Finance and Economics Discussion Series
Divisions of Research & Statistics and Monetary Affairs
Federal Reserve Board, Washington, D.C.

Identification of First-Price Auctions With Biased Beliefs

Serafin J. Grundl and Yu Zhu

2015-056

Please cite this paper as:
Grundl, Serafin J., and Yu Zhu (2015). "Identification of First-Price Auctions With Biased Beliefs," Finance and Economics Discussion Series 2015-056. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.056.

Identification of First-Price Auctions With Biased Beliefs

Serafin Grundl and Yu Zhu*

July 23, 2015

Abstract

This paper exploits variation in the number of bidders to separately identify the valuation distribution and the bidders' belief about the valuation distribution in first-price auctions with independent private values. By exploiting variation in auction volume, the result is extended to environments with risk averse bidders. In an illustrative application we fail to reject the null hypothesis of correct beliefs.

*We are very grateful to Amit Gandhi for his advise and many useful suggestions which helped to improve this paper. We would also like to thank Andres Aradillas-Lopez, Francesco Decarolis, Ken Hendricks, Antonio Penta, Jack Porter, Dan Quint and Ricardo Serrano-Padial for their helpful comments. We thank Doug McDonald for many discussions about timber auctions.
Earlier drafts of this paper were circulated as 'Identification and Estimation of First-Price Auctions without Assuming Correct Beliefs' (Grundl and Zhu (2012)) and 'Identification and Estimation of First-Price Auctions Under Ambiguity' (Grundl and Zhu (2013)).
Serafin Grundl: Federal Reserve Board of Governors, serafin.j.grundl@frb.gov. Yu Zhu: Department of Economics, University of Leicester, yz317@leicester.ac.uk. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the staff, by the Board of Governors, or by the Federal Reserve Banks.

1 Introduction

Estimation of games assumes that players have correct beliefs about the distribution of rival actions, so they coincide with the distribution of rival actions observed by the econometrician and are therefore identified. Identification of preferences then exploits that players best respond to these beliefs. In the context of first-price auctions with independent private values this was demonstrated by Guerre, Perrigne, and Vuong (2000).

The assumption of correct beliefs is plausible in environments where the players have access to similar information as the econometrician.[1] In environments where this is not the case however, it remains an empirical question whether players have correct beliefs and if not, how biased beliefs affect their behavior. Biased beliefs can have important policy implications in first-price auctions. If biased beliefs lead to aggressive bidding the seller has no incentive to reveal more information about bids in previous auctions which could help the bidders to correct their belief. Biased beliefs can also affect the optimal reserve price policy. If biased beliefs lead to aggressive bidding the seller does not have to encourage aggressive bidding with a high reserve price.

This paper studies the identification of the first-price auction model with independent private values without assuming that bidders have correct beliefs. We maintain all assumptions made in the theoretical analysis of first-price auctions as Bayesian Games. Bidders share a common belief about the distribution of valuations which is common knowledge and play the unique symmetric monotone Bayesian Nash Equi-

[1]For first-price auctions this intuition was formalized by Esponda (2008) using the concept of a self-confirming equilibrium. In a self-confirming equilibrium beliefs must be consistent with the distribution of outcomes players can observe. Self-confirming equilibria can therefore be interpreted as the outcome of a learning dynamic. Esponda showed that in a private value environment it is sufficient if bidders observe the two highest bids to ensure that they have correct beliefs.

librium. The observable bid distribution is generated by the valuation distribution and the equilibrium bid function, which depends on the belief.[2]

Without imposing further restrictions the model is not identified because different combinations of the belief and the valuation distribution can generate the same bid distribution. To separately identify the valuation distribution and the belief we exploit variation in the number of bidders. Variation in the number of bidders is useful because the effect of the belief on bidding depends on the number of opponents a bidder is facing. We show that the valuation distribution and the belief can both be non-parametrically point-identified if they have a common support and do not vary with the number of bidders.[3] The assumption can be relaxed if an instrument for the number of bidders is available.

In an extension we consider risk averse bidders because risk aversion can have similar effects on bidding as biased beliefs and Guerre, Perrigne, and Vuong (2009) use variation in the number of bidders to identify a model with risk-averse bidders and correct beliefs. Assuming that bidders have constant relative risk aversion the main identification result can be extended under a support restriction. Variation in the volume or the number of units of the good at sale can be exploited to separately identify the utility function non-parametrically under the additional restriction that valuations are proportional to volume.

We also show that the valuation distribution can be bounded if bidders have "pessimistic beliefs" about the highest rival bid which leads to aggressive bidding.

[2]For a discussion of the common prior assumption see Morris (1995). See also Harsanyi (1968) and Harsanyi (1995). In an extension we relax the assumption of a common belief about the valuation distribution and equilibrium play and instead assume that bidders have "pessimistic beliefs" about rival bids.

[3]This assumption is common in the structural auction literature and often called exogenous participation (e.g. Guerre, Perrigne, and Vuong (2009) and Aradillas-López, Gandhi, and Quint (2013)). Footnote 10 discusses that exogenous participation holds in entry models where bidders receive independent signals before their entry decision.

This result does not rely on an exclusion restriction and allows for heterogeneity of beliefs and risk preferences. The bounds can be tightened if the valuation distribution is stochastically increasing in the number of bidders.

In an illustrative application we study US Forest Service timber auctions in the Pacific Northwest. After the National Forest Management Act of 1976 the Forest Service switched from ascending auctions to first-price auctions as the main sale method in this region. As less than two percent of the sales before the end of 1976 were first-price auctions, the period after the policy change is a suitable environment to relax the assumption of correct beliefs.[4]

We propose a simulated method of moments estimator based on Bierens and Song (2012) which allows us to efficiently combine auctions with various numbers of bidders and to incorporate an unobserved auction characteristic. Moreover, we develop a test of the null hypothesis that bidders have correct beliefs. Our point estimates suggest that bidders with low valuations and therefore a low objective probability of winning have optimistic beliefs while bidders with high valuations and a high probability of winning have pessimistic beliefs but we cannot reject the null hypothesis of correct beliefs.

This paper is related to Aguirregabiria and Magesan (2012) (AM) who relax the assumption of equilibrium beliefs in the estimation of dynamic discrete games.[5] While

[4]Knowledge of the bid distribution in ascending auctions would allow the bidders to uncover the true valuation distribution under the button auction model where every bidder bids her value. Under the milder restrictions in Haile and Tamer (2003) however the bidders could only bound the valuation distribution which could then give rise to a multiple prior model. In section 2 we show that bidding under a multiple prior model with maxmin preferences (Gilboa and Schmeidler (1989)) is observationally equivalent to a model with an appropriately chosen single prior.

[5]AM achieve point identification of heterogeneous beliefs about rival actions and payoff functions for all periods in two player games under the following two assumptions. First, there is a special state variable which enters the payoff of one player but not her rivals. Second, the belief about the distribution of rival actions is known at some points in the support of the special state variable.

The identification argument starts in the last period when the continuation value is zero. Exploiting that the belief is assumed to be known at some points in the state space the last period payoff function can be identified because the player best responds to the known belief. Once the payoff

AM are motivated by strategic uncertainty where players are uncertain about the strategies of their opponents, this paper is mostly concerned with structural uncertainty where players are uncertain about a primitive of the game but beliefs about strategies are in equilibrium.[6] Therefore AM consider identification of heterogeneous beliefs about the distribution of rival actions whereas our main identification result considers the belief about the valuation distribution which is common knowledge among bidders. For this result we exploit that the belief about the distribution of rival actions is endogenously determined which allows us to link data from auctions with different numbers of bidders where bidders have different beliefs about the bid distribution. As the belief about rival actions is endogenously determined we can consider counterfactual policies which affect the belief about rival actions such as changes of the reserve price.

In an extension we show how the valuation distribution can be bounded if the bidders have "pessimistic", possibly heterogenous beliefs about rival actions. This approach is similar in spirit to Haile and Tamer (2003) who show that the valuation distribution can be bounded from bid data in ascending auctions with independent private values under mild restrictions. We believe that robustness concerns are equally important in in the estimation of first-price auctions where optimal bidding depends

function is identified best response behavior can be used to identify the last period belief at other points where it is potentially biased. Therefore the continuation value for the penultimate period is known and the identification argument proceeds backwards in the same fashion.

An analogous approach could be used to achieve point identification of the valuation distribution and heterogeneous beliefs about the bid distribution in first-price auctions. We believe that this would be impractical in the context of our application. First, it is difficult to find a variable which can take the role of the special state variable. Second, we would have to find a subset of auctions for which the beliefs are known or for which the bidders have correct beliefs. Third, in our application we observe only few bids for each bidder but hundreds of bidders. Therefore it would not be feasible to estimate the belief of each bidder.

[6] Strategic uncertainty is an important issue in dynamic games because they often have multiple equally plausible equilibria. In first-price auctions however, there is a unique symmetric monotone equilibrium in pure strategies under fairly weak restrictions. Therefore the equilibrium strategy can be found by "introspection" without learning through repeated play. Hence, strategic uncertainty might be less important in first-price auctions with high stakes.

on the beliefs of players and their risk preferences.

Ambiguity aversion is one important reason why bidders could have "biased beliefs". Biased beliefs are observationally equivalent to a model where bidders have a set of belief distributions and are ambiguity averse with maxmin expected utility preferences (MEU). As briefly laid out in section 2, bidding in the MEU model depends only on the lower contour of the prior set and is therefore identical to a model with an appropriately chosen single belief distribution. Identification and estimation of the MEU model is considered in Aryal and Kim (2014) and Grundl and Zhu (2013). In contrast to Aryal and Kim (2014) this paper illustrates the identification result in an application to field data and also considers non-parametric identification of the utility function and partial identification under "pessimistic beliefs" about rival bids. While Aryal and Kim (2014) develop a bayesian estimator we consider estimation and testing in a frequentist framework.

The remainder of this paper is organized as follows. Section 2 introduces the model and presents the identification results. Section 3 is an illustrative application to timber auctions held by the US Forest Service. Section 4 concludes. Omitted proofs and a discussion of estimation and inference can be found in the Appendix.

2 Identification

Consider a first-price auction with independent private values and I bidders. The bidders share a common prior belief F^b which describes the distribution of valuations and is common knowledge. The density f^b is bounded away from zero on the support $\left[\underline{v}^b, \overline{v}^b\right]$ in $\Re^+$. The bid strategy s in the unique, symmetric, monotone equilibrium is characterized by the first order condition

$$s'(v) = (v - s(v))(I-1)\frac{f^b(v)}{F^b(v)} \tag{1}$$

for $v \in (\underline{v}^b, \overline{v}^b]$ and the boundary condition $s(\underline{v}^b) = \underline{v}^b$.[7]

Valuations are iid draws from the distribution F with density f which is bounded away from zero on the support $[\underline{v}, \overline{v}] \subseteq [\underline{v}^b, \overline{v}^b]$ in $\Re^+$. The support restriction ensures that bidders do not rule out any possible valuations. The bid distribution is $G(b) = F(s^{-1}(b))$, where s^{-1} is the inverse bid function.

In a closely related model the belief is a closed and convex set of distribution functions Δ and the bidders have maxmin expected utility preferences (Gilboa and Schmeidler (1989)) which was first studied by Lo (1998). Bidders choose b to maximize $\min_{F_\Delta \in \Delta} (v-b) F_\Delta (s^{-1}(b))^{I-1}$. If F^b coincides with the lower contour of Δ ($F^b(v) = \min_{F_\Delta \in \Delta} F_\Delta(v)$) both models generate the same bid functions.[8]

Under the assumption that the bidders know the valuation distribution ($F^b = F$) the classic identification result in Guerre, Perrigne, and Vuong (2000) applies. The key to their identification argument is that the bidders bid as if they best respond to the bid distribution observed by the econometrician.[9] Therefore the inverse bidding strategy is identified and can be used to uncover the unobservable valuations from observed bids. This identification strategy is no longer valid if $F \neq F^b$, because the bidders might not best respond to the bid distribution observed by the econometrician. Indeed the model cannot be identified without imposing further restrictions.

Theorem 1 (Nonidentification)**.** *Let G be any bid distribution with a density which*

[7] The equilibrium is unique up to bids made by a measure zero set of types. See McAdams (2007) Theorem 1 and footnote 3. See also Athey and Haile (2007) Theorem 2.1 (ii) and footnote 12.

[8] Bose, Ozdenoren, and Pape (2006) and Bodoh-Creed (2012) study the mechanism design problem in this environment. See also Aryal and Kim (2014) and Grundl and Zhu (2013) for more detailed discussions.

[9] Importantly the identification strategy does not require that the bidders understand the equilibrium. It is sufficient that the bidders know the bid distribution and best respond to it.

is bounded away from zero and infinity on the support $\left[\underline{b}, \overline{b}\right]$*. Then* G *can be generated by a model where* F^b *is the uniform distribution on* $\left[\underline{b}, \underline{b} + \frac{I\left(\overline{b}-\underline{b}\right)}{I-1}\right]$ *so the bid function is* $s\left(v\right) = \underline{b} + \frac{I-1}{I}\left(v - \underline{b}\right)$ *and* $F\left(v\right) = G\left(\underline{b} + \frac{I-1}{I}\left(v - \underline{b}\right)\right)$*.*

This result is not surprising as we cannot hope to uncover the valuation distribution and the belief distribution from a single bid distribution.

In light of this nonidentification result we exploit variation in the number of bidders to identify both distributions. Variation in the number of bidders can provide additional information because the effect of beliefs on bid shading depends on the number of opponents a bidder is facing. To use this identification strategy we rely on an exclusion restriction.

Assumption 1. *F and F^b do not depend on I.*

If in addition $F = F^b$ this assumption reduces to the standard form used in the literature which is often called exogenous participation (e.g. Guerre, Perrigne, and Vuong (2009) and Aradillas-López, Gandhi, and Quint (2013)).[10]

Under Assumption 1 the two first order conditions from auctions with different numbers of bidders define a system of two functional equations involving F and F^b. If this system has a unique solution F and F^b are point identified. Showing uniqueness

[10] Exogenous participation holds in entry models with independent signals. Suppose N potential bidders observe some signal s_i which is iid across bidders but might be informative about their value before they enter. In equilibrium a bidder's entry decision is determined by a cutoff $\overline{s}$ such that the expected payoff from entry is zero for the cutoff bidder. For any number of entering bidders the valuation distribution is then simply the valuation distribution of a potential bidder conditional on $s_i \geq \overline{s}$. If this is reflected in the bidders' beliefs Assumption 1 is satisfied. See Grundl and Zhu (2015) for a formal exposition of this argument.

The equilibrium cutoff signal depends on N. If N varies across auctions and is not observed the identification strategy is no longer valid. An, Hu, and Shum (2010) discuss the case where N is an unobservable auction characteristic.

Variation in the number of actual bidders created by the entry stage is useful for identification if the bidders know the number of actual bidders at the bidding stage. This is a common assumption in the literature (e.g. Athey, Levin, and Seira (2011) and Krasnokutskaya and Seim (2011)). For an alternative approach relying on variation in the number of potential bidders see Gentry, Li, and Lu (2015).

in a system of functional equations is a difficult problem in general. In this particular case we can exploit the form of the first order conditions to show that F and F^b are identified. Moreover the identification proof is constructive and leads to a closed form for the inverse bid function.

Let s_I be the bid function in an I bidder auction and G_I the corresponding bid distribution with density g_I.First use the monotonicity of the bid function to write $G_I\left(s_I\left(v\right)\right) = F\left(v\right)$. Differentiating this with respect to v yields $g_I\left(s_I\left(v\right)\right)s_I'\left(v\right) = f\left(v\right)$. Therefore the first order condition can be rewritten as follows

$$(v - s_I(v))\frac{f^b(v)}{F^b(v)}\frac{F\left(v\right)}{f\left(v\right)} = \frac{G_I\left(s_I\left(v\right)\right)}{g_I\left(s_I\left(v\right)\right)\left(I-1\right)},\ v \in \left(\underline{v}, \overline{v}\right]. \tag{2}$$

If the bidders know the valuation distribution ($F^b = F$) they best respond to the bid distribution and the right hand side is the bid shading of the bidder with valuation v (Guerre, Perrigne, and Vuong (2000)). If the bidders have biased beliefs the right hand side has to be adjusted by the ratio of the reverse hazard rates $\frac{f(v)}{F(v)}$ and $\frac{f^b(v)}{F^b(v)}$ to obtain the bid shading. Our identification strategy exploits the fact that this adjustment factor does not vary across auctions with different numbers of bidders. Take the ratio of the two first order conditions (2) with I_1 and I_2 bidders and solve for the underlying valuation as

$$v = \frac{s_2(v)\left(I_2-1\right)g_2\left(s_2(v)\right) - s_1(v)\left(I_1-1\right)g_1\left(s_1(v)\right)}{\left(I_2-1\right)g_2\left(s_2(v)\right) - \left(I_1-1\right)g_1\left(s_1(v)\right)},\ v \in \left[\underline{v}, \bar{v}\right]. \tag{3}$$

We use the subscripts 1 and 2 to index the strategies and bid distributions instead of I_1 and I_2. Corresponding bids with same underlying valuation can be coupled through $G_2\left(s_2(v)\right) = F(v) = G_1\left(s_1(v)\right)$, because the bid function is strictly increasing. As a result, the inverse bid function in the I_1 bidder auction is identified as follows

$$s_1^{-1}(b) = \frac{b_2(b)(I_2-1)\,g_2(b_2(b)) - b(I_1-1)\,g_1(b)}{(I_2-1)\,g_2(b_2(b)) - (I_1-1)\,g_1(b)},\ b \in [s_1(\underline{v}), s_1(\bar{v})], \tag{4}$$

where $b_2(b) = G_2^{-1}(G_1(b))$. This immediately implies that F is identified.

The first order condition can now be solved for

$$\begin{aligned} F^b(v|v \le \overline{v}) &= \exp\left(-\int_v^{\bar{v}} \frac{f(u)}{(u - s_1(u))\,g_1(s_1(u))(I_1-1)}du\right) \\ &= \exp\left(-\mathrm{E}_V\left[\frac{1\{V > v\}}{(V - s_1(V))\,g_1(s_1(V))(I_1-1)}\right]\right),\ v \in (\underline{v}, \bar{v}], \end{aligned} \tag{5}$$

where the expectation is taken with respect to F.

Theorem 2 (Point Identification)**.** *Under Assumption 1 F and $F^b(\cdot|v \le \overline{v})$ are point identified on $[\underline{v}, \overline{v}]$. If in addition $\overline{v}^b = \overline{v}$, F^b is identified on $[\underline{v}, \overline{v}]$.*

Remark 1. Notice that the bid function can be expressed in terms of $F^b(\cdot|v \le \overline{v})$ on $[\underline{v}, \overline{v}]$. On the one hand this implies that without the support condition $\overline{v}^b = \overline{v}$ we can only identify the "conditional belief distribution" $F^b(\cdot|v \le \overline{v})$ rather than F^b. On the other hand knowledge of $F^b(\cdot|v \le \overline{v})$ is sufficient for most counterfactuals. As $F^b(\cdot|v \le \overline{v})$ bounds F^b from above we can detect if bidders have overly pessimistic beliefs without the support condition.

Remark 2. The identification result can be extended to allow for an unobserved auction characteristic if multiple bids from the same auction are observed. Accounting for an unobserved characteristic is always important in the estimation of first-price auctions because otherwise variation in bids across auctions due to variation in the unobservable characteristic will be wrongly attributed to private information. Accounting for an unobserved characteristic is however particularly important if variation in the number of bidders is exploited for identification as in this paper. As explained in footnote 10 Assumption 1 is satisfied in entry models with independent signals. However it is important to apply Theorem (2) to the bid distributions conditional on the unobserved characteristic and to let the distribution of the unobserved characteristic vary with the number of bidders. Otherwise, differences in the bid distribution across auctions with different numbers of bidders due to selective entry will be wrongly attributed to the bidders' belief.[11]

If the unobserved characteristic enters additively or multiplicatively as in Krasnokutskaya (2011) observing two bids per auction is sufficient. If the unobserved characteristic enters in a nonseparable way as in Hu, McAdams, and Shum (2013) at least three bids are required. These papers show how to identify the bid distribution conditional on the unobserved characteristic. The identification result in Theorem (2) can then be applied to the conditional bid distributions.[12]

Remark 3. If Assumption 1 is violated, point identification requires a different exclusion restriction. If an instrument for the number of bidders is available identification follows from a similar result as Guerre, Perrigne, and Vuong (2009, Corollary 2). For details see Appendix A.1.

Remark 4. With a binding reserve price r, F and $F^b\left(\cdot|v \leq \overline{v}\right)$ can only be identified on $[r, \bar{v}]$. For details, please refer to Appendix A.2.

Remark 5. So far we maintained all the assumptions embedded in the Bayesian Nash Equilibrium concept. Bidders share a common knowledge belief about the valuation distribution and the belief about rival actions is endogenously determined. This is exploited in the identification argument because it allows us to link auctions with different numbers of bidders where bidders have different beliefs about rival actions. It also enables us to consider counterfactuals which affect the belief about rival actions

[11]This is explained in more detail in Grundl and Zhu (2015).

[12]Grundl and Zhu (2013) discusses the separable case in more detail.

such as changes in the reserve price.

In some environments the researcher might not be willing to maintain the assumption of Bayesian Nash Equilibrium. Theorem 4 in Appendix A.3 considers identification under fairly mild restrictions. The valuation distribution can be bounded if bidders have "pessimistic" beliefs about the highest rival bid. We show that bidders with pessimistic beliefs bid more aggressively than with correct beliefs, which allows us to bound the valuation underlying some observed bid from above. As the underlying valuation is also bounded from below by the bid itself the valuation distribution can be bounded from both sides. More generally these bounds are robust to all deviations from the standard model which lead to more aggressive bidding, for example risk aversion. The bounds can be tightened if the valuation distribution is stochastically increasing in the number of bidders. This approach allows for heterogeneous beliefs and does not require equilibrium play.

2.1 Risk Aversion

In this section we consider risk averse bidders because risk aversion and biased beliefs can have similar effects on bidding. Moreover Guerre, Perrigne, and Vuong (2009) use variation in the number of bidders to identify a model with risk-averse bidders and correct beliefs.

Unlike risk aversion biased beliefs do not always lead to more aggressive bidding. Moreover bid distributions generated by a model with biased beliefs and risk neutral bidders cannot always be rationalized by a model with correct beliefs and risk averse bidders.[13]

[13]Both of these claims can be shown with an example adapted from Bodoh-Creed (2012) where $I = 2$, F is the standard uniform distribution and $F^b(v) = \frac{1}{2}\left(3v - v^2\right)$ on $[0,1]$. In this case $s(v) = \left(\frac{3}{2}v - \frac{2}{3}v^2\right)/(3-v) < \frac{v}{2}$ for $v \in (0,1]$ so biased beliefs lead to less aggressive bidding. A sufficient condition to ensure that biased beliefs lead to more aggressive bidding is that beliefs

In the remainder of this section we ask whether risk aversion can be separately identified from the other primitives of the model.

Constant Relative Risk Aversion Let $\gamma \in [0,1)$ be the coefficient of relative risk aversion. The first order condition becomes

$$s'(v) = \begin{cases} (v - s(v)) \frac{(I-1)}{1-\gamma} \frac{f^b(v)}{F^b(v)} & v \in \left(\underline{v}^b, \overline{v}^b\right] \\ \frac{I-1}{I-\gamma} & v = \underline{v}^b \end{cases} \tag{6}$$

The form of the derivative at the lower bound is the limit of $s'(v)$ as v goes to $\underline{v}^b$ and has first been derived in Guerre, Perrigne, and Vuong (2009, Theorem 1).

The inverse bid function 4 remains unchanged because γ cancels out when we take the ratio of the two first order conditions. Therefore F is identified.

Assumption 2. $\underline{v}^b = \underline{v}$.

This restriction allows us to exploit the form of the first order condition at the lower bound, where it does not depend on the belief, to identify γ. Take the ratio of two first order conditions 6 at $\underline{v}$:

$$\frac{I_2-\gamma}{I_2-1}\frac{I_1-1}{I_1-\gamma} = \frac{s_1'(\underline{v})}{s_2'(\underline{v})} = \frac{g_2(\underline{v})}{g_1(\underline{v})} \iff \gamma = \frac{(I_2-1)\, I_1 g_2(\underline{v}) - (I_1-1)\, I_2 g_1(\underline{v})}{(I_2-1)\, g_2(\underline{v}) - (I_1-1)\, g_1(\underline{v})} \tag{7}$$

Identification of the conditional belief distribution follows from analogous arguments as in the risk neutral case. Integrating the first order condition yields:

---

are "pessimistic" in the sense of Reverse Hazard Rate Dominance relative to the true valuation distribution ($F^b(v)/f^b(v) \leq F(v)/f(v)$ for all v). See Appendix A.3.

Furthermore, the bid distributions generated by this example cannot be rationalized by a model with correct beliefs and risk averse bidders. If we would follow the identification strategy in Guerre, Perrigne, and Vuong (2009) we would recover a utility function which is not concave. Matlab code to illustrate this is available from the authors.

$$F^b\left(v|v\leq\overline{v}\right)^{\frac{1}{1-\gamma}}=\exp\left(-\mathrm{E}_V\left[\frac{1\{V>v\}}{\left(V-s_1\left(V\right)\right)g_1\left(s_1\left(V\right)\right)\left(I_1-1\right)}\right]\right)\ v\in\left(\underline{v},\bar{v}\right]$$

Corollary 1 (Constant Relative Risk Aversion)**.** *Under Assumptions 1 and 2* γ *is identified and* F*,* $F^b\left(\cdot|v\leq\overline{v}\right)$ *are identified on* $[\underline{v},\bar{v}]$*.*

Nonparametric Utility Function In many applications there is variation in the volume of the auctioned good or the number of auctioned units x which can be exploited to achieve nonparametric identification of the utilty function. Intuitively variation in the auction volume allows us to distinguish decreasing, constant and increasing relative risk aversion.

Assumption 3. $F\left(\cdot|x\right)=F\left(\cdot\right)$ *and* $F^b\left(\cdot|x\right)=F^b\left(\cdot\right)$ *for all* x*.*

Let U be a von Neumann-Morgenstern utility function with $U\left(0\right)=0$, $U\left(1\right)=1$, $U'\left(\cdot\right)>0$ and $U''\left(\cdot\right)\leq 0$.

Assumption 4. $\left|\sup_z \frac{U(z)}{U'(z)z}\right|\leq M<\infty$*.*

Theorem 3 (Nonparametric Risk Aversion)**.** *Suppose Assumptions 1, 2, 3 and 4 hold and there is sufficient variation in* x *then* F *and* $F^b\left(\cdot|v\leq\overline{v}\right)$ *are identified on* $[\underline{v},\bar{v}]$ *and* U *is identified on some interval which depends on the observed range of* x*.*

The proof can be found in Appendix B. The exclusion restriction in Assumption 3 requires that the distribution of values per unit of the auctioned good does not depend the auction volume and this is reflected in the bidders' belief. This is an important restriction. For example, it rules out higher per-unit valuations for larger x because of scale economies. Alternatively we could try to use other restrictions on how some observable enters F and F^b to nonparametrically identify the utility function (see

Campo, Guerre, Perrigne, and Vuong (2011)). We focus on Assumption 3 as it has a natural interpretation and there is variation in auction volume in many auction data sets.

Assumption 4 is a technical restriction to ensure that the per-unit bid shading is bounded away from zero. Loosely speaking, if the per-unit bid shading is zero bidders do not care about how many units they win and we can therefore not make use of variation in the auction volume.

Theorem 3 is an extension of Theorem 2 under additional restrictions. An interesting question left for future research is whether the model primitives can be identified if only one of the exclusion restrictions, either for the number of bidders or the auction volume is imposed.

3 Empirical Application

This section presents an illustrative application in an environment where bidders had very little prior experience with first-price auctions. We study US Forest Service timber auctions in the Pacific Northwest from December 1976 to December 1978. Prior to the National Forest Management Act of 1976 the Forest Service relied almost exclusively on ascending auctions in the Pacific Northwest. Less than two percent of the sales were first-price sealed bid auctions during that time. Due to concerns that ascending auctions raise less revenue this policy was changed and starting in December 1976 almost all sales were first-price sealed bid auctions.[14] Many bidders opposed this change and the Forest Service gradually returned to using ascending auctions. In the years after 1978 less than five percent of the sales were sealed bid auctions. Figure I shows the number of first-price auctions every month between 1973 and 1979.

[14]One concern was that ascending auctions are vulnerable to collusion, see Baldwin, Marshall, and Richard (1997).

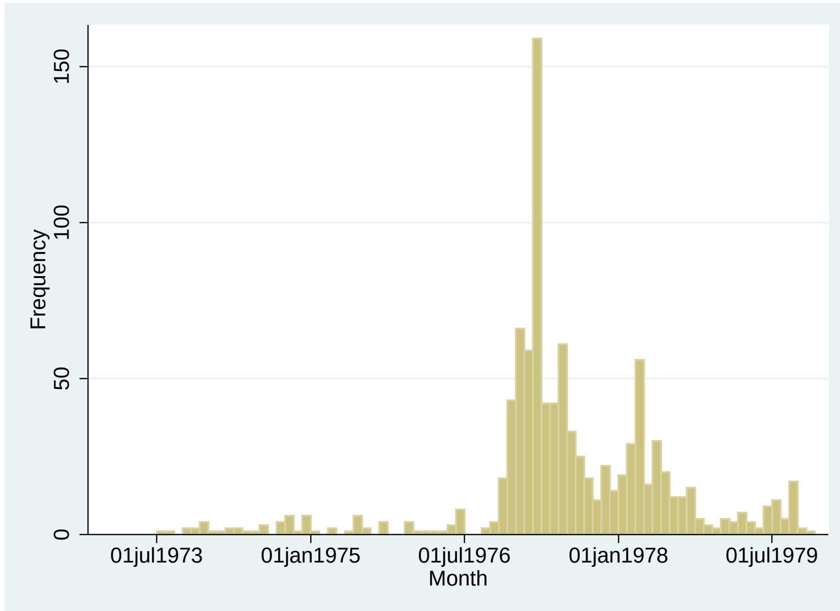

Figure I: Number of first-price auctions in the Pacific Northwest region of the US Forest Service from 1973 to 1979. Data Sources: US Forest Service timber auction data, accessible from http://www.econ.yale.edu/ pah29/timber/timber.htm.

A potential concern with estimating a static model in this environment is that beliefs could evolve as bidders gain more experience with first-price auctions during the sample period. To investigate this we regress bids on bidder experience with fixed effects for auctions and bidders. Bidder experience is defined as the number of first-price auctions a bidder has entered before participating in the auction under consideration.[15] The median experience level in our sample is two which reflects the small size of most bidding firms. We find no significant effect of experience on bidding and conclude that estimating a static model is appropriate.[16]

[15]This includes experience from first-price auctions before the beginning of the sample period.

[16]The regression includes all first-price auctions during the sample period with 2-11 bidders. The sample contains 2957 bids from 1068 auctions submitted by 595 bidders. We regress log(bid) on log(1+experience) with bidder and auction fixed effects. The estimated coefficient on log(1+experience) is 0.002 with a standard error of 0.017. The point estimate suggests that a 1% increase in 1+experience corresponds to a 0.002% increase in bids; the 95% confidence interval ranges from -0.031% to 0.035%. In light of the limited experience which bidders acquire during the sample period (the median experience level at the time of bidding is 2 auctions, the 75 percentile 7 auctions and the 90th percentile 16 auctions) we conclude that the effect of bidder experience is economically

While the bidders had very limited experience in first-price auctions they could rely on their experience in ascending auctions to form their beliefs about the distribution of valuations. Knowledge of the bid distribution in ascending auctions however only allows the bidders to bound the valuation distribution without imposing strong assumptions (Haile and Tamer (2003)). This could then give rise to a multiple prior model where the bidders consider all distributions in the identified set to be reasonable. This model is observationally equivalent to a model with a single belief distribution which equals the lower contour of the prior set as briefly explained in section 2.

Specification, Estimation and Results We illustrate how to implement the identification result in Theorem 2. We allow for an unobserved auction characteristic u which enters valuations in a multiplicative way $v_i = uv_i^*$ as in Krasnokutskaya (2011).[17] From now on we let F and F^b denote the distribution of the private information component v_i^* and the corresponding belief. We assume that the bidders know the support of f so the results can be interpreted as estimates of F^b rather than of $F^b\left(\cdot|v \leq \overline{v^*}\right)$.

To estimate f, f^b and the density of u f_I^u we adapt the estimator in Bierens and Song (2012) which matches the joint characteristic functions of log bids from the

negligible. We have also tried specifications where experience is binned and found that the coefficients frequently change signs as the experience level increases and the effect of experience is not statistically significant.

One explanation for this finding is simply that the bidders have correct beliefs without prior experience in first-price auctions and therefore do not need to learn as they gain experience. An alternative explanation is that learning through experience is slow in this environment because the auctioned timber tracts differ in many of their characteristics (e.g. appraisal value, timber volume, acreage, timber density, species composition etc.) and the bidders do not acquire sufficient experience during the sample period to change their initial beliefs significantly.

[17] Several papers stress the importance of allowing for unobserved characteristics in USFS timber data (e.g. Athey, Levin, and Seira (2011), Aradillas-López, Gandhi, and Quint (2013) or Roberts and Sweeting (2013)). We find that the multiplicative form fits the data well: While log bids from the same auction remain highly correlated after controlling for many observables, the differences between two log bids are essentially uncorrelated.

same auction.[18] The unknown densities are approximated with 5-th order Legendre polynomials. We also provide a testing procedure for the hypothesis that bidders have correct beliefs $F = F^b$. Details are provided in Appendix C.

It is difficult to use all auctions in estimation because we have to estimate a separate f_I^u for every I. Identification comes from the difference in competitiveness for auctions with different numbers of bidders. Therefore combining two and four bidder auctions yields more precise estimates than combining two and three bidder auctions for example. Moreover adding an additional bidder to a two bidder auction doubles the number of competitors a bidder is facing whereas adding an additional bidder to a five bidder auction increases the number of competitors by only one quarter. In light of these considerations we combine auctions with two, four and five bidders to estimate the model.[19]

In a first step we regress the log of the bids on the log of observable timber tract characteristics and apply the identification strategy to the exponential of the residual of this regression. This approch is valid if the the observable characteristics enter the valuations multiplicatively because then they also enter the bids multiplicatively (Krasnokutskaya (2011)). We include the two observables we found to be most important, the appraisal value and the timber volume.[20]

[18]As the identification argument yields a closed form for the inverse bid function a two step estimator in the spirit of Guerre, Perrigne, and Vuong (2000) is a natural alternative. However if there is an unobserved auction characteristic we can no longer directly invert the observed bids to uncover the underlying valuations. Moreover, a two step estimator does not allow us to combine auctions with various numbers of bidders in an efficient manner. Grundl and Zhu (2013) discusses the asymptotic properties of a two step estimator to implement the point identification result in Theorem 2. They also propose a matching estimator for the CRRA coefficient to implement Corollary 1 which does not suffer from the curse of dimensionality as the number of observable covariates increases. The convergence rate of the estimator for the CRRA coefficient approaches the parametric rate as the smoothness of the model primitives increases.

[19]Grundl and Zhu (2013) contains a Monte Carlo study illustrating that combining 2 and 4 bidder auctions yields more precise estimates than combining 2 and 3 bidder auctions or 3 and 4 bidder auctions.

[20]The regression results are available from the authors.

Figure II shows the estimates of F and F^b.[21] These point estimates suggest that bidders with low valuations and therefore a low probability of winning have optimistic beliefs while bidders with high valuations and a high probability of winning have pessimistic beliefs. This leads to underbidding for bidders with low valuations and overbidding for bidders with high valuations. If bidders do have biased beliefs this can have important policy implications for the seller's information policy which could help bidders to correct their beliefs and the optimal reserve price policy. However we do not analyze the policy implications of the point estimates as we fail to reject that the bidders have correct beliefs (test statistic = 0.1227 and p-value = 0.27). For the same reason we also do not consider the extensions to risk averse bidders in Corollary 1 and Theorem 3 or the partial identification result in Theorem 4.

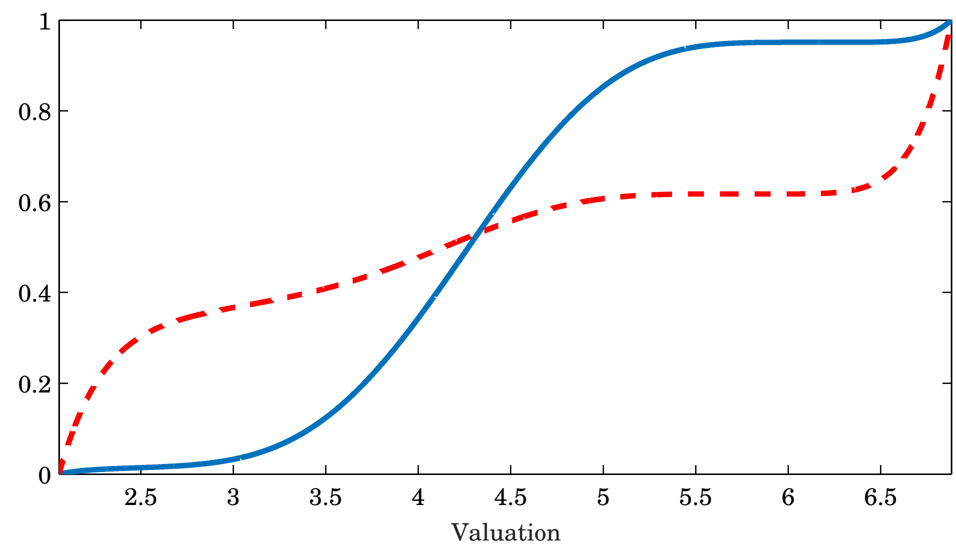

Figure II: Estimates of F (solid blue line) and F^b (dashed red line). On the horizontal axis is the bidder's private information v_i^*.

[21]The estimates for the distributions of the unobservable characteristic are not reported. We find that these distributions are fairly similar for auctions with different numbers of bidders. This is consistent with the distribution of observable characteristics which are also similar for auctions with different numbers of bidders.

4 Conclusion

This paper considers environments where the bidders do not have access to similar information as the econometrician at the time of bidding. We argue that in such an environment it is an empirical question whether bidders have correct beliefs.

We exploit variation in the number of bidders to separately identify the valuation distribution and the bidders' belief about the valuation distribution. We also show how to exploit variation in auction volume to separately identify the bidders' utility function. In an application we illustrate how to estimate the model and how to test the assumption of correct beliefs.

Relaxing some of the maintained assumptions such as symmetric bidders and independent private valuations are challenging avenues for future research.

A Remarks about Theorem 2

A.1 Instrument for the Number of Bidders

Suppose F and F^b vary with the number of bidders. In this case we can use a different exclusion restriction to achieve identification. We require an instrument which affects participation, but does not affect the the valuation distribution or the belief.

Formally, let F and F^b depend on an unobservable ϵ and a vector of observable auction characteristics, which we partition into Z_1 and Z_2, where Z_2 will take the role of the instrument. The following result mirrors Guerre, Perrigne, and Vuong (2009, Corollary 2).

Corollary 2. *Suppose* $I = I(Z_1, Z_2, \epsilon)$, $F(\cdot|Z_1, Z_2, \epsilon) = F(\cdot|Z_1, \epsilon)$ *and* $F^b(Z_1, Z_2, \epsilon) = F^b(Z_1, \epsilon)$*. Let* $[\underline{v}(Z_1, \epsilon), \overline{v}(Z_1, \epsilon)]$ *be the support of* $f(\cdot|Z_1, \epsilon)$*. If either of the following conditions hold*

(i). $\epsilon = I - \mathrm{E}[I|Z_1, Z_2]$

(ii). $Z_2 = m(X, \epsilon)$ *where* $m(\cdot, \cdot)$ *is strictly increasing in* ϵ *with* $X \perp \epsilon$ *and* $X \not\subseteq Z_1$*.*

$F(\cdot|Z_1, \epsilon)$ *and* $F^b(\cdot|Z_1, \epsilon, v \leq \overline{v}(Z_1, \epsilon))$ *are identified on* $[\underline{v}(Z_1, \epsilon), \overline{v}(Z_1, \epsilon)]$ *for each* Z_1, ϵ*.*

In case (i) ϵ is the residual of the nonparametric regression of I on Z_1, Z_2. For case (ii), Matzkin (2003) shows how to recover ϵ under a normalization, exploiting the monotonicity of m. Hence, ϵ can be treated like an observable. Using the two bid distributions $G_1(\cdot|Z_1, \epsilon)$ and $G_2(\cdot|Z_1, \epsilon)$ for $I_1 < I_2$, identification follows as in Theorem 2.

A.2 Identification With a Reserve Price

With a binding reserve price $r > \underline{v}$, only bidders who draw valuations above r become active and place a bid. Suppose Assumption 1 holds and we observe the number of active bidders I_1^*, I_2^* and the truncated bid distributions G_1^*, G_2^* from auctions with two different numbers of potential bidders $1 < I_1 < I_2$. I_1^* and I_2^* follow a binomial distribution with parameters $[I_1, 1 - F(r)]$ and $[I_2, 1 - F(r)]$. Consequently, I_1, I_2 and $F(r)$ are identified. The bid function with I_1 potential bidders can be inverted using equation (3), where the bid densities are replaced with their truncated counterparts. Since $G_1^*(s_1(v,r)) = \frac{F(v)-F(r)}{1-F(r)}$ F is identified on $[r, \bar{v}]$. We can also identify $F^b(\cdot|v \leq \overline{v})$ on $[r, \bar{v}]$ using equation (5).

Corollary 3 (Reserve Price)**.** *Suppose we know the truncated bid distributions* G_1^* *and* G_2^* *from auctions with two different numbers of potential bidders* $1 < I_1 < I_2$ *and a binding reserve price* $r > \underline{v}$*. Then under Assumption 1* F *and* $F^b(\cdot|v \leq \overline{v})$ are identified on $[r, \bar{v}]$.

A.3 Partial Identification With Pessimistic Beliefs

This section shows that type specific valuation distributions can be bounded if bidders have "pessimistic beliefs" about the highest rival bid without relying on exclusion restrictions or equilibrium play and allowing for heterogeneity of beliefs and risk aversion. Bidder type t believes that her probability of winning if she bids b is $H_t^b(b)$with density h_t^b. Her valuations are drawn from F_t and her utility function is U_t with $U_t'(\cdot) > 0$ and $U_t''(\cdot) \leq 0$, $U_t(0) = 0$, $U_t(1) = 1$ and $\lambda_t(\cdot) = U_t(\cdot)/U_t'(\cdot)$. Let H_t be t's objective probability of winning. We assume that bidders have pessimistic beliefs in the following sense.

Assumption 5 (Reverse Hazard Rate Dominance (RHRD))**.** $\frac{H_t^b(b)}{h_t^b(b)} \leq \frac{H_t(b)}{h_t(b)}$ *for all* b*.*

This assumption is partly motivated by the large experimental literature devoted to the overbidding puzzle.[22] The overbidding puzzle is the finding that bidders in laboratory experiments bid more aggressively than predicted by the risk neutral Bayesian Nash Equilibrium. As we show below, RHRD implies more aggressive bidding than under correct beliefs about the highest rival bid.

In the Bayesian Nash Equilibrium with a common belief about the valuation distribution we have $\frac{H^b(b)}{h^b(b)} = \frac{F_I^b\left(s^{-1}(b)\right)s'\left(s^{-1}(b)\right)}{f_I^b(s^{-1}(b))(I-1)}$ and $\frac{H(b)}{h(b)} = \frac{F_I\left(s^{-1}(b)\right)s'\left(s^{-1}(b)\right)}{f_I(s^{-1}(b))(I-1)}$. Therefore a sufficient condition for Assumption (5) is that the common belief about the valuation distribution satisfies RHRD: $F_I^b\left(v\right)/f_I^b(v) \leq F_I\left(v\right)/f_I\left(v\right)$ for all v.

The first order condition for bidder i is

$$v = \lambda_t^{-1}\left(\frac{H_t^b\left(b\right)}{h_t^b\left(b\right)}\right) + b \leq \frac{H_t\left(b\right)}{h_t\left(b\right)} + b \equiv \xi_t\left(b\right) \tag{8}$$

The inequality follows from Assumption 5 and the fact that λ_t^{-1} is increasing with $\lambda_t^{-1}\left(z\right) \leq z$ for $z \geq 0$.

.

Let G_t be the bid distribution of type t. If we observe many auctions where bidder type t participates H_t and G_t are identified from bid data and we can bound the valuation distribution as follows:

$$G_t\left(\xi_t^{-1}\left(v\right)\right) \leq F_t\left(v\right) \leq G_t\left(v\right) \tag{9}$$

The upper bound is given by the bid distribution because no bidder bids more than her value. The lower bound is given by the distribution of the inverse bid function assuming risk neutrality and correct beliefs as in Guerre, Perrigne, and Vuong (2000), because risk aversion and biased beliefs satisfying RHRD lead to more aggressive bid-

[22] See for example Kagel and Levin (2008) and the references therein.

ding. More generally the lower bound is valid whenever bidders bid more aggressively than a bidder who maximizes $H_t(b)(v_i - b)$, for example due to loss aversion.

Assumption 6 (Stochastically Increasing Valuations)**.** *$F_{I_2,t}(v) \leq F_{I_1,t}(v)$ for all v and $I_1 < I_2$ and all i.*

This relaxes Assumption 1 and says that valuations are stochastically increasing in the number of bidders in the sense of first-order stochastic dominance. This restriction allows us to tighten the bounds of the valuation distribution. We construct $\xi_{I,t}(b) = H_t(b|I)/H_t(b|I) + b$ and let $G_{I,t}$ denote t's bid distribution in I bidder auctions:

$$\max_{J \geq I} G_{J,t}\left(\xi_{J,t}^{-1}(v)\right) \leq F_{I,t}(v) \leq \min_{J \leq I} G_{J,t}(v) \tag{10}$$

If $G_{I,t}\left(\xi_{I,t}^{-1}(v)\right)$ is not contained in these tighter bounds we can conclude that bidder i has biased beliefs or is risk averse.

Theorem 4 (Partial Identification)**.**

(i). Bidder type t's valuation distribution can be bounded as in (9) *under Assumption* (5).

(ii). These bounds can be tightened as in (10) *under Assumptions* (5) *and* (6).

B Proof of Theorem 3

Suppose we observe I_1 and I_2 bidder auctions with $x \in [\underline{x}, \overline{x}]$ where $\underline{x} > 0$.[23] Under Assumption 3 the first-order condition becomes

$$\frac{\partial}{\partial v} s(v,x) = \begin{cases} \frac{(I-1)\lambda'(0)}{(I-1)\lambda'(0)+1} & v = \underline{v} \\ \frac{(I-1)f^b(v)\lambda(x(v-s(v,x)))}{F^b(v)x} & v > \underline{v} \end{cases} \tag{11}$$

, where x is volume or the number of units of the auctioned good, v is the per-unit valuation, s is the per-unit bid function $F^b(\cdot|x)$ is the per-unit belief distribution with density $f^b(\cdot|x)$and $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. Let $v(\alpha)$ denote the α-th quantile of the per-unit valuation distribution and $b_I(\alpha, x)$ the corresponding per-unit bid in a I bidder auction with volume x.

Lemma 1. *If Assumption 4 is satisfied there exist $\beta_1(\alpha)$ and $\beta_2(\alpha)$ such that $0 < \beta_1(\alpha) \leq v(\alpha) - b_I(\alpha, x) \leq \beta_2(\alpha)$ for $\alpha \in (0,1]$.*

Proof. First we show that if $\lambda_1(z) \leq \lambda_2(z)$ for $z \geq 0$ then the corresponding bid functions are ordered such that $s_1(v,x) \leq s_2(v,x)$ for all v. Suppose that $s_1(v^*, x) = s_2(v^*, x)$ for some $v* > \underline{v}$ then $\frac{\partial}{\partial v} s_2(v^*, x) \geq \frac{\partial}{\partial v} s_1(v^*, x)$ so $s_1(v,x) \leq s_2(v,x)$ for all $v \geq v^*$. It remains to rule out that $s_1(v,x) > s_2(v,x)$ for $v \in (v, v+\epsilon]$. This would require that $\lambda_1'(0) > \lambda_2'(0)$ and therefore $\lambda_1(z) > \lambda_2(z)$ for some z sufficiently close to 0 which is a contradiction.

Now set the lower bound equal to the per-unit bid shading if $\lambda(z) = zM$ for all z and the upper bound equal to the per-unit bid shading of a risk neutral bidder with $\lambda(z) = z$:

[23] If $x \in (0, \overline{x}]$ the identification argument can be considerably simplified but in many applications we do not observe auctions for x close to zero.

$$\beta_1(\alpha) = \int_{\underline{v}}^{v(\alpha)} \left[\frac{F^b(t)}{F^b(v(\alpha))}\right]^{M(I-1)} dt$$
$$\beta_2(\alpha) = \int_{\underline{v}}^{v(\alpha)} \left[\frac{F^b(t)}{F^b(v(\alpha))}\right]^{I-1} dt$$

Differentiate $\lambda(z)$ to see that $u''(\cdot) \leq 0$ implies $\lambda(z) \geq z$ and by Assumption 4 $\lambda(z) \leq Mz$ for $z \geq 0$. This completes the proof of Lemma 1. □

It follows from Lemma 1 that if $[\underline{x}, \overline{x}]$ is sufficiently wide and Assumptions 3 and 4 are satisfied then for each quantile $\alpha \in (0,1]$ we can find $x_1 \neq x_2$ such that

$$\begin{aligned} x_1(v(\alpha) - s_1(v(\alpha), x_1)) &= x_2(v(\alpha) - s_2(v(\alpha), x_2)) \\ g_1(\alpha|x_1)(I_1 - 1)/x_1 &= g_2(\alpha|x_2)(I_2 - 1)/x_2 \end{aligned} \tag{12}$$

, where g_1 and g_2 are the conditional bid densitites at the α-th quantile for I_1 and $I_2 < I_1$ bidder auctions. The first line of equation 12 says that a bidder with value $v(\alpha)$ is indifferent between winning an I_1 bidder auction for x_1 units and winning an I_2 bidder auction for x_2 units. The second line of equation 12 expresses the condition solely in terms of observables. We can solve equation 12 for $v(\alpha)$ as follows:

$$v(\alpha) = [x_2 b_2(\alpha|x_2) - x_1 b_1(\alpha|x_1)] / (x_2 - x_1)$$

Therefore F is identified.

Now fix $\alpha_0 \in (0,1]$. If $[\underline{x}, \overline{x}]$ is sufficiently wide then $\overline{x}/\underline{x} > \sup_{\alpha>0} [\beta_2(\alpha)/\beta_1(\alpha)]$.[24]

[24]Using l'Hï¿œpital's Rule we obtain $\lim_{\alpha\to 0} \beta_2(\alpha)/\beta_1(\alpha) = \lim_{\alpha\to 0} \beta_2'(\alpha)/\beta_1'(\alpha) = \left[1 - s_2'(\underline{v})\right] / \left[1 - s_1'(\underline{v})\right]$ and we know that $1 > s_2'(\underline{v}) > s_1'(\underline{v})$.

Therefore for all α

$$\overline{x}\left[v\left(\alpha\right)-b_{I}\left(\alpha,\bar{x}\right)\right]>\underline{x}\left[v\left(\alpha\right)-b_{I}\left(\alpha,\underline{x}\right)\right]$$

The bid shading goes to zero as α goes to zero so we can recursively define a decreasing α sequence such that

$$\overline{x}\left[v\left(\alpha_{t+1}\right)-b_{I}\left(\alpha_{t+1},\bar{x}\right)\right]=\underline{x}\left[v\left(\alpha_{t}\right)-b_{I}\left(\alpha_{t},\underline{x}\right)\right]$$

If $[\underline{x},\overline{x}]$ is sufficiently wide then there is $c<1$ such that

$$v\left(\alpha_{t+1}\right)-b_{I}\left(\alpha_{t+1},\underline{x}\right)<c^{t+1}\left(v\left(\alpha_{0}\right)-b_{I}\left(\alpha_{0},\underline{x}\right)\right)$$

Therefore the per-unit bid shading converges to zero as t becomes large. Define for $z\geq 0$

$$\lambda_{t}\left(z\right)=\lambda\left(z\right)\frac{f^{b}\left(v\left(\alpha_{t}\right)\right)}{F^{b}\left(v\left(\alpha_{t}\right)\right)f\left(v\left(\alpha_{t}\right)\right)}$$

The functions $\lambda_{t}\left(z\right)$ are identified for

$$z\in\left[\min_{x\in[\underline{x},\overline{x}]}x\left(v\left(\alpha_{t}\right)-b_{I}\left(\alpha_{t},x\right)\right),\max_{x\in[\underline{x},\overline{x}]}x\left(v\left(\alpha_{t}\right)-b_{I}\left(\alpha_{t},x\right)\right)\right] \tag{13}$$

, because

$$\lambda_{t}\left(x\left(v\left(\alpha_{t}\right)-b_{I}\left(\alpha_{t},x\right)\right)\right)=\frac{x}{g\left(b_{I}\left(\alpha_{t},x\right),x\right)\left(I-1\right)}$$

The two consecutive functions λ_t and λ_{t+1} differ only by a multiplicative constant and have overlapping identification regions (13). Therefore the constant can be identified and the identification region of λ_t can be extended to the identification region of λ_{t+1}.

By choosing t sufficiently large λ_0 is therefore identified on

$$z \in \left(0, \max_{x \in [\underline{x}, \overline{x}]} x \left(v\left(\alpha_0\right) - b_I\left(\alpha_0, x\right)\right)\right] \tag{14}$$

As we also know that $\lambda_0(0) = 0$ we can uncover $\lambda_0^{'}(0)$. To identify λ we exploit the form of the first-order condition 11 at the lower bound. As $g(\underline{v}, x) \frac{\partial}{\partial v} s(\underline{v}, x) = f(\underline{v})$ we can back out $\lambda'(0)$. We can now solve for

$$\frac{f^b\left(v\left(\alpha_0\right)\right)}{F^b\left(v\left(\alpha_0\right)\right) f\left(v\left(\alpha_0\right)\right)} = \lambda_0^{'}(0) / \lambda^{'}(0)$$

Hence λ is identified on the region (14). Lastly, choose $I = I_1$ and vary α_0 so λ is identified on

$$z \in \left[0, \max_{x \in [\underline{x}, \overline{x}], \alpha \in [0,1]} x\left(v(\alpha) - b_1(\alpha, x)\right)\right] \tag{15}$$

As $U(1)$ is normalized to 1 we can solve $\lambda(z) = U(z)/U'(z)$ for U. The assumption that $[\underline{x}, \overline{x}]$ is wide enough ensures that $\lambda(1)$ is identified.

Identification of $F^b(\cdot|v \leq \overline{v})$ follows from the same argument as in the risk neutral case. The first order condition is a differential equation with boundary condition $F^b(\overline{v}|v \leq \overline{v}) = 1$ which can be solved for $F^b(v|v \leq \overline{v})$ for all $v \in [\underline{v}, \overline{v}]$.

C Estimation and Inference

Let $\theta = \left(f, f^b, \{f_I^u\}_{I \in \mathbf{N}}\right)$ where f, f^b and f_I^u are densities of the private values, the belief and an unobserved characteristic. The true parameter $\theta_0 = \left(f_0, f_0^b, \left\{f_{0,n}^u\right\}_{n \in \mathbf{N}}\right)$ is assumed to live in a known parameter space Θ. The data available contain L_I I-bidder auctions and L auctions overall. Let I_ℓ and $\mathbf{b}_\ell = (b_{1,\ell}, b_{2,\ell}, \cdots, b_{I_\ell,\ell})$ denote the number of bidders and the vector of all the bids in the ℓ-th auction. The econometrician observes $(\mathbf{y}_\ell, X_\ell, I_\ell)$, where $\mathbf{y}_\ell = (y_{1,\ell}, y_{2,\ell}, \cdots, y_{I_\ell,\ell})$ and X_ℓ is a vector of observable auction characteristics. From $\mathbf{y}_\ell$ and X_ℓ, the econometrician can obtain an estimate of $\mathbf{b}_\ell$ denoted by $\hat{\mathbf{b}}_\ell = \left(\hat{b}_{1,\ell}, \hat{b}_{2,\ell}, \cdots, \hat{b}_{I_\ell,\ell}\right)$. This formulation allows for cases in which data are generated in a first-stage estimation. For example, in our application the bids are generated in a first-stage regression to remove the effect of observed characteristics. Krasnokutskaya and Seim (2011) showed that if the unobserved characteristics enters a bidder's value in a multiplicative fashion it also enters the bids multiplicatively. A bidder with a value uv^* bids:

$$u \times s_I^*(v^*, \theta) = u \times \left(v^* - \int_{\underline{v}^*}^{v^*} \left[\frac{F^b(s)}{F^b(v^*)}\right]^{I-1} ds\right).$$

C.1 Estimation

Our estimator extends the simulated integrated moments estimator proposed by Bierens and Song (2012). For a given θ, the model generates joint distributions of bids. Under the true parameter, the model implied joint distributions should be close to the data. Therefore, an estimate of θ_0 can be obtained by matching the model implied distributions with the data. Rather than directly matching the joint bid distributions, we match the joint characteristic functions because they are easier to compute as we explain later.

It suffices to focus on characteristic functions of two randomly selected log bids from the same auction (Krasnokutskaya (2011), Kotlarski (1967)). Let $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$ and $\Psi_I(\theta, \mathbf{t})$ be the model implied joint characteristic function in an I-bidder auction under θ and evaluated at $\mathbf{t}$. Let

$$\psi_{I_\ell}(\mathbf{b}_\ell, \mathbf{t}) = \frac{1}{I_\ell(I_\ell - 1)} \sum_{i \neq j} \exp\left[\mathbf{i}t_1 \log b_{i,\ell} + \mathbf{i}t_2 \log b_{j,\ell}\right].$$

$E\left[\psi_{I_\ell}\left(\mathbf{b}_\ell, \mathbf{t}\right)\mid I_\ell = I\right]$ is the population joint characteristic function of two log bids in an auction with I bidders. Notice that ψ_{I_ℓ} uses all the bids in an auction. The model restrictions imply the following moment conditions.

$$\Psi_I\left(\theta_0, \mathbf{t}\right) = E\left[\psi_{I_\ell}\left(\mathbf{b}_\ell, \mathbf{t}\right)\mid I_\ell = I\right] \ \forall I \in \mathbf{N}, \mathbf{t} \in \mathbf{T} \tag{16}$$

$\mathbf{T}$ is a compact subset of $\mathbf{R}^2$ on which we match the model implied characteristic functions with the data. It contains 0 in its interior. We choose $\mathbf{T} = \left[-\kappa, \kappa\right]^2$ where κ is some positive constant.

(16) can be written as $Q\left(\theta_0\right) = 0$ where Q is the population criterion function defined as

$$Q\left(\theta\right) = \sum_{I \in \mathbf{N}} \alpha_I \int_{\mathbf{T}} \left|\Psi_I\left(\theta, \mathbf{t}\right) - \Psi_I\left(\theta_0, \mathbf{t}\right)\right|^2 d\mu\left(\mathbf{t}\right). \tag{17}$$

Here the α_Is are positive constants that add up to 1. μ is a known probability measure on $\mathbf{T}$ which weights moment conditions in (16). For our application, any μ with strictly positive density on $\mathbf{T}$ suffices.

Our estimator is based on a sample analogue of (17). To construct it, we first replace $E\left[\psi_{I_\ell}\left(\mathbf{b}_\ell, \mathbf{t}\right)\mid I_\ell = I\right]$ with its sample average

$$\hat{\psi}_I\left(\mathbf{t}\right) = \frac{1}{L_I} \sum_{\ell}^{L} \mathbf{1}\left(I_\ell = I\right) \psi_{I_\ell}\left(\hat{\mathbf{b}}_\ell, \mathbf{t}\right).$$

$\Psi_I\left(\theta, \mathbf{t}\right)$ is a complicated function of θ which does not have a closed form. A natural way to compute it is by simulation. It is computationally burdensome to simulate $\Psi_I\left(\theta, \cdot\right)$ directly because it is two-dimensional function. We decompose $\Psi_I\left(\theta, \cdot\right)$ into two one-dimensional characteristic functions as follows:

$$\begin{aligned}
\Psi_I(\theta,\mathbf{t}) &= E\left[\psi_{I_\ell}(\mathbf{b}_\ell,\mathbf{t})\middle| I_\ell = I\right] \\
&= E\exp\left[\mathbf{i}t_1\log s_I^*(v_1^*,\theta)+\mathbf{i}t_2\log s_I^*(v_2^*,\theta)+\mathbf{i}(t_1+t_2)\log u\right] \\
&= E\left\{\exp\left[\mathbf{i}t_1\log s_I^*(v_1^*,\theta)\right]\exp\left[\mathbf{i}t_2\log s_I^*(v_2^*,\theta)\right]\exp\left[\mathbf{i}(t_1+t_2)\log u\right]\right\} \\
&= E\exp\left[\mathbf{i}t_1\log s_I^*(v_1^*,\theta)\right]E\exp\left[\mathbf{i}t_2\log s_I^*(v_2^*,\theta)\right]E\exp\left[\mathbf{i}(t_1+t_2)\log u\right] \quad (18)
\end{aligned}$$

The second equality of (18) exploits the separable form of the equilibrium bidding strategy. The last equality holds because v_1^*, v_2^* and u are independent. Notice that $E\exp\left[\mathbf{i}t_1\log s_I^*(v_1^*,\theta)\right]$ and $E\exp\left[\mathbf{i}t_2\log s_I^*(v_2^*,\theta)\right]$ are the same characteristic function evaluated at different points. Therefore, we only need to simulate two one-dimensional characteristic functions which is much easier. It is worth noting that joint bid distributions do not allow a similar decomposition and are therefore more difficult to simulate.

Next, we discuss how to simulate $E\exp\left[\mathbf{i}t\log s_I^*(v^*,\theta)\right]$. $E\exp\left[\mathbf{i}t\log u\right]$ can be simulated in a similar fashion.

(i). Randomly draw ϵ_i from a standard uniform distribution.

(ii). Compute $v_i^* = \inf\{v: \ F(v) > \epsilon_i\}$ where F is the CDF of f.

(iii). Compute the corresponding bids using the bidding strategy

$$s_I^*(v^*,\theta) = v^* - \int_{\underline{v}^*}^{v^*}\left[\frac{F^b(s)}{F^b(v^*)}\right]^{I-1} ds.$$

(iv). Repeat 1-3 for D_L times and approximate $E\exp\left[\mathbf{i}t\log s_I^*(v^*,\theta)\right]$ with $\frac{1}{D_L}\sum_{i=1}^{D_L}\exp\left(\mathbf{i}ts_I^*(v_i^*,\theta)\right)$.

Let $\bar{\Psi}_I$ be the simulated characteristic function in I-bidder auctions. Define the sample criterion as

$$\widehat{Q}_L(\theta) = \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left| \bar{\Psi}_I(\theta, \mathbf{t}) - \hat{\psi}_I(\mathbf{t}) \right|^2 d\mu(\mathbf{t}). \tag{19}$$

Here $\alpha_{I,L}$ are some (stochastic) weights with $\sum_{I \in \mathbf{N}} \alpha_{I,L} = 1$ and $\alpha_{I,L}$ converges in probability to α_I for every I.

Let $\left\{\Theta_{k(L)}\right\}_{L=1}^{\infty}$ be a sequence of sieve spaces that approximate Θ as $L \to \infty$. Our estimator is defined as the minimizer of the above sample criterion on the sieve space.

$$\hat{\theta}_L = \arg \max_{\theta \in \Theta_{k(L)}} \widehat{Q}_L(\theta).$$

C.1.1 Consistency

Define the sup-norm on Θ as $\|\theta\|_\infty = \max\left\{\|f\|_\infty, \left\|f^b\right\|_\infty, \{\|f_I^u\|_\infty\}_{I \in \mathbf{N}}\right\}$ with $\|f\|_\infty = \max_x |f(x)|$

Assumption 7. *The parameter space and the sieve space satisfy:*

(i). Θ *is compact under* $\|\cdot\|_\infty$.

(ii). $\Theta \subset \mathcal{F}(\eta)^{2+|\mathbf{N}|}$ *where* $|\mathbf{N}|$ *is the cardinality of the set* $\mathbf{N}$ *.* $\mathcal{F}(\eta)$ *is the set of continuously differentiable density functions with bounded positive interval support. The lower bound of the support is no less than* $\eta > 0$ *. In addition, the density functions in* $\mathcal{F}(\eta)$ *are bounded by an integrable function* $\mathcal{E}$.

(iii). There exists $\theta_{0,L} \in \Theta_{k(L)}$ *such that* $Q(\theta_{0,L}) = o(1)$ *and* $\|\theta_{0,L} - \theta_0\|_\infty = o(1)$ *as* $L \to \infty$.

(iv). $\sup_{(\theta,\mathbf{t}) \in \Theta_{k(L)} \times \mathbf{T}} \left|\Psi_I(\theta, \mathbf{t}) - \bar{\Psi}_I(\theta, \mathbf{t})\right| = o_p(1)$ for all $I \in \mathbf{N}$ as $L \to \infty$.

(v). $\sup_\ell \left\|\hat{\mathbf{b}}_\ell - \mathbf{b}_\ell\right\|_E \xrightarrow{p} 0$ as $L \to \infty$ where $\|\cdot\|_E$ is the Euclidean norm.

Assumption 7(i) ensures that the parameter space is compact. It rules out the possibility of inconsistency due to the well-known ill-posed inverse problem. Assumption 7(ii) guarantees that θ_0 is point-identified and the population criterion function $Q(\theta)$ is continuous in θ. Assumption 7(iii) states that the sieve space approximates θ_0 well enough. Assumption 7(iv) requires that the simulated moments approximate the true moments well on the sieve space. It is guaranteed if a large number of simulation draws is used. Assumption 7(v) requires that the error coming from the first-stage estimation is negligible as the sample size gets large. In our application, it is satisfied if the covariates have enough variation and are bounded away from 0 and ∞.

Theorem 5. *Under Assumption 7,* $\left\|\hat{\theta}_L - \theta_0\right\|_\infty = o_p(1)$ *as* $L \to \infty$.

Proof. We first establish several facts.

Fact 1: $\Psi_I(\theta, \mathbf{t})$ is continuous in θ under $\|\cdot\|_\infty$ for every $\mathbf{t}$. To see this, notice by Lemma F.2 from Grundl and Zhu (2015), $s_I^*(v, \theta)$ is continuous in θ given v. By Assumption 7(ii),

$$E \exp\left[\mathbf{i}t \log s_I^*(v^*, \theta)\right] = \int \exp\left[\mathbf{i}t \log s_I^*(v^*, \theta)\right] f(v^*)\, dv^*$$

has an integrand which is continuous in θ and bounded by the integrable function $\mathcal{E}(x)$. The dominance convergence theorem implies that $E \exp\left[\mathbf{i}t \log s_I^*(v^*, \theta)\right]$ is continuous in θ for each t. Similarly, we can establish the continuity of $E \exp\left[\mathbf{i}t \log u\right]$ with respect to θ. Therefore, $\Psi_I(\theta, \mathbf{t})$ is continuous in θ for every $\mathbf{t}$.

Fact 2: Let $\widetilde{\psi}_I(\mathbf{t}) = \frac{1}{L_I} \sum_\ell^L \mathbf{1}(I_\ell = I)\, \psi_{I_\ell}(\mathbf{b}_\ell, \mathbf{t})$.

$$\sup_{\mathbf{t} \in \mathbf{T}} \left|\hat{\psi}_I(\mathbf{t}) - \widetilde{\psi}_I(\mathbf{t})\right| \leq C \frac{1}{L_I} \sum_\ell^L \mathbf{1}(I_\ell = I) \frac{1}{I_\ell(I_\ell - 1)} \sum_{i \neq j} \left|b_{i,\ell} - \hat{b}_{i,\ell}\right| \xrightarrow{p} 0$$

for some constant C by Assumption 7(v).

Fact 3: $\sup_{\mathbf{t}\in\mathbf{T}}\left|\widetilde{\psi}_I(\mathbf{t})-\Psi_I(\theta_0,\mathbf{t})\right|=o_p(1)$. To see this, notice that

$$\mathrm{Re}\widetilde{\psi}_I(\mathbf{t})=\frac{1}{L_I}\sum_{\ell}^{L}\mathbf{1}(I_\ell=I)\,\mathrm{Re}\psi_{I_\ell}(\mathbf{b}_\ell,\mathbf{t})$$

and $\mathrm{Re}\psi_{I_\ell}(\mathbf{b}_\ell,\mathbf{t})=\frac{1}{I_\ell(I_\ell-1)}\sum_{i\neq j}\cos[t_1\log b_{i,\ell}+t_2\log b_{j,\ell}]$, where Re denotes the real part. $\theta_0\in\Theta$ implies that private values and the unobserved characteristic are both bounded from below by η and from above by some constant $M<\infty$. Hence, $b_{i,\ell}$ are no smaller than η^2 and no larger than M^2. There exists a constant C independent of $\mathbf{b}_\ell$ and $\mathbf{t}$, such that

$$\left|\mathrm{Re}\psi_{I_\ell}(\mathbf{b}_\ell,\mathbf{t}_1)-\mathrm{Re}\psi_{I_\ell}(\mathbf{b}_\ell,\mathbf{t}_2)\right|\leq C\max\{2|\log M|,2|\log\eta|\}\,\|\mathbf{t}_1-\mathbf{t}_2\|_E.$$

Then by Theorem 2.7.11 from van der Vaart and Wellner (2000), $\mathrm{Re}\psi_I(\mathbf{b},\mathbf{t})$ indexed by $\mathbf{t}$ is a Glivenko-Cantelli class. Hence, $\sup_{\mathbf{t}\in\mathbf{T}}\left|\mathrm{Re}\widetilde{\psi}_I(\mathbf{t})-\mathrm{Re}\Psi_I(\theta_0,\mathbf{t})\right|=o_p(1)$. Similarly, the imaginary part satisfies the same condition. This establishes Fact 3.

Now we are ready to establish the consistency. By Fact 1, $Q(\theta)$ is continuous in θ. Because θ_0 is point identified and Θ is compact, for any $\epsilon>0$, there must exist an $\delta>0$ such that $\min_{\theta\in\Theta:\|\theta-\theta_0\|_\infty\geq\epsilon}Q(\theta)\geq\delta$. By Assumption 7(iv), Fact 2 and Fact 3,

$$\begin{aligned}
&\sup_{\theta\in\Theta_{k(L)}}\left|\widehat{Q}_L(\theta)-Q(\theta)\right|\\
\leq\;&\sup_{\theta\in\Theta_{k(L)}}\sum_{I\in\mathbf{N}}4\alpha_I\int_{\mathbf{T}}\left|\left[\Psi_I(\theta,\mathbf{t})-\bar{\Psi}_L(\theta,\mathbf{t})\right]-\left[\Psi_I(\theta_0,\mathbf{t})-\hat{\psi}_I(\mathbf{t})\right]\right|d\mu(\mathbf{t})+4\sum_{I\in\mathbf{N}}|\alpha_{I,L}-\alpha_I|\\
\leq\;&\sup_{\theta\in\Theta_{k(L)}}\sum_{I\in\mathbf{N}}4\alpha_I\int_{\mathbf{T}}\left|\left[\Psi_I(\theta,\mathbf{t})-\bar{\Psi}_L(\theta,\mathbf{t})\right]\right|d\mu(\mathbf{t})+o_p(1)=o_p(1).
\end{aligned}\qquad(20)$$

Because $\hat{\theta}_L$ is the minimizer of $\widehat{Q}_L(\theta)$ on $\Theta_{k(L)}$ and $\theta_{0,L}\in\Theta_{k(L)}$, $\widehat{Q}_L\left(\hat{\theta}_L\right)-\widehat{Q}_L(\theta_{0,L})<$

0. This fact together with (20) and Assumption 7(iii) implies that

$$\begin{aligned}
& P\left(\left\|\hat{\theta}_L-\theta_0\right\|_{\infty} \geq \epsilon\right)=P\left(\left\|\hat{\theta}_L-\theta_0\right\|_{\infty} \geq \epsilon\right) \leq P\left(Q\left(\hat{\theta}_L\right) \geq \delta\right) \\
= \; & P\left(Q\left(\hat{\theta}_L\right)-\widehat{Q}_L\left(\hat{\theta}_L\right)+\widehat{Q}_L\left(\hat{\theta}_L\right)-\widehat{Q}_L\left(\theta_{0,L}\right)+\widehat{Q}_L\left(\theta_{0,L}\right)-Q\left(\theta_{0,L}\right)+Q\left(\theta_{0,L}\right) \geq \delta\right) \\
= \; & P\left(o_p(1)+\widehat{Q}_L\left(\hat{\theta}_L\right)-\widehat{Q}_L\left(\theta_{0,L}\right)+o_p(1)+o(1) \geq \delta\right) \\
= \; & P\left(o_p(1) \geq \delta\right) \rightarrow 0
\end{aligned}$$

□

C.2 Inference

We provide a bootstrap procedure to test the hypothesis that bidders have correct beliefs:

$$H_0: \; f_0^b = f_0 \text{ vs } \; H_1: f_0^b \neq f_0.$$

This procedure adapts the bootstrap statistic proposed in Zhu and Grundl (2014) to our setup. The key difference is that the null hypothesis is defined by equalities instead of inequalities as in Zhu and Grundl (2014). Hence, we can construct a different bootstrap statistic which has less tuning parameters and does not diverge under the alternative.

The test statistic is the minimum of our sample criterion on the sieve space under the constraint that $f^b = f$, i.e.

$$T_L = \min_{\theta \in \Theta_{k(L)} \cap R} L\widehat{Q}_L(\theta)$$

with $R = \left\{\theta \in \Theta : f^b = f\right\}$. Let the minimizer be $\widetilde{\theta}_L$. Rejection occurs if T_L is large enough. The critical value can be constructed by bootstrap. Suppose we obtain bootstrap bids $\hat{\mathbf{b}}^*$ from the original sample. Let $\hat{\psi}_I^*(\mathbf{t}) = \frac{1}{L_I}\sum_{\ell}^{L} \mathbf{1}(I_\ell = I)\, \psi_{I_\ell}\left(\hat{\mathbf{b}}_\ell^*, \mathbf{t}\right)$,

$\mathbb{G}^*_{I,L}(\cdot)$ be the law of $\sqrt{L}\left(\hat{\psi}^*_I(\cdot) - \hat{\psi}_I(\cdot)\right)$ under the empirical measures and $B^{\sigma_L}\left(\widetilde{\theta}_L\right)$ be the σ_L closed neighborhood of $\widetilde{\theta}_L$. The bootstrap statistic is defined as

$$T^*_L = \min_{\theta \in \Theta_{k(L)} \cap R \cap B^{\sigma_L}\left(\widetilde{\theta}_L\right)} \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left| \mathbb{G}^*_{I,L}(\mathbf{t}) - \sqrt{L}\left(\bar{\Psi}_I(\theta, \mathbf{t}) - \bar{\Psi}_I\left(\widetilde{\theta}_L, \mathbf{t}\right)\right)\right|^2 d\mu(\mathbf{t})$$

The critical value for an α significance level test is the $1-\alpha$-th quantile of T^*_L denoted by $c(1-\alpha)$.

Now we provide a justification of this testing procedure. We begin with introducing some notation and assumptions. Let $\mathbb{G}_{I,L}(\cdot)$ be the law of $\sqrt{L}\left(\hat{\psi}_I(\cdot) - \Psi_I(\theta_0, \cdot)\right)$ under the true parameter and $\mathbb{G}_I(\cdot)$ be its limit. Define the directional derivative at θ_1 in the direction of θ_2 as

$$\frac{d\Psi_I(\theta_1, \mathbf{t})}{d\theta}[\theta_2 - \theta_1] = \left.\frac{d\Psi_I(\theta_1 + \tau(\theta_2 - \theta_1), \mathbf{t})}{d\tau}\right|_{\tau=0}.$$

Assumption 8. *(i). θ_0 lives in the interior of Θ.*

(ii). $\mathbb{G}^*_{I,L}(\cdot)$ converges in law to $\mathbb{G}_I(\cdot)$ as $L \to \infty$ almost surely.

(iii). $\max_{\theta_1 \in \Theta} \min_{\theta_2 \in \Theta_{k(L)}} \|\theta_1 - \theta_2\|_\infty = o\left(1/\sqrt{L}\right)$.

(iv). $\sup_{(\theta,\mathbf{t}) \in \Theta_{k(L)} \times \mathbf{T}} \left|\Psi_I(\theta, \mathbf{t}) - \bar{\Psi}_I(\theta, \mathbf{t})\right| = o_p\left(1/\sqrt{L}\right)$ for all $I \in \mathbf{N}$.

(v). $\frac{d\Psi_I(\theta_1,\mathbf{t})}{d\theta}[\theta_2 - \theta_1]$ exists and is uniformly bounded for all $\theta_1, \theta_2 \in \Theta$ and $\mathbf{t} \in \mathbf{T}$.

(vi). There exists $\epsilon > 0$ such that $\forall \theta_1, \theta_2, \theta_3, \theta_4 \in B^\epsilon(\theta_0)$,

$$\left|\frac{d\Psi_I(\theta_1, \mathbf{t})}{d\theta}[\theta_3 - \theta_4] - \frac{d\Psi_I(\theta_2, \mathbf{t})}{d\theta}[\theta_3 - \theta_4]\right| \le C\|\theta_3 - \theta_4\|_\infty \|\theta_2 - \theta_1\|^\omega_\infty$$

for some constant C and $\omega > 0$.

Assumption 8(i) requires that the true parameter lies in the interior of the parameter space. This is without loss of generality. Section C.3 shows a way to transform the parameter space to a subset of a linear space. After the transformation, the true parameter lies in the interior of the transformed parameter space. Assumption 8(ii) requires that the empirical process based on the bootstrap sample converges in law to the limiting process $\mathbb{G}_I(\cdot)$. This is satisfied in our application if we redraw $\mathbf{y}$ and X in pairs and generate bids using the bootstrap sample. Assumption 8(iii) requires that the sieve spaces approximate the parameter space well enough. This can be guaranteed by first restricting the parameter space to the set of functions that are smooth enough and then choosing the sieve space correspondingly, see Newey (1997). Assumption 8(iv) says that the error coming from simulating the moments is negligible compared to the sampling error. It is satisfied if the number of simulation draws is large enough. Assumption 8(v) and Assumption 8(vi) put smoothness restrictions on the characteristic functions. They guarantee that the difference in characteristic functions at different θ can be approximated by a first-order expansion.

Theorem 6. *Under Assumption 8 and the null hypothesis, if* $\left\|\widetilde{\theta}_L - \theta_0\right\|_\infty = o_p\left(\frac{1}{\log L^{1/2\omega}}\right)$ *and* $\sigma_L = O\left(\sqrt{\frac{\log L}{L}}\right)$, *then* $\limsup_{L\to\infty} P\left(T_L > c\left(1-\alpha\right)\right) \leq \alpha$. *Under the alternative,* $P\left(T_n > c\left(1-\alpha\right)\right) \to 1$ *as* $L \to \infty$.

Proof. Under the null, by Lemma 3.1 in Zhu and Grundl (2014) and the fact that θ_0 is point identified

$$T_L \leq \min_{\theta \in B^{\delta_L}(\theta_0)\cap\Theta_{k(L)}\cap R} \sum_{I\in\mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left|\mathbb{G}_I\left(\mathbf{t}\right) - \sqrt{L}\frac{d\Psi_I\left(\theta_{0,L},\mathbf{t}\right)}{d\theta}\left[\theta - \theta_{0,L}\right]\right|^2 d\mu\left(\mathbf{t}\right) + o_p\left(1\right)$$

where $\theta_{0,L}$ is the projection of θ_0 onto $\Theta_{k(L)}$ and $\delta_L = o\left(L^{-1/2(1+\omega)}\right)$ is a sequence of positive numbers such that $\sigma_L = o\left(\delta_L\right)$. This choice of δ_L guarantees that we can use a first-order expansion around $\theta_{0,L}$ to approximate $\Psi_I(\theta, \mathbf{t})$ if $\theta \in B^{\delta_L}(\theta_0)$. Similarly,

$$T_L^* = \min_{\theta \in \Theta_{k(L)} \cap R \cap B^{\sigma_L}\left(\widetilde{\theta}_L\right)} \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left| \mathbb{G}_I(\mathbf{t}) - \sqrt{L} \frac{d\Psi_I\left(\widetilde{\theta}_L, \mathbf{t}\right)}{d\theta} \left[\theta - \widetilde{\theta}_L\right] \right|^2 d\mu(\mathbf{t}) + o_{p^*}(1).$$

Here p^* denotes the probability under empirical measures. For any $\theta \in \Theta_{k(L)} \cap R \cap B^{\sigma_L}\left(\widetilde{\theta}_L\right)$, $\theta_1 = \theta_{0,L} + \left(\theta - \widetilde{\theta}_L\right) \in B^{\delta_L}(\theta_0) \cap \Theta_{k(L)} \cap R$. To see this, notice that $\left\|\theta - \widetilde{\theta}_L\right\|_\infty = O(\sigma_L)$ and $\|\theta_{0,L} - \theta_0\|_\infty = o\left(1/\sqrt{L}\right)$ by Assumption 8(iii). Therefore, $\|\theta_1 - \theta_0\|_\infty = O(\sigma_L) + o\left(1/\sqrt{L}\right) = o(\delta_L)$. For such θ_1, Assumption 8(vi), $\left\|\widetilde{\theta}_L - \theta_0\right\|_\infty = o_p\left(1/\log L^{1/2\omega}\right)$ and $\sigma_L = O\left(\sqrt{\log L/L}\right)$ imply

$$\left| \sqrt{L} \frac{d\Psi_I(\theta_{0,L}, \mathbf{t})}{d\theta} [\theta_1 - \theta_{0,L}] - \sqrt{L} \frac{d\Psi_I\left(\widetilde{\theta}_L, \mathbf{t}\right)}{d\theta} \left[\theta - \widetilde{\theta}_L\right] \right| \leq C\sqrt{L}\sigma_L \left\|\widetilde{\theta}_L - \theta_{0,L}\right\|_\infty = o_p(1).$$

This suggests that for any $\theta \in \Theta_{k(L)} \cap R \cap B^{\sigma_L}\left(\widetilde{\theta}_L\right)$, we can find an $\theta_1 \in B^{\delta_L}(\theta_0) \cap \Theta_{k(L)} \cap R$ such that

$$\begin{aligned} & \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left| \mathbb{G}_I(\mathbf{t}) - \sqrt{L} \frac{d\Psi_I(\theta_{0,L}, \mathbf{t})}{d\theta} [\theta_1 - \theta_{0,L}] \right|^2 d\mu(\mathbf{t}) \\ = & \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left| \mathbb{G}_I(\mathbf{t}) - \sqrt{L} \frac{d\Psi_I\left(\widetilde{\theta}_L, \mathbf{t}\right)}{d\theta} \left[\theta - \widetilde{\theta}_L\right] \right|^2 d\mu(\mathbf{t}) + o_p(1). \end{aligned}$$

Therefore, we must have $T_L^* > T_L + o_p(1)$ which implieslim $\sup_{L \to \infty} P(T_L > c(1-\alpha)) \leq \alpha$. Under the alternative, the population criterion function is positive on the restriction set R. Hence, $T_L \to \infty$ by Lemma 3.1 from Zhu and Grundl (2014). But $T^* \leq \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} |\mathbb{G}_I(\mathbf{t})|^2 d\mu(\mathbf{t}) + o_{p^*}(1)$ is bounded. Consequently, $P(T_n > c(1-\alpha)) \to 1$ as $L \to \infty$. □

The additional assumption required for size control is that under the null, $\widetilde{\theta}_L$

converges to θ_0 at a rate faster than $\log L^{1/2\omega}$. Under very mild conditions, the converges rate will be polynomial and therefore faster than the log rate.

The bootstrap critical value is valid if $\mathbb{G}^*_{I,L}(\mathbf{t})$ converges in law to $\mathbb{G}_I$ almost surely. This condition can be guaranteed by a pairwise bootstrap where$\mathbf{y}$ and X are treated as pairs. Each time we redraw an auction, we redraw a pair of $\mathbf{y}$ and X. The following is a step-by-step guide.[25]

(i). Randomly redraw L_I I-bidder auctions from the original sample with replacement for each $I \in \mathbf{N}$ to form a bootstrap sample.

(ii). Generate $\hat{\mathbf{b}}^*$ based on the bootstrap sample using the first-step regression.

(iii). Compute

$$t^*_b = \min_{\theta \in \Theta_{k(L)} \cap R \cap B^{\sigma_L}\left(\widetilde{\theta}_L\right)} L \sum_{I \in \mathbf{N}} \alpha_{I,L} \int_{\mathbf{T}} \left| \left(\hat{\psi}^*_I(\mathbf{t}) - \hat{\psi}_I(\mathbf{t}) \right) - \left(\bar{\Psi}_I(\theta, \mathbf{t}) - \bar{\Psi}_I\left(\widetilde{\theta}_L, \mathbf{t}\right) \right) \right|^2 d\mu(\mathbf{t})$$

(iv). Repeat 1-3 for B times and collect all t^*_b the for $b = 1, 2, \cdots, B$. Then the critical value $c(1-\alpha)$ can be estimated by $\hat{c}(1-\alpha) = \inf\left\{x : \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\left(t^*_b > x\right) < \alpha\right\}$.

C.3 Implementation

Sieve Spaces We assume that f^b_0 and f_0 share the same support and $f^u_{0,I}$ has the same support for all I. In addition, all the densities have compact interval support with densities bounded away from 0 on their support. If $\delta > 0$, any densities with support $[0, \delta]$ can be expressed as $g(\delta x)/\delta$ where g is some density function with support $[0, 1]$. Notice that there is a unique $r(x) \geq -1$ supported on $[0, 1]$ with $\int_0^1 r(x)\, dx = 0$ and $\int_0^1 r(x)^2\, dx < \infty$ such that $g = \frac{[1+r(x)]^2}{1+\int_0^1 r(s)^2 ds}$. We follow Bierens and Song (2012) to approximate $r(x)$ with Legendre polynomials $\sum_{i=1}^{k_L} \phi_i(x)\beta_i$ where

[25] A standard argument can show that $\mathbb{G}^*_{I,L}(\mathbf{t})$ converges in law to $\mathbb{G}_I$ almost surely for the pairwise bootstrap. Hence, the proof is omitted.

ϕ_i is the i-th order polynomial. We exclude the constant function ϕ_0 to guarantee that $\int_0^1 r(x)\,dx = 0$. The lower bound of the support of $f_{0,I}^u$ is normalized to 1. Therefore, our estimation problem reduces to estimating the coefficients β_i, the lower bound of v^*, the length of the support for f_0^* and the length of the support for $f_{0,n}^u$. In the application, we use 5-th order Legendre polynomials for each density function.

Notice that we have transformed the parameter into $\left(\underline{v}, \overline{v}, \bar{u}, r, r^b, \{r_I\}_{I\in\mathbf{N}}\right)$. Here $\underline{v}, \overline{v}$ are the lower and upper bound of v^*, $\bar{u}$ is the length of the support of f_I^u, and r, r^b and r_I are all functions defined on $[0,1]$ which are greater or equal than -1. In addition, their integrals on $[0,1]$ are 0. Being in the interior of the parameter space means that the r functions are strictly greater than -1, $\underline{v} < \overline{v}$ and $\bar{u} > 0$. This is equivalent to requiring that the densities are strictly positive on their support. In addition, the parameter space after the transformation is a subset of a linear space.

Simulation Draws D_L must go to infinity sufficiently fast to make the simulation error negligible. This can be ensured if $D_L = L^{1+a}$ for some $a > 0$. We choose $a = 0.5$.

Weights and σ_L We choose $\alpha_{I,L}$ to be the proportion of bids from auctions with I bidders and $\sigma_L = \sqrt{\log L/L}$.

Probability Measure We choose μ to be uniform and $\mathbf{T} = [-8, 8]^2$. The integral is approximated by the average on 1225 evenly spaced points in $\mathbf{T}$.

References

Aguirregabiria, V., and A. Magesan (2012): "Identification and estimation of dynamic games when players' beliefs are not in equilibrium," *Available at SSRN 2117055.* 4

An, Y., Y. Hu, and M. Shum (2010): "Estimating first-price auctions with an unknown number of bidders: A misclassification approach," *Journal of Econometrics*, 157(2), 328–341. 8

Aradillas-López, A., A. Gandhi, and D. Quint (2013): "Identification and inference in ascending auctions with correlated private values," *Econometrica*, 81(2), 489–534. 3, 8, 17

Aryal, G., and D.-H. Kim (2014): "Empirical Relevance of Ambiguity in First Price Auction Models," Discussion paper, Working Paper. 6, 7

Athey, S., and P. Haile (2007): "Nonparametric approaches to auctions," *Handbook of Econometrics*, 6, 3847–3965. 7

Athey, S., J. Levin, and E. Seira (2011): "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions," *The Quarterly Journal of Economics*, 126(1), 207. 8, 17

Baldwin, L., R. Marshall, and J. Richard (1997): "Bidder Collusion at Forest Service Timber Sales," *The Journal of Political Economy*, 105(4), 657–699. 15

Bierens, H. J., and H. Song (2012): "Semi-nonparametric estimation of independently and identically repeated first-price auctions via an integrated simulated moments method," *Journal of Econometrics*, 168(1), 108–119. 4, 17, 29, 39

Bodoh-Creed, A. (2012): "Ambiguous beliefs and mechanism design," *Games and Economic Behavior*. 7, 12

Bose, S., E. Ozdenoren, and A. Pape (2006): "Optimal auctions with ambiguity," *Theoretical Economics*, 1(4), 411–438. 7

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. 15

Esponda, I. (2008): "Information feedback in first price auctions," *The RAND Journal of Economics*, 39(2), 491–508. 2

Gentry, M. L., T. Li, and J. Lu (2015): "Identification and estimation in first-price auctions with risk-averse bidders and selective entry," *Available at SSRN*. 8

Gilboa, I., and D. Schmeidler (1989): "Maxmin expected utility with non-unique prior," *Journal of mathematical economics*, 18(2), 141–153. 4, 7

Grundl, S., and Y. Zhu (2012): "Identification and Estimation of First-Price Auctions Without Assuming Correct Beliefs," Discussion paper, https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=NASM2012&paper_id=581. 1

——— (2013): "Identification and Estimation of First-Price Auctions Under Ambiguity," Discussion paper, http://ssrn.com/abstract=2288105. 1, 6, 7, 11, 18

——— (2015): "Inference of Risk Aversion in First-Price Auctions With Unobserved Auction Heterogeneity," Discussion paper. 8, 11, 33

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574. 2, 7, 9, 18, 23

——— (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica*, 77(4), 1193–1227. 3, 8, 11, 12, 13, 21

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, pp. 1–51. 4, 5, 17

Harsanyi, J. (1968): "Games with incomplete information played by" Bayesian" players, i-iii. Part iii. The basic probability distribution of the game," *Management Science*, pp. 486–502. 3

——— (1995): "Games with incomplete information," *The American Economic Review*, 85(3), 291–303. 3

Hu, Y., D. McAdams, and M. Shum (2013): "Identification of first-price auctions with non-separable unobserved heterogeneity," *Journal of Econometrics*, 174(2), 186–193. 11

Kagel, J., and D. Levin (2008): "Auctions: A Survey of Experimental Research, 1995-2008," *Handbook of Experimental Economics*, 2. 23

Kotlarski, I. (1967): "On characterizing the gamma and normal distribution," *Pacific Journal of Mathematics*, 20(1), 69–76. 29

Krasnokutskaya, E. (2011): "Identification and estimation of auction models with unobserved heterogeneity," *The Review of Economic Studies*, 78(1), 293. 11, 17, 18, 29

Krasnokutskaya, E., and K. Seim (2011): "Bid preference programs and participation in highway procurement auctions," *The American Economic Review*, 101(6), 2653–2686. 8, 29

Lo, K. (1998): "Sealed bid auctions with uncertainty averse bidders," *Economic Theory*, 12(1), 1–20. 7

Matzkin, R. (2003): "Nonparametric estimation of nonadditive random functions," *Econometrica*, 71(5), 1339–1375. 21

McAdams, D. (2007): "Uniqueness in symmetric first-price auctions with affiliation," *Journal of Economic Theory*, 136(1), 144–166. 7

Morris, S. (1995): "The common prior assumption in economic theory," *Economics and philosophy*, 11(02), 227–253. 3

Newey, W. K. (1997): "Convergence rates and asymptotic normality for series estimators," *Journal of Econometrics*, 79(1), 147 – 168. 37

Roberts, J. W., and A. Sweeting (2013): "When Should Sellers Use Auctions?," *American Economic Review*, 103(5), 1830–61. 17

van der Vaart, A., and J. Wellner (2000): *Weak Convergence and Empirical Processes: With Applications to Statistics (Springer Series in Statistics)*. Springer, corrected edn. 34

Zhu, Y., and S. Grundl (2014): "Nonparametric Tests in Moment Equality Models with an Application to Infer Risk Aversion in First-Price Auctions," Discussion paper. 35, 37, 38