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

Identification and Estimation of Risk Aversion in First Price Auctions With Unobserved Auction Heterogeneity

Serafin J. Grundl and Yu Zhu

2015-089

Please cite this paper as:
Grundl, Serafin J., and Yu Zhu (2015). “Identification and Estimation of Risk Aversion in First Price Auctions With Unobserved Auction Heterogeneity,” Finance and Economics Discussion Series 2015-089. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.089.

Identification and Estimation of Risk Aversion in First-Price Auctions with Unobserved Auction Heterogeneity

Serafin Grundl and Yu Zhu *

September 29, 2015

Abstract

We extent the point-identification result in Guerre, Perrigne, and Vuong (2009) to environments with one-dimensional unobserved auction heterogeneity. In addition, we also show a robustness result for the case where the exclusion restriction used for point identification is violated: We provide conditions to ensure that the primitives recovered under the violated exclusion restriction still bound the true primitives in this case. We propose a new Sieve Maximum Likelihood Estimator, show its consistency and illustrate its finite sample performance in a Monte Carlo experiment. We investigate the bias in risk aversion estimates if unobserved auction heterogeneity is ignored and explain why the sign of the bias depends on the correlation between the number of bidders and the unobserved auction heterogeneity. In an application to USFS timber auctions we find that the bidders are risk neutral, but we would reject risk neutrality without accounting for unobserved auction heterogeneity.

*We are very grateful to Amit Gandhi and Jack Porter for their advise and many helpful suggestions. We would also like to thank Andrés Aradillas-López, Bruce Hansen and Xiaoxia Shi for their helpful comments. Serafin Grundl: Federal Reserve Board of Governors, serafin.j.grundl@frb.gov. Yu Zhu: Department of Economics, University of Leicester, yz317@leicester.ac.uk.

1 Introduction

Risk aversion plays an important role in auction theory. Risk aversion leads to more aggressive bidding in first-price auctions with independent private values whereas bidding in English auctions is not affected. Therefore first-price auctions generate higher revenues than English auctions if the bidders are risk averse (Holt (1980)).[1] Matthews (1987) compares auction formats from the perspective of risk averse bidders and Maskin and Riley (1984) study the optimal auction mechanism under risk aversion. In first price auctions risk aversion also tends to reduce the optimal reserve price because aggressive bidding does not have to be induced with the help of a high reserve price (Riley and Samuelson (1981), Hu, Matthews, and Zou (2010)). Knowing bidders' risk aversion is therefore crucial for auction design.

Guerre, Perrigne, and Vuong (2009) showed that risk aversion can be nonparametrically identified from bid data in first-price auctions under an exclusion restriction, by exploiting variation in the number of bidders. This paper tries to bridge the gap between this identification result and applications to field data.

We begin by showing identification in environments with one-dimensional unobserved auction heterogeneity, i.e. one of the auction characteristics is observed by the bidders but not by the econometrician. Identification proceeds in two steps. In the first step multiple bids from the same auction are used to identify the bid distribution conditional on the unobserved characteristic. This step builds on results of Krasnokutskaya (2011), Hu, McAdams, and Shum (2013) and d'Haultfoeuille and Février (2010b), who apply techniques from the measurement error literature. Intuitively, the bid distributions conditional on the unobserved characteristic can be identified using the dependence among bids from the same auction created by the unobserved characteristic. Applying the techniques from the measurement error literature to first-price auctions with risk averse bidders creates a technical challenge, because it requires the highest bid to be strictly increasing in the unobserved characteristic.

[1] This result holds for a given number of risk averse bidders. The revenue ranking is preserved in the entry model of Levin and Smith (1994a) but Smith and Levin (1996) show that it can be reversed with endogenous entry and decreasing absolute risk aversion.

We provide new comparative statics results for auctions with risk averse bidders to establish this monotonicity condition.

In the second step we apply the results of Guerre, Perrigne, and Vuong (2009) to the bid distributions conditional on the unobserved characteristic to identify the primitives. As we condition on the unobserved characteristic we can exploit variation in the number of bidders for identification even if the entry decision depends on the unobserved characteristic. The exclusion restriction required for point identification is that the distribution of valuations conditional on the unobserved characteristic does not depend on the number of bidders. We also discuss the case where the exclusion restriction is violated such that the (conditional) valuation distribution in an auction with more bidders first-order stochastically dominates the valuation distribution with fewer bidders. We provide a condition for the bid distributions that guarantees robustness with respect to this violation in the following sense: The primitives recovered under the violated exclusion restriction still bound the true primitives in this case and risk neutrality remains testable.

Next, we turn to estimation and inference. In light of the typical sample size available in applications we consider a semi-parametric specification with constant relative risk aversion and multiplicative unobserved auction heterogeneity. We propose a new Sieve Maximum Likelihood Estimator and show its consistency under low level conditions.[2] Monte Carlo experiments show that the estimator performs well with sample sizes commonly found in applications.

The Monte Carlo study also presents estimates if the unobserved characteristic is ignored, which has two opposing effects on risk aversion estimates. First, if auctions with a better unobservable characteristic attract more bidders this increases the shift of the (unconditional) bid distributions as the number of bidders increases. Risk aversion tends to decrease the

[2] Deriving the asymptotic distribution of the estimator is difficult due to the non-regular likelihood function and the semi-parametric specification. Ackerberg, Chen, and Hahn (2012) show that for a regular likelihood function, treating the problem as parametric is numerically identical to using the asymptotic formula for semi-parametric estimation. Unfortunately this result does not apply here because the likelihood function is non-regular. Simulation results suggest however that treating the problem as parametric works very well in practice.

shift of the bid distribution as the number of bidders increases, because bids are close to valuations even without fierce competition. Therefore this effects leads us to underestimate risk aversion. Second, the unobserved characteristic increases the dispersion of bids. Risk aversion also increases the dispersion of bids because the bid function at the lower bound of the valuation distribution not affected by risk aversion while bidders with higher values bid more aggressively. Hence, this effect leads us to to overestimate risk aversion. Which of the two effects dominates, and therefore the bias of the risk aversion estimates, depends on how strongly the number of bidders is correlated with the unobserved characteristic.

In an illustrative application we study US Forest Service timber auctions. We find that the bidding firms are close to risk neutral, but we would reject risk neutrality without allowing for unobserved auction heterogeneity.

This paper connects two separate strands of the structural auction literature - unobserved auction heterogeneity and risk averse bidders.

Krasnokutskaya (2011) and Krasnokutskaya (2012) considers identification and estimation with separable unobserved auction heterogeneity in first-price auctions while Hu, McAdams, and Shum (2013) consider identification in the non-separable case. Several papers have documented unobserved auction heterogeneity in US Forest Service timber auctions (e.g. Aradillas-López, Gandhi, and Quint (2013a), Aradillas-López, Gandhi, and Quint (2013b), Roberts and Sweeting (2010), Roberts and Sweeting (2013) and Athey, Levin, and Seira (2011)).

The empirical literature on risk aversion in first price auctions started with laboratory experiments where risk aversion has been proposed as an explanation of the overbidding puzzle.[3] Bajari and Hortacsu (2005) apply structural auction methods to experimental data and conclude that the canonical auction model with risk averse bidders fits experimental data better than some other models which give up the assumption of Bayesian Nash Equilibrium.

[3]The overbidding puzzle refers to the common finding in laboratory experiments that bidders bid more aggressively than predicted by the risk neutral Bayesian Nash Equilibrium. See for example Cox, Smith, and Walker (1988). For more references see the excellent surveys Kagel (1995) and Kagel and Levin (2010)).

Several papers found the bidders in US Forest Service timber auctions to be risk averse. Their results cannot be directly compared to this paper, because they rely on different restrictions for the identification of risk aversion.[4]

Kim (2015b) and Zincenko (2014) propose nonparametric estimators to implement the main identification result in Guerre, Perrigne, and Vuong (2009). As we are interested to applications to field data we also consider identification with unobserved auction heterogeneity and partial identification if the exclusion restriction is relaxed. In an extension Guerre, Perrigne, and Vuong (2009) also consider the case where an unobserved auction characteristic affects the number of bidders and the distribution of valuations. They provide conditions such that the model can still be identified if an instrument is available, which affects the number of bidders but not the distribution of valuations. They consider two alternative conditions to achieve identification: Under the first condition there is a monotone mapping between the number of bidders and the unobserved characteristic. Under the second condition, there is a monotone mapping between the instrument and the unobserved characteristic. These monotonicity assumptions allow the econometrician to back out the unoberved characteristic in a first step and then proceed as if the unobserved characteristic is observed, to identify the distribution of valuations and the utility function.

In a complementary paper to our's Gentry, Li, and Lu (2015) also consider identification and estimation of risk aversion in first-price auctions. In contrast to this paper, they consider a model where the bidders do not know the number of entrants when they submit a bid. Therefore the result of Guerre, Perrigne, and Vuong (2009) does no longer apply in their model and identification is more challenging. They show that a parametric restriction on the copula governing entry usually restores point identification, while a parametric restriction of the utility function leads to partial identification. Allowing for unobserved auction heterogeneity in their framework would be an interesting avenue for future research.

[4]For example Lu and Perrigne (2008) use variation in the auction format, while Campo, Guerre, Perrigne, and Vuong (2011) impose mild parametric restrictions to identify risk aversion. For risk aversion in timber auction see also Baldwin (1995) and Athey and Levin (2001). Campo (2012) finds evidence of risk aversion in construction procurement auctions.

This rest of the paper is organized as follows. Section 2 presents the identification results. In section 3, we propose a semi-parametric Sieve Maximum Likelihood estimator. Section 4 conducts Monte Carlo experiments to evaluate the finite sample performance of the estimator. The Monte Carlo study also illustrates that the bias of risk aversion estimates if the unobserved characteristic is ignored. We explain why the sign of the bias depends the correlation between the unobserved characteristic and the number of bidders. Section 5 is an application to USFS timber auctions.

2 Identification

There are $n \geq 2$ active bidders with independent private values. Their values v are independent draws from the distribution $F(\cdot|u,n)$ with a continuous density $f(\cdot|u,n)$ supported on $[\underline{v}(u), \overline{v}(u)]$ where $0 \leq \underline{v}(u) < \overline{v}(u) \leq \infty$. The econometrician does not observe the one-dimensional auction characteristic u which follows the distribution $F^u(\cdot|n)$. The bidders share a common utility function U with $U'(\cdot) \geq 0$, $U''(\cdot) \leq 0$ and U'' is continuous. The utility function is normalized such that $U(0) = 0$ and $U(1) = 1$.[5] Define $\lambda(\cdot) = U(\cdot)/U'(\cdot)$. The equilibrium bidding strategy $s_n(\cdot, u)$ is characterized by the following first-order condition

$$\frac{\partial s_n'(v,u)}{\partial v} = (n-1)\frac{f(v|u,n)}{F(v|u,n)}\lambda(v - s_n(v,u)),$$

with the boundary condition $s_n(\underline{v}(u), u) = \underline{v}(u)$.

Assumption 1. $F(\cdot|u,n) = F(\cdot|u)$.

Guerre, Perrigne, and Vuong (2009, Proposition 3) showed that if u is observed, F and U are point identified from bid data under this assumption. This assumption holds in many common entry models under fairly general conditions as shown in Appendix B.1.[6] In Theorem

[5] As we normalize the utility function such that $U(1) = 1$ we implicitely assume that $\max_{v \in [\underline{v}(u), \overline{v}(u)]} v - s_n(v,u) \geq 1$. If this condition is violated identification of λ and F is not affected, but we would have to choose a smaller point for the normalization to solve the differential equation $\lambda(\cdot) = U(\cdot)/U'(\cdot)$ for U.

[6] Notice that, like most of the literature, we assume that the bidders know how many of their rivals decided to enter the auction when they bid. Intuitively, Assumption 1 is not very restrictive in this case, because once

3 we provide a robustness result for the case where the condition is violated.

In applications to field data we have to confront the possibility that u is not observed. Previous work studying such environments assumes that bidders are risk neutral and focuses on the identification of $F(\cdot|u)$ (Krasnokutskaya (2011) and Hu, McAdams, and Shum (2013)). The identification arguments exploit the fact that the data contain more than one bid for each auction. The unobserved characteristic creates dependence among bids from the same auction, which allows the researcher to separately identify the distribution of u and the bidders' private information. We combine this strategy with Guerre, Perrigne, and Vuong (2009). The first result is an extension of Krasnokutskaya (2011) which considers cases where valuations consist of two independent and separable components.

Theorem 1. *Suppose that Assumption 1 holds and we observe at least two randomly selected bids from auctions with $n_1, n_2 \geq 2$ bidders. Suppose one of the following conditions holds:*

(1). $F(v|u) = F^*(v-u)$ for all v and u, for some F^* with density f^*. In addition Assumption 7(1) (Appendix A.1) holds.

(2). $F(v|u) = F^*(v/u)$ for all v and u, for some F^* with density f^*. Bidders have constant relative risk aversion (CRRA) with CRRA-coefficient $\sigma \in [0,1)$. In addition Assumption 7(2) (Appendix A.1) holds.

Normalize the lower bound of the support of f^* to 1. Then U, F^*, $F^u(\cdot|n_1)$ and $F^u(\cdot|n_2)$ are identified.

One insight from this result is that there is an important distinction between additive and multiplicative auction heterogeneity if the bidders are risk averse. If the unobserved characteristic enters valuations additively, it also enters the equilibrium bid function additively - regardless of the utility function. If the unobserved characteristic enters valuations

we condition on all the variables observed by the potential bidders - including the unobserved characteristic and if it varies also the number of potential bidders - variation in the number of entrants is “exogenous”. See Gentry, Li, and Lu (2015) for identification of risk aversion in first-price auctions if the bidders do not know how many of their rivals decided to enter the auction when they bid.

multiplicatively and the bidders have constant relative risk aversion, it also enters the bid function multiplicatively. If the utility function is not of the CRRA form however, the bidding strategy is not separable in u and the deconvolution techniques in Kotlarski (1967) can therefore no longer be applied.[7]

The result requires a location normalization. To see why consider the additive case 1(1). If F^* is shifted to the right by 1 while $F^u(\cdot|n_1)$ and $F^u(\cdot|n_2)$ are shifted to the left by 1, the distribution of v and therefore the bid data remains unchanged. Hence this shifted set of primitives is observationally equivalent to the original set of primitives. An analogous argument can be made for the multiplicative case in Theorem 1(1).

Besides allowing for risk aversion, Theorem 1 also generalizes Krasnokutskaya (2011) to accommodate an unbounded unobserved characteristic and unbounded private values. This is achieved by building on an extension of Kotlarski (1967) by Evdokimov and White (2012).

If the unobserved characteristic does not enter in a separable way establishing identification is more involved. Hu, McAdams, and Shum (2013) show how to achieve identification if bidders are risk neutral and u takes on a finite number different values under the following monotonicity restriction on F.

Assumption 2. *$F(v|u_1, n) \leq F(v|u_2, n)$ for all v, $u_1 > u_2$ and n and there exists v such that $F(v|u_1, n) < F(v|u_2, n)$.*

Proposition 1. *Suppose Assumption 2 holds and $\bar{v}(u) < \infty$ for every u, then $s_n(\bar{v}(u_1), u_1) > s_n(\bar{v}(u_2), u_2)$.*

This result says that the highest bid in an auction with unobserved characteristic u_1 is strictly higher than with u_2. Hu, McAdams, and Shum (2013) establish this property by exploiting the closed form of bidding strategy if the bidders are risk neutral. If bidders are risk averse the bidding strategy does typically not have a closed form and establishing strict monotonicity of the highest bid is therefore more involved. [8]

[7] This case will be covered in Theorem 2.

[8] To the best of our knowledge this is a new comparative static result for auctions with risk averse bidders.

Theorem 2. *Suppose that Assumptions 1 and 2 hold and we observe three randomly selected bids from each auction with $n_1, n_2 \geq 3$ bidders. Then U and F are identified if one of the following two conditions is satisfied:*

(1). Discrete u: The support of u is $1, 2, ...K$ with $K < \infty$ for n_1 and n_2.

(2). Continuous u:

(a) $[\underline{u}(n_1), \overline{u}(n_1)] \cap [\underline{u}(n_2), \overline{u}(n_2)] \neq \emptyset$.

(b) $\underline{v}(u)$ is strictly increasing in u.

(c) Assumption 8 (Appendix A.1) holds.

(d) $u = \underline{v}(u)$.

Here $[\underline{u}(n), \overline{u}(n)]$ is the support of the unobserved characteristic in an n bidder auction. Theorem 2(1) extends the result of Hu, McAdams, and Shum (2013) for discrete u. Theorem 2(2) builds on d'Haultfoeuille and Février (2010a) and applies to cases where u is continuous.

The condition for Theorem 2(1) can be broken up into three parts. First, the support of u has finitely many points. Second, the support is the same for n_1 and n_2. Third, the support is normalized to $1, 2, ...K$. Next we turn to the condition for Theorem 2(2). First, we require that for some u we observe n_1 and n_2 bidder auctions - otherwise we could not exploit variation in the number of bidders conditional on u for identification. Second, we assume that $\underline{v}(u)$ is strictly increasing in u. Together with Proposition 1 this implies that the lowest and the highest bid are both strictly increasing in u. The third assumption is a smoothness condition. The fourth assumption is a normalization of u.[9]

It is important that Theorems 1 and 2 allow the distribution of u to depend on the number of bidders. Intuitively, if the bidders observe u before they make their entry decision, then

To show identification we only need to establish strict monotonicity of the highest bid in u, but the proof in Appendix A.3 shows that the whole bid distribution is (weakly) shifted to the right as u increases.

[9] To see why the the normalizations of u are required in Theorem 2 consider $\widetilde{u} = h(u)$ for some increasing function h and $\widetilde{F}(\cdot|u, \cdot) = F\left(\cdot|h^{-1}(u), \cdot\right)$, which leads to the same distribution of valuations and bids as the true primitives.

auctions with a better unobserved characteristic should attract more bidders. In Appendix B.3 we confirm this intution for the case we consider in our application where the unobserved characteristic enters valuations multiplicatively and bidders have constant relative risk aversion. Formally, we show that the distribution of the unobserved characteristic is increasing in n in the sense of first-order stochastic dominance.

Next, we relax Assumption 1 such that valuations are increasing in n in the sense of first-order stochastic dominance.

Assumption 3. *$F(v|u, n_1) \geq F(v|u, n_2)$ for all v, u and $n_1 < n_2$.*

Define

$$R_i(\alpha, u) = \frac{1}{n_i - 1} \frac{\alpha}{g(b_{n_i}(\alpha, u)|u, n_i)},$$

where $i = 1, 2$, $\alpha \in [0, 1]$, $g(\cdot|u, n)$ is the bid density and $b_n(\alpha, u)$ is the α-th quantile of the bid distribution.

Condition 1. Let $n_1 < n_2$. There is u^* such that

(1). $b_{n_1}(0, u^*) = b_{n_2}(0, u^*)$.

(2). $R_1(\alpha, u^*) > R_2(\alpha, u^*)$ for all $\alpha > 0$.

This is not an assumption on primitives but a condition for the bid distribution. Therefore, it can be checked once the bid distribution conditional on u has been recovered. The first part of this condition states that the lowest bid in n_1 and n_2 bidder auctions is the same. To interpret the second part note that the first-order condition for an i bidder auction can be written as $R_i(\alpha, u) = \lambda(v(\alpha, u) - b_{n_i}(\alpha, u))$. Therefore, the condition says that bid shading is larger at the αth quantile in an n_1 bidder auction than in the more competitive n_2 bidder auction.

Let $\widetilde{\lambda}$ with $\widetilde{\lambda}(0) = 0$ be consistent with the bid distributions given u^* if we (incorrectly) impose Assumption 1 for n_1 and n_2. Let $\overline{x} = \widetilde{\lambda}^{-1}\left(\max_{\alpha \in [0,1]} R_1(\alpha, u^*)\right)$. Let $\widetilde{U}(x) = \exp\left(\int_x^1 \log\left(\widetilde{\lambda}(t)\right)\right) dt$ for $x \in [0, 1]$ and $\widetilde{U}(x) = \exp\left(-\int_1^x \log\left(\widetilde{\lambda}(t)\right)\right) dt$ for $x \in [1, \overline{x})$.

Theorem 3. Suppose that that u is observed and that Assumption 3 and Condition 1 hold, then

(1). $\lambda(x) \geq \widetilde{\lambda}(x)$ *for* $x \in [0, \overline{x})$, $U(x) \geq \widetilde{U}(x)$ *for* $x \in [0,1]$, *and* $U(x) \leq \widetilde{U}(x)$ *for* $x \in [1, \overline{x})$.

(2). $b_{n_i}(\alpha, u^*) \leq F^{-1}(\alpha | u^*, n_i) \leq \widetilde{\lambda}^{-1}(R_i(\alpha, u^*)) + b_{n_i}(\alpha, u^*)$ *for* $i = 1, 2$.

To shorten the statement of the result it is assumed that u is observed, but the extension to unobserved u along the lines of Theorems 1 and 2 is straightforward.

The first part of the result shows that $\widetilde{\lambda}$ bounds the true λ from below. By integrating up $\lambda(\cdot) = U(\cdot)/U'(\cdot)$ with $U(1) = 1$, this bound can be translated into a bound on U. The second part shows that the valuations are bounded from below by the bids and from above by the inverse bid function consistent with $\widetilde{\lambda}$.

This is a robustness result. It provides conditions to ensure that the primitives recovered under Assumption 1 remain meaningful as bounds even if the assumption is violated. For example suppose we estimate $\hat{\lambda}$ under Assumption 1 and conclude that the bidders are risk averse because $\hat{\lambda}(x) > x$ for some x. This conclusion remains valid if Assumption 1 is violated, but Assumption 3 and Condition 1 are satisfied. The primitives can be partially identified under Assumption 3, even if Condition 1 does not hold. In this case the bounds do however no longer coincide with the primitives recovered under Assumption 1.

3 Estimation

In light of the typical sample size in applications we consider a semi-parametric specification with constant relative risk aversion and auction characteristics (observed and unobserved), which enter valuations multiplicatively. A bidder's valuation is $v = v^* u \exp[\log(X)\gamma]$. The bidder's private value v^* follows the distribution F^* with density f^*.[10] To simplify the notation, let F_n^u denote the distribution of the unobservable characteristic $F^u(\cdot|u)$ and let

[10]Notice that this notation implicitly imposes Assumption 1.

f_n^u be its density. The private values v^* and the unobserved characteristic u are independent of each other.[11] The p-dimensional vector X contains observable auction characteristics. We assume that X is independent of both $v^*$and u. Bidders share a CRRA utility function with coefficient σ. Following Proposition 1 in Krasnokutskaya (2011) it can be shown that $u \exp\left[\log\left(X\right)\gamma\right]$ enters the bidding strategy multiplicatively.

The data contain L auctions. Let L_n denote the number of auctions with $n \geq 2$ active bidders. Let $\mathbf{N}$ be the set of n such that $L_n > 0$. For the ℓ-th auction, we observe $Z_\ell = (\mathbf{b}_\ell, X_\ell, n_\ell)$. Here $\mathbf{b}_\ell$ is the vector of all bids, X_ℓ is the vector of observed auction characteristics and n_ℓ is the number of active bidders. We also denote the i-th element of $\mathbf{b}_\ell$ as $b_{i,\ell}$. The primitives of the model are $\left(\sigma, \gamma, f^*, \{f_n^u\}_{n\in\mathbf{N}}\right)$. This specification satisfies the assumptions of Theorem 1(2) if $\mathbf{N}$ has at least two elements.

To the best of our knowledge, estimation of bidders' risk aversion in first-price auctions with unobserved auction heterogeneity has not been discussed in the literature. In this paper, we develop a Sieve Maximum Likelihood Estimator (Sieve MLE) based on the joint densities of all the bids from the same auction. We propose a computationally feasible method to compute the joint bid densities. We also show that the estimator is consistent.[12]

It is worth pointing out that there are several other methods in the literature that can be generalized to handle this estimation problem. One possibility is to combine the method suggested by Krasnokutskaya (2011) with Bajari and Hortacsu (2005) to form a two-step estimator.[13] Another candidate is the Simulated Method of Moments (SMM) proposed by

[11] It is worth noting that this assumption is imposed on bidders who decided to enter the auction. In Appendix B.2, we impose the same assumption on potential bidders and ask for which entry models separability and independence of u and v^* carries over to entrants. We show that if the potential bidders observe a signal for v^* *or* u independence carries over to entrants, but not if they observe both.

[12] Formally deriving the asymptotic distribution of the estimator is beyond the scope of this paper. The major difficulty is that our likelihood is non-regular because the support of the bid densities depends on the parameters. Therefore, the results from Ackerberg, Chen, and Hahn (2012), which are valid for regular models, do not directly apply here. The Monte Carlo experiments show that treating the model as parametric and using the asymptotic results from Smith (1985) performs well in practice. It is worth noting that here the bid density does not jump at the boundary of its support, so the results in Donald and Paarsch (1993), Chernozhukov and Hong (2004) and Hirano and Porter (2003) do not apply.

[13] In the first step, use Krasnokutskaya (2011) to separate out the effect of unobserved auction heterogeneity. In the second-step, use Bajari and Hortacsu (2005) to estimate the CRRA coefficient and the distribution of private values.

Bierens and Song (2011) which is extended to the case with unobserved auction heterogeneity by Grundl and Zhu (2015). A third possibility is to extend a Bayesian estimator proposed by Kim (2015a). The Sieve MLE method is our preferred choice for the following reasons. First, compared to the first two alternatives, MLE is more efficient. Second, the Bayesian method considered in Kim (2015a) does not have efficiency advantages in our problem and at the same time involves choosing several tuning parameters for discretization and priors.[14]

3.1 Parameter Space

The support of the densities of unobserved heterogeneity and private values are $[\mu, \overline{u} + \mu]$ and $[1, \overline{v}^* + 1]$ with $\overline{u} > 0$, $\overline{v}^* > 0$ known.[15] Here $\overline{u}$ and $\overline{v}^*$ are the length of the support which may be infinity. μ is the unknown lower bound of the support of u which has to be estimated. It lies in some known closed interval $\mathcal{I} \subset \mathbb{R}$ with lower bound greater than 0. Without loss of generality, the lower bound of v^* is normalized to be 1.

Instead of working directly with primitives, we transform them into the parameter $\theta = \left(\sigma, \gamma, \mu, \psi^*, \{\psi_n^u\}_{n \in \mathbf{N}}\right)$ where μ is the lower bound of unobserved heterogeneity and the ψs are functions supported on $[0,1]$ which take on values no less than -1 and integrate up to 0. $f^*, \{f_n^u\}_{n \in \mathbf{N}}$ can be expressed in terms of ψ functions. To do so, first choose some base density functions h^u and h^* supported on $[0, \overline{u}]$ and$[1, \overline{v}^* + 1]$, respectively. Let H^* and H^u be their corresponding distributions. With some abuse of notation, let the densities given θ be $f^*(x;\theta) = [T\psi^*](H^*(x))\, h^*(x)$ and $f_n^u(x;\theta) = [T\psi_n^u](H^u(x))\, h^u(x)$ where

$$[T\psi](x) = \frac{[1 + \psi(x)]^2}{1 + \int \psi(x)^2\, dx}.$$

It is easy to show that for any primitives f^* and f_n^u, we can find θ such that $f^*(\cdot) =$

[14] Unlike in Kim (2015a), the joint bid densities are continuous due to the unobserved auction heterogeneity. In a parametric model where the densities are continuous at the boundary of the support, the maximum likelihood estimator is efficient even if the support of the densities depends on the parameters, see Smith (1985). Therefore, we choose MLE to avoid picking several additional tuning parameters which is required in Kim (2015a).

[15] Alternatively, we could assume that $\overline{u}$ and $\overline{v}^*$ are unknown but finite and treat them as parameters.

$f^*(\cdot;\theta)$ and $f_n^u(\cdot) = f_n^u(\cdot - \mu;\theta)$. This transformation allows us to work only with functions supported on $[0,1]$.[16]

Let $\theta_0 = \left(\sigma_0, \gamma_0, \mu_0, \psi_0^*, \left\{\psi_{0,n}^u\right\}_{n\in\mathbf{N}}\right)$ be the true parameter under h^* and h^u which lives in a known space $\Theta = \Sigma \times \mathcal{K}^p \times \mathcal{I} \times \mathcal{A}$. $\Sigma = [0, 1-\eta]$, $\mathcal{K} \subset \mathbb{R}$ is a compact set and $\mathcal{I}$ is a closed interval with lower bound greater than 0. $\mathcal{A} = \Psi(B)^{n+1}$ where

$$\Psi(B) = \left\{\psi \in C^q[0,1] : \int \psi(x)\,dx = 0, \int \psi^2(x)\,dx < \infty, \psi + 1 \geq \eta, \sum_{0\leq k\leq q} \int \psi^{(k)}(x)^2\,dx \leq B\right\},$$

where η is some arbitrarily small positive number. B is a known positive constant and q is a positive integer. Notice that $\Psi(B)$ only contains functions that are smooth enough to guarantee that $\Psi(B)$ is compact under the sup-norm. Therefore we avoid inconsistency problem due to a ill-posed inverse problem.[17]

We further define $\alpha = \left(\psi^*, \{\psi_n^u\}_{n\in\mathbf{N}}\right)$, hence $\theta = (\sigma, \gamma, \mu, \alpha)$.With some abuse of notation, let $\|\psi\|_\infty = \sup_{x\in[0,1]} |\psi(x)|$ and

$$\|\alpha\|_\infty = \max\left\{\|\psi^*\|_\infty, \max_{n\in\mathbf{N}}\left\{\|\psi_n^u\|_\infty\right\}\right\},$$

where$\|\cdot\|_E$ is the standard Euclidean norm. One can show that Θ is a compact space under $\|\cdot\|_s$ where

$$\|\theta_1 - \theta_2\|_s = \max\left\{|\sigma_1 - \sigma_2|, |\mu_1 - \mu_2|, \|\gamma_1 - \gamma_2\|_E, \|\alpha_1 - \alpha_2\|_\infty\right\}.$$

3.2 Sieve Maximum Likelihood Estimator

One difficulty in constructing the Sieve MLE is to compute the joint bid densities. These possibly high dimensional objects are complicated functions of θ and have no closed forms. We compute the bid densities numerically by exploiting the separable form of the bidding

[16]This transformation follows Bierens and Song (2012a).
[17]This regularization follows Santos (2012).

function. Let $g_n(\cdot;\theta)$ be the joint density of bids given θ in n-bidder auctions if $\log X = 0$.

$$g_n(\mathbf{b};\theta) = \int \frac{1}{u^n} \prod_{i=1}^{n} g_n^*(b_i/u;\theta) f_n^u(u-\mu;\theta)\, du. \tag{1}$$

Here g_n^* is the marginal bid distribution in an auction with n-bidders whose value density is $f^*(\cdot;\theta)$. $g_n^*(b^*;\theta)$ can be obtained by exploiting the first-order condition of the bidding strategy. Notice that

$$g_n^*(b^*;\theta) = \begin{cases} \frac{1-\sigma}{n-1} \frac{F^*\left(s_n^{*-1}(b^*;\theta);\theta\right)}{s_n^{*-1}(b^*;\theta)-b^*} & \text{if } 1 < b^* \leq s_n^*(\bar{v}^*;\theta) \\ 0 & \text{Otherwise} \end{cases}$$

where $s_n^{*-1}(\cdot;\theta)$ is the inverse of the bidding strategy

$$s_n^*(v;\theta) = v - \int_1^v \left[\frac{F^*(x;\theta)}{F^*(v;\theta)}\right]^{\frac{n-1}{1-\sigma}} dx.$$

The likelihood function can be written as

$$l(Z_\ell;\theta) = l(Z_\ell;(\sigma,\gamma,\mu,\alpha)) = \sum_{n\in\mathbf{N}} \mathbf{1}_{\{n_\ell=n\}} \log g_n\left(\exp\left(\log \mathbf{b}_\ell - \log X_\ell \gamma\right);\theta\right).$$

The Sieve Maximum Likelihood Estimator is defined as

$$\widehat{\theta}_L = \arg\max_{\theta\in\Theta_{k_L}} \frac{1}{L} \sum l(Z_\ell;\theta). \tag{2}$$

$\Theta_{k_L} = \Sigma \times \mathcal{K}^p \times \mathcal{I} \times \mathcal{A}_{k_L}$ is the sieve space where $\mathcal{A}_{k_L}$ is a sequence of finite dimensional spaces that grows with the sample size. The estimator of the CRRA coefficient $\widehat{\sigma}_L$ is the first element of $\widehat{\theta}_L$. Let E_0 be the expectation under the true primitives.

Assumption 4. *(1). h^u and h^* are bounded and strictly bigger than 0 in the interior of their support and they have bounded continuous derivatives.*

(2). $\lim_{v \downarrow 1} h^*(v) / (v-1)^{\epsilon} = C$ *as* $v \downarrow 1$ *for some* $\epsilon \geq 0$ *and* $C > 0$.

(3). $\limsup_{v \to \infty} h^*(v) v^{2+\delta} < C$ *and* $\limsup_{v \to \infty} h^u(v) v^{2+\delta} < C$ *for some* $C, \delta > 0$.

(4). Under h^* *and* h^u, θ_0 *lives in the interior of* Θ.

Assumption 5. *The sieve space satisfies*

(1). $\{\mathcal{A}_{k_L}\}_{L=1}^{\infty}$ *is an increasing sequence of closed subsets of* $\mathcal{A}$.

(2). $\sup_{\alpha \in \mathcal{A}_{k_L}} \|\alpha - \mathcal{A}\|_{\infty} = o(1)$.

Assumption 6. $E_0\left[\log X^T \log X\right]$ *has eigenvalues bounded away from* 0 *and* ∞.

Assumption 4(1)-(3) are requirements for the choice of h^* and h^u. Many commonly used density functions satisfy these requirements. Assumption 4(4) rules out the possibility that θ_0 is a boundary point of Θ. It implies that the densities in the primitives are their corresponding base densities multiplied by functions bounded from above and bounded away from 0.[18] Under this assumption, identification is guaranteed by Theorem 1 if there is no co-variate. Assumption 5(1) requires that the sieve space is closed and increasing so that the maximization problem in (2) is well-defined. Assumption (5)(2) requires that $\mathcal{A}_{k_L}$ approximates $\mathcal{A}$ well enough. In Assumption 6 X^T is the transpose of X. This assumption guarantees γ_0 is identified.

Proposition 2 (Consistency)**.** *If Assumptions 4, 5 and 6 hold,* $\widehat{\theta}_L \xrightarrow{p} \theta_0$ *as* $L \to \infty$ *under* $\|\cdot\|_s$. *In particular,* $\widehat{\sigma}_L \xrightarrow{p} \sigma_0$.

The proof is based on Theorem 5.14 in van der Vaart and Wellner (2000) and generalizes Wald's consistency proof to the Sieve MLE. The complication in this case is that the expected log likelihood function can take on the value $-\infty$ for some θ. Bierens (2014) considers a similar case but he requires the parameters at which the expected log likelihood is greater

[18]Therefore, we rule out densities with unconnected support, unbounded first moment and unbounded values.

than $-\infty$ to be dense in the parameter space. One can show that in the case considered here the set of θ such that $E_0 l(Z_\ell, \theta) = -\infty$ has interior points. It is worth noting that Assumptions 4, 5 and 6 are low level conditions. A key step to prove consistency is to show that under these low level conditions, the likelihood function and the sieve spaces satisfy certain regularity conditions. In particular, we need to show that $l(Z;\theta)$ is upper semi-continuous in θ, Z-a.e. and that there exists $\theta_{0,k_L} \in \Theta_{k_L}$ such that $\|\theta_{0,k_L} - \theta_0\|_s \to 0$ and $E_0 l(Z_\ell, \theta_{0,k_L}) \to E_0 l(Z_\ell, \theta_0)$. Lemmas that establish these regularity conditions are collected in Appendix D.

4 Monte Carlo Experiments

4.1 Setup

Each generated sample has 900 auctions and the number of bidders n ranges from 2 to 5.[19] We consider three different data generating processes (DGPs). In all DGPs $v = v^* u X^{\gamma_0}$ with $\log X \overset{iid}{\sim} N(0,1)$ and $\gamma_0 = 0.9$. The unobserved characteristic is drawn from a χ^2 distribution. In DGP 1, there is no selection on u and the χ^2 parameter is 2 for all n. In DGP 2, there is weak selection on u and the χ^2 parameter increases from 2 for $n = 2$ to 2.6 for $n = 5$. In DGP 3 there is strong selection on u and the χ^2 parameter increases from 2 for $n = 2$ to 6.5 for $n = 5$. In all DGPs, bidders' private values v^* are drawn from a χ^2-distribution with parameter 3. We consider the CRRA coefficients $\sigma_0 = 0,\ 0.1,\ 0.2,\ 0.3$, to assess how well the estimation method can distinguish risk neutrality and moderate levels of risk aversion. We repeat the Monte Carlo experiment 1000 times.

4.2 Estimators

Results for two estimators are reported. First, the Sieve MLE estimator proposed in section 3.2. H^* and H^u are both exponential with parameter 8. ψ^* and ψ_n^* are both 4-th

[19] The shares are 36% for $n = 2$, 27% for $n = 3$, 21% for $n = 4$ and 16% for $n = 5$.

order Legendre polynomials. We compute bidding strategies at 3000 points and interpolate linearly.[20]

Second, as a benchmark, we also estimate the CRRA coefficient without taking unobserved heterogeneity into account, following the method used in Bajari and Hortacsu (2005) (BH). This estimator is computationlly light and therefore a natural choice for a specification without unobserved auction heterogeneity. First, we estimate the following equation by OLS:

$$\log b_{i,\ell} = c + \gamma \log X_\ell + \epsilon_{i,\ell}$$

Let $\hat{\gamma}$ be the OLS estimate. Then we construct the residual bids $\hat{b}^*_{i,\ell} = \exp\left(b_{i,\ell} - \hat{\gamma}\log X_\ell\right)$. Next we estimate the following equation by OLS:

$$\hat{b}^*_{n_1}(q) - \hat{b}^*_{n_2}(q) = (1-\sigma)\left(\frac{q_i}{\hat{g}_{n_2}\left(\hat{b}^*_{n_2}(q)\right)(n_2-1)} - \frac{q_i}{\hat{g}_{n_1}\left(\hat{b}^*_{n_1}(q)\right)(n_1-1)}\right) \tag{3}$$

Here $q \in [0,1]$ and $\hat{b}^*_n(q)$ is the q-th quantile in the empirical distribution of $\hat{b}^*_{i,\ell}$ given $n_\ell = n$ and $\hat{g}_n\left(\hat{b}^*_n(q)\right)$ is the corresponding density. A Gaussian kernel with the rule-of-thumb bandwidth is used to estimate $\hat{g}_n$. Equation 3 is estimated at 100 equally spaced quantiles ranging from 0.25 to 0.75.[21] We restrict the estimates to be between 0 and 1. We report results for $n_1 = 2$ and $n_2 = 4$.[22]

4.3 Results

The discussion focuses on the results for the CRRA coefficient shown in Table 1. [23]

[20] The grid points are chosen such that the grid is finer for low values, because the bidding strategy there can be very nonlinear.

[21] To avoid boundary effects we exclude quantiles close to 0 and 1. We experimented with different quantiles ranges and found similar results (available upon request).

[22] The two-step estimator does not allow us to combine more than two n in an efficient manner. Results for other pairs of n are similar (available upon request).

[23] The Online Appendix shows the results for the value distribution and the distribution of the unobserved characteristic if unobserved heterogeneity is taken into account (Sieve MLE). These findings suggest that the Sieve MLE estimator is might be preferable to two-step estimation procedures (Krasnokutskaya (2011)) or Simulated Method of Moments estimators (Bierens and Song (2012b) and Grundl and Zhu (2015)) even if

First, consider the results if unobserved auction heterogeneity is taken into account using the Sieve MLE estimator shown in the upper half of Table 1. The estimator works well for all three DGPs. The bias is very small (at most 0.014) if $\sigma_0 \neq 0.$, but if the parameter is on the boundary of the parameter space ($\sigma_0 = 0$) it is somewhat larger (up to 0.048). The standard deviation is at most 0.102.[24]

Now consider the result if we ignore unobserved heterogeneity, using the two-step BH estimator shown in the lower half of Table 1. Interestingly, the sign of the bias depends on the DGP.[25] The CRRA coefficient is significantly over-estimated under DGPs 1 (no selection) and 2 (moderate selection), but under-estimated under DGP 3 (strong selection). Section 4.4 provides some intution to understand why the sign of the bias depends on the correlation between the number of bidders and the unobserved characteristic.

We also test risk neutrality using the sieve MLE estimator $H_0 : \sigma_0 = 0$, $H_1 : \sigma_0 > 0$. To construct the test we treat the model as parametric and use the asymptotic distribution of the estimator.[26] Notice that under the null hypothesis, σ_0 is on the boundary of the parameter space. Following the insight from Andrews (1999), $\hat{\sigma}$ is asymptotically truncated normal. Therefore it is still valid for the one sided test to reject the null hypothesis if $\hat{\sigma}$ divided by the standard error exceeds the corresponding quantiles of a standard normal random variable.

Table 2 shows the results for testing risk neutrality. We consider significance levels of 5% and 10%. The test has good size control. For all three DGPs the rejection probability is close to the significance level if $\sigma_0 = 0$. The test also performs well in terms of power. The rejection probability for a 5% significance level increases from about 20% if $\sigma_0 = 0.1$, to about 55% if $\sigma_0 = 0.2$ and about 85% $\sigma_0 = 0.3$. In light of the sample size and the flexibility of the model, it is not surprising that it is difficult to distinguish $\sigma_0 = 0.1$ from risk neutrality.

the researcher assumes that the bidders are risk neutral.

[24]We found that the standard error of the estimator is about 60% smaller than for a Simulated Method of Moments Estimator (Bierens and Song (2012b) and Grundl and Zhu (2015)).

[25]As expected, the bias gets smaller as the variance of the unobserved characteristic is reduced. These Monte Carlo results are available upon request.

[26]To estimate the asymptotic distribution, we need the joint bid densities to vanish smoothly at the boundary of their supports. Hence, we require that at least that the f_n^u vanish smoothly at their boundaries.

		$\sigma_0 = 0$	$\sigma_0 = 0.1$	$\sigma_0 = 0.2$	$\sigma_0 = 0.3$
		Allowing for Unobserved Heterogeneity			
DGP 1: No Selection	Mean	0.046	0.114	0.202	0.292
	Std	0.069	0.086	0.102	0.094
DGP 2: Weak Selection	Mean	0.041	0.108	0.188	0.284
	Std	0.063	0.090	0.096	0.097
DGP 3: Strong Selection	Mean	0.048	0.109	0.198	0.288
	Std	0.074	0.093	0.105	0.102
		Ignoring Unobserved Heterogeneity			
DGP 1: No Selection	Mean	0.698	0.714	0.737	0.754
	Std	0.205	0.160	0.146	0.138
DGP 2: Weak Selection	Mean	0.540	0.554	0.578	0.606
	Std	0.232	0.193	0.174	0.156
DGP 3: Strong Selection	Mean	0.019	0.007	0.001	0.000
	Std	0.134	0.083	0.032	0.000

Table 1: This table shows results of the Monte Carlo study for two estimators of the CRRA coefficient σ. The upper half of the table shows results if unobserved auction heterogeneity is taken into account using the Sieve MLE described in section 3. The lower half of the table shows results if unobserved heterogeneity is ignored using the two-step estimator proposed by Bajari and Hortacsu (2005).

Sig. Level	$\sigma_0 = 0$	$\sigma_0 = 0.1$	$\sigma_0 = 0.2$	$\sigma_0 = 0.3$
DGP 1: No Selection				
10	8.8	34.3	72.5	93.8
5	5.7	24.4	57.6	88.2
DGP 2: Weak Selection				
10	7.6	27.7	69.5	91.9
5	4.9	19.8	53.1	85.9
DGP 3: Strong Selection				
10	10.1	30.1	71.7	92.2
5	7.1	21.2	58.3	84.7

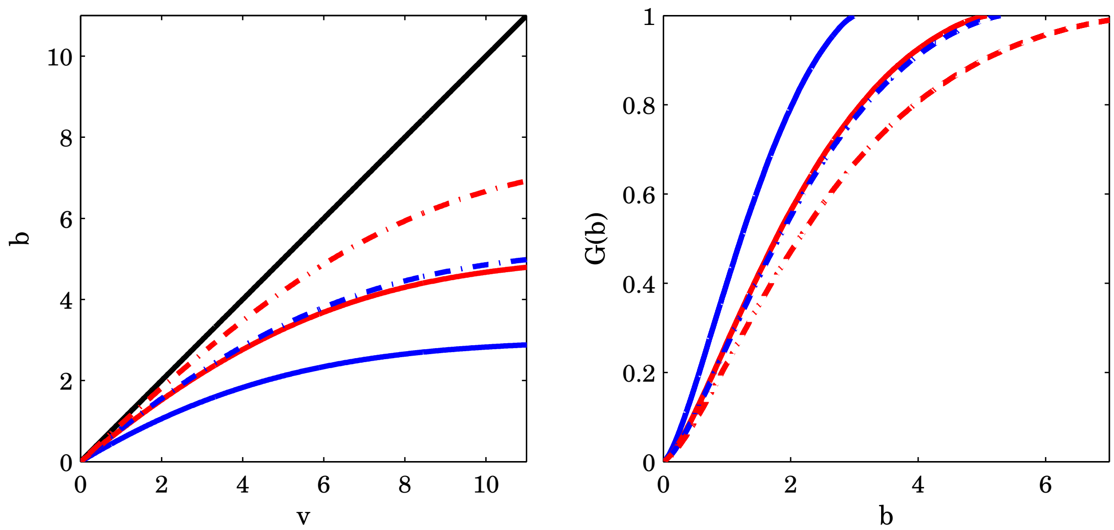
Table 2: This table shows the probability (in %) that risk neutrality ($\sigma_0 = 0$) is rejected if the unobserved characteristic is taken into account (Sieve MLE).

4.4 Understanding the Bias if Unobserved Heterogeneity is Ignored

Here we provide some intution for the bias in risk aversion estimates if unobserve auction heterogeneity is ignored.

Figure I(a) shows bid functions of risk neutral and risk averse bidders in two and four bidder auctions. Private values are on the horizontal axis and the corresponding bids on the vertical axis. The solid blue line and the solid red line depict a risk neutral bidder's strategies in two and four-bidder auctions, respectively. The dashed lines depict a risk averse bidder's strategies. Figure I(b) shows the corresponding bid distributions.

Consider risk neutral bidders first. Their bid shading depends only on the distribution of valuations. Intuitively, the bidders shade their bids more if the values are more dispersed and the bidders have more private information and thereby more market power. If the number of competitors increases, market power declines and the bidders shade their bids less. This shift in the bid function is smaller if the values are are not very dispersed, because then the bids are close to values even for a small number of competitors. Hence, the bid distribution tends



(a) Bid functions (b) Bid distributions

Figure I: This graph illustrates how risk aversion is identified by variation in the number of bidders. The left panel shows bid functions and the right panel the corresponding bid distributions. Solid lines depict risk neutral bidders and dashed lines show risk averse bidders. Blue lines show two bidder auctions and red lines show four bidder auctions.

to respond more to changes in n if the values (and therefore the bids) are more dispersed.

Now consider risk averse bidders who bid more aggressively, which affects how much the bid distribution responds to changes in n and the dispersion of bids. Risk averse bidders respond less to changes in n because the bids are close to values even for a small number of competitors. Second, the dispersion of their bids is larger, because risk aversion has no effect for bidders at the lower bound of the valuation distribution but increases the bids of bidders with higher values. Therefore we conclude that the bidders are risk averse if the bid distribution does not respond much to increases in n relative to the dispersion of the bid distributions.

If unobserved auction heterogeneity is ignored the (unconditional) bid distributions appear very dispersed, as variation in bids due to the unobservable characteristic is attributed to bidders' private information. In addition, if auctions with a higher unobserved characteristic attract more bidders, this increases the shift of the (unconditional) bid distribution as n increases. The first effect increases the dispersion of bids and therefore leads to over-estimation of risk aversion. The second effect increases the shift of the bid distribution as n increases and therefore leads to under-estimation of risk aversion. Which of these two effects dominates, and therefore the sign of the bias, depends on how strongly the number of bidders is correlated with the unobserved characteristic.

5 Empirical Application

5.1 Data Description

We estimate the risk aversion of bidders in USFS timber auctions.[27] The data can be downloaded from Phil Haile's website.[28] Lu and Perrigne (2008) and Campo, Guerre, Perrigne, and Vuong (2011) found the bidding firms to be risk averse.[29] Other work documented unobserved heterogeneity in these auctions (e.g. Aradillas-López, Gandhi, and Quint (2013a), Aradillas-López, Gandhi, and Quint (2013b), Roberts and Sweeting (2010), Roberts and Sweeting (2013) and Athey, Levin, and Seira (2011)).

Following Haile and Tamer (2003), we construct a sub-sample of scaled sales with contract length of less than one year between 1982 and 1990, for which the assumption of private values is plausible.[30] Geographically, we focus on timber tracts from the Southern Region, ranging from Texas and Oklahoma to Florida and Virginia, where most of the first-price auctions take place.

To limit the number of parameters in the distributions of the unobserved characteristic,

[27] Baldwin, Marshall, and Richard (1997, Appendix A) contains a detailed description of the auction procedure and background on the timber industry.

[28] http://www.econ.yale.edu/~pah29/timber/timber.htm

[29] The findings in these papers cannot be directly compared to the findings in this paper, because they do not rely on variation in the number of bidders for the identification of risk aversion. Lu and Perrigne (2008) use variation in the auction format, while Campo, Guerre, Perrigne, and Vuong (2011) impose mild parametric restrictions. For risk aversion in timber auctions see also Baldwin (1995) and Athey and Levin (2001).

[30] In scaled sales, bidders pay only for the timber that is actually harvested; this insures the bidders against the risk of overestimating the volume of timber and reduces the common value component in the valuations. Short term contracts with a contract length of less than one year limit resale opportunities and thereby reduce the common value component generated by the resale market. In 1981 the Forest Service introduced new policies designed to limit subcontracting and speculative bidding (Haile (2001)). Therefore, only auctions after 1981 are included. The data does not contain sales after 1990 for this region.

we further restrict the sample to auctions with 2 to 5 bidders. Intuitively, auctions with few competiors contain most information about risk preferences. As the number of competitors increases the effect of risk aversion on bids becomes small, because competition drives bids close to the values even for risk neutral bidders. To reduce the influence of the extreme bids, we also discard 8 auctions with bids more than 8 times the appraisal value.[31] The final sample includes 370 2-bidder, 263 3-bidder, 172 4-bidder, and 105 5-bidder auctions.

Our estimates condition only on the appraisal value provided by the US Forest Service, but not on other observed timber tract characteristics such as the timber volume, acreage or the species composition. As the appraisal value is meant to capture all the relevant information it can plausibly be treated as a sufficient statistic for the various timber tract characteristics. If some observed characteristics contain important information, which is not captured by the appraisal value, this affects only the precision of the estimator but not its consistency, as we allow for unobserved auction heterogeneity.

5.2 Results and Discussion

The point estimate for the CRRA coefficient is 0.0018. The p-value for testing risk neutrality is 0.4914 and the 95% confidence interval for σ_0 is $[0, 0.163]$. Hence, we reject high levels of risk aversion.

For comparison table 3 shows results if risk aversion is ignored using the estimator in Bajari and Hortacsu (2005) as described in section 4. We report results for different pairs of auction sizes. To assess the robustness of the results, we report estimates based on three choices of quantiles. The bandwidth for the bid density estimators are chosen to be $std(b)L_n^{-1/4}$.[32] The point estimates for the CRRA coefficient range from 0.547 to 0.708. The estimated confidence intervals do not cover any values below 0.324.

Hence we find that the bidders are close to risk neutral if we allow for an unobserved

[31] The remaining bids are all less than 4 times the appraisal value. Therefore we believe that the bids above 8 times the appraisal value can plausibly be considered outliers.

[32] The results are robust to different bandwidth choices we tried.

auction characteristic, but reject risk neutrality in a specification without unobserved auction heterogeneity. This pattern is consistent with a low correlation between the unobserved characteristic and the number of bidders as explained in section 4.4. Indeed we find that the distribution of the unobserved characteristic for different numbers of bidders is fairly similar. A possible explanation is that the unobserved characteristic is observed by the bidder only after they decided to enter the auction. For example, characteristics that are only observable for entrants who typically cruise the auctioned tract, but not for potential bidders.

We follow most of the structural auction literature in assuming that the bidders know the number of their opponents, who also decided to enter, when they submit their bid.[33] Intuitively, a violation of this assumption would bias our risk aversion estimates upwards. To see this consider the case where the number of potential bidders is the same for all auctions.[34] In this case the bid distribution would not vary with the number of entrants. Through the lense of our model this is consistent with extreme levels of risk aversion, such that the bids are very close to valuation regardless of the number of bidders.

	2 and 3 bidders		2 and 4 bidders		2 and 5 bidders	
Quantiles	$\hat{\sigma}$	95% CI	$\hat{\sigma}$	95% CI	$\hat{\sigma}$	95% CI
$[0.20, 0.80]$	0.708	$[0.501, 1.000]$	0.666	$[0.480, 0.898]$	0.694	$[0.552, 0.912]$
$[0.25, 0.75]$	0.652	$[0.406, 1.000]$	0.606	$[0.398, 0.913]$	0.635	$[0.450, 0.913]$
$[0.30, 0.70]$	0.615	$[0.333, 1.000]$	0.547	$[0.324, 0.891]$	0.568	$[0.357, 0.870]$

Table 3: Estimates of the CRRA coefficient σ in a specification without unobserved auction heterogeneity.

[33] Athey, Levin, and Seira (2011) argues that this is a reasonable assumption for timber auctions as the bids are highly correlated with the number of active bidders even after controlling for a variety of variables including the number of potential bidders. See Gentry, Li, and Lu (2015) for identification of risk aversion in first-price auctions if this assumption does not hold.

[34] Alternatively, we could condition number of potential bidders if it is observed.

6 Conclusion

This paper extends the point identification result in Guerre, Perrigne, and Vuong (2009) to environments with unobserved auction heterogeneity and provides conditions to ensure that the primitives recovered under the exclusion restriction for the number of bidders remain meaningful as bounds of the true primitives even if the exclusion restriction is violated. We propose a Sieve Maximum Likelihood Estimator and show its consitency under low level conditions. We explain why the bias in risk aversion estimates if unobserved auction heterogentiy is ignored depends on the correlation between the number of bidders and the unobserved auction heterogeneity. The application underscores the importance of accounting for unobserved heterogeneity as we find the bidders to be risk neutral, but would reject risk neutrality if unobserved heterogeneity is ignored.

We see two avenues for future research. First, relaxing the assumptions of symmetric, independent and private values are important extensions for many applications. Relaxing the assumption of independent values is perhaps most pertinent, because this creates an additional source of correlation among bids from the same auction. The researcher then faces the challenging task disentangle which part of this correlation can be attributed to the unobserved characteristic and which part to the correlation of values conditional on the unobserved characteristic. Second, allowing for unobserved heterogeneity in the framework of Gentry, Li, and Lu (2015) where bidders do not know the number of entrants. For this extension we would have to confirm that the conditions to apply the techniques from the measurment error literature are still satisfied.

A Identification

A.1 Technical Assumptions

Assumption 7. *Technical Assumptions for Theorem 1.*

(1). Additive Case:

(a) The density f^*has non-negative interval support and $f^*(x) < a_1 \exp(-a_2|x|)$ for some constants $a_1, a_2 > 0$. In addition,$\int |u|\, dF^u(u|n) < \infty$ for all n.

(b) $\lambda(x) < \exp(a_3 x)$ for some $a_3 > 0$. In addition, either $\exists a_4 > 0$ such that $\liminf_{x\to\infty} \lambda(x) / \exp(a_4 x) > 0$ or $a_3 < a_2$.

(2). Multiplicative Case: The density f^* has positive interval support and $\int |v|\, dF^*(v) < \infty$. In addition, $\int |\log u|\, dF^u(u|n) < \infty$ for all n.

Assumption 8. *(1). $F^u(\cdot|n)$ has a continuous density $f^u(\cdot|n)$ supported on $[\underline{u}(n), \overline{u}(n)]$.*

(2). $F(\cdot|\cdot, n)$ is continuously differentiable on $\{(v, u) :\ v \in [\underline{v}(u), \overline{v}(u)], u \in [\underline{u}(n), \overline{u}(n)]\}$.

A.2 Proof of Theorem 1

In Theorem 1 (1) bids are additive in u and in Theorem 1 (2) log bids are additive in $\log(u)$. This follows from a slight generalization of Proposition 1 in Krasnokutskaya (2011) presented in section A.2.1. The main identification proof is presented in section A.2.2.

A.2.1 Bidding Strategy

Lemma 1. *Let $s_n(v, u)$ be the bidding strategy for a bidder with value v in an auction with unobserved heterogeneity u and s_n^* be the bidding strategy under F^*.*

(1). If $F(v|u) = F^*(v - u)$, then $s_n(v, u) = s_n^*(v - u) + u$ for all $u \geq 0$ and $v \geq u + 1$.

(2). If $F(v|u) = F^*(v/u)$ and the bidders have constant relative risk aversion, then $s_n(v,u) = s_n^*(v/u)\,u$ for all $u > 0$ and $v \geq u$.

Proof. The bidding strategy under F^* is given by the boundary condition $s_n^*(1) = 1$ and the first-order condtion

$$\frac{ds_n^*(v)}{dv} = (n-1)\frac{f^*(v)}{F^*(v)}\lambda(v - s_n^*(v)).$$

If $F(v|u) = F^*(v-u)$, then $s_n(v,u) = s_n^*(v-u)+u$ satisfies the initial condition $s_n(\underline{v}(u), u) = \underline{v}(u)$ and the first order condition holds:

$$\frac{\partial s_n(v,u)}{\partial v} = \frac{ds_n^*(v-u)}{dv} = (n-1)\frac{f^*(v-u)}{F^*(v-u)}\lambda(v-u-s_n^*(v-u)) = (n-1)\frac{f^*(v|u)}{F^*(v|u)}\lambda(v - s_n(v,u)).$$

If $F(v|u) = F^*(v/u)$, then $s_n(v,u) = s_n^*(v/u)\,u$ satisfies the initial condition $s_n(\underline{v}(u), u) = \underline{v}(u)$ and:

$$\frac{\partial s_n(v,u)}{\partial v} = \frac{ds_n^*\left(\frac{v}{u}\right)}{dv} = (n-1)\frac{f^*\left(\frac{v}{u}\right)}{F^*\left(\frac{v}{u}\right)}\lambda\left(\frac{v}{u} - s_n^*\left(\frac{v}{u}\right)\right) = (n-1)\frac{f^*(v|u)}{F^*(v|u)}\lambda\left(\frac{v - s_n(v,u)}{u}\right)u$$

Notice that $\lambda(\cdot/u)\,u = \lambda(\cdot)$ if and only if the utility function is of the CRRA form. Therefore $s_n(v,u) = s_n^*(v/u)\,u$ satisfies the first-order condidtion with CRRA in this case. □

A.2.2 Proof of Theorem 1

Proof. Let G_n^* and g_n^* be the bid distribution and the corresponding bid density in an n bidder auction if $u = 0$ in Theorem 1 (1) or if $u = 1$ in Theorem 1 (2).

The proof proceeds in two steps. First, we identify $g_{n_1}^*$ and $g_{n_2}^*$ building on Lemma 2 in Evdokimov and White (2012). Second, we identify the model primitives from $g_{n_1}^*$ and $g_{n_2}^*$ building on Proposition 3 in Guerre, Perrigne, and Vuong (2009). Please refer to Evdokimov and White (2012) and Guerre, Perrigne, and Vuong (2009) for these results. Here, we only show that the joint bid distributions satisfy the conditions in Lemma 2 of Evdokimov and White (2012) for both cases in Theorem 1.

For Theorem 1 (1), we can rewrite the model as $v = v^* + u$ with v^* independent of u.

The bidding strategy in an auction with n bidders u is $u + s_n^*(v^*)$ where $s_n^*(1) = 1$ and for $v^* > 1$

$$\frac{ds_n^*(v^*)}{dv^*} = (n-1)\frac{f^*(v^*)}{F^*(v^*)}\lambda(v^* - s_n^*(v^*)).$$

To apply Lemma 2 from Evdokimov and White (2012), we need to show that (a) $E[|u| + |s_n^*(v^*)|] < \infty$ and (b) g_n^* has a tail bounded by an exponential function. Condition (a) is guaranteed by the fact that $E|s_n^*(v^*)| < E|v^*| < \infty$ and the assumption $\int |u|\, dF^u(u|n) < \infty$. For condition (b), notice that by assumption $\exists C > 0$ such that for $v^* > C > 0$ we have

$$\frac{ds_n^*(v^*)}{dv^*} = (n-1)\frac{f^*(v^*)}{F^*(v^*)}\lambda(v^* - s_n^*(v^*)) < (n-1)\, 2a_1 \exp(-a_2 v^*) \exp(a_3(v^* - s_n^*(v^*))).$$

The inequality uses the exponential bound for f^* and λ. Let $s_n^1(v^*)$ be a function that solves

$$\frac{ds_n^1(v^*)}{dv^*} = (n-1)\, 2a_1 \exp(-a_2 v^*) \exp\left(a_3\left(v^* - s_n^1(v^*)\right)\right) \tag{4}$$

with $s_n^1(C) = s_n^*(C)$. Then $s_n^1(v^*) > s_n^*(v^*)$ if $v^* > C$.[35]

If $a_3 < a_2$, it is easy to see that s_n^1 is bounded, so g_n^* has bounded support and is bounded by an exponential tail.

If $a_2 < a_3$, (4) has the solution $\exp(s_n^1(v^*)) = c_1 \exp\left(\frac{a_3 - a_2}{a_3} v^*\right) + c_2$ where $c_1 > 0$ and c_2 are constants. As $\frac{a_3 - a_2}{a_3} < 1$, $s_n^*(v^*) < s_n^1(v^*) < c_2 v^*$ with $0 < c_2 < 1$ for v^* large enough. Then from the first order condition, the density of $s_n^*(v^*)$ satisfies

$$\begin{aligned} g_n^*(s_n^*(v^*)) &= \frac{f^*(v^*)}{\frac{ds_n^*(v^*)}{dv^*}} = \frac{F^*(v^*)}{(n-1)\lambda(v^* - s_n^*(v^*))} \\ &< \frac{F^*(v^*)}{(n-1)\exp(a_4(1-c_2)v^*)} < \frac{1}{(n-1)\exp\left(\frac{a_4(1-c_2)}{c_2} s_n^*(v^*)\right)} \end{aligned}$$

The first inequality follows from the assumption that $\lambda(x) > \exp(a_4 x)$ for x large enough. Hence, g_n^* has an exponential bound.

For Theorem 1 (2), we can rewrite the model as $v = v^* u$ with v^* independent of u . The

[35]This follows from a standard contradiction argument. The proof is available upon request.

bidding strategy is $us_n^*(v^*)$ with

$$\frac{ds_n^*(v^*)}{dv^*} = (n-1)\frac{f^*(v^*)}{F^*(v^*)}(1-\sigma)(v^* - s_n^*(v^*)) < (n-1)\frac{f^*(v^*)}{F^*(v^*)}(1-\sigma)v^*$$

Now we need to show that (a) $E[|\log u| + |\log s_n^*(v^*)|] < \infty$ and that (b) $\log s_n^*(v^*)$ has a density with a tail bounded by an exponential function. First, let $\underline{v}^*$ be the lower bound of v^*. Then $s_n^*(v^*) \leq (n-1)\int_{\underline{v}^*}^{v^*} \frac{f^*(v)}{F^*(v)}(1-\sigma)vdv$ is bounded from above by the assumption that $\int vf^*(v)dv < \infty$. In addition, the bidding function is bounded away from 0. Hence, the density of $\log s_n^*(v^*)$ has a bounded support. Hence, the density satisfies (b) which also suggests $E|\log s_n^*(v^*)| < \infty$. In addition, $E|\log u| < \infty$ by assumption which implies (a) is satisfied.

We normalize the lower bound of the support of $f^*$and thereby the lower bounds of the supports of g_n^* for all n to one. It follows from Lemma 2 in Evdokimov and White (2012) that $g_n^*$and $f^u(\cdot|n)$ are identified for $n = n_1, n_2$.

Next we apply Proposition 3 in Guerre, Perrigne, and Vuong (2009) to $g_{n_1}^*$ and $g_{n_2}^*$. This allows us to identify f^* and U. □

A.3 Proof of Proposition 1

To simplify the notation, let $F_i(\cdot) = F(\cdot|u_i)$, $v_i(\alpha) = F_i^{-1}(\alpha)$ and $s_n^i(\cdot)$ be the bidding strategy under F_i for $i = 1, 2$. In addition, $b_n^i(\alpha) = s_n^i(v_i(\alpha))$ is the αth quantile of the bid distribution.

As $v'(\alpha)f(v(\alpha)) = 1$ we can rewrite the first order condition as follows:

$$\frac{db_n^i(\alpha)}{d\alpha} = \begin{cases} (n-1)\frac{1}{\alpha}\lambda(v_i(\alpha) - b_n^i(\alpha)) & \text{if } \alpha > 0 \\ \\ \frac{(n-1)\lambda'(0)}{(n-1)\lambda'(0)+1}\frac{1}{f_i(v_i(0))} & \text{if } \alpha = 0 \end{cases} \tag{5}$$

Before we prove Proposition 1, we illustrate the main idea of the proof by showing that the stronger assumption $v_1(\alpha) > v_2(\alpha)$ for all α implies that $b_n^1(\alpha) > b_n^2(\alpha)$ for all α. To see this notice that $b_n^1(0) > b_n^2(0)$. Now suppose towards contradiction that for some $\alpha > 0$ we have $b_n^1(\alpha) \leq b_n^2(\alpha)$. By continuity of the bid functions there exists $\alpha_1 = \min\{\alpha : b_n^1(\alpha) = b_n^2(\alpha)\} > 0$. Notice that $b_n^1(\alpha) > b_n^2(\alpha)$ for $\alpha < \alpha_1$ by construction. At the same time we have $\frac{\partial}{\partial\alpha} b_n^1(\alpha_1) > \frac{\partial}{\partial\alpha} b_n^2(\alpha_1)$ because $v_1(\alpha_1) > v_2(\alpha_1)$. Therefore there exists some α slightly smaller than α_1 such that $b_n^1(\alpha) < b_n^2(\alpha)$ which is a contradiction. The proof of Proposition 1 follows a similar idea but is more involved.

Lemma 2. *Under Assumption 2,* $b_n^1(\alpha) \geq b_n^2(\alpha)$ *for all* $\alpha \in [0,1]$.

Proof. First, notice that $b_n^1(0) \geq b_n^2(0)$ as $v_1(0) \geq v_2(0)$. Now suppose towards contradiction that there is $\alpha_2 > 0$ such that $b_n^1(\alpha_2) < b_n^2(\alpha_2)$. Define $\alpha_1 = \max\{\alpha : b_n^1(\alpha) \geq b_n^2(\alpha), \alpha \leq \alpha_2\}$. By construction, $b_n^1(\alpha) < b_n^2(\alpha)$ for $\alpha \in (\alpha_1, \alpha_2)$. As $v_1(\alpha) \geq v_2(\alpha)$ for all α we have $\frac{db_n^1(\alpha)}{d\alpha} > \frac{db_n^2(\alpha)}{d\alpha}$ for all $\alpha \in (\alpha_1, \alpha_2)$. This implies that $b_n^1(\alpha_2) = b_n^1(\alpha_1) + \int_{\alpha_1}^{\alpha_2} \frac{db_n^1(\alpha)}{d\alpha} > b_n^2(\alpha_2) = b_n^2(\alpha_1) + \int_{\alpha_1}^{\alpha_2} \frac{db_n^2(\alpha)}{d\alpha}$ which is a contradiction. $\square$

Lemma 3. *Under Assumption 2 , if* $v_1(\alpha) > v_2(\alpha)$ *then* $b_n^1(\alpha) > b_n^2(\alpha)$*, for* $\alpha \in [0,1]$.

Proof. For $\alpha = 0$ this holds because $b_n^i(0) = v_i(0)$ for $i = 1,2$. Now suppose towards contradiction that $b_n^1(\alpha) \leq b_n^2(\alpha)$ for some $\alpha \in (0,1]$ such that $v_1(\alpha) > v_2(\alpha)$. This implies that $\frac{db_n^1(\alpha)}{d\alpha} > \frac{db_n^2(\alpha)}{d\alpha}$. Therefore we can find α_1 slightly smaller than α such that $b_n^1(\alpha_1) < b_n^2(\alpha_1)$ which contradicts Lemma 2. $\square$

Proof of Proposition 1. Suppose towards contradiction that $b_n^1(1) = b_n^2(1)$ and $v_1(1) = v_2(1)$. This is the only case left to be ruled out, because the remaining cases where $b_n^1(1) \leq b_n^2(1)$ are covered by Lemmas 2 and 3. Define $\Delta b(\alpha) = b_n^1(\alpha) - b_n^2(\alpha)$, $\Delta v(\alpha) = v_1(\alpha) - v_2(\alpha)$ and let $\underline{\alpha} = \inf\{\alpha : v_1(\alpha) = v_2(\alpha) \text{ on } [\alpha, 1]\} > 0$. Notice that $\Delta b(\alpha) = 0$ for all $\alpha \in [\underline{\alpha}, 1]$.[36] Take the difference of the first-order conditions for b_n^1 and b_n^2 and apply the mean value the-

[36] On this region both bid functions can be derived by solving the same differential equation given by equation 5 and the end condition $b_n^1(1) = b_n^2(1)$.

orem twice to obtain

$$\begin{aligned}
\alpha \Delta b'(\alpha) &= (n-1)\,\lambda\left(v_1(\alpha) - b_n^1(\alpha)\right) - (n-1)\,\lambda\left(v_2(\alpha) - b_n^2(\alpha)\right) \\
&= (n-1)\,\lambda'(\bar{r}(\alpha))\,(\Delta v(\alpha) - \Delta b(\alpha)) \\
&= (n-1)\,\lambda'\left(v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})\right)(\Delta v(\alpha) - \Delta b(\alpha)) \\
&\quad + (n-1)\,\lambda''(\tilde{r}(\alpha))\left(\bar{r}(\alpha) - \left(v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})\right)\right)(\Delta v(\alpha) - \Delta b(\alpha)) \qquad (6) \\
&= (n-1)\,(c + \delta(\alpha))\,(\Delta v(\alpha) - \Delta b(\alpha))
\end{aligned}$$

Here $\bar{r}(\alpha)$ is some value between $v_1(\alpha) - b_n^1(\alpha)$ and $v_2(\alpha) - b_n^2(\alpha)$, $\tilde{r}(\alpha)$ is some value between $\bar{r}(\alpha)$ and $v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})$, $c = \lambda'(v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})) \geq 1$ and $\delta(\alpha) = \lambda''(\tilde{r}(\alpha))(\bar{r}(\alpha) - (v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})))$ .If $\alpha \to \underline{\alpha}$ then $v_i(\alpha) - b_n^i(\alpha) \to v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})$ for $i = 1, 2$. Consequently, $\bar{r}(\alpha) \to v_1(\underline{\alpha}) - b_n^1(\underline{\alpha})$ and $\delta(\alpha) \to 0$ as $\alpha \to \underline{\alpha}$. As $c \geq 1$, we can find an $\epsilon > 0$ such that $\underline{\alpha} - \epsilon > 0$ and $c + \delta(\alpha) > 0$ for all $\alpha \in [\underline{\alpha} - \epsilon, \underline{\alpha}]$. Suppose we know δ then we can solve the differential equation (6) for Δb on $[\underline{\alpha} - \epsilon, \underline{\alpha}]$ with the end condition $\Delta b(\underline{\alpha}) = 0$. The closed form solution is

$$\Delta b(\alpha) = -\frac{\int_{\alpha}^{\underline{\alpha}} [c + \delta(w)]\,\Delta v(w) \exp \int_{\underline{\alpha}-\epsilon}^{w} \frac{c+\delta(z)}{z} dz dw}{\exp \int_{\underline{\alpha}-\epsilon}^{\alpha} \frac{c+\delta(z)}{z} dz} < 0.$$

This contradicts Lemma 2. Therefore, $b_n^1(1) > b_n^2(1)$.

$\square$

A.4 Proof of Theorem 2

Proof. The bid distribution given u in an n-bidder auction is denoted by $G_n(\cdot|u)$ and the corresponding density by $g_n(\cdot|u)$. First, consider the case where u is discrete. As the support of u does not depend on the number of bidders, we can normalize u such that it takes values on $1, 2, 3, \cdots, K$ for n_1 and n_2. Hu, McAdams, and Shum (2013) shows that $G_n(\cdot|u)$ is identified if the highest bid is strictly increasing in u. This is satisfied by Proposition 1. We then pair $G_{n_1}(\cdot|u)$ and $G_{n_2}(\cdot|u)$ to identify U and F by applying Proposition 3 in Guerre, Perrigne, and Vuong (2009).

Now consider the case where where u is continuous. We show that the relevant conditions for Steps 1 and 2 in the proof of Theorem 2.1 in d'Haultfoeuille and Février (2010b) are satisfied: First, the highest bid given u is strictly increasing in u by Proposition 1. Second, the lowest bid is assumed to be strictly increasing in u. Third, $G_n(\cdot|\cdot)$, $s_n(\bar{v}(u), u)$ and $s_n(\underline{v}(u), u)$ are continuously differentiable. To see this, notice that $F(\cdot|\cdot)$ and the utility function U are both continuously differentiable by Assumption 8. By Theorem 1 in Campo, Guerre, Perrigne, and Vuong (2011), the bidding strategy $s_n(v, u)$ is continuously differentiable on the support of $F(\cdot|\cdot)$. Hence the highest bid $s_n(\bar{v}(u), u)$ is continuously differentiable with respect to u and $G_n(\cdot|\cdot)$ is continuously differentiable. Therefore, $s_n(\underline{v}(u), u) = \underline{v}(u)$ is continuously differentiable. We normalize $u = \underline{v}(u)$. Now we can apply Theorem 2.1 from d'Haultfoeuille and Février (2010b) to show that $G_n(\cdot|\underline{v}(u))$ is identified for n_1 and n_2. As the supports of $f^u_{n_1}$ and $f^u_{n_2}$ overlap, we can find some $\underline{v}(u)$ such that we observe $G_n(\cdot|\underline{v}(u))$ for $n = n_1, n_2$. We then invoke Proposition 3 in Guerre, Perrigne, and Vuong (2009) to identify U and F. □

A.5 Proof of Theorem 3

Proof. Let G_1 and G_2 be the bid distributions from n_1 and n_2 bidder auctions. We first prove that the bid distribution G_2 first order stochastically dominates G_1. To simplify the notation we suppress u^* from now on. Suppose towards contradiction that G_2 does not first order stochastically dominate G_1. Let $\underline{v}$ denote the common lower bound of the support of both bid distributions (Condition 1(1)). Guerre, Perrigne, and Vuong (2009, Theorem 1) establish that the slope of the bid function at $\underline{v}$ is strictly higher in the n_2 bidder auction. Moreover Assumption 2 implies that the density of valuations in the n_2 bidder auction is weakly lower at $\underline{v}$. Therefore $g_2(\underline{v}) < g_1(\underline{v})$ and at the smallest point $\widetilde{b} > \underline{v}$ such that $G_1\left(\widetilde{b}\right) = G_2\left(\widetilde{b}\right) = \alpha < 1$ we must have $g_2\left(\widetilde{b}\right) \geq g_1\left(\widetilde{b}\right)$. The first order condition of the bidding strategy can be written as

$$g_i\left(\tilde{b}\right) = \frac{1}{n_i - 1}\frac{\alpha}{\lambda\left(v_i\left(\alpha\right) - \tilde{b}\right)} \text{ for } i = 1, 2.$$

, where $v_i\left(\alpha\right)$ is the α-th quantile of the valuation distribution. As $n_2 > n_1$ and $v_2\left(\alpha\right) \geq v_1\left(\alpha\right)$, we must have $g_2\left(\widetilde{b}\right) < g_1\left(\widetilde{b}\right)$ which is a contradiction. Therefore, G_2 must first order stochastically dominate G_1.

As in Guerre, Perrigne, and Vuong (2009), construct a decreasing sequence of αs such that $R_1\left(\alpha_t\right) = R_2\left(\alpha_{t-1}\right)$ with $R_1\left(\alpha_0\right) = x$. As $R_1\left(\alpha_{t-1}\right) > R_2\left(\alpha_{t-1}\right)$, $R_1\left(0\right) = 0 < R_2\left(\alpha_{t-1}\right)$ and R_1 is continuous, there exits an $\alpha_t \in \left(0, \alpha_{t-1}\right)$ such that $R_1\left(\alpha_t\right) = R_2\left(\alpha_{t-1}\right)$ by the intermediate value theorem. Therefore, such a decreasing sequence of αs exists. In addition, this sequence converges to 0. This can be shown by contradiction. First, notice that the sequence is decreasing and bounded from below by 0. Hence, it must converge to some non-negative number c. Suppose towards contradiction that $c > 0$. As R_1 and R_2 are both continuous,

$$R_1\left(c\right) = R_1\left(\lim_{t\to\infty}\alpha_t\right) = \lim_{t\to\infty}R_1\left(\alpha_t\right) = \lim_{t\to\infty}R_2\left(\alpha_{t-1}\right) = R_2\left(\lim_{t\to\infty}\alpha_{t-1}\right) = R_2\left(c\right)$$

This violates the condition that $R_1\left(\alpha\right) > R_2\left(\alpha\right)$ for $\alpha > 0$.

We want to bound $\lambda^{-1}\left(x\right)$. We define $\widetilde{\lambda}$ as the strictly increasing function satisfying $\widetilde{\lambda}^{-1}\left(R_1\left(\alpha\right)\right) - \widetilde{\lambda}^{-1}\left(R_2\left(\alpha\right)\right) = b_2\left(\alpha\right) - b_1\left(\alpha\right)$ for $\alpha \in [0, 1]$ with $\widetilde{\lambda}\left(0\right) = 0$. Notice that if Assumption 1 is violated the existence of this function is no longer guaranteed. We assume that it exists henceforth. Using the α sequence and recursive subsitution $\widetilde{\lambda}$ can be expressed as follows:

$$\widetilde{\lambda}^{-1}\left(x\right) = \sum_{t=0}^{\infty}\left[b_2\left(\alpha_t\right) - b_1\left(\alpha_t\right)\right] = \sum_{t=0}^{\infty}b_2\left(\alpha_t\right) - b_1\left(\alpha_t\right) = \sum_{t=0}^{\infty}\Delta b\left(\alpha_t\right)$$

with $x \in \left[0, \max_{\alpha \in [0,1]} R_1(\alpha)\right)$. This infinite sum exists because for any finite T, $\sum_{t=0}^{T} \Delta b(\alpha_t) \leq \widetilde{\lambda}^{-1}(x)$ and $\Delta b(\alpha_t) \geq 0$ by the first order stochastic dominance of bid distributions shown above. The true λ satisfies the first-order condition $R_i(\alpha) = \lambda(v_i(\alpha) - b_i(\alpha))$ for $i = 1, 2$ so

$$\begin{aligned} \sum_{t=0}^{\infty} b_2(\alpha_t) - b_1(\alpha_t) &\geq \sum_{t=0}^{\infty} [b_2(\alpha_t) - b_1(\alpha_t) + v_1(\alpha_t) - v_2(\alpha_t)] \\ &= \sum_{t=0}^{\infty} \left[\lambda^{-1}(R_1(\alpha_t)) - \lambda^{-1}(R_2(\alpha_t))\right] \\ &= \lambda^{-1}(R_1(\alpha_0)) - \lim_{t \to \infty} \lambda^{-1}(R_2(\alpha_t)) = \lambda^{-1}(x) \end{aligned}$$

The inequality follows from Assumption 2. The last equality uses the fact that $\lim_{t \to \infty} \lambda^{-1}(R_2(\alpha_t)) = 0$ because the α sequence converges to zero. Hence $\widetilde{\lambda}^{-1}(\cdot) \geq \lambda^{-1}(\cdot)$ and therefore $\widetilde{\lambda}(\cdot) \leq \lambda(\cdot)$.

The bounds for the utility function are obtained by solving the differential equation $\lambda(x) = \frac{U(x)}{U'(x)}$ with the boundary condition $U(1) = 1$.

The underlying valuations recovered under Assumption 1 bound the actual valuations from above: $F^{-1}(\alpha | u, n_i) = \lambda^{-1}(R_i(\alpha, u)) + b_i(\alpha, u) \leq \widetilde{\lambda}^{-1}(R_i(\alpha, u)) + b_i(\alpha, u)$ for $i = 1, 2$. Moreover, the valuations are bounded from below by the bids: $F^{-1}(\alpha | u, n_i) \geq b_i(\alpha, u)$ for $i = 1, 2$. □

B Entry

B.1 Entry and Assumption 1

In this section we show that Assumption 1 is satisfied in a fairly general entry framework with (conditionally) independent signals. There are N potential entrants. Prior to bidding, a potential bidder i observes a vector of auction characteristics X including the number of potential entrants and a private signal ξ_i (possibly multi-dimensional). Bidder $i's$ entry strategy is a function $\phi_i : (X, \xi_i) \to \{0, 1\}$. If ϕ_i takes the value 1, the potential bidder enters. Most commonly considered entry models fit into this framework for some ϕ_i. To simplify the notation, we suppress the argument of ϕ_i from now on.

Let $\overrightarrow{v} = (v_1, v_2, \cdots, v_N)$, $\overrightarrow{\xi} = (\xi_1, \xi_2, \cdots, \xi_N)$ and $\overrightarrow{\phi} = (\phi_1, \phi_2, \cdots, \phi_N)$. We use the subscript $-i$ to denote the vector of variables from bidders other than i. Let $F_{A,B|C}(\cdot,\cdot|\cdot)$ denote the joint distribution of A and B conditional on C. For instance, $F_{\overrightarrow{v},\overrightarrow{\xi}|X}(\cdot,\cdot|\cdot)$ is the joint distribution of the value and signal for bidder i given X.[37] The next lemma provides a sufficient condition to ensure that an entrant's value distribution depends only on his identity and X. Therefore, if the bidders are symmetric, an active bidder's value distribution does not vary with n and Assumption 1 holds.

Lemma 4. *If $(v_i, \xi_i) \underset{X}{\perp} \overrightarrow{\xi}_{-i}$ for all i, then for every i, $F_{v_i|\phi_i,\phi_{-i},X}(v|1,\cdot,x) = F_{v_i|\phi_i,X}(v|1,x)$, x-a.e.*

Proof. Notice that $(v_i, \xi_i) \underset{X}{\perp} \overrightarrow{\xi}_{-i}$ implies that $(v_i, \phi_i) \underset{X}{\perp} \overrightarrow{\phi}_{-i}$. By conditional independence

$$F_{v_i|\phi_i,\phi_{-i},X}(v|1,\cdot,x) = \frac{F_{v_i,\phi_i,\phi_{-i}|,X}(v,1,\cdot|x)}{P(\phi_i = 1, \phi_{-i} = \cdot|x)} = \frac{F_{v_i,\phi_i,\phi_{-i}|,X}(v,1|x)\,P(\phi_{-i} = \cdot|x)}{P(\phi_i = 1|x)\,P(\phi_{-i} = \cdot|x)} = F_{v_i|\phi_i,X}(v|1,x)$$

□

B.2 Independence of v^* and u with Entry

Proposition 3. *Suppose v^* and u are independent for potential bidders, then $v^*$and u are independent conditional on n if one of the following conditions holds:*

(1). Potential bidders do not observe any information about v^* or u.

(2). Potential bidders observe only u.

(3). Potential bidders observe a signal s_i of v^*, which is independent across bidders and independent of u.

[37]For example, in the AS model studied in Li, Lu, and Zhao (2015), entry can be described by $\phi_i(X, \xi_i) = 1(\xi_i > c(X))$ where $c(X)$ be a cutoff determined by the model.

Proof. We must allow for the possibility of mixed entry strategies. Bidder i uses a randomization device which generates a random variable ϵ_i independent of (u, v^*) to implement the mixed strategy.

For Proposition 3(1) the entry strategy can be described by some function $g(\epsilon_i)$ which takes the value 0 or 1. Therefore, the joint distribution of (u, v^*) for entrants is

$$\Pr(u \leq x, v^* \leq y | g(\epsilon_i) = 1) = \Pr(u \leq x, v^* \leq y) = \Pr(u \leq x) P(v^* \leq y)$$

.

For Proposition 3(2) the entry strategy can be described by $g(u, \epsilon_i)$, the joint distribution for entrants is

$$\begin{aligned}
\Pr(u \leq x, v^* \leq y | g(u, \epsilon_i) = 1) &= E\{E[1(u \leq x, v^* \leq y)|u, \epsilon_i]|g(u, \epsilon_i) = 1\}. \\
&= E[1(u \leq x) E[1(v^* \leq y)|u, \epsilon_i]|g(u, \epsilon_i) = 1] \\
&= \Pr(v^* \leq y) \Pr(u \leq x | g(u, \epsilon_i) = 1)
\end{aligned}$$

The last step follows because v^* is independent of (u, ϵ_i).

Lastly, following an analogous argument one can show for Proposition 3(3) that

$$P(u \leq x, v^* \leq y | g(s_i, \epsilon_i) = 1) = P(v^* \leq y | g(s_i, \epsilon_i) = 1) P(u \leq x).$$

□

This result does not cover the case where bidders obtain some information about the private information v^* and about the unobserved characteristic u before they enter. In this case $v^*$and u are generally no longer independent for entrants.

B.3 Selection on u

With some abuse of notation, let F^* and F^u denote the distributions of v^* and u for potential bidders and $F^*(\cdot|\, n)$ and $F^u(\cdot|\, n)$ the distributions of v^* and u for a given number of active bidders n. Let N be the number of potential bidders. For simplicity, assume that f^u has a bounded interval support $[\underline{u}, \overline{u}]$. The entry cost is k. Potential bidders share a wealth level $W > k$ and a utility function U. Define $U(x) = \mathrm{U}(x + W - k) - \mathrm{U}(W - k)$ as the re-centered utility function. We assume that at the bidding stage the bidders know their own v^*, u and n.

Proposition 4. *Suppose that $v = uv^*$ where u and v^* are independent, and $U(x) = x^{1-\sigma}$. Potential bidders observe only u, but no signal for v^*. Then $F^u(\cdot|\, n)$ is first-order stochastically increasing in n.*

To prove this we use the following lemma.

Lemma 5. *Let $p(u) \in [0, 1]$ be a weakly increasing function on $[\underline{u}, \overline{u}]$. Let*

$$F^u(x|n) = \int_{\underline{u}}^{x} p(u)^n [1 - p(u)]^{N-n} f^u(u)\, du / C(n)$$

with $C(n) = \int_{\underline{u}}^{\bar{u}} p(u)^n [1 - p(u)]^{N-n} f^u(u)\, du$. Then if $n < n'$, $F^u(x|n') \leq F^u(x|n)$ for $x \in [\underline{u}, \overline{u}]$.

Proof. This is proved by contradiction. As $F^u(\underline{u}|n') = F^u(\underline{u}|n) = 0$ any violation of first order stochastic dominance must be such that $F^u(x|n') \leq F^u(x|n)$ until a point $\overline{x} \in [\underline{u}, \overline{u})$ and then $F^u(x|n') > F^u(x|n)$ for $x \in (\overline{x}, \overline{x} + \epsilon)$ with $\epsilon > 0$. If this is the case, we must have $\frac{\partial}{\partial x} F^u(\overline{x}|n') = f^u(\overline{x}|n') \geq f^u(\overline{x}|n) = \frac{\partial}{\partial x} F^u(\overline{x}|n)$. Now consider

$$\frac{f^u(x|n)}{f^u(x|n')} = \frac{C(n)\, p(x)^n [1 - p(x)]^{N-n}}{C(n')\, p(x)^{n'} [1 - p(x)]^{N-n'}} = \frac{C(n)}{C(n')} \left[\frac{1 - p(x)}{p(x)}\right]^{n'-n}.$$

As $n' > n$ and p is weakly increasing, the ratio above is weakly decreasing in x. Consequently, $f^u(x|n') \geq f^u(x|n)$ for all $x > \overline{x}$. As $F^u(x|n') > F^u(x|n)$ for $x \in (\overline{x}, \overline{x} + \epsilon)$ this implies that

$F^u\left(\overline{u}|n'\right) > F^u\left(\overline{u}|n\right)$. This is a contradiction because by definition $F^u\left(\overline{u}|n'\right) = F^u\left(\overline{u}|n\right) = 1$. $\square$

Proof of Proposition 4. At the bidding stage bidder i solves the following problem:

$$
\begin{aligned}
& \max_{b_i} \left[\mathrm{U}\left(v_i^* u - b_i + W - k\right) - \mathrm{U}\left(W-k\right)\right] P\left(S_n\left(\max_{j\neq i} v_j^*, u\right) \leq b_i \middle| u\right)^{n-1} + \mathrm{U}\left(W-k\right) \\
= \ & \max_{b_i} \left(v_i^* u - b_i\right)^{1-\sigma} P\left(S_n\left(v_j^*, u\right) \leq b_i \middle| u\right)^{n-1} + \mathrm{U}\left(W-k\right) \\
= \ & u^{1-\sigma}\left(v_i^* - s_n^*\left(v_i^*\right)\right)^{1-\sigma} F^*\left(v_i^*\right)^{n-1} + \mathrm{U}\left(W-k\right). \qquad (7)
\end{aligned}
$$

Here S_n is the bidding strategy and s_n^* is the bidding strategy if $u=1$. Let

$$
\pi_n\left(u, v_i^*\right) = u^{1-\sigma}\left(v_i^* - s_n^*\left(v_i^*\right)\right)^{1-\sigma} F^*\left(v_i^*\right)^{n-1}
$$

Let $\Pi_n\left(u\right) = \int \pi_n\left(u, v_i^*\right) dF^*\left(v_i^*\right)$. Notice that $\Pi_n\left(u\right) = u^{1-\sigma}\Pi_n\left(1\right)$.

In equilibrium, potential bidders enter randomly with some probability which is a function of u. The symmetric entry probability $p\left(u\right)$ satisfies

1. $p\left(u\right) = 1$ if $u^{1-\sigma_0}\Pi_N\left(1\right) > \mathrm{U}\left(W\right) - \mathrm{U}\left(W-k\right)$
2. $p\left(u\right) = 0$ if $u^{1-\sigma_0}\Pi_1\left(1\right) < \mathrm{U}\left(W\right) - \mathrm{U}\left(W-k\right)$.
3. Otherwise,

$$
\sum_{n=1}^{N}\left[1-p\left(u\right)\right]^{N-n} p\left(u\right)^{n-1}\Pi_n\left(1\right) = \frac{\mathrm{U}\left(W\right) - \mathrm{U}\left(W-k\right)}{u^{1-\sigma}}. \qquad (8)
$$

We only need to show that $p\left(u\right)$ is weakly increasing in u then Lemma 5 implies that $F^u\left(\cdot|\, n\right)$ is increasing in n in the sense of first-order stochastic dominance . This takes two steps. First, notice that $\Pi_n\left(1\right)$ is decreasing in n. To see this, recall that

$$
\Pi_n\left(1\right) = \int \left(v_i^* - s_n^*\left(v_i^*\right)\right)^{1-\sigma} F^*\left(v_i^*\right)^{n-1} dF^*\left(v_i^*\right).
$$

Here $v^* - s_n^*\left(v^*\right)$ and $F^*\left(v^*\right)^{n-1}$ are both decreasing in n. Hence $\Pi_n\left(1\right)$ is decreasing in n as the integrand is always decreasing in n. Second, notice that a higher u reduces $\frac{\mathrm{U}(W)-\mathrm{U}(W-k)}{u^{1-\sigma}}$ which corresponds to the entry cost in Lemma 1 from Levin and Smith (1994b). Apply this lemma to conclude that $p(u)$ is increasing in u if $p(u)$ solves (8). If $u^{1-\sigma}\Pi_N(1) > \mathrm{U}(W) - \mathrm{U}(W-k)$ or $u^{1-\sigma}\Pi_1(1) < \mathrm{U}(W) - \mathrm{U}(W-k)$, $\frac{d}{du}p(u) = 0$. Then $p(u)$ is weakly increasing in u. □

Proposition 4 considers the specification used in the application with constant relative risk aversion and multiplicative unobserved heterogeneity. The result can be extended to additive unobserved heterogeneity and in this case the restriction to constant relative risk aversion is no longer needed.

C Proof of Proposition 2

Proof. Assumption 6 and the identification result imply that $E_0 l(Z_\ell; \theta)$ is uniquely maximized at θ_0. $l(Z_\ell; \theta)$ is bounded from above by a constant because by Lemma 9, g_n^* is bounded. Then $E_0 l(Z_\ell; \theta)$ is upper semi-continuous by Lemma 11 and the Reverse Fatou Lemma,

$$\limsup_{k\to\infty} E_0 l(Z_\ell; \theta_k) \le E_0 \limsup_{k\to\infty} l(Z_\ell; \theta_k) \le E_0 l(Z_\ell; \theta).$$

As Θ is compact, for any $\epsilon > 0$, there exists a $\delta > 0$ such that

$$E_0 l(Z_\ell, \theta_0) - \sup_{\|\theta-\theta_0\|_s \ge \epsilon} E_0 l(Z_\ell, \theta) > \delta.$$

Define $\Theta(\epsilon) = \{\theta \in \Theta : \|\theta - \theta_0\|_s \ge \epsilon\}$. For any $\theta \in \Theta(\epsilon)$, let $\mathcal{N}(\theta)$ be a closed ball around it. Let $l_{\mathcal{N}(\theta)}(Z_\ell) = \sup_{\theta' \in \mathcal{N}(\theta)} l\left(Z_\ell, \theta'\right)$. By Lemma 11, $l(Z_\ell, \theta)$ is continuous in θ Z_ℓ-a.s. Hence if $\mathcal{N}(\theta) \downarrow \theta$, $l_{\mathcal{N}(\theta)}(Z_\ell) \downarrow l(Z_\ell, \theta)$ a.s. By the monotone convergence theorem, $E_0 l_{\mathcal{N}(\theta)}(Z_\ell) \downarrow E_0 l(Z_\ell, \theta)$. Therefore, for any $\theta \in \Theta(\epsilon)$, there exists $\mathcal{N}(\theta)$ such that $E_0 l(Z_\ell, \theta_0) - E_0 l_{\mathcal{N}(\theta)}(Z_\ell) > \delta/2$. Then $\Theta(\epsilon) \subseteq \cup_{\theta\in\Theta(\epsilon)} \mathcal{N}(\theta)$. Because $\Theta(\epsilon)$ is a closed subset of a compact space Θ, it is also compact. Hence, there exist $\mathcal{N}_j = \mathcal{N}(\theta_j)$, $j = 1, 2, \cdots, J$ that cover $\Theta(\epsilon)$. By Lemma 6 and Lemma 9, $g_n(\mathbf{b}, \theta)$ is bounded from above, so is $l(Z_\ell, \theta)$ and $l_{\mathcal{N}_j}(Z_\ell)$. Therefore, we can still apply law of large numbers even if the expectation may be $-\infty$.

$$\sup_{\theta\in\Theta(\epsilon)}\frac{1}{L}\sum l\left(Z_\ell,\theta\right)\le\sup_j\frac{1}{L}\sum l_{\mathcal{N}_j}\left(Z_\ell\right)\to^{a.s.}\sup_j E_0 l_{\mathcal{N}_j}\left(Z_\ell\right)<E_0 l\left(Z_\ell,\theta_0\right)-\delta/2. \tag{9}$$

There exists sequence $\{\theta_{0,k_L}\}_{k_L=1}^{\infty}$, $\theta_{0,k_L}\in\Theta_{k_L}$ such that $\lim_{k_L\to\infty}\left\|\theta_{0,k_L}-\theta_0\right\|_s=0$, and $E_0 l\left(Z_\ell,\theta_{0,k_L}\right)-E_0 l\left(Z_\ell,\theta_0\right)\to 0$ by Lemma 12. Therefore, we can find K large enough such that $\left|E_0 l\left(Z_\ell,\theta_{0,k_L}\right)-E_0 l\left(Z_\ell,\theta_0\right)\right|<\delta/4$ and $\theta_K\in\Theta\left(\epsilon\right)^c\cap\Theta_{k_L}$ for all $k_L\ge K$. By this definition and (9),

$$E_0 l\left(Z_\ell,\theta_K\right)-\sup_j E_0 l_{\mathcal{N}_j}\left(Z_\ell\right)>\delta/4 \tag{10}$$

For L large enough

$$\begin{aligned}\left\{\widehat{\theta}_L\in\Theta\left(\epsilon\right)\right\} &\subseteq \left\{\sup_{\theta\in\Theta(\epsilon)\cap\Theta_{k_L}}\frac{1}{L}\sum l\left(Z_\ell,\theta\right)\ge\sup_{\theta\in\Theta(\epsilon)^c\cap\Theta_{k_L}}\frac{1}{L}\sum l\left(Z_\ell,\theta_L\right)\right\}\\ &\subseteq \left\{\sup_{\theta\in\Theta(\epsilon)}\frac{1}{L}\sum l\left(Z_\ell,\theta\right)\ge\sup_{\theta\in\Theta(\epsilon)^c\cap\Theta_{k_L}}\frac{1}{L}\sum l\left(Z_\ell,\theta_L\right)\right\}\\ &\subseteq \left\{\sup_{\theta\in\Theta(\epsilon)}\frac{1}{L}\sum l\left(Z_\ell,\theta\right)\ge\frac{1}{L}\sum l\left(Z_\ell,\theta_K\right)\right\}\end{aligned}$$

The probability of $\left\{\sup_{\theta\in\Theta(\epsilon)}\frac{1}{L}\sum l\left(Z_\ell,\theta\right)\ge\frac{1}{L}\sum l\left(Z_\ell,\theta_K\right)\right\}$ converges to 0 because

$$\begin{aligned}&\limsup_{L\to\infty}P\left(\sup_{\theta\in\Theta(\epsilon)}\frac{1}{L}\sum l\left(Z_\ell,\theta\right)\ge\frac{1}{L}\sum l\left(Z_\ell,\theta_K\right)\right)\le\limsup_{L\to\infty}P\left(\sup_j\frac{1}{L}\sum l_{\mathcal{N}_j}\left(Z_\ell\right)\ge\frac{1}{L}\sum l\left(Z_\ell,\theta_K\right)\right)\\ \le\ & P\left(\sup_j E_0 l_{\mathcal{N}_j}\left(Z_\ell\right)+\delta/8\ge E_0 l\left(Z_\ell,\theta_K\right)-\delta/8\right)+\limsup_{L\to\infty}P\left(\frac{1}{L}\sum l\left(Z_\ell,\theta_K\right)-E_0 l\left(Z_\ell,\theta_K\right)<-\delta/8\right)\\ &+\limsup_{L\to\infty}P\left(\sup_j\frac{1}{L}\sum l_{\mathcal{N}_j}\left(Z_\ell\right)-\sup_j E_0 l_{\mathcal{N}_j}\left(Z_\ell\right)>\delta/8\right)\\ \le\ & P\left(E_0 l\left(Z_\ell,\theta_K\right)-\sup_j E_0 l_{\mathcal{N}_j}\left(Z_\ell\right)\le\delta/4\right)+\limsup_{L\to\infty}P\left(\sup_j\frac{1}{L}\sum l_{\mathcal{N}_j}\left(Z_\ell\right)-\sup_j E_0 l_{\mathcal{N}_j}\left(Z_\ell\right)>\delta/8\right)\\ &+\limsup_{L\to\infty}P\left(\frac{1}{L}\sum l\left(Z_\ell,\theta_K\right)-E_0 l\left(Z_\ell,\theta_K\right)<-\delta/8\right)=0\end{aligned}$$

The last step follows from (10) and the Law of Large Numbers. Hence,

$$P\left(\widehat{\theta}_L \in \Theta\left(\epsilon\right)\right) \leq P\left(\sup_{\theta\in\Theta(\epsilon)} \frac{1}{L}\sum l\left(Z_\ell, \theta\right) \geq \frac{1}{L}\sum l\left(Z_\ell, \theta_K\right)\right) \to 0.$$

$\square$

D Lemmas for the Proof of Proposition 2 (not for publication)

Throughout this section, C and any C with a subscript are generic finite positive constants which may take different values at different places.

Lemma 6. *$f^*(\cdot;\theta)$ and $f_n^u(\cdot;\theta)$ satisfies the following*

(1). $\frac{\eta^2}{1+B}h^*(\cdot) \le f^*(\cdot;\theta) \le \left(1+3\sqrt{B}\right)^2 h^*(\cdot)$, $\frac{\eta^2}{1+B}h^u(\cdot) \le f_n^u(\cdot;\theta) \le \left(1+3\sqrt{B}\right)^2 h^u(\cdot)$. *Hence, $f^*(\cdot,\theta)$ and $f_n^u(\cdot,\theta)$ are uniformly bounded from above.*

(2). $\frac{\eta^2}{1+B}H^*(\cdot) \le F^*(\cdot;\theta) \le \left(1+3\sqrt{B}\right)^2 H^*(\cdot)$, $\frac{\eta^2}{1+B}H^u(\cdot) \le F_n^u(\cdot;\theta) \le \left(1+3\sqrt{B}\right)^2 H^u(\cdot)$.

(3). $\sup_x \|f^*(x;\theta_1) - f^*(x;\theta_2)\|_\infty \le C\|\theta_1-\theta_2\|_s$ $\forall\theta_1,\theta_2 \in \Theta$ *for some $C<\infty$. The same holds for $f_n^u(\cdot;\theta)$ $\forall n \in \mathbf{N}$.*

(4). $\sup_x \|F^*(x;\theta_k) - F^*(x;\theta)\|_\infty \le C\|\theta_1-\theta_2\|_s$ $\forall\theta_1,\theta_2 \in \Theta$ *for some $C<\infty$.The same holds for $F_n^u(\cdot;\theta)$ $\forall n \in \mathbf{N}$.*

Proof. We start with the first claim. $\psi \in \Psi(B)$ implies $|\psi(x)| \le 3\sqrt{B}$. To see this, notice $|\psi(0)| \le 2\sqrt{B}$ for any $\psi \in \Psi(B)$. If not, we can find a $\psi \in \Psi(B)$ such that $|\psi(0)| > 2\sqrt{B}$. Without loss of generality, we can assume that $\psi(0) > 2\sqrt{B}$. Then

$$\psi(x) = \psi(0) + \int_0^x \psi'(y)\,dy \ge \psi(0) - \int_0^x |\psi'(y)|\,dy \ge \psi(0) - \sqrt{\int_0^1 |\psi'(y)|^2\,dy} > \sqrt{B}$$

which suggests$\psi \notin \Psi(B)$. This is a contradiction. Therefore, $\psi(0) \le 2\sqrt{B}$ and

$$|\psi(x)| = \left|\psi(0) + \int_0^x \psi'(y)\,dy\right| \le |\psi(0)| + \int_0^x |\psi'(y)|\,dy \le |\psi(0)| + \sqrt{\int_0^1 |\psi'(y)|^2\,dy} \le 3\sqrt{B} \quad (11)$$

Because $|\psi(x)| \le 3\sqrt{B}$,

$$\frac{\eta^2}{1+B} \le \frac{(1+\psi(x))^2}{1+\int\psi(x)^2\,dx} \le \left(1+3\sqrt{B}\right)^2 \quad (12)$$

Then we have

$$\left(1+3\sqrt{B}\right)^2 h^u(x) \geq f_n^u(x;\theta) = \left(T\psi_n^u\right)\left[H^u(x)\right] h^u(x) \geq \frac{\eta^2}{1+B} h^u(x),$$

The same inequalities hold for $f^*(x;\theta)$. The second claim holds by integrating the above inequalities.

Next, we prove the third claim. We only need to show that the mapping $T\psi = \frac{[1+\psi(\cdot)]^2}{1+\int \psi^2(x)dx}$ is continuous in ψ under $\|\cdot\|_\infty$ on $\Psi(B)$. If $\|\psi_k - \psi\|_\infty \to 0$ as $k \to \infty$,

$$\begin{aligned}
\|T\psi_1 - T\psi_2\|_\infty &= \sup_{x\in[0,1]} \left| \frac{[1+\psi_1(x)]^2}{1+\int \psi_1^2(x)\,dx} - \frac{[1+\psi_2(x)]^2}{1+\int \psi_2^2(x)\,dx} \right| \\
&\leq \sup_{x\in[0,1]} \left| \frac{[1+\psi_1(x)]^2}{1+\int \psi_1^2(x)\,dx} - \frac{[1+\psi_2(x)]^2}{1+\int \psi_1^2(x)\,dx} \right| \\
&\quad + \sup_{x\in[0,1]} \left| \frac{[1+\psi_2(x)]^2}{1+\int \psi_1^2(x)\,dx} - \frac{[1+\psi_2(x)]^2}{1+\int \psi_2^2(x)\,dx} \right| \\
&\leq \|\psi_1 - \psi_2\|_\infty \left(2\|\psi_1\|_\infty + 2\|\psi_2\|_\infty + 4\left(1+\|\psi_2\|_\infty\right)^2\right).
\end{aligned}$$

By (11), $\|T\psi_1 - T\psi_2\|_\infty \leq \left(6\sqrt{B} + 4\left(1+3\sqrt{B}\right)^2\right)\|\psi_1 - \psi_2\|_\infty \equiv C_2\|\psi_1-\psi_2\|_\infty$ which is Hölder continuous in ψ. Trivially

$$\begin{aligned}
\left|f_n^u(x;\theta_1) - f_n^u(x;\theta_2)\right| &= \left|\left(T\psi_{n,1}^u\right)\left[H^u(x)\right]h_u(x) - \left(T\psi_{n,2}^u\right)\left[H^u(x)\right]h_u(x)\right| \\
&\leq C_2\|h^u\|_\infty \|\psi_1-\psi_2\|_\infty \leq C_2\|h^u\|_\infty \|\theta_1-\theta_2\|_s
\end{aligned}$$

which is Hölder continuous in θ. Notice that the bound of the above expression does not depend on x. The same holds for $f^*(x,\theta)$. For the last claim, just notice that $\left|F_n^u(x;\theta_1) - F_n^u(x;\theta_2)\right| \leq \int \left|f_n^u(x;\theta_1) - f_n^u(x;\theta_2)\right| dx \leq C\|\theta_1-\theta_2\|_\infty$. The same inequality holds for $F^*(\cdot;\theta)$ □

Lemma 7. $\sup_{v\in[1,\bar{v}^*+1]} \left|s_n^*(v;\theta_k) - s_n^*(v;\theta)\right| \to 0$ for *if* $\|\theta_k - \theta\|_s \to 0$.

Proof. First, suppose $\overline{v}^* < \infty$. We first show that that $s_n^*(\cdot;\theta_k)$ converges to $s_n^*(\cdot;\theta)$ uniformly on $[1+\epsilon, \overline{v}^*+1]$ for any $\epsilon > 0$.

$$\begin{aligned}
|s_n^*(x;\theta_k) - s_n^*(x;\theta)| &= \left| \int_1^x \left[\frac{F^*(v;\theta_k)}{F^*(x;\theta_k)}\right]^{\frac{n-1}{1-\sigma_k}} dv - \int_1^x \left[\frac{F^*(v;\theta)}{F^*(x;\theta)}\right]^{\frac{n-1}{1-\sigma}} dv \right| \\
&\leq \left| \int_1^x \left[\frac{F^*(v;\theta_k)}{F^*(x;\theta_k)}\right]^{\frac{n-1}{1-\sigma_k}} dv - \int_1^x \left[\frac{F^*(v;\theta)}{F^*(x;\theta_k)}\right]^{\frac{n-1}{1-\sigma_k}} dv \right| \\
&\quad + \left| \int_1^x \left[\frac{F^*(v;\theta)}{F^*(x;\theta_k)}\right]^{\frac{n-1}{1-\sigma_k}} dv - \int_1^x \left[\frac{F^*(v;\theta)}{F^*(x;\theta_k)}\right]^{\frac{n-1}{1-\sigma}} dv \right| \\
&\quad + \left| \int_1^x \left[\frac{F^*(v;\theta)}{F^*(x;\theta_k)}\right]^{\frac{n-1}{1-\sigma}} dv - \int_1^x \left[\frac{F^*(v;\theta)}{F^*(x;\theta)}\right]^{\frac{n-1}{1-\sigma}} dv \right| \\
&= A_1(x) + A_2(x) + A_3(x).
\end{aligned}$$

$A_1(x)$,$A_2(x)$ and $A_3(x)$ have bounds independent of x.

$$\begin{aligned}
A_1(x) &\leq \frac{1}{F^*(x;\theta_k)^{\frac{n-1}{1-\sigma_k}}} \int_1^x \left| F^*(v;\theta_k)^{\frac{n-1}{1-\sigma_k}} - F^*(v;\theta)^{\frac{n-1}{1-\sigma_k}} \right| dv \\
&\leq \frac{1}{F^*(x;\theta_k)^{\frac{n-1}{1-\sigma_k}}} \int_1^{\bar{v}^*+1} \left| F^*(v;\theta_k)^{\frac{n-1}{1-\sigma_k}} - F^*(v;\theta)^{\frac{n-1}{1-\sigma_k}} \right| dv
\end{aligned}$$

Because $x > \epsilon$, $F^*(x) > 2\delta$ for some $\delta > 0$. $F^*(v;\theta_k) \to F^*(v;\theta)$ uniformly by Lemma 6. Hence, there exists an K such that for all $k > K$, $F_k^*(x) > \delta$ for all $x > 1+\epsilon$. Therefore, for k large enough,

$$A_1(x) \leq \frac{1}{\delta^{\frac{n-1}{1-\sigma_k}}} \int_1^{\bar{v}^*+1} \left| F^*(v;\theta_k)^{\frac{n-1}{1-\sigma_k}} - F^*(v;\theta)^{\frac{n-1}{1-\sigma_k}} \right| dv.$$

In addition, notice $\sigma_k \to \sigma$, $\delta^{\frac{n-1}{1-\sigma_k}} \to \delta^{\frac{n-1}{1-\sigma}}$. Therefore, for k large enough

$$A_1(x) \leq \frac{2}{\delta^{\frac{n-1}{1-\sigma}}} \int_0^{\overline{v}^*+1} \left| F^*(v;\theta_k)^{\frac{n-1}{1-\sigma_k}} - F^*(v;\theta)^{\frac{n-1}{1-\sigma_k}} \right| dv.$$

Because $\left| F^*(v;\theta_k)^{\frac{n-1}{1-\sigma_k}} - F^*(v;\theta)^{\frac{n-1}{1-\sigma_k}} \right| \to 0$ uniformly, $A_1(x) \to 0$ uniformly in x. We can apply a similar argument to $A_2(x)$ and $A_3(x)$ to conclude they converges to 0 uniformly in x. Therefore, $|s_n^*(x;\theta_k) - s_n^*(x;\theta)|$ converges to 0 uniformly on $[1+\epsilon, 1+\overline{v}^*]$. Because

$s_n^*(1;\theta_k) = s_n^*(1;\theta) = 1$ and bid functions are continuous and increasing, for any $\delta > 0$ we can find an ϵ such that $s_n^*(\epsilon;\theta) < \delta/3$. There exists a K such that for all $k > K$,

$$|s_n^*(x;\theta_k) - s_n^*(x;\theta)| < \delta/3, \ \forall x \in [1+\epsilon, 1+\overline{v}^*].$$

If $x \in [1, 1+\epsilon]$, one can easily show that f $k > K$,

$$|s_n^*(x;\theta_k) - s_n^*(x;\theta)| \leq \sup\left(|s_n^*(1;\theta_k) - s_n^*(\epsilon;\theta)|, |s_n^*(\epsilon;\theta_k) - s_n^*(1;\theta)|\right) < \delta$$

Because the above inequality holds for all $\delta > 0$ and it is independent of x, we can conclude $s_n^*(\cdot;\theta_k)$ converges to $s_n^*(\cdot;\theta)$ uniformly on $[1, 1+\bar{v}^*]$ if $\|\theta_k - \theta\|_s \to 0$ if $\bar{v}^* < \infty$.

If $\bar{v}^* = \infty$, by the above argument, for any $c < \infty$, $\sup_{v\in[1,c+1]} |s_n^*(v;\theta_k) - s_n^*(v;\theta)| \to 0$. By the FOC, for any $\theta \in \Theta$

$$\frac{\partial s_n^*(v;\theta)}{\partial v} = \frac{n-1}{1-\sigma}\frac{f^*(v;\theta)}{F(v;\theta)}\left(v - s_n^*(v;\theta)\right) \leq C\frac{n-1}{\eta}\frac{h^*(v)}{H(v)}v$$

for some constant C. The last inequality holds because of Lemma 6 and $s_n^*(v;\theta) \geq 0$. By Assumption 4, $h^*(v)$ has a tail bounded by $C/v^{2+\delta}$ for some $\delta > 0$, for $v_1 > v_2$ large enough,

$$|s_n^*(v_2;\theta) - s_n^*(v_1;\theta)| \leq C\int_{v_2}^{v_1}\frac{h^*(v)}{H(v_2)}vdv \leq \frac{C}{H(v_2)}v_2^{-\delta}.$$

Hence, for any $\epsilon > 0$, there exists a $v(\epsilon) < \infty$ such that for $v > v(\epsilon)$, $|s_n^*(v(\epsilon);\theta) - s_n^*(v;\theta)| \leq \epsilon$. Therefore,

$$\begin{aligned}
& \sup_{v\geq v(\epsilon)} |s_n^*(v;\theta_k) - s_n^*(v;\theta)| \\
\leq\ & \sup_{v\geq v(\epsilon)} |s_n^*(v;\theta_k) - s_n^*(v(\epsilon);\theta_k)| + \sup_{v\geq v(\epsilon)} |s_n^*(v;\theta) - s_n^*(v(\epsilon);\theta)| \\
& + |s_n^*(v(\epsilon);\theta_k) - s_n^*(v(\epsilon);\theta)| \leq 2\epsilon + |s_n^*(v(\epsilon);\theta_k) - s_n^*(v(\epsilon);\theta)|
\end{aligned}$$

Consequently, $\sup_{v\in[1,\bar{v}^*+1]} |s_n^*(v;\theta_k) - s_n^*(v;\theta)| \to 0$, because the following holds for any

$\epsilon > 0$

$$
\begin{aligned}
\sup_{v\in[1,\infty]}\left|s_n^*(v;\theta_k)-s_n^*(v;\theta)\right| &\leq \sup_{v\in[1,v(\epsilon)]}\left|s_n^*(v;\theta_k)-s_n^*(v;\theta)\right|+\sup_{v\geq v(\epsilon)}\left|s_n^*(v;\theta_k)-s_n^*(v;\theta)\right| \\
&\leq 2\epsilon+2\sup_{v\in[1,v(\epsilon)]}\left|s_n^*(v;\theta_k)-s_n^*(v;\theta)\right|\to 2\epsilon.
\end{aligned}
$$

□

Lemma 8. *For every $n \in \mathbf{N}$ and $\theta_k, \theta \in \Theta$, if $\|\theta_k-\theta\|_s \to 0$ then $g_n^*(x;\theta_k) \to g_n^*(x;\theta)$ x-a.e.*

Proof. Because $s_n^*(v;\theta)$ is strictly increasing on $[1,\overline{v}^*+1]$ for all $\theta\in\Theta$, by Lemma 7, for any $1<x<s_n^*(\bar{v}^*;\theta)$, $s_n^{*-1}(x;\theta_k)\to s_n^{*-1}(x;\theta)$. If not, we can find an $v\neq s_n^{*-1}(x;\theta)$ such that $s_n^*(v;\theta)=x$ which violates the strictly increasing property. For such an x, $s_n^{*-1}(x;\theta_k)-x\to s_n^{*-1}(x;\theta)-x>0$,

$$
\begin{aligned}
& \left|F^*\left(s_n^{*-1}(x;\theta_k);\theta_k\right)-F^*\left(s_n^{*-1}(x;\theta);\theta\right)\right| \\
\leq\; & \left|F^*\left(s_n^{*-1}(x;\theta_k);\theta_k\right)-F^*\left(s_n^{*-1}(x;\theta);\theta_k\right)\right|+\left|F^*\left(s_n^{*-1}(x;\theta);\theta_k\right)-F^*\left(s_n^{*-1}(x;\theta);\theta\right)\right| \\
\leq\; & \|f^*(\cdot;\theta_k)\|_\infty\left|s_n^{*-1}(x;\theta_k)-s_n^{*-1}(x;\theta)\right|+\|F^*(\cdot;\theta)-F^*(\cdot;\theta_k)\|_\infty \\
\leq\; & C\left(\left|s_n^{*-1}(x;\theta_k)-s_n^{*-1}(x;\theta)\right|+\|\theta_k-\theta\|_s\right)\to 0
\end{aligned}
$$

In addition, $\frac{1-\sigma_k}{n-1}\to\frac{1-\sigma}{n-1}$

$$
g_n^*(x;\theta_k)=\frac{1-\sigma_k}{n-1}\frac{F^*\left(s_n^{*-1}(x;\theta_k);\theta_k\right)}{s_n^{*-1}(x;\theta_k)-x}\to\frac{1-\sigma}{n-1}\frac{F^*\left(s_n^{*-1}(x;\theta);\theta\right)}{s_n^{*-1}(x;\theta)-x}=g_n^*(x;\theta).
$$

It is easy to see for any $x<1$, $g_n(x;\theta_k)=g_n(x;\theta)=0$. In addition, if $x>s_n^*(\overline{v}^*+1;\theta)$, $x>s_n^*(\overline{v}^*+1;\theta_k)$ for k large enough by Lemma 7. Hence $g_n(x;\theta_k)=g_n(x;\theta)=0$ for all k large enough. If $x=1$, $g_n(1;\theta_k)=\frac{f^*(1;\theta_k)(1-\sigma_k)}{n-\sigma_k}\to g_n(1;\theta)$. Hence $g_n^*(x;\theta_k)\to g_n^*(x;\theta)$ x-a.e. □

Lemma 9. There exists a constant $C>0$ such that, $g_n^*(\cdot;\theta)\leq C$ for all $\theta\in\Theta$.

Proof. For any $v \in [1, \overline{v}^* + 1]$, the first order condition implies

$$g_n^*\left(s_n^*\left(v;\theta\right);\theta\right) = \frac{1-\sigma}{n-1}\frac{F^*\left(v;\theta\right)}{v - s_n^*\left(v;\theta\right)} = \frac{1-\sigma}{n-1}\frac{F^*\left(v;\theta\right)}{\int_1^v \left[\frac{F^*(s;\theta)}{F^*(v;\theta)}\right]^{\frac{n-1}{1-\sigma}} ds} = \frac{1-\sigma}{n-1}\frac{F^*\left(v;\theta\right)^{\frac{n-\sigma}{1-\sigma}}}{\int_1^v F^*\left(s;\theta\right)^{\frac{n-1}{1-\sigma}} ds} \quad (13)$$

By Lemma 6, for $C_1 = \frac{\eta^2}{1+B}$ and $C_2 = \left(1 + 3\sqrt{B}\right)^2$,

$$g_n^*\left(s_n^*\left(v;\theta\right);\theta\right) \leq \frac{1-\sigma}{n-1}\frac{C_2^{\frac{n-\sigma}{1-\sigma}}}{C_1^{\frac{n-\sigma}{1-\sigma}}}\frac{H^*\left(v\right)^{\frac{n-\sigma}{1-\sigma}}}{\int_0^v H^*\left(s\right)^{\frac{n-1}{1-\sigma}} ds} \leq C_3 \frac{H^*\left(v\right)^{\frac{n-\sigma}{1-\sigma}}}{\int_1^v H^*\left(s\right)^{\frac{n-1}{1-\sigma}} ds} \leq C_3 \frac{H^*\left(v\right)^{1+\frac{n-1}{\eta}}}{\int_1^v H^*\left(s\right)^{\frac{n-1}{\eta}} ds} \quad (14)$$

Notice that

$$\lim_{v \to 1} \frac{H^*\left(v\right)^{1+\frac{n-1}{\eta}}}{\int_1^v H^*\left(s\right)^{\frac{n-1}{\eta}} ds} = \left(\frac{n-1}{\eta} + 1\right) h^*\left(1\right)$$

and let v_1 be $H^*\left(v_1\right) = 1/2$, for any $v > v_1$

$$\frac{H^*\left(v\right)^{1+\frac{n-1}{\eta}}}{\int_1^v H^*\left(s\right)^{\frac{n-1}{\eta}} ds} < \frac{2^{\frac{n-1}{\eta}}}{v_1}$$

Therefore,

$$g_n^*\left(s_n^*\left(v;\theta\right);\theta\right) \leq C_3 \max\left\{\left(\frac{n-1}{\eta} + 1\right) h^*\left(1\right), \max_{v \in [1, v_1]} \frac{H^*\left(v\right)^{1+\frac{n-1}{\eta}}}{\int_1^v H^*\left(s\right)^{\frac{n-1}{\eta}} ds}\right\} = C$$

$\max_{v \in [1, v_1]} \frac{H^*(v)^{1+\frac{n-1}{\eta}}}{\int_1^v H^*(s)^{\frac{n-1}{\eta}} ds}$ is bounded because it is continuous in v on $(1, v_1]$ and smaller than $\left(\frac{n-1}{\eta} + 1\right) h^*\left(1\right)$ if v approaches 1. □

Lemma 10. *$g_n\left(\log \mathbf{b} / \exp\left(\log X\gamma\right);\theta\right)$ is continuous in θ on $\Theta\left(\mathbf{b}, X\right)$-a.e..*

Proof. $\forall\theta\in\Theta$ and $\theta_k\to\theta$ under $\|\cdot\|_s$,

$$\begin{aligned}
& \left| g_n\left(\frac{\mathbf{b}}{\exp(\log X\gamma_k)};\theta_k\right) - g_n\left(\frac{\mathbf{b}}{\exp(\log X\gamma)};\theta\right)\right| \\
= & \left| \int \frac{1}{u^n}\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu)\exp(\log X\gamma)};\theta\right) f_n^u(u;\theta)\,du \right.\\
& \left. -\int \frac{1}{u^n}\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu_k)\exp(\log X\gamma_k)};\theta_k\right) f_n^u(u;\theta_k)\,du\right| \\
\le & \; C\int \prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu_k)\exp(\log X\gamma_k)};\theta_k\right)\left|f_n^u(u;\theta_k)-f_n^u(u;\theta)\right| du + \\
& C\int\left|\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu_k)\exp(\log X\gamma_k)};\theta_k\right)-\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu)\exp(\log X\gamma)};\theta\right)\right| f_n^u(u;\theta)\,du \\
= & \; A_1+A_2.
\end{aligned}$$

The first inequality holds because μ has to be strictly greater than some positive number. Hence, $u\ge\mu$ is bounded from below. Consequently, there exists $C>1/u^n$ for all u. Next, we show that A_1 and A_2 converge to 0.

$$A_1\le C\|\theta_k-\theta\|_s\int\prod_{i=1}^n g_n^*\left(b_i/(u+\mu_k)\exp(\log X\gamma_k);\theta_k\right)du\le C\|\theta_k-\theta\|_s\to 0.$$

The first inequality holds by Lemma 6 and the second inequality holds by the fact that g_n^* is bounded and it is a density function.

$$\begin{aligned}
A_2 \le & \int\left|\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu_k)\exp(\log X\gamma_k)};\theta_k\right)-\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu_k)\exp(\log X\gamma_k)};\theta\right)\right| f_n^u(u;\theta)\,du \\
& +\int\left|\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu_k)\exp(\log X\gamma_k)};\theta\right)-\prod_{i=1}^n g_n^*\left(\frac{b_i}{(u+\mu)\exp(\log X\gamma)};\theta\right)\right| f_n^u(u;\theta)\,du \\
= & \; B_1+B_2
\end{aligned}$$

First, use a change of variables to rewrite

$$B_1 = \int\left|\prod_{i=1}^n g_n^*(b_i/u;\theta_k)-\prod_{i=1}^n g_n^*(b_i/u;\theta)\right| f_n^u\left(u/\exp(\log X\gamma_k)-\mu_k;\theta\right)du.$$

$\left|\prod_{i=1}^n g_n^*(b_i/u;\theta_k)-\prod_{i=1}^n g_n^*(b_i/u;\theta)\right|\to 0$ u-a.e. by Lemma 8. In addition, by Lemma 9 for some $C<\infty$

$$\left|\prod_{i=1}^{n} g_n^*(b_i/u;\theta_k)-\prod_{i=1}^{n} g_n^*(b_i/u;\theta)\right| f_n^u(u/\exp(\log X\gamma_k)-\mu_k;\theta)\leq Cf_n^u(u/\exp(\log X\gamma_k)-\mu_k;\theta)$$

Notice that $Cf_n^u(u/\exp(\log X\gamma_k)-\mu_k;\theta)\to Cf_n^u(u/\exp(\log X\gamma)-\mu;\theta)$ u-a.e. and

$$\int Cf_n^u(u/\exp(\log X\gamma_k)-\mu_k;\theta)\,du=C\exp(X\gamma_k)\to C\exp(X\gamma)=\int Cf_n^u(u/\exp(\log X\gamma)-\mu;\theta)$$

The generalized Dominated Convergence Theorem implies $B_1\to 0$. Similarly, the Dominated Convergence Theorem implies $B_2\to 0$. Consequently, $g_n(\log\mathbf{b}/\exp(\log X\gamma),\theta)$ is continuous in θ $(\mathbf{b},X)$-a.e.. □

Lemma 11. *$l(Z,\theta)$ is continuous in θ on Θ Z-a.e..*

Proof. $l(Z,\theta)$ is the log of a continuous function g_n by Lemma 10. Therefore, it is continuous in θ, Z-a.e. □

Lemma 12. *Under Assumptions 4 and 5, there exists a sequence $\{\theta_{0,k_L}\}_{k_L=1}^{\infty}$ such that $\theta_{0,k_L}\in\Theta_{k_L}$, $\lim_{k_L\to\infty}\|\theta_{0,k_L}-\theta_0\|_s=0$, and $E_0 l(Z_\ell,\theta_{0,k_L})-E_0 l(Z_\ell,\theta_0)\to 0$.*

Proof. By Assumption 5, there exists $\alpha_{0,k_L}\in\mathcal{A}_{k_L}$ such that $\|\alpha_{0,k_L}-\alpha_0\|_\infty\to 0$. Now we find $\sigma_{0,k_L}\downarrow\sigma_0$ such that $\theta_{0,k_L}=(\sigma_{0,k_L},\gamma_0,\mu_0,\alpha_{0,k_L})$ which satisfies $s_n^*(\cdot;\theta_0)\leq s_n^*(\cdot;\theta_{0,k_L})\leq C(s_n^*(\cdot;\theta_0)-1)+1$ for some $C>1$. Then we show that this sequence of θ_{0,k_L} is the desirable sequence.

Recall the first order condition

$$\frac{\partial s_n^*(v;\theta)}{\partial v}=\frac{n-1}{1-\sigma}\frac{f^*(v;\theta)}{F^*(v;\theta)}(v-s_n^*(v;\theta)).$$

Therefore $s_n^*(\cdot;\theta_0)\leq s_n^*(\cdot;\theta_{0,k_L})$ if $\frac{n-1}{1-\sigma_0}\frac{f^*(v;\theta_0)}{F^*(v;\theta_0)}\leq\frac{n-1}{1-\sigma_{0,k_L}}\frac{f^*\left(v;\theta_{0,k_L}\right)}{F^*\left(v;\theta_{0,k_L}\right)}$, or $\frac{1-\sigma_0}{n-1}\frac{F^*(v;\theta_0)}{f^*(v;\theta_0)}\geq\frac{1-\sigma_{0,k_L}}{n-1}\frac{F^*\left(v;\theta_{0,k_L}\right)}{f^*\left(v;\theta_{0,k_L}\right)}$ which in turn can be written as

$$\frac{1-\sigma_0}{n-1}\frac{\int_1^v(T\psi_0^*)\circ H^*(s)\,h^*(s)\,ds}{(T\psi_0^*)\circ H^*(v)\,h^*(v)}\geq\frac{1-\sigma_{0,k_L}}{n-1}\frac{\int_1^v\left(T\psi_{0,k_L}^*\right)\circ H^*(s)\,h^*(s)\,ds}{\left(T\psi_{0,k_L}^*\right)\circ H^*(v)\,h^*(v)}$$

Because $\left\|\psi_{0,k_L}^{*}-\psi_{0}^{*}\right\|_{\infty}\to 0$, by Lemma 6,

$$\sup_{v}\left|\left(T\psi^{*}\right)\circ H^{*}\left(v\right)-\left(T\psi_{k_L}^{*}\right)\circ H^{*}\left(v\right)\right|=\epsilon_{k_L}\to 0$$

By (12), $\frac{\eta^2}{1+B}\leq\left(T\psi^{*}\right)\circ H^{*}\left(v\right)\leq\left(1+3\sqrt{B}\right)^{2}$. Therefore, for k_L large enough

$$\left[1-\frac{\epsilon_{k_L}\left(1+B\right)}{\eta^{2}}\right]\left(T\psi_{0}^{*}\right)\circ H^{*}\left(v\right)h^{*}\left(v\right)\leq\left(T\psi_{0,k_L}^{*}\right)\circ H^{*}\left(v\right)h^{*}\left(v\right)\leq\left[1+\frac{\epsilon_{k_L}\left(1+B\right)}{\eta^{2}}\right]\left(T\psi_{0}^{*}\right)\circ H^{*}\left(v\right)h^{*}\left(v\right)$$

which suggests

$$\left[1-\frac{\epsilon_{k_L}\left(1+B\right)}{\eta^{2}}\right]f^{*}\left(v;\theta_{0}\right)\leq f^{*}\left(v;\theta_{0,k_L}\right)\leq\left[1+\frac{\epsilon_{k_L}\left(1+B\right)}{\eta^{2}}\right]f^{*}\left(v;\theta_{0}\right)\tag{15}$$

$$\left[1-\frac{\epsilon_{k_L}\left(1+B\right)}{\eta^{2}}\right]F^{*}\left(v;\theta_{0}\right)\leq F^{*}\left(v;\theta_{0,k_L}\right)\leq\left[1+\frac{\epsilon_{k_L}\left(1+B\right)}{\eta^{2}}\right]F^{*}\left(v;\theta_{0}\right)\tag{16}$$

(15) and (16) together imply

$$\frac{1-\sigma_0}{n-1}\frac{F^{*}\left(v;\theta_{0}\right)}{f^{*}\left(v;\theta_{0}\right)}\geq\frac{1-\sigma_0}{n-1}\frac{\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]F^{*}\left(v;\theta_{0,k_L}\right)}{\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]f^{*}\left(v;\theta_{0,k_L}\right)}\geq\frac{1-\sigma_0}{n-1}\frac{F^{*}\left(v;\theta_{0}\right)}{f^{*}\left(v;\theta_{0}\right)}\frac{\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]^{2}}{\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]^{2}}$$

Define $\sigma_{k_L}=1-\left(1-\sigma_0\right)\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]/\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]\downarrow\sigma_0$ as $\epsilon_{k_L}\to 0$. Then let $C>\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]^{2}/\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^{2}}\right]^{2}$ for k_L large enough, then because $s_n^{*}\left(v;\theta_0\right)\leq s_n^{*}\left(v;\theta_{0,k_L}\right)$

$$C\frac{\partial s_n^{*}\left(v;\theta\right)}{\partial v}=C\frac{n-1}{1-\sigma_0}\frac{f^{*}\left(v;\theta_0\right)}{F^{*}\left(v;\theta_0\right)}\left(v-s_n^{*}\left(v;\theta_0\right)\right)\geq\frac{n-1}{1-\sigma_{0,k_L}}\frac{f^{*}\left(v;\theta_{0,k_L}\right)}{F^{*}\left(v;\theta_{0,k_L}\right)}\left(v-s_n^{*}\left(v;\theta_{0,k_L}\right)\right)=\frac{\partial s_n^{*}\left(v;\theta_{0,k_L}\right)}{\partial v}$$

Therefore, $s_n^{*}\left(\cdot;\theta_{k_L}\right)\leq C\left(s_n^{*}\left(\cdot;\theta_0\right)-1\right)+1$ where 1 comes from the initial condition $s^{*}\left(1;\theta\right)=1$. Notice θ_{0,k_L} converges to θ_0 under $\left\|\cdot\right\|_s$.

It is left to show that $E_0 l\left(Z_\ell;\theta_{0,k_L}\right)-E_0 l\left(Z_\ell;\theta_0\right)\to 0$. First, notice $l\left(Z;\theta_{0,k_L}\right)\to l\left(Z;\theta_0\right)$ Z-a.e. by Lemma 11. Then if $\left|l\left(Z;\theta_{0,k_L}\right)\right|$ is bounded by a integrable function, the Dominated Convergence Theorem implies that $E_0 l\left(Z_\ell;\theta_{0,k_L}\right)-E_0 l\left(Z_\ell;\theta_0\right)\to 0$. First notice, Lemma 9 implies that $g_n^{*}\left(\cdot;\theta\right)<C$. Therefore $g_n^{*}\left(\mathbf{b};\theta\right)=\int\frac{1}{u^n}\prod_{i=1}^{n}g_n^{*}\left(b_i;\theta\right)f_n^{u}\left(u;\theta\right)du<C$ for some constant C . Because $l\left(Z;\theta_{0,k_L}\right)$ is the log of g_n^{*}, it is bounded from above by a constant.

Then it suffices to show that $l\left(Z;\theta_{0,k_L}\right)$ is bounded from below by an integrable function. To this end we show $g_n\left(b;\theta_{0,k_L}\right) > C_1 g_n\left(b;\theta_0\right)$ for some constant $C_1 > 0$, which implies $l\left(Z;\theta_{0,k_L}\right) \geq l\left(Z;\theta_0\right) + C_1$.

We first show $b \in [1, s_n^*\left(1+\bar{v}^*;\theta_0\right)]$, $g_n^*\left(b;\theta_{0,k_L}\right) \geq C_2 g_n^*\left(b;\theta_0\right)$ for some constant $C_2 > 0$. Use (13) to obtain

$$\frac{g_n^*\left(b;\theta_{0,k_L}\right)}{g_n^*\left(b;\theta_0\right)} = \frac{1-\sigma_{k_L}}{1-\sigma_0}\frac{F^*\left(s_n^{*-1}\left(b;\theta_{0,k_L}\right);\theta_{0,k_L}\right)^{\frac{n-\sigma_{0,k_L}}{1-\sigma_{0,k_L}}}}{F^*\left(s_n^{*-1}\left(b;\theta_0\right);\theta_0\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}\frac{\int_1^{s_n^{*-1}(b;\theta_0)} F^*\left(s;\theta_0\right)^{\frac{n-1}{1-\sigma_0}}\,ds}{\int_1^{s_n^{*-1}\left(b;\theta_{0,k_L}\right)} F^*\left(s;\theta_{0,k_L}\right)^{\frac{n-1}{1-\sigma_{0,k_L}}}\,ds}$$

By Lemma 6 and the fact $\sigma_{0,k_L} \geq \sigma_0$, there exists a constant C_3 such that

$$\begin{aligned}
\frac{g_n^*\left(b;\theta_{0,k_L}\right)}{g_n^*\left(b;\theta_0\right)} &\geq C_3\frac{H^*\left(s_n^{*-1}\left(b;\theta_{k_L}\right)\right)}{H^*\left(s_n^{*-1}\left(b;\theta_0\right)\right)}\frac{\int_1^{s_n^{*-1}(b;\theta_0)}\left[\frac{H^*(s)}{H^*\left(s_n^{*-1}(b;\theta_0)\right)}\right]^{\frac{n-1}{1-\sigma_0}}ds}{\int_1^{s_n^{*-1}\left(b;\theta_{k_L}\right)}\left[\frac{H^*(s)}{H^*\left(s_n^{*-1}\left(b;\theta_{k_L}\right)\right)}\right]^{\frac{n-1}{1-\sigma_{0,k_L}}}ds} \\
&\geq C_3\frac{H^*\left(s_n^{*-1}\left(b;\theta_{0,k_L}\right)\right)}{H^*\left(s_n^{*-1}\left(b;\theta_0\right)\right)}\frac{\int_1^{s_n^{*-1}(b;\theta_0)}\left[\frac{H^*(s)}{H^*\left(s_n^{*-1}(b;\theta_0)\right)}\right]^{\frac{n-1}{1-\sigma_0}}ds}{\int_1^{s_n^{*-1}\left(b;\theta_{0,k_L}\right)}\left[\frac{H^*(s)}{H^*\left(s_n^{*-1}\left(b;\theta_{0,k_L}\right)\right)}\right]^{\frac{n-1}{1-\sigma_0}}ds} \\
&= C_3\frac{H^*\left(s_n^{*-1}\left(b;\theta_{0,k_L}\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}{H^*\left(s_n^{*-1}\left(b;\theta_0\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}\frac{\int_1^{s_n^{*-1}(b;\theta_0)} H^*\left(s\right)^{\frac{n-1}{1-\sigma_0}}\,ds}{\int_1^{s_n^{*-1}\left(b;\theta_{0,k_L}\right)} H^*\left(s\right)^{\frac{n-1}{1-\sigma_0}}\,ds}
\end{aligned}$$

First, because $s_n^{*-1}\left(b;\theta_0\right) \geq s_n^{*-1}\left(b;\theta_{0,k_L}\right)$,

$$\frac{g_n^*\left(b;\theta_{0,k_L}\right)}{g_n^*\left(b;\theta_0\right)} \geq C_3\frac{H^*\left(s_n^{*-1}\left(b;\theta_{0,k_L}\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}{H^*\left(s_n^{*-1}\left(b;\theta_0\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}$$

If $b = s_n^*\left(\bar{v}^*+1;\theta_0\right)$,

$$\frac{g_n^*\left(b;\theta_{0,k_L}\right)}{g_n^*\left(b;\theta_0\right)} \geq C_3\frac{H^*\left(s_n^{*-1}\left(b;\theta_{k_L}\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}{H^*\left(\bar{v}^*+1\right)^{\frac{n-\sigma_0}{1-\sigma_0}}} \geq C_3/2^{1+\frac{n-1}{\eta}} \tag{17}$$

for large k_L. This is because we can find an $\epsilon > 0$ such that $s_n^{*-1}\left(b-\epsilon;\theta_0\right) > v_H\left(1/2\right)$ where $H^*\left(v_H\left(1/2\right)\right) = 1/2$. Because $s_n^{*-1}\left(b-\epsilon;\theta_{0,k_L}\right) \to s_n^{*-1}\left(b-\epsilon;\theta_0\right)$ for any $\epsilon > 0$ and H^* is continuous, $H^*\left(s_n^{*-1}\left(b;\theta_{0,k_L}\right)\right) > 1/2$. Because $\frac{n-\sigma_0}{1-\sigma_0} \leq 1 + \frac{n-1}{\eta}$, the last inequality in (17) holds. In addition, because $s_n^*(\cdot;\theta_{k_L}) \le C(s_n^*(\cdot;\theta_0)-1)+1$

$$\frac{g_n^*(b;\theta_{k_L})}{g_n^*(b;\theta_0)} \ \ge\ C_3\frac{H^*\left(s_n^{*-1}(b;\theta_{k_L})\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}{H^*\left(s_n^{*-1}(b;\theta_0)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}} \ge C_3\frac{H^*\left(s_n^{*-1}\left(\frac{b-1}{C}+1;\theta_0\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}{H^*\left(s_n^{*-1}(b;\theta_0)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}$$

By Assumption 4,

$$\begin{aligned}\lim_{b\downarrow 1}\frac{H^*\left(s_n^{*-1}\left(\frac{b-1}{C}+1;\theta_0\right)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}}{H^*\left(s_n^{*-1}(b;\theta_0)\right)^{\frac{n-\sigma_0}{1-\sigma_0}}} &= \lim_{b\downarrow 1}\frac{\left(s_n^{*-1}\left(\frac{b-1}{C}+1;\theta_0\right)-1\right)^{\frac{n-\sigma_0}{1-\sigma_0}(1+\epsilon)}}{\left(s_n^{*-1}(b;\theta_0)-1\right)^{\frac{n-\sigma_0}{1-\sigma_0}(1+\epsilon)}}\\ &= \lim_{b\downarrow 1}\left[\frac{\frac{b-1}{C}\frac{1}{s_n^{*\prime}(1;\theta_0)}}{(b-1)\frac{1}{s_n^{*\prime}(1;\theta_0)}}\right]^{\frac{n-\sigma_0}{1-\sigma_0}(1+\epsilon)} = 1/C^{\frac{n-\sigma_0}{1-\sigma_0}(1+\epsilon)} = C_4\end{aligned}$$

Because $\frac{g_n^*\left(b;\theta_{0,k_L}\right)}{g_n^*(b;\theta_0)} > 0$ for all $b\in(1, s_n^*(\bar{v}^*+1;\theta_0))$ and it is continuous, there exists an $C_2>0$ such that $\frac{g_n^*\left(b;\theta_{0,k_L}\right)}{g_n^*(b;\theta_0)} \ge C_2$ on $[1, s_n^*(\bar{v}^*+1;\theta_0)]$. In addition, $g_n^*(b;\theta)=0$ if $b\notin[1, s_n^*(\bar{v}^*+1;\theta_0)]$. $g_n^*(b;\theta_{0,k_L}) \ge C_2 g_n^*(b;\theta_0)$ for k_L large enough.By Lemma 6, $f_n^u(u;\theta_{k_L}) > C_5 f_n^u(u;\theta_0)$

$$\begin{aligned}g_n(\mathbf{b};\theta_{0,k_L}) &= \int\frac{1}{u^n}\prod_{i=1}^n g_n^*(b_i/u;\theta_{0,k_L})\, f_n^u(u;\theta_{0,k_L})\,du \ge C_5\int\frac{1}{u^n}\prod_{i=1}^n g_n^*(b_i/u;\theta_k)\, f_n^u(u;\theta_0)\,du\\ &\ge C_5C_2\int\frac{1}{u^n}\prod_{i=1}^n g_n^*(b_i/u;\theta_0)\, f_n^u(u;\theta_0)\,du = C_1 g_n(\mathbf{b};\theta_0)\end{aligned}$$

Therefore, $l(Z_\ell;\theta_{0,k_L}) \ge l(Z_\ell;\theta_0)+\log C_1$. Then by the Dominated Convergence Theorem, $E_0 l(Z_\ell;\theta_{0,k_L}) \to E_0 l(Z_\ell;\theta_0)$.

$\square$

References

Ackerberg, D., X. Chen, and J. Hahn (2012): "A practical asymptotic variance estimator for two-step semiparametric estimators," *Review of Economics and Statistics*, 94(2), 481–498. 3, 12

Andrews, D. W. (1999): "Estimation when a parameter is on a boundary," *Econometrica*, 67(6), 1341–1383. 19

Aradillas-López, A., A. Gandhi, and D. Quint (2013a): "Identification and inference in ascending auctions with correlated private values," *Econometrica*, 81(2), 489–534. 4, 23

——— (2013b): "A Simple Test for Moment Inequality Models with an Application to English Auctions," . 4, 23

Athey, S., and J. Levin (2001): "Information and Competition in US Forest Service Timber Auctions," *Journal of Political Economy*, 109(2). 5, 23

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. 4, 23, 25

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. 4, 12, 18, 20, 24

Baldwin, L. (1995): "Risk averse bidders at oral ascending-bid Forest Service timber sales," *Unpublished Manuscript, RAND Corporation.[1193]*. 5, 23

Baldwin, L., R. Marshall, and J. Richard (1997): "Bidder Collusion at Forest Service Timber Sales," *The Journal of Political Economy*, 105(4), 657–699. 23

Bierens, H., and H. Song (2011): "Nonparametric Identification of the First-Price Auction Model," . 13

Bierens, H. J. (2014): "CONSISTENCY AND ASYMPTOTIC NORMALITY OF SIEVE ML ESTIMATORS UNDER LOW-LEVEL CONDITIONS," *Econometric Theory*, 30, 1021–1076. 16

Bierens, H. J., and H. Song (2012a): "Semi-nonparametric estimation of independently and identically repeated first-price auctions via an integrated simulated moments method," *Journal of Econometrics*, 168(1), 108 – 119. 14

Bierens, H. J., and H. Song (2012b): "Semi-nonparametric estimation of independently and identically repeated first-price auctions via an integrated simulated moments method," *Journal of Econometrics*, 168(1), 108–119. 18, 19

Campo, S. (2012): "Risk aversion and asymmetry in procurement auctions: identification, estimation and application to construction procurements," *Journal of Econometrics*, 168(1), 96–107. 5

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. 5, 23, 33

Chernozhukov, V., and H. Hong (2004): "Likelihood Estimation and Inference in a Class of Nonregular Econometric Models," *Econometrica*, 72(5), 1445–1480. 12

Cox, J. C., V. L. Smith, and J. M. Walker (1988): "Theory and individual behavior of first-price auctions," *Journal of Risk and uncertainty*, 1(1), 61–99. 4

d'Haultfoeuille, X., and P. Février (2010a): "Identification of a Class of Adverse Selection Models with Contracts Variation," Discussion paper, Mimeo. 9

——— (2010b): "Identification of mixture models using support variations," *Manuscript, ENSAE*. 2, 33

Donald, S. G., and H. J. Paarsch (1993): "Piecewise pseudo-maximum likelihood estimation in empirical models of auctions," *International Economic Review*, pp. 121–148. 12

Evdokimov, K., and H. White (2012): "Some Extensions of a Lemma of Kotlarski," *Econometric Theory*, 28(04), 925–932. 8, 28, 29, 30

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*. 5, 7, 25, 26

Grundl, S., and Y. Zhu (2015): "Inference of Risk Aversion in First-Price Auctions With Unobserved Auction Heterogeneity," Discussion paper. 13, 18, 19

Guerre, E., I. Perrigne, and Q. Vuong (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions," *Econometrica*, 77(4), 1193–1227. 1, 2, 3, 5, 6, 7, 26, 28, 30, 32, 33, 34

Haile, P. (2001): "Auctions with Resale Markets: An Application to US Forest Service Timber Sales," *The American Economic Review*, 91(3), 399–427. 23

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, pp. 1–51. 23

Hirano, K., and J. R. Porter (2003): "Asymptotic Efficiency in Parametric Structural Models with Parameter-Dependent Support," *Econometrica*, 71(5), pp. 1307–1338. 12

Holt, C. A. (1980): "Competitive bidding for contracts under alternative auction procedures," *The Journal of Political Economy*, pp. 433–445. 2

Hu, A., S. A. Matthews, and L. Zou (2010): "Risk aversion and optimal reserve prices in first-and second-price auctions," *Journal of Economic Theory*, 145(3), 1188–1202. 2

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. 2, 4, 7, 8, 9, 32

Kagel, J., and D. Levin (2010): "Auctions: A Survey of Experimental Research, 1995-2010," *Handbook of Experimental Economics*, 2. 4

Kagel, J. H. (1995): "Auctions: A survey of experimental research. John H. Kagel, Alvin E. Roth, eds., The Handbook of Experimental Economics," . 4

Kim, D.-H. (2015a): "Flexible Bayesian analysis of first price auctions using a simulated likelihood," *Quantitative Economics*, forthcoming. 13

——— (2015b): "Nonparametric estimation of utility function in first-price sealed-bid auctions," *Economics Letters*, 126, 101–106. 5

Kotlarski, I. (1967): "On characterizing the gamma and normal distribution," *Pacific Journal of Mathematics*, 20(1), 69–76. 8

Krasnokutskaya, E. (2011): "Identification and estimation of auction models with unobserved heterogeneity," *The Review of Economic Studies*, 78(1), 293. 2, 4, 7, 8, 12, 18, 27

Krasnokutskaya, E. (2012): "Identification and Estimation of Auction Model with Two-Dimensional Unobserved Heterogeneity," *International Economic Review*, 53(3), 659–692. 4

Levin, D., and J. L. Smith (1994a): "Equilibrium in Auctions with Entry," *The American Economic Review*, 84(3), pp. 585–599. 2

——— (1994b): "Equilibrium in Auctions with Entry," *The American Economic Review*, 84(3), pp. 585–599. 40

Li, T., J. Lu, and L. Zhao (2015): "Auctions with Selective Entry and Risk Aversion Bidders: Theory and Evidence," *Rand Journal of Economics*, forthcoming. 36

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. 5, 23

Maskin, E., and J. Riley (1984): "Optimal auctions with risk averse buyers," *Econometrica: Journal of the Econometric Society*, pp. 1473–1518. 2

Matthews, S. (1987): "Comparing auctions for risk averse buyers: A buyer's point of view," *Econometrica: Journal of the Econometric Society*, pp. 633–646. 2

Riley, J., and W. Samuelson (1981): "Optimal auctions," *The American Economic Review*, 71(3), 381–392. 2

Roberts, J. W., and A. Sweeting (2010): "Bailouts and the Preservation of Competition," Discussion paper, National Bureau of Economic Research. 4, 23

——— (2013): "When Should Sellers Use Auctions?," *American Economic Review*, 103(5), 1830–61. 4, 23

Santos, A. (2012): "Inference in Nonparametric Instrumental Variables With Partial Identification," *Econometrica*, 80(1), 213–275. 14

Smith, J. L., and D. Levin (1996): "Ranking auctions with risk averse bidders," *Journal of Economic Theory*, 68(2), 549–561. 2

Smith, R. L. (1985): "Maximum Likelihood Estimation in a Class of Nonregular Cases," *Biometrika*, 72(1), pp. 67–90. 12, 13

van der Vaart, A., and J. Wellner (2000): *Weak Convergence and Empirical Processes: With Applications to Statistics (Springer Series in Statistics)*. Springer, corrected edn. 16

Zincenko, F. (2014): "Nonparametric Estimation of First-Price Auctions with Risk-Averse Bidders," *Unpublished Manuscript, University of Pittsburgh.* 5