v + q$ , one plugs the affine transformations into equations (8) and (9) and observes that  $p$  and  $q$  drop out.

To see that  $(\beta_\alpha^{\text{IND}}, \alpha(l_0; \mathbf{y}), \alpha(\mu_0; \mathbf{y}), \alpha(u_0; \mathbf{y}))$  is a moment equilibrium for distribution  $F_\alpha$ , observe that

$$\begin{aligned} \int_{\underline{v}}^{\bar{v}} \beta(\alpha(v; \mathbf{y}) | \alpha(l_0; \mathbf{y}), \alpha(\mu_0; \mathbf{y}), \alpha(u_0; \mathbf{y})) dF_\alpha(\alpha(v; \mathbf{y})) \\ = \int_{\underline{v}}^{\bar{v}} p(\mathbf{y}) \beta(v | l_0, \mu_0, u_0) + q(\mathbf{y}) dF(v) = \alpha(\mu_0; \mathbf{y}). \end{aligned}$$

An analogous proof works for aggregate moment beliefs.

## B Incentives in Procurement Auctions

In procurement auctions, the incentives are reversed compared to the usual bidder-as-buyer auction. As bidders are sellers in procurement auctions, the bidder with the

lowest bid wins. The private information is the privately known cost of producing the good. Bidders incur this cost conditional on winning the auction. In a bidder-as-seller auction, the expected utility of bidder  $i$  when bidding  $b_i$  is

$$\begin{aligned}\mathbb{P}(b_i < \min b_{-i})(b_i - c_i) &= (1 - \mathbb{P}(\min b_{-i} < b_i))(b_i - c_i) \\ &= (1 - (1 - (1 - B(b_i))^{n-1}))(b_i - c_i) \\ &= (1 - B(b_i))^{n-1}(b_i - c_i),\end{aligned}$$

where the bids of the other bidders are drawn from distribution  $B$ .

We first consider the incentives with aggregate moment beliefs. Let  $l$  be the lowest bid and  $u$  the highest bid of the other bidders. Let  $m$  denote the expected winning bid, i.e., the expected minimum of  $n$  independent bids. The worst-case bid distribution has the elements  $x_1$  and  $x_2$  in the support,  $l \leq x_1 \leq m \leq x_2 \leq u$ . The moment constraint implies  $q \cdot x_1 + (1 - q) \cdot x_2 = m$ , so  $q(x_1, x_2) = \frac{x_2 - m}{x_2 - x_1}$ . The lowest bid of  $n$  independent bids is  $x_1$  if not all bid  $x_2$ , i.e.,  $q = 1 - \mathbb{P}(b_j = x_2)^n = 1 - (1 - \mathbb{P}(b_j = x_1))^n$ . The probability that a competing bidder bids  $x_1$  is therefore  $\mathbb{P}(b_j = x_1) = 1 - (1 - q)^{\frac{1}{n}}$ . The probability that  $x_1$  is the lowest of  $n - 1$  independent draws is then  $1 - (1 - q)^{\frac{n-1}{n}}$ . The probability that a bid  $b \in (x_1, x_2)$  is winning is then  $(1 - q)^{\frac{n-1}{n}}$ .

The maximal loss conditional on bidding too high is  $\max\{(1 - q(l, b))^{\frac{n-1}{n}}(b - c), m - c\}$  if  $m < b$  and  $(b - c)(1 - (1 - q(b, u))^{\frac{n-1}{n}})$  if  $b \leq m$ . The maximal loss conditional on bidding too low is  $(1 - q(l, u))^{\frac{n-1}{n}}(u - b)$ . The minimax bid equates the maximal loss conditional on bidding too high and the maximal loss conditional on bidding too low. The inverse minimax bidding function  $c^{\text{AGG}}(b|l, m, u)$  is well defined in a separating aggregate moment equilibrium.

We now consider individual moment beliefs in a procurement context. If other bidders only use two bids,  $x_1$  and  $x_2$ , then the probability that bid  $b \in (x_1, x_2)$  wins if all  $n - 1$  other bidders bid  $x_2$ . In a procurement auction, the bid is too low if a higher bid would not decrease the probability of winning. The worst-case bid distribution is then  $\pi[l] + (1 - \pi)x_2^P$ , where  $x_2^P = \min\{u, \max\{m, \hat{x}_2^P\}\}$  and

$$\hat{x}_2^P = \begin{cases} \frac{(n-1)b-l}{n-2} & \text{if } n \geq 3 \\ u & \text{if } n = 2. \end{cases}$$

Maximal loss conditional on bidding too high equals  $\left(\frac{m-l}{x_2^P-l}\right)^{n-1}(x_2^P - b)$ . The worst case when bidding too high comes from bidding marginally too high. A lower bid would have won. If  $m < b$ , then the worst-case bid distribution is  $\pi[l] + (1 - \pi)[b - \epsilon]$  so

that maximal loss is  $\left(\frac{m-l}{b-l}\right)^{n-1} (b-c)$ . If  $b \leq m$ , then the worst-case bid distribution is  $\pi[b - \epsilon] + (1 - \pi)[u]$ . Maximal loss equals  $(b-c)(1 - \left(\frac{m-b}{u-b}\right)^{n-1})$ . In a separating equilibrium, the (inverse) minimax bidding function equalizes the maximal loss from bidding too high with the maximal loss from bidding too low and equals

$$c^{\text{IND}}(b|l, m, u) = \begin{cases} b - \left(\frac{b-l}{x_2^P - l}\right)^{n-1} (x_2^P - b) & \text{if } m < b \\ b - \left(\frac{(m-l)(u-b)}{x_2^P - l}\right)^{n-1} \frac{x_2^P - b}{(u-b)^{n-1} - (m-b)^{n-1}} & \text{if } b \leq m \end{cases} \quad (19)$$

for  $b \in [l, u]$ .

The upper bound belief  $u$  must equal the highest cost  $\bar{c}$  in a moment equilibrium.

## References

- Athey, S. and P. Haile (2007). *Nonparametric Approaches to Auctions*, Volume 6A of *Handbook of Econometrics*, Chapter 60, pp. 3847–3965. Elsevier.
- Bajari, P., S. Houghton, and S. Tadelis (2014, April). Bidding for incomplete contracts: An empirical analysis of adaptation costs. *American Economic Review* 104(4), 1288–1319.
- Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2015, February). Self-confirming equilibrium and model uncertainty. *American Economic Review* 105(2), 646–77.
- Battigalli, P., S. Cerreia-Vioglio, F. Maccheroni, and M. Marinacci (2016). Analysis of information feedback and selfconfirming equilibrium. *Journal of Mathematical Economics* 66, 40 – 51.
- Bergemann, D. and S. Morris (2005, November). Robust mechanism design. *Econometrica* 73(6), 1771–1813.
- Bergemann, D. and K. H. Schlag (2011). Robust monopoly pricing. *Journal of Economic Theory* 146(6), 2527–2543.
- Carrasco, V., V. F. Luz, N. Kos, M. Messner, P. Monteiro, and H. Moreira (2018). Optimal selling mechanisms under moment conditions. *Journal of Economic Theory* 177, 245 – 279.
- Carroll, G. (2015, February). Robustness and linear contracts. *American Economic Review* 105(2), 536–563.
- Carroll, G. and I. Segal (2019). Robustly optimal auctions with unknown resale opportunities. *Review of Economic Studies* 86(4), 1527–1555.



- Chen, Y., P. Katusčák, and E. Ozdenoren (2007). Sealed bid auctions with ambiguity: Theory and experiments. *Journal of Economic Theory* 136(1), 513 – 535.
- Donald, S. G. and H. J. Paarsch (2002). Superconsistent estimation and inference in structural econometric models using extreme order statistics. *Journal of Econometrics* 109, 305–340.
- Dow, J. and S. R. D. C. Werlang (1994). Nash equilibrium under knightian uncertainty: Breaking down backward induction. *Journal of Economic Theory* 64(2), 305 – 324.
- Eichberger, J. and D. Kelsey (2000). Non-additive beliefs and strategic equilibria. *Games and Economic Behavior* 30(2), 183 – 215.
- Elstrodt, J. (2018). *Maß- und Integrationstheorie* (8 ed.). Springer.
- Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics* 18(2), 141 – 153.
- Gretschko, V. and H. Mass (2018). Worst-case subjective-belief equilibria in first-price auctions. *Working paper*.
- Guerre, E., I. Perrigne, and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica* 68(3), 525–574.
- Guo, Y. and E. Shmaya (2019). Robust monopoly regulation. *Working paper*.
- Guo, Y. and E. Shmaya (2021). Project choice from a verifiable proposal. *Working paper*.
- Hong, H., P. Haile, and M. Shum (2003). Nonparametric tests for common values in first-price sealed-bid auctions. *Working paper*.
- Hortaçsu, A., F. Luco, S. L. Puller, and D. Zhu (2019, December). Does strategic ability affect efficiency? evidence from electricity markets. *American Economic Review* 109(12), 4302–42.
- Hyafil, N. and C. Boutilier (2004). Regret minimizing equilibria and mechanisms for games with strict type uncertainty. In *Proceedings of the Twentieth Annual Conference on Uncertainty in Artificial Intelligence*, pp. 268–277.
- Kasberger, B. and K. H. Schlag (2022). Robust bidding in first-price auctions: How to bid without knowing what others are doing. *Working paper*.
- Klibanoff, P. (1996). Uncertainty, decision, and normal form games. *MIT Working Paper*.
- Krasnokutskaya, E. (2011, 01). Identification and Estimation of Auction Models with Unobserved Heterogeneity. *The Review of Economic Studies* 78(1), 293–327.

- Lehrer, E. (2012, February). Partially specified probabilities: Decisions and games. *American Economic Journal: Microeconomics* 4(1), 70–100.
- Linhart, P. and R. Radner (1989). Minimax-regret strategies for bargaining over several variables. *Journal of Economic Theory* 48(1), 152 – 178.
- Lo, K. C. (1996). Equilibrium in beliefs under uncertainty. *Journal of Economic Theory* 71(2), 443 – 484.
- Lo, K. C. (1998). Sealed bid auctions with uncertainty averse bidders. *Economic Theory* 12(1), 1–20.
- Marinacci, M. (2000). Ambiguous games. *Games and Economic Behavior* 31(2), 191 – 219.
- Ollár, M. and A. Penta (2017, August). Full implementation and belief restrictions. *American Economic Review* 107(8), 2243–77.
- Osborne, M. J. and A. Rubinstein (2003). Sampling equilibrium, with an application to strategic voting. *Games and Economic Behavior* 45(2), 434 – 441. Special Issue in Honor of Robert W. Rosenthal.
- Perakis, G. and G. Roels (2008). Regret in the newsvendor model with partial information. *Operations Research* 56(1), 188–203.
- Popescu, I. (2005). A semidefinite programming approach to optimal-moment bounds for convex classes of distributions. *Mathematics of Operations Research* 30(3), 632–657.
- Renou, L. and K. H. Schlag (2010). Minimax regret and strategic uncertainty. *Journal of Economic Theory* 145(1), 264 – 286.
- Roughgarden, T. (2020). Transaction fee mechanism design for the ethereum blockchain: An economic analysis of eip-1559. *Working paper*.
- Salant, Y. and J. Cherry (2020). Statistical inference in games. *Econometrica* 88(4), 1725–1752.
- Savage, L. J. (1951). The theory of statistical decision. *Journal of the American Statistical Association* 46(253), 55–67.
- Savage, L. J. (1954). *The Foundations of Statistics*. Wiley.
- Schlag, K. H. and A. Zapechelnnyuk (2020). Compromise, don't optimize: A prior-free alternative to bayesian nash equilibrium. *Working paper*.
- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* 57(3), 571–587.
- Simon, C. P. and L. Blume (1994). *Mathematics for Economists*. W. W. Norton & Company.

- Smith, J. E. (1995). Generalized chebychev inequalities: Theory and applications in decision analysis. *Operations Research* 43(5), 807–825.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Addison-Wesley.
- Wilson, R. (1987). Game-theoretic analyses of trading processes. In T. F. Bewley (Ed.), *Advances in Economic Theory*, Fifth World Congress, pp. 33–70. Cambridge University Press.
- Winkler, G. (1988). Extreme points of moment sets. *Mathematics of Operations Research* 13(4), 581–587.