Robust Inference in First-Price Auctions: Overbidding as an Identifying Restriction

Serafin Grundl and Yu Zhu*

December 22, 2021

Abstract

Laboratory experiments find consistently that bidding in first-price auctions tends to be more aggressive than predicted by the risk-neutral Bayesian Nash Equilibrium (RNBNE) - a finding known as the overbidding puzzle. Several alternative models can explain the overbidding puzzle, but no canonical alternative to RNBNE has emerged. Instead of estimating a particular model of overbidding, we use the overbidding restriction itself for identification, which allows us to bound the valuation distribution and counterfactuals including the seller's revenue function in first- and second-price auctions in the spirit of Haile and Tamer (2003). These bounds are consistent with RNBNE and all models of overbidding, and the bounds remain valid even if there is unobserved heterogeneity in bidding strategies. We evaluate the validity of the bounds numerically and in experimental data.

JEL Codes: C57, D44, C14

Keywords: First-Price Auction, Robust Inference, Experimental Findings, Structural Estimation, Partial Identification

*We thank Douglas Dyer, John H. Kagel and Dan Levin for generously sharing their data with us. We thank Victor Aguirregabiria, Yonghong An, Emmanuel Guerre, Brent Hickman, Yingyao Hu, Jack Porter, Tong Li, Elie Tamer and seminar participants at Vanderbilt University, University of Toronto, IIOC (2017), Shanghai University of Finance and Economics, Penn State, Federal Reserve Board, and Tsinghua University for helpful comments. Serafin Grundl: Federal Reserve Board of Governors, serafin.j.grundl@frb.gov. Yu Zhu: Bank of Canada, yuzhu@bank-banque-canada.ca.

1 Introduction

Identification and estimation of first-price auctions is typically based on the assumption that bidders play the risk neutral Bayesian Nash Equilibrium (RNBNE) (Guerre, Perrigne, and Vuong (2000)). However, there are two broad reasons why one may want to relax the RNBNE assumption. First, a long series of laboratory experiments with independent private valuations have consistently found that bidders tend to bid more aggressively than predicted by the RNBNE - a finding known as the "overbidding puzzle". Second, in the theoretical literature several more general bidding models have been proposed that can explain the bidding puzzle, such as risk aversion, ambiguity aversion, loss aversion, level-k, quantal response equilibrium or models of regret.

One approach to relax the assumption of RNBNE bidding in structural estimation is therefore to study the identification of alternative bidding models. Indeed, under additional restrictions, identification results have been established for models with risk aversion (Perrigne and Vuong (2007), Lu and Perrigne (2008), Guerre, Perrigne, and Vuong (2009), Campo, Guerre, Perrigne, and Vuong (2011), Campo (2012), Gentry, Li, and Lu (2015), Kong (2017)), ambiguity aversion (Aryal, Grundl, Kim, and Zhu (2018)), and the level-k model (An (2017), Gillen (2009)).

It is however unclear which alternative bidding model should be used in structural estimation for two reasons. First, it is very difficult, if not impossible, to distinguish among the alternative models using bid data. This paper provides two theoretical results on the non-testability among the models. Moreover, even in the experimental literature, where valuations are observed in addition to bids so the models can be distinguished more easily, no canonical alternative to RNBNE has emerged. Second, these alternative models are not mutually exclusive. For example, bidders can have risk aversion, ambiguity aversion and/or loss aversion all at the same time. This further complicates the problem of searching for the "right model".

In light of this difficulty, this paper proposes a more robust approach. Instead of assuming a particular model of bidding we make two general assumptions. First, we assume that bidders tend to overbid compared to RNBNE in the sense that their bid distribution first-order stochastically dominates the bid distribution implied by RNBNE bidding. Second, we assume that the valuation distribution first-order stochastically dominates the bid distribution. These two assumptions are consistent with the overbidding puzzle in the experimental literature. An important class of alternative bidding models in the theoretical literature predicts overbidding by all bidders. These models include the risk aversion model, the Choquet Expected Utility model of ambiguity aversion, the loser regret model and the loss aversion model. The assumptions are also consistent with models which may predict overbidding by some bidders and underbidding by others, such as the level-k model, if these models are parameterized such that there is an overall tendency towards overbidding.[1]

These two assumptions allow us to partially identify the valuation distribution. Although these bounds can be wide, they can be tightened considerably under the assumption that the density of valuations is unimodal. They can be tightened further by exploiting variation in the number of bidders under assumptions that are common in the structural auction literature. After the tightening, we can obtain very informative bounds.

The identification results are not only robust to overbidding, but perhaps equally important, also allow for unobserved heterogeneity in bidding strategies across bidders, as long as the identifying restrictions are satisfied. Such heterogeneity can reflect unobserved heterogeneity in preferences and/or beliefs, which can be relevant both in the field and in lab experiments. For instance, Cox, Smith, and Walker (1988) proposed a model with heterogeneous risk aversion as the explanation of the overbidding puzzle. Robustness to unobservable bidder heterogeneity is an important advantage compared to most point-identification results in the literature. Notice however that we cannot allow for bidder heterogeneity that violates the identifying assumptions. For instance, if most bidders underbid compared to RNBNE such that the bid distribution does not first-order stochastically dominate the valuation distribution, the bounds would no longer be valid.

One challenge for the partial identification approach is to construct counterfactuals. We focus on the revenue in a first-price auction with a binding reserve price and show how to obtain sharp bounds under an additional assumption. This in turn allows us to bound the optimal reserve price. Because the sharp bounds can be computationally difficult to obtain, we then propose slightly more conservative bounds which are much easier to compute. We evaluate the performance of the bounds in numerical examples and in experimental data from Dyer, Kagel, and Levin (1989). We also show how to obtain sharp bounds on the seller's revenue in a second-price auction with a reserve price.

This paper is motivated by a large experimental literature on the overbidding puzzle. The overbidding puzzle goes back at least to Coppinger, Smith, and Titus (1980), followed by a series of papers by Cox et al. (Cox, Roberson, and Smith (1982), Cox, Smith, and Walker (1983a,b, 1985, 1988)), who proposed a model of heterogeneous bidders with constant relative risk aversion (CRRA) as the explanation of the puzzle. Harrison (1989) and Harrison (1990) criticized the risk aversion explanation, which sparked a controversial discussion about the correct interpretation of the overbidding puzzle in the December 1992 issue of the American Economic Review (Friedman (1992), Kagel and Roth (1992), Cox, Smith, and Walker (1992),

[1]We discuss alternative bidding models in Appendix C.

Merlo and Schotter (1992) and Harrison (1992)). More recently, Kirchkamp and Reiß (2011) found that deviations from RNBNE occur because bidders deviate from the risk-neutral best response, not because they have incorrect beliefs. This finding suggests that overbidding may be a persistent phenomenon that would not disappear if bidders gradually update their expectations.[2]

It is sometimes argued that while overbidding may be common in the laboratory, it is not relevant for applications to field data, because there the bidders are more experienced. There is however evidence that is consistent with overbidding in the field.[3] Moreover, in some important applications that have been studied extensively in the structural auction literature many of the bidders have little experience in first-price auctions.[4] Even if all bidders are experienced it is not clear whether overbidding would become irrelevant. We find no evidence that experience reduces overbidding in the experimental data by Dyer, Kagel, and Levin (1989).[5] Indeed, if overbidding is due to preferences (e.g. risk aversion), it should not disappear with more experience.[6] Overall, we argue that overbidding in the

[2]An incomplete list of further references includes Kagel and Levin (1993), Selten, Buchta, et al. (1994), Chen and Plott (1998), Isaac and James (2000), Dufwenberg and Gneezy (2002), Goeree, Holt, and Palfrey (2002), Dorsey and Razzolini (2003), Ockenfels and Selten (2005), Neugebauer and Selten (2006), Engelbrecht-Wiggans and Katok (2007), Neugebauer and Perote (2008), Engelbrecht-Wiggans and Katok (2009), Shachat and Wei (2012), Astor, Adam, Jähnig, and Seifert (2013), Georganas, Levin, and McGee (2015), Ratan (2015) and Füllbrunn, Janssen, and Weitzel (2018). See also the surveys of Kagel (1995) and Kagel and Levin (2010), and the references therein.

[3]In field data evidence for or against overbidding is necessarily less direct than in data from laboratory experiments, because valuations are not observed. There is however evidence from field data that is consistent with overbidding. For example, Lu and Perrigne (2008), Campo, Guerre, Perrigne, and Vuong (2011), Campo (2012) and Kong (2017) find evidence of risk aversion, which leads to overbidding.Aryal, Grundl, Kim, and Zhu (2018) find evidence of ambiguity aversion, which in this case also leads to overbidding. Athey and Levin (2001) also find evidence that is consistent with bidder risk aversion in a study that focuses on skewed bidding. Lastly, Burkart (1995) and de Bodt, Cousin, and Roll (2018) find overbidding in takeover contests.

[4]This can be the case either because the auctions are held infrequently, because the auction format changes over time, or because many bidders are small firms that participate only infrequently in auctions. For example in USFS timber auctions many bidders are small logging firms that bid infrequently. Moreover, many of the timber sales are not first-price auctions, but English auctions or so-called "direct sales", so bidders do not accumulate experience in first price auctions in all sales.

[5]Bidders in their experiments participated in up to 20 rounds, which allowed them to gain substantially more experience than the median experience of bidders in the US Forest Service data. In this experimental data valuations are observable so we can regress the difference between bids and the RNBNE prediction on experience. In simple regressions that include bidder and experiment fixed effects, we find no evidence that experience lowers overbidding. Experience is associated with slightly more overbidding in both three and six bidder auctions. The experience effect in the regressions is statistically significant at the 5% level for six bidder auctions, but not for three bidder auctions. Regression tables are available upon request. Figure 1 in Dyer, Kagel, and Levin (1989) also shows how overbidding in all three experiments varied across rounds of bidding, and there is no apparent effect of experience.

[6]Thus our experience regressions may suggest a preference-based explanation for the overbidding puzzle, rather than a belief based explanation. Following this interpretation, the identification results in this paper can be viewed as a way to allow for risk averse bidders that have heterogenous utility functions without assuming exogenous variation in the number of bidders, which is allows for point identification of the model

field cannot be ruled out by the available evidence, and therefore it is desirable to develop inference that is robust to overbidding.

This paper is close in spirit to Haile and Tamer (2003), who show how to bound the valuation distribution in English auctions if the assumptions of the point-identified button model are replaced with two weak behavioral restrictions. The papers are similar in that they do not rely on a particular model of bidding and the estimates are therefore robust to model misspecification, but they differ regarding the source of identifying restrictions. Haile and Tamer (2003) relax the assumptions of the canonical model of English auctions in a very intuitive way and still obtain informative bounds. It is more difficult to follow this approach in first-price auctions. For example, Aradillas-Lopez and Tamer (2008) find that a similar approach in first-price auctions (level-k rationalizability) bounds the valuation distribution only from above.[7] In this paper, we therefore rely on restrictions from experiments to bound the valuation distribution from both sides.

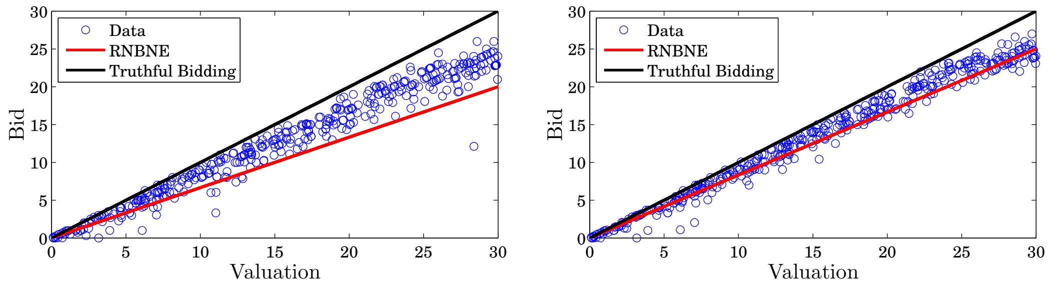
The remainder of this paper is organized as follows. Section 2 documents the extent of overbidding in the experimental data by Dyer, Kagel, and Levin (1989) to motivate our identifying assumptions. Section 3 presents the non-testability results and the baseline identification results for the valuation distribution, and also shows how to tighten the bounds by assuming a unimodal valuation density and/or exploiting variation in the number of bidders. Section 4 shows how to obtain bounds for counterfactuals in first-price and second-price auctions. Section 5 illustrates the identification results through numerical examples. We discuss estimation in Section 6. Section 7 applies the estimators to experimental data, and Section 8 concludes. Some proofs and details on computation are presented in the Appendix.

2 Overbidding in Lab Experiments

Before presenting the identification results we briefly motivate our identifying restriction using the experimental bid data from Dyer, Kagel, and Levin (1989), which was also studied by Bajari and Hortacsu (2005). The subjects in their experiments were MBA students at the University of Houston. As in other experimental work on auctions, the bidders were assigned valuations randomly and then asked to bid given the valuations they drew. The valuations were drawn from a uniform distribution from \$0 to \$30. The payoff for the winning bidder is the difference between bid and valuation. The data set contains bids for auctions with $n = 3$ and $n = 6$ bidders. After discarding bids from training rounds there are 414 bids for $n = 3$ and 414 bids for $n = 6$.

with homogenous risk averse bidders (Guerre, Perrigne, and Vuong (2009)).

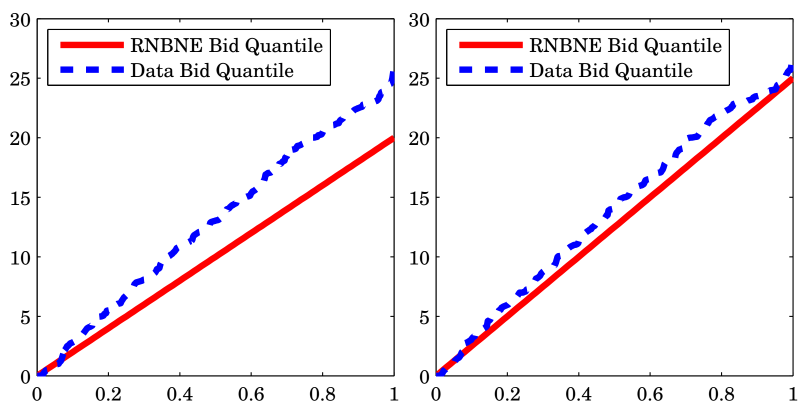
[7] See Kosenkova (2017) for identification under level-k rationalizability with asymmetric bidders.



(a) $n = 3$ (b) $n = 6$

Figure 1: **Overbidding Compared to RNBNE:** Experimental bid data from Dyer, Kagel, and Levin (1989). Our first identifying restriction is that the bids (blue circles) are between the RNBNE (red line) and their valuations (black line).

The RNBNE bidding strategy is $\sigma_n(t) = \frac{n-1}{n}t$, where t is the valuation of a bidder for the auctioned good. Figure 1 shows scatter plots of the bids (blue circles), the RNBNE strategy (red line), and "truthful bidding", i.e., bidding one's valuation (black line). Figure 1(a) shows bids for $n = 3$ and Figure 1(b) shows bids for $n = 6$. For $n = 3$, 91.3 percent of the bidders overbid compared to the RNBNE. For $n = 6$, 74.4 percent overbid. Overbidding tends to be more prevalent in auctions with fewer opponents. This is perhaps because in auctions with many opponents, the RNBNE prediction is already close to bidding one's valuation, so there is less scope for overbidding while still bidding less than one's valuation.



(a) $n = 3$ (b) $n = 6$

Figure 2: Bid Quantile Functions

While most bidders overbid, there is a nontrivial extent of underbidding as well. For $n = 3$ 8.7% of the bidders underbid, and for $n = 6$ 25.6% of bidders underbid. For $n = 3$ underbidding occurs almost exclusively for bidders with low valuations. This pattern is common in the experimental literature, as described by Kagel and Levin (2010). For $n = 6$

underbidding is almost equally prevalent across bidders with different valuations. However, the extent of underbidding is considerably larger for bidders with low valuations than those with high valuations.[8]

Despite of the nontrivial underbidding, there is still a general tendency towards overbidding in the sense that the actual bid distributions first-order stochastically dominate the RNBNE distribution, or equivalently, actual bid quantile functions lie above the RNBNE bid quantile functions, as shown in Figure 2. There are slight violations of first-order stochastic dominance for a small set of points with $\alpha < 0.1$, but Kolmogorov-Smirnov tests do not reject that the observed bid distributions first-order stochastically dominate the RNBNE distributions at any commonly used significance level. Tests do reject however that the bid distributions are generated by the RNBNE. This motivates our identifying restriction in the next section.

Notice also that no bidder bids above his valuation. We will use this restriction to bound the valuation distribution from the other side. Notice also that the bidding strategies are very heterogeneous, ranging from somewhere below the RNBNE strategy to the truthful bidding strategy. Therefore, robustness to heterogeneous bidding strategies is desirable.

3 Identification of Valuation Distribution

We consider an independent private-value environment. There are $n \geq 2$ bidders participating in a first-price auction for an indivisible good. Each of them simultaneously submits a sealed bid. The bidder with the highest bid wins the good and pays the bid. Bidders' valuations for the good are independently drawn from a distribution F_n with a density f_n. The bid distribution is denoted by G_n with a density g_n. We assume that both f_n and g_n are supported on bounded intervals, and bounded away from 0 on their supports. Without loss of generality, we assume that the support of f_n has a lower bound of 0. We allow bidders to use different bidding strategies although their valuations are drawn from the same distribution. Bidder i with valuation t bids $s_n^i(t)$ and the bidders' identities are unknown to the econometrician. This can occur if bidders have heterogenous risk aversion and/or heterogeneous beliefs that are unobservable to econometricians.[9] The available information is G_n

[8] For $n = 6$ underbidding bidders with valuations above 15 underbid on average by 2.9% (median 2.0%), whereas bidders with valuations below 15 underbid on average by 21.3% (median 8.2%).

[9] One common example is that bidders have heterogeneous risk aversion. There can be three scenarios. First, bidders are unaware of this heterogeneity and assume other bidders have the same risk aversion as them. Second, the level of risk aversion is private information and the distribution of risk aversion is common knowledge. Third, risk aversion of every bidder is common knowledge. In the first scenario, bidders adopt symmetric bidding strategies but with different risk aversion levels and RNBNE overbidding is satisfied. In the second and third scenarios, there is no general proof that equilibrium bidding strategies are higher

and the number of bidders. The goal is to identify F_n using this information. Before moving to the identification results, we first present two non-testability results, which highlight the difficulties in distinguishing among the alternative overbidding models.

Theorem 1. *A bid distribution G_n with a continuous and positive density g_n supported on a bounded interval can be rationalized by an Ambiguity Aversion model (Maxmin Expected Utility or Choquet Expected Utility), the Risk Aversion model, the Level-k model, the Loser Regret model, and the Loss Aversion model, if $b + G_n(b)/g_n(b)$ is strictly increasing.*

Theorem 1 shows that if bids are from auctions with the same n, it is not possible to figure out which model generates the data. This result holds even if bidders use the same bidding strategy.

In some situations, the econometrician may observe auctions with different n, which may be exploited to distinguish among models. If auctions with different n are generated by different primitives, then Theorem 1 still applies. To exploit variation in n, it is necessary to impose assumptions on how bidders select into auctions. One common assumption used in the literature is that the valuation distribution is independent of n. Even under this assumption, the following result shows that we cannot distinguish among the models.

Theorem 2. *It remains impossible to distinguish all models in Theorem 1 even if auctions with different n are observed and the valuation distribution is independent of n.*

Theorem 2 holds because other primitives such as preferences or beliefs are allowed to vary with n. It does not rule out the possibility of distinguishing the models if we further strengthen the assumption to require that all the primitives are independent of n. Pursuing this is beyond the scope of this paper and in the rest of the paper, we do not impose assumptions beyond those in Theorem 2 to exploit variation in n.

Theorem 1 and 2 are concerned only with models that have a single overbidding motive. They are not concerned with, for example, distinguishing an ambiguity aversion model from a model with both risk aversion and loss aversion. The latter is an even more daunting task because there can be many combinations of overbidding motives. Our partial identification approach circumvents these difficulties.

3.1 Bounds for the Valuation Quantile Function

Structural analysis of first-price auctions typically assumes that the bidders play the RNBNE. Under this assumption a bidder with valuation t bids $\sigma_n(t)$, which satisfies the first-order

than the RNBNE strategy. However, there are papers showing this holds in certain cases. For example, Cox, Smith, and Walker (1988) show that under the second scenario, RNBNE overbidding is satisfied if the valuation distribution is uniform.

condition,

$$\sigma_n'(t) = (n-1)\left[t - \sigma_n(t)\right]\frac{f_n(t)}{F_n(t)}, \tag{1}$$

and the boundary condition, $\sigma_n(0) = 0$. This differential equation has the following closed form solution:

$$\sigma_n(t) = t - \int_0^t \left[\frac{F_n(x)}{F_n(t)}\right]^{n-1} dx.$$

Notice that because bidders have a homogeneous valuation distribution, the symmetric RNBNE is the only RNBNE (Maskin and Riley (2003)). Instead of assuming that bidders follow the strategy σ_n, we rely on weaker assumptions motivated by the overbidding puzzle. To proceed, one can assume that every bidder bids more than the RNBNE strategy and less than his valuation, i.e., $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t. However, these conditions may be too strong in light of the experimental findings shown in the last section. Therefore, we work with the following two conditions.

Definition 1 (RNBNE Overbidding in Distribution)**.** Bidders overbid compared to RNBNE in distribution if $G_n(b) \leq F_n\left[\sigma_n^{-1}(b)\right]$ for all b.

Definition 2 (Rationality in Distribution)**.** Bidders are rational in distribution if $G_n(b) \geq F_n(b)$ for all b.

RNBNE overbidding in distribution means that the true bid distribution weakly first-order stochastically dominates the RNBNE bid distribution. Rationality in distribution means that the valuation distribution weakly first-order stochastically dominates the bid distribution.

One can show that if $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t, RNBNE overbidding and rationality in distribution follow. However, these distributional definitions are weaker than requiring $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t. In particular, RNBNE overbidding in distribution can hold even if some bidders underbid. In this case, the true bid distribution is a mixture of the bid distribution from bidders who overbid and that from bidders who do not. The former distribution first-order stochastically dominates the RNBNE bid distribution while the latter does not. If the effect of the former distribution dominates, the resulting mixture can still first-order stochastically dominate the RNBNE bid distribution.[10] This is important because, as discussed in Section 2, the experimental data show that some bidders underbid

[10] For instance, if valuations are uniformly distributed on $[0,1]$ and there is a fraction p of bidders who overbid by o percent, and a fraction $1-p$ of bidders who underbid by u percent, then the bounds remain valid if $\frac{p}{1+o} + \frac{1-p}{1-u} \leq 1$. If we set $u = o = x$ then this can be simplified to $\frac{1+x}{2} \leq p$. For example, if the bidders over- or underbid by ten percent ($x = 0.1$) then $p \geq 0.55$ so the fraction of underbidders must be below 45%.

compared to RNBNE. Moreover, some bidding theories, such as level-k bidding can predict underbidding (Gillen (2009)).

Although RNBNE overbidding and rationality in distribution are weaker than requiring $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t, they are observationally equivalent if the only available information is the bid distribution. In other words, a valuation distribution can rationalize a bid distribution under RNBNE overbidding and rationality in distribution if and only if it can rationalize the bid distribution when $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t. We omit the term "in distribution" when referring to Definitions 1 and 2 in the rest of the paper. The following lemma summarizes the preceding discussion.

Lemma 1. *If $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t, RNBNE overbidding and Rationality in distribution follow. Moreover, G_n can be rationalized by a valuation distribution F_n under $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t if and only if it can be rationalized by F_n under RNBNE overbidding and rationality in distribution.*

Proof. We start with the first claim. We focus on the case with a finite number of bidder types. A similar proof applies to the case with an infinite number of types. Let δ_i be the fraction of type i bidders in the data. Then $G_n(b) = \sum_i \delta_i P(s_n^i(t) \leq b) \geq \sum_i \delta_i P(t \leq b) = F_n(b)$. The inequality holds because $t \geq s_n^i(t)$ by assumption. Similarly, $G_n(b) \leq \sum_i \delta_i P(\sigma_n(t) \leq b) = F_n(\sigma_n^{-1}(b))$. This proves the first part of the lemma. For the second part, we only need to show that if F_n rationalizes G_n under RNBNE overbidding and rationality in distribution, it also rationalizes G_n under the assumption that $t \geq s_n^i(t) \geq \sigma_n(t)$ for all i and t. To see this, one can define $s_n^i(t) = s_n(t) = G_n^{-1}(F_n(t))$. One can check that $t \geq s_n(t) \geq \sigma_n(t)$ under RNBNE overbidding and rationality in distribution. This concludes the proof. □

We now show how to bound the valuation distribution under RNBNE overbidding and rationality. It turns out that it is more convenient to work with the quantile functions than the distribution functions. Define $v_n(\alpha) = F_n^{-1}(\alpha)$ for $\alpha \in [0,1]$ to be the valuation quantile function. Because f_n has a support with a lower bound of 0, $v_n(0) = 0$. Let $b_n(\alpha) = G_n^{-1}(\alpha)$ for $\alpha \in [0,1]$ be the bid quantile function. Bounding v_n is equivalent to bounding F_n. Hence, we focus on v_n.

First notice that rationality directly implies $v_n(\cdot) \geq b_n(\cdot)$. Therefore, $\underline{v}_n(\cdot) = b_n(\cdot)$ is a valid lower bound for $v_n(\cdot)$. It is also sharp because it can be attained by a model with constant relative risk aversion (CRRA) as the CRRA coefficient approaches 1.

Next, we construct the sharp upper bound for $v_n(\cdot)$. The idea is that if we can obtain the set of all valuation quantile functions that are consistent with b_n and satisfy RNBNE overbidding and rationality, then the upper contour of this set is the sharp pointwise upper bound for $v_n(\cdot)$. If a quantile function $q(\cdot)$ is consistent with $b_n(\cdot)$ under rationality, then $q(\alpha) \geq b_n(\alpha)$ for all $\alpha \in [0,1]$. To check RNBNE overbidding, the following lemma will be useful.

Lemma 2. *A valuation quantile function $q(\cdot)$ is consistent with $b_n(\cdot)$ and RNBNE overbidding if and only if* $\beta_n(\alpha, q) \leq b_n(\alpha)$ for all $\alpha \in [0,1]$, where

$$\beta_n(\alpha, q) = \begin{cases} \frac{1}{\alpha^{n-1}} \int_0^\alpha q(t)\, dt^{n-1} & \text{if } \alpha > 0; \\ q(0) & \text{if } \alpha = 0. \end{cases} \tag{2}$$

Proof. By definition, RNBNE overbidding holds if and only if the observed bid distribution first-order stochastically dominates the RNBNE bid distribution under q. This is equivalent to that the RNBNE bid quantile function under q lies below the observed bid quantile function b_n. This lemma follows because the RNBNE bid quantile function under q is $\beta_n(\alpha, q)$ by Gimenes and Guerre (2016). □

In light of Lemma 2, an upper bound for $v_n(\alpha)$ can be defined as

$$\bar{v}_n(\alpha) = \sup_{q \in \Theta} q(\alpha) \text{ s.t. } \beta_n(t, q) \leq b_n(t) \leq q(t), \forall t \in [0,1], \tag{3}$$

where Θ is the set of non-negative, non-decreasing and continuous functions that are supported on $[0,1]$ and differentiable almost everywhere. The inequalities impose RNBNE overbidding and rationality. The supremum is well-defined because at least $q(\cdot) = b_n(\cdot)$ satisfies all the constraints. By construction, $\bar{v}_n(\alpha)$ is the sharp upper bound. The following result characterizes the properties of the identified set for $v_n(\alpha)$.

Proposition 1. *If the bidders are rational and overbid compared to RNBNE, then the sharp identified set for $v_n(\alpha)$ is $[\underline{v}_n(\alpha), \bar{v}_n(\alpha)]$, where $\bar{v}_n(\alpha)$ is defined by* (3) and $\underline{v}_n(\cdot) = b_n(\cdot)$. Moreover, $\bar{v}_n(\cdot)$ *satisfies*

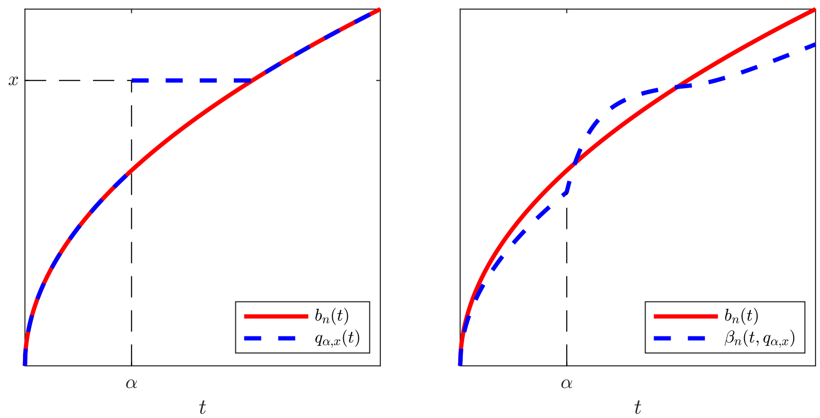
(1) $\bar{v}_n(\alpha) < b_n(1) / (1-\alpha)$ for any $\alpha < 1$;

(2) $\exists C > 0$ such that $\bar{v}_n(\alpha) > C / (1-\alpha)$ for α close to 1;

(3) $\bar{v}_n$ is strictly increasing and continuous.

Proposition 1 shows that $v_n(\alpha)$ is interval identified. The most surprising part of this proposition is perhaps that $\bar{v}_n(\alpha)$ is bounded for any $\alpha < 1$ but it diverges to infinity as $\alpha \to 1$. The easiest way to understand this is to start with a quantile function q and investigate what happens to its RNBNE bid quantile if we increase $q(\alpha)$ for some $\alpha < 1$.

First notice that this increase does not change the RNBNE bidding behavior of bidders with valuations smaller than $q(\alpha)$. Therefore, only the bid quantile function evaluated at points larger than α can provide information to bound $q(\alpha)$ from above. If $q(\alpha)$ is increased then the bidder with valuation $q(\alpha)$ increases his bid to increase his winning probability. In the RNBNE, a bidder with a valuation higher than $q(\alpha)$ never bids lower than a bidder with valuation $q(\alpha)$. Therefore bidders with valuations higher than $q(\alpha)$ have to increase their bids as well. These upward shifts of the RNBNE bid function could drive the RNBNE bid quantile function above the observed bid quantile function on $[\alpha, 1)$. This bounds $v_n(\alpha)$ from above if $\alpha < 1$. However, a bidder with the highest valuation $v_n(1)$ does not have an incentive to bid higher even if his valuation increases, because he wins with probability 1 even with the current bid. Therefore, $v_n(1)$ is not bounded from above.



(a) $q_{\alpha,x}(t)$ (b) Bid Quantiles

Figure 3: Alternative Formulation

Solving (3) appears to be a daunting task. In the online appendix, we show how it can be approximated by a linear programming problem, exploiting the fact that both the objective function and the constraints are linear in q. Here, we show that the problem can be recast as a one-dimensional maximization problem, which simplifies computation and also provides further intuition for $\bar{v}_n$. If we ignore the continuity requirement, the lowest valuation quantile that satisfies rationality and takes value x at a some $\alpha \in [0, 1]$ is

$$q_{\alpha,x}(t) = \begin{cases} b_n(t) & \text{if } t < \alpha \\ \max\{x, b_n(t)\} & \text{if } t \geq \alpha \end{cases}. \tag{4}$$

Figure 3(a) shows that $q_{\alpha,x}(t)$ (dashed blue line) coincides with $b_n(t)$ (red solid line) for $t < \alpha$. At $t = \alpha$, $q_{\alpha,t}(t)$ jumps to x and stays constant until $b_n(t)$ reaches x. After that, $q_{\alpha,x}(t)$ is identical to $b_n(t)$. If x is too big, the resulting RNBNE bid quantile function,

$\beta_n(t, q_{\alpha,x})$, lies above $b_n(t)$ for some $t > \alpha$ and violates RNBNE overbidding, as shown in Figure 3(b). It is not difficult to see that $\beta_n(\cdot, q)$ is weakly increasing in q, i.e., if $q_1(t) \geq q_2(t)$ for every t, then $\beta_n(t, q_1) \geq \beta_n(t, q_2)$ for every t.[11] Therefore, if $\beta_n(\cdot, q_{x,\alpha})$ does not lie below $b_n(\cdot)$, there cannot exist a q with $q(\alpha) = x$ that satisfies rationality and RNBNE overbidding. Hence, if there were no continuity requirement on q, $\bar{v}_n(\alpha)$ can be obtained by finding the maximum x such that $\beta_n(\cdot, q_{x,\alpha})$ does not exceed $b_n(\cdot)$. Because any function with a jump can be approximated by a continuous function, we arrive at the following equivalent formulation.

Proposition 2. *The sharp upper bound for the valuation quantile can be expressed as follows:*

$$\bar{v}_n(\alpha) = \bar{v}_n^a(\alpha) = \max_{x \geq b_n(\alpha)} x, \quad s.t.\ \beta_n(t, q_{\alpha,x}) \leq b_n(t), \quad \forall t \in [\alpha, 1]. \tag{5}$$

Notice that only $t \in [\alpha, 1]$ enter the expression for $\bar{v}_n^a(\alpha)$, which confirms, as discussed earlier, that only points larger than α can provide information to bound $v_n(\alpha)$ from above.

3.2 Tightening Bounds under Unimodality

Proposition 2 shows that the upper bound of $v_n(\alpha)$ is achieved with a quantile function that jumps up at α and remains constant for quantiles that are slightly larger than α. This implies a very extreme density: it takes value 0 in the interior of its support and then jumps to infinity. It is then natural to consider imposing shape restrictions to rule out such extreme distributions, which may tighten the upper bound.

We consider imposing that the valuation density is unimodal. It is an interesting shape restriction for the following reasons. First, it is natural as most commonly used distributions have this property. Second, it leads to linear constraints on the quantile function, which are easy to incorporate into (3). Third, it rules out the valuation quantile functions defined by (4), which have at least two modes. Therefore, it may significantly tighten the upper bound.

Assumption 1 (Unimodality)**.** *There exists an $\alpha^* \in [0, 1]$ such that $f_n(v_n(\alpha))$ is increasing if $\alpha < \alpha^*$ and is decreasing if $\alpha > \alpha^*$.*

Because $f_n(v_n(\alpha)) = \frac{1}{v_n'(\alpha)}$, unimodality implies that $v_n'(\alpha)$ is decreasing if $\alpha < \alpha^*$ and

[11] See Lemma 7 in the appendix.

increasing if $\alpha > \alpha^*$. If we knew the value of α^*, we could obtain the upper bound by solving

$$\begin{aligned} \bar{q}_n^u(\alpha, \alpha^*) = \sup_{q \in \Theta} q(\alpha) \ , \\ \text{s.t. } \beta_n(t, q) \le b_n(t) \le q(t), \forall t \in [0,1] \\ q'(\alpha_1) \ge q'(\alpha_2) \ \text{ if } \alpha_1 > \alpha_2 > \alpha^* \text{ or } \alpha_1 < \alpha_2 < \alpha^*. \end{aligned} \tag{6}$$

However, since α^* is unknown, we need to take the supremum over α^* and define

$$\bar{v}_n(\alpha) = \sup_{\alpha^* \in [0,1]} \bar{q}_n^u(\alpha, \alpha^*). \tag{7}$$

Similarly, we can obtain the lower bound by

$$\underline{v}_n(\alpha) = \inf_{\alpha^* \in [0,1]} \underline{q}_n^u(\alpha, \alpha^*), \tag{8}$$

where

$$\begin{aligned} \underline{q}_n^u(\alpha, \alpha^*) = \inf_{q \in \Theta} q(\alpha) \ , \\ \text{s.t. } \beta_n(t, q) \le b_n(t) \le q(t), \forall t \in [0,1] \\ q'(\alpha_1) \ge q'(\alpha_2) \ \text{ if } \alpha_1 > \alpha_2 > \alpha^* \text{ or } \alpha_1 < \alpha_2 < \alpha^*. \end{aligned} \tag{9}$$

Unlike in the case without the shape restriction, $\underline{v}_n(\cdot)$ may not be the same as $b_n(\cdot)$ because the latter may have a density that is not unimodal. Obviously, if b_n has a unimodal density function, then $\underline{v}_n(\alpha) = b_n(\alpha)$.

Proposition 3. *Suppose that Assumption 1 holds. Then under rationality and RNBNE overbidding, the sharp identified set for $v_n(\alpha)$ is $[\underline{v}_n(\alpha), \bar{v}_n(\alpha)]$, where $\bar{v}_n(\alpha)$ and $\underline{v}_n(\alpha)$ are defined by (7) and (8), respectively. Moreover, if b_n has a unimodal density, $\underline{v}_n(\alpha) = b_n(\alpha)$.*

Proof. The bounds are valid and sharp by construction. We only need to show that $v_n(\alpha)$ is interval identified. To achieve this, we need to show that for any $x \in (\underline{v}(\alpha), \bar{v}_n(\alpha))$, there exists a q with $q(\alpha) = x$ that satisfies Assumption 1 and rationalizes b_n under rationality and overbidding. Notice that the constraints in the problems for $\bar{q}_n^u(\alpha, \alpha^*)$ and $\underline{q}_n^u(\alpha, \alpha^*)$ define a convex set of q. Therefore, if $x \in \left(\underline{q}_n^u(\alpha, \alpha^*), \bar{q}_n^u(\alpha, \alpha^*)\right)$, these exists q that satisfies all the constraints and $q(\alpha) = x$. Because

$$[\underline{v}_n(\alpha), \bar{v}_n(\alpha)] = \cup_{\alpha^* \in [0,1]} \left[\underline{q}_n^u(\alpha, \alpha^*), \bar{q}_n^u(\alpha, \alpha^*)\right],$$

any $x \in (\underline{v}(\alpha), \bar{v}_n(\alpha))$ lies in $\left(\underline{q}_n^u(\alpha, \alpha^*), \bar{q}_n^u(\alpha, \alpha^*)\right)$ for some α^*. Hence, there exists a q with $q(\alpha) = x$ that satisfies Assumption 1 and rationalizes b_n under rationality and overbidding.

□

3.3 Tightening Bounds with Variation in n

In many applications, auctions with different number of bidders are observed. This variation can be used to tighten the bounds. To exploit this variation, we must restrict how the valuation distribution varies with n. To this end, we consider the following two assumptions:

Assumption 2. $v_n(\cdot) = v(\cdot)$ *for all* $n \in \mathbb{N}$*, where* $\mathbb{N}$ *is a finite set of integers.*

Assumption 3. $v_n(\alpha)$ *is weakly increasing in* n *for all* $\alpha \in [0, 1]$*.*

Assumption 2 requires that the valuation distribution is independent of n. This assumption is sometimes called exogenous participation (EP) in the literature. It holds trivially if bidders are randomly assigned to auctions. However, even if the bidders decide which auctions to enter, it holds in many cases.[12] EP is used in many papers to either improve inference and/or to achieve identification.[13] Testing EP empirically in field data is usually difficult due to the presence of unobserved auction level heterogeneity.

Assumption 3 is usually referred to as stochastically increasing valuations (SIV). It allows for some positive selection and thereby relaxes Assumption 2. This assumption is also considered in Aradillas-López, Gandhi, and Quint (2011) and Grundl and Zhu (2016). Notice that we do not restrict how bidding behavior changes with n. This is important for robustness because different overbidding models have different predictions on how bidding changes with n.

We now develop a unified formulation of the bounds that is valid under various assumptions on how v_n change with n. Let $\mathbb{N}_1(n)$ be the set of integers such that if $m \in \mathbb{N}_1(n)$, then $v_m(\alpha) \leq v_n(\alpha)$ for all $\alpha \in [0, 1]$. Similarly, let $\mathbb{N}_2(n)$ be the set of integers such that if $m \in \mathbb{N}_2(n)$, then $v_m(\alpha) \geq v_n(\alpha)$ for all $\alpha \in [0, 1]$. If Assumption 2 is imposed, $\mathbb{N}_1(n) = \mathbb{N}_2(n) = \mathbb{N}$. If Assumption 3 is imposed, $\mathbb{N}_1(n) = \{m \in \mathbb{N} : m \leq n\}$ and $\mathbb{N}_2(n) = \{m \in \mathbb{N} : m \geq n\}$. If variation in n is not exploited, $\mathbb{N}_1(n) = \mathbb{N}_2(n) = \{n\}$. The

[12] Grundl and Zhu (2016) show that this assumption holds very generally as long as potential bidders observe independent signals before making their entry decisions. If there is auction heterogeneity, we need to condition on all auction level heterogeneity (observed and unobserved) that is observed by the bidders, including the number of potential bidders.

[13] An incomplete list includes Haile and Tamer (2003), Haile, Hong, and Shum (2003), Guerre, Perrigne, and Vuong (2009), Aradillas-López, Gandhi, and Quint (2011) and Aryal, Grundl, Kim, and Zhu (2018).

sharp upper bound for the valuation quantile function can then be written as

$$\bar{v}_n(\alpha) = \sup_{q \in \Theta} q(\alpha), \text{ s.t. } b_m(t) \leq q(t) \ \ \forall m \in \mathbb{N}_1(n), \ \ \forall t \in [0,1], \qquad (10)$$
$$\beta_m(t, q) \leq b_m(t), \forall m \in \mathbb{N}_2(n), \forall t \in [0,1].$$

The first constraint in (10) imposes rationality. It holds only for auctions with a valuation quantile function weakly lower than that in n-bidder auctions. Otherwise, one cannot guarantee that the bid quantile function is bounded by the valuation quantile function in n-bidder auctions. The second constraint imposes RNBNE overbidding. It holds only for auctions with a valuation quantile function weakly higher than that in n-bidder auctions. Because the RNBNE bid quantile function is increasing in the valuation quantile, the auctions with a weakly higher valuation quantile have a weakly higher RNBNE bid quantile.

Similarly, the sharp lower bound can be written as

$$\underline{v}_n(\alpha) = \inf_{q \in \Theta} q(\alpha), \text{ s.t. } b_m(t) \leq q(t) \ \ \forall m \in \mathbb{N}_1(n), \ \ \forall t \in [0,1], \qquad (11)$$
$$\beta_m(t, q) \leq b_m(t), \forall m \in \mathbb{N}_2(n), \forall t \in [0,1].$$

One can show that this reduces to

$$\underline{v}_n(\alpha) = \max_{m \in \mathbb{N}_1(n)} \{b_m(\alpha)\}. \qquad (12)$$

We can simply take the maximum of the bid quantile functions from auctions with a valuation quantile weakly lower than that in n-bidder auctions.

Proposition 4. *If bidders are rational and overbid compared to RNBNE, the sharp identified set for $v_n(\alpha)$ is $[\underline{v}_n(\alpha), \bar{v}_n(\alpha)]$, where $\bar{v}_n(\alpha)$ and $\underline{v}_n(\alpha)$ are defined in (10) and (12).*

Proof. The bounds are sharp by construction and $v_n(\alpha)$ is interval identified because the set of quantile functions that satisfy overbidding and rationality is convex. □

Similar to Proposition 2, we can rewrite (10) as a one-dimensional optimization problem. Define

$$q_{\alpha,x}(t) = \begin{cases} \max_{m \in \mathbb{N}_1(n)} b_m(t) & \text{if } t < \alpha \\ \max\left\{x, \max_{m \in \mathbb{N}_1(n)} b_m(t)\right\} & \text{if } t \geq \alpha \end{cases}. \qquad (13)$$

By an argument similar to that proceeding Proposition 2, one can obtain the following result.

Proposition 5. *With variation in the number of bidders, the upper bound for the valuation quantile can be expressed as:*

$$\bar{v}_n(\alpha) = \bar{v}_n^a(\alpha) = \max_{x \geq \max_{m \in \mathbb{N}_1(n)} b_n(\alpha)} x, \;\; s.t. \; \beta_m(t, q_{\alpha,x}) \leq b_m(t), \;\; \forall t \in [\alpha, 1], \forall n \in \mathbb{N}_2(n). \tag{14}$$

It is straightforward to impose unimodality in addition to variation in n. One only needs to add the unimodality restriction to the constraints in (10) and (11). In the interest of space, we omit the details but would like to comment that the resulting bounds are sharp and that $\underline{v}_n(\alpha) = \max_{m \in \mathbb{N}_1(n)} \{b_m(\alpha)\}$ if $\max_{m \in \mathbb{N}_1(n)} \{b_m(\alpha)\}$ has a unimodal density.

4 Bounds for Counterfactuals

A fundamental challenge to the partial identification approach is counterfactual policy analysis.[14] There are two types of counterfactuals in our context: (1) counterfactuals that depend only on the valuation distributions; and (2) counterfactuals that also depend on other unknown parameters. The first type includes the average valuation of the good and the counterfactual profit in second-price auctions.[15] Another example would be the expected revenue of the seller by making a sequence of take-it-or-leave-it offers to the bidders. The second type includes the profit in first-price auctions with a binding reserve price, where counterfactual bid strategies depend on unknown parameters other than the valuation distribution. Formally, a counterfactual of the first type can be considered as a known function of the valuation quantile function $\Phi(q)$. We can obtain bounds for the counterfactual by finding the supremum and the infimum of $\Phi(q)$ over the set of q that rationalize the data under rationality and overbidding. A counterfactual of the second type can be considered as a known function of the valuation quantile function and other unknown parameters γ: $\Phi(q, \gamma)$. One may obtain bounds for $\Phi(q, \gamma)$ by optimizing over the set of (q, γ) that can rationalize the observed bid quantile functions under rationality, overbidding and potentially

[14]This challenge arises not because the partial identification approach does not impose a particular model when estimating the parameters, but because it does not impose a particular model on how agents behave under counterfactual scenarios. For example, in first-price auctions it is common to assume that the bidders play the risk-neutral symmetric Bayes Nash equilibrium with and without a binding reserve price. At the estimation stage this assumption helps to obtain point identification of the valuation distribution. However, this assumption also guarantees point identification of the counterfactual profit function and the optimal reserve price.

[15]In this section we assume that the seller's valuation of the auctioned good, c, is known. While an actual seller may know c, it is typically unknown in most applications in the structural auction literature. In this case one has to focus on the counterfactual revenue by setting the seller's cost to $c = 0$. We provide bounds for the counterfactual profits that depend on c in this section for generality, and set $c = 0$ in the later sections.

additional assumptions. For the bounds to be informative, the observed data need to contain sufficient information on γ. In this section, we first show how to construct bounds on the profit with a binding reserve price in first-price auctions. We characterize the sharp bounds and also propose conservative but easy-to-compute bounds. As will be shown in the next section, the difference between the sharp and the easy-to-compute bounds are negligible in many cases. We then show how to construct sharp bounds on the profit in second-price auctions.

4.1 Sharp Profit Bounds in First-Price Auctions

If we relax the assumption that bidders play the risk neutral Bayes Nash equilibrium, the bidding strategy under a binding reserve price r is unknown. To proceed, one can be very general and does not need to impose more than rationality and overbidding. In that case, we cannot rule out the possibility that the bidders bid their valuation for all $r > 0$, which leads to the sharp upper bound on profit. Similarly, we cannot rule out the possibility that bidders bid their RNBNE strategy for all $r > 0$, which leads to the sharp lower bound. However, these bounds are in general very wide. Therefore, we impose a slightly stronger assumption that is robust to most of the overbidding models. To this end, let $s_n(t,r)$ be the bidding strategy under a reserve price r.

Assumption 4. *For any reserve price $r \geq 0$, (1) if $t < r$, the bidding strategy $s_n(t,r)$ satisfies $s_n(t,r) = 0$; (2) if $t \geq r$, s_n solves the following differential equation with the initial condition $s_n(r,r) = r$,*

$$\frac{\partial s_n(t,r)}{\partial t} = (n-1)\lambda_n[t - s_n(t,r)]H_n(t)\frac{f_n(t)}{F_n(t)}$$

where $H_n(\cdot) > 0$ and $\lambda_n(\cdot)$ is a non-negative, weakly increasing function with $\lambda_n(0) = 0$.

Assumption 4 implicitly requires that bidders use the same strategy but is otherwise very general and satisfied by most first-price auction models that can explain overbidding. The functions λ_n and H_n allow for deviations from RNBNE that can lead to overbidding, where λ_n captures deviations of bidder preferences from risk neutrality and H_n captures deviations from correct beliefs. They allow multiple overbidding motives to exist at the same time. Different auction models imply different λ_n and H_n. If bidders play RNBNE, both λ_n and H_n are the identity function. If bidders are instead risk averse with a utility function U and ambiguity averse with a maxmin expected utility, then $\lambda_n = U/U'$ and H_n captures the difference between the true valuation distribution and bidders' most pessimistic belief on the valuation distribution. We prove the following lemma in Appendix C.[16]

[16]Assumption 4 requires the bidding strategy to be a fixed point. Therefore, it implicitly assumes that

Lemma 3. *Models with Risk Aversion, Ambiguity Aversion (Maxmin Expected Utility or Choquet Expected Utility), Loss Aversion, Loser Regret all satisfy Assumption 4.*

Assumption 4 alone does not improve the bounds on the valuation quantile function. Bidders may bid even lower than the RNBNE bidding strategy under this assumption. However, with the help of the assumption that bidders overbid with respect to the RNBNE in distribution at every reserve price r, it improves the bounds on the valuation quantile.[17] Interestingly, the improvement is very small or even negligible if unimodality and variation in n are used, as will be shown in the next section. Therefore, we focus on how to obtain the bounds for profit under this assumption. Assumption 4 allows us to link bidding strategies under different r through λ_n and H_n. This additional structure helps to produce informative profit bounds. Our identification strategy does not require knowing λ_n and H_n, but instead optimizes over the set of (q, λ_n, H_n) that rationalizes the data under the identifying restrictions. This procedure may be greatly simplified by exploiting two implications of Assumption 4: first, $s_n(t, r)$ increases with r if $t > r$; second, $\partial s_n(t, 0) / \partial t$ is an upper bound of $\partial s_n(t, r) / \partial t$ for any $r > 0$ and $t > r$.

To start, notice that the profit function in an n-bidder auction under a reserve price r, a valuation quantile function q and $\gamma = (\lambda_n, H_n)$ is

$$\Phi_n(q, \gamma, r) = \int_{q(\alpha) \geq r} [\rho_n(\alpha, q, \gamma, r) - c] \, d\alpha^n + c,$$

where c is the seller's valuation of the good and $\rho_n(\alpha, q, \gamma, r)$ is the bid quantile function under q, γ and r. If $q(\alpha) \geq r$, it solves the differential equation

$$\frac{\partial}{\partial \alpha} \rho_n(\alpha, q, \gamma, r) = (n-1) \frac{\lambda_n [q(\alpha) - \rho_n(\alpha, q, \gamma, r)]}{\alpha} \mathcal{H}_n(\alpha) \tag{15}$$

with the initial condition $\rho_n(q^{-1}(r), q, \gamma, r) = r$, where $\mathcal{H}_n(\alpha) = H_n(q(\alpha))$. Otherwise,

bidders have infinite level of rationality and excludes Level-k rationality.

[17] One can impose RNBNE overbidding in two ways. First, bidders overbid compared to the RNBNE strategy only in the observed data. Second, bidders overbid compared to the RNBNE strategy in the data and under counterfactual reserve prices. Under Assumption 4, the former does not lead to additional restrictions on the valuation quantile while the latter does. We focus on the latter in the rest of the paper. But if one wants to impose the former, one can simply remove the constraints that involve counterfactual reserve prices.

$\rho_n(\alpha, q, \gamma, r) = 0$. The sharp upper bound on profit can be expressed as

$$\bar{\pi}_n(r) = \sup_{q,\gamma} \Phi_n(q, \gamma, r), \tag{16}$$

$$\text{s.t. } q \text{ and } \gamma \text{ rationalizes } b_n \text{ under rationality and overbidding,}$$
$$q \text{ and } \gamma \text{ consistent with rationality and overbidding } \forall r > 0,$$
$$\text{other restrictions,}$$

where other restrictions include unimodality and restrictions introduced by variation in n. We impose unimodality throughout this section, but one can remove it by removing corresponding constraints. The sharp lower profit bound, $\underline{\pi}_n(r)$, can be expressed in the same way except that supremum is replaced by infimum. These bounds are conceptually straightforward but difficult to compute. The main challenge is that they involve nonlinear optimization over three functions: q, λ_n and H_n. We now show how to reduce these problems to simpler problems. To ease the presentation, we assume that Assumption 4 is imposed only for the n at which the counterfactual profit is calculated.

We first show how $\bar{\pi}_n(r)$ can be obtained from an optimization problem that involves only q. Because λ_n is weakly increasing by Assumption 4, $\partial s_n(t, r)/\partial t \leq \partial s_n(t, 0)/\partial t$ for any $r \geq 0$ and $t \geq r$. Therefore, if q and λ_n rationalize $b_n(\cdot)$ under Assumption 4 with some H_n, $\partial \rho_n(\alpha, q, \gamma, r)/\partial \alpha \leq b_n'(\alpha)$ for all $\alpha \geq q^{-1}(r)$. Moreover, the assumption that $\lambda_n(0) = 0$ guarantees rationality, i.e., $\rho_n(\alpha, q, \gamma, r) \leq q(\alpha)$. Then $\rho_n(\alpha, q, \gamma, r) \leq \bar{\rho}_n(\alpha, q, r)$ where $\bar{\rho}_n(\alpha, q, r) = 0$ if $q(\alpha) < r$; $\bar{\rho}_n(\alpha, q, r) = r$ if $q(\alpha) = r$; and

$$\frac{\partial}{\partial \alpha} \bar{\rho}_n(\alpha, q, r) = \begin{cases} \min\{b_n'(\alpha), q'(\alpha)\} & \text{if } \bar{\rho}_n(\alpha, q, r) = q(\alpha) \\ b_n'(\alpha) & \text{if otherwise} \end{cases}, \tag{17}$$

if $q(\alpha) > r$. It turns out that $\bar{\rho}_n(\alpha, q, r)$ is the sharp upper bound of the bid quantile under q if q rationalizes $b_n(\cdot)$ with some λ_n and H_n under Assumption 4. Intuitively, such a q can rationalize $b_n(\cdot)$ with a λ_n which is constant except in an arbitrarily small neighborhood of 0, and $H_n[q(\alpha)] = b_n'(\alpha)/\{\lambda_n[q(\alpha) - b_n(\alpha)](n-1)\}$. Under such λ_n and H_n, the bid quantile function under $r > 0$ either stays arbitrarily close to $q(\alpha)$ or has a derivative equal to $b_n'(\alpha)$ according to Assumption 4. Therefore, it is arbitrarily close to $\bar{\rho}_n(\alpha, q, r)$ because by definition, $\bar{\rho}_n(\alpha, q, r)$ either equals $q(\alpha)$ or has a derivative equal to $b_n'(\alpha)$. Notice that if $q'(\cdot) \geq b_n'(\cdot)$, then $\bar{\rho}_n(\alpha, q, r)$ reduces to $b_n(\alpha) - b_n(q^{-1}(r)) + r$. Because $\bar{\rho}_n(\alpha, q, r)$ does not depend on γ, we only need to optimize with respect to q to obtain $\bar{\pi}_n(r)$:

$$\begin{aligned}\bar{\pi}_n(r) = &\sup_{q\in\Theta}\int_{q(\alpha)\geq r}[\bar{\rho}_n(\alpha,q,r)-c]\,d\alpha^n + c \\ &\text{s.t. } b_m(t)\leq q(t), \quad \forall t\in[0,1] \text{ and } m\in\mathbb{N}_1(n) \\ &\quad \beta_m(t,q)\leq b_m(t) \quad \forall t\in[0,1] \text{ and } m\in\mathbb{N}_2(n) \\ &\quad \bar{\rho}_n(t,q,x)\geq\beta_n(t,q,x), \quad \forall x>0 \text{ and } t\in[0,1] \\ &\quad q \text{ has a unimodal density,}\end{aligned} \tag{18}$$

where $\beta_n(\cdot,q,x)$ is the RNBNE bid quantile function under q and a reserve price x; and $\mathbb{N}_1(n)$ and $\mathbb{N}_2(n)$ are defined in Section 3.3. The first two constraints impose rationality, RNBNE overbidding, and variation in n. The third constraint imposes overbidding for all positive reserve prices. If Assumption 4 is not imposed, this restriction reduces to the rationality restriction because we cannot rule out the possibility that bidders bid their valuations. Unimodality can be imposed using the method discussed in Section 3.2.

Next, we move to $\underline{\pi}_n(r)$. Notice that q and γ can rationalize $b_n(\cdot)$ under Assumption 4 if and only if

$$b_n'(\alpha) = (n-1)\frac{\lambda_n[q(\alpha)-b_n(\alpha)]}{\alpha}\mathcal{H}_n(\alpha)$$

and $b_n(0)=q(0)$. We can use this expression to eliminate $\mathcal{H}_n$ in (15) and obtain that if $r>0$ and $q(\alpha)>r$, ρ_n satisfies

$$\rho_n'(\alpha,q,\gamma,r) = \frac{\lambda_n[q(\alpha)-\rho_n(\alpha,q,\lambda_n,r)]\,b_n'(\alpha)}{\lambda_n[q(\alpha)-b_n(\alpha)]}, \tag{19}$$

and the initial condition $\rho_n(q^{-1}(\alpha),q,\gamma,r)=r$. Notice that ρ_n depends only on q and λ_n because rationalizing the data implies a particular H_n. Then the sharp lower bound for profit is

$$\begin{aligned}\underline{\pi}_n(r) = &\inf_{q\in\Theta,\lambda_n\in\Lambda}\Phi_n(q,\gamma,r) \\ &\text{s.t. } b_m(t)\leq q(t), \quad \forall t\in[0,1] \text{ and } m\in\mathbb{N}_1(n) \\ &\quad \beta_m(t,q)\leq b_m(t) \quad \forall t\in[0,1] \text{ and } m\in\mathbb{N}_2(n) \\ &\quad \rho_n(t,q,\gamma,x)\geq\beta_n(t,q,x), \quad \forall x>0 \text{ and } t\in[0,1], \\ &\quad q \text{ has a unimodal density.}\end{aligned} \tag{20}$$

Unlike the upper bound, it is in general not possible to obtain the sharp lower profit bound from a problem that involves only q. But if no variation in n is used and $b_n(\cdot)$ has a unimodal density, we can further simplify $\underline{\pi}_n$. By Assumption 4, if q and γ rationalize $b_n(\cdot)$ under rationality and overbidding, then $\rho_n(\alpha, q, \gamma, r) \geq b_n(\alpha)$, which implies

$$\Phi_n(q, \gamma, r) \geq \int_{q(\alpha)\geq r} [b_n(\alpha) - c]\, d\alpha^n + c \geq \int_{b_n(\alpha)\geq r} [b_n(\alpha) - c]\, d\alpha^n + c,$$

where the second inequality holds because $b_n(\alpha)$ is the lowest valuation quantile function consistent with the data under rationality and overbidding. Next, suppose that $H_n = 1$ and λ_n is generated by a CRRA utility function. The resulting bidding strategy satisfies rationality and overbidding under every $r > 0$. As the CRRA coefficient approaches 1, $q(\cdot)$ needs to be arbitrarily close to $b_n(\cdot)$ in order to rationalize the data. If $r > 0$, bids are close to valuations if valuations are higher than r. Therefore, $\rho_n(\alpha, q, \gamma, r)$ is arbitrarily close to $b_n(\alpha)$ if $q(\alpha) > r$, implying

$$\underline{\pi}_n(r) = \int_{b_n(\alpha)\geq r} [b_n(\alpha) - c]\, d\alpha^n + c. \tag{21}$$

Proposition 6. *Suppose that bidders are rational and overbid compared to RNBNE for any reserve price r. If Assumption 4 and unimodality hold, then $\bar{\pi}_n(r)$ and $\underline{\pi}_n(r)$ defined by (18) and (20) are the sharp upper and lower bounds of $\pi_n(r)$. If no variation in n is used and $b_n(\cdot)$ has a unimodal density, $\underline{\pi}_n(r)$ reduces to (21).*

4.2 Easy-to-Compute Profit Bounds in First-Price Auctions

Although calculating the sharp profit bounds is feasible, it can be time consuming. This is particularly true for the sharp lower profit bound if variation in n and unimodality are used, where one needs to solve many optimization problems involving both q and λ_n. To remedy this problem, we develop easy-to-compute profit bounds. They are based directly on bounds for v_n and do not involve optimization.

We start with the upper bound. In the previous section, we have established that $\int_{q(\alpha)\geq r} [\bar{\rho}_n(\alpha, q, r) - c]\, d\alpha^n + c$ is a sharp upper bound on profit if the true valuation quantile is q. To obtain the easy-to-compute bound, we replace $\bar{\rho}_n(\alpha, q, r)$ with its form when $q'(\cdot) \geq b_n'(\cdot)$ and replace q with a pointwise upper bound for v_n. Define

$$\tilde{\pi}_n(r) = \int_{\bar{v}_n(\alpha)\geq r} [\tilde{\rho}_n(\alpha, \bar{v}_n, r) - c]\, d\alpha^n + c,$$

where

$$\tilde{\rho}_n(\alpha, q, r) = \min\left\{q(\alpha), b_n(\alpha) - b_n\left(q^{-1}(r)\right) + r\right\},$$

and $\bar{v}_n$ is the sharp pointwise upper bound for the valuation quantile function without imposing Assumption 4. For example, if one does not want to impose unimodality or variation in n, $\bar{v}_n$ is defined by (3). If one wants to impose only unimodality, $\bar{v}_n$ is defined by (7). The following lemma implies that $\tilde{\pi}_n(r)$ is valid upper bound for profit.

Lemma 4. *If $q_1(\alpha) \geq q_2(\alpha)$ for all $\alpha \in [0,1]$, then $\bar{\rho}_n(\alpha, q_1, r) \geq \bar{\rho}_n(\alpha, q_2, r)$ for all $\alpha \in [0,1]$.*

Notice that $\tilde{\pi}_n$ is not sharp for two reasons. First, $\bar{v}_n$ is the sharp pointwise upper bound for v_n, which is in general larger than the valuation quantile function that solves (18). Second, $\tilde{\rho}_n(\alpha, \bar{v}_n, r)$ may be larger than $\bar{\rho}_n(\alpha, \bar{v}_n, r)$. However, in our numerical examples and in the empirical application, $\bar{v}_n'(\cdot) > b_n'(\cdot)$ and $\tilde{\rho}_n(\alpha, \bar{v}_n, r) = \bar{\rho}_n(\alpha, \bar{v}_n, r)$.

Now we move to the lower bound. Because bidders overbid compared to the RNBNE in distribution, the bid quantile function under $r > 0$ is at least $\beta_n(\cdot, \underline{v}_n, r)$, where $\underline{v}_n(\cdot)$ is the sharp pointwise lower bound for the valuation quantile without Assumption 4. Moreover, Assumption 4 implies that the bid quantile under $r > 0$ is no less than $b_n(\cdot)$. Therefore, a lower profit bound is

$$\underset{\sim}{\pi}_n(r) = \int_{\underline{v}_n(\alpha) \geq r} \left[\max\left\{\beta_n(\alpha, \underline{v}_n, r), b_n(\alpha)\right\} - c\right] d\alpha^n + c.$$

Several comments are in order. First, we do not impose Assumption 4 when calculating the sharp bounds for the valuation quantile used in the profit bounds. This helps to reduce the computational burden, which is particularly important for the lower bound. Without Assumption 4, $\underline{v}_n(\alpha) = \max_{m \in \mathbb{N}_1(n)} b_m(\alpha)$ if $\max_{m \in \mathbb{N}_1(n)} b_m(\alpha)$ has a unimodal density. But with Assumption 4, we need to solve many optimization problems with respect to q and λ_n to obtain $\underline{v}_n(\alpha)$. Second, to prove the validity of $\tilde{\pi}_n(r)$ it is important that all bidders use the same bidding strategies both in the data ($r = 0$) and in counterfactuals ($r > 0$), although $\tilde{\pi}_n(r)$ appears to be valid even if bidders adopt heterogeneous strategies, as shown in Section 5. Notice that Assumption 4 implies only that the derivative of the bid function under $r = 0$ bounds that of the bid function at the same valuation under $r > 0$. But in $\tilde{\pi}_n(r)$, we use the derivative of the bid quantile function at $r = 0$ to bound that of the bid quantile function at the same quantile under $r > 0$. This step is valid because symmetry in bidding strategy under $r = 0$ and under $r > 0$ guarantees that the α-th quantiles of the bid distributions under $r = 0$ and under $r > 0$ correspond to the same valuation. Third, unlike $\underline{\pi}_n(r)$, $\underset{\sim}{\pi}_n(r)$ uses only one implication of Assumption 4: a higher r induces active bidders to bid more. In fact, $\underset{\sim}{\pi}_n(r)$ is a sharp lower bound if we replace Assumption 4 with this implication. This is because without additional assumptions, we cannot rule out the possibility that if $r > 0$, bidders bid the maximum of the RNBNE strategy and the strategy without a reserve price. One advantage of not using all implications of Assumption 4 is robustness. The easy-to-compute lower profit bound is valid if the bidding strategies are heterogeneous, but the sharp lower bound may be invalid, as will be shown later.

Proposition 7. *Suppose bidders are rational and overbid compared to RNBNE for any reserve price r. If active bidders bid more under a higher r, $\underset{\sim}{\pi}_n(r)$ is the sharp lower bound for $\pi_n(r)$. If Assumption 4 holds, $\tilde{\pi}_n(r) \geq \pi_n(r) \geq \underset{\sim}{\pi}_n(r)$.*

4.3 Second-Price Auctions

In a second-price auction with a reserve price r, the bidder with the highest bid wins the auction if his bid is higher than r and pays the maximum of r and the second highest bid. The dominant strategy is "truthful bidding", i.e. to bid one's valuation. Because the bidding strategy is known, the profit of the seller depends only on the valuation quantile (distribution) function. The counterfactual profit under a reserve price r and a quantile function q in an n-bidder auction is

$$\Phi_n(q,r) = cq^{-1}(r)^n + nq^{-1}(r)^{n-1}\left[1 - q^{-1}(r)^{n-1}\right] r + \int_{q(\alpha)>r} q(\alpha)\, d\left[\alpha^n + n\alpha^{n-1}(1-\alpha)\right].$$

One can obtain the lower and upper bounds for $\Phi_n(q)$ by searching over the set of valuation quantile functions that can rationalize the data with overbidding and rationality. The resulting method is the same as the method used in Section 3 for the valuation quantile function except the objective function is different.Formally, the sharp upper profit bound is

$$\begin{aligned}
\bar{\Phi}_n(r) = & \sup_{q\in\Theta} \Phi_n(q,r) \\
& \text{s.t. } b_m(t) \leq q(t), \quad \forall t \in [0,1] \text{ and } m \in \mathbb{N}_1(n) \\
& \qquad \beta_m(t,q) \leq b_m(t) \quad \forall t \in [0,1] \text{ and } m \in \mathbb{N}_2(n) \\
& \qquad q \text{ has a unimodal density.}
\end{aligned}$$

The sharp lower profit bound $\underline{\Phi}_n(r)$ can be obtained in the same way except supremum is replaced with infimum. Note that we do not need additional assumptions on bidding in second-price auctions because we implicitly assume that bidders adopt the dominant strategy.

5 Illustrating the Identification Results

In this section we illustrate the identification results using numerical examples. First, we consider the case where bidders adopt the same bidding strategy. They play the risk neutral

Bayesian Nash equilibrium and their valuations are drawn from a mixture of a beta distribution with parameters $\alpha = 2$ and $\beta = 8$ and a uniform distribution on $[0, 1]$. The mixture weights are 99 percent on the beta distribution and 1 percent on the uniform distribution. The small weight on the uniform distribution has a negligible effect on the shape of the mixture, but ensures that the density is strictly positive on $[0, 1]$. We observe auctions with $n = 2$ and $n = 7$.

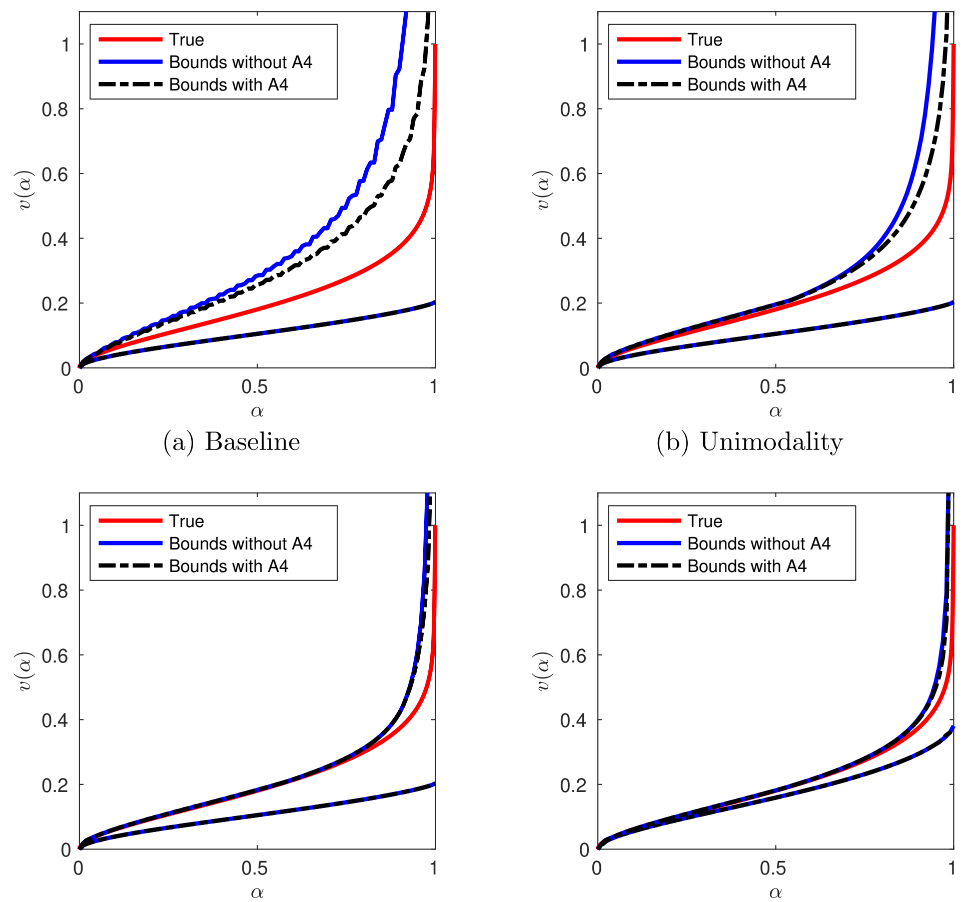


Figure 4: **Bounding the Valuation Quantile Function:** These graphs show how to bound the valuation quantile for $n = 2$. Panels (a) and (b) only use bid data from $n = 2$. Panels (c)-(d) also use bid data from $n = 7$ to tighten the bounds. Bounds with A4 impose Assumption 4.

5.1 Valuation Quantile Function

Figure 4 illustrates the identification results for the valuation quantile function with $n = 2$. Figures 4(a) and 4(b) do not use variation in the number of bidders. Figures 4(c) and 4(d) use Assumption 3 (SIV), and the stronger Assumption 2 (EP), to tighten the bounds. They also use bid data with $n = 7$ in addition to data with $n = 2$.[18] Figure 4(a) does not impose unimodality while the rest imposes unimodality. Each figure shows bounds with and without Assumption 4 to assess how much it affects the valuation bounds.

Baseline First consider the baseline case in Figure 4(a). The true valuation quantile function $v_2(\alpha)$ is shown in red, the solid blue line shows the baseline bounds that only impose overbidding and rationality. The black dashed lines show the bounds if Assumption 4 is imposed as well, which tightens the upper bound considerably. The true valuation quantile lies within the bounds but the bounds can be wide.

Unimodality Next consider Figure 4(b), where Assumption 1 (Unimodality) is imposed. This assumption tightens the upper bounds substantially. The lower bound is not affected because the bid density is unimodal. Also notice that after Assumption 1 is imposed the gap between the bounds with and without Assumption 4 is much smaller than in the baseline case.

Variation in n under SIV Now consider Figure 4(c), which uses bid data from $n = 7$ and imposes Assumption 3 (SIV) to further tighten the bounds. Notice that bid data from $n = 7$ do not allow us to tighten the lower bound for valuation quantile function under $n = 2$, i.e., $\underline{v}_2(\alpha) = b_2(\alpha)$. Intuitively, observing that bidders in auctions with $n = 7$ bid more aggressively does not allow us to tighten the lower bound for $v_2(\alpha)$, because $v_7(\alpha)$ may be higher than $v_2(\alpha)$ under SIV. The upper bound, however, can be tightened considerably. Moreover, the gap between the bounds with and without Assumption 4 has almost disappeared.

Variation in n under EP Figure 4(d) shows that Assumption 2 (EP) allows us to tighten the lower bound relative to Assumption 3 (SIV) (Figure 4(c)). The upper bound is also tightened further. The resulting bounds are valid and very close to the true valuation quantile function. This shows that our method is very informative in this case. Moreover, the bounds with and without Assumption 4 are indistinguishable.

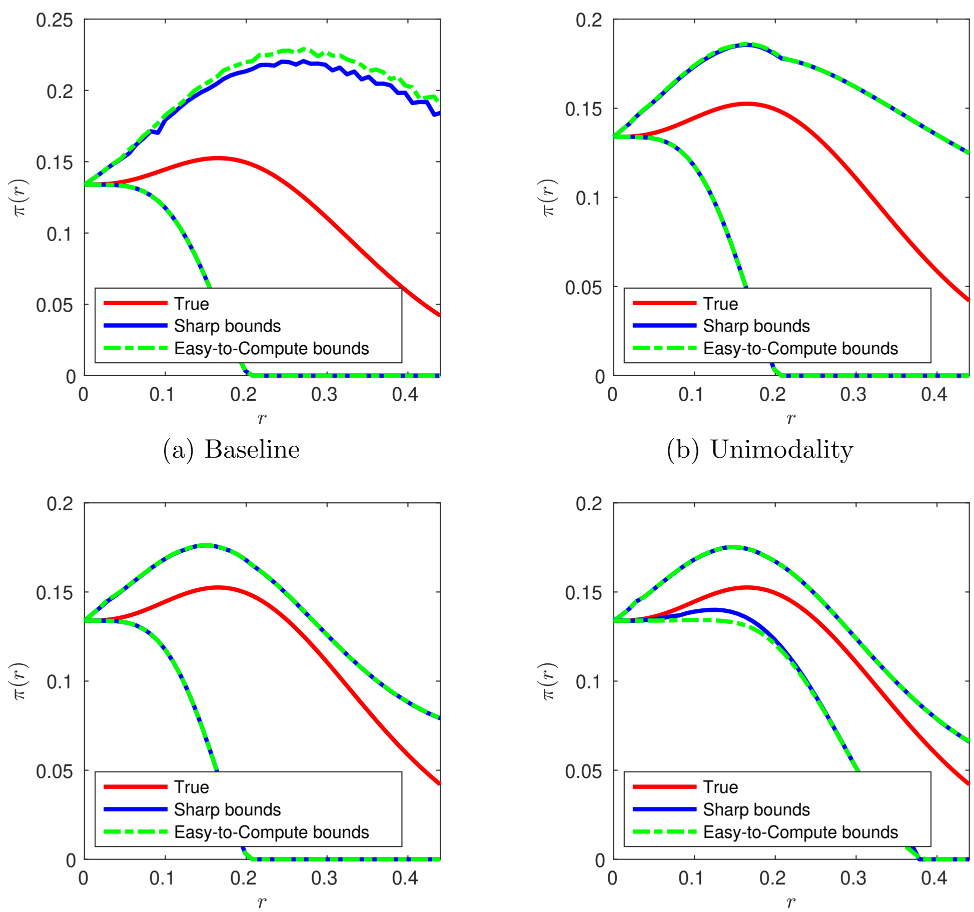
5.2 Revenue

Figure 5 illustrates bounds for the seller's revenue, which is obtained by setting $c = 0$. In each panel we show the true revenue function in red, the sharp revenue bounds in blue and

[18] Thus, the graphs only use variation in n from "one side" to tighten the bounds, i.e., from auctions with more bidders. In some cases this only allows us to tighten one of the bounds. This allows us to better illustrate the workings of the identification results than, for example, using data from $n = 2$ and $n = 7$ to tighten the bounds for $n = 3$.

the easy-to-compute bounds in green.

As for the valuation quantile function, Assumptions 1 (Unimodality), 3 (SIV), and 2 (EP) tighten the revenue bounds substantially compared to the baseline bounds which neither impose Assumption 1 nor exploit variation in n. The bounds are quite informative after imposing unimodality and EP. Except for the baseline case in panel (a), the easy-to-compute upper bounds are almost indistinguishable from the sharp upper bounds. There is a small but visible difference between the sharp lower bound and the easy-to-compute lower bound if Assumptions 1 (Unimodality) and 2 (EP) are imposed. This may not be surprising because unlike the easy-to-compute upper bound, the easy-to-compute lower bound exploits only one implication of Assumption 4.



(a) Baseline

(b) Unimodality

(c) Unimodality and SIV

(d) Unimodality and EP

Figure 5: **Bounding the Seller's Revenue:** These graphs show how to bound the seller's revenue for $n = 2$. Panels (a) and (b) only use bid data from $n = 2$. Panels (c)-(d) also use bid data from $n = 7$ to tighten the bounds. Notice that the range for the vertical axis is larger in panel (a). Each panel shows the true revenue function in red, the sharp revenue bounds in blue and the easy-to-compute bounds that are easier to compute in green.

Optimal Reserve Price Figure 6 illustrates how to bound the optimal reserve price. The blue solid lines are the sharp revenue bounds that impose Assumptions 1 (Unimodality) and 2 (EP). The vertical dashed black lines demarcate the range of reserve prices that may be optimal.

How can we translate the revenue bounds into the bounds for the optimal reserve price? If the maximum revenue under some reserve price r_1 is smaller than the minimum revenue under some alternative reserve price r_2, i.e., $\bar{\pi}_2(r_1) < \underline{\pi}_2(r_2)$, then r_1 cannot be optimal. The set of potentially optimal reserve prices contains all reserve prices that cannot be ruled out in this fashion.

This is shown in Figure 6. The horizontal dashed black line in the graph indicates the maximum of the lower revenue bound, $\underline{\pi}_2^{\max}$. Any reserve price r such that $\bar{\pi}_2(r) < \underline{\pi}_2^{\max}$ can therefore be ruled out. The reserve prices between the two vertical black lines satisfy $\bar{\pi}_2(r) \geq \underline{\pi}_2^{\max}$ and therefore cannot be ruled out. In this example, the optimal reserve price is bounded between 0.018 and 0.268. If we instead use the easy-to-compute bounds, the optimal reserve price is between 0.001 and 0.2796. This interval is about 11% wider than the one based on the sharp bounds. Notice that both bounds rule out $r = 0$ to be optimal. In terms of computing time, calculating the sharp revenue bounds with unimodality is infeasible for a desktop computer, while one can obtain the easy-to-compute bounds in less than a second once the bounds for the valuation quantile function have been obtained.

Which reserve price inside the bounds should the seller choose? The procedure outlined above only bounds the set of reserve prices, but ultimately the seller has to choose a particular reserve price. One way to choose the reserve price in partially identified models is to maximize the minimum revenue, i.e., maximize the revenue under a worst case scenario (Aryal and Kim (2013)). The resulting reserve price is 0.1267 using the sharp lower bound and 0.1089 using the easy-to-compute lower bound.

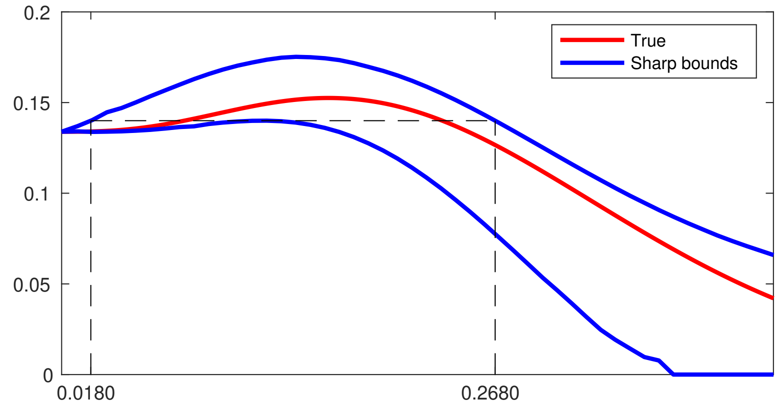


Figure 6: **Bounding the Optimal Reserve Price:** This graphs show how to bound the optimal reserve price for $n = 2$ using the bounds imposing the SP and EP restrictions. The vertical dashed black lines demarcate the range of reserve prices that may be optimal. The horizontal dashed purple line is the maximum revenue under the lower revenue bound.

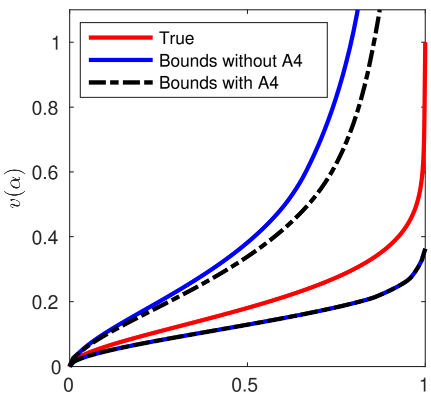
5.3 Heterogeneous Bidding Behavior

Lastly, we consider the case of heterogeneous bidding behavior. Suppose there are three bidder groups, each accounting for one-third of the bids in auctions with any given n. The first group overbids moderately compared to RNBNE. These bidders use the bid function of the symmetric BNE where bidders have CRRA with a coefficient of 0.1. The second group overbids more, using the bid function of the symmetric BNE with a CRRA coefficient of 0.5. The third group overbids the most, using the bid function of the symmetric BNE with a CRRA coefficient of 0.8.

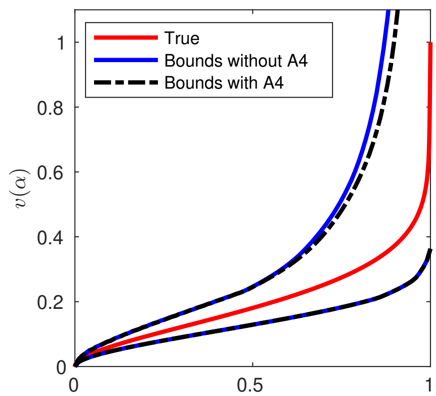
Figure 7 shows bounds for valuation quantile functions. Theory predicts that all bounds without Assumption 4 remain valid with heterogeneous bidders. This is indeed the case in this example. Similar to the case with homogeneous bidders, unimodality and variation in n tighten the bounds dramatically. Assumption 4 has limited impact on the bounds once unimodality is imposed. In particular, if both unimodality and variation in n are used, bounds with and without Assumption 4 are almost indistinguishable. Therefore, the bounds with Assumption 4 are also valid in this example. But it is worth pointing out that Assumption 4 implicitly imposes homogeneity, which is violated in this example. Therefore, the bounds with Assumption 4 are not guaranteed to be valid.

Figure 8 show bounds for revenues. The true revenue function assumes that bidders continues to use the BNE bidding strategies with binding reserve prices under corresponding CRRA coefficients. This again violates Assumption 4. Therefore, the sharp revenues bounds may not be valid. It turns out that the sharp lower bound is not valid if Assumption 2 is exploited as shown in Figure 8(d). The lower blue curve is above the red curve if r is small.

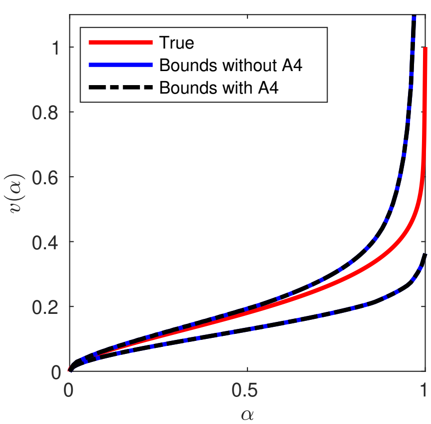
The easy-to-compute lower bound, however, stays valid. This is consistent with Proposition 7, which shows the easy-to-compure lower profit bounds is valid if bids increase with the reserve price. Interestingly, the sharp upper bound is valid in all cases. We conjecture that it is generally valid with heterogeneous bidding strategies if every strategy satisfies Assumption 4. Again, the easy-to-compute upper bound is indistinguishable from the sharp upper bound once unimodality is imposed.



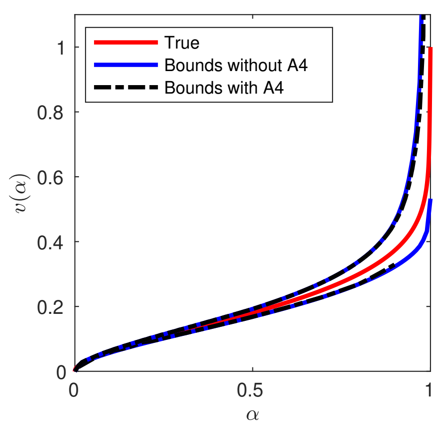
(a) Baseline



(b) Unimodality

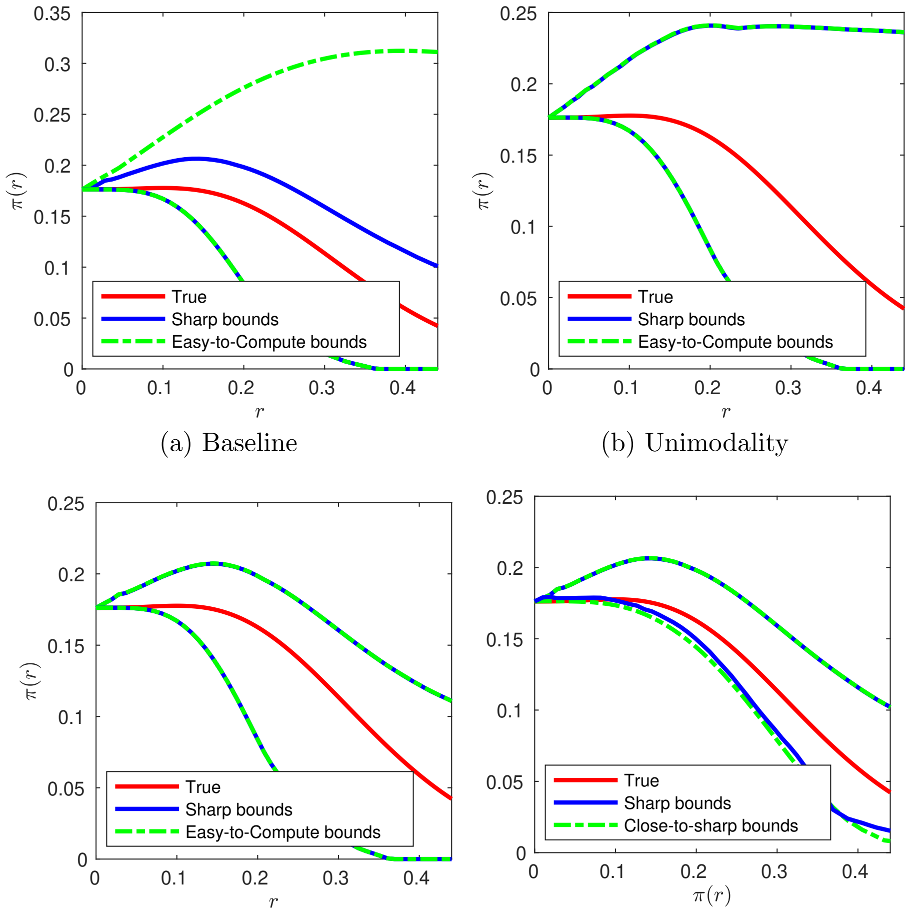


(c) Unimodality and SIV



(d) Unimodality and EP

Figure 7: **Bounding the Valuation Quantile Function (Bidder Heterogeneity):** These graphs show how to bound the valuation quantile functionfor $n = 2$. Bidders have heterogeneous bid functions: One third of the bidders has a CRRA coefficient of 0.3, one third has 0.5, and one third has 0.8. The graphs in the upper row only use bid data from $n = 2$. The graphs in the lower row also use bid data from $n = 7$ to tighten the bounds. Bounds with A4 impose Assumption 4.



(a) Baseline (b) Unimodality

(c) Unimodality and SIV (d) Unimodality and EP

Figure 8: **Bounding the Seller's Revenue (Bidder Heterogeneity):** These graphs show how to bound the seller's revenue for $n = 2$. Bidders have heterogeneous bid functions: One third of the bidders has a CRRA coefficient of 0.3, one third has 0.5, and one third has 0.8. The graphs in the upper row do not use variation in the number of bidders and only use bid data from $n = 2$. The in the lower row also use bid data from $n = 7$ to tighten the bounds. Notice that the range of the vertical axis is wider for panel (a).

6 Estimation

Suppose we observe N_n bids, $\{y_{i,n}\}_{i=1}^{N_n}$, from n-bidder auctions for each $n \in \mathbb{N}$. Let $N = \sum_{n\in\mathbb{N}} N_n$ be the number of total observations and assume that $\lim_{N\to\infty} N_n/N > 0$ for every $n \in \mathbb{N}$. The goal is to estimate the bounds for the valuation quantile function and the profit function. Estimators of bounds on revenue can be obtained by setting c to 0. Because unimodality tightens the bounds dramatically as shown in the previous section, we impose unimodality throughout this section.

We start with estimating the bounds for the valuation quantile function. If unimodality is not used, one can replace the bid quantile function in (5) by its sample counterpart and solve a one-dimensional problem to obtain an estimator for the upper bound. This procedure is simple and fast. However, it is not clear how to generalize it to incorporate the unimodality restriction. Therefore, we instead focus on estimators based on formulations like (3). To proceed, we first put restrictions on Θ so that it contains only sufficiently smooth functions.

Let $\|q\|_{sob}^2 = \sum_{m\leq M}\int_0^1 q^{(m)}(x)^2\,dx$ be the Sobolev norm, where m is an integer, $M \geq 2$ and $q^{(m)}(x)$ is the m-th derivative of q at x. Define the Sobolev space to be

$$\mathcal{W}(B) = \{q \in C[0,1] : \|q\|_{sob} \leq B\},$$

where $C[0,1]$ is the set of continuous functions defined on $[0,1]$ and B is a pre-specified positive constant. Notice that $\mathcal{W}(B)$ is compact under $\|\cdot\|_{sob}$. Next, define

$$\Theta = \{q \in \mathcal{W}(B) : q \text{ is weakly increasing and } q(0) \geq 0\}.$$

Therefore, Θ contains only quantile functions with bounded second-order derivatives. Notice that Θ is a closed subset of $\mathcal{W}(B)$ and hence compact under $\|\cdot\|_{sob}$. Define $\mathbf{b} = \{b_n\}_{n\in\mathbb{N}}$ to be the set of population bid quantile functions and

$$\begin{aligned}\Theta_R(\mathbf{b}) = \Big\{ & q \in \Theta : b_m(\cdot) \leq q(\cdot), \forall m \in \mathbb{N}_1(n); \\ & \beta_m(\cdot, q) \leq b_m(\cdot),\ \forall\ m \in \mathbb{N}_2(n);\ \text{and } q \text{ is unimodal}\Big\},\end{aligned} \tag{22}$$

where $\mathbb{N}_1(n)$ and $\mathbb{N}_2(n)$ are defined in Section 3.3 and depend on the assumption on variation in n. Then $\bar{v}_n(\alpha) = \max_{q\in\Theta_R(\mathbf{b})} q(\alpha)$ and $\underline{v}_n(\alpha) = \min_{q\in\Theta_R(\mathbf{b})} q(\alpha)$. If we have an estimator for $\Theta_R(\mathbf{b})$, denoted by $\hat{\Theta}_R \subset \Theta$, we can define the estimators for $\bar{v}_n(\alpha)$ and $\underline{v}_n(\alpha)$ to be $\hat{\bar{v}}_n(\alpha) = \max_{\hat{\Theta}_R} q(\alpha)$ and $\hat{\underline{v}}_n(\alpha) = \min_{\hat{\Theta}_R} q(\alpha)$. Also, define $\|q\|_\infty = \sup_{x\in[0,1]}|q(x)|$ and

$$d_\infty(q, A) = \inf_{\tilde{q}\in A} d_\infty(q, \tilde{q}).$$

One problem is that if $\hat{\Theta}_R$ is an arbitrary consistent estimator of $\Theta_R(\mathbf{b})$, $\hat{\bar{v}}_n$ and $\hat{\underline{v}}_n$ may not be defined with positive probability even if the sample size is large. This is because such an $\hat{\Theta}_R$ may be empty due to estimation error even if $\Theta_R(\mathbf{b})$ is not empty. Therefore, we introduce the following high-level assumption.

Assumption 5. *As $N \to \infty$, (1) $P\left(\Theta_R(\mathbf{b}) \subseteq \hat{\Theta}_R\right) \to 1$; (2) $\sup_{q\in\hat{\Theta}_R} d_\infty(q, \Theta_R(b)) \to^p 0$.*

Assumption 5(1) guarantees that $\hat{\bar{v}}_n$ and $\hat{\underline{v}}_n$ are well-defined with probability approaching

1 and Assumption 5(2) implies that $\hat{\Theta}_R$ converges to $\Theta_R(\mathbf{b})$ under the one-sided Hausdorff metric. They together imply that $\hat{\Theta}_R$ is a consistent estimator of $\Theta_R(b)$ under the two-sided Hausdorff metric. Notice that for the lower bound, if one is willing to assume that the quantile function $\max_{m\in\mathbb{N}_1(n)} b_m(\alpha)$ has a unimodal density, then $\underline{v}_n(\alpha) = \max_{m\in\mathbb{N}_1(n)} b_m(\alpha)$ and we can instead use $\hat{\underline{v}}_n^a(\alpha) = \max_{m\in\mathbb{N}_1(n)} \hat{b}_m(\alpha)$ as an estimator for $\underline{v}_n(\alpha)$.

Proposition 8. *Under Assumption 5, as $N \to \infty$,*

(1) $\sup_{\alpha\in[0,1]} \left|\hat{\bar{v}}_n(\alpha) - \bar{v}_n(\alpha)\right| \to^p 0$ *and* $\sup_{\alpha\in[0,1]} \left|\hat{\underline{v}}_n(\alpha) - \underline{v}_n(\alpha)\right| \to^p 0$*;*

(2) If $\max_{m\in\mathbb{N}_1(n)} b_m(\alpha)$ *has a unimodal density,* $\sup_{\alpha\in[0,1]} \left|\hat{\underline{v}}_n^a(\alpha) - \underline{v}_n(\alpha)\right| \to^p 0$*;*

Notice that convergence of $\hat{\bar{v}}_n(\alpha)$ is uniform on $[0,1]$. This is because our restriction on the parameter space implies that the true valuation quantile function is bounded away from infinity. Also notice that both $\hat{\underline{v}}_n(\alpha)$ and $\hat{\underline{v}}_n^a(\alpha)$ are valid lower bounds, although the latter may be conservative asymptotically if the density of $\max_{m\in\mathbb{N}_1(n)} b_m(\alpha)$ is not unimodal.

We next discuss how to choose $\hat{\Theta}_R$. Suppose we have estimators for bid quantile functions, which are denoted as $\hat{b}_n$. These estimators satisfy the following assumption.

Assumption 6. *As $N \to \infty$, (1)* $\sup_{m\in\mathbb{N}} \sup_{t\in[0,1]} \left|\hat{b}_m(t) - b_m(t)\right| \to 0$ *almost surely; (2)* $\sup_{m\in\mathbb{N}} \sup_{t\in[0,1]} \left|\hat{b}_m(t) - b_m(t)\right| = O_p\left(N^{-1/2}\right)$.

This assumption is very mild. One can show that $\hat{b}_m(t) = \inf\left\{b : \hat{G}_m(b) \geq t\right\}$, where $\hat{G}_m$ is the empirical bid distribution in auctions with m bidders, satisfies this assumption. One natural idea to obtain $\hat{\Theta}_R$ is to replace the bid quantile functions in $\Theta_R(\mathbf{b})$ by $\hat{b}_n$. However, the resulting set may not satisfy Assumption 5(1). To guarantee Assumption 5(1), we need to relax the constraints. To this end, let $\eta_N > 0$ be a sequence of real numbers that converges to 0. Define

$$\begin{aligned}\hat{\Theta}_R = \Big\{ & q \in \Theta : \hat{b}_m(\cdot) - \eta_N \leq q(\cdot), \forall m \in \mathbb{N}_1(n); \\ & \beta_m(\cdot, q) \leq \hat{b}_m(\cdot) + \eta_N,\ \forall\ m \in \mathbb{N}_2(n);\ \text{and } q \text{ is unimodal} \Big\}\end{aligned} \tag{23}$$

Lemma 5. *If $\eta_N \to 0$, $\sqrt{N}\eta_N \to \infty$ and Assumption 6 holds, then $\hat{\Theta}_R$ satisfies Assumption 5.*

We briefly discuss how to implement the estimator using $\hat{\bar{v}}_n$ as an example. First, approximate Θ by Θ_K, which is the space spanned by B-splines with K knots. We choose a large K so that the approximation error is negligible. Impose the constraints in $\hat{\Theta}_R$ on a fine grid in $[0,1]$. To impose unimodality, we first impose the restriction that $q'(\alpha)$ is decreasing if $\alpha < \alpha^*$ and increasing if $\alpha > \alpha^*$, where α^* takes values on a fine grid in $[0,1]$. This leads to a linear programming problem for each value of α^*. Then define $\hat{\tilde{v}}_n$ to be the largest value of the objective functions of all the linear programming problems. For more details, please see Section B in the online appendix.

For generality, we propose estimators for the profit bounds. The bounds on revenues can be obtained by setting c to 0. One can construct estimators for the sharp profit bounds following a strategy similar to the bounds for the valuation quantile function. The resulting estimators can be difficult to compute because they involve solving a large number of high-dimensional nonlinear optimization problems once unimodality is imposed. Moreover, they may not provide significant improvement over the easy-to-compute bounds once variation in n is used. Therefore, we recommend using the easy-to-compute bounds in practice and focus on the corresponding estimators. Define

$$\hat{\tilde{\pi}}_n(r) = \int_{\hat{\tilde{v}}_n(\alpha)\geq r} \left[\hat{\tilde{\rho}}_n\left(\alpha, \hat{\tilde{v}}_n(\cdot), r\right) - c\right] d\alpha^n + c$$

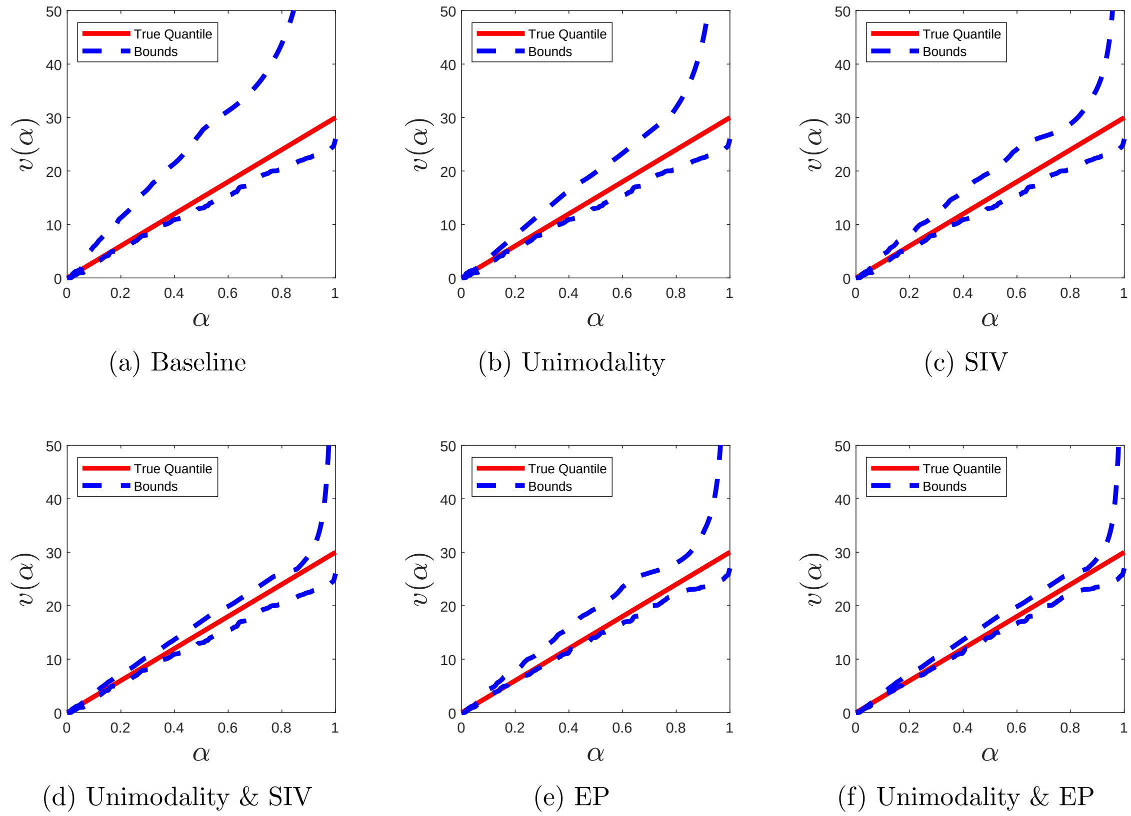
$$\hat{\underset{\sim}{\pi}}_n(r) = \int_{\hat{\underline{v}}_n(\alpha)\geq r} \left[\max\left\{\beta_n\left(\alpha, \hat{\underline{v}}_n(\cdot), r\right), \hat{b}_n(\alpha)\right\} - c\right] d\alpha^n + c.$$

Proposition 9. *Under Assumption 6, $\hat{\tilde{\pi}}_n(r) \to^p \tilde{\pi}_n(r)$ and $\hat{\underset{\sim}{\pi}}_n(r) \to^p \underset{\sim}{\pi}_n(r)$ for all $r > 0$ as $N \to \infty$.*

7 Application to Experimental Data

In this section we apply the estimators developed in the last section to the experimental data from Dyer, Kagel, and Levin (1989). As discussed in Section 2, although a non-negligible fraction of bidders bid lower than the RNBNE bidding strategy, RNBNE overbidding in distribution holds except for some slight violations at a small set of points with $\alpha < 0.1$. Statistically, Kolmogorov-Smirnov tests do not reject that the observed bid distributions first-order stochastically dominate the RNBNE distributions at any commonly used significance level, supporting our identifying restriction. They strongly reject that the bid distributions are generated by the RNBNE. As the valuation distribution does not vary with n, we can exploit variation in the number of bidders to tighten the bounds.[19]

[19] We set $\eta_N = 0.12$, which is about 1% of the average bids in the data under $n = 3$. Results are not sensitive to η_N.



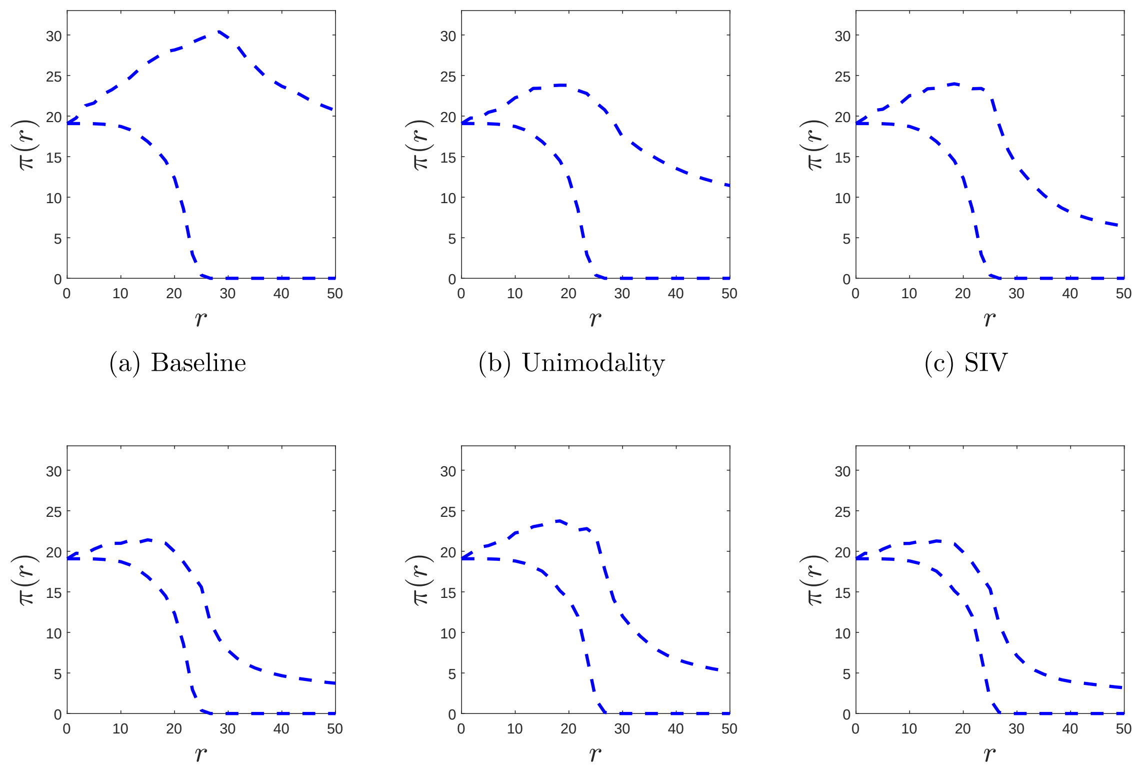
(a) Baseline (b) Unimodality (c) SIV

(d) Unimodality & SIV (e) EP (f) Unimodality & EP

Figure 9: **Bounding the Valuation Quantile Function:** Bounds on the valuation quantile function for auctions with three bidders. Panels (a) and (b) do not use variation in the number of bidders. Panels (c)-(f) use bid data from $n = 6$ to tighten the bounds under Assumptions 3 (SIV) and 2 (EP).

Valuation Quantile Function Figure 9 shows the estimated bounds for the valuation quantile function for $n = 3$. The true valuation quantile function is shown in red and the bounds in blue. Panels (a) and (b) do not use variation in n. Panels (c)-(f) use bid data from $n = 6$ to tighten the bounds under Assumptions 3 (SIV) and 2 (EP).

The graphs confirm that all the estimated bounds are valid even though some bidders bid lower than the RNBNE strategy. The baseline estimates in Figure 9(a) only use the restrictions of overbidding and rationality. While the lower bound is fairly tight, the upper bound is not. However, the upper bound can be tightened considerably even without using variation in n under Assumption 1 (Unimodality), as shown in Figure 9(b). Using variation in n helps to tighten the bounds further, as shown in Figures 9 (c)-(f). By combining Assumptions 1 (Unimodality) and 2 (EP) in Figure 9(f) we obtain tight bounds at most quantiles, except for α close to one.



(a) Baseline (b) Unimodality (c) SIV

(d) Unimodality and SIV (e) EP (f) Unimodality and EP

Figure 10: **Bounding the Seller's Revenue:** Bounds on the revenue function for $n = 3$. Figures (a) and (b) do not use variation in the number of bidders. Figures (c)-(f) use bid data from $n = 6$ to tighten the bounds under Assumptions 3 (SIV) and 2 (EP).

Seller Revenue Figure 10 shows the estimated bounds for the seller's revenue. All graphs start with a single point at $r = 0$, because the revenue without a reserve price is point identified from the data. For positive reserve prices the bounds in the baseline case (Figure 10(a)) can be tightened considerably by imposing additional restrictions. For instance, the baseline estimates suggest that the seller's revenue could be increased above 30 by imposing a high reserve price of about 30. However, the estimates in Figures 10(b) - (f) show that setting such a high reserve price would result in a revenue that is even smaller than not setting any reserve price at all. In fact the seller's revenue can at best be increased slightly by setting a positive reserve price.

Reserve Price Figure 11 illustrates the estimated bounds for the optimal reserve price under Assumptions 1 (Unimodality) and 2 (EP) for $n = 3$. Reserve prices between 0 and 20 may be optimal as indicated by the two dashed vertical lines. Reserve prices above 20 can be ruled out because the maximum revenue under these reserve prices is lower than the minimum revenue under some other reserve price, as indicated by the dashed horizontal line.

To choose among reserve prices between 0 and 20 requires some assumption about the preferences of the seller. For example, a maxmin seller would choose a reserve price of 0 because this yields the highest minimum revenue whereas a maxmax seller would choose a reserve price of 17 because this yields the highest maximum revenue.

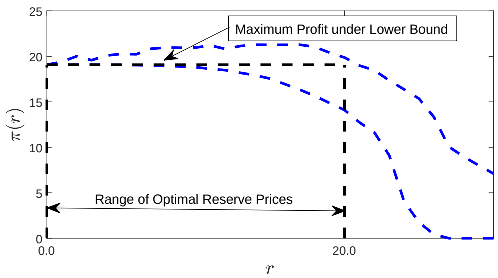


Figure 11: **Bounding the Reserve Price:** This graph shows the bounds for the optimal reserve price under Assumptions 1 (Unimodality) and 2 (EP) for $n = 3$. Reserve prices between 0 and 20 may be optimal as indicated by the two vertical lines. Reserve prices above 20 can be ruled out because the maximum revenue under these reserve prices is lower than the minimum revenue under some other reserve price, as indicated by the horizontal line.

8 Conclusion

This paper suggests a new link between experimental economics and structural estimation. Usually experiments are conducted to compare their findings to theoretical predictions. Hence, experimental work affects structural work only indirectly if it sparks the development of new models. However, in this paper we demonstrate that experimental findings can be used directly in structural estimation without imposing a particular model. This approach is particularly useful in cases where there is no agreement which model is best supported by the experimental findings, as is the case for first-price auctions.

In the context of first-price auctions, imposing the restrictions from experiments allows us to relax the assumptions of the canonical model, which appears to be inconsistent with findings in experiments and some field studies, and still obtain informative bounds on primitives. In other contexts it could be the case that while the canonical model is not rejected by experimental findings, it is also not identified from field data. In this case, imposing the restrictions from experiments in addition to the restrictions from the canonical model could help in identification.

We see two broad avenues for future work. First, structural researchers can ask whether using existing experimental findings to inform structural estimation is helpful in other contexts, for example, for identification and estimation of bargaining models. Second, experimental researchers can ask whether it is useful to conduct experiments specifically to inform structural estimation rather than to test theory.

A Proofs (For Online Publication)

A.1 Proofs of the Impossibility Results

For a description of the models discussed in Theorem 1 and Theorem 2 see Appendix C. The following lemma is handy in proving some of our results.

Lemma 6. *Suppose a first-price auction model leads to an equilibrium bidding strategy s_n that satisfies $s_n(0) = 0$ and for all $t > 0$*

$$s_n'(t) = (n-1)\lambda_n[t - s_n(t)]\frac{f_n(t)}{F_n(t)}H_n(t) \tag{24}$$

where $\lambda_n(0) = 0$, $\lambda_n(\cdot) \geq 0$, $H_n(\cdot) > 0$ and they satisfy a set of conditions. Suppose that G_n is a bid distribution that has a strictly positive and differentiable density function g_n supported on a bounded interval with a lower bound of 0. Then G_n can be rationalized by this auction model if there exist a $\tilde{\lambda}$ and $\mathbf{h}(\cdot) > 0$ such that

$$\xi_n(\alpha) = b_n(\alpha) + \tilde{\lambda}^{-1}\left[\frac{1}{n-1}\frac{\alpha}{g_n(b(\alpha))}\mathbf{h}(\alpha)\right] \tag{25}$$

is strictly increasing in α and $\tilde{\lambda}$ and $\tilde{H}(t) = 1/\mathbf{h}\left(\xi_n^{-1}(t)\right)$ satisfy the same set of conditions as λ_n and H_n.

Proof. Define the valuation distribution to be $\tilde{F}_n(t) = \xi_n^{-1}(t)$ and the bid function $\tilde{s}_n(t) = b_n\left(\tilde{F}_n(t)\right)$. Because $\tilde{\lambda}(0) = 0$, $\tilde{F}_n(0) = 0$. Moreover, $\xi_n(\cdot)$ is strictly increasing. Therefore, $\tilde{F}_n$ is a distribution function. Obviously, $\tilde{F}_n$ and $\tilde{s}_n$ are consistent with b_n. We next show that $\tilde{s}_n$ is a bidding strategy that is consistent with the model, i.e., $\tilde{s}_n(0) = 0$ and $\tilde{s}_n'$ satisfies the FOC with $\tilde{F}_n$, $\tilde{\lambda}$ and $\tilde{H}$. First, because the support of G_n has a lower bound of 0, $\tilde{s}_n(0) = b_n(0) = 0$. Next,

$$\tilde{s}_n'(t) = b_n'\left(\tilde{F}_n(t)\right)\tilde{f}_n(t) = \frac{\tilde{f}_n(t)}{g_n\left(b_n\left(\tilde{F}_n(t)\right)\right)}. \tag{26}$$

By (25),

$$(n-1)\frac{\tilde{\lambda}\left[\xi_n(\alpha)-b_n(\alpha)\right]}{\alpha}\frac{1}{\mathbf{h}(\alpha)}=\frac{1}{g_n(b_n(\alpha))}. \tag{27}$$

Use (26) to eliminate g_n from (27) and obtain

$$\begin{aligned}\tilde{s}_n'(t) &= (n-1)\frac{\tilde{\lambda}\left[\xi_n\left(\tilde{F}_n(t)\right)-b_n\left(\tilde{F}_n(t)\right)\right]}{\tilde{F}_n(t)}\tilde{f}_n(t)\\ &= (n-1)\tilde{\lambda}\left[t-s_n(t)\right]\frac{\tilde{f}_n(t)}{\tilde{F}_n(t)}\frac{1}{\mathbf{h}\left(\tilde{F}_n(t)\right)}.\end{aligned}$$

Define $\tilde{H}(t)=1/\mathbf{h}\left(\tilde{F}_n(t)\right)$ and then $\tilde{F}_n,\tilde{\lambda},\tilde{H}$ rationalize G_n under this first-price auction model. □

Proof of Theorem 1. Without loss of generality, we assume that the support of g_n has a lower bound of 0. We start with the Ambiguity Aversion models.

(1) The Choquet Expected Utility model (CEU) generates an equilibrium bidding strategy that satisfies $s_n(0)=0$ and (24) with $\lambda_n(x)=x$

$$H_n(t)=\frac{\phi'\left(F_n(t)^{n-1}\right)F_n(t)^{n-1}}{\phi\left(F_n(t)^{n-1}\right)} \tag{28}$$

for some convex function ϕ. We next construct $\tilde{\lambda}$ and $\tilde{H}$ that rationalize G_n. Suppose that $\tilde{\phi}(x)=x^{1/c}$ for some constant $0<c<1$ and $\tilde{\lambda}(x)=x$. Obviously, $\tilde{\phi}$ is a convex function. Define $\mathbf{h}(x)=c$, which is a constant function. By assumption,

$$\xi_n(\alpha)=b_n(\alpha)+\tilde{\lambda}^{-1}\left[\frac{1}{n-1}\frac{\alpha}{g_n(b(\alpha))}\mathbf{h}(\alpha)\right]=b_n(\alpha)+\frac{c}{n-1}\frac{\alpha}{g_n(b_n(\alpha))}$$

is increasing. Let $\tilde{H}(t)=\mathbf{h}\left(\xi_n^{-1}(t)\right)^{-1}=1/c$. By Lemma 6, CEU rationalizes G_n because for any $\tilde{F}_n$,

$$\tilde{H}(t)=\frac{1}{c}=\frac{\tilde{\phi}'\left(\tilde{F}_n(t)^{n-1}\right)\tilde{F}_n(t)^{n-1}}{\tilde{\phi}\left(\tilde{F}_n(t)^{n-1}\right)}$$

and $\tilde{\phi}$ is convex. In other words, $\tilde{H}$ satisfies the set of restrictions on H_n imposed by the model.

(2) The Maxmin Expected Utility model (MEU) generates an equilibrium bidding strategy

that satisfies $s_n(0) = 0$ and (24) with $\lambda_n(x) = x$ and

$$H_n(t) = \frac{F_n(t)}{f_n(t)} \frac{f_n^{\min}(t)}{F_n^{\min}(t)}, \tag{29}$$

where F_n and $F_n^{\min}$ are distribution functions and $F_n^{\min}$ first-order stochastically dominates F_n. Again, let $\tilde{\lambda}(x) = x$ and $\mathbf{h}(x) = c$ with $0 < c < 1$. Notice that $\tilde{\lambda}$ and $\mathbf{h}$ rationalize G_n with some $\tilde{F}_n$. We only need to show that $\tilde{H} = 1/c$ is consistent with the model. Define

$$\tilde{F}_n^{\min}(t) = \tilde{F}_n(t)^{1/c}.$$

Obviously, $\tilde{H}(t) = \frac{\tilde{F}_n(t)}{\tilde{f}_n(t)} \frac{\tilde{f}_n^{\min}(t)}{\tilde{F}_n^{\min}(t)} = 1/c$ and $\tilde{F}_n^{\min}$ first-order stochastically dominates $\tilde{F}_n$. Hence, G_n is consistent with MEU.

(3) For the Risk Aversion model, just notice that if the utility function has constant relative risk aversion (CRRA), then $\lambda^{-1}(x) = 1 - \theta$ where θ is the CRRA coefficient. Because $b + \frac{c}{n-1} \frac{G_n(b)}{g_n(b)}$ is increasing in b if $0 < c < 1$ by assumption, G_n can be rationalized by the Risk Aversion model with CRRA coefficient $\theta = 1 - c$.

(4) The Loss Aversion model implies a bidding strategy that satisfies the assumption of Lemma 6 with $\lambda(x) = x$ and $H_n(t) = \left(1 + \frac{\chi F_n(t)^{n-1}}{1-\chi\left(1-F_n(t)^{n-1}\right)}\right)$ if $\chi \le 1$. Now suppose $\chi = 1$, which implies that $\tilde{H}(t) = 2$ and $\mathbf{h}(\alpha) = 1/2$. Because $\tilde{H}(t) = 2$ satisfies the model restrictions and

$$\xi_n(\alpha) = b_n(\alpha) + \frac{1/2}{n-1} \frac{\alpha}{g_n(b_n(\alpha))}$$

is strictly increasing in α by assumption, G_n can be generated by the Loss Aversion model by Lemma 6.

(5) In a Loser Regret model, the equilibrium strategy is defined by $s_n(0) = 0$ and

$$s_n'(t) = (n-1)\left[t - s_n(t) + h_n(t - s_n(t))\right] \frac{f_n(t)}{F_n(t)}, \tag{30}$$

where the regret function h_n satisfies $h_n(0) = 0$ and $h_n' \ge 0$. Define $\tilde{\lambda}(x) = \frac{1+c}{c} x$ for some $c > 0$ and $\mathbf{h}(\alpha) = 1$. Obviously, the resulting $\xi_n(\alpha)$ is increasing in α. Then we only need to check that $\tilde{\lambda}$ and $\mathbf{h}$ is consistent with the Loser Regret model. Notice that $\mathbf{h}$ leads to $\tilde{H} = 1$ and $\tilde{\lambda}(x)$ is consistent with $h_n(x) = x/c$. Therefore, G_n is consistent with a Loser Regret model.

(6) Lastly, consider the Level-k model. In this case G_n is consistent with a model where all bidders are level-1 who best respond to level-0 bidders whose bids are drawn from a uniform distribution. Specifically, the bidding strategy of the bidders is $s_n(t) = \frac{(n-1)t}{n\bar{t}}$ where $\bar{t}$ is the highest valuation. Then G_n is consistent with $\tilde{F}_n(t) = G_n\left(\frac{(n-1)t}{n\bar{t}}\right)$, where $\bar{t} = \frac{n}{n-1}\bar{b}$ and $\bar{b}$ is the upper bound of the support of g_n.

$\square$

Proof of Theorem 2. We first show that if G_n can be rationalized by a risk aversion model with valuation distribution F and $\lambda_n'(0) = 1$ for all n, then it can also be rationalized by an MEU model. We only need to show that there exists $F_n^{\min}$ that can generate the same bidding function as the Risk Aversion model with λ_n. Then $F_n^{\min}$ and F also generate the bid distribution. The bidding strategy under F and λ_n satisfies $s_n(0) = 0$ and

$$s_n'(t) = (n-1)\lambda_n[t - s_n(t)]\frac{f(t)}{F(t)} = (n-1)[t - s_n(t)]\frac{\lambda_n[t - s_n(t)]}{t - s_n(t)}\frac{f(t)}{F(t)}.$$

Define

$$\frac{f_n^{\min}(t)}{F_n^{\min}(t)} = \frac{\lambda_n[t - s_n(t)]}{t - s_n(t)}\frac{f(t)}{F(t)}.$$

Then s_n satisfies the FOC of the MEU model with $F_n^{\min}$ and

$$F_n^{\min}(t) = \exp\left\{-\int_t^{\bar{t}}\frac{\lambda_n[t - s_n(t)]}{t - s_n(t)}\frac{f(t)}{F(t)}dt\right\},$$

where $\bar{t}$ is the upper bound of the support of f. Notice that $F_n^{\min}$ is a valid distribution function because it is strictly increasing and if $t \to 0$, $F_n^{\min}(t)$ converges to 0. Also notice that by Guerre, Perrigne, and Vuong (2009) $s_n'(0) = (n-1)/n$ in both cases. Therefore, both models generate the same bid density at 0, and they both rationalize G_n. Lastly, because $\lambda_n' \geq 1$ and $\lambda_n(0) = 0$ in a Risk Aversion model, $\lambda_n[t - s_n(t)] \geq t - s_n(t)$. Therefore,

$$\frac{f_n^{\min}(t)}{F_n^{\min}(t)} \geq \frac{f(t)}{F(t)},$$

which means that $F_n^{\min}$ dominates F in reverse hazard rate. Hence, $F_n^{\min}$ first-order stochastically dominates F. There then exists a prior set of beliefs that contains F and has a lower envelope $F_n^{\min}$. Hence, the bid distribution can also be rationalized by $F_n^{\min}$ and F under a MEU model.

Although the above proof suffices to prove the theorem, we also show that if G_n are generated by a Loss Aversion model with $\chi_n < 1$ and with a strictly increasing bid shading for each n, they can be rationalized by a Risk Aversion model. Recall that the bidding strategy satisfies $s_n(0) = 0$ and (24) with $H_n(t) = \left(1 + \frac{\chi_n F_n(t)^{n-1}}{1-\chi_n\left(1-F_n(t)^{n-1}\right)}\right)$ for all n. Then we can define

$$\lambda_n [t - s_n(t)] = [t - s_n(t)] H_n(t).$$

By definition, $H_n(t) \geq 1$ and $H_n'(t) > 0$ and $\lambda_n(x) = xH_n[y_n^{-1}(x)]$ where $y_n(t) = t - s_n(t)$ is increasing in t by assumption. Obviously, $\lambda(0) = 0$, $\lambda'(0) = 1$ and

$$\lambda'(x) = H_n\left[y_n^{-1}(x)\right] + xH_n'\left[y_n^{-1}(x)\right]\frac{1}{y_n'\left[y_n^{-1}(x)\right]} > H_n\left[y_n^{-1}(x)\right] > 1.$$

Hence, we can conclude F and λ_n rationalize G_n. One can also prove the following, which we omit in the interest of space.

(1) If G_n are generated by a CEU model, then they can also be rationalized by an MEU model.

(2) If G_n are generated by a Loss Aversion model, then they can be rationalized by an MEU model.

(3) If G_n are generated by an MEU model where bid shadings are increasing in valuations for each n, and $\frac{F_n(t)}{f_n(t)}\frac{f_n^{\min}(t)}{F_n^{\min}(t)} \geq 1$ is increasing in t, then they can be rationalized by a Risk Aversion model, or a CEU model.

□

A.2 Proof of Results in Section 3.1

We first prove Proposition 2 because it will be handy in proving Proposition 1.

Proof of Proposition 2. Notice that $\bar{v}_n^a(0) = \bar{v}_n(0) = b_n(0)$. Therefore, the statement holds for $\alpha = 0$.

Now consider $\alpha > 0$. First we show that $\bar{v}_n(\alpha) \leq \bar{v}_n^a(\alpha)$. If this is not the case there exists some q that satisfies overbidding and rationality such that $q(\alpha) > \bar{v}_n^a(\alpha)$. Then $q_{\alpha,q(\alpha)}(\cdot)$ satisfies rationality by construction and RNBNE overbidding because $b_n(\cdot) \geq \beta_n(\cdot, q) \geq \beta_n\left(\cdot, q_{\alpha,q(\alpha)}\right)$. By the definition of $\bar{v}_n^a$, $\bar{v}_n^a(\alpha) \geq q_{\alpha,q(\alpha)}(\alpha) = q(\alpha)$, which contradicts $\bar{v}_n^a(\alpha) < q(\alpha)$.

Next, we show that $\bar{v}_n(\alpha) \geq \bar{v}_n^a(\alpha)$. Define

$$q_{\alpha,\epsilon,\delta}(t) = \begin{cases} b_n(t) & \text{if } t < \alpha - \epsilon \\ \frac{\bar{v}_n^a(\alpha) - \delta - b_n(\alpha - \epsilon)}{\epsilon}(t - \alpha + \epsilon) + b_n(\alpha - \epsilon) & \text{if } t \in [\alpha - \epsilon, \alpha], \\ \max\{\bar{v}_n^a(\alpha) - \delta, b_n(t)\} & \text{if } t \geq \alpha \end{cases} \tag{31}$$

where ϵ and δ are two positive numbers. Notice that $q_{\alpha,\epsilon,\delta}$ is weakly increasing. We will show that for every sufficiently small δ, there exists $\epsilon > 0$ such that $\beta(\cdot, q_{\alpha,\epsilon,\delta}) \leq b_n(\cdot) \leq q_{\alpha,\epsilon,\delta}(\cdot)$ and $q_{\alpha,\epsilon,\delta}$ is continuous. This implies that $\bar{v}_n(\alpha) \geq q_{\alpha,\epsilon,\delta}(\alpha) = \bar{v}_n^a(\alpha) - \delta$ and hence $\bar{v}_n(\alpha) = \bar{v}_n^a(\alpha)$ as δ can be arbitrarily small.

We first show that if $\delta > 0$ is sufficiently small, there exists $\epsilon > 0$ such that $b_n(\cdot) \leq q_{\alpha,\epsilon,\delta}(\cdot)$. Notice that $\beta_n\left(t, q_{\alpha,b_n(\alpha)+x}\right)$ equals $\beta_n(t, b_n)$ if $t < \alpha$ and it converges uniformly to $\beta_n\left(\cdot, q_{\alpha,b_n(\alpha)}\right) = \beta_n(\cdot, b_n)$ on $[\alpha, 1]$. Moreover, for all $t \geq \alpha$

$$\begin{aligned} \beta_n(t, b_n) &= \frac{1}{t^{n-1}} \int_0^t b_n(s)\, ds^{n-1} \\ &= b_n(t) - \frac{1}{t^{n-1}} \int_0^t s^{n-1} db_n(s) < b_n(t) - \int_0^\alpha s^{n-1} db_n(s). \end{aligned}$$

Because $b_n(\cdot)$ is strictly increasing, $\int_0^\alpha s^{n-1} db_n(s) > 0$ and

$$\min_{t \geq \alpha} [b_n(t) - \beta_n(t, b_n)] > 0. \tag{32}$$

This inequality holds for any $\alpha > 0$. Therefore, if x is sufficiently small $\beta_n\left(\cdot, q_{\alpha,b_n(\alpha)+x}\right)$ lies below $b_n(\cdot)$, which implies that $\bar{v}_n^a(\alpha) > b_n(\alpha) + x > b_n(\alpha)$ by the definition of $\bar{v}_n^a(\alpha)$. Therefore, we can find $\bar{\delta} > 0$ such that $\bar{v}_n^a(\alpha) - \bar{\delta} > b_n(\alpha)$. Because g_n is bounded from 0 in its support, $|b_n'(t)| < M$ for all $t \in [0,1]$ for some $M > 0$. Let $\bar{\epsilon} = \left[\bar{v}_n^a(\alpha) - \bar{\delta} - b_n(\alpha)\right]/M$, $0 < \epsilon < \bar{\epsilon}$ and $0 < \delta < \bar{\delta}$. If $t \in [\alpha - \epsilon, \alpha]$,

$$q_{\alpha,\epsilon,\delta}(t) > M(t - \alpha + \epsilon) + b_n(\alpha - \epsilon) \geq \int_{\alpha-\epsilon}^t b_n'(s)\, ds + b_n(\alpha - \epsilon) = b_n(t).$$

Moreover, $q_{\alpha,\epsilon,\delta}(t) \geq b_n(t)$ for all $t < \alpha - \epsilon$ and $t \geq \alpha$ by definition. Therefore, for any $\delta < \bar{\delta}$, $b_n(\cdot) \leq q_{\alpha,\epsilon,\delta}(\cdot)$ if $\epsilon < \bar{\epsilon}$. Also notice that because $\bar{v}_n^a(\alpha) - \delta > b_n(\alpha)$, $q_{\alpha,\epsilon,\delta}(\cdot)$ is continuous.

Next, we show that for every δ sufficiently small, there exists $\epsilon > 0$ such that $\beta(\cdot, q_{\alpha,\epsilon,\delta}) \leq b_n(\cdot)$. First notice that for any $\alpha > \underline{\alpha} > 0$,

$$\sup_{t \geq \underline{\alpha}} \left|\beta(t, q_{\alpha,\epsilon,\delta}) - \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)\right| \leq \frac{1}{\underline{\alpha}^{n-1}} \int_0^1 \left|q_{\alpha,\epsilon,\delta}(s) - q_{\alpha,\bar{v}_n^a(\alpha)-\delta}(s)\right| ds^{n-1}.$$

The right hand side converges to 0 as $\epsilon \to 0$ by the dominated convergence theorem because $q_{\alpha,\epsilon,\delta}(\cdot)$ converges to $q_{\alpha,\bar{v}_n^a(\alpha)-\delta}(\cdot)$ almost surely and $q_{\alpha,\epsilon,\delta}$, $q_{\alpha,\bar{v}_n^a(\alpha)-\delta}(\cdot)$ are bounded. This implies that $\beta(\cdot, q_{\alpha,\epsilon,\delta})$ converges uniformly to $\beta\left(\cdot, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)$ as $\epsilon \to 0$ on $[\underline{\alpha}, 1]$. We now show that

$$\min_{t \geq \underline{\alpha}} \left[b_n(t) - \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)\right] > 0. \tag{33}$$

First notice that $\beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right) = \beta_n(t, b_n)$ if $\underline{\alpha} \leq t \leq \alpha$. By (32),

$$\min_{\alpha \geq t \geq \underline{\alpha}} \left[b_n(t) - \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)\right] \geq \min_{t \geq \underline{\alpha}} \left[b_n(t) - \beta(t, b_n)\right] > 0.$$

We now show $\min_{t \geq \alpha}\left[b_n(t) - \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)\right] > 0$. Suppose towards contradiction, $b_n(t) = \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)$ for some $t > \alpha$. Then $\beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta/2}\right) > \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right) > b_n(t)$, which implies that $q_{\alpha,\bar{v}_n^a(\alpha)-\delta/2}$ violates RNBNE overbidding. However, this contradicts the definition of $\bar{v}_n^a(\alpha)$. Therefore, $\min_{t \geq \alpha}\left|b_n(t) - \beta\left(t, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)\right| > 0$ and (33) holds. Because $\beta(\cdot, q_{\alpha,\epsilon,\delta})$ converges uniformly to $\beta\left(\cdot, q_{\alpha,\bar{v}_n^a(\alpha)-\delta}\right)$ as $\epsilon \to 0$ on $[\underline{\alpha}, 1]$, $\beta(\cdot, q_{\alpha,\epsilon,\delta}) < b_n(\cdot)$ on $[\underline{\alpha}, 1]$ if ϵ is sufficiently small. Moreover, $\beta(\cdot, q_{\alpha,\epsilon,\delta}) = \beta_n(\cdot, b_n) \leq b_n(\cdot)$ on $[0, \underline{\alpha}]$, which implies $\beta(\cdot, q_{\alpha,\epsilon,\delta}) \leq b_n(\cdot)$. This concludes the proof. $\square$

Proof of Proposition 1. First, to see that $v_n(\alpha)$ is interval identified, just notice that all the quantile functions that satisfy overbidding and rationality form a convex set. This implies that if q_1 and q_2 satisfies overbidding and rationality, for any $\epsilon \in (0,1)$, we can find a q_3 that satisfies overbidding, rationality and $\epsilon q_1(\alpha) + (1-\epsilon) q_2(\alpha) = q_3(\alpha)$. Therefore, the identified set of $v_n(\alpha)$ is convex and hence an interval. By definition, $\bar{v}_n(\alpha)$ is an upper bound for the true valuation quantile. Now we show that it is a sharp bound. Given α, by the definition of $\bar{v}_n$, for any $\eta > 0$, there exists a $\tilde{q}$ that satisfies all the constraints and $\tilde{q}(\alpha) > \bar{v}_n(\alpha) - \eta/2$. Define $\tilde{\tilde{q}} = (1-\delta)\tilde{q} + \delta b_n$ for some $\delta \in (0,1)$. By assumption, $\tilde{\tilde{q}} \geq b_n$ and it is strictly increasing. Because $\tilde{\tilde{q}} < \tilde{q}$, by Lemma 7, the RNBNE bid quantile function of $\tilde{\tilde{q}}$ is weakly lower than that of $\tilde{q}$ and hence weakly lower than b_n. This means that $\tilde{\tilde{q}}$ is a valid candidate quantile function for any $\delta \in (0,1)$. Then pick a sufficiently small δ such that $\tilde{\tilde{q}}(\alpha) > \tilde{q}(\alpha) - \eta/2 > \bar{v}_n(\alpha) - \eta$. Therefore, $\bar{v}_n(\alpha) - \eta$ cannot be a valid upper bound for any $\eta > 0$. This shows that $\bar{v}_n(\alpha)$ is sharp. The lower bound is sharp because we cannot rule out the case where bidders have arbitrarily high risk aversion. This would induce them to bid arbitrarily close to their valuations.

Proposition 1(1) Next we show that $\bar{v}_n(\alpha) < b_n(1)/(1-\alpha)$ for any $\alpha < 1$. By the RNBNE constraint at 1, $\int_0^1 q(t)\,dt^{n-1} \leq b_n(1)$. As $q(t) \geq 0$

$$\int_\alpha^1 q(t)\,dt^{n-1} \leq \int_0^1 q(t)\,dt^{n-1} \leq b_n(1).$$

Because $q(t)$ is non-decreasing, $q(\alpha)(1-\alpha^{n-1}) \leq \int_\alpha^1 q(t)\,dt^{n-1}$. Therefore,

$$q(\alpha) \leq \frac{b_n(1)}{1-\alpha^{n-1}} < \frac{b_n(1)}{1-\alpha}.$$

The last inequality follows because $\alpha < 1$.

Proposition 1(2) By Proposition 2, this is equivalent to proving that $\exists C > 0$ such that $\bar{v}_n^a(\alpha) > C/(1-\alpha)$ for α close to 1. Define $2y = b_n(1) - \beta_n(1, b_n) > 0$. By continuity, there exists η sufficiently small such that $b_n(\alpha) - \beta_n(\alpha, b_n) > y$ for all $\alpha \geq 1-\eta$. Define $C = y/n$ and consider the quantile function $q_{\alpha,\frac{C}{1-\alpha}}$ if $\alpha > 1-\eta$. Notice that if η is sufficiently close to 0, $\frac{C}{1-\alpha} > b_n(1)$ for all $\alpha \geq 1-\eta$. Then

$$q_{\alpha,\frac{C}{1-\alpha}}(t) = \begin{cases} b_n(t) & \text{if } t < \alpha \\ \frac{C}{1-\alpha} & \text{if } t \geq \alpha \end{cases}.$$

Therefore,

$$\beta_n\left(1, q_{\alpha,\frac{C}{1-\alpha}}\right) - \beta_n\left(\alpha, q_{\alpha,\frac{C}{1-\alpha}}\right) \leq \frac{1}{\alpha^{n-1}} \int_\alpha^1 q_{\alpha,\frac{C}{1-\alpha}}(t)\,dt^{n-1} \leq \frac{C}{(1-\alpha)\alpha^{n-1}}\left(1-\alpha^{n-1}\right). \quad (34)$$

Since $\frac{1}{(1-\alpha)\alpha^{n-1}}(1-\alpha^{n-1})$ converges to $n-1$ as $\alpha \to 1$, it is less than n for all $\alpha > 1-\eta$ if η is sufficiently small. Because $C = y/n$, if $\alpha \geq 1-\eta$ and η sufficiently small

$$\beta_n\left(1, q_{\alpha,\frac{C}{1-\alpha}}\right) - \beta_n\left(1-\alpha, q_{\alpha,\frac{C}{1-\alpha}}\right) < \frac{y}{n}\frac{1-\alpha^{n-1}}{(1-\alpha)\alpha^{n-1}} < y.$$

Moreover, because $q_{\alpha,\frac{C}{1-\alpha}}(t) = b_n(t)$ for all $t < \alpha$, $\beta_n\left(t, q_{\alpha,\frac{C}{1-\alpha}}\right) = \beta_n(t, b_n)$ if $t \leq \alpha$. For every $t \geq \alpha \geq 1-\eta$,

$$\begin{aligned} \beta_n\left(t, q_{\alpha,\frac{C}{1-\alpha}}\right) \leq& \beta_n\left(1, q_{\alpha,\frac{C}{1-\alpha}}\right) - \beta_n\left(\alpha, q_{\alpha,\frac{C}{1-\alpha}}\right) + \beta_n\left(\alpha, q_{\alpha,\frac{C}{1-\alpha}}\right) \\ \leq& y + \beta_n\left(\alpha, q_{\alpha,\frac{C}{1-\alpha}}\right) = y + \beta_n(\alpha, b_n) \leq b_n(t). \end{aligned} \quad (35)$$

This proves that $q_{\alpha,\frac{C}{1-\alpha}}$ satisfies RNBNE overbidding. As a result, $\bar{v}_n^a(\alpha) \geq q_{\alpha,\frac{C}{1-\alpha}}(\alpha) = C/(1-\alpha)$, which proves Proposition 1(2).

Proposition 1(3) First, we prove that $\bar{v}_n(\cdot)$ is strictly increasing. Notice that it is weakly increasing because it is the supremum of a set of increasing functions. Suppose towards contradiction that $\bar{v}_n(\cdot)$ is not strictly increasing. There must exist $0 = \alpha_1 < \alpha_2 < 1$ such that $\bar{v}_n(\alpha_1) = \bar{v}_n(\alpha_2) = x$. By definition, $x \geq b_n(\alpha_2) > b_n(\alpha_1)$ and $q_{\alpha_1,x}(t) > q_{\alpha_2,x}(t)$ if $t \in (\alpha_1, \alpha_2)$ because $b_n(\cdot)$ satisfies overbidding and rationality and is strictly increasing. By Proposition 2, both $q_{\alpha_1,x}$ and $q_{\alpha_2,x}$ satisfies RNBNE overbidding and rationality. Notice that $b_n(t) \geq \beta_n(t, q_{\alpha_1,x}) > \beta_n(t, q_{\alpha_2,x})$ if $t \in [\alpha_2, 1]$. Because both b_n and $\beta(\cdot, q_{\alpha_2,x})$ are continuous, there exists $\epsilon > 0$ such that $b_n(t) > \beta_n(t, q_{\alpha_2,x}) + \epsilon$ if $t \in [\alpha_2, 1]$. Then there exists $\delta > 0$ such that $\beta_n(t, q_{\alpha_2,x+\delta}) \leq b_n(t)$ if $t \in [\alpha_2, 1]$ and $\beta_n(t, q_{\alpha_2,x+\delta}) = \beta_n(t, b_n) \leq b_n(t)$. This implies that $q_{\alpha_2,x+\delta}$ satisfies overbidding and rationality, which contradicts the assumption that $\bar{v}_n(\alpha_1) = \bar{v}_n(\alpha_2) = x$. Therefore, $\bar{v}_n(\cdot)$ is strictly increasing.

To see that $\bar{v}_n$ is continuous, notice that for any $\epsilon > 0$, we can find a q that satisfies all the constraints and $q(\alpha) > \bar{v}_n(\alpha) - \epsilon/2$ by the definition of supreme. Because q is continuous, we can find a $\delta > 0$ such that $|q(x) - q(\alpha)| < \epsilon/2$ for all $x \in (\alpha - \delta, \alpha]$. This means that $\bar{v}_n(x) - \bar{v}_n(\alpha) > q(x) - q(\alpha) - \epsilon/2 > -\epsilon$, which in turn implies that $|\bar{v}_n(x) - \bar{v}_n(\alpha)| < \epsilon$ for all $x \in (\alpha - \delta, \alpha]$ because $\bar{v}_n(x)$ is weakly increasing. This shows that $\bar{v}_n(\cdot)$ is left continuous.

Now we prove $\bar{v}_n(\cdot)$ is right continuous at any $\alpha \in (0,1)$. Let $\{\alpha_k\}_{k=1}^{\infty}$ be a decreasing sequence that converges to α. Let $x = \lim_{k\to\infty} \bar{v}_n(\alpha_k)$. Because $\bar{v}_n(\cdot)$ is weakly increasing, $\bar{v}_n(\alpha) \leq x$. Notice that $q_{\alpha,x}$ satisfies rationality by construction. Now we show that $q_{\alpha,x}$ satisfies RNBNE overbidding. First, as $k \to \infty$,

$$\sup_{t\in[\alpha,1]} \left|\beta_n(t, q_{\alpha,x}) - \beta_n\left(t, q_{\alpha_k,\bar{v}_n(\alpha_k)}\right)\right| \leq \frac{1}{\alpha^{n-1}} \int_0^1 \left|q_{\alpha,x}(t) - q_{\alpha_k,\bar{v}_n(\alpha_k)}(t)\right| dt^{n-1} \to 0.$$

As a result, $\beta_n(t, q_{\alpha,x}) \leq b_n(t)$ for all $t \in [\alpha, 1]$ because by Proposition 2, $q_{\alpha_k,\bar{v}_n(\alpha_k)}(t) \leq b_n(t)$ for all $t \in [0,1]$. And by construction, $q_{\alpha,x}$ satisfies RNBNE overbidding on $[0,\alpha]$. Therefore, $\bar{v}_n(\alpha) \geq q_{\alpha,x}(\alpha) = x$. This proves $\bar{v}_n(\alpha)$ is right continuous if $\alpha \in (0,1)$. Now consider the case where $\alpha = 0$. Suppose that $\lim_{k\to\infty} \bar{v}_n(\alpha_k) = x > b_n(0)$. Then for any $\epsilon > 0$, by RNBNE overbidding,

$$b_n(\epsilon) \geq \lim_{k\to\infty} \beta\left(\epsilon, q_{\alpha_k,\bar{v}_n(\alpha_k)}\right) \geq \lim_{k\to\infty} \beta\left(\epsilon, q_{\alpha_k,x/2}\right) \geq x/2 > b_n(0),$$

which contradicts the continuity of b_n. In addition, $\lim_{k\to\infty} \bar{v}_n(\alpha_k) \geq b_n(0)$ by rationality. As a result, $\lim_{k\to\infty} \bar{v}_n(\alpha_k) = b_n(0) = \bar{v}_n(0)$. □

A.3 Proof of Results in Section 4

Here we prove Proposition 6, but relegate the proof of Lemma 3 to Section C.

Proof of Proposition 6. We only need to show that $\bar{\pi}_n(r)$ defined by (18) is the sharp upper bound for profit. We start with showing $s_n(t,r) \geq s_n(t,0)$ under Assumption 4, which implies that $\partial s_n(t,r)/\partial t \leq \partial s_n(t,0)/\partial t$. Suppose toward contradiction there exists $t > r$ such that $s_n(t,r) < s_n(t,0)$. Because bidders never bid more than their valuations by rationality, $s_n(r,r) = r \geq s_n(r,0)$. There must then exist an interval $[x, x+\epsilon]$ such that $s_n(x,r) = s_n(x,0)$ and $s_n(y,r) < s_n(y,0)$ for all $y \in (x, x+\epsilon]$. Then, there exists at least one $\tilde{y} \in [x, x+\epsilon]$ such that $\partial s_n(\tilde{y},r)/\partial t < \partial s_n(\tilde{y},0)/\partial t$, because otherwise $s_n(y,r) \geq s_n(y,0)$ for all $y \in (x, x+\epsilon]$. By Assumption 4, $\partial s_n(\tilde{y},r)/\partial t < \partial s_n(\tilde{y},0)/\partial t$ implies $s_n(\tilde{y},r) > s_n(\tilde{y},0)$ because $\lambda_n(\cdot)$ is weakly increasing. Then we reach a contradiction. Therefore, $\bar{\rho}_n(\alpha,q,r)$ is a valid upper bound for the bid quantile function under q and r and (18) is a valid upper bound on profit.

To see that it is the sharp bound, we only need to show $\bar{\rho}_n(\alpha,q,r)$ is the sharp upper bound for the bid quantile function under q and r if q can rationalize b_n under rationality, overbidding, unimodality, assumptions on variation in n and Assumption 4. Given such a q, let $\lambda_n(x)$ be δ if $x > \delta$ and linear between 0 and δ. One can back out H_n that rationalizes b_n with q and λ_n accordingly. Under this λ_n and H_n, $\partial s_n(t,r)/\partial t = \partial s_n(t,0)/\partial t$ if $t - s_n(t,r) > \delta$, which implies that $\rho_n'(\alpha,q,\lambda_n,r) = b_n'(\alpha)$. Otherwise, $\rho_n(\alpha,q,\lambda_n,r)$ is bigger than $q(\alpha) - \delta$ and no more than $q(\alpha)$ because $\lambda_n(0) = 0$. As δ converges to 0, $\rho_n(\alpha,q,\lambda_n,r)$ becomes arbitrarily close to $\bar{\rho}_n(\alpha,q,r)$, implying that $\bar{\rho}_n(\alpha,q,r)$ is the sharp upper bound for the bid quantile function under q and r. Therefore, $\bar{\pi}_n(r)$ defined by (18) is the sharp upper bound for revenue. □

Proof of Lemma 4. Because $q_1(\alpha) \geq q_2(\alpha)$ for all $\alpha \in [0,1]$, $q_1^{-1}(r) \leq q_2^{-1}(r)$. Therefore, $\bar{\rho}_n(\alpha,q_1,r) \geq \bar{\rho}_n(\alpha,q_2,r)$ for all $\alpha \leq q_2^{-1}(r)$. Now suppose toward contradiction, $\bar{\rho}_n(\alpha,q_1,r) < \bar{\rho}_n(\alpha,q_2,r)$ for some $\alpha > q_2^{-1}(r)$. Because $\bar{\rho}_n\left(q_2^{-1}(r), q_1, r\right) \geq \bar{\rho}_n\left(q_2^{-1}(r), q_2, r\right)$, there must exits $\bar{\alpha} \in \left[q_2^{-1}(r), \alpha\right]$ such that $\bar{\rho}_n(\bar{\alpha},q_1,r) = \bar{\rho}_n(\bar{\alpha},q_2,r)$ and $\bar{\rho}_n(\cdot,q_1,r) < \bar{\rho}_n(\cdot,q_2,r)$ on $(\bar{\alpha}, \bar{\alpha}+\epsilon)$ for some $\epsilon > 0$. This implies that there exists $\tilde{\alpha} \in [\bar{\alpha}, \bar{\alpha}+\epsilon)$ such that $\partial\bar{\rho}_n(\tilde{\alpha},q_1,r)/\partial\alpha < \partial\bar{\rho}_n(\tilde{\alpha},q_2,r)/\partial\alpha$. If $\tilde{\alpha} > \bar{\alpha}$, then $\bar{\rho}_n(\tilde{\alpha},q_1,r) < \bar{\rho}_n(\tilde{\alpha},q_2,r)$. Because $q_1(\tilde{\alpha}) \geq q_2(\tilde{\alpha})$, this implies $\partial\bar{\rho}_n(\tilde{\alpha},q_1,r)/\partial\alpha = b_n'(\alpha) \geq \partial\bar{\rho}_n(\tilde{\alpha},q_2,r)/\partial\alpha$. Therefore, it must be $\tilde{\alpha} = \bar{\alpha}$, $\bar{\rho}_n(\tilde{\alpha},q_1,r) = \bar{\rho}_n(\tilde{\alpha},q_2,r) = q_1(\tilde{\alpha}) = q_2(\tilde{\alpha})$ and $q_1'(\tilde{\alpha}) < q_2'(\tilde{\alpha}) \leq b_n'(\alpha)$. However, this contradicts the assumption that $q_1(\alpha) \geq q_2(\alpha)$ for all $\alpha \in [0,1]$. Therefore, $\bar{\rho}_n(\alpha,q_1,r) \geq \bar{\rho}_n(\alpha,q_2,r)$ for all $\alpha > q_2^{-1}(r)$. □

A.4 Proof of Results in Section 6

Proof of Proposition 8. We first show that $\left|\hat{\bar{v}}_n(\alpha) - \bar{v}_n(\alpha)\right| \to^p 0$ if $\alpha \in [0,1]$. Because $\Theta \subseteq \mathcal{W}(B)$, $\bar{v}_n(\alpha)$ is defined and bounded by B for every $\alpha \in [0,1]$. By Assumption 5(1), $\hat{\bar{v}}_n(\cdot)$ is defined on $[0,1]$ with probability approaching 1 and $P\left(\hat{\bar{v}}_n(\alpha) \geq \bar{v}_n(\alpha)\right) \to 1$ as $N \to \infty$. We next show that $P\left(\hat{\bar{v}}_n(\alpha) \leq \bar{v}_n(\alpha) + \epsilon\right) \to 1$ for any $\epsilon > 0$. To see this, notice that by Assumption 5(2), as $N \to \infty$, there exists a $\tilde{q} \in \Theta_R(\mathbf{b})$ for every $q \in \hat{\Theta}_R$ such that $d_\infty(q, \tilde{q}) < \epsilon$ with probability approaching 1. This implies that $\hat{\bar{v}}_n(\alpha) \leq \bar{v}_n(\alpha) + \epsilon$ with probability approaching 1. Because ϵ can be arbitrarily small, $P\left(\hat{\bar{v}}_n(\alpha) \geq \bar{v}_n(\alpha)\right) \to 1$ and $P\left(\hat{\bar{v}}_n(\alpha) \leq \bar{v}_n(\alpha) + \epsilon\right) \to 1$ imply $\left|\hat{\bar{v}}_n(\alpha) - \bar{v}_n(\alpha)\right| \to^p 0$.

Because $\Theta \subseteq \mathcal{W}(B)$, both $\hat{\bar{v}}_n(\alpha)$ and $\bar{v}_n(\alpha)$ are increasing and continuous. Define $t_i = (i-1)/k$ for $i = 1, 2, \cdots, k+1$. If $\alpha \in (t_i, t_{i+1})$

$$\begin{aligned}
\left|\hat{\bar{v}}_n(\alpha) - \bar{v}_n(\alpha)\right| &\leq \left|\hat{\bar{v}}_n(\alpha) - \hat{\bar{v}}_n(t_i)\right| + \left|\hat{\bar{v}}_n(t_i) - \bar{v}_n(t_i)\right| + \left|\bar{v}_n(t_i) - \bar{v}_n(\alpha)\right| \\
&\leq \left|\hat{\bar{v}}_n(t_{i+1}) - \hat{\bar{v}}_n(t_i)\right| + \left|\hat{\bar{v}}_n(t_i) - \bar{v}_n(t_i)\right| + \left|\bar{v}_n(t_i) - \bar{v}_n(t_{i+1})\right| \\
&\leq \left|\hat{\bar{v}}_n(t_{i+1}) - \bar{v}_n(t_{i+1})\right| + 2\left|\hat{\bar{v}}_n(t_i) - \bar{v}_n(t_i)\right| + 2\left|\bar{v}_n(t_i) - \bar{v}_n(t_{i+1})\right|,
\end{aligned}$$

where the second inequality holds because $\hat{\bar{v}}_n(t)$ and $\bar{v}_n(t)$ are both increasing in t. Moreover, because $\bar{v}_n(t)$ is continuous on $[0,1]$, it is uniformly continuous on $[0,1]$. Therefore, for any $\kappa > 0$, $\sup_i \left|\bar{v}_n(t_i) - \bar{v}_n(t_{i+1})\right| < \kappa$ if k is sufficiently large. This implies that

$$\sup_{\alpha \in [0,1]} \left|\hat{\bar{v}}_n(\alpha) - \bar{v}_n(\alpha)\right| \leq 3 \sup_i \left|\hat{\bar{v}}_n(t_i) - \bar{v}_n(t_i)\right| + \kappa.$$

As $N \to \infty$, the right hand side converges to κ almost surely. Because κ is arbitrary,

$$\sup_{\alpha \in [0,1]} \left|\hat{\bar{v}}_n(\alpha) - \bar{v}_n(\alpha)\right| \to^p 0. \tag{36}$$

One can apply a similar argument to obtain $\sup_{\alpha \in [0,1]} \left|\hat{\underline{v}}_n(\alpha) - \underline{v}_n(\alpha)\right| \to^p 0$. If $\max_{m \in \mathbb{N}_1(n)} b_m(\alpha)$ has a unimodal density, $\underline{v}_n(\alpha) = \max_{m \in \mathbb{N}_1(n)} b_m(\alpha)$. Obviously, $\sup_{\alpha \in [0,1]} \left|\hat{\underline{v}}^a_n(\alpha) - \underline{v}_n(\alpha)\right| \to^p 0$. □

Proof of Lemma 5. Under this assumption, $\hat{b}_n$ converges to b_n uniformly at the rate $\sqrt{N}$. This implies that $\hat{b}_n(\cdot) - \eta_N \leq b_n(\cdot) \leq \hat{b}_n(\cdot) + \eta_N$ for all $n \in \mathbb{N}$ with probability approaching 1 as $N \to \infty$. As a result, if $q \in \Theta_R(\mathbf{b})$, then with probability approaching 1, $q \geq b_n \geq \hat{b}_n(\cdot) - \eta_N$ and $\beta_n(\cdot, q) \leq b_n(\cdot) \leq \hat{b}_n(\cdot) + \eta_N$ for all $n \in \mathbb{N}$. This implies that Assumption 5(1) is satisfied. We next prove Assumption 5(2) by deriving a contradiction. First notice that $\hat{b}_n$ converges to b_n uniformly almost surely by Assumption 6(1). Suppose towards contradiction Assumption 5(2) does not hold. This implies that for some data realization under which $\hat{b}_n$ converges to b_n for all $n \in \mathbb{N}$, there exists $\epsilon > 0$ and $q_N \in \hat{\Theta}_R$ such that $d_\infty (q_N, \Theta_R (\mathbf{b})) > \epsilon$. Because Θ is compact under $\|\cdot\|_{sob}$, there exists a subsequence q_{N_k} that converges under the norm $\|\cdot\|_{sob}$. Denote the limit as q^*. By assumption, q^* satisfies $b_n (\cdot) \leq q^* (\cdot), \beta_n (\cdot, q^*) \leq b_n (\cdot)$ for all $n \in \mathbb{N}$. Moreover, its density satisfies unimodality. To see this, notice that if otherwise, the density has at least two local minima. There then exist $\alpha_1 < \alpha_2$ such that for all sufficiently small $\epsilon > 0$, $q^{*\prime} (\alpha_1 - \epsilon) > q^{*\prime} (\alpha_1) > q^{*\prime} (\alpha_1 + \epsilon)$ and $q^{*\prime} (\alpha_2 - \epsilon) > q^{*\prime} (\alpha_2) > q^{*\prime} (\alpha_2 + \epsilon)$. Because q_{N_k} converges to q^* under $\|\cdot\|_{sob}$ and $M \geq 2$, q'_{N_k} converges to $q^{*\prime}$ uniformly. This implies that for some $0 < \epsilon < \frac{\alpha_2 - \alpha_1}{2}$ sufficiently small, there exists a sufficiently large k such that, $q'_{N_k} (\alpha_1 - \epsilon) > q'_{N_k} (\alpha_1) > q'_{N_k} (\alpha_1 + \epsilon)$ and $q'_{N_k} (\alpha_2 - \epsilon) > q'_{N_k} (\alpha_2) > q'_{N_k} (\alpha_2 + \epsilon)$. This contradicts the fact that q_{N_k} is unimodal. Therefore, q^* is unimodal. This implies that $q^* \in \Theta_R (\mathbf{b})$, which contradicts the assumption that $d_\infty (q^*, \Theta_R (\mathbf{b})) = \lim_{N \to \infty} d_\infty (q_N, \Theta_R (\mathbf{b})) > \epsilon$. Consequently, we conclude that for every realization such that $\hat{b}_n$ converges to b_n for all $n \in \mathbb{N}$, $\sup_{q \in \hat{\Theta}_R} d_\infty (q, \Theta_R (b)) \to 0$. Therefore, the latter holds almost surely. As a result, Assumption 5(2) holds. This proves the lemma. □

Proof of Proposition 9. Notice that $\hat{\tilde{\pi}}_n (r)$ is continuous in $\hat{\bar{v}}_n$ under the sup norm. By the continuous mapping theorem, $\hat{\tilde{\pi}}_n (r) \to^p \tilde{\pi}_n (r)$ because $\sup_{\alpha \in [0,1]} \left| \hat{\bar{v}}_n (\alpha) - \bar{v}_n (\alpha) \right| \to^p 0$. A similar argument applies to $\hat{\underset{\sim}{\pi}}_n (r)$. □

A.5 A Useful Lemma

Lemma 7. *Let $\sigma_n^l (\cdot, r)$ be the RNBNE bid quantile function under valuation quantile v^l and reserve price r in n bidder auctions. If $v^1 (\cdot) \geq v^2 (\cdot)$, then $\sigma_n^1 (\cdot, r) \geq \sigma_n^2 (\cdot, r)$.*

Proof. Under RNBNE, the bid quantile function satisfies

$$\frac{\partial \sigma_n^l (\alpha, r)}{\partial \alpha} = (n-1) \frac{v^l (\alpha) - \sigma_n^l (\alpha, r)}{\alpha},$$

with the initial condition $\sigma_n^l (\alpha, r) = r$ if $\sigma^l (\alpha) = r$. And if $v^l (\alpha) < r$, $\sigma_n^l (\alpha) = 0$. Let $v^1 (\alpha_1) = r$ and $v^2 (\alpha_2) = r$. Because $v^1 (\cdot) \geq v^2 (\cdot)$, $\alpha_1 \leq \alpha_2$. In addition, $\sigma_n^1 (\alpha, r) \geq r$ $\forall \alpha \geq \alpha_1$. Therefore, $\sigma_n^1 (\cdot, r) \geq \sigma_n^2 (\cdot, r)$ $\forall \alpha \in [0, \alpha_2)$. Next, we conclude the proof by showing that the statement holds $\forall \alpha \in [\alpha_2, 1]$ using a contradiction argument. Suppose there is some $\alpha_3 \in [\alpha_2, 1]$ such that $\sigma_n^1 (\alpha_3, r) < \sigma_n^2 (\alpha_3, r)$. Because $\sigma_n^1 (\alpha_2, r) \geq r = \sigma_n^2 (\alpha_2, r)$, by the continuity of the bid function, there must be some $\alpha_3 > \alpha_4 \geq \alpha_2$ such that $\sigma_n^1 (\alpha_4, r) = \sigma_n^2 (\alpha_4, r)$ and $\sigma_n^1 (\alpha, r) < \sigma_n^2 (\alpha, r)$ for all $\alpha \in (\alpha_4, \alpha_3]$. Then by the first order condition, $\frac{\partial}{\partial \alpha} \sigma_n^1 (\alpha, r) > \frac{\partial}{\partial \alpha} \sigma_n^2 (\alpha, r)$. This implies $\sigma_n^1 (\alpha_3, r) > \sigma_n^2 (\alpha_3, r)$, which is a contradiction. □

B Computation Methods (For Online Publication)

This section discusses how to compute the bounds proposed in the paper. We focus on the case where the population bid quantile functions are observed. For estimation, the computation method stays unchanged except that we replace the population bid quantile function by the estimated bid quantile function and relax the overbidding and rationality constraints by η_N as described in Section 6.

We start with the upper bound on the bound for the valuation quantile function. The idea is to approximate quantile functions with linear sieves. Let $P_K = (p_1, p_2, \cdots, p_K)$ be a vector of basis functions supported on $[0, 1]$. We recommend using B-splines. Then $q(\alpha) \approx P_K(\alpha)\theta$, where θ is a K-dimensional column vector. The constraints can be imposed on a set of points $0 = t_1 < t_2 < \cdots t_J = 1$. The upper bound for $v_n(\alpha)$ can be approximated by

$$\max_{\theta} P_K(\alpha)\theta, \text{ s.t. } P_K(t_j)\theta \geq b_n(t_j) \geq A_K(t_j)\theta, \; P_K(0)\theta \geq 0, \; P_K'(t_j)\theta \geq 0 \; \forall j, \tag{37}$$

where $A_K(\alpha) = \frac{1}{\alpha^{n-1}} \int_0^\alpha P_K(t)\, dt^{n-1}$ if $\alpha > 0$ and $A_K(0) = 0$. This is a linear programming problem that can be solved quickly even if K and J are both very large. If unimodality is imposed, we solve

$$\begin{aligned} & \max_{l \in \{1,2,\cdots,J\}} \max_{\theta} P_K(\alpha)\theta, \\ \text{s.t. } \; & P_K(t_j)\theta \geq b_n(t_j) \geq A_K(t_j)\theta, \; \forall 1 \leq j \leq J, \\ & P_K(0)\theta \geq 0, \; {P'}_K(0)\theta \geq 0, P_K'(1)\theta \geq 0 \\ & P_K'(t_j)\theta \geq P_K'(t_{j+1})\theta \; \forall j+1 < l, \\ & P_K'(t_j)\theta \leq P_K'(t_{j+1})\theta \; \forall j > l. \end{aligned} \tag{38}$$

For a given l, this is a linear programming problem. We solve this problem for J different values of l and then take the maximum. Because it is fast to solve each of the linear programming problems, the method is computationally manageable even if J is very large. For the lower bound, one can replace max in (38) by min. If b_n has a unimodal density, which is the case in our numerical analysis, $\underline{v}_n(\alpha) = b_n(\alpha)$ is the sharp lower bound. Bounds with variation in n are calculated in the same way by adding constraints from auctions with different number of bidders.

Sharp bounds on seller's profit can be approximated in a similar way. If unimodality is not impose, one can solve the following problem to obtain the sharp upper bound on profit under a reserve price r:

$$\begin{aligned} & \max_{\theta} \int_{P_K(\alpha)\theta \geq r} \left[\bar{\rho}_n\left(\alpha, P_K\left(\cdot\right)\theta, r\right) - c\right] d\alpha^n + c \qquad (39) \\ \text{s.t. } & P_K\left(t_j\right)\theta \geq b_n\left(t_j\right) \geq A_K\left(t_j\right)\theta,\ P_K\left(0\right)\theta \geq 0,\ P_K'\left(t_j\right)\theta \geq 0, \forall 1 \leq j \leq J \\ & \bar{\rho}_n\left(t_j, P_K\left(\cdot\right)\theta, P_K\left(t_m\right)\theta\right) \geq \beta_n\left(t_j, P_K\left(\cdot\right)\theta, P_K\left(t_m\right)\theta\right), \quad \forall j \geq m > 1. \end{aligned}$$

The first line of the constraints are the same as in (37), which imposes overbidding and rationality without a reserve price. The second line of the constraints imposes overbidding if the reserve price is $P_K(t_m)\theta$, where $m = 2, 3, \cdots, J$. If unimodality is imposed, we simply follow the approach for the upper bound of the valuation quantile function: first, calculate the highest profit if the valuation density peaks at t_l for all $1 \leq l \leq J$; second, take the maximum of all these profits to obtain the upper profit bound. If variation in n is impose, we only need to add additional constraints to the optimization problem.

The same approach can be used to construct the sharp lower profit bound. But one needs to solve a problem like (20) that involves both q and λ_n. We approximate both q and λ_n by B-splines and impose all the constraints on a grid as before. One additional challenge is that we need to compute the bidding strategy under (q, λ_n). Following Hubbard, Kirkegaard, and Paarsch (2013), we use Chebyshev collocation to obtain the bidding strategy. For estimation, we use the same computational method except that we replace the population bid quantile functions by their sample counterparts and relax some of the constraints, i.e., we replace the constraint $P_K(t_j)\theta \geq b_n(t_j) \geq A_K(t_j)\theta$ by $P_K(t_j)\theta + \eta_N \geq \hat{b}_n(t_j) \geq A_K(t_j)\theta - \eta_N$.

C Overbidding Models (For Online Publication)

We first introduce sufficient conditions for a bidding strategy $s_n(t)$ to satisfy RNBNE overbidding.

Lemma 8. *If $\exists \tilde{t} \in [0, \bar{t}_n]$ such that (1) $s_n(t) = t$ for $t \leq \tilde{t}$; (2)*

$$s_n'(t) = (n-1)\lambda_n(t - s_n(t)) H_n(t) f_n(t) / F_n(t)$$

for $t > \tilde{t}$, where $H_n(t) \geq 1$ and $\lambda_n(0) = 0$, $\lambda_n'(\cdot) \geq 1$, then $s_n(t) \geq \sigma_n(t)$ for all t.

Proof. If $t \leq \tilde{t}$, $\sigma_n(t) \leq t = s_n(t)$. Next, suppose $t > \tilde{t}$. Because $\lambda_n(x) \geq x$ and $H_n(t) \geq 1$, $s_n'(t) \geq \sigma_n'(t)$ if $s_n(t) = \sigma_n(t)$. This implies that s_n does not cross σ_n from the above. Moreover, $s_n(t) \geq \sigma_n(t)$ if $t = \tilde{t}$, which implies $s_n(t) \geq \sigma_n(t)$ for all t. □

Notice that models satisfying the conditions of Lemma 8 also satisfy Assumption 4, which we use to bound counterfactual bidding behavior under alternative reserve prices.

Next, we show that most of the alternative bidding models satisfy the conditions in Lemma 8 and therefore lead to overbidding. The exceptions are the Maxmin Expected Utility (MEU) model and the Level-k model. The MEU model does not necessarily generate RNBNE overbidding. However, an intuitively assumption on the prior set guarantees the model to generate RNBNE overbidding. The level-k model, on the other hand, is unique in the sense that it allows for heterogeneous bidders and does not involve a fixed point problem. We show that even if all bidders have the same level, there is no guarantee that they overbid or underbid compared to the RNBNE. Notice that the following analysis also serves as a proof of Lemma 3 and specifies λ_n and H_n for each of the models.

Risk Aversion Bidders share a utility function U_n with $U_n(0) = 0$, $U'_n(\cdot) > 0$ and $U''_n(\cdot) \leq 0$. The equilibrium strategy solves $s_n(t) = \arg\max_b U_n(t-b) F_n(s_n^{-1}(b))^{n-1}$, where s_n^{-1} is the inverse of s_n. The equilibrium bidding strategy is characterized by the boundary condition $s_n(0) = 0$ and

$$s'_n(t) = (n-1)\lambda_n [t - s_n(t)] \frac{f_n(t)}{F_n(t)}, \tag{40}$$

where $\lambda_n(\cdot) = U_n(\cdot)/U'_n(\cdot)$ satisfies $\lambda_n(0) = 0$ and $\lambda'_n(\cdot) \geq 1$. This model satisfies the assumptions in Lemma 8 with $\tilde{t} = 0$, λ_n, and $H_n(t) = 1$. Hence, $s_n(t) \geq \sigma_n(t)$ for all t.

Loser Regret Filiz-Ozbay and Ozbay (2007) propose loser regret as the explanation of the overbidding puzzle. As in the standard model, the revenue of a bidder with valuation t who wins the auction with bid b is $t - b$. However, if the bidder loses the auction to another bidder who bids $b^W < t$, she experiences regret and gets a revenue of $-h_n\left(t - b^W\right)$. The regret function h_n satisfies $h_n(0) = 0$ and $h'_n \geq 0$. The equilibrium strategy is defined by $s_n(0) = 0$ and

$$s'_n(t) = (n-1)\left[t - s_n(t) + h_n(t - s_n(t))\right] \frac{f_n(t)}{F_n(t)}. \tag{41}$$

The model satisfies the assumptions in Lemma 8 with $\tilde{t} = 0$, $\lambda_n(x) = x + h_n(x)$ and $H_n(t) = 1$. Hence, $s_n(t) \geq \sigma_n(t)$ for all t.

Choquet Expected Utility Salo and Weber (1995) apply the Choquet Expected Utility (CEU) model to first-price auctions. A CEU bidder chooses b to maximize $(t-b)\,\phi\left(F(s_n^{-1}(b))^{n-1}\right)$, where ϕ is convex and $\phi(0) = 0$. The equilibrium is defined by $s_n(0) = 0$ and

$$s_n'(t) = (n-1)[t - s_n(t)] \frac{\phi'\left(F_n(t)^{n-1}\right) F_n(t)^{n-1}}{\phi\left(F_n(t)^{n-1}\right)} \frac{f_n(t)}{F_n(t)}. \tag{42}$$

Now let $\tilde{t} = 0$, $\lambda_n(x) = x$, and $H_n(t) = \phi'\left(F_n(t)^{n-1}\right) F_n(t)^{n-1} / \phi\left(F_n(t)^{n-1}\right)$. As ϕ is convex and $\phi(0) = 0$, it follows that $H_n(t) \geq 1$. Hence, $s_n(t) \geq \sigma_n(t)$ for all t.

Loss Aversion Lange and Ratan (2010) show that loss aversion can also generate overbidding if bidders are loss averse in money.[20] The bidding strategy satisfies the following conditions: $\exists \tilde{t}_n \in [0, \bar{t}_n]$ such that if $t \leq \tilde{t}_n$, $s_n(t) = t$ and if $t > \tilde{t}_n$, $s_n(t)$ solves the differential equation

$$s_n'(t) = (n-1)[t - s_n(t)] \frac{f_n(t)}{F_n(t)} \left[1 + \frac{\chi F_n(t)^{n-1}}{1 - \chi\left(1 - F_n(t)^{n-1}\right)}\right]$$

with the initial condition $s_n\left(\tilde{t}_n\right) = \tilde{t}_n$. Here $\chi \geq 0$ is the loss aversion parameter. The assumptions in Lemma 8 are satisfied with $\lambda_n(x) = x$ and $H_n(t) = 1 + \frac{\chi F_n(t)^{n-1}}{1-\chi\left(1-F_n(t)^{n-1}\right)} \geq 1$.

Maxmin Expected Utility The MaxMin Expected Utility (MEU) model of ambiguity aversion, which was axiomatized by Gilboa and Schmeidler (1989), can also explain overbidding. It was introduced in the literature on first-price auction by Lo (1998). MEU bidders do not know the true valuation distribution F_n and consider all distributions in some prior set Δ_n plausible. They maximize $\min_{\widetilde{F} \in \Delta_n} (t - b) \widetilde{F}\left(s_n^{-1}(b)\right)^{n-1}$. Let $F_n^{\min}(t) = \inf_{\tilde{F} \in \Delta_n} \widetilde{F}(t)$ for all t be the lower contour of the prior set Δ_n, which is itself a distribution function, and let $f_n^{\min}$ be the corresponding density. If $F_n \in \Delta_n$ then $F_n^{\min}$ first-order stochastically dominates F_n. The equilibrium bidding strategy is $s_n(0) = 0$ and

$$s_n'(t) = (n-1)[t - s_n(t)] \frac{f_n(t)}{F_n(t)} \frac{F_n(t)}{f_n(t)} \frac{f_n^{\min}(t)}{F_n^{\min}(t)}. \tag{43}$$

Thus $H_n(t) = \frac{F_n(t)}{f_n(t)} \frac{f_n^{\min}(t)}{F_n^{\min}(t)}$. The fact that $F_n^{\min}$ first-order stochastically dominates $F_n(t)$ does not guarantee that $H_n(t) \geq 1$. As a result, this model in general does not satisfy the conditions of Lemma 8. However, if $F_n^{\min}$ dominates F_n in the reverse hazard rate, then $H_n(t) \geq 1$ and $s_n(t) \geq \sigma_n(t)$ for all t.

[20] See Corollary 2 in Lange and Ratan (2010). They also consider the case where bidders are loss averse in the auctioned good. In this case loss aversion does not generate overbidding.

Level-k In the level-k model, bidders with level $k+1$ best respond to the bidding behavior of bidders with level k. The bidding behavior of bidders with $k>0$ crucially depends on the bidding behavior of $k=0$, which is not determined by the model. We will show that depending on the choice for the bidding behavior of $k=0$, the model can either generate under- or overbidding. We first construct an example such that level-1 bidders follow the RNBNE strategy and then deviate from the example to generate under- or overbidding.

Denote the bid function of a bidder with level k by s_k^{lk}.[21] Consider the most commonly studied case in which level $k=0$ bidders are bidding "truthfully", i.e., $s_0^{lk}(v)=v$. Now consider the best response of bidder level $k=1$, who chooses b to maximize

$$G_0^{n-1}(b)(v-b)=F^{n-1}(b)(v-b).$$

As $s_0(v)=v$, $G_0=F$, i.e. the bid distribution of level $k=0$ bidders is equal to the valuation distribution. The first-order condition is

$$v-s_1^{lk}(v) \quad = \quad \frac{F\left(s_1^{lk}(v)\right)}{f\left(s_1^{lk}(v)\right)(n-1)}.$$

Now suppose F is the uniform distribution on $[0,1]$. Then we can solve the first-order condition for

$$s_1^{lk}(v) \quad = \quad \frac{n-1}{n}v,$$

which is equal to the RNBNE strategy in this case. Hence if F is uniform and $k=0$ bidders bid truthfully, then bidders with $k>0$ use the RNBNE strategy, so they neither underbid nor overbid.

Next, we show that deviations from the case where F is uniform will generate either underbidding or overbidding compared to RNBNE. Instead of assuming that f is constant, we assume that f is either linearly increasing or decreasing on $[0,1]$, i.e., $f(v)=1-c+2cv$ for some constant c. If $c=0$, then f is the uniform distribution; if $c<0$ then f is increasing; and for $c>0$, f is increasing. The cdf associated with f is $F=(1-c)v+cv^2$. For $c\neq 0$,

[21] Notice that in this section we use the subscript of the bidding function s to denote the level k of the bidder rather than the number of bidders. We also drop the subscript n for the valuation distribution F.

the first-order condition yields a quadratic equation that can be solved for

$$\begin{aligned} s_1^{lk}(v) &= \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_0} \\ a_0 &= 2c(n-1) + c \\ a_1 &= n(1-c) - 2(n-1)cv \\ a_2 &= 4(n-1)(1-c)v. \end{aligned}$$

In Figure 12 we plot the difference between $s_1^{lk}(v)$ and the RNBNE bidding strategy $\sigma_n(v)$. On the left-hand side we can see that $c = -0.3 < 0$ leads to overbidding, whereas $c = 0.3 > 0$ on the right-hand side leads to underbidding.

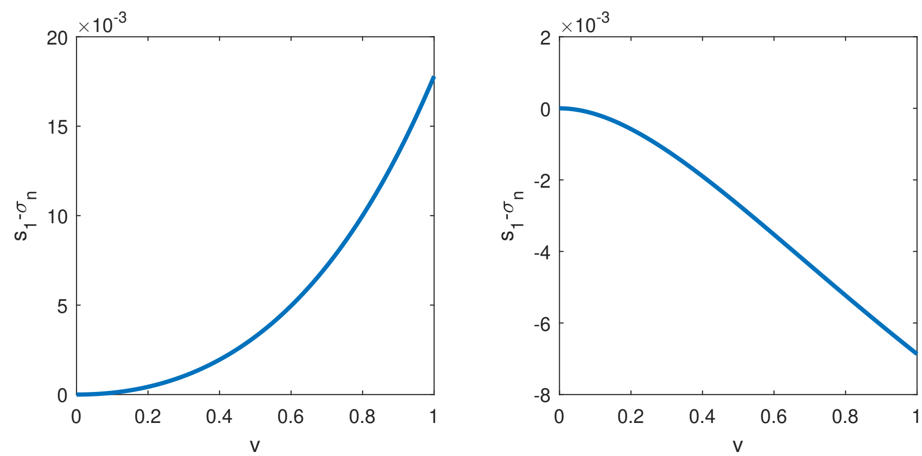


Figure 12: This graph shows $s_1^{lk}(v) - \sigma_n(v)$ if level-0 bidders bid truthfully and $f(v) = 1 - c + 2cv$ for some constant c. The graph on the left shows overbidding for $c = -0.3 < 0$ on the graph on the right shows underbidding for $c = 0.3 > 0$.

References

An, Y. (2017): "Identification of First-Price Auctions with Non-Equilibrium Beliefs: A Measurement Error Approach," *Journal of Econometrics*, 200(2), 326–343.

Aradillas-López, A., A. Gandhi, and D. Quint (2011): "Identification and Inference in Ascending Auctions with Correlated Private Values," *Econometrica*.

Aradillas-Lopez, A., and E. Tamer (2008): "The Identification Power of Equilibrium in Simple Games," *Journal of Business and Economic Statistics*, 26(3), 261–283.

Aryal, G., S. Grundl, D.-H. Kim, and Y. Zhu (2018): "Empirical Relevance of Ambiguity in First-Price Auctions," *Journal of Econometrics*, 204(2), 189–206.

Aryal, G., and D.-H. Kim (2013): "A point decision for partially identified auction models," *Journal of Business & Economic Statistics*, 31(4), 384–397.

Astor, P. J., M. T. Adam, C. Jähnig, and S. Seifert (2013): "The Joy of Winning and the Frustration of Losing: A Psychophysiological Analysis of Emotions in First-Price Sealed-Bid Auctions.," *Journal of Neuroscience, Psychology, and Economics*, 6(1), 14.

Athey, S., and J. Levin (2001): "Information and Competition in US Forest Service Timber Auctions," *Journal of Political Economy*, 109(2).

Bajari, P., and A. Hortacsu (2005): "Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data," *Journal of Political Economy*, 113(4), 703–741.

Burkart, M. (1995): "Initial Shareholdings and Overbidding in Takeover Contests," *The Journal of Finance*, 50(5), 1491–1515.

Campo, S. (2012): "Risk Aversion and Asymmetry in Procurement Auctions: Identification, Estimation and Application to Construction Procurements," *Journal of Econometrics*, 168(1), 96–107.

Campo, S., E. Guerre, I. Perrigne, and Q. Vuong (2011): "Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders," *The Review of Economic Studies*, 78(1), 112.

Chen, K.-Y., and C. R. Plott (1998): "Nonlinear behavior in sealed bid first price auctions," *Games and Economic Behavior*, 25(1), 34–78.

Coppinger, V. M., V. L. Smith, and J. A. Titus (1980): "Incentives and Behavior in English, Dutch and Sealed-Bid Auctions," *Economic inquiry*, 18(1), 1–22.

Cox, J. C., B. Roberson, and V. L. Smith (1982): "Theory and Behavior of Single Object Auctions," *Research in experimental economics*, 2(1), 1–43.

Cox, J. C., V. L. Smith, and J. M. Walker (1983a): "A test that discriminates between two models of the Dutch-first auction non-isomorphism," *Journal of Economic Behavior & Organization*, 4(2-3), 205–219.

——— (1983b): "Tests of a heterogeneous bidders theory of first price auctions," *Economics Letters*, 12(3-4), 207–212.

——— (1985): "Experimental development of sealed-bid auction theory; calibrating controls for risk aversion," *The American Economic Review*, 75(2), 160–165.

——— (1988): "Theory and Individual Behavior of First-Price Auctions," *Journal of Risk and Uncertainty*, 1(1), 61–99.

——— (1992): "Theory and Misbehavior of First-Price Auctions: Comment," *The American Economic Review*, pp. 1392–1412.

de Bodt, E., J.-G. Cousin, and R. Roll (2018): "Empirical Evidence of Overbidding in M&A Contests," *Journal of Financial and Quantitative Analysis*, 53(4), 1547–1579.

Dorsey, R., and L. Razzolini (2003): "Explaining overbidding in first price auctions using controlled lotteries," *Experimental Economics*, 6(2), 123–140.

Dufwenberg, M., and U. Gneezy (2002): "Information disclosure in auctions: an experiment," *Journal of Economic Behavior & Organization*, 48(4), 431–444.

Dyer, D., J. H. Kagel, and D. Levin (1989): "Resolving Uncertainty about The Number of Bidders in Independent Private-Value Auctions: an Experimental Analysis," *Rand Journal of Economics*, 20(2), 268–279.

Engelbrecht-Wiggans, R., and E. Katok (2007): "Regret in Auctions: Theory and Evidence," *Economic Theory*, 33(1), 81–101.

——— (2009): "A Direct Test of Risk Aversion and Regret in First Price Sealed-Bid Auctions," *Decision Analysis*, 6(2), 75–86.

Filiz-Ozbay, E., and E. Y. Ozbay (2007): "Auctions with anticipated regret: Theory and experiment," *The American Economic Review*, 97(4), 1407–1418.

Friedman, D. (1992): "Theory and Misbehavior of First-Price Auctions: Comment," *The American Economic Review*, 82(5), 1374–1378.

Füllbrunn, S., D.-J. Janssen, and U. Weitzel (2018): "Risk Aversion and Overbidding in First Price Sealed Bid Auctions: New Experimental Evidence," *Economic Inquiry*.

Gentry, M., T. Li, and J. Lu (2015): "Identification and Estimation in First-Price Auctions with Risk-Averse Bidders and Selective Entry," *Available at SSRN*.

Georganas, S., D. Levin, and P. McGee (2015): "A New Perspective on Overbidding in Private-Value Auctions," Discussion paper, Working Paper.

Gilboa, I., and D. Schmeidler (1989): "Maxmin Expected Utility with Non-Unique Prior," *Journal of mathematical economics*, 18(2), 141–153.

Gillen, B. (2009): "Identification and Estimation of Level-k Auctions," *Unpublished Manuscript*.

Gimenes, N., and E. Guerre (2016): "Quantile regression methods for first-price auction: A signal approach," *mimeo*.

Goeree, J. K., C. A. Holt, and T. R. Palfrey (2002): "Quantal response equilibrium and overbidding in private-value auctions," *Journal of Economic Theory*, 104(1), 247–272.

Grundl, S., and Y. Zhu (2016): "Inference of Risk Aversion in First-Price Auctions With Unobserved Auction Heterogeneity," Discussion paper.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

——— (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica*, 77(4), 1193–1227.

Haile, P., H. Hong, and M. Shum (2003): "Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions," .

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, pp. 1–51.

Harrison, G. W. (1989): "Theory and Misbehavior of First-Price Auctions," *The American Economic Review*, pp. 749–762.

——— (1990): "Risk Attitudes in First-Price Auction Experiments: A Bayesian Analysis," *The Review of Economics and Statistics*, pp. 541–546.

Harrison, G. W. (1992): "Theory and Misbehavior of First-Price Auctions: Reply," *The American Economic Review*, 82(5), 1426–1443.

Hubbard, T. P., R. Kirkegaard, and H. J. Paarsch (2013): "Using Economic Theory to Guide Numerical Analysis: Solving for Equilibria in Models of Asymmetric First-Price Auctions," *Computational Economics*, 42(2), 241–266.

Isaac, R. M., and D. James (2000): "Just Who Are You Calling Risk Averse?," *Journal of Risk and Uncertainty*, 20(2), 177–187.

Kagel, J. H. (1995): "Auctions: A Survey of Experimental Research," *The Handbook of Experimental Economics.*

Kagel, J. H., and D. Levin (1993): "Independent Private Value Auctions: Bidder Behaviour in First-, Second-and Third-price Auctions with Varying Numbers of Bidders," *The Economic Journal*, 103(419), 868–879.

Kagel, J. H., and D. Levin (2010): "Auctions: A Survey of Experimental Research, 1995-2010," *Handbook of Experimental Economics*, 2.

Kagel, J. H., and A. E. Roth (1992): "Theory and Misbehavior in First-Price Auctions: Comment," *The American Economic Review*, 82(5), 1379–1391.

Kirchkamp, O., and J. P. Reiss (2011): "Out-Of-Equilibrium Bids in First-Price Auctions: Wrong Expectations or Wrong Bids," *The Economic Journal*, 121(557), 1361–1397.

Kong, Y. (2017): "Selective Entry in Auctions: Estimation and Evidence," mimeo.

Kosenkova, L. (2017): "Nonparametric inference in asymmetric first-price auctions with k-rationalizable beliefs," .

Lange, A., and A. Ratan (2010): "Multi-dimensional Reference-Dependent Preferences in Sealed-Bid Auctions: How (most) Laboratory Experiments Differ From the Field," *Games and Economic Behavior*, 68(2), 634 – 645.

Lo, K. (1998): "Sealed Bid Auctions with Uncertainty Averse Bidders," *Economic Theory*, 12(1), 1–20.

Lu, J., and I. Perrigne (2008): "Estimating Risk Aversion from Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data," *Journal of Applied Econometrics*, 23(7), 871–896.

Maskin, E., and J. Riley (2003): "Uniqueness of equilibrium in sealed high-bid auctions," *Games and Economic Behavior*, 45(2), 395–409.

Merlo, A., and A. Schotter (1992): "Theory and Misbehavior of First-Price Auctions: Comment," *The American Economic Review*, 82(5), 1413–1425.

Neugebauer, T., and J. Perote (2008): "Bidding as if Risk Neutral in Experimental First Price Auctions without Information Feedback," *Experimental Economics*, 11(2), 190–202.

Neugebauer, T., and R. Selten (2006): "Individual Behavior of First-Price Auctions: The Importance of Information Feedback in Computerized Experimental Markets," *Games and Economic Behavior*, 54(1), 183–204.

Ockenfels, A., and R. Selten (2005): "Impulse balance equilibrium and feedback in first price auctions," *Games and Economic Behavior*, 51(1), 155–170.

Perrigne, I., and Q. Vuong (2007): "Identification and Estimation of Bidders' Risk Aversion in First-Price Auctions," *The American Economic Review*, 97(2), 444–448.

Ratan, A. (2015): "Does Displaying Probabilities Affect Bidding in First-Price Auctions?," *Economics Letters*, 126, 119–121.

Salo, A., and M. Weber (1995): "Ambiguity Aversion in First-Price Sealed-bid Auctions," *Journal of Risk and Uncertainty*, 11(2), 123–137.

Selten, R., J. Buchta, et al. (1994): "Experimental Sealed Bid First Price Auctions with Directly Observed Bid Functions," Discussion paper, University of Bonn, Germany.

Shachat, J., and L. Wei (2012): "Procuring Commodities: First-Price Sealed-bid or English Auctions?," *Marketing Science*, 31(2), 317–333.