BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel 2016-23

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

by Serafin Grundl and Yu Zhu

Bank of Canada Staff Working Paper 2016-23

May 2016

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

by

Serafin Grundl[1] and Yu Zhu[2]

[1]Federal Reserve Board of Governors
serafin.j.grundl@frb.gov

[2]Funds Management and Banking Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
zhuy@bankofcanada.ca

ISSN 1701-9397

Acknowledgements

We are very grateful to Amit Gandhi and Jack Porter for their advice and many helpful suggestions. We would also like to thank Andrés Aradillas-López, Emmanuel Guerre, Bruce Hansen and Xiaoxia Shi for their helpful comments.

Abstract

This paper shows point identification in first-price auction models with risk aversion and unobserved auction heterogeneity by exploiting multiple bids from each auction and variation in the number of bidders. The required exclusion restriction is shown to be consistent with a large class of entry models. If the exclusion restriction is violated, but weaker restrictions hold instead, the same identification strategy still yields valid bounds for the primitives. We propose a sieve maximum likelihood estimator. A series of Monte Carlo experiments illustrate that the estimator performs well in finite samples and that ignoring unobserved auction heterogeneity can lead to a significant bias in risk-aversion estimates. In an application to U.S. Forest Service timber auctions we find that the bidders are risk neutral, but we would reject risk neutrality without accounting for unobserved auction heterogeneity.

JEL classification: C57, C14, D44, L00
Bank classification: Econometric and statistical methods

Résumé

Nous présentons une méthode d'identification ponctuelle dans le cadre de modèles d'enchères au premier prix où sont prises en compte l'aversion au risque et une hétérogénéité non observée des enchères. La méthode d'identification proposée s'appuie sur de multiples offres tirées de chaque enchère et sur le nombre variable d'offreurs. L'hypothèse d'exclusion retenue est compatible avec une classe étendue de modèles formalisant les décisions d'entrée. Si cette hypothèse d'exclusion est violée mais que des restrictions moins strictes demeurent, la même stratégie d'identification aboutit à des bornes valides pour les primitives. Nous proposons un estimateur du maximum de vraisemblance par tamisage local. À partir d'une série de simulations de Monte-Carlo, il est montré que cet estimateur donne de bons résultats sur des échantillons finis et que ne pas prendre en compte l'hétérogénéité non observée des enchères peut causer un biais significatif des estimations de l'aversion au risque. En appliquant notre méthode aux enchères organisées par le Service des forêts des États-Unis pour l'adjudication de bois d'œuvre, nous constatons que les offreurs sont neutres à l'égard du risque; cependant, en ignorant l'hétérogénéité non observée des enchères, nous rejetterions cette neutralité.

Classification JEL : C57, C14, D44, L00
Classification de la Banque : Méthodes économétriques et statistiques

Non-Technical Summary

Bidders' risk attitude is crucial to auction design. It greatly influences the optimal format of the auction, as well as the optimal reserve price in first-price auctions. Previous papers show that one can estimate bidders' risk attitude from bid data, and they find that bidders are risk averse using U.S. Forest Service (USFS) timber auctions. However, these papers ignore unobserved heterogeneity of auctions, which refers to characteristics of auctioned objects that are observed by the bidders but not by the econometricians. Unobserved auction heterogeneity is common in many auction environments and can potentially bias the estimates.

This paper shows that by exploiting multiple bids from the same auction and variation in the number of bidders, risk attitudes can be identified in auctions with unobserved heterogeneity. We propose a sieve maximum likelihood estimator to estimate the bidders' risk attitudes. Evidence from the USFS timber auctions shows that bidders are close to risk neutrality and that ignoring unobserved heterogeneity leads to significant overestimation of bidders' risk aversion.

1 Introduction

Estimating the risk aversion of bidders is important for auction design. 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] In first-price auctions, risk aversion reduces 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)).[2]

This paper studies identification and estimation of risk aversion in first-price auctions with unobserved auction heterogeneity, which is ubiquitous in applications to field data. We consider the workhorse model with symmetric independent private values and one-dimensional unobserved auction heterogeneity. We begin by showing point identification under an exclusion restriction, which 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 auction heterogeneity. 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 auction heterogeneity can be identified using the dependence among bids from the same auction that is created by the unobserved auction heterogeneity. Applying the techniques from the non-separable measurement error literature to first-price auctions with risk-averse bidders is not straightforward because it requires the highest bid to be strictly increasing in the unobserved auction heterogeneity. We provide new comparative statics results for auctions with risk-averse bidders to establish this monotonicity condition.

[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. Matthews (1987) compares auction formats from the perspective of risk-averse bidders.

[2] Maskin and Riley (1984) study the optimal auction mechanism under risk aversion.

In the second step, we apply the identification result of Guerre, Perrigne, and Vuong (2009) to the bid distributions conditional on the unobserved auction heterogeneity to recover the primitives. The exclusion restriction required for point identification is that the distribution of valuations conditional on the unobserved auction heterogeneity does not depend on the number of bidders. Hence, bidders are allowed to select into auctions based on the unobserved auction heterogeneity. We show that the exclusion restriction is satisfied under a wide range of entry models if the entry signals and valuations are independent across potential bidders, conditional on the unobserved auction heterogeneity. Intuitively, if the entry signals and valuations are independent across potential bidders, so are their entry decisions. Therefore, after conditioning on unobserved auction heterogeneity, an entrant's valuation is independent of other potential bidders' entry decisions. Hence, the distribution of valuations conditional on the unobserved auction heterogeneity does not depend on the number of entrants.

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 if risk neutrality is rejected, this conclusion remains valid.

Next, we turn to estimation. 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.[3] We propose a sieve maximum likelihood estimator and show its consistency under low-level conditions.[4] Monte Carlo experiments show that the estimator performs well with sample sizes commonly found in applications.

[3]Earlier applied work also considered multiplicative unobserved auction heterogeneity (e.g., Krasnokutskaya (2011) or Athey, Levin, and Seira (2011)).

[4]Deriving the asymptotic distribution of the estimator is beyond the scope of this paper 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. While this result does not apply here, the Monte Carlo results suggest that treating the problem as parametric works well in practice.

The Monte Carlo study also shows that ignoring unobserved auction heterogeneity can lead to a significant bias in risk-aversion estimates. Interestingly, the sign of the bias depends on the correlation between the unobserved auction heterogeneity and the number of bidders, because there are two opposing effects. First, if auctions with better unobserved auction heterogeneity attract more bidders, the shift of the (unconditional) bid distribution as the number of bidders increases is amplified. This effect leads us to underestimate risk aversion because risk aversion has the opposite effect on the bid distribution.[5] Second, the unobserved auction heterogeneity increases the dispersion of bids. This effect leads us to overestimate risk aversion because risk aversion has the same effect on the bid distributions.[6] Which of the two effects dominates depends on how strongly the number of bidders is correlated with the unobserved auction heterogeneity.

In an application, we study U.S. Forest Service (USFS) timber auctions. We find that the bidders 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) consider 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 USFS timber auctions (e.g., Aradillas-López, Gandhi, and Quint (2013a), Aradillas-López, Gandhi, and Quint (2013b), Roberts and Sweeting (2013), Roberts and Sweeting (forthcoming) and Athey, Levin, and Seira (2011)).

The empirical literature on risk aversion in first-price auctions started with laboratory experiments where risk aversion had been proposed as an explanation of the overbidding

[5]Risk aversion tends to attenuate the shift of the bid distribution as the number of bidders increases, because it leads to aggressive bidding. Therefore, bids are close to valuations even for a low number of competitors, and the bid distribution cannot shift much as the number of bidders increases.

[6]Risk aversion also increases the dispersion of bids, because the bid function at the lower bound of the valuation distribution is not affected by risk aversion, while bidders with higher values bid more aggressively.

puzzle.[7] 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 alternative models, which give up the assumption of Bayesian Nash Equilibrium. Several papers found the bidders in USFS timber auctions to be risk averse, relying on different restrictions for the identification of risk aversion.[8]

The only other paper we are aware of that considers nonparametric identification with risk-averse bidders and unobserved auction heterogeneity is Guerre, Perrigne, and Vuong (2009).[9] They provide conditions to ensure that the model can be identified if an instrument is available, which affects the number of bidders but not the distribution of valuations.[10] As such an instrument is difficult to find in many applications, we exploit multiple bids from the same auction to achieve identification with unobserved auction heterogeneity.

In a complementary paper to ours, 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 their bid. Therefore, the result of Guerre, Perrigne, and Vuong (2009) no longer applies in their model, and identification is more challenging. They show that a parametric restriction on the copula governing entry usually restores the point identification of all primitives, while a parametric restriction of the utility function leads to point identification of the utility function and to

[7] 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 (e.g. Cox, Smith, and Walker (1988)). For further references, see the excellent surveys by Kagel (1995) and Kagel and Levin (2010).

[8] 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 auctions, see also Baldwin (1995) and Athey and Levin (2001). Campo (2012) finds evidence of risk aversion in construction procurement auctions. Kong (2015) finds moderate levels of risk aversion in oil and gas auctions which explains the revenue difference between first-price and ascending auctions.

[9] Kim (2015b) and Zincenko (2014) propose nonparametric estimators to implement their main identification result without unobserved auction heterogeneity.

[10] 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 auction heterogeneity. Under the second condition, there is a monotone mapping between the instrument and the unobserved auction heterogeneity. These monotonicity assumptions allow the econometrician to identify the unobserved auction heterogeneity for each auction, and then proceed as if the unobserved auction heterogeneity is observed, in order to identify the distribution of valuations and the utility function. In contrast, our approach does not allow us to recover the unobserved auction heterogeneity for each auction, but identifying the bid distribution conditional on the unobserved auction heterogeneity is sufficient to identify the utility function.

partial identification of the remaining primitives.

The 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 and to illustrate the bias of risk-aversion estimates if unobserved auction heterogeneity is ignored. Section 5 is an application to USFS timber auctions, and Section 6 concludes.

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, where $U'(\cdot) \geq 0$, $U''(\cdot) \leq 0$, and U'' is continuous. The utility function is normalized such that $U(0) = 0$ and $U(1) = 1$.[11] 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 $s_n(\underline{v}(u), u) = \underline{v}(u)$. Guerre, Perrigne, and Vuong (2009, Proposition 3) showed that if u is observed, F and U are point identified from bid data under the following restriction:

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

In a model of entry, $F(\cdot|u,n)$ is generated by the equilibrium of an entry game among potential bidders.[12] If the potential bidders observe u before making their entry decision,

[11] As we normalize the utility function such that $U(1) = 1$, we implicitly 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.

[12] The primitive of this bigger model is the distribution of valuations conditional on u and the number of

they will select into auctions based on u. Therefore, the distribution of valuations without conditioning on u generally does vary with n. Once we condition on all the variables that are observed by the potential bidders, however, the distribution of valuations for entrants generally does not depend on n.[13] Indeed, we show in Appendix B.1 that Assumption 1 holds in many common entry models, including the Affiliated Signal Entry Model (Ye (2007), Gentry and Li (2014)) and its two polar cases considered in Levin and Smith (1994a) and Samuelson (1985).[14] The exclusion restriction is therefore consistent with selective entry once we condition on u. Even if Assumption 1 is violated, the estimated primitives still bound the true primitives under weaker restrictions, as shown in Theorem 3.

In applications to field data, we have to confront the possibility that u is not observed. Previous work studying such environments has assumed that bidders are risk neutral and focused on the identification of $F(\cdot|u)$ (Krasnokutskaya (2011) and Hu, McAdams, and Shum (2013)). The identification strategy exploits the fact that the data contain more than one bid for each auction. The unobserved auction heterogeneity 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) that considers cases where valuations consist of two independent and separable components.[15]

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, As-

potential bidders, whereas $F(\cdot|u,n)$ is no longer a primitive.

[13] We also have to condition on the number of potential bidders if it varies across auctions.

[14] Notice that, like most of the literature, we assume that the bidders know when they bid how many of their rivals decided to enter the auction. See Gentry, Li, and Lu (2015) for identification of risk aversion in first-price auctions if the bidders do not know when they bid how many of their rivals decided to enter the auction.

[15] It is worth noting that the assumption of independence between u and v^* 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 independence of u and v^* carries over to entrants. We show that independence is preseved if the potential bidders observe a signal for v^* *or* u, but generally not if they observe both.

sumption 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 auction heterogeneity enters valuations additively, then it also enters the equilibrium bid function additively – regardless of the utility function. If the unobserved auction heterogeneity enters valuations multiplicatively and the bidders have CRRA utility, then it also enters the bid function multiplicatively. If the utility function is not of the CRRA form, however, then the bidding strategy is not separable in u and the deconvolution techniques in Kotlarski (1967) can therefore no longer be applied.[16]

The result requires a location normalization. To see why, consider the additive case (Theorem 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(2).

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

If the unobserved auction heterogeneity does not enter in a separable way, establishing identification is more involved. Hu, McAdams, and Shum (2013) show identification under

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

the following monotonicity restriction on F if bidders are risk neutral and u takes on a finite number of different values .

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 is strictly increasing in u. This is an important requirement to apply the techniques from the non-separable measurement error literature. Hu, McAdams, and Shum (2013) establish this property by exploiting the closed form of the bidding strategy if the bidders are risk neutral. If the bidders are risk averse the bidding strategy typically does not have a closed form, and establishing strict monotonicity of the highest bid is therefore more involved.[17]

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

[17] To the best of our knowledge, this is a new comparative statics result for auctions with risk-averse bidders. 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.

Here, $[\underline{u}(n), \overline{u}(n)]$ is the support of the unobserved auction heterogeneity 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 a finite number of 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, which is required because observationally equivalent primitives can be constructed by applying monotone transformations to u.[18]

In both cases, the support restrictions for u allow us to match bid distributions from n_1- and n_2-bidder auctions on u. If u is discrete, we match bid distributions based on their first-order stochastic dominance ranking. To guarantee that the bid distributions with the same ranking correspond to the same u, the support of u must be invariant. If u is continuous, we can match the bid distributions based on the lower bound of their support due to the additional assumption that $\underline{v}(u)$ is strictly increasing. Therefore, it is sufficient if the two supports overlap.[19]

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 auctions with better unobserved auction heterogeneity might attract more bidders. In Appendix B.3, we confirm this intuition for the separable case covered in Theorem 1.

[18]Formally, 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 lead to the same distribution of valuations and bids.

[19]It is worth noting that the model with discrete u could also be identified with strictly increasing $\underline{v}(u)$ and overlapping support.

Formally, we show that the distribution of the unobserved auction heterogeneity 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 α-th 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 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 $\lambda(\cdot) = U(\cdot)/U'(\cdot)$ 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, however, the bounds 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, multiplicative observable auction characteristics and multiplicative unobserved auction heterogeneity. 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^*.[20] To simplify the notation, let F_n^u denote the distribution of the unobserved auction heterogeneity and let f_n^u be its density. The private values v^* and the unobserved auction heterogeneity u are independent of each other. 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

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

$u \exp[\log(X)\gamma]$ enters the bidding strategy multiplicatively (see Appendix A.2.1).

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.

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 under low-level conditions.[21,22]

3.1 Parameter Space

The supports of the densities of unobserved auction heterogeneity and the private values are $[\mu, \overline{u} + \mu]$ and $[1, \overline{v}^* + 1]$, with $\overline{u} > 0$ and $\overline{v}^* > 0$ known.[23] Here, $\overline{u}$ and $\overline{v}^*$ are the lengths of the supports, which may be infinite, and μ is the unknown lower bound of the support of u. It lies in some known closed interval $\mathcal{I} \subset \mathbb{R}$ with a lower bound greater than 0. Without loss

[21] Formally deriving the asymptotic distribution of the estimator is beyond the scope of this paper. The major difficulty is that the likelihood is non-regular, because the support of the bid densities depends on the parameters. Therefore, the results from Ackerberg, Chen, and Hahn (2012) do not 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 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.

[22] An alternative frequentist estimator would be the simulated method of moments estimator proposed by Bierens and Song (2011), which has been extended to the case with unobserved auction heterogeneity by Grundl and Zhu (2015). We found that the standard errors for estimates of the CRRA coefficient with the sieve MLE are about 60% smaller than with the simulated method of moments with an exponential or uniform weight function (results available upon request). Consequently, the test of risk neutrality has more power. The precision with the simulated method of moments could be improved by estimating the optimal weight function. We were not able to obtain satisfactory risk-aversion estimates with two-step estimators where the effect of unobserved auction heterogeneity is separated out in a first step as in Krasnokutskaya (2011). For Bayesian estimation approaches for first-price auctions, see Kim (2015a), Kim (2015b) and Aryal, Grundl, Kim, and Zhu (2015).

[23] Alternatively, we could assume that $\overline{u}$ and $\overline{v}^*$ are unknown but finite and treat them as parameters. In practice, it is difficult to distinguish a large upper bound with thin density from a small lower bound.

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 auction 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 $f^*(\cdot) = f^*(\cdot;\theta)$ and $f_n^u(\cdot) = f_n^u(\cdot - \mu;\theta)$. This transformation allows us to work with functions supported on $[0,1]$.[24]

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 a lower bound greater than 0. $\mathcal{A} = \Psi(B)^{n+1}$ where

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

and 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 the inconsistency problem due to an ill-posed inverse problem.[25]

Define $\alpha = \left(\psi^*, \{\psi_n^u\}_{n\in\mathbf{N}}\right)$, so $\theta = (\sigma, \gamma, \mu, \alpha)$.With some abuse of notation, let $\|\psi\|_\infty =$

[24] This transformation follows Bierens and Song (2012).
[25] This regularization follows Santos (2012).

$\sup_{x\in[0,1]} |\psi(x)|$ and

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

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 computing the joint bid densities. These potentially 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 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

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

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

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

The sieve maximum likelihood estimator is defined as

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

$\Theta_{k_L}=\Sigma\times\mathcal{K}^p\times\mathcal{I}\times\mathcal{A}_{k_L}$ 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). $\lim_{v\downarrow 1}h^*\left(v\right)/\left(v-1\right)^\epsilon=C$ as $v\downarrow 1$ for some $\epsilon\geq 0$ and $C>0$.

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

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

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

Assumption 4 includes requirements for the choice of h^* and h^u. Many commonly used density functions satisfy these requirements. Assumption 4 and the definition of Θ imply that the densities of the primitives are their corresponding base densities multiplied by functions bounded from above and bounded away from 0.[26] Assumption 5(1) requires that the sieve space is closed and increasing so that the maximization problem in equation (2) is well

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

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 that 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 than $-\infty$ to be dense in the parameter space. One can show that in the case considered here, the set of θ such that $E_0 l\left(Z_\ell, \theta\right) = -\infty$ 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 $l\left(Z; \theta\right)$ is upper semi-continuous in θ, Z-a.e., and that there exists $\theta_{0,k_L} \in \Theta_{k_L}$ such that $\left\|\theta_{0,k_L} - \theta_0\right\|_s \to 0$ and $E_0 l\left(Z_\ell, \theta_{0,k_L}\right) \to E_0 l\left(Z_\ell, \theta_0\right)$. Lemmas that establish these regularity conditions are collected in the Appendix.

4 Monte Carlo Experiments

4.1 Setup

Each generated sample has 900 auctions, and the number of bidders n ranges from 2 to 5.[27] We consider three different data-generating processes (DGPs). In all DGPs, $v = v^* u X^{\gamma_0}$, with $\log X \overset{iid}{\sim} N\left(0, 1\right)$ and $\gamma_0 = 0.9$. The unobserved auction heterogeneity is drawn from an 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 $n = 2$

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

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 $\sigma_0 = 0,\ 0.1,\ 0.2$ and 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 1,000 times.

4.2 Estimators

Results for two estimators are reported. The first estimator is 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 fourth-order Legendre polynomials. We compute bidding strategies at 3,000 points and interpolate linearly.[28]

As a benchmark, we also obtain estimates without taking unobserved heterogeneity into account, following the method used in Bajari and Hortacsu (2005) (BH estimator).[29] This estimator is computationally light and therefore a natural choice for a specification without unobserved auction heterogeneity. It is a two-step estimator. First, we estimate the following equation by ordinary least squares regression (OLS):

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

Let $\hat{\gamma}$ be the OLS estimate. We then 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:

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

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 that

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

[29]Bajari and Hortacsu (2005) used this estimator for experimental data without unobserved auction heterogeneity.

$n_\ell = n$ 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.[30] We restrict the estimates to be between 0 and 1. We report results for $n_1 = 2$ and $n_2 = 4$.[31]

4.3 Results

The discussion in this section focuses on the results for the CRRA coefficient shown in Table 1. Appendix D.2 contains the results for the value distribution and the distribution of the unobserved auction heterogeneity.

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.

Now consider the result if we ignore unobserved auction heterogeneity, using the two-step BH estimator shown in the lower half of Table 1. Ignoring unobserved auction heterogeneity can lead to a significant bias in risk-aversion estimates. Interestingly, the sign of the bias depends on the DGP. The CRRA coefficient is significantly overestimated under DGPs 1 (no selection) and 2 (weak selection), but it is underestimated under DGP 3 (strong selection). Section 4.4 provides some intuition to understand why the sign of the bias depends on the correlation between the number of bidders and the unobserved auction heterogeneity.

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.[32] Notice that under the null hypothesis, σ_0 is on the boundary of the

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

[31] 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.

[32] To estimate the asymptotic distribution, we need the joint bid densities to vanish smoothly at 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 10% significance level increases from about 30% if $\sigma_0 = 0.1$ to about 70% if $\sigma_0 = 0.2$ and about 92% if $\sigma_0 = 0.3$. 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.

		$\sigma_0 = 0$	$\sigma_0 = 0.1$	$\sigma_0 = 0.2$	$\sigma_0 = 0.3$
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: 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 auction heterogeneity is ignored using the two-step estimator proposed by Bajari and Hortacsu (2005).

boundary of their supports. Hence, we require that at least the f_n^u vanish smoothly at their boundaries.

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

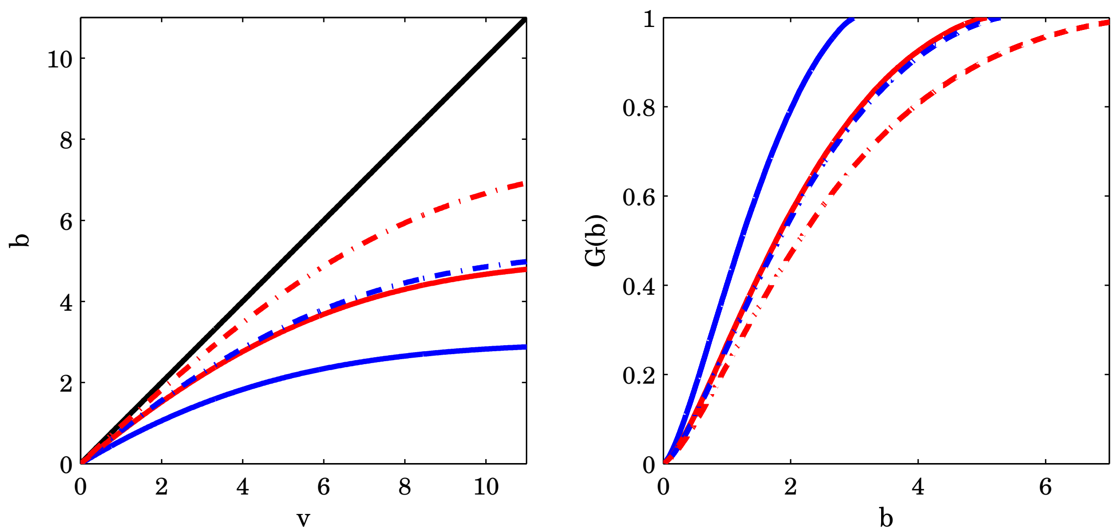
Table 2: This table shows the probability (in %) that risk-neutrality ($\sigma_0 = 0$) is rejected if the unobserved auction heterogeneity is taken into account (Sieve MLE).

4.4 Understanding the Bias If Unobserved 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 and the number of opponents. 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 not very dispersed, because then the bids are close to values even for a small number of competitors. Hence, the bid distribution tends 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. Risk aversion affects how much the bid distribution responds to changes in n and the dispersion of bids. Risk-averse



(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.

bidders respond less to changes in n because the bids are close to values even for a small number of competitors. 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, the econometrician concludes that the bidders are risk averse if the bid distribution does not respond much to increases in n relative to the dispersion of the bids.

If unobserved auction heterogeneity is ignored, the (unconditional) bid distributions appear very dispersed, as variation in bids due to unobserved auction heterogeneity is attributed to bidders' private information. In addition, if auctions with higher unobserved auction heterogeneity 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 overestimation of risk aversion. The second effect increases the shift of the bid distribution as n increases and therefore leads to underestimation 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 auction heterogeneity.

5 Empirical Application

5.1 Data Description

We estimate the risk aversion of bidders in USFS timber auctions.[33] The data can be downloaded from Phil Haile's website.[34] Lu and Perrigne (2008) and Campo, Guerre, Perrigne, and Vuong (2011) found the bidders to be risk averse.[35] Other work documented unobserved auction 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 subsample of scaled sales with contract lengths of less than one year between 1982 and 1990, for which the assumption of private values is plausible.[36] 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 auction heterogeneity, we further restrict the sample to auctions with two to five bidders. Intuitively, auctions with few competitors contain the most information about risk preferences. As the number of competitors increases, the effect of risk aversion on bids becomes small because

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

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

[35] 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).

[36] 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 do not include sales after 1990 for this region.

competition drives bids close to the values even for risk-neutral bidders. To reduce the influence of the extreme bids, we also discard eight auctions with bids more than eight times the appraisal value.[37] The final sample includes 370 two-bidder, 263 three-bidder, 172 four-bidder, and 105 five-bidder auctions. Our estimates condition on the appraisal value provided by the US Forest Service, which is designed to summarize all relevant information about the timber tract.

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 unobserved auction heterogeneity 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 quantile ranges. The bandwidth for the bid density estimators are chosen to be $std(b)L_n^{-1/4}$.[38] 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 unobserved auction heterogeneity, but reject risk neutrality in a specification without unobserved auction heterogeneity. This pattern is consistent with a low correlation between the unobserved auction heterogeneity and the number of bidders, as explained in Section 4.4. Indeed, we find that the distribution of the unobserved auction heterogeneity for different numbers of bidders is fairly similar. A possible explanation is that the unobserved auction heterogeneity is observed by the bidders only after they decided to enter the auction. For example, some characteristics are only observable to entrants who typically cruise the auctioned tract, but not to potential

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

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

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.[39] Intuitively, a violation of this assumption would bias our risk-aversion estimates upward. To see this, consider the case where the number of potential bidders is the same for all auctions.[40] In this case, the bid distribution would not vary with the number of entrants. Through the lens 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.

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 consistency under low-level conditions. We explain why the bias in risk-aversion estimates, if unobserved auction het-

[39] Athey, Levin, and Seira (2011) argue 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.

[40] Alternatively, we could condition on some measure for the number of potential bidders.

erogeneity 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 that the bidders are risk neutral, but we would reject risk neutrality if unobserved heterogeneity is ignored.

We see several 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 of disentangling which part of this correlation can be attributed to the unobserved auction heterogeneity and which part to the correlation of values conditional on the unobserved auction heterogeneity. A second avenue would be 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 measurement-error literature are still satisfied. Lastly, it would be useful to develop an estimator for the case with non-separable unobserved auction heterogeneity. Maximum likelihood estimation is challenging in this case, because we can no longer exploit the separability to reduce the computational burden when the likelihood function is evaluated.

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. *Technical Assumptions for Theorem 2.*

(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 condition

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

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

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

The first-order condition is satisfied by $s_n(v,u) = s_n^*(v/u)\,u$ only if the bidders have CRRA utility because otherwise $\lambda(\cdot/u)\,u \neq \lambda(\cdot)$. □

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 $v^* > C$.[41]

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 large enough x. 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

[41] This follows from a standard contradiction argument.

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 that (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 toward contradiction that for some $\alpha > 0$, we have $b_n^1(\alpha) \leq b_n^2(\alpha)$. By the continuity of the bid functions, there exists $\alpha_1 = \min\{\alpha : b_n^1(\alpha) = b_n^2(\alpha)\} > 0$. Notice that by construction, $b_n^1(\alpha) > b_n^2(\alpha)$ for $\alpha < \alpha_1$. 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 toward 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. □

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 toward 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. □

Proof of Proposition 1. Suppose toward 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 $\alpha \in [\underline{\alpha}, 1]$.[42] Take the difference of the first-order conditions for b_n^1 and b_n^2 and apply the mean value theorem

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

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

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

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

□

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) show 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 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 notation, we suppress u^* from now on. Suppose toward 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 $G_1\left(\widetilde{b}\right) = G_2\left(\widetilde{b}\right) = \alpha < 1$, so 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 $v_i(\alpha)$ is the α-th quantile of the valuation distribution. As $n_2 > n_1$ and $v_2(\alpha) \geq v_1(\alpha)$, 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), we construct a decreasing sequence of αs such that $R_1(\alpha_t) = R_2(\alpha_{t-1})$, with $R_1(\alpha_0) = x$. As $R_1(\alpha_{t-1}) > R_2(\alpha_{t-1})$, $R_1(0) = 0 < R_2(\alpha_{t-1})$, and R_1 is continuous, there exists an $\alpha_t \in (0, \alpha_{t-1})$ such that $R_1(\alpha_t) = R_2(\alpha_{t-1})$ 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,

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

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

We want to bound $\lambda^{-1}(x)$. We define $\widetilde{\lambda}$ as the strictly increasing function satisfying $\widetilde{\lambda}^{-1}(R_1(\alpha)) - \widetilde{\lambda}^{-1}(R_2(\alpha)) = b_2(\alpha) - b_1(\alpha)$ for $\alpha \in [0,1]$, with $\widetilde{\lambda}(0) = 0$. 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 substitution, $\widetilde{\lambda}$ can be expressed as follows:

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

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

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

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 $i = 1, 2$. $\square$

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.[43] 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. 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,

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

□

Notice that Lemma 4 holds even if potential bidders do not observe all components of X at the entry stage. In this case ϕ_i depends only on a subvector of X. Conditional on this subvector, i.e. from the perspective of the potential bidders, the signals are therefore correlated. Such situations are not considered in the Affiliated Signal Entry Model (Ye (2007), Gentry and Li (2014)).

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 among entrants conditional on* n *if one of the following conditions holds:*

[43]For example, in the Affiliated Signal Entry model studied in Li, Lu, and Zhao (2015), bidders observe X in the entry stage. Entry strategy can be described by $\phi_i(X, \xi_i) = 1(\xi_i > c(X))$ where $c(X)$ is a cutoff that makes potential bidders indifferent between entering and not entering.

(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.

Proof. We must allow for the possibility of mixed entry strategies. Bidder i uses a randomization device that 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 $\phi(\epsilon_i)$ that takes the value 0 or 1. Therefore, the joint distribution of (u, v^*) for entrants is

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

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

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

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

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

□

This result does not cover the case where bidders obtain some information about the private information v^* and about the unobserved auction heterogeneity 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 u is separable as in Theorem 1. Potential bidders observe only u, but no signal for v^*. Let $n < n'$. Then, $F^u(u|n') \leq F^u(u|n)$ for all $u \in [\underline{u}, \overline{u}]$ and $F^u(u|n') < F^u(u|n)$ for all u such that the entry probability is not 0 or 1.*

Proof of Proposition 4. We begin with the multiplicative case with CRRA utility. At the bidding stage, bidder i solves the following problem:

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

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

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

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

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

1. $p(u) = 0$ if $u \leq \left[\frac{\mathrm{U}(W) - \mathrm{U}(W-k)}{\Pi_1(1)}\right]^{\frac{1}{1-\sigma_0}} = u_1$.

2. $p(u) = 1$ if $u \geq \left[\frac{\mathrm{U}(W)-\mathrm{U}(W-k)}{\Pi_N(1)}\right]^{\frac{1}{1-\sigma_0}} = u_2$.

3. For $u \in (u_1, u_2)$, $p(u)$ solves

$$\sum_{n=1}^{N} \begin{pmatrix} N-1 \\ n-1 \end{pmatrix} [1-p(u)]^{N-n} p(u)^{n-1} \Pi_n(1) = \frac{\mathrm{U}(W) - \mathrm{U}(W-k)}{u^{1-\sigma}}. \tag{8}$$

Next, we show that p is strictly increasing in the third case by differentiating (8):[44]

$$p'(u) = -(1-\sigma) \frac{\mathrm{U}(W) - \mathrm{U}(W-k)}{u^{2-\sigma}(A_1 + A_2)},$$

where

$$\begin{aligned} A_1 &= -\sum_{n=1}^{N-1} \begin{pmatrix} N-1 \\ n-1 \end{pmatrix} (N-n)[1-p(u)]^{N-n-1} p(u)^{n-1} \Pi_n(1) \\ A_2 &= \sum_{n=2}^{N} \begin{pmatrix} N-1 \\ n-1 \end{pmatrix} (n-1)[1-p(u)]^{N-n} p(u)^{n-2} \Pi_n(1). \end{aligned}$$

Notice that

$$\begin{aligned} (1-p(u))\,p(u)\,A_1 &= -\sum_{n=1}^{N-1} \begin{pmatrix} N-1 \\ n \end{pmatrix} n[1-p(u)]^{N-n} p(u)^{n} \Pi_n(1) \\ (1-p(u))\,p(u)\,A_2 &= \sum_{n=1}^{N-1} \begin{pmatrix} N-1 \\ n \end{pmatrix} n[1-p(u)]^{N-n} p(u)^{n} \Pi_{n+1}(1) \end{aligned}$$

Therefore,

$$A_1 + A_2 = \frac{1}{(1-p(u))\,p(u)} \sum_{n=1}^{N-1} \begin{pmatrix} N-1 \\ n \end{pmatrix} n[1-p(u)]^{N-n} p(u)^{n} (\Pi_{n+1}(1) - \Pi_n(1)).$$

[44] This part of the proof is similar to the proof of Lemma 1 in Levin and Smith (1994b).

Now, we show that $\Pi_n(1)$ is decreasing in n. To see this, recall that

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

Here, $v^* - s_n^*(v^*)$ and $F^*(v^*)^{n-1}$ are both decreasing in n. Therefore, the integrand and $\Pi_n(1)$ are decreasing in n. Consequently, $A_1 + A_2 < 0$ and $p'(u) > 0$.

The distribution of u conditional on n is:

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

with $C(n) = \int_{\underline{u}}^{\bar{u}} p(x)^n [1 - p(x)]^{N-n} f^u(x)\, dx$. Notice that the support of $f^u(u|n)$ is $[u_1, u_2]$ for $n < N$ and $[u_1, \overline{u}]$ for $n = N$.

Now consider

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

As $n' > n$ and p is strictly increasing on (u_1, u_2), the density ratio $f^u(u|n)/f^u\left(u|n'\right)$ is strictly decreasing in u on (u_1, u_2). Moreover, as u goes to u_1 from above, $p(u)$ goes to zero and the ratio of densities goes to infinity. Hence, $f^u(u_1 + \epsilon|n) > f^u\left(u_1 + \epsilon|n'\right)$ for sufficiently small $\epsilon > 0$. We know that $F^u(u_1|n) = 0$ for all n, so this implies $F^u(u_1 + \epsilon|n) > F^u\left(u_1 + \epsilon|n'\right)$ for sufficiently small $\epsilon > 0$.

Now, suppose toward contradiction that $F^u\left(u|n'\right) \geq F^u(u|n)$ for some $u \in (u_1, u_2)$ and denote the smallest one of these points by $\tilde{u}$. As $F^u(u_1 + \epsilon|n) > F^u\left(u_1 + \epsilon|n'\right)$ for sufficiently small $\epsilon > 0$, $f^u(\tilde{u}|n) \leq f^u(\tilde{u}|n)$. As $f^u(u|n)/f^u\left(u|n'\right)$ is strictly decreasing, $f^u(u|n) < f^u(u|n)$ for $u \in (\tilde{u}, u_2)$. Therefore, this would imply that $F^u(u_2|n) < F^u\left(u_2|n'\right)$. This is a contradiction because we know that $F^u(u_2|n) = 1$ for $n < N$ and $F^u(u_2|N) \leq 1$. Hence, $F^u\left(u|n'\right) < F^u(u|n)$ for all $u \in (u_1, u_2)$.

To summarize, $F^u(u|n) = 0$ for all n and $u \leq u_1$, $F^u\left(u|n'\right) < F^u(u|n)$ for all $u \in (u_1, u_2)$

and $n^{'} > n$, and $F^u(u_2|n) = 1$ for $n < N$ and $F^u(u_2|N) \leq 1$.

Next consider the additive case. If $n > 1$, the bidding strategy for an entrant with v^* in an n-bidder auction with u is $s_n(v^*, u) = s_n^*(v^*) + u$, where $s_n^*(v^*) = s_n(v^*, 0)$. If $n = 1$, $s_n(v^*, u) = 0$. Therefore,

$$\Pi_n(u) \quad = \quad E_{v^*}\left[\mathrm{U}\left(v^* + u - u - s_n^*(v^*) + W - k\right) - \mathrm{U}(W - k)\right] F^*(v^*)^{n-1} = \Pi_n(0)$$

for all $n > 1$ while $\Pi_1(u) = E_{v^*}\left[\mathrm{U}(v^* + u + W - k) - \mathrm{U}(W - k)\right]$. Notice that $\Pi_1(u)$ is strictly increasing in u.

All potential bidders enter for all u if

$$\Pi_N(0) + \mathrm{U}(W - k) > \mathrm{U}(W).$$

In this case there is no variation in n, so we assume henceforth that this condition does not hold. Then no bidder enters for

$$u < \Pi_1^{-1}\left[\mathrm{U}(W) - \mathrm{U}(W - k)\right] = u_1.$$

For $u > u_1$ the entry probability solves

$$(1 - p(u))^{N-1} \Pi_1(u) + \sum_{n=2}^{N} \begin{pmatrix} N-1 \\ n-1 \end{pmatrix} (1 - p(u))^{N-n} p(u)^{n-1} \Pi_n(0) + \mathrm{U}(W - k) - \mathrm{U}(W) = 0.$$

Applying a similar argument as for the multiplicative case we obtain

$$p'(u) = -\frac{p(u)(1 - p(u))^N \Pi_1'(u)}{\sum_{n=1}^{N-1} \begin{pmatrix} N-1 \\ n \end{pmatrix} n\left[1 - p(u)\right]^{N-n} p(u)^n \left(\Pi_{n+1}(u) - \Pi_n(u)\right)} > 0.$$

Here we used the fact that $\Pi_n(u) = \Pi_n(0)$ for $n > 1$ and that $\Pi_n(u)$ is decreasing in n.

Notice that while p is strictly increasing in u it never reaches 1.

Following the same argument as in the multiplicative case, we can now show that $F^u(u|n') < F^u(u|n)$ for $u > u_1$ and $n' > n$. $\square$

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 8, g_n^* is bounded. Then $E_0 l(Z_\ell;\theta)$ is upper semi-continuous by Lemma 10 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 10, $l(Z_\ell,\theta)$ is continuous in θ Z_ℓ-a.s. Hence, if $\mathcal{N}(\theta) \downarrow \theta$, $l_{\mathcal{N}(\theta)}(Z_\ell) \downarrow l(Z_\ell,\theta)$ almost surely. 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 sets $\mathcal{N}_j = \mathcal{N}(\theta_j)$ indexed by $j = 1, 2, \cdots, J$ that cover $\Theta(\epsilon)$. By Lemma 5 and Lemma 8, $g_n(\mathbf{b},\theta)$ is bounded from above, as is $l(Z_\ell,\theta)$ and $l_{\mathcal{N}_j}(Z_\ell)$. Therefore, we can still apply the law of large numbers even if the expectation may be $-\infty$:

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

There exists sequence $\{\theta_{0,k_L}\}_{k_L=1}^{\infty}$, $\theta_{0,k_L} \in \Theta_{k_L}$ such that $\lim_{k_L\to\infty} \|\theta_{0,k_L} - \theta_0\|_s = 0$, and $E_0 l(Z_\ell, \theta_{0,k_L}) - E_0 l(Z_\ell, \theta_0) \to 0$ by Lemma 11. Therefore, we can find K large enough such that $|E_0 l(Z_\ell, \theta_{0,k_L}) - E_0 l(Z_\ell, \theta_0)| < \delta/4$ and $\theta_K \in \Theta(\epsilon)^c \cap \Theta_{k_L}$ for all $k_L \geq K$. By this definition and (9),

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

For large enough L,

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

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

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

The last step follows from (10) and the law of large numbers:

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

□

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. 4, 14

Andrews, D. W. (1999): "Estimation When a Parameter Is on a Boundary," *Econometrica*, 67(6), 1341–1383. 21

Aradillas-López, A., A. Gandhi, and D. Quint (2013a): "Identification and Inference in Ascending Auctions with Correlated Private Values," *Econometrica*, 81(2), 489–534. 5, 24

——— (2013b): "A Simple Test for Moment Inequality Models with an Application to English Auctions," . 5, 24

Aryal, G., S. Grundl, D.-H. Kim, and Y. Zhu (2015): "Empirical Relevance of Ambiguity in First Price Auctions," Discussion paper, Working Paper. 14

Athey, S., and J. Levin (2001): "Information and Competition in US Forest Service Timber Auctions," *Journal of Political Economy*, 109(2). 6, 24

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, 5, 24, 26

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. 6, 19, 21, 25

Baldwin, L. (1995): "Risk Averse Bidders at Oral Ascending-Bid Forest Service Timber Sales," *Unpublished Manuscript, RAND Corporation.[1193]*. 6, 24

Baldwin, L., R. Marshall, and J. Richard (1997): "Bidder Collusion at Forest Service Timber Sales," *The Journal of Political Economy*, 105(4), 657–699. 24

Bierens, H., and H. Song (2011): "Nonparametric Identification of the First-Price Auction Model," Discussion paper. 14

Bierens, H. J. (2014): "Consistency and Asymptotic Normality of Sieve Ml Estimators under Low-Level Conditions," *Econometric Theory*, 30, 1021–1076. 18

Bierens, H. J., and H. Song (2012): "Semi-Nonparametric Estimation of Independently and Identically Repeated First-Price Auctions via an Integrated Simulated Moments Method," *Journal of Econometrics*, 168(1), 108–119. 15

Campo, S. (2012): "Risk Aversion and Asymmetry in Procurement Auctions: Identification, Estimation and Application to Construction Procurements," *Journal of Econometrics*, 168(1), 96–107. 6

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. 6, 24, 34

Chernozhukov, V., and H. Hong (2004): "Likelihood Estimation and Inference in a Class of Nonregular Econometric Models," *Econometrica*, 72(5), 1445–1480. 14

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. 6

d'Haultfoeuille, X., and P. Février (2010a): "Identification of a Class of Adverse Selection Models with Contracts Variation," Discussion paper, Mimeo. 11

——— (2010b): "Identification of Mixture Models Using Support Variations," *Manuscript, ENSAE*. 3, 34

Donald, S. G., and H. J. Paarsch (1993): "Piecewise Pseudo-Maximum Likelihood Estimation in Empirical Models of Auctions," *International Economic Review*, pp. 121–148. 14

Evdokimov, K., and H. White (2012): "Some Extensions of a Lemma of Kotlarski," *Econometric Theory*, 28(04), 925–932. 9, 29, 30, 31

Gentry, M., and T. Li (2014): "Identification in Auctions with Selective Entry," *Econometrica*, 82(1), 315–344. 8, 37

Gentry, M., T. Li, and J. Lu (2015): "Identification and Estimation in First-Price Auctions with Risk-Averse Bidders and Selective Entry," *Available at SSRN*. 6, 8, 26, 27

Grundl, S., and Y. Zhu (2015): "Identification of First-Price Auctions With Biased Beliefs," *Finance and Economics Discussion Series 2015-056. Washington: Board of Governors of the Federal Reserve System*. 14

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. 4, 6, 7, 8, 26, 29, 31, 33, 34, 35

Haile, P. (2001): "Auctions with Resale Markets: An Application to US Forest Service Timber Sales," *The American Economic Review*, 91(3), 399–427. 24

Haile, P., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *Journal of Political Economy*, pp. 1–51. 24

Hirano, K., and J. R. Porter (2003): "Asymptotic Efficiency in Parametric Structural Models with Parameter-Dependent Support," *Econometrica*, 71(5), pp. 1307–1338. 14

Holt, C. A. (1980): "Competitive Bidding for Contracts under Alternative Auction Procedures," *The Journal of Political Economy*, pp. 433–445. 3

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. 3

Hu, Y., D. McAdams, and M. Shum (2013): "Identification of First-Price Auctions with Non-Separable Unobserved Heterogeneity," *Journal of Econometrics*, 174(2), 186–193. 3, 5, 8, 9, 10, 11, 33

Kagel, J. H. (1995): "Auctions: A Survey of Experimental Research," *The Handbook of Experimental Economics*. 6

Kagel, J. H., and D. Levin (2010): "Auctions: A Survey of Experimental Research, 1995-2010," *Handbook of Experimental Economics*, 2. 6

Kim, D.-H. (2015a): "Flexible Bayesian Analysis of First Price Auctions Using a Simulated Likelihood," *Quantitative Economics*, forthcoming. 14

——— (2015b): "Nonparametric Estimation of Utility Function in First-Price Sealed-Bid Auctions," *Economics Letters*, 126, 101–106. 6, 14

Kong, Y. (2015): "Endogenous Entry and Risk Aversion in Oil and Gas Lease Auctions," Discussion paper. 6

Kotlarski, I. (1967): "On Characterizing the Gamma and Normal Distribution," *Pacific Journal of Mathematics*, 20(1), 69–76. 9

Krasnokutskaya, E. (2011): "Identification and Estimation of Auction Models with Unobserved Heterogeneity," *The Review of Economic Studies*, 78(1), 293. 3, 4, 5, 8, 9, 13, 14, 28

——— (2012): "Identification and Estimation of Auction Model with Two-Dimensional Unobserved Heterogeneity," *International Economic Review*, 53(3), 659–692. 5

Levin, D., and J. L. Smith (1994a): "Equilibrium in Auctions with Entry," *The American Economic Review*, 84(3), pp. 585–599. 3, 8

——— (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. 37

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. 6, 24

Maskin, E., and J. Riley (1984): "Optimal Auctions with Risk Averse Buyers," *Econometrica: Journal of the Econometric Society*, pp. 1473–1518. 3

Matthews, S. (1987): "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View," *Econometrica: Journal of the Econometric Society*, pp. 633–646. 3

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

Roberts, J. W., and A. Sweeting (2010): "Bailouts and the Preservation of Competition," Discussion paper, National Bureau of Economic Research. 24

——— (2013): "When Should Sellers Use Auctions?," *American Economic Review*, 103(5), 1830–61. 5, 24

——— (forthcoming): "Bailouts and the Preservation of Competition: The Case of the Federal Timber Contract Payment Modification Act," *American Economic Journal: Microeconomics*. 5

Samuelson, W. F. (1985): "Competitive Bidding with Entry Costs," *Economics Letters*, 17(1-2), 53–57. 8

Santos, A. (2012): "Inference in Nonparametric Instrumental Variables with Partial Identification," *Econometrica*, 80(1), 213–275. 15

Smith, J. L., and D. Levin (1996): "Ranking Auctions with Risk Averse Bidders," *Journal of Economic Theory*, 68(2), 549–561. 3

Smith, R. L. (1985): "Maximum Likelihood Estimation in a Class of Nonregular Cases," *Biometrika*, 72(1), pp. 67–90. 14

van der Vaart, A., and J. Wellner (2000): *Weak Convergence and Empirical Processes: with Applications to Statistics (Springer Series in Statistics)*. Springer, corrected edn. 18

Ye, L. (2007): "Indicative Bidding and a Theory of Two-Stage Auctions," *Games and Economic Behavior*, 58(1), 181–207. 8, 37

Zincenko, F. (2014): "Nonparametric Estimation of First-Price Auctions with Risk-Averse Bidders," *Unpublished Manuscript, University of Pittsburgh*. 6

D Supplementary Materials

D.1 Lemmas for the Proof of Proposition 2

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

Lemma 5. *$f^*(\cdot;\theta)$ 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

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

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

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

Then we have

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

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,

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

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 $|F_n^u(x;\theta_1) - F_n^u(x;\theta_2)| \leq \int |f_n^u(x;\theta_1) - f_n^u(x;\theta_2)|\,dx \leq C\|\theta_1-\theta_2\|_\infty$. The same inequality holds for $F^*(\cdot;\theta)$. □

Lemma 6. $\sup_{v\in[1,\bar{v}^*+1]} |s_n^*(v;\theta_k) - s_n^*(v;\theta)| \to 0$ *if* $\|\theta_k-\theta\|_s \to 0$.

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

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

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

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

Because $x > \epsilon$, $F^*(x) > 2\delta$ for some $\delta > 0$. $F^*(v;\theta_k) \to F^*(v;\theta)$ uniformly by Lemma 5. Hence, there exists a K such that for all $k > K$, $F_k^*(x) > \delta$ for all $x > 1+\epsilon$. Therefore, for sufficiently large k,

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

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

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

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 that they converge to 0 uniformly in x. Therefore, $\left|s_n^*(x;\theta_k) - s_n^*(x;\theta)\right|$ 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 if $k > K$,

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

Because the above inequality holds for all $\delta > 0$ and 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 $\|\theta_k - \theta\|_s \to 0$ and 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 first-order condition, for any $\theta \in \Theta$,

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

for some constant C. The last inequality holds because of Lemma 5 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 large enough $v_1 > v_2$,

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

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 7. *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 6, 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 a $v\neq s_n^{*-1}(x;\theta)$ 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 that 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 large enough k by Lemma 6. Hence, $g_n(x;\theta_k)=g_n(x;\theta)=0$ for all large enough k. 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 8. 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 5, 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 9. *$g_n \left(\log \mathbf{b} / \exp \left(\log X\gamma \right);\theta \right)$ is continuous in θ on $\Theta \left(\mathbf{b}, X \right)$-a.e.*

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

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

The first inequality holds because μ has to be strictly greater than some positive number. Hence, $u \geq \mu$ 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.

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

The first inequality holds by Lemma 5, and the second inequality holds by the fact that g_n^* is bounded and is a density function.

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

First, use a change of variables to rewrite

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

$\left|\prod_{i=1}^{n} g_n^*\left(b_i/u;\theta_k\right) - \prod_{i=1}^{n} g_n^*\left(b_i/u;\theta\right)\right| \to 0$ u-a.e. by Lemma 7. In addition, by Lemma 8, for some $C<\infty$,

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

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

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

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

Lemma 10. *$l\left(Z,\theta\right)$ is continuous in θ on Θ Z-a.e.*

Proof. $l\left(Z,\theta\right)$ is the log of a continuous function g_n by Lemma 9. Therefore, it is continuous in θ, Z-a.e. □

Lemma 11. *Under Assumptions 4 and 5, there exists a sequence $\{\theta_{0,k_L}\}_{k_L=1}^{\infty}$ such that $\theta_{0,k_L}\in\Theta_{k_L}$, $\lim_{k_L\to\infty}\left\|\theta_{0,k_L}-\theta_0\right\|_s=0$, and $E_0 l\left(Z_\ell,\theta_{0,k_L}\right)-E_0 l\left(Z_\ell,\theta_0\right)\to 0$.*

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

Recall the first-order condition

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

Therefore, $s_n^*\left(\cdot;\theta_0\right)\leq s_n^*\left(\cdot;\theta_{0,k_L}\right)$ 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

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

Because $\left\|\psi^*_{0,k_L}-\psi^*_0\right\|_\infty \to 0$, by Lemma 5,

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

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

$$\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right](T\psi^*_0)\circ H^*(v)\,h^*(v) \leq \left(T\psi^*_{0,k_L}\right)\circ H^*(v)\,h^*(v) \leq \left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right](T\psi^*_0)\circ H^*(v)\,h^*(v),$$

which suggests

$$\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right] f^*(v;\theta_0) \leq f^*(v;\theta_{0,k_L}) \leq \left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right] f^*(v;\theta_0) \tag{15}$$

$$\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right] F^*(v;\theta_0) \leq F^*(v;\theta_{0,k_L}) \leq \left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right] F^*(v;\theta_0). \tag{16}$$

(15) and (16) together imply

$$\frac{1-\sigma_0}{n-1}\frac{F^*(v;\theta_0)}{f^*(v;\theta_0)} \geq \frac{1-\sigma_0}{n-1}\frac{\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right]F^*(v;\theta_{0,k_L})}{\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right]f^*(v;\theta_{0,k_L})} \geq \frac{1-\sigma_0}{n-1}\frac{F^*(v;\theta_0)}{f^*(v;\theta_0)}\frac{\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right]^2}{\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right]^2}.$$

Define $\sigma_{k_L} = 1-(1-\sigma_0)\left[1-\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right]/\left[1+\frac{\epsilon_{k_L}(1+B)}{\eta^2}\right] \downarrow \sigma_0$ 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 large enough k_L, then because $s^*_n(v;\theta_0) \leq s^*_n(v;\theta_{0,k_L})$,

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

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

It is left to show that $E_0 l(Z_\ell;\theta_{0,k_L}) - E_0 l(Z_\ell;\theta_0) \to 0$. First, notice $l(Z;\theta_{0,k_L}) \to l(Z;\theta_0)$ Z-a.e. by Lemma 10. Then, if $|l(Z;\theta_{0,k_L})|$ is bounded by an integrable function, the Dominated Convergence Theorem implies that $E_0 l(Z_\ell;\theta_{0,k_L}) - E_0 l(Z_\ell;\theta_0) \to 0$. First, notice that Lemma 8 implies that $g^*_n(\cdot;\theta) < C$. Therefore, $g^*_n(\mathbf{b};\theta) = \int \frac{1}{u^n}\prod_{i=1}^n g^*_n(b_i;\theta)\, f^u_n(u;\theta)\,du < C$ for some constant C . Because $l(Z;\theta_{0,k_L})$ 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 $b \in [1, s_n^*\left(1+\overline{v}^*;\theta_0\right)]$, $g_n^*\left(b;\theta_{0,k_L}\right) \geq C_2 g_n^*\left(b;\theta_0\right)$ 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 5 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 enough 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 $s_n^*(\cdot;\theta_{k_L}) \le C\left(s_n^*(\cdot;\theta_0)-1\right)+1$,

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

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 $C_2 > 0$ 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 large enough k_L. By Lemma 5, $f_n^u(u;\theta_{k_L}) > C_5 f_n^u(u;\theta_0)$

$$\begin{aligned}
g_n(\mathbf{b};\theta_{0,k_L}) &= \int \frac{1}{u^n}\prod_{i=1}^n g_n^*(b_i/u;\theta_{0,k_L}) f_n^u(u;\theta_{0,k_L})\,du \ge C_5 \int \frac{1}{u^n}\prod_{i=1}^n g_n^*(b_i/u;\theta_k) f_n^u(u;\theta_0)\,du \\
&\ge C_5 C_2 \int \frac{1}{u^n}\prod_{i=1}^n g_n^*(b_i/u;\theta_0) f_n^u(u;\theta_0)\,du = C_1 g_n(\mathbf{b};\theta_0).
\end{aligned}$$

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

D.2 Additional Monte Carlo Results

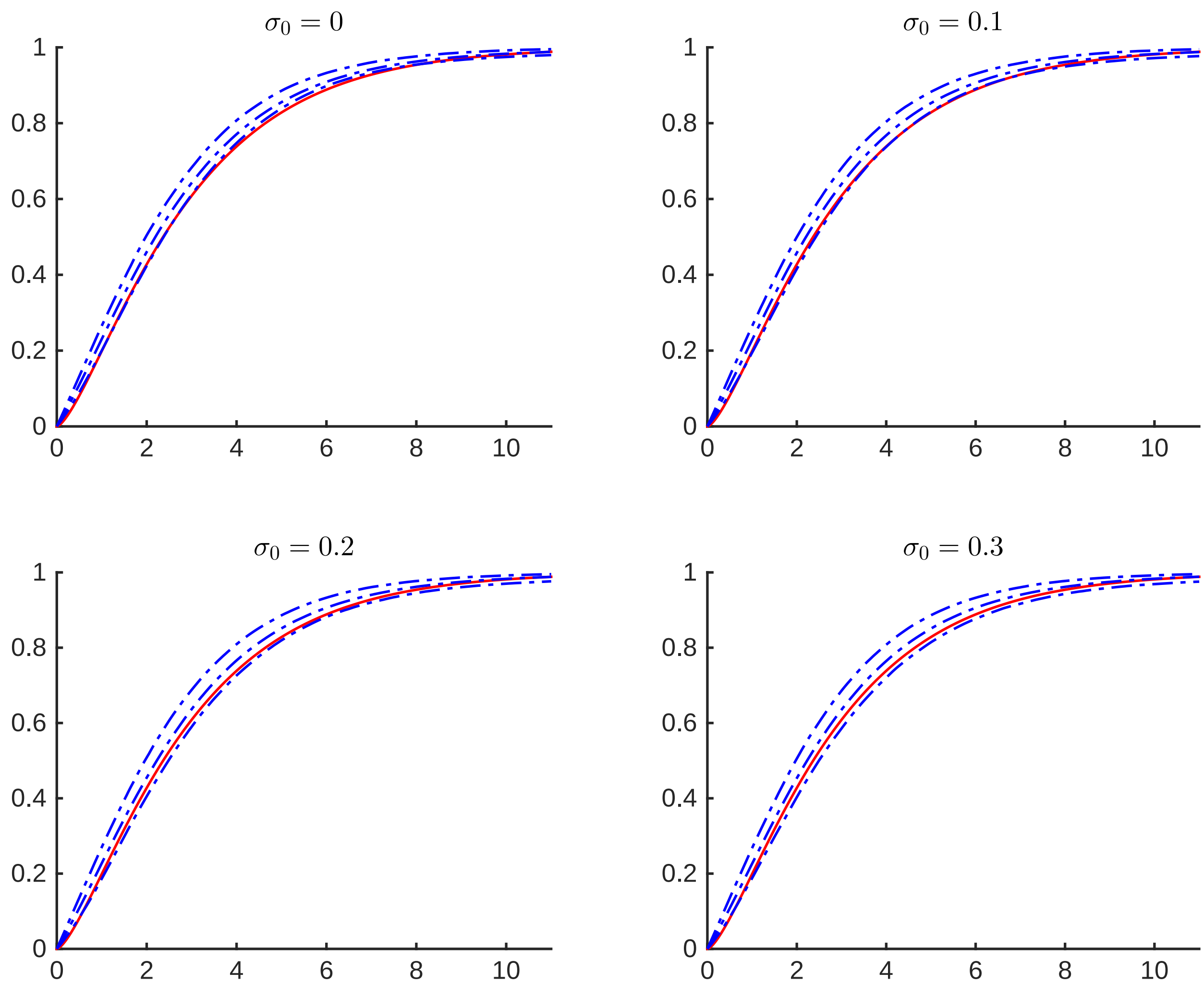
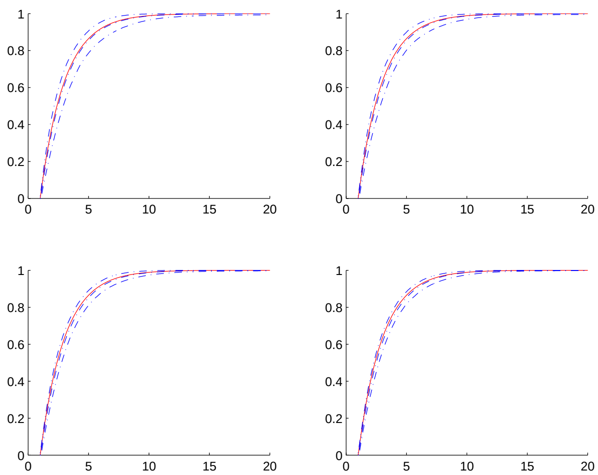
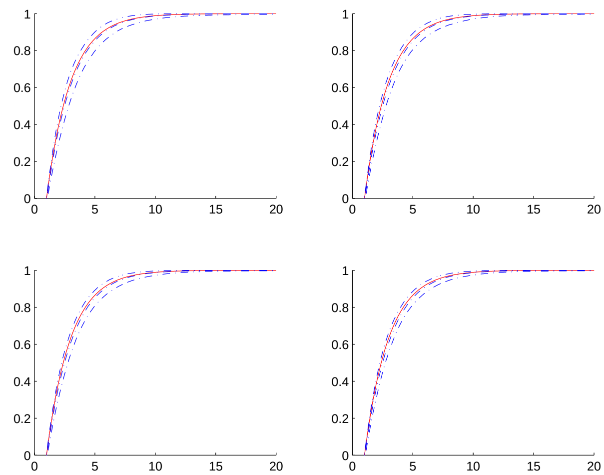


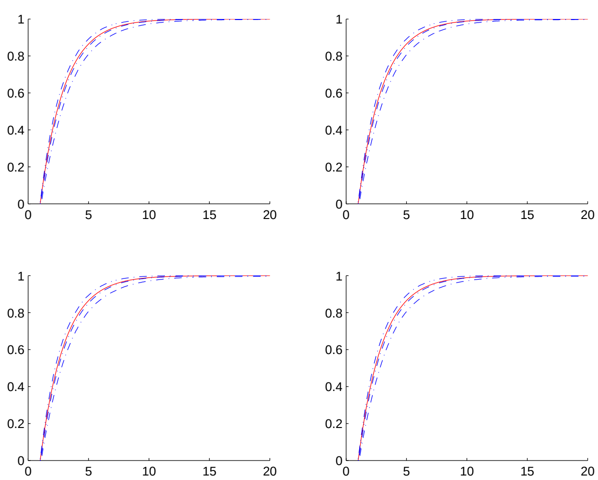
Figure II: Value Distribution. No Selection (DGP 1). Red solid line: True distribution. Blue dot-dash lines: 2.5%, 50%, 97.5% quantiles of the estimates.



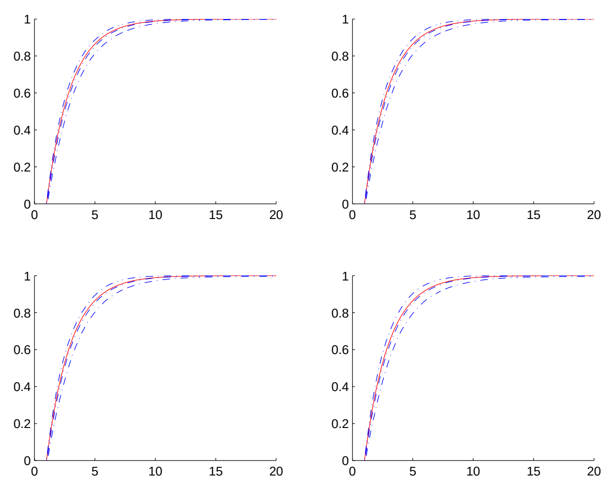
(a) $\sigma_0 = 0$



(b) $\sigma_0 = 0.1$



(c) $\sigma_0 = 0.2$



(d) $\sigma_0 = 0.3$

Figure III: Distributions of Unobserved Auction Heteorgeneity. No Selection (DGP 1). Red solid line: True distribution. Blue dot-dash lines: 2.5%, 50%, 97.5% quantiles of the estimator. In each sub-figure, the first row is for 2 and 3-bidder auctions, and the second row is for 4 and 5-bidder auctions.

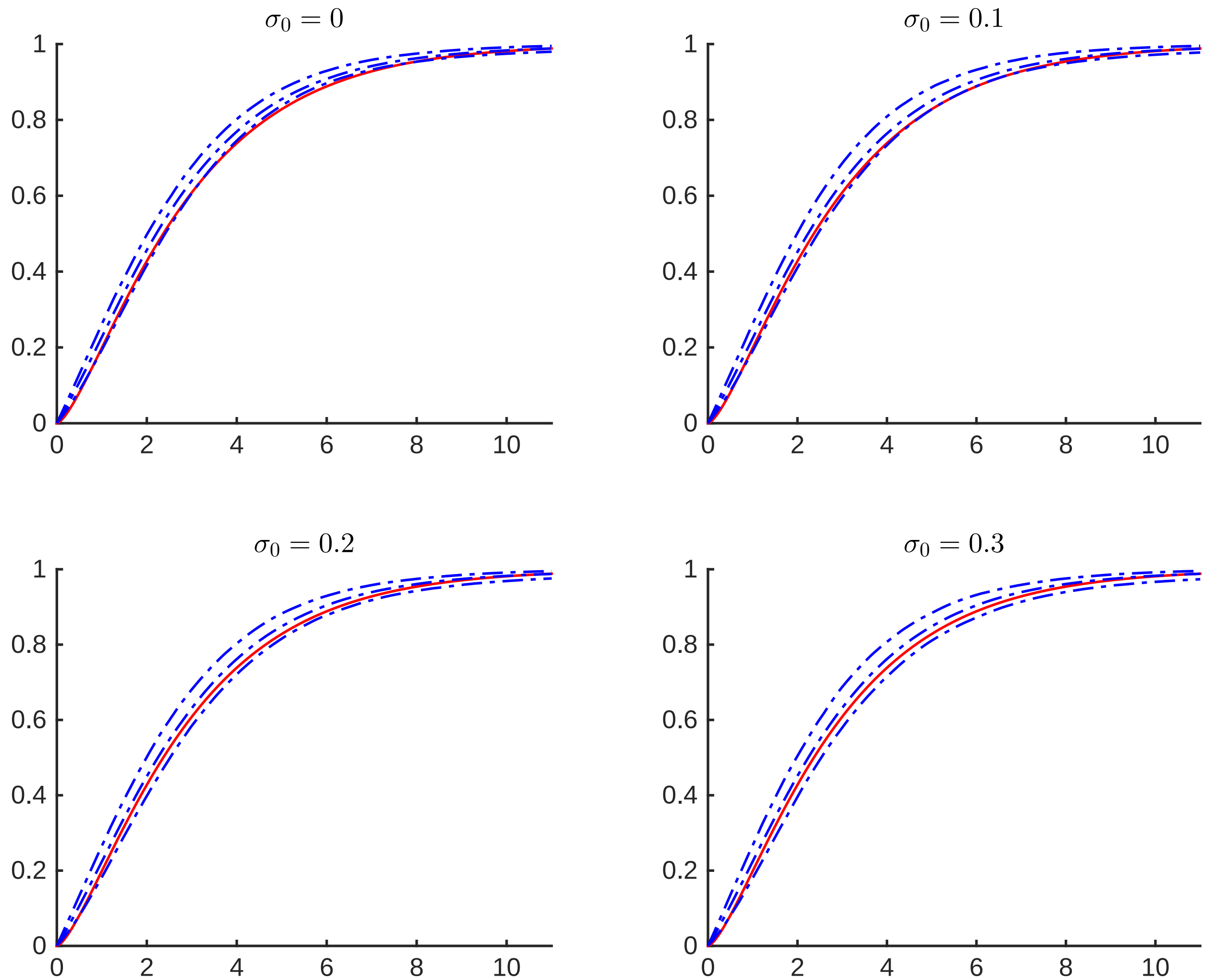
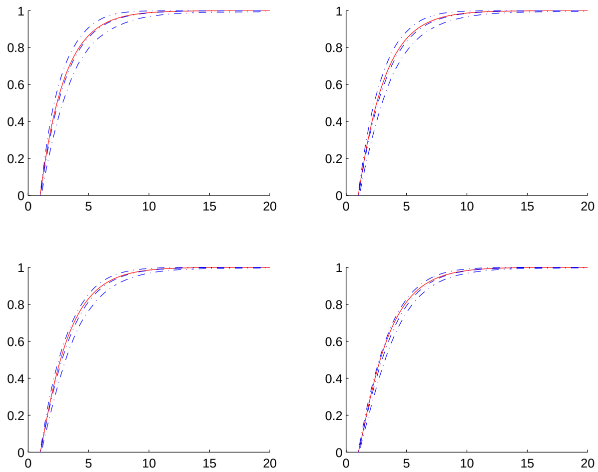
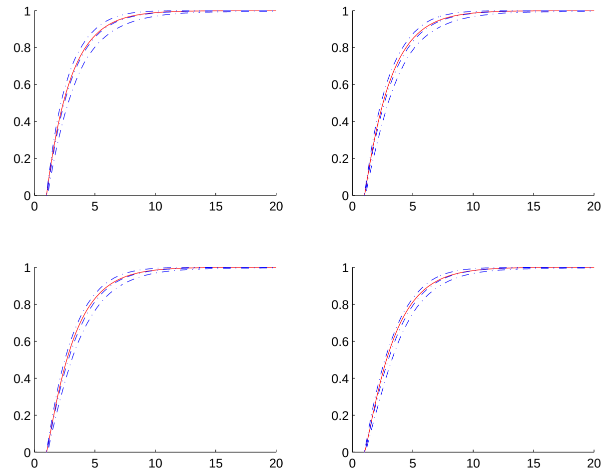


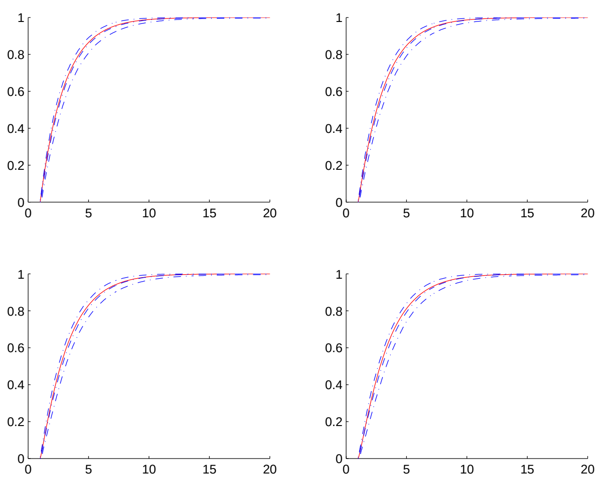
Figure IV: Value Distribution. Weak Selection (DGP 2). Red solid line: True distribution. Blue dot-dash lines: 2.5%, 50%, 97.5% quantiles of the estimator



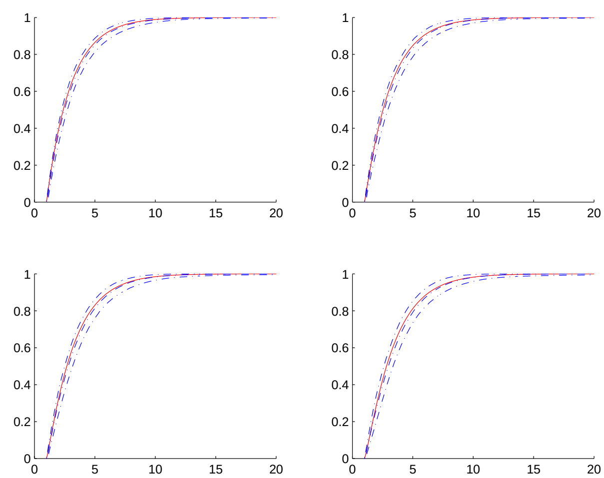
(a) $\sigma_0 = 0$



(b) $\sigma_0 = 0.1$



(c) $\sigma_0 = 0.2$



(d) $\sigma_0 = 0.3$

Figure V: Distributions of Unobserved Auction Heterogeneity: Weak Selection (DGP 2). Red solid line: True distribution. Blue dot-dash lines: 2.5%, 50%, 97.5% quantiles of the estimator. In each sub-figure, the first row is for 2 and 3-bidder auctions, and the second row is for 4 and 5-bidder auctions.

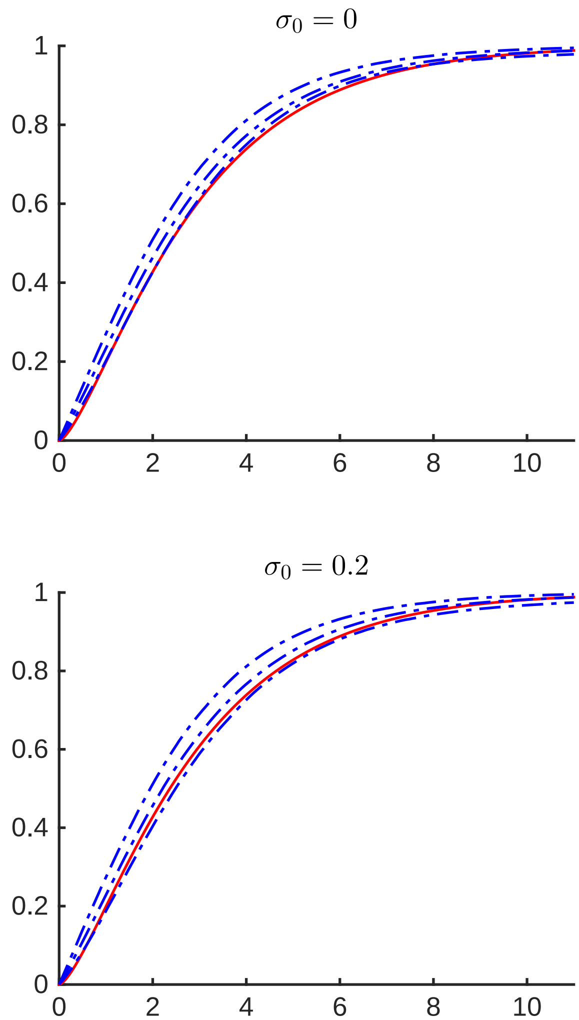
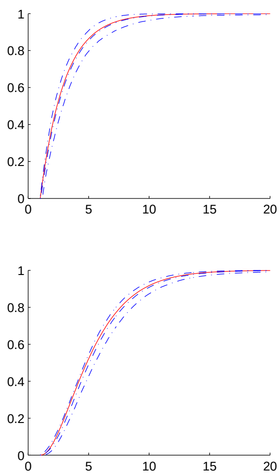
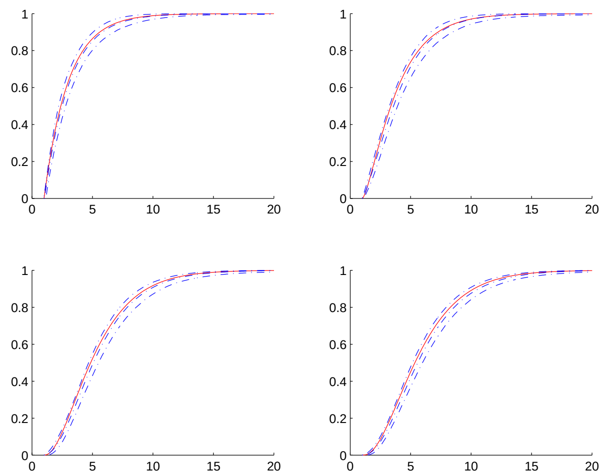


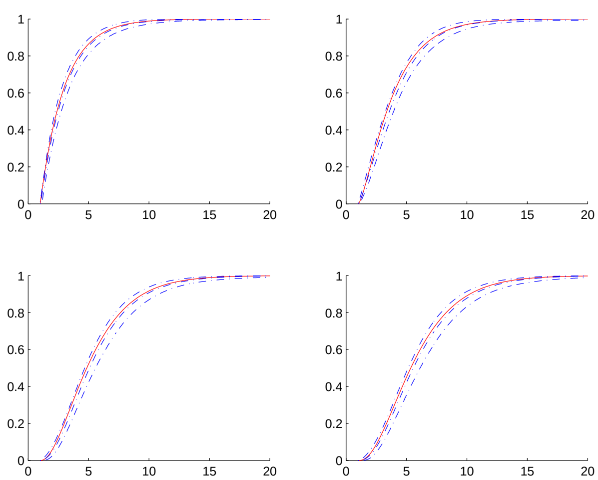
Figure VI: Value Distribution: Strong Selection. Red solid line: True distribution. Blue dot-dash lines: 2.5%, 50%, 97.5% quantiles of the estimator



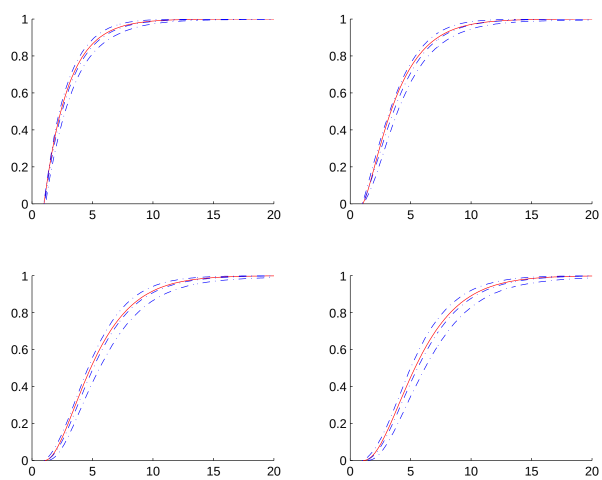
(a) $\sigma_0 = 0$



(b) $\sigma_0 = 0.1$



(c) $\sigma_0 = 0.2$



(d) $\sigma_0 = 0.3$

Figure VII: Distributions of Unobserved Auction Heterogeneity: Strong Selection. Red solid line: True distribution. Blue dot-dash lines: 2.5%, 50%, 97.5% quantiles of the estimator. In each sub-figure, the first row is for 2 and 3-bidder auctions, and the second row is for 4 and 5-bidder auctions.