cemmap
centre for microdata methods and practice

Identification and estimation in first-price auctions with risk-averse bidders and selective entry

Matthew Gentry
Tong Li
Jingfeng Lu

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP16/15

Identification and Estimation in First-Price Auctions with Risk-Averse Bidders and Selective Entry*

Matthew Gentry[†] Tong Li[‡] Jingfeng Lu[§]

March 2015

Abstract

We study identification and estimation in first-price auctions with risk averse bidders and selective entry, building on a flexible entry and bidding framework we call the Affiliated Signal with Risk Aversion (AS-RA) model. This framework extends the AS model of Gentry and Li (2014) to accommodate arbitrary bidder risk aversion, thereby nesting a variety of standard models as special cases. It poses, however, a unique methodological challenge – existing results on identification with risk aversion fail in the presence of selection, while the selection-robust bounds of Gentry and Li (2014) fail in the presence of risk aversion. Motivated by this problem, we translate excludable variation in potential competition into identified sets for AS-RA primitives under various classes of restrictions on the model. We show that a single parametric restriction – on the copula governing selection into entry – is typically sufficient to restore point identification of all primitives. In contrast, a parametric form for utility yields point identification of the utility function but only partial identification of remaining primitives. Finally, we outline a simple semiparametric estimator combining Constant Relative Risk Aversion utility with a parametric signal-value copula. Simulation evidence suggests that this estimator performs very well even in small samples, underscoring the practical value of our identification results.

Keywords: Auctions, endogenous participation, risk aversion, identification.

1 Introduction

Risk aversion and endogenous entry both play major roles in shaping real-world auction performance. As is well known, risk aversion influences answers to a wide range of

*We thank Li Zhao for for many helpful comments which have significantly improved the paper. We also thank seminar participants at the University of Chicago and University College London for valuable comments and discussion.

[†]Matthew Gentry: Department of Economics, London School of Economics,Houghton Street, London, UK WC2A 2AE. Email: m.l.gentry@lse.ac.edu.

[‡]Tong Li: Department of Economics, Vanderbilt University, Nashville, TN 37235-1819. Email: tong.li@vanderbilt.edu.

[§]Jingfeng Lu: Department of Economics, National University of Singapore, Singapore 117570. Email: ecsljf@nus.edu.sg.

fundamental questions in auction design, such as choice of auction format (Maskin and Riley (1984)), structure of the optimal mechanism (Matthews (1987)), and whether to disclose reserve prices (Li and Tan (2000)). This in turn has motivated a substantial body of empirical work on risk aversion in real-world auctions, with available evidence strongly confirming its relevance in practice.[1] Similarly, although perhaps less widely known, endogenous entry can also overturn core predictions of standard auction theory; for instance, endogenous entry can lead a seller to prefer *less* potential competition (Li and Zheng (2009)) or a zero reserve price (Levin and Smith (1994)). Building on these observations, a substantial recent literature has developed on structural analysis of auctions with entry, with findings confirming that entry is an empirically important feature of most widely studied auction markets.[2] Taken together, these literatures strongly suggest that risk aversion and entry both matter for practical and policy analysis of real-world auction markets.

The importance of integrating risk aversion and entry in a unified analytical framework was first highlighted in a pioneer theoretical analysis by Smith and Levin (1996) who illustrate that ignoring either factor could lead to misleading predictions on revenue comparison across standard auction formats.[3] Nevertheless, research analyzing both factors together remains very sparse. Furthermore, the small body of work which

[1]For example, Baldwin (1995) and Athey and Levin (2001) find that bidders diversify risk across species in U.S. Forest Service timber auctions, Ackerberg, Hirano, and Shahriar (2006) use bidder risk aversion to rationalize the use of buy-it-now options in eBay auctions, and Bajari and Hortacsu (2005) find risk aversion to be the best explanation of bidder behavior in experiments. Using more structural approaches, Lu and Perrigne (2008) and Campo, Guerre, Perrigne, and Vuong (2011) find substantial risk aversion in U.S. Forest Service timber auctions. Finally, Li, Lu, and Zhao (2014) use differences in entry between auction formats to test for risk aversion in timber auctions, again finding significant support for risk aversion.

[2]For instance, Hendricks, Pinkse, and Porter (2003) report that less than 25 percent of eligible bidders participate in U.S. Minerals Management Service "wildcat auctions" held from 1954 to 1970. Li and Zheng (2009) find that only about 28 percent of planholders in Texas Department of Transportation mowing contracts actually submit bids. Similar patterns have been reported for timber auctions (Athey, Levin, and Seira (2011), Li and Zhang (2014, 2010), Roberts and Sweeting (2013)), online auctions (Bajari and Hortacsu (2003)), highway procurement (Krasnokutskaya and Seim (2011)) and corporate takeover markets (Gentry and Stroup (2014)) among others.

[3]In particular, Smith and Levin (1996) show that entry in conjunction with decreasing average risk aversion can lead ascending auctions to outperform first-price auctions, thereby reversing Maskin and Riley (1984)'s prediction of first-price dominance. This result also contrasts with the revenue equivalence between standard auctions with risk neutral bidders and endogenous entry (Levin and Smith (1994), Gentry and Li (2012a)).

does exist (e.g. Smith and Levin (1996), Fang and Tang (2014), Li, Lu, and Zhao (2014)) focuses primarily on theory and testing rather than estimation.[4] This is due at least in part to the fact that little is presently known about identification in auctions with both risk averse bidders and selective entry, particularly in settings where the number of entrants is disclosed only after the auction concludes.[5] More precisely, existing results on nonparametric identification with risk averse bidders turn on two classes of exclusion restrictions: either invariance of the latent distribution of values among bidders to the seller's choice of auction format (Lu and Perrigne (2008)), or invariance of the latent distribution of values among bidders to the set of competitors faced (Guerre, Perrigne, and Vuong (2009)). When entry is potentially selective, however, both variation in auction format (Li, Lu, and Zhao (2014)) and variation in the set of potential competitors (Gentry and Li (2014)) will endogenously shift the distribution of valuations among bidders choosing to enter, thereby violating the key exclusion restrictions needed for nonparametric identification. Thus little is presently known about identification in environments with both risk averse bidders and selective entry. Given the qualitative and quantitative importance of both risk aversion and entry in real-world auction design, we view this as a substantial constraint on empirical analysis of auction markets.[6]

Motivated by this gap in the literature, we explore identification and estimation in first-price auctions with risk averse bidders and selective entry, building on a flexi-

[4]For instance, Fang and Tang (2014) develop a nonparametric test for risk aversion in ascending auctions based on entry and bidding data plus information on entry costs, and Li, Lu, and Zhao (2014) test predictions the framework we consider here using data on entry in U.S. Forest Service timber auctions. We discuss both papers in detail below.

[5]We view the assumption of known potential competition but unknown actual competition as best reflecting typical institutional practices in sealed-bid procurement markets, where the auctioneer may reveal a set of planholders prior to bidding but generally discloses the set of entrants only after the auction concludes. In settings where a known number of rivals is considered a preferable assumption, identification would be considerably simpler since – conditional on the set of potential competitors faced – variation in actual entry may be effectively exogenous.

[6]Correction for selection has been a central focus in econometrics since at least the work of Heckman (1976). The primary motivation for this literature is the concern that economic actors may base choices on unobserved information regarding their private types – precisely the concern our entry model seeks to address. Thus while we frame analysis in the context of an auction game, our results also contribute to the literature on selection more broadly.

ble framework we label the Affiliated Signal with Risk Aversion (AS-RA) model. First proposed by Li, Lu, and Zhao (2014), the AS-RA model considers a set of symmetric potential bidders with wealth preferences described by a smooth concave von Neumman-Morganstern (vNM) utility function U competing for a single indivisible object via a first-price auction with entry. Potential bidders have independent private values, observe signals of their values prior to entry, and choose whether to incur a fixed entry cost, with entrants learning their values and submitting bids. This framework flexibly nests a wide range of existing models as special cases, including the affiliated-signal (AS) models of Marmer, Shneyerov, and Xu (2013) and Gentry and Li (2014) (which build on the indicative bidding model of Ye (2007)), the mixed-strategy entry model of Levin and Smith (1994), the perfectly selective entry model of Samuelson (1985), and models with risk averse bidders but exogenous entry including Guerre, Perrigne, and Vuong (2009) and Campo, Guerre, Perrigne, and Vuong (2011). It thereby represents a natural focal point for researchers seeking to understand the structural interaction between risk aversion and entry. The results we develop here provide a formal foundation for this program of research.

Working within the flexible AS-RA model, we make several key contributions to the econometric analysis of auction data. We begin by studying identification in auctions with both risk aversion and selection, mapping excludable variation in potential competition through restrictions generated by the bidding model to characterize the set of primitives consistent with observed bidding behavior. As in Gentry and Li (2014), this set will not be a singleton, although numeric analysis suggests that bounds on primitives are both reasonably tight and economically meaningful in that – for instance – the null hypothesis of risk neutrality is typically outside the identified set when the true process involves risk aversion. We then proceed to consider semiparametric identification under two natural classes of restrictions on model primitives: first assuming a parametric family for the utility function U, then assuming a parametric family for the copula C linking pre-entry signals to post-entry values. We show that either class of restrictions is typically sufficient to restore semiparametric identification of U, with parametric U yielding partial identification and parametric C yielding point identification of remaining primitives.[7] Finally, building on this analysis, we outline a simple semiparametric estimator combining Constant Relative Risk Aversion (CRRA) utility with a parametric signal-value copula. Monte Carlo analysis suggests that this estimator performs well even in small samples, underscoring the practical nature of our identification results.

Within the literature on structural analysis of auction data, our work relates most closely to three prior studies. The first of these is Guerre, Perrigne, and Vuong (2009) (henceforth GPV (2009)), who study identification in auctions with nonparametric U but exogenous participation. In this setting, GPV (2009) show that excludable variation in the number of bidders n yields nonparametric identification of model primitives, where "excludable" in the sense of GPV (2009) means that neither U nor the equilibrium distribution of values among bidders depend on the realization of n. We parallel GPV (2009) in considering excludable variation in auction-level competition as a source of identifying information. In our context, however, "excludable" means invariance of *ex ante* primitives with respect to the number of *potential* competitors N, with both entry and the distribution of valuations among entrants responding endogenously to the set of competitors faced. Hence exclusion no longer implies invariance of the latent distribution of values with respect to N, thereby undermining the key hypothesis of the GPV (2009) identification argument.[8] We show, however, that given any *candidate* for the copula C, there exists an identified map h such that *reindexing bid quantiles by* h restores GPV (2009) style quantile invariance. Applying this key insight, arguments similar to GPV (2009) then yield nonparametric identification of primitives up to C, with a parametric family for C leading to overidentification of copula parameters (and

[7]Intuitively, a parametric signal-value copula explicitly links bidding behavior across competition levels, thereby resolving the main challenge induced by selection. Meanwhile, while parametric utility helps to identify the preference dimension of the model, it leaves the entry dimension effectively unrestricted.

[8]As noted above, we find the assumption of known N but unknown n as a natural formalization of information in many sealed-bid procurement contexts, where the number of potential competitors may be common knowledge but the number of actual bidders is typically revealed only after the auction concludes. In settings where n is observed prior to bidding, however, variation in n conditional on N may provide additional GPV (2009) style identifying restrictions. Our analysis here focuses on the more challenging – but in our view more realistic – case where n is observed only after bidding.

hence point identification of the model) even with fully nonparametric U. We thereby extend the fundamental insight of GPV (2009) to a substantially more general class of models accommodating endogenous and arbitrarily selective entry.

Second, our work extends Gentry and Li (2014) (henceforth GL (2014)), who study identification in auctions with arbitrarily selective entry but risk-neutral bidders. Working within essentially the same AS entry framework we consider here, GL (2014) show how excludable variation in potential competition translates into sharp nonparametric bounds on model primitives.[9] While this analysis conveys insights useful in our context, the addition of an unknown utility function U to the AS entry model radically transforms (and complicates) the identification problem. In particular, the *point of departure* for GL (2014)'s analysis is the hypothesis that distributions of values among entrants at each competition level are identified – assuming risk neutral bidders, this follows immediately from standard results in the literature (e.g. Guerre, Perrigne, and Vuong (2000), Athey and Haile (2005)). In contrast, our *main problem* is to establish joint identification of U and distributions of values among entrants from distributions of bids observed at each competition level – in other words, to reach the point of departure for GL (2014). Hence although this study explores essentially the same entry framework as GL (2014), our identification analysis is almost entirely novel.

Finally, our work builds on Campo, Guerre, Perrigne, and Vuong (2011) (henceforth CGPV (2011)), who study semiparametric identification and estimation in auctions with risk averse bidders but with exogenous participation. In particular, given a set of covariates Z varying across auctions, CGPV (2011) show that a parametric form for U plus parameterization of one quantile of the distribution of private values (as a function of Z) yields semiparametric identification of all primitives. Although motivated by a substantially different problem, our analysis of identification in the AS-RA model under restrictions on U ultimately links back to CGPV (2011) in the following sense: we derive a sharp characterization of restrictions generated by the bidding model under any class

[9] We impose two minor technical restrictions (differentiability and invariant support) not maintained by GL (2014), but apart from these our entry frameworks are identical.

of assumptions on U, one element of which involves invariance of the top quantile of latent valuations across N. When U is assumed parametric, this restriction reduces to a system of equations in the unknown parameters of U, which essentially parallels the identifying system of CGPV (2011).[10] Under regularity conditions analogous to those in CGPV (2011), this system will be sufficient to recover the parameters in U, yielding point identification of U and distributions of values among entrants. In contrast to CGPV (2011), this information is insufficient to point identify remaining primitives, instead yielding partial identification of the model as in GL (2014).[11] Combining these observations, we ultimately obtain a new result on semiparametric identification in the AS-RA model with parametric utility, stated formally as Proposition 1.

Our results also contribute to the literature on structural analysis of auctions with risk averse bidders and / or selective entry more broadly. Broadly speaking, studies in this literature fall into one of three major categories. First, a substantial body of work exists on auctions with risk averse bidders but without entry. In addition to the studies cited above, notable contributions to this branch of the literature include Lu and Perrigne (2008), who explore identification and estimation of risk aversion based on comparisons between first- and second-price auctions, Zincenko (2014), who develops a procedure for nonparametric sieve estimation within the identification framework of GPV (2009), and Zhu and Grundl (2014), who propose a test for risk aversion in auctions with multiplicative unobserved auction-level heterogeneity.[12] Apart from Zhu and Grundl (2014), studies in this literature typically find substantial evidence of risk aversion, motivating our investigation here.

Second, a smaller but growing body of empirical work explores the role of selective entry within risk-neutral models of bidding. For instance, Marmer, Shneyerov, and Xu (2013) develop nonparametric specification tests for the perfectly selective (Samuelson

[10]The presence of entry implies that optimal bidding is characterized by a slightly different first order condition, but the systems are otherwise identical.

[11]Indeed, the risk neutrality assumption of GL (2014) can be interpreted as a (strong) parametric form for utility, with even this assumption insufficient for point identification.

[12]Note that as demonstrated by Li, Lu, and Zhao (2014), differences in auction format will generally induce differential selection into entry. Hence in our setting variation in auction format typically will not yield identification.

(1985)), non-selective (Levin and Smith (1994)), and affiliated signal (AS) entry models and apply these in the context of Texas Department of Transportation roadside mowing auctions, finding that the AS model fits substantially better than either polar alternative. Roberts and Sweeting (2013) and Bhattacharya, Roberts, and Sweeting (2014) apply parametric variants of the risk-neutral AS model to ascending U.S. Forest Service timber auctions and first-price Michigan Department of Transportation highway procurement auctions respectively, finding evidence of substantial selection in both settings. Finally, Bhattacharya and Sweeting (2014) numerically explore the implications of selection for auction design, finding (again in a risk-neutral context) that failure to account for selection can substantially distort counterfactual policy analysis.

Third, we are aware of at least two studies exploring both risk aversion and selection, although these focus primarily on testing. The first of these is Fang and Tang (2014), who propose a nonparametric test for risk aversion in ascending auctions based on entry and bidding data, which can be extended to accommodate selective entry given data on entry costs. The second is Li, Lu, and Zhao (2014), who develop predictions of the AS-RA model and test these using data on U.S. Forest Service timber auctions, finding substantial evidence of risk aversion. Our framing of the AS-RA model follows Li, Lu, and Zhao (2014), and our analysis draws heavily upon their theoretical results. To our knowledge, however, our study is the first to explore identification and estimation within the AS-RA model, thereby supporting structural analysis unifying risk aversion, endogenous entry, and selection.

The rest of this paper is organized as follows. Section 2 outlines the AS-RA model and characterizes its key predictions. Section 3 formalizes the identification problem arising when variation in potential competition is excludable and establishes several preliminary results. Section 4 analyzes identification under restrictions on utility, first characterizing restrictions generated by the bidding model under any given family for U, showing in particular that a parametric form for utility is typically sufficient to point identify U but supports only partial identification of remaining primitives. Meanwhile, Section 5 analyzes identification under restrictions on the copula, showing (in contrast to Section 4) that a parametric form for C is typically sufficient to point identify all primitives in the model. Section 6 translates our identification analysis into a simple semiparametric estimator assuming both a parametric copula and CRRA utility, and Section 7 analyzes performance of this estimator in finite samples. Finally, Section 8 concludes. We collect additional results in three appendices: Appendix A presents technical proofs, Appendix B explores the bounds implied by our characterization of nonparametric restrictions imposed by the bidding model, and Appendix C explores semiparametric estimation with a parametric copula but nonparametric utility.

2 The AS-RA model

Following Li, Lu, and Zhao (2014), we consider allocation of an indivisible good among N (≥ 2) symmetric potential bidders via a two-stage auction game, where bidders have private values for the object being sold. Timing of the game is as follows. First, in Stage 1, each potential bidder i receives a private signal S_i of her (unknown) private value V_i, and all potential bidders simultaneously choose whether to enter the auction at cost c. Next, in Stage 2, the n bidders who chose to enter in Stage 1 learn their true private values v_i and submit bids for the object being sold. Finally, the object is allocated among these bidders through a first-price sealed-bid auction with a non-binding reserve price $r = 0$. Higher Stage 1 signal realizations are "good news" in the sense that the distribution of V_i given $S_i = s_i$ is stochastically increasing in s_i, with value-signal pairs (V_i, S_i) drawn independently across bidders from a symmetric joint distribution $F_{vs}(v, s)$. Without loss of generality, we normalize Stage 1 signals to have a standard uniform distribution: $S_i \sim U[0, 1]$.

Risk aversion Potential bidders are risk averse with risk preferences described by some symmetric concave Bernoulli utility function $u(w)$, where w is net post-auction wealth. To avoid negative post-auction wealth, we assume bidders are endowed with common initial wealth $w_0 \geq c$. To simplify the analysis, we follow Li, Lu and Zhao

(2014) in defining a *centered utility function* $U(\cdot)$ as follows:

$$U(w) \equiv u(w + w_0 - c) - u(w_0 - c),$$

where we normalize $U(1) = u(1 + w_0 - c) - u(w_0 - c) \equiv 1$ without loss of generality. As noted by Li, Lu and Zhao (2014), centered utility $U(\cdot)$ belongs to the same category of Arrow-Pratt absolute risk aversion (increasing, constant, or decreasing) as initial utility $u(\cdot)$. Furthermore, as we show below, knowledge of U is equivalent to knowledge of (u, w_0) in terms of characterizing equilibrium entry and bidding behavior. We thus frame our subsequent analysis in terms of centered utility U.

Information structure As usual, the entry cost c, centered utility function U, and joint value-signal distribution F_{vs} are known to all potential bidders, with value-signal realizations (v_i, s_i) being private information revealed with timing described above. We take the number of potential competitors N to be common knowledge prior to entry, but assume that the number of entrants n is revealed only after the auction concludes.[13] In our view, this informational structure best reflects institutional practices typical in sealed-bid markets, where auctioneer announcements or industry experience convey knowledge of potential competition but actual bids are announced only after bids are received.[14] Known n would substantially change details of the derivation, but in general would strengthen identification results.[15]

[13]While not universal in the literature, the assumption of unknown n is common in applied studies: see, for instance, Li and Zheng (2009), Marmer, Shneyerov, and Xu (2013), GL (2014) and Li, Lu, and Zhao (2014) among others. For closely related studies assuming both N and n are common knowledge, see for example Levin and Smith (1994) or Smith and Levin (1996) among others.

[14]For example, in US highway procurement markets, the auctioneer will typically publish a list of planholders (potential entrants) on each contract prior to the letting date. But only a small fraction of planholders actually submit bids (Li and Zheng (2009)), and the set of bids received is only disclosed after the letting concludes. We view such auctions as naturally modeled by the assumption of known N but unknown n.

[15]In circumstances where known n is considered a preferable assumption, one would condition bidding strategies on both N and n. Maintaining the assumption that (V_i, S_i) pairs are independent across bidders, this would substantially simplify identification: conditional on N, realizations of n would be effectively random, allowing for direct application of GPV (2009) identification arguments.

2.1 Definitions

We follow GPV (2009) in defining the following (weak) regularity classes for the centered utility function U and the marginal distribution of private valuations V_i:

Definition 1. Let $\mathcal{U}$ be the set of normalized utility functions $U(\cdot)$ such that:

1. $U : [0, \infty] \to [0, \infty]$, $U(0) = 0$, and $U(1) = 1$.
2. $U(\cdot)$ is continuous on $[0, \infty]$ and admits three continuous derivatives on $(0, \infty)$, with $U'(\cdot) > 0$ and $U''(\cdot) \leq 0$ on $(0, \infty)$.
3. $\lim_{x \downarrow 0} \lambda^{(r)}$ is finite for $1 \leq r \leq 2$, where $\lambda(x) \equiv u(x)/u'(x)$ and $\lambda^{(r)}$ is the rth derivative of $\lambda(\cdot)$.

Definition 2. Let $\mathcal{F}$ be the set of distributions $F(\cdot)$ such that:

1. $F(\cdot)$ is a cumulative distribution function (c.d.f.) with support of the form $[0, \bar{v}]$, where $0 < \bar{v} < \infty$.
2. $F(\cdot)$ is twice continuously differentiable on $[0, \bar{v}]$.
3. $f(\cdot) > 0$ on $[0, \bar{v}]$.

Finally, to close the model, we describe the bivariate copula C linking pre-entry signals S_i to post-entry values V_i. Given a continuous marginal c.d.f. F for V_i and normalizing S_i to be marginal uniform as above, we know by Sklar's theorem that for every bivariate c.d.f. F_{vs} there exists a unique bivariate copula C such that $F_{vs}(v, s) = C(F(v), s)$ for all v, s. Our focus on C (rather than F_{vs}) is thus without loss of generality. Specifically, we introduce the following regularity class for C:

Definition 3. Let $\mathcal{C}$ be the set of bivariate copula functions $C(\cdot)$ such that:

1. $C(\cdot)$ is a joint c.d.f. on $[0, 1] \times [0, 1]$.
2. C is continuous on $[0, 1] \times [0, 1]$ and twice differentiable on $(0, 1) \times (0, 1)$.
3. For all $s \in (0, 1)$, $\partial^2 C(a, s)/\partial a\, \partial s$ satisfies $\partial^2 C(a, s)/\partial a\, \partial s > 0$ for all $a \in [0, 1]$.
4. For all $s \in (0, 1)$, $\partial^2 C(a, s)/\partial s^2$ satisfies $\partial^2 C(a, s)/\partial s^2 \leq 0$ for all $a \in [0, 1]$.

While Definitions 1 and 2 involves only weak regularity conditions, Definition 3 imposes two nontrivial restrictions on the joint distribution of signals and values. First, Condition 3 of Definition 3 ensures that the support of V_i is invariant to the signal realization $S_i = s$ drawn by bidder i. In particular, this rules out the Samuelson (1985) assumption of perfect pre-entry information, although the model can approach this arbitrarily closely as a limit. Second, noting that $F(v|s) = \frac{\partial C(F(v),s)}{\partial s}$, Condition 4 of Definition 3 implies $F(v|s') \leq F(v|s)$ for all $s' > s$. In other words, the distribution of V_i conditional on S_i is stochastically ordered in S_i in the sense that higher realizations of pre-entry signals lead prospective entrants to expect (weakly) stochastically increasing distributions of post-entry values.

Two further comments on this structure should be noted here. First, the assumptions on F_{vs} embedded in Definitions 2 and 3 closely parallel the assumptions on F_{vs} maintained in GL (2014), with one notable difference: we assume that F_{vs} admits a positive joint density on $[0, \bar{v}] \times [0, 1]$, whereas GL (2014) impose somewhat weaker smoothness restrictions on F_{vs}. The main practical implication is that GL (2014) formally nest the perfectly selective Samuelson (1985) model whereas we only approach it as a limit. Second, although to maintain consistency with prior work (Ye (2007), Marmer, Shneyerov, and Xu (2013), GL (2014)) we use the "Affiliated Signal" label, in fact we neither need nor assume affiliation between signals and values. The weaker assumption of stochastic ordering is sufficient for all results.

2.2 Equilibrium

We seek a symmetric monotone Bayesian Nash equilibrium in our two-stage auction game. Suppose that Stage 1 entry involves an entry threshold $\bar{s}$ such that bidder j enters if and only if $S_j \geq \bar{s}$; note that any monotone equilibrium must involve such an entry rule. For each $\bar{s} \in [0, 1)$, we seek a strictly increasing bidding strategy $\beta(\cdot|N, \bar{s})$ such that bidder i with valuation v_i optimally bids $\beta(v_i|N, \bar{s})$ when facing $N - 1$ rivals who enter according to $\bar{s}$ and bid according to $\beta(\cdot|N, \bar{s})$.

Let $\Psi(\cdot|N, \bar{s})$ be the c.d.f. of the maximum valuation among rival entrants when

i's $N-1$ rivals enter according to threshold $\bar{s}$:

$$\Psi(y|N,\bar{s}) = [\bar{s} + (1-\bar{s})F(y|s_j \geq \bar{s})]^{N-1},$$

where $F(\cdot|s_j \geq \bar{s})$ denotes the c.d.f. of rival j's valuation conditional on choosing to enter at threshold $\bar{s}$:

$$F(y|s_j \geq \bar{s}) = \frac{1}{1-\bar{s}} \int_{\bar{s}}^{1} F(y|t)\, dt.$$

Let $\pi_i(y_i, v_i; \bar{s})$ be the expected interim profit of an entrant with valuation v_i who bids as if his type were y_i against rivals entering according to $\bar{s}$ and bidding according to $\beta(\cdot|N,\bar{s})$. Following Li, Lu and Zhao (2014), we can express $\pi_i(y_i, v_i; \bar{s})$ as follows:

$$\pi_N(y_i, v_i; \bar{s}) = U(v_i - \beta(y_i|N,\bar{s}))\Psi(y_i|N,\bar{s}) + u(w_0 - c).$$

Taking a first-order condition with respect to y_i, enforcing the equilibrium condition $y_i = v_i$, and solving for $\beta_v(\cdot|N,\bar{s})$, we conclude that $\beta(\cdot|N,\bar{s})$ must satisfy

$$\beta_v(v|N,\bar{s}) = \lambda(v_i - \beta(v_i|N,\bar{s}))\frac{\Psi_v(v|N,\bar{s})}{\Psi(v|N,\bar{s})}, \tag{1}$$

where as above $\lambda(x) \equiv U(x)/U'(x)$. Note that we can rewrite

$$\frac{\Psi_v(v|N,\bar{s})}{\Psi(v|N,\bar{s})} = \frac{(N-1)(1-\bar{s})f(v|s_j \geq \bar{s})}{\bar{s} + (1-\bar{s})F(v|s_j \geq \bar{s})}.$$

Substituting into the differential equation (1) and imposing the boundary condition $\beta(r|N,\bar{s}) = r$, we obtain an initial value problem characterizing $\beta(\cdot|N,\bar{s})$:

$$\begin{aligned} &\beta(r|N,\bar{s}) = r, \\ &\beta_v(v|N,\bar{s}) = \lambda(v - \beta(v|N,\bar{s}))\frac{(N-1)(1-\bar{s})f(v|s_j \geq \bar{s})}{\bar{s} + (1-\bar{s})F(v|s_j \geq \bar{s})}, \quad v \in [r, \bar{v}]. \end{aligned} \tag{2}$$

Li, Lu, and Zhao (2014) show that (2) yields a unique solution $\beta(\cdot|N,\bar{s})$ which is increasing in v, increasing in N, and decreasing in $\bar{s}$. From this, it follows that expected *equilibrium* Stage 2 profit $\pi_N^*(v_i;\bar{s}) \equiv \pi_N(v_i, v_i;\bar{s})$ will be increasing in v_i, decreasing in N, and increasing in $\bar{s}$.

Now consider the Stage 1 entry decision of potential bidder i with signal s_i facing $N-1$ potential rivals who enter according to $\bar{s}$ and bid according to $\beta(\cdot|N,\bar{s})$. Recalling that i earns expected payoff $u(w_0)$ from staying out, the *change* in payoff i expects from entry is

$$\Pi(s_i,\bar{s},N) = \int_0^{\bar{v}} U(v-\beta(v|N,\bar{s}))\Psi(v|N,\bar{s})\,dF(v|s_i) + u(w_0-c) - u(w_0),$$

which we may equivalently rewrite as

$$\Pi(s_i,\bar{s},N) = \int_0^{\bar{v}} U(v-\beta(v|N,\bar{s}))\Psi(v|N,\bar{s})\,dF(v|s_i) - U(c). \tag{3}$$

Finally, let s_N^* be the signal threshold characterizing *equilibrium* Stage 1 entry at competition N. Clearly, a bidder with $\Pi(s_i, s_N^*, N) > 0$ will enter with certainty, and conversely for $\Pi(s_i, s_N^*, N) < 0$. At any equilibrium with nontrivial entry, s_N^* must therefore be such that a bidder with signal $s_i = s_N^*$ is just indifferent to entry:

$$\Pi(s_N^*, s_N^*, N) = 0. \tag{4}$$

Li, Lu, and Zhao (2014) show that $\Pi(s_i,\bar{s},N)$ is increasing in s_i, strictly increasing in $\bar{s}$, and decreasing in N. Hence Equation (4) will uniquely determine s_N^*.[16] Furthermore, given rival entry according to $s_N^* \in [0,1)$, equilibrium bidding behavior is uniquely described by the strategy $\beta(\cdot|N, s_N^*)$ derived above. We thereby conclude:

Theorem 1 (Li, Lu and Zhao (2014))**.** *Suppose that $U \in \mathcal{U}$, $F \in \mathcal{F}$, and $C \in \mathcal{C}$. Then there exists a unique symmetric monotone pure strategy Bayesian Nash equilibrium in the two-stage auction game. The equilibrium bidding strategy $\beta(\cdot|N, s_N^*)$ is the unique solution to the initial value problem* (2) *with $\bar{s} = s_N^*$. The equilibrium entry threshold*

[16]More precisely, either $\Pi(0,0,N) > 0$ and all bidders enter, $\Pi(1,1,N) < 0$ and no bidder enters, or there is a unique solution s_N^* to (4).

s_N^* *is uniquely determined as follows:*

- *If* $\Pi(0,0,N) > 0$*, then* $s_N^* = 0$ *and all bidders enter.*
- *If* $\Pi(1,1,N) > 0$*, then* $s_N^* = 1$ *and no bidder enters.*
- *Otherwise,* s_N^* *is the unique solution to the breakeven condition* $\Pi(s_N^*, s_N^*, N) = 0$.

Furthermore, if $N' > N$ *then* $s_{N'}^* \geq s_N^*$*, and in particular* $s_{N'}^* \in (s_N^*, 1)$ *if* $s_N^* \in (0,1)$.

Proof. See Li, Lu, and Zhao (2014). $\square$

We next outline the specific identification problem analyzed in this paper.

3 The identification problem with excludable N

As in GPV (2009), CGPV (2011), and GL (2014), any meaningful analysis of identification must begin by imposing some form of exclusion restriction; without this we could identify neither risk aversion nor selection, let alone both. As a point of departure for our identification analysis, we follow GL (2014) in assuming that variation in potential competition N is *excludable* in the sense that model primitives are invariant to realizations of N.[17] Excludable N directly extends the core identifying restriction of GPV (2009) to environments with entry. It also follows several prior studies using variation in N for auction-related hypothesis testing: for instance, Haile, Hong, and Shum (2003), use it to test for affiliated values and Marmer, Shneyerov, and Xu (2013) use it to test competing entry specifications. This section describes the identification problem arising under excludable variation in N, focusing on predictions generated by the bidding model. The next two sections translate these predictions into identified sets for model primitives under restrictions on U and C respectively.

[17] Gentry and Li (2014) also introduce a factor Z assumed to shift entry costs but not the joint signal-value distribution. In a risk-neutral context, they show that variation in this factor can support point identification of model primitives. When bidders are risk averse, however, variation in entry costs will also affect the utility function characterizing second-stage bidding. While it may be possible to extend the arguments developed here to establish identification in this case, we have not attempted to do so as we believe the underlying exclusion restriction is of primarily theoretical interest. As we discuss in more detail in the conclusion, however, in environments with asymmetric bidders one could also take Z to be the set of rival types, and in this case continuous variation in Z could yield point identification.

3.1 Identifying assumptions

Suppose that the econometrician has access to a large cross section of auctions from data generating process $\mathcal{L}$, where for each auction l the following variables are observed: number of potential competitors N_l, number of entrants n_l, the vector of submitted bids $\mathbf{b}_l$. As usual, all results extend immediately conditional on any further set of auction-level covariates X_l.

As described above, we here assume that potential competition is *excludable* in the sense that model primitives are invariant to realizations of N:

Assumption 1. For all $N \in \mathcal{N}$, $U(x|N) = U_0(x)$, $F(v|N) = F_0(v)$, $C(a,s|N) = C_0(a,s)$, and $c(N) = c_0(z)$.

We interpret observed entry and bidding outcomes as arising from symmetric Bayesian Nash Equilibrium play in a cross-section of auctions identical up to observables. We impose the following (weak) regularity conditions on equilibrium bidding behavior:

Assumption 2. $U_0 \in \mathcal{U}$, $F_0 \in \mathcal{F}$, and $C_0 \in \mathcal{C}$.

Assumption 3. For all distinct $N, M \in \mathcal{L}$, equilibrium bid strategies $\beta(\cdot; N, s_N^*)$ and $\beta(\cdot; M, s_M^*)$ have no more than finitely many points of intersection.

Following GL (2014), the structure we consider here can readily be relaxed to accommodate unobserved auction heterogeneity, interpreted as an auction-level value shifter known to bidders but not the econometrician. So long as some variation in N remains after conditioning on the auction-level unobservable, all results below extend to this much more general case. As the details of this extension closely parallel GL (2014), we simply sketch main ideas in Section 8.

3.2 Directly identified objects

Let K denote the cardinality of $\mathcal{N}$ and $N_1, ..., N_K$ denote the elements of $\mathcal{N}$, ordered such that $N_1 < N_2 < ...N_K$. For each $k \in \mathcal{K} \equiv \{1, ..., K\}$, let s_k be the entry threshold, $F_k = F(\cdot|s_i \geq s_k)$ be the distribution of valuations among entrants, and G_k be the distribution of bids generated by play of the the symmetric Bayesian Nash Equilibrium of the AS-RA model at competition N_k and primitives (U_0, C_0, F_0, c_0). Similarly, for each $k \in \mathcal{K}$, let $b_k(\cdot)$ be the quantile function of $G_k(\cdot)$ and $v_k(\cdot)$ be the quantile function of F_k. Finally, let v_0 be the quantile function of the ex ante value distribution $F_0(\cdot)$. We assume $K \geq 3$ throughout; while slightly more restrictive than GPV (2009) (who require only $K \geq 2$), this constraint is unlikely to be binding in applications.

As usual, observation of bids at each competition level N_k will directly identify $G_k(\cdot)$ for each $k \in \mathcal{K}$. Similarly, by our normalization $S_i \sim U[0,1]$, we have

$$s_k = 1 - \frac{E(n|N_k)}{N_k}.$$

Process $\mathcal{L}$ thus directly identifies the equilibrium bid distribution $G_k(\cdot)$ and the equilibrium entry threshold s_k prevailing at each competition level $k \in \mathcal{K}$. For purposes of our identification analysis, we take these objects as known.

3.3 The bid-stage identification problem

Formally, the sharp identification problem to characterize the set of primitives (U, C, F, c) consistent with observed entry and bidding behavior given excludable variation in N. In practice, however, knowledge of c_0 follows directly from knowledge of (U_0, C_0, F_0) through the equilibrium entry condition (4). Furthermore, while in principle the entry condition (4) does convey some information on bid-stage primitives (U_0, C_0, F_0), GL (2014)'s analysis of the risk-neutral case suggests that this information is typically of little practical value. We thus focus here on what we call the *bid-stage identification problem*: recovering bid-relevant primitives from identified bid distributions taking observed entry behavior as given. This turns out to lead to a much cleaner and (in our view) more useful characterization of primitives consistent with bid-stage data.

In particular, given our focus on bid-stage identification, we can reframe the problem as follows. Taking entry behavior as given, U_0 matters for bid-stage behavior only through $\lambda_0^{-1} \equiv U_0/U_0'$. We therefore follow GPV (2009) in framing bid-stage identification in terms of λ_0^{-1} rather than U_0; since U_0 is identified from λ_0^{-1} up to location and scale normalizations, this involves no loss of generality. Formally, let Λ^{-1} be the set of functions λ^{-1} such that $\lambda^{-1} = [U/U']^{-1}$ for some $U \in \mathcal{U}$. We call each tuple $(\lambda^{-1}, C, F) \in \Lambda^{-1} \times \mathcal{C} \times \mathcal{F}$ a *candidate bidding structure* for process $\mathcal{L}$. We say that a candidate bidding structure *rationalizes bidding behavior* if for each $k \in \mathcal{K}$ the structure (λ^{-1}, C, F) rationalizes bid distribution G_k taking s_k and N_k as given. The bid-stage identification problem is then to characterize the set of candidate bidding structures (λ^{-1}, C, F) rationalizing bidding behavior within the primitive space $\Lambda^{-1} \times \mathcal{C} \times \mathcal{F}$, with (bid-stage) point identification following if this set is a singleton.

3.4 Quantile inverse bidding function

We next derive the key equilibrium restriction linking the observed bid distributions $G_1, ..., G_K$ to the latent value distributions $F_1, ..., F_K$ prevailing at competition levels $N_1, ..., N_K$. Toward this end, we first rearrange Equation (2) to obtain

$$\lambda_0(v - \beta(v; N_k, s_k)) = \frac{s_k + (1 - s_k)F_k(v)}{(N_k - 1)(1 - s_k)f_k(v)} \beta_v(v|N_k, s_k).$$

Now following GPV (2009), apply the change of variables $b_i = \beta(v_i; N_k, s_k)$ to obtain

$$\lambda_0(v_i - b_i) = \frac{s_k + (1 - s_k)G_k(b_i)}{(N_k - 1)(1 - s_k)g_k(b_i)}.$$

Concavity of U_0 implies $\lambda_0'(x) = 1 - U_0(x)/U_0(x)'' > 1$, so we can invert $\lambda_0(\cdot)$ to obtain an equilibrium inverse bid function of the form:

$$v_i = b_i + \lambda_0^{-1}\left(\frac{s_k + (1 - s_k)G_k(b_i)}{(N_k - 1)(1 - s_k)g_k(b_i)}\right). \tag{5}$$

Finally, re-expressing Equation (5) in terms of quantiles, we obtain an equilibrium quantile inverse bidding function paralleling GPV (2009):

$$v_k(\alpha) = b_k(\alpha) + \lambda_0^{-1}(R_k(\alpha)), \tag{6}$$

where $R_k(\alpha)$ indexes the equilibrium bid markup set by an entrant drawing value $v_i = v_k(\alpha)$ against competition N_k:

$$R_k(\alpha) \equiv \frac{s_k + (1 - s_k)\alpha}{(N_k - 1)(1 - s_k)g_k(b_k(\alpha))}. \tag{7}$$

Note that identification of s_k, G_k implies identification of b_k, R_k and hence identification of the right-hand side of (6) up to λ_0^{-1} for all $k = 1..., K$.

4 Identification under restrictions on utility

In exploring bid-stage identification of the AS-RA model, we begin with the case in which the researcher is willing to impose some structure on the latent utility function U_0. In particular, suppose that U_0 is such that the true markup function λ_0^{-1} can be represented as a member of some known family $\Lambda_\Gamma^{-1} \subset \Lambda^{-1}$, with elements of Λ_Γ^{-1} indexed by some (finite- or infinite-dimensional) parameter vector $\gamma \in \Gamma$:

Assumption 4. $\lambda_0^{-1} = \lambda^{-1}(\cdot; \gamma_0)$ for some $\gamma_0 \in \Gamma$, with Γ a known finite- or infinite-dimensional parameter space.

Note that this notation imposes no structure on Γ and hence in fact is fully general; to nest the special case of no restrictions on λ_0^{-1}, we can simply take $\Gamma = \Lambda^{-1}$. In this case the analysis below will characterize the sharp nonparametric bid-stage identified set. We introduce the generic parameter space Γ to permit a unified treatment of identification under any class of restrictions on λ_0^{-1}.

4.1 Bid-stage identified set with excludable N

Assuming $\lambda_0^{-1} \in \Lambda_\Gamma^{-1}$, we now turn to consider restrictions on (γ_0, C_0, F_0) induced by bid-stage behavior when variation in N is excludable. Toward this end, take $\gamma \in \Gamma$ as given, and let $\tilde{v}_1, ..., \tilde{v}_K$ be the unique candidates for $v_1, ..., v_K$ induced by the quantile inverse bidding function (6) under the hypothesis $\gamma = \gamma_0$:

$$\tilde{v}_k(a;\gamma) \equiv b_k(a) + \lambda^{-1}(R_k(a);\gamma) \text{ for all } k \in \{1, ..., K\}. \tag{8}$$

Observe that $\tilde{v}_1, ..., \tilde{v}_K$ are identified up to γ and well-defined for any $\gamma \in \Gamma$, and by construction $\gamma = \gamma_0$ implies $\tilde{v}_k = v_k$ for all k. Furthermore, from above, equilibrium bidding implies $R_k(0) = 0$, $b_k(0) = 0$, and $b_k(\cdot)$ and $R_k(\cdot)$ differentiable for all k. Hence for all $\gamma \in \Gamma$, $\tilde{v}_k(\cdot;\gamma)$ will be differentiable with $\tilde{v}_k(0;\gamma) = 0$ for each $k = 1, ..., K$.

Now observe that if we take λ_0^{-1} and entry behavior as given, primitives (C_0, F_0) influence bidding behavior only through the latent quantile functions $v_1, ..., v_K$, with the quantile inverse bidding function (6) equivalent to the differential equation (1) defining equilibrium bid strategies. Hence to determine whether any conjectured parameter $\gamma \in \Gamma$ is consistent with bid-stage observables, it is sufficient to determine whether there exists a structure $(C, F) \in \mathcal{C} \times \mathcal{F}$ consistent with the candidate quantile functions $\tilde{v}_1, ..., \tilde{v}_K$ generated by γ through (8). This turns out to reduce to a set of three directly verifiable restrictions on $\tilde{v}_1, ..., \tilde{v}_K$, leading to a relatively simple characterization of the sharp bid-stage identified set. We state this result formally as follows:

Definition 4. For any $\gamma \in \Gamma$, we say the pair $(C, F) \in \mathcal{C} \times \mathcal{F}$ *rationalizes bid-stage observables at* γ if the triple (γ, C, F) rationalizes bid-stage observables.

Theorem 2. *Fix $\gamma \in \Gamma$, let $\tilde{v}_1, ..., \tilde{v}_K$ be the candidate quantile functions derived from γ through* (8)*, and let $\tilde{v}_1^{-1}, ..., \tilde{v}_K^{-1}$ be pseudo-inverses of $\tilde{v}_1, ..., \tilde{v}_K$ respectively. Then for each $\gamma \in \Gamma$, there exists a structure $(C, F) \in \mathcal{C} \times \mathcal{F}$ rationalizing bid-stage observables at γ if and only if all of the following hold:*

1. *$\tilde{v}_k$ is strictly increasing for all $k = 1, ..., K$, with $\tilde{v}_1(1) = \tilde{v}_2(1) = ... = \tilde{v}_K(1)$.*
2. *For all $k = 1, .., K-1$, the functions $\tilde{v}_k^{-1}$, $\tilde{v}_{k+1}^{-1}$ are such that*

$$(1 - s_k) \cdot \frac{d}{dy}\tilde{v}_k^{-1}(y) > (1 - s_{k+1}) \cdot \frac{d}{dy}\tilde{v}_{k+1}^{-1}(y) \text{ for all } y \in (\tilde{v}_k(0), \tilde{v}_k(1)).$$

3. For all $k = 2, ..., K-1$, the functions $\tilde{v}_{k-1}^{-1}$, $\tilde{v}_k^{-1}$, and $\tilde{v}_{k+1}^{-1}$ are such that

$$1 \geq \frac{(1-s_{k-1})\tilde{v}_{k-1}^{-1} - (1-s_k)\tilde{v}_k^{-1}}{s_k - s_{k-1}} \geq \frac{(1-s_k)\tilde{v}_k^{-1} - (1-s_{k+1})\tilde{v}_{k+1}^{-1}}{s_{k+1} - s_k} \geq \tilde{v}_K^{-1} \geq 0.$$

Furthermore, in this case, the set of $(\tilde{C}, \tilde{F}) \in \mathcal{C} \times \mathcal{F}$ rationalizing bid-stage observables at γ is the set of $(\tilde{C}, \tilde{F}) \in \mathcal{C} \times \mathcal{F}$ such that for all $k = 1, ..., K$:

$$\tilde{v}_k^{-1}(y;\gamma) = \frac{\tilde{F}(y) - \tilde{C}(F(y), s)}{1 - s_k} \text{ for all } y \in [\tilde{v}_k(0;\gamma), \tilde{v}_k(1;\gamma)]; \tag{9}$$

i.e. the set of structures $(\tilde{C}, \tilde{F})$ generating ex post quantile functions $\tilde{v}_1, ..., \tilde{v}_K$.

Proof. In Appendix. □

Recall that $\tilde{v}_1, ..., \tilde{v}_K$ are differentiable for any $\gamma \in \Gamma$. Hence Condition 1 of Theorem 2 implies that the functions $\tilde{v}_1^{-1}, ..., \tilde{v}_K^{-1}$ are continuous on $[\tilde{v}_1(0), \tilde{v}_1(1)]$ and differentiable on $(\tilde{v}_1(0), \tilde{v}_1(1))$. This in turn ensures that Conditions 2 and 3 are well defined.

Now consider how the characterization of Theorem 2 helps to simplify analysis of the sharp bid-stage identified set. Recall that if we set $\Gamma = \Lambda^{-1}$ then Theorem 2 in fact summarizes nonparametric restrictions generated by the bidding model. While construction of the identified set corresponding to these restrictions is nontrivial, Theorem 2 implies that we can reduce the problem of search over the full primitive space $\Lambda^{-1} \times \mathcal{C} \times \mathcal{F}$ to that of search over a set of one-dimensional functions Λ^{-1}. Although still challenging, the latter problem is orders of magnitude easier, especially since for any $\lambda^{-1} \in \Lambda^{-1}$ existence of $(\tilde{C}, \tilde{F})$ rationalizing bid-stage behavior is equivalent to a set of readily verifiable restrictions on the directly identified objects $\tilde{v}_1, ..., \tilde{v}_K$. This in turn provides a basis for numerical approximation of the sharp bid-stage identified set for λ_0^{-1}, for instance by search over a sequence of sieve space approximating Λ^{-1} for the set of elements satisfying the conditions of Theorem 2. Each element of this set will then correspond to a set of tuples $(C, F) \in \mathcal{C} \times \mathcal{F}$ satisfying Equation (9), with the set of all such tuples approximating the sharp bid-stage identified set.

To illustrate the practical contribution of Theorem 2, we have constructed a series of numerical simulations implementing the sieve approximation algorithm sketched above. In these simulations, we first construct a sieve space $\tilde{\Lambda}^{-1}$ approximating Λ^{-1}: here the space of shape-constrained Bernstein polynomials considered by Zincenko (2014). We then analyze the set of polynomials $\tilde{\lambda}^{-1}$ within $\tilde{\Lambda}^{-1}$ at which Conditions 1-3 of Theorem 2 approximately hold. Results of this exercise strongly confirm the relevance of the bounds in Theorem 2: while λ_0^{-1} is clearly set identified, bounds in most cases are surprisingly tight, and (in particular) are clearly bounded away from risk neutrality for even moderately concave $U(\cdot)$. While we do not explore inference based on Theorem 2 in detail here, we believe this represents a promising direction for future research, with our simulation results in particular suggesting the feasibility of constructing tests for risk aversion robust even in the presence of endogenous and arbitrarily selective entry. We refer interested readers to Appendix B for further discussion.

4.2 Bid-stage identification with finite-dimensional γ_0

Now to the structure outlined above add the assumption that the parameter vector γ is finite-dimensional; i.e. that $\Gamma \subset \mathbb{R}^Q$ for some $Q < \infty$. Let $\hat{\Gamma}$ denote the set of parameters $\gamma \in \Gamma$ consistent with bid-stage observables. From Theorem 2, we can have $\gamma \in \hat{\Gamma}$ only if $\tilde{v}_1(1,\gamma) = \tilde{v}_k(1,\gamma)$ for all $k = 1, ..., K$. Defining $\bar{b}_k \equiv b_k(1)$ and taking $\bar{v} \equiv \tilde{v}_1(1)$ as an auxiliary parameter to be identified, we can express this restriction as

$$\bar{v} = \bar{b}_k + \lambda^{-1}(R_k(1); \gamma_0), \quad k = 1, ..., K.$$

But recalling that $R_k(1) \equiv 1/[(N-1)(1-s_k)g_k(\bar{b}_k)]$ and rearranging, this system is precisely the system of identifying restrictions considered in CGPV (2011), extended to accommodate endogenous and selective entry:

$$(N_k - 1)(1 - s_k)g_k(\bar{b}_k) = \frac{1}{\lambda(\bar{v} - \bar{b}_k; \gamma_0)}, \quad k = 1, ..., K. \tag{10}$$

The left-hand side of (10) is identified for each $k = 1, ..., K$, with the right-hand side identified up to $\bar{v}$ and the unknown parameters γ_0. Pooling restrictions of the form (10) across k, we thus obtain a system of K equations in the $Q+1$ unknowns $(\gamma_0, \bar{v})$, which for $K \geq Q+1$ will generally overidentify $(\gamma_0, \bar{v})$. For purposes of this subsection, we follow CGPV (2011) in maintaining this as a regularity condition:

Condition 1. The system of equations

$$(N_k - 1)(1 - s_k) g_k(\bar{b}_k) = \frac{1}{\lambda(\bar{v} - \bar{b}_k; \gamma_0)}, \quad k = 1, ..., K$$

has a unique solution $(\gamma_0, \bar{v}) \in \Gamma \times (0, \infty)$.

In other words, under assumptions paralleling those in CGPV (2011), equilibrium behavior at the top quantile of ex post valuations will be sufficient to identify γ_0 for any form of selection into entry, with identification of γ_0 yielding identification of $F_1, ..., F_K$ through the quantile inverse bid function (8). As in GL (2014), this knowledge is insufficient to identify F_0, C_0, and c_0, and hence the model as a whole will be only set identified. Given knowledge of γ_0 and $F_1, ..., F_K$, however, the problem of *partially* identifying remaining primitives reduces in essence to that analyzed in GL (2014). Extending Propositions 3 and 4 developed there, we therefore conclude:

Proposition 1. *Under Assumptions 1-2 and Condition 1, γ_0 and $F_1, ..., F_K$ are point identified while remaining primitives are set identified. In particular, identified bounds on the conditional c.d.f. $F(v|S_i = s)$ and the entry cost c_0 may be obtained as follows:*

- *Let $F^+(\cdot|\cdot)$ and $F^-(\cdot|\cdot)$ be defined as in Proposition 3 of Gentry and Li (2014). Then $F^+(\cdot|s)$ and $F^-(\cdot|s)$ are identified, describe distributions over $[0, \bar{v}]$, and bound $F(\cdot|s_i = s)$ for all $s \in [0, 1)$:*

$$F^+(\cdot|s) \geq F(\cdot|S_i = s) \geq F^-(\cdot|s) \text{ for all } s \in [0, 1).$$

- *For each $k \in \mathcal{K}$, let constants $\hat{U}_k^+$ and $\hat{U}_k^-$ be defined as follows:*

$$\hat{U}_k^+ = \int_0^{\bar{v}} U(\lambda^{-1}(R_k(F_k(v)); \gamma_0); \gamma_0) \cdot \Psi(v|N_k, s_k)\, dF^+(v|s_k)$$
$$\hat{U}_k^- = \int_0^{\bar{v}} U(\lambda^{-1}(R_k(F_k(v)); \gamma_0); \gamma_0) \cdot \Psi(v|N_k, s_k)\, dF^+(v|s_k).$$

Then $\hat{U}_k^+$ and $\hat{U}_k^-$ are identified and yield identified bounds on c_0:

$$\min_k U^{-1}(\hat{U}_k^+; \gamma_0) \geq c_0 \geq \max_k U^{-1}(\hat{U}_k^-; \gamma_0).$$

Recall that the system (10) underlying Proposition 1 yields identification of γ_0 without respect to (C_0, F_0). Given a parametric form for U_0 (e.g. Constant Absolute Risk Aversion or Constant Relative Risk Aversion), the system (10) thus provides a basis for testing the null hypothesis of risk aversion under *any* assumptions on the nature of selection into entry. Furthermore, following Gentry and Li (2012b), the bounds on primitives in Proposition 1 can be shown to imply identified bounds on expected revenue under a variety of counterfactual policy choices, such as reserve prices and entry fees. We point interested readers to Gentry and Li (2013) for further details.[18]

5 Identification under restrictions on C_0

While parametric assumption on utility are widely employed, other classes of restrictions may also prove fruitful in applications. In particular, as noted by GL (2014) in a risk-neutral context, a parametric family for the signal-value copula C_0 may have considerable identifying power in auctions with selective entry: intuitively, such structure helps to link latent quantile functions $v_1, ..., v_K$ across competition levels k, thereby directly addressing the main challenge induced by selection. Motivated by this observa-

[18]In principle, one could also exploit bounds on $F(\cdot|s_i = s)$ to derive bounds on deep structural primitives F_0 and C_0. Observe, however, that γ_0, $F(\cdot|s_i = s)$ and c_0 are precisely the quantities needed to analyze counterfactual entry and bidding behavior; once identified sets for these objects are known, implied bounds on F_0 and C_0 will add nothing to policy analysis. We thus conclude the analysis in this section at the statement of Proposition 1, without developing bounds on F_0 and C_0 in detail.

tion, this section explores semiparametric identification in the AS-RA model assuming the true copula C_0 belongs to a known family indexed by parameter vector θ:

Assumption 5. $C_0(a,s) = C_{\theta_0}(a,s)$, with the parameter vector θ_0 a member of some known parameter space Θ.

GL (2014) proposed a parametric signal-value copula as a way to obtain point identification in the AS model with risk neutral bidders; see footnote 18 in GL (2014) for details. We here show that this assumption is in fact substantially more powerful than initially envisioned by GL (2014), with a standard finite-dimensional family for C_θ typically sufficient to restore point identification of model primitives not only under risk neutrality but also in the presence of an arbitrary (nonparametric) markup function λ_0^{-1}.

Note that while our primary interest (Sections 5.3 and 5.4) will be cases where Θ is finite-dimensional, Assumption 5 imposes no structure on the parameter space Θ and hence in fact is fully general. For instance, if C_0 is assumed to be Archimedean, then one could take Θ to be the space of Archimedean generator functions, and if C_0 is simply assumed to be unrestricted, one could trivially set $\Theta = \mathcal{C}$. In this sense the results in Sections 5.1 and 5.2 apply without loss of generality.

5.1 Bid-stage identification up to θ_0

Pooling quantile inverse bidding functions of the form (6) across $k \in \mathcal{K}$, we obtain a system of K restrictions induced by equilibrium bidding with selective entry:

$$\begin{aligned} v_1(\alpha) &= b_1(\alpha) + \lambda_0^{-1}(R_1(\alpha)), \\ &\vdots \\ v_K(\alpha) &= b_K(\alpha) + \lambda_0^{-1}(R_K(\alpha)). \end{aligned} \tag{11}$$

If entry were non-selective, then we would then have $v_0(\alpha) = v_k(\alpha)$ for all $k \in \mathcal{K}$ and the identification problem would be trivial: we would need only rearrange (11) to obtain the *compatibility condition*

$$b_l(\alpha) + \lambda_0^{-1}(R_l(\alpha)) = b_k(\alpha) + \lambda_0^{-1}(R_k(\alpha)) \text{ for all } k, l \in \mathcal{K}. \tag{12}$$

The arguments in GPV (2009) would then yield identification of λ_0^{-1} on the support of the data, with identification of F_0 following through the left-hand side of (11).

Unfortunately, in the presence of selection this simple argument is no longer feasible: the αth quantile of bids at competition N_k now corresponds to latent value $v_k(\alpha)$ rather than latent value $v_0(\alpha)$, where we know only that $v_k(\alpha) \leq v_l(\alpha)$ for $k < l$. Hence in contrast to GPV (2009), we can no longer compare bids across different competition levels directly. But observe that for each $k \in \mathcal{K}$ we can rewrite the latent distribution F_k corresponding to latent quantile function v_k as follows:

$$\begin{aligned} F_k(y) \equiv F(y|s_i \geq s_k) &= \frac{1}{1-s_k}\int_{s_k}^{1} F(y|t)dt \\ &= \frac{F_0(y) - F_{vs}(v, s_k)}{1-s_k} \\ &= \frac{F_0(y) - C_0(F_0(y), s_k)}{1-s_k}. \end{aligned} \tag{13}$$

Applying the change of variables $y = v_0(a)$ on both sides of this equation gives:

$$F_k(v_0(a)) = \frac{a - C_0(a, s_k)}{1-s_k} \equiv h_0^k(a), \quad k = 1, ..., K. \tag{14}$$

Inverting F_k on both sides of (14) yields $v_0(a) \equiv v_k(h_0^k(a))$ for all $k = 1, ..., K$. Hence applying the change of variables $\alpha = h_0^k(a)$ in each line of (11) we ultimately obtain:

$$v_0(a) = b_k(h_0^k(a)) + \lambda_0^{-1}(R_k(h_0^k(a))), \quad k = 1, ..., K. \tag{15}$$

But the left-hand side of (15) is now invariant to k! In other words, *reindexing* each inverse bid function of the form (6) by its corresponding quantile mapping h_0^k, we transform the initially incompatible system (11) into a system for which GPV (2009) style compatibility holds:

$$b_k(h_0^k(a)) + \lambda_0^{-1}(R_k(h_0^k(a))) = b_l(h_0^l(a)) + \lambda_0^{-1}(R_k(h_0^l(a))) \text{ for all } k, l \in \mathcal{K}. \tag{16}$$

In practice, of course, C_0 is unknown, hence the quantile mappings $h_0^1, ..., h_0^K$ are also unknown and direct application of (15) is infeasible. But *taking* $\theta \in \Theta$ *as given*, there exists a unique set of quantile transformations $(h_\theta^1, ..., h_\theta^K)$ consistent with the hypothesis $\theta = \theta_0$:

$$h_\theta^k(a) \equiv \frac{a - C_\theta(a, s_k)}{1 - s_k}, \quad k = 1, ..., K.$$

Clearly $h_{\theta_0}^k \equiv h_0^k$ by construction, so under the hypothesis $\theta = \theta_0$, we must have:

$$b_k(h_\theta^k(a)) + \lambda_0^{-1}(R_k(h_\theta^k(a))) = b_l(h_\theta^l(a)) + \lambda_0^{-1}(R_l(h_\theta^l(a))) \text{ for all } k, l \in \mathcal{K}. \tag{17}$$

Since these identities turn on a particular hypothesis regarding θ_0, we refer to restrictions of the form (17) as *conjectured compatibility conditions* implied by the bidding model. Recall that the form of C_θ is known (up to θ_0) by Assumption 5, with $s_1, ..., s_K$ directly identified by equilibrium entry. Hence taking $\theta \in \Theta$ as given both sides of (17) are identified up to λ_0^{-1}.

Note that for $\theta \neq \theta_0$ the conjectured compatibility condition (17) will *misspecify* the true equilibrium bidding relationship (16). Hence in contrast to GPV (2009), for arbitrary $\theta \in \Theta$ and $k, l \in \mathcal{K}$ there need not exist a function $\lambda_{kl,\theta}^{-1} \in \Lambda^{-1}$ satisfying (17). But for any $\theta \in \Theta$ for which such a function exists, the problem of recovering this function from the conjectured compatibility condition (17) closely parallels the problem of recovering the true markup function λ_0^{-1} from the true compatibility condition (16) – a problem which we already know how to solve following GPV (2009). Building on this intuition, we obtain this section's main result: at any θ for which there exists a function $\lambda_{kl,\theta}^{-1} \in \Lambda^{-1}$ satisfying (17), this $\lambda_{kl,\theta}^{-1}$ will be *unique* and *constructively identified* on its domain in the data. Noting that λ_0^{-1} satisfies (17) at $\theta = \theta_0$ and that identification of $(\theta_0, \lambda_0^{-1})$ yields identification of F_0 through (15), this in turn implies that *identification* *of* $(\theta_0, \lambda_0^{-1}, F_0)$ *is equivalent to identification of* θ_0.

More formally, define $\bar{r}_k \equiv \sup_{a\in[0,1]} R_k(a)$, $\bar{r}_{kl} \equiv \max\{\bar{r}_k, \bar{r}_l\}$, and $\bar{r} \equiv \max_k \bar{r}_k$, and let $\Lambda^{-1}[0, r]$ be the set of functions obtained by restricting elements of Λ^{-1} to the domain $[0, r]$. We are now in position to state this section's main result:

Theorem 3. *Consider any distinct* $k, l \in \mathcal{K}$ *and any* $\theta \in \Theta$ *such that:*

1. *There exists a function* $\lambda_{kl,\theta}^{-1} \in \Lambda^{-1}[0, \bar{r}_{kl}]$ *at which* (17) *holds for all* $a \in [0, 1]$*;*
2. *Reindexed bid quantiles* $b_k(h_\theta^k(\cdot))$, $b_l(h_\theta^l(\cdot))$ *satisfy the same finite intersection condition as equilibrium bids (Assumption 3).*

Then the function $\lambda_{kl,\theta}^{-1}$ *satisfying* (17) *is* unique *within* $\Lambda^{-1}[0, \bar{r}_{kl}]$ *and* constructively identified *on* $[0, \bar{r}_{kl}]$ *by bid-stage behavior at* k, l *given* θ.

Note that if equilibrium bid strategies satisfy finite intersection (Assumption 3) then Conditions 1 and 2 of Theorem 3 are in fact *necessary* for $\theta = \theta_0$: if $\theta = \theta_0$, then $\lambda_0^{-1} \in \Lambda^{-1}$ clearly satisfies (17), and since $v_k(h_{\theta_0}^k(a)) \equiv v_0(a)$ and $v_0(\cdot)$ is strictly increasing finite intersection of $\beta(\cdot|N_k, s_k)$, $\beta(\cdot|N_l, s_l)$ implies finite intersection of $b_k(h_{\theta_0}^k(\cdot))$, $b_l(h_{\theta_0}^l(\cdot))$. Hence in practice these conditions involve no loss of generality.

For our purposes, Theorem 3 has two major implications. First, recall that identification of λ_0^{-1} implies identification of F_0 through (15). Hence Theorem 3 implies that given any viable candidate θ for θ_0 there exists exactly one structure $(\lambda_{kl,\theta}^{-1}, F_{kl,\theta}) \in \Lambda^{-1}[0, \bar{r}_{kl}] \times \mathcal{F}$ consistent with bid-stage behavior at any $k, l \in \mathcal{K}$, with this structure constructively identified in terms of bid-stage observables. In other words, the problem of identifying primitives $\theta_0 \times \lambda_0^{-1} \times F_0$ within the infinite-dimensional space $\Theta \times \Lambda^{-1}[0, \bar{r}] \times \mathcal{F}$ reduces to the vastly simpler problem of identifying viable candidates for θ_0 within the (often finite-dimensional) set Θ. Second, at $\theta = \theta_0$ the *same* structure (λ_0^{-1}, F_0) must rationalize bid-stage behavior at all competition pairs $k, l \in \mathcal{K}$. Since for arbitrary $\theta \neq \theta_0$ the candidate $\lambda_{kl,\theta}^{-1}$ obtained at competition pair kl will typically strictly diverge from the candidate $\lambda_{jl,\theta}^{-1}$ obtained at competition pair jl, this in turn represents a very powerful identifying restriction on the underlying copula parameter θ_0. The next two subsections develop the implications of this restriction in detail.

5.2 Bid-stage identified set for θ_0

In this section we seek to characterize the *bid-stage identified set* for θ_0; that is, the set of $\theta \in \Theta$ potentially consistent with bid-stage observables. Building on the discussion in Section 5.1, we formally define this set as follows:

Definition 5. Let the *bid-stage identified set* Θ_I for θ_0 be the set of $\theta \in \Theta$ such that:

1. For all distinct $k, l \in \mathcal{K}$, the functions $b_k(h_\theta^k(\cdot))$, $b_l(h_\theta^l(\cdot))$ satisfy finite intersection;
2. There exist $(\lambda^{-1}, F) \in \Lambda^{-1} \times \mathcal{F}$ such that for all $k \in \mathcal{K}$:

$$F^{-1}(a) = b_k(h_\theta^k(a)) + \lambda^{-1}(R_k(h_\theta^k(a))) \text{ for all } a \in [0,1]. \tag{18}$$

Now consider restrictions on Θ_I generated by bid-stage behavior. Choose $l \in \mathcal{K}$ such that $\bar{r}_l = \bar{r}$, and for each $k \in \mathcal{K}$ let $\tilde{\Theta}_{kl}$ be the set of $\theta \in \Theta$ satisfying the hypotheses of Theorem 3: i.e. such that (i) there exists a function $\phi \in \Lambda^{-1}[0, \bar{r}_{kl}]$ satisfying the candidate compatibility condition (17) and (ii) reindexed bid functions satisfy the same finite intersection condition as equilibrium bids. Clearly $\Theta_I \in \tilde{\Theta}_{kl}$ for all $k, l \in \mathcal{K}$, and by Theorem 3 we know that for each $k \in \mathcal{K}$ and each $\theta \in \tilde{\Theta}_{kl}$ there exists a unique, constructively identified candidate $\lambda_{kl,\theta}^{-1}$ for λ_0^{-1} determined up to θ by bid-stage behavior. Furthermore, by hypothesis, the *same* function λ_0^{-1} must rationalize bid-stage behavior for all $k \in \mathcal{K}$. In other words, at $\theta = \theta_0$ we must have

$$\sup_{r \in [0,\bar{r}]} |\lambda_{jl,\theta}^{-1}(r) - \lambda_{kl,\theta}^{-1}(r)| = 0 \text{ for all } j, k \in \mathcal{K} \backslash \{l\}. \tag{19}$$

Define $\tilde{\Theta} \equiv \cap_{k \in 1}^{K} \tilde{\Theta}_{kl}$; note that $\tilde{\Theta}$ is determined by observables (hence identified) with $\Theta_I \subset \tilde{\Theta}$. Let $\hat{\Theta}$ be the set of $\theta \in \tilde{\Theta}$ satisfying (19). We know that for all $\theta \in \tilde{\Theta}$ the functions $\lambda_{kl,\theta}^{-1}$, $\lambda_{jl,\theta}^{-1}$ are unique and constructively identified, hence $\hat{\Theta}$ is identified. Furthermore, since (18) implies (19), we know $\Theta_I \in \hat{\Theta}$. The set $\hat{\Theta}$ thus provides a fully general, directly identified outer bound on the bid-stage identified set Θ_I:

Proposition 2. *Define $\tilde{\Theta} = \cap_{k \in 1}^{K} \tilde{\Theta}_{kl}$ and let $\hat{\Theta}$ be the set of $\theta \in \tilde{\Theta}$ satisfying* (19). *Then $\tilde{\Theta}$ and $\hat{\Theta}$ are identified with $\theta_0 \in \Theta_I \subset \hat{\Theta}$ $(\subset \tilde{\Theta})$.*

The outer bound $\hat{\Theta}$ is fully general, applying under any class of restrictions on C_0 and informative so long as $K \geq 3$. One important special case of Proposition 2, however, arises when the system (19) has a unique solution: i.e. $\hat{\Theta} = \{\theta_0\}$. In this case identification of $\hat{\Theta}$ implies point identification of θ_0 and therefore bid-stage identification of $(\theta_0, \lambda_0^{-1}, F_0)$ through Theorem 3.

As evident in the proof of Theorem 3, the system (19) is highly nonlinear, and as usual with nonlinear equations it is difficult to provide formal sufficient conditions under which (19) will have a unique solution. But note that so long as $K \geq 3$ the restrictions embodied in (19) in fact define an *infinite* system of identifying restrictions on the unknown vector θ_0: formally, a set of $K-2$ equations in functions identified up to θ_0, each of which must hold pointwise on a continuum in $\mathbb{R}^+$. If C_0 is left fully unrestricted, even this infinite system (a set of $K-2$ restrictions on the continuum $[0, \bar{r}]$) will still be of much lower cardinality than the parameter space Θ (a set of two-dimensional functions from $[0,1] \times [0,1]$ to $[0,1]$), hence θ_0 will be only set identified. But given any standard finite-dimensional family for C_0, we expect the infinite system (19) to be greatly overdetermined, leading to point identification of θ_0 and therefore bid-stage identification of $(\theta_0, \lambda_0^{-1}, F_0)$. We next develop the intuition underlying this claim.

5.3 Bid-stage identification with finite-dimensional θ_0

Now to Assumption 5 add two further restrictions: the parameter space Θ is finite-dimensional, and the family $\mathcal{C}_\Theta$ containing C_0 is *regular* in the following sense:

Definition 6. For each $\theta \in \Theta$, let $h_\theta^{1,-1}, ..., h_\theta^{K,-1}$ denote inverses of $h_\theta^1, ..., h_\theta^K$ respectively.[19] We say $\mathcal{C}_\Theta$ is *regular* if for all $\theta, \theta' \in \Theta$ such that $\theta \neq \theta'$, the set

$$\mathcal{H}(\theta, \theta') \equiv \{a \in [0,1] : h_{\theta'}^{k,-1}(h_\theta^k(a)) \neq h_{\theta'}^{l,-1}(h_\theta^l(a)) \text{ for all } k, l \in \mathcal{K}\}$$

has positive Lebesgue measure.

[19]Recall that $h_\theta^k(\cdot)$ is strictly increasing, so $h_\theta^{k,-1}(\cdot)$ exists.

"Regularity" essentially requires the family $\mathcal{C}_\Theta$ to be such that different choices of θ lead to differential changes in selection across different observed competition levels. This property should hold trivially for any standard single-parameter copula family: e.g. Gaussian or families in the Archimedean class (Gumbel, Frank, Clayton, etc). It might fail in cases where, for instance, $\mathcal{C}_\Theta$ is specified such that some elements of θ matter for $C(\cdot; s_k)$ at only a subset of competition levels $k \in \mathcal{K}$. For our purposes, regularity of $\mathcal{C}_\Theta$ is useful mainly as a sufficient primitive condition guaranteeing that the difference $v_k(h_\theta^k(\cdot)) - v_l(h_\theta^l(\cdot))$ will vanish only at $\theta = \theta_0$:

Lemma 1. *Suppose that $\mathcal{C}_\Theta$ is regular. Then $\theta \neq \theta_0$ implies $v_k(h_\theta^k(\cdot)) \neq v_l(h_\theta^l(\cdot))$ on a set of positive measure in* $[0,1]$.

Now consider the implications of this lemma for the system (19) underlying Proposition 2. Recall that in constructing $\lambda_{kl,\theta}^{-1}(\cdot)$ for given $k, l \in \mathcal{K}$ we start from the hypothesis that the conjectured compatibility condition (17) reflects the true bidding relationship (15). At $\theta = \theta_0$, this hypothesis holds by construction and therefore $\lambda_{kl,\theta}^{-1}(\cdot)$ recovers $\lambda_0^{-1}(\cdot)$. At $\theta \neq \theta_0$, however, the conjectured relationship (17) will be *systematically misspecified*; whereas we construct $\lambda_{kl,\theta}^{-1}(\cdot)$ to satisfy

$$b_k(h_\theta^k(a)) + \lambda_{kl,\theta}^{-1}(R_k(h_\theta^k(a))) = b_l(h_\theta^l(a)) + \lambda_{kl,\theta}^{-1}(R_l(h_\theta^l(a))),$$

the true equilibrium bidding relationship will be given by

$$b_k(h_\theta^k(a)) + \lambda_0^{-1}(R_k(h_\theta^k(a))) = b_l(h_\theta^l(a)) + \lambda_0^{-1}(R_l(h_\theta^l(a))) + \left(v_k(h_\theta^k(a)) - v_l(h_\theta^l(a))\right).$$

If $\mathcal{C}_\Theta$ is regular, then for $\theta \neq \theta_0$ we will have $v_k(h_\theta^k(a)) \neq v_l(h_\theta^l(a))$ on a set of positive measure in $[0,1]$. Hence at $\theta \neq \theta_0$ the candidate markup function $\lambda_{kl,\theta}^{-1}$ solving (17) at each k, l must differ systematically from λ_0^{-1}. Furthermore, and more importantly, the misspecification $v_k(h_\theta^k(\cdot)) - v_l(h_\theta^l(\cdot))$ latent in (17) will also vary systematically with the particular k, l pair considered. Hence for $\theta \neq \theta_0$ and $j \neq k$ we will generally have $\lambda_{\theta,jl}^{-1}(\cdot) \neq \lambda_{\theta,kl}^{-1}(\cdot)$ on a set of positive measure. In other words, each functional equation of the form (19) will in fact induce a continuum of restrictions *binding* on each element of θ_0, and the resulting system will in general be greatly overdetermined in the unknown parameters θ_0. Identification of θ_0 then reduces to the (weak) regularity condition that the overdetermined system (19) has at most one solution, which in practice could naturally be maintained as an assumption:

Condition 2. The system (19) has at most one solution in the parameter space Θ.

If one is willing to maintain Condition 2, then Theorem 3 and the inverse bidding function (15) immediately yield bid-stage identification of $(\theta_0, \lambda_0^{-1}, F_0)$, with identification of c_0 following through the breakeven condition (4). We thus view the system (19) as a natural starting point for semiparametric inference within the AS-RA model, which given any regular family for C_0 will typically support point identification of all primitives even when U and F are left completely nonparametric.

Several further comments on Condition 2 are worth noting here. First, note that assumptions paralleling Condition 2 are widely maintained in applications; see, for instance, Assumption A1(iv) in CGPV (2011) and virtually all work on nonlinear moment condition models. Condition 2 simply formalizes the intuition that (barring an extraordinarily pathological coincidence) an overdetermined system will have at most one solution.[20] Second, recall that the set of solutions $\hat{\Theta}$ to (19) is identified. Hence in large samples Condition 2 is directly verifiable (e.g. by a simple grid search over Θ). Third, in every simulation we have conducted to date, point identification of θ_0 in fact follows from a much *weaker* condition than Condition 2: at any k, l such that $s_k, s_l > 0$, the set of $\theta \in \Theta$ at which there exists *any* function $\lambda_{kl,\theta}^{-1}$ satisfying (17) – i.e. $\tilde{\Theta}_{kl}$ defined above – has been a singleton.[21] We thus see Condition 2 as both a (very)

[20]Note that for $K \geq 3$ a parametric copula in fact induces a substantially greater degree of overdetermination than arises under parametric utility in CGPV (2011): $K-2$ continuum restrictions in P parameters versus K restrictions in $P+1$ parameters.

[21]Intuitively, this holds because if entry is endogenous then equilibrium bid strategies for different N_k, N_l will typically satisfy single crossing; in particular, for $N_l > N_k$, we will typically have $\beta_l(v|N_l, s_l) < \beta_k(v|N_k, s_k)$ for v close to 0 but $\beta_k(v|N_k, s_k) < \beta_l(v|N_l, s_l)$ for v close to $\bar{v}$. (See Li and Zheng (2009) for a detailed discussion in a risk-neutral context.) In turn, this single crossing property aids identification because for there to exist a function $\lambda_{kl,\theta}^{-1}$ satisfying the compatibility condition (17), we must have $R_k(h_\theta^k(a)) = R_l(h_\theta^l(a))$ for any a such that $b_k(h_\theta^k(a)) = b_l(h_\theta^l(a))$ and vice versa. Furthermore, by construc-

weak regularity condition on the underlying process and a natural starting point for empirical research. In the event that search over Θ suggests multiple solutions in a particular application, one could in principle fall back on inference about the identified set $\hat{\Theta}$. In light of our simulation results, however, we believe it would be difficult (if not impossible) to construct an example where such set inference is required in practice.[22]

5.4 Discussion

To conclude this section, it may be helpful to relate the analysis here to that of GPV (2009). With exogenous entry, GPV (2009) show that bidding behavior at two competition levels n_k, n_l is sufficient to nonparametrically identify the markup function λ_0^{-1}. Our Theorem 3 generalizes this finding to environments with selective entry, showing that for any two competition levels N_k, N_l and any *conjectured* copula C_θ there exists at most one markup function $\lambda_{kl,\theta}^{-1}$ capable of rationalizing observed behavior at N_k, N_l, with this function $\lambda_{kl,\theta}^{-1}$ identified (up to θ) by bidding behavior at N_k, N_l. Proposition 2 extends this observation to accommodate a third competition level N_m, noting that if $\theta = \theta_0$ the candidate markup functions $\lambda_{kl,\theta}^{-1}$ and $\lambda_{km,\theta}^{-1}$ must *coincide*. This in turn leads to a continuum of overidentifying restrictions which under regularity conditions will point identify θ_0. In other words, GPV (2009) establish point identification of the model without entry based on only two competition levels; we use the overidentifying third level to pin down the parameters θ governing entry. The key to this extension is the quantile mapping (14), which allows us to recast the problem with selection in a form analogous to GPV (2009). To our knowledge, this insight is novel in the literature.

It is also instructive to compare the analysis here with that of GL (2014). Working

tion, if b_k, b_l and R_k, R_l cross at any point, they must do so in opposite directions. Finally, changing θ always shifts a given pair b_k, R_k in parallel. Hence if (as in our examples) bid functions cross exactly once, at least one of b_k, b_l or R_k, R_l will cross for any $\theta \in \Theta$, and so long as at least one of b_k, b_l or R_k, R_l cross then both equations above must bind. This in turn induces a system of equations whose unknowns are the parameters θ_0 plus the unknown crossing point $\bar{a}_{kl}$, which in practice is almost always sufficient to pin down θ_0.

[22]Indeed, one goal of this study was initially to develop a semiparametric inference procedure robust to set identification of θ_0. After working through the analysis above, however, we ultimately concluded that such a procedure would be fundamentally uninteresting – at least in the setting considered here, we believe it would be all but impossible to construct an example in which set inference would be required.

within the AS entry framework but assuming risk-neutral bidders, GL (2014) show that exogenous variation in N yields only partial identification of the copula dimension of the model. They propose a parametric copula as one way this challenge could be addressed. We have shown that this restriction is in fact far more powerful than initially envisioned by GL (2014): it can restore point identification not only under risk neutrality but under arbitrary nonparametric utility as well. This in turn opens the door to empirical analysis unifying risk aversion and entry in ways to date unprecedented.

Finally, recall that a parametric form for C_0 in fact leads to a set of continuum restrictions on the unknown vector θ_0. This suggests at least two intriguing possibilities not explored in detail here. First, in cases where θ_0 is finite-dimensional, access to a large set of overidentifying restrictions suggests that the form of C_0 is testable. Second, given access to a *continuum* of overidentifying restrictions, one may in fact be able to relax the assumption of finite θ_0: by, for instance, taking C_0 to be an Archimedean copula with θ_0 the unknown nonparametric generator function for C_0. We leave both extensions for future research.

6 A simple OLS-CRRA estimator

To illustrate how the formal identification results above map into practical strategies for structural estimation, we next outline a simple semiparametric estimator for auctions with both risk-averse bidders and selective entry. As above, we consider estimation based on a sample of independent auctions identical up to potential competition N, with N excludable in the sense of Assumption 1. Following Section 5.3, we further assume that C_0 belongs to a known parametric family: i.e. estimation maintaining Assumption 5 above. As we show in Section 5.3 this structure alone is sufficient for semiparametric identification of remaining primitives. In practice, however, we expect additional structure on U to substantially improve estimation performance in finite samples. For the moment, therefore, we further assume that bidder preferences exhibit constant relative risk aversion (CRRA) with risk-aversion parameter ρ_0:

Assumption 6 (CRRA utility)**.** $U(x) = x^{1-\rho_0}$, with $c_0 = w_0$ for all bidders.

The normalization $c_0 = w_0$ ensures that post-entry bidding behavior is well-described by a standard CRRA bidding model. This parallels the structure typically maintained in applications without entry, e.g. CGPV (2011). We consider estimation with a parametric copula but fully nonparametric utility in Appendix C.

6.1 Nonparametric estimation of s_k, b_k, R_k

Given a sample of auctions $\ell = 1, ..., L$, we estimate s_k via a simple sample average:

$$\hat{s}_k = 1 - \frac{1}{N_k} \sum_{\ell=1}^{L} n_\ell.$$

To obtain first-step estimates of b_k, R_k, we first apply the local polynomial quantile estimator of Fan, Li, and Pesendorfer (2015) (henceforth FLP) to obtain estimates $\hat{b}_k(\cdot), \hat{b}'_k(\cdot)$ for $b_k(\cdot), b'_k(\cdot)$ respectively. Using the identity $b'_k(a) = 1/g_k(b_k(a))$, we then plug in estimators $\hat{s}_k$, $\hat{b}_k(\cdot)$, and $\hat{b}'_k(\cdot)$ to obtain a consistent first-step estimator for R_k:

$$\hat{R}_k(a) = \frac{\hat{s}_k + (1 - \hat{s}_k)a}{(N-1)(1 - \hat{s}_k)} \hat{b}'_k(a).$$

In implementing the FLP procedure, we employ a triweight kernel specification:

$$\kappa(x) = \frac{35}{32}(1 - x^2)^3 \mathbf{1}[|x| \leq 1].$$

We default to bandwidths of the form $h(a) = \text{bwc} \cdot S^{\frac{1}{3}}$, where bwc is a scaling constant and S is the number of observations in the relevant bid sample. For moderate to large sample sizes, the resulting estimator performs well even at the boundaries so long as the model is not greatly over-smoothed. One could in principle perform more robust boundary correction as in Hickman and Hubbard (2014); we leave this as an extension.

6.2 Minimum distance estimation of (ρ_0, θ_0)

Under Assumption 6, it is straightforward to show that $\lambda_0^{-1}(x)$ is linear in x:

$$\lambda_0(x) = \frac{u(x)}{u'(x)} = \frac{x}{1-\rho_0},$$

and therefore $\lambda_0^{-1}(x) = (1-\rho_0)x$. For any $k, l \in \mathcal{K}$, the candidate compatibility condition (17) therefore simplifies to:

$$\tilde{b}_{k,\theta_0}(a) + (1-\rho_0)\tilde{R}_{k,\theta_0}(a) = \tilde{b}_{l,\theta_0}(a) + (1-\rho_0)\tilde{R}_{l,\theta_0}(a), \tag{20}$$

where $\tilde{b}_{k,\theta}(a) \equiv \hat{b}_k(h_\theta^k(a))$, $\tilde{R}_{k,\theta}(a) \equiv \hat{R}_k(h_\theta^k(a))$, and as defined above

$$h_\theta^k(a) \equiv \frac{a - C_\theta(a, \hat{s}_k)}{1-\hat{s}_k}.$$

Minimum distance criterion Since Equation (20) must hold for all $k, l \in \{1, ..., K\}$, we can in principle form a wide range of estimation criteria based on (20). As a baseline, we here simply average the left-hand side of (20) across l:

$$\begin{aligned}
\tilde{b}_{k,\theta_0}(a) + (1-\rho_0)\tilde{R}_{k,\theta_0}(a) &= \tilde{b}_{1,\theta_0}(a) + (1-\rho_0)\tilde{R}_{1,\theta_0}(a), \\
&\vdots \\
\tilde{b}_{k,\theta_0}(a) + (1-\rho_0)\tilde{R}_{k,\theta_0}(a) &= \tilde{b}_{K,\theta_0}(a) + (1-\rho_0)\tilde{R}_{K,\theta_0}(a),
\end{aligned}$$

and therefore

$$\Delta\tilde{b}_{k,\theta_0}(a) + (1-\rho_0)\Delta\tilde{R}_{k,\theta_0}(a) = 0, \tag{21}$$

where $\Delta\tilde{b}_{k,\theta_0}(a)$ and $\Delta\tilde{R}_{k,\theta_0}(a)$ denote differences between $\tilde{b}_{k,\theta_0}(a)$, $\tilde{R}_{k,\theta_0}(a)$ and their respective means across k:

$$\Delta\tilde{b}_{k,\theta_0}(a) = \tilde{b}_{k,\theta_0}(a) - \frac{1}{K}\sum_{l=1}^{K}\tilde{b}_{l,\theta_0}(a),$$
$$\Delta\tilde{R}_{k,\theta_0}(a) = \tilde{R}_{k,\theta_0}(a) - \frac{1}{K}\sum_{l=1}^{K}\tilde{R}_{l,\theta_0}(a).$$

Squaring both sides of Equation (21), integrating across a, and summing across k, we obtain a minimum-squared-error criterion for the unknown parameter (θ_0, ρ_0):

$$Q(\rho_0, \theta_0) = \sum_{k=1}^{K}\int_0^1 (\Delta\tilde{b}_{k,\theta_0}(a) + (1-\rho_0)\Delta\tilde{R}_{k,\theta_0}(a))^2\, da. \tag{22}$$

It would of course be straightforward to extend (22) to accommodate weights differing by k or a. For now, however, we take the unweighted criterion (22) as a baseline.

OLS implementation To implement minimum distance estimation based on this intuition, we first specify a discrete grid $A \subset [0,1]$ on which to approximate each integral in the minimum-distance criterion (22). This yields the discrete approximation

$$\tilde{Q}_A(\rho_0, \theta_0) = \sum_{k\in\mathcal{K}}\sum_{a\in A}\left(\Delta\tilde{b}_{k,\theta}(a) + (1-\rho)\Delta\tilde{R}_{k,\theta}(a)\right)^2. \tag{23}$$

Note that since the identity $\Delta\tilde{b}_{k,\theta_0}(a) + (1-\rho_0)\Delta\tilde{R}_{k,\theta_0}(a) = 0$ must hold pointwise, this discretized criterion is in fact a valid basis for estimation in its own right (i.e. we need not require the grid to become arbitrarily fine). Further, for any θ, the parameter $\rho(\theta)$ minimizing (23) given θ takes a simple OLS form:

$$\rho(\theta) = 1 - \left(\sum_{k\in\mathcal{K}}\sum_{a\in A}\Delta\tilde{R}_{k,\theta}(a)^2\right)^{-1}\left(\sum_{k\in\mathcal{K}}\sum_{a\in A}-\Delta\tilde{R}_{k,\theta}(a)\Delta\tilde{b}_{k,\theta}(a)\right).$$

This in turn suggests a potential nested minimization algorithm. First, given θ, obtain $\rho(\theta)$ via OLS of $\tilde{b}_{k,\theta}(a)$ on $-\tilde{R}_{k,\theta}(a)$ and return the value of the criterion (23). Then, in the outer loop, search over θ to find the overall minimum of (23). This algorithm is fast, stable, and supports both derivative- and grid-based search over θ. Unfortunately, in practice it may also be sensitive to noise in first-step estimates for $\Delta \tilde{R}_{k,\theta}(a)$ and $\Delta \tilde{b}_{k,\theta}(a)$. While this measurement error is clearly not classical, Monte Carlo analysis suggests it has similar effects in terms of attenuating inner-loop OLS estimates.

As one way to ameliorate bias due to first-step noise, note that first-step estimates of $b_k'(\cdot)$ will typically contain substantially more noise than first-step estimates of $b_k(\cdot)$. Furthermore, while the analogy is imperfect, intuition from classical measurement error suggests that noise in the OLS dependent variable will be much less problematic than noise in the OLS regressor. We therefore ultimately implement estimation based on the following (identification-equivalent) transformation of (23):

$$\tilde{Q}(\rho, \theta) = \sum_{k \in \mathcal{K}} \sum_{a \in A} (\Delta \tilde{R}_{k,\theta}(a) + (1-\rho)^{-1} \Delta \tilde{b}_{k,\theta}(a))^2. \tag{24}$$

The inner loop of our final algorithm thus involves OLS regression of $\Delta \tilde{R}_{k,\theta}(a)$ on $-\Delta \tilde{b}_{k,\theta}(a)$, with the coefficient on $\Delta \tilde{b}_{k,\theta}(a)$ interpreted as $(1-\rho(\theta))^{-1}$. In the outer loop, we substitute $(1-\rho(\theta))^{-1}$ into (24) and minimize the resulting objective over θ. This reframed algorithm retains the speed and stability benefits outlined above, but Monte Carlo analysis suggests that the reframed inner loop dramatically reduces bias in finite samples. Note further that attenuation in the reframed problem tends to bias $(1-\rho(\theta))^{-1}$ downward and therefore $\rho(\theta)$ toward zero (risk-neutrality). Given that some first-step bias is inevitable, this seems the appropriate direction in which to err.

7 Monte Carlo experiments

Finally, we explore a series of Monte Carlo experiments designed to evaluate the performance of the simple OLS-CRRA estimator outlined in Section 6. Toward this end, we adopt the following model specification. Bidders have CRRA utility with initial wealth equal to entry cost: $u_0(x) = x^{1-\rho_0}$ and $c_0 = w_0$. Dependence between signals and values is characterized by a Gumbel copula with dependence parameter θ_0:

$$C_0(F, s) = \psi^{-1}(\psi(F; \theta_0) + \psi(s; \theta_0); \theta_0),$$

where

$$\psi(u; \theta) = (-\log(u))^{\theta} \text{ for } \theta \in [1, \infty).$$

Valuations are drawn from a truncated normal distribution with mean $\mu_0 = 5$, support $[0, 10]$, and variance parameter σ_0:

$$F_0(v) = \text{TN}(v; 5, \sigma_0, 0, 10).$$

For simplicity, we assume the econometrician has access to a sample of auctions differing only in potential competition, with N varying exogenously on $\mathcal{N} = [2, 4, 6, 8]$. We choose the number of auctions L_k observed at each competition level N_k such that the number of bids observed is approximately constant across k: $L_k \approx \frac{S}{(1-s_k)N_k}$, where S is a pre-specified constant indexing average number of bids observed at each N_k. We vary primitives θ_0, ρ_0, c_0, and σ_0 across experiments as indicated below.

Tables 1-3 report results derived from applying our OLS-CRRA estimator to a baseline data generating process with $\rho_0 = 0.5$, $\theta_0 = 0.5$, $c_0 = 0.5$, $\sigma_0 = 2$, and sample scale S varying on $\{500, 1000, 2000\}$. As above, potential competition N varies exogenously on $\mathcal{N} = \{2, 4, 6, 8\}$, inducing entry thresholds {0.2487, 0.6288, 0.7542, 0.8158} respectively. To reduce contamination from boundary estimates, we take A to be an evenly spaced 200-point grid on the interval $[0.1, 0.9]$.[23]

[Table 1 about here.]

[Table 2 about here.]

[Table 3 about here.]

[23]Note that given our CRRA assumption A need not span the interval $[0, 1]$.

On balance, Tables 1-3 suggest that our simple OLS-CRRA estimator performs very well. Estimates $\hat{\theta}$ for θ_0 are typically quite precise, with no clear bias pattern. As expected, estimates $\hat{\rho}$ for ρ_0 are somewhat biased toward zero, but this bias diminishes substantially as sample size increases. Note that downward bias in $\hat{\rho}$ will tend to lead to under-rejection of risk neutrality; this is in our view the appropriate direction in which to err. Encouragingly, results are relatively stable across different bandwidth sizes, with larger scale factors reducing downward bias in $\hat{\rho}$ and increasing downward bias in $\hat{\theta}$. Moderate scale factors (bwc = 1.5 and bwc = 2.0) seem on net to produce the best estimates across specifications, though our estimators' relatively stable performance (even in small samples) suggests that choice of bandwidth is a second-order concern.

Tables 4-6 report results from Monte Carlo experiments for three further auction processes: selection but no risk aversion ($\rho_0 = 0$, $\theta_0 = 1.5$), risk aversion but no selection ($\rho_0 = 0.5$, $\theta_0 = 1$), and minimal risk aversion with minimal selection ($\rho_0 = 0.2$, $\theta_0 = 1.2$). All are reported for $S = 2000$. As indicated by Tables 4-5, our simple estimator perform very well in DGPs with either no risk aversion ($\rho_0 = 0$) or no selection ($\theta_0 = 1$), suggesting that relatively little is lost by explicitly modeling both factors. As a percentage of ρ_0, bias in $\hat{\rho}$ is somewhat worse in models with little risk aversion. From the model's perspective, however, recall that the relevant parameter is not ρ_0 but $1 - \rho_0$. Applying intuition from measurement error, we expect downward bias in $1 - \hat{\rho}$ to be roughly proportional to $1 - \rho_0$, and this prediction is in fact borne out quite closely in practice. We conclude that our proposed two-step CRRA estimator performs well across a variety of DGPs and model specifications, with this performance relatively insensitive to choice of bandwidth.

[Table 4 about here.]

[Table 5 about here.]

[Table 6 about here.]

8 Conclusion

We study identification and estimation in auctions with both risk averse bidders and selective entry. In the process, we make several core contributions to the related literature. We first develop a flexible structural framework accommodating both risk aversion and selection, building on the recent theoretical contributions of Li, Lu, and Zhao (2014). We then proceed to study identification in this framework when variation in potential competition is excludable, focusing on restrictions generated by the bidding model. We begin with identification under restrictions on utility, deriving a sharp characterization of the set of primitives consistent with bidding behavior corresponding to any given class of candidate utility functions. This system nests the identification restrictions of CGPV (2011), and (as there) is generally sufficient for point identification of utility parameters under parametric restrictions on utility. In contrast to CGPV (2011), however, point identification of utility parameters leads to only partial identification of remaining primitives as in Gentry and Li (2014). We then turn to consider identification under restrictions on the copula linking signals to values, showing that parametric structure on this copula is typically sufficient for point identification even when other primitives are left fully unrestricted. Finally, building on these results, we propose a simple semiparametric estimator combining CRRA utility with a parametric copula, showing that this simple estimator performs very well even in small samples. We thereby provide a formal framework within which to analyze interaction between risk aversion and entry, helping to unify the extensive empirical literatures analyzing these factors in isolation.

While we have here assumed access to a sample of auctions identical up to observables, our identification results can be readily extended to environments with auction-level unobservable heterogeneity following Gentry and Li (2014). Such an extension would allow us to relax the framework described above in two important dimensions. First, rather than assuming independence of signal-value pairs across bidders, we require only independence conditional on the auction-level unobservable. Second, in place of Assumption 1, we require only exclusion conditional on realizations of auction-level unobserved heterogeneity. The first step in the identification argument would be to recover distributions of bidding and entry decisions following Section 4 of Gentry and Li (2014), with the remainder of the analysis proceeding as developed here.

Finally, while we focus here on the case of symmetry, our core insights extend immediately to environments with asymmetric bidders. In fact, so long as rival types do not affect own primitives, we expect asymmetry to improve identification. In particular, if rival types are continuous, then multidimensional variation in rival type vectors could potentially be used to obtain nonparametric identification. The essence of this argument is as follows: suppose there exist two distinct sets of rivals inducing the *same* entry probability by bidder i; say, market structure Z with two similar rivals versus market structure Z' with one strong and one weak. Then the latent distribution of i's private values would be the same at Z and Z', but the distribution of the maximum bid among i's rivals could be distinct. If so, we could obtain identification in two steps: first apply the argument of GPV (2009) to bids at Z and Z' to recover i's utility function, then apply the argument of Gentry and Li (2014) to recover entry-related primitives.[24] At the same time, however, asymmetric types would also induce a substantial complication: in general there may be many type-symmetric equilibria, in which case observed play could involve mixing over these equilibria. This problem is well beyond the scope of the current paper, and we leave it for future research.

Appendix A: Proofs

Proof of Theorem 2. We first establish the "only if" direction: for any $(C, F) \in \mathcal{C} \times \mathcal{F}$, the ex post quantile functions $v_1, v_2, ..., v_K$ satisfy Conditions 1-3 in Theorem 2. Noting that Condition 1 follows directly from the invariant support assumption on C_0, we proceed to establish Conditions 2 and 3.

First consider Condition 2. As a copula, C must be d-increasing: i.e. for any a, a', s, s' such that $1 \geq a' \geq a \geq 0$ and $1 \geq s' \geq s \geq 0$, we must have

$$C(a', s') - C(a, s') - C(a', s) + C(a, s) \geq 0.$$

[24]Meanwhile, in environments with discrete types, nonparametric point identification would remain infeasible, but variation in rival types could still enrich the set of entry variables used to identify the model.

Note that we can rearrange the definition of F_k to obtain the identity:

$$C(a, s_k) = a - [F_k F^{-1}](a)(1 - s_k) \text{ for all } k = 1, ..., K.$$

Hence setting $s' = s_{k+1}$, $s = s_k$ and rearranging the inequality above:

$$\begin{aligned}
& C(a', s_{k+1}) - C(a, s_{k+1}) - C(a', s_k) + C(a, s_k) \geq 0 \\
\Leftrightarrow \quad & C(a', s_{k+1}) - C(a, s_{k+1}) - C(a', s_k) + C(a, s_k) \geq 0 \\
\Leftrightarrow \quad & \frac{\partial C}{\partial a}(a, s_{k+1}) \geq \frac{\partial C}{\partial a}(a, s_k) \\
\Leftrightarrow \quad & 1 - [F_{k+1} F^{-1}]'(a)(1 - s_{k+1}) \geq 1 - [F_k F^{-1}]'(a)(1 - s_k) \\
\Leftrightarrow \quad & (1 - s_k)[F_k F^{-1}]'(a) \geq (1 - s_{k+1})[F_{k+1} F^{-1}]'(a) \text{ for all } a,
\end{aligned}$$

where the third line follows since v_k is differentiable in a for all k and the fourth follows by differentiating the identity $C(a, s_k) = a - [F_k F^{-1}](a)(1 - s_k)$ with respect to a. But note that $[F_k F^{-1}]'(a) = F_k'(F^{-1}(a)) \cdot F^{-1,\prime}(a)$. The final inequality thus holds if and only if $(1 - s_k)F_k' \geq (1 - s_{k+1})F_{k+1}'$. This in turn implies Condition 2.

Finally consider Condition 3. Rearranging the definition of F_k,

$$F_k(y)(1 - s_k) = \int_{s_k}^{1} F(v|t)dt.$$

Now consider the distribution of V_i conditional on the event $S_i \in [s_k, s_{k+1}]$:

$$F(y|S_i \in [s_k, s_{k+1}]) \equiv \frac{1}{s_{k+1} - s_k} \int_{s_k}^{s_{k+1}} F(v|t)dt = \frac{F_k(y)(1 - s_k) - F_{k+1}(y)(1 - s_{k+1})}{s_{k+1} - s_k}.$$

Clearly $1 \geq F(y|S_i \in [s_k, s_{k+1}])$, and by stochastic ordering $F(y|S_i \in [s_k, s_{k+1}])$ is decreasing in k and satisfies $F(y|S_i \in [s_k, s_{k+1}]) \geq F_{k+1}(y)$ for all $k = 1, ..., K - 1$. Hence for all $k = 2, ..., K_1$ we must have

$$1 \geq \frac{F_{k-1}(1 - s_{k-1}) - F_k(1 - s_k)}{s_k - s_{k-1}} \geq \frac{F_k(1 - s_k) - F_{k+1}(1 - s_{k+1})}{s_{k+1} - s_k} \geq F_K \geq 0,$$

i.e. the condition to be shown. Note that this latter property is *stronger* than $F_k \geq F_{k+1}$: rearranging the identities above,

$$\begin{aligned}
(1 - s_k)F_k(y) &= (s_{k+1} - s_k)F(y|S_i \in [s_k, s_{k+1}]) + (1 - s_{k+1})F_{k+1}(y) \\
&\geq (s_{k+1} - s_k)F_{k+1}(y) + (1 - s_{k+1})F_{k+1}(y) \\
&= (1 - s_k)F_{k+1}(y).
\end{aligned}$$

Hence it would be insufficient simply to assume $F_k \geq F_{k+1}$ for all $k < K$.

"If" direction We next establish the "if" direction: for any λ^{-1} such that $\tilde{v}_1, ..., \tilde{v}_K$ satisfy Conditions 1-3 of Theorem 2, there exists a candidate bidding structure (λ^{-1}, C, F) rationalizing bid-stage observables. Since (C, F) influence bidding behavior only through $v_1, ..., v_K$, this is equivalent to showing existence of (C, F) generating $\tilde{v}_1, ..., \tilde{v}_K$.

Toward this end, first define a candidate F for F_0 as follows:

$$F(y) = s_1 \frac{\tilde{v}_1^{-1}(y)(1-s_1) - \tilde{v}_2^{-1}(y)(1-s_2)}{s_2 - s_1} + (1-s_1)\tilde{v}_1^{-1}(y).$$

Recall that $\tilde{v}_1$ and $\tilde{v}_2$ are differentiable, strictly increasing, and satisfy $\tilde{v}_1(0) = \tilde{v}_2(0) = b_1(0) = 0$ by construction, with $\tilde{v}_1(1) = \tilde{v}_2(1)$ by identical support. Hence $F(0) = 0$ and $F(y) = 1$ for $y \geq \tilde{v}_1(1)$. Furthermore, by distribution consistence, we have $v_1^{-1,\prime}(y)(1-s_1) > v_2^{-1,\prime}(y)(1-s_2)$. Thus F is a differentiable, strictly increasing c.d.f. on $[0, \tilde{v}_1(1)]$, hence $F \in \mathcal{F}[0, \tilde{v}_1(1)]$.

Now construct a candidate C for C_0 piecewise in s as follows. For $s < s_1$ set

$$C(a,s) = s\frac{[\tilde{F}_1 F^{-1}](a)(1-s_1) - [\tilde{F}_2 F^{-1}](a)(1-s_2)}{s_2 - s_1},$$

for $s \geq s_K$ set $C(a,s) = a - [\tilde{F}_K F^{-1}](a)(1-s)$, and for each $s \in [s_k, s_{k+1})$ define $C(a,s)$ by linear interpolation (in s) between $a - [\tilde{F}_k F^{-1}](a)(1-s_k)$ and $a - [\tilde{F}_{k+1} F^{-1}](a)(1-s_{k+1})$. Note that $C(0,s) = C(a,0) = 0$, $C(1,s) = s$, $C(a,1) = a$, and

$$\begin{aligned}
\lim_{s \uparrow s_1} C(F(y), s) &= s_1 \frac{\tilde{F}_1(y)(1-s_1) - \tilde{F}_2(y)(1-s_2)}{s_2 - s_1} \\
&= s_1 \frac{\tilde{F}_1(y)(1-s_1) - \tilde{F}_2(y)(1-s_2)}{s_2 - s_1} + (1-s_1)\tilde{F}_1(y) - (1-s_1)\tilde{F}_1(y) \\
&\equiv F(y) - (1-s_1)\tilde{F}_1(y) \\
&\equiv \lim_{s \downarrow s_1} C(F(y), s).
\end{aligned}$$

Thus C is a continuous function satisfying the limit properties of a joint distribution which reproduces the candidates $\tilde{F}_1, ..., \tilde{F}_K$: i.e. such that for each $k = 1, ..., K$ we have

$$\frac{F(y) - C(F(y), s_k)}{1 - s_k} \equiv \tilde{F}_k(y).$$

We now verify that C is d-increasing (hence a joint c.d.f) and implies a conditional distribution $F(y|s)$ satisfying stochastic ordering.

First show that C is d-increasing. Since C is continuous, it is sufficient to restrict attention to rectangles such that either $s, s' \leq s_1$, $s, s' \geq s_K$, or $s, s' \in [s_k, s_{k+1}]$. If $s, s' \leq s_1$, then $\frac{\partial C}{\partial a \partial s} = 1$, and if $s, s' \geq s_K$, then $\frac{\partial C}{\partial a \partial s} = [\tilde{F}_K F^{-1}]'(a) \geq 0$. Hence C is d-increasing for rectangles in both regions. Finally, if $s, s' \in [s_k, s_{k+1}]$ then $C(a, \cdot)$ is a linear interpolation in s between $C(a, s_k) \equiv a - [\tilde{F}_k F^{-1}](a)(1-s_k)$ and $C(a, s_{k+1}) \equiv a - [\tilde{F}_{k+1} F^{-1}](a)(1-s_{k+1})$. Hence it is sufficient to verify that

$$C(a', s_{k+1}) - C(a, s_{k+1}) - C(a', s_k) + C(a, s_k) \geq 0.$$

But reversing the arguments used in the "only if" direction above, this is equivalent to Condition 2 of Theorem 2. Hence C is d-increasing.

Next show that C satisfies stochastic ordering. Recall that $F(y|s) = \frac{\partial C(F(y),s)}{\partial s}$ where this derivative exists; we fill in endpoints in the construction above by taking right derivatives where necessary. We wish to show that $F(y|s)$ is decreasing in s.

Toward this end, first consider $s < s_1$. By definition of C, we then have

$$F(y|s) = \frac{\tilde{v}_1^{-1}(y)(1-s_1) - \tilde{v}_2^{-1}(y)(1-s_2)}{s_2 - s_1}.$$

Now consider $s \in [s_k, s_{k+1})$. For such s we have $C(a, \cdot)$ a linear interpolation (in s) between $C(a, s_k) \equiv a - [\tilde{F}_k F^{-1}](a)(1-s_k)$ and $C(a, s_{k+1}) \equiv a - [\tilde{F}_{k+1} F^{-1}](a)(1-s_{k+1})$. Hence

$$F(y|s) = \frac{C(F(y), s_{k+1}) - C(F(y), s_k)}{s_{k+1} - s_k} = \frac{\tilde{F}_k(y)(1-s_k) - \tilde{F}_{k+1}(y)(1-s_{k+1})}{s_{k+1} - s_k}.$$

Finally consider $s \geq s_K$. Then by construction we have

$$F(y|s) = \frac{\partial}{\partial s}[F(y) - \tilde{F}_K(y)(1-s)] = \tilde{F}_K(y).$$

Hence $F(y|s)$ is decreasing in s if and only if for all $k = 1, ..., K-1$.

$$\frac{F_{k-1}(1-s_{k-1}) - F_k(1-s_k)}{s_k - s_{k-1}} \geq \frac{F_k(1-s_k) - F_{k+1}(1-s_{k+1})}{s_{k+1} - s_k} \geq F_K.$$

But is precisely Condition 3 of Theorem 2. Hence C satisfies stochastic ordering.

We have thus constructed a candidate structure (C, F) reproducing $\tilde{v}_1, ..., \tilde{v}_K$ and satisfying all properties on (C_0, F_0) except twice differentiability of C_0. By slightly perturbing $C(a, s)$ to smooth transitions at each $s_1, ..., s_K$, one can construct a copula satisfying the properties above plus the twice differentiability; i.e. all conditions required by $\mathcal{C}$. This establishes the claim. □

Proof of Proposition 1. Assumption 1 implies a unique solution to (10) in the parameter space Γ, hence identification of γ_0, λ_0^{-1}, and U_0 under Assumption 4.

Identification of λ_0^{-1} implies identification of $v_1, ..., v_K$ through the quantile inverse bidding function (6), therefore identification of ex post value distributions $F_1, ..., F_K$. While knowledge of $F_1, ..., F_K$ is insufficient to identify $C(\cdot)$ and F_0, Theorem 3 of Gentry and Li (2014) can be applied to yield identified bounds F^+ and F^- on the conditional distribution $F(v|s)$. This establishes the second part of Proposition 1.

Finally, we know from Theorem 1 that at each $k \in \mathcal{K}$ the entry threshold s_k must satisfy the relationship

$$U_0(c_0) = \int_r^{\bar{v}} U_0(v - \beta(v|N_k, s_k)) \cdot \Psi(v|N_k, s_k)\, dF(v|s_k).$$

But from the inverse bidding function (6) we have $\lambda_0^{-1}(R_k(a)) = v_k(a) - b_k(a)$, or equivalently $\lambda_0^{-1}(R_k(F_k(v))) = v - \beta(v|N_k, s_k)$. Therefore

$$U_0(c_0) = \int_r^{\bar{v}} U_0(\lambda_0^{-1}(R_k(F_k(v)))) \cdot \Psi(v|N_k, s_k)\, dF(v|s_k). \tag{25}$$

From above, U_0, λ_0^{-1}, F_k, R_k, and $\Psi(\cdot|N_k, s_k)$ are identified, so the only unknown on the right-hand side is $F(\cdot|s_k)$. But from Theorem 3 in Gentry and Li (2014) we obtain identified distributions $F^+(\cdot|s_k), F^-(\cdot|s_k)$ bounding $F(\cdot|s_k)$, and by the arguments underlying Theorem 1 we know $U_0(\lambda_0^{-1}(R_k(F_k(v)))) \cdot \Psi(v|N_k, s_k)$ is increasing in v. Hence (25) represents the expectation of an identified, increasing function with respect to the unknown distribution $F(\cdot|s_k)$. Substituting the identified, stochastically ordered distributions $F^+(\cdot|s_k)$ and $F^-(\cdot|s_k)$ for $F(\cdot|s_k)$ in (25) thus yields identified constants U_k^+ and U_k^- such that $U_k^+ \geq U_0(c_0) \geq U_k^-$. Monotonicity of U_0 establishes the claim. $\square$

Proof of Theorem 3. Consider any k, l and θ satisfying the hypotheses of Theorem 3. For each $k = 1, ..., K$, let $\tilde{b}_{k,\theta}(\cdot)$, $\tilde{R}_{k,\theta}(\cdot)$ be the functions obtained when $b_k(\cdot)$, $R_k(\cdot)$ are reindexed according to $h_\theta^k(\cdot)$:

$$\tilde{b}_{k,\theta}(a) \equiv b_k(h_\theta^k(a)),$$
$$\tilde{R}_{k,\theta}(a) \equiv R_k(h_\theta^k(a)).$$

Define $\bar{r}_k \equiv \sup_a \tilde{R}_{k,\theta}(a)$, $\bar{r}_{kl} \equiv \max\{\bar{r}_k, \bar{r}_l\}$ as in the main text, and let functions $\bar{R}_{kl,\theta}(\cdot)$, $\underline{R}_{kl,\theta}(\cdot)$ be the pointwise maximum and minimum of $\tilde{R}_{k,\theta}(\cdot)$, $\tilde{R}_{l,\theta}(\cdot)$ respectively:

$$\bar{R}_{kl,\theta}(a) \equiv \max\{\tilde{R}_{k,\theta}(a), \tilde{R}_{l,\theta}(a)\}, \tag{26}$$
$$\underline{R}_{kl,\theta}(a) \equiv \min\{\tilde{R}_{k,\theta}(a), \tilde{R}_{l,\theta}(a)\}. \tag{27}$$

Finally, for each $r \in [0, \bar{r}_{kl}]$, let $\mathcal{A}_{kl,\theta}(r)$ be the set all decreasing sequences $\{\alpha^t\}_{t=1}^\infty$ satisfying the recursive relationship

$$\bar{R}_{kl,\theta}(\alpha^0) \equiv r, \quad \bar{R}_{kl,\theta}(\alpha^t) = \underline{R}_{kl,\theta}(\alpha^{t-1}) \text{ for } t = 1, 2, ... \tag{28}$$

and define the set $A_{kl,\theta}$ as follows:

$$A_{kl,\theta} \equiv \{a \in [0,1] : \tilde{R}_{k,\theta}(a) = \tilde{R}_{l,\theta}(a)\}.$$

Note the following properties of $\mathcal{A}_{kl,\theta}(r)$:

Lemma 2. *For any $k, l \in \{1, ..., K\}$ and any $r \in [0, \bar{r}_{kl}]$, $\mathcal{A}_{kl,\theta}(r)$ is nonempty. Furthermore, for all sequences $\{\alpha^t\}_{t=1}^\infty \in \mathcal{A}_{kl,\theta}(r)$, $\lim_{t\to\infty} \alpha^t \in A_{kl,\theta}$.*

Proof. First show that $0 \in A_{kl,\theta}$. By Theorem 1, we have $\beta(0|N_k, \bar{s}) = 0$ for all $\bar{s} \in [0, 1)$ and $k \in \mathcal{K}$. Hence in any equilibrium $R_k(0) = 0$ for all $k \in \mathcal{K}$. Furthermore, for any $k, l \in \mathcal{K}$ and any θ, we have $h_{k,\theta}(0) = h_{l,\theta}(0) = 0$ and therefore $0 \in A_{kl,\theta}$.

Next, following GPV (2009), observe that both $\bar{R}_{kl,\theta}$ and $\underline{R}_{kl,\theta}$ are continuous, with $\bar{R}_{kl,\theta}$ having range $[0, \bar{r}_{kl}]$. Choose any $r_0 \in [0, \bar{r}_{kl}]$. Since $r_0 \in [0, \bar{r}_{kl}]$, by the Intermediate Value Theorem there exists $\alpha \in [0, 1]$ such that $\bar{R}_{kl,\theta} = \alpha$. Choose any such α, set $\alpha_0 = \alpha$, and set $r_1 = \underline{R}_{kl,\theta}(\alpha_0)$. Note that $\bar{R}_{kl,\theta}$ is continuous on $[0, \alpha_0]$, with $\bar{R}_{kl,\theta}(\alpha_0) \geq r_1$. Hence again by the intermediate value theorem there exists $\alpha_1 \in [0, \bar{\alpha_0}]$ such that $\bar{R}_{kl,\theta}(\alpha_1) = r_1$ and $\alpha_1 \leq \alpha_0$. Iterating the argument establishes existence of a decreasing sequence $\{\alpha_t\}_{t=0}^\infty \in \mathcal{A}_{kl,\theta}$.

Finally show that any sequence $\{\alpha_t\}_{t=0}^\infty \in \mathcal{A}_{kl,\theta}$ converges to a limit $\bar{a} \in A_{kl,\theta}$. Clearly, if $\{\alpha_t\}_{t=0}^\infty \in \mathcal{A}_{kl,\theta}$ then $\{\alpha_t\}_{t=0}^\infty$ is a decreasing sequence bounded below by 0. Hence $\{\alpha_t\}_{t=0}^\infty$ converges to some limit $\bar{a}$. Furthermore, by definition, we must have $\lim_{t\to\infty} \bar{R}_{kl,\theta}(\alpha_t) = \lim_{t\to\infty} \underline{R}_{kl,\theta}(\alpha_t)$. Hence $\bar{a} \in A_{kl,\theta}$, establishing the claim.

$\square$

Now let ϕ be any continuous, increasing, zero-at origin function on $[0, \bar{r}_{kl}]$ satisfying the compatibility condition

$$\tilde{b}_{k,\theta}(a) + \phi(\tilde{R}_{k,\theta}(a)) = \tilde{b}_{l,\theta}(a) + \phi(\tilde{R}_{l,\theta}(a))\, \forall a \in [0,1]. \tag{29}$$

If no such ϕ exists, the theorem is true by construction. Otherwise, choose any such ϕ and rearrange (29) to obtain for any $a \in [0,1]$

$$\phi(\tilde{R}_{k,\theta}(a)) - \phi(\tilde{R}_{l,\theta}(a)) = \tilde{b}_{l,\theta}(a) - \tilde{b}_{k,\theta}(a).$$

ϕ is continuous, increasing, and satisfies (29), this expression in turn implies

$$\phi(\bar{R}_{kl,\theta}(a)) = |\tilde{b}_{k,\theta}(a) - \tilde{b}_{l,\theta}(a)| + \phi(\underline{R}_{kl,\theta}(a)).$$

Now choose $r \in [0, \bar{r}_{kl}]$, and let $\{\alpha^t\}_{t=0}^{\infty}$ be any element of $\mathcal{A}_{kl,\theta}(r)$. Recall that by definition $\{\alpha^t\}_{t=0}^{\infty}$ satisfies $\bar{R}_{kl,\theta}(\alpha^{t+1}) = \underline{R}_{kl,\theta}(\alpha^t)$ for all t. Thus for any t

$$\phi(\bar{R}_{kl,\theta}(\alpha^t)) = |\tilde{b}_{k,\theta}(\alpha^t) - \tilde{b}_{l,\theta}(\alpha^t)| + \phi(\bar{R}_{kl,\theta}(\alpha^{t+1}). \tag{30}$$

Noting that $r \equiv \bar{R}_{kl,\theta}(\alpha^0)$ and recursively substituting into (30), we therefore conclude

$$\phi(r) = \sum_{t=0}^{\infty} |\tilde{b}_{k,\theta}(\alpha^t) - \tilde{b}_{l,\theta}(\alpha^t)| + \lim_{t\to\infty} \phi(\bar{R}_{kl,\theta}(\alpha^t)). \tag{31}$$

Under our maintained assumptions $A_{kl,\theta}$ is a finite set containing zero. If $A_{kl,\theta} = \{0\}$, then $\alpha^t \to 0$ for any $\{\alpha^t\}_{t=0}^{\infty} \in \mathcal{A}_{kl,\theta}$, the final term vanishes, and we have

$$\phi(r) = \sum_{t=0}^{\infty} |\tilde{b}_{k,\theta}(\alpha^t) - \tilde{b}_{l,\theta}(\alpha^{t+1})|$$

for all $r \in [0, \bar{r}_{kl}]$. Otherwise, let $\{0, a_1, a_2, ..., a_M\}$ index the elements of $A_{kl,\theta}$, and partition the interval $[0, \bar{r}_{kl}]$ into subintervals as follows:

$$\begin{aligned}
r \in [0, r_1) &\quad \text{if} \quad \inf\{\alpha \in [0,1] | r = \bar{R}_{kl,\theta}(\alpha)\} \in [0, a_1);\\
r \in [r_1, r_2) &\quad \text{if} \quad \inf\{\alpha \in [0,1] | r = \bar{R}_{kl,\theta}(\alpha)\} \in [a_1, a_2);\\
&\quad \vdots\\
r \in [r_{M-1}, r_M) &\quad \text{if} \quad \inf\{\alpha \in [0,1] | r = \bar{R}_{kl,\theta}(\alpha)\} \in [a_{M-1}, a_M);\\
r \in [r_M, \bar{r}_{kl}] &\quad \text{if} \quad \inf\{\alpha \in [0,1] | r = \bar{R}_{kl,\theta}(\alpha)\} \in [a_M, 1].
\end{aligned}$$

Note that under finite intersection $r_M > \cdots > r_1 > 0$. If $r \in [0, r_1)$, then by Lemma 2 $\alpha^t \to 0$. Hence $\phi(r)$ is identified on $[0, r_1)$ as above, and by continuity of $\phi(\cdot)$ we can take limits to extend this definition to establish uniqueness on $[0, r_1]$. Now consider (r_1, r_2). Then by Lemma 2 we must have either $\alpha^t \to 0$ or $\alpha^t \to r_1$. In the former case we can express $\phi(r)$ as above. In the latter case, we can express $\phi(r)$ as

$$\phi(r) = \sum_{t=0}^{\infty} |\tilde{b}_{k,\theta}(\alpha^t) - \tilde{b}_{l,\theta}(\alpha^t)| + \phi(r_1),$$

with $\phi(r_1)$ identified by limits of sums on $[0, r_1)$. In either case $\phi(\cdot)$ is identified on $[0, r_2)$, which again invoking continuity extends immediately to identification on $[0, r_2]$. Iterating these arguments as necessary (with $r \in [r_2, r_3)$ implying α^t convergent to either $0, r_1$ or r_2 and so forth), we can express $\phi(\cdot)$ in terms of identified objects on the whole interval $[0, \bar{r}_{kl}]$.

Finally, observe that if $\phi(\cdot)$ satisfies (29) than any construction satisfying the rules above must return the *same* function $\phi(\cdot)$; i.e. the relationships above must hold for any selection $\alpha(\cdot)$ from $\mathcal{A}_{kl,\theta}(\cdot)$ (or else we would obtain a contradiction). But given any selection $\alpha(\cdot)$ from $\mathcal{A}_{kl,\theta}(\cdot)$ ϕ can be represented as an identified transformation of observables. Thus ϕ is unique and identified. Hence ϕ is identified. Furthermore, since the selection $\alpha(\cdot)$ is arbitrary, one could select $\alpha(\cdot)$ according to a known rule: e.g. fastest "greedy descent". This would yield constructive identification of ϕ in terms of observables, establishing the claim.

□

Proof of Lemma 1. Consider any distinct $k, l \in \mathcal{K}$. By definition, $v_0(a) = v_k(h^k_{\theta_0}(a)) = v_l(h^l_{\theta_0}(a))$, so $v_k(a) = v_0(h^{k,-1}_{\theta_0}(a))$ and $v_l(a) = v_0(h^{l,-1}_{\theta_0}(a))$. Thus if $h^{k,-1}_{\theta_0}(h^k_\theta(a)) \neq h^{l,-1}_{\theta_0}(h^l_\theta(a))$ then $v_k(h^k_\theta(\cdot)) = v_0(h^{k,-1}_{\theta_0}(h^k_\theta(\cdot))) \neq v_0(h^{l,-1}_{\theta_0}(h^l_\theta(a))) = v_l(h^l_\theta(a))$ and vice versa. But by regularity, if $\theta \neq \theta_0$ then $h^{k,-1}_{\theta'}(h^k_\theta(a)) \neq h^{l,-1}_{\theta'}(h^l_\theta(a))$ on a set of positive measure. This establishes the claim. □

Appendix B: Nonparametric bid-stage identified set

As one way to quantify nonparametric information induced by exogenous variation in competition, this appendix explores the sharp bounds on λ_0^{-1} induced by the characterization of bid-stage restrictions given in Theorem 2. Toward this end, we implement a variant of the algorithm sketched in Section 4: first restrict attention to a class of finite-dimensional sieve spaces approximating Λ^{-1}, then numerically analyze the set of functions ϕ consistent with Conditions (1)-(3) of Theorem 2 within a given element of this space.

Specifics of our procedure are as follows. We approximate Λ^{-1} with a variant of the shape-constrained Bernstein polynomial sieve space considered by Zincenko (2012, working paper). Let $\bar{r} = \max\{\bar{r}_1, ..., \bar{r}_K\}$ be the upper limit on the empirically relevant domain of λ^{-1} and $p_{J,j}(x)$ denote the Bernstein polynomial basis on the interval $[0, \bar{r}]$:

$$p_{J,j}(r) \equiv \binom{J}{j} \left(\frac{r}{\bar{r}}\right)^j \left(1 - \frac{r}{\bar{r}}\right)^{J-j}.$$

Let $\{J^n\}_{n \in \mathbb{N}}$ be an increasing divergent sequence of positive integers, and $\mathcal{P}^{(n)}$ be the space of degree-P^n Bernstein polynomials defined as follows.

Definition 7. Let $\Gamma^{(n)} \subset \mathbb{R}^{H_n+1}$ be the set of vectors $\gamma = \{\gamma_j : j = 0, 1, ..., H_n\} \in \mathbb{R}^{H_n+1}$ such that $\gamma_0 = 0$ and

$$\bar{r} J_n^{-2} \leq \gamma_{j+1} - \gamma_j \leq \bar{r} J_n^{-1} \text{ for } 0 \leq j \leq J_n - 1.$$

Let $\mathcal{P}^{(n)}$ be the set of Bernstein polynomials $P : [0, \bar{r}] \to \mathbb{R}_+$ of degree J_n generated by coefficient vectors $\gamma \in \Gamma^{(n)}$: i.e. the set of functions $P(\cdot)$ such that

$$P(r) = \sum_{j=0}^{J_n} \gamma_j p_{J_n,j}(r) \text{ for some } \gamma \in \Gamma^{(n)}.$$

We refer to $\mathcal{P}^{(1)}, \mathcal{P}^{(2)}, ...$ as *shape-constrained Bernstein sieves* for Λ^{-1}.

As noted by Zincenko (2012, working paper), we have $\mathcal{P}^{(n)} \subset \Lambda^{-1}$ for each n, with $\mathcal{P}^{(n)}$ becoming dense in Λ^{-1} (with respect to the L_1 norm) as $n \to \infty$. The shape-constrained Bernstein sieves $\mathcal{P}^{(1)}, \mathcal{P}^{(2)}, ...$ thus provide a natural space within which to analyze restrictions generated by the bidding model. Now fix $n \in \mathbb{N}$, and consider the set of $P \in \mathcal{P}^{(n)}$ satisfying Conditions 1-3 of Theorem 2. In practice, we approximate this set as follows.

Approximation algorithm Starting from a known data generating process, fix a grid A of points in $[0, 1]$ and for each $k \in \mathcal{K}$ compute equilibrium bids $b_k(a)$ and markup functions $R_k(a)$ corresponding to each $a \in A$. For each $a \in A$, let $\mathbf{p}^k_{J_n}(a)$ be the $1 \times (J_n + 1)$ vector of Bernstein polynomial basis functions evaluated at $r = R_k(a)$:

$$\mathbf{p}^{(n)}_k(a) \equiv (p_{J_n,0}(R_k(a)), ..., p_{J_n,J_n}(R_k(a))).$$

Note that for purposes of characterizing the identified set we may take $\mathbf{p}^{(n)}_k(a)$ as known and fixed across iterations.

For each $k \in \mathcal{K}$, let $\tilde{v}_k(\cdot; \gamma)$ be the ex post quantile function implied by $\gamma \in \Gamma^{(n)}$:

$$\tilde{v}_k(a; \gamma) = b_k(a) + \mathbf{p}^k_{J_n}(a) \cdot \gamma. \tag{32}$$

Now consider Condition 1 of Theorem 2:

$$\tilde{v}_1(1; \gamma) = \tilde{v}_2(1; \gamma) = \cdots = \tilde{v}_K(1; \gamma). \tag{33}$$

Let $\bar{b}(a) = \frac{1}{K} \sum_{k=1}^{K} b_k(a)$ and $\bar{\mathbf{p}}^{(n)}(a) = \frac{1}{K} \sum_{k=1}^{K} \mathbf{p}^{(n)}(a)$ denote means of $b_k(a)$ and $\mathbf{p}^{(n)}_k(a)$ across $k \in \mathcal{K}$ respectively, and define:

$$\Delta\mathbf{P}^{(n)}(1) = \begin{bmatrix} \mathbf{p}^{(n)}_2(1) - \bar{\mathbf{p}}^{(n)}(1) \\ \mathbf{p}^{(n)}_3(1) - \bar{\mathbf{p}}^{(n)}(1) \\ \vdots \\ \mathbf{p}^{(n)}_K(1) - \bar{\mathbf{p}}^{(n)}(1). \end{bmatrix}, \qquad \Delta\mathbf{b}(1) = \begin{bmatrix} b_1(1) - \bar{b}(1) \\ b_2(1) - \bar{b}(1) \\ \vdots \\ b_K(1) - \bar{b}(1) \end{bmatrix}.$$

Substituting from (32) and stacking across k, the restriction (33) is then equivalent to

$$\Delta\mathbf{b}(1) = -\Delta\mathbf{P}^{(n)}(1) \cdot \gamma, \tag{34}$$

This in turn corresponds to a system of $K - 1$ distinct linear restrictions on the $J_n \times 1$ vector γ. Since we are here interested in high-order approximations to Λ^{-1} (i.e. $J_n >> K - 1$), this system will not uniquely determine γ. Rather, we exploit the linear structure of (34) to simplify approximation of the identified set.

In particular, let $\tilde{\Gamma}^{(n)}$ be the set of elements $\gamma \in \Gamma^{(n)}$ such that (34) holds for all k. We then proceed via the following two-step algorithm:

Step 1 Obtain a large random sample of coefficients γ^s from the feasible set $\tilde{\Gamma}^{(n)}$.

Step 2 For each coefficient γ^s drawn in Step 1, evaluate Conditions 2 and 3 of Theorem 2 implied by γ^s. If Conditions 2 and 3 are (approximately) satisfied at γ^s, keep γ^s as a feasible parameter; else reject γ^s.

We now describe details of each step.

Step 1 By definition, $\tilde{\Gamma}^{(n)}$ is the intersection of the linear subspace of $\mathbb{R}^{J_n}$ satisfying (34) with the compact, convex set $\Gamma^{(n)} \subset \mathbb{R}^{J_n}$. To exploit this fact, we first construct an orthonormal basis $\mathbf{B}^{(n)}$ for the null space $\mathbf{N}^{(n)}$ of $\Delta\mathbf{P}^{(n)}(1)$. Letting γ^0 be any point in $\tilde{\Gamma}^{(n)}$, the set of $\gamma \in \mathbb{R}^{J_n}$ satisfying (34) can then be represented as

$$\{\gamma \in \mathbb{R}^{J_n} : \gamma = \gamma^0 + \mathbf{B}^{(n)} \cdot \nu \text{ for some } \nu \in \mathbb{R}^{J_n-(K-1)}\}.$$

Building on this observation, we draw γ^s as follows. Starting from an initial point $\gamma^0 \in \tilde{\Gamma}^{(n)}$, sample $\tilde{\nu}$ from a multivariate uniform distribution on $-[\kappa, \kappa]$, with $\kappa > 0$ a tuning parameter chosen by the researcher. Construct $\tilde{\gamma} = \gamma^0 + \mathbf{B}^{(n)} \cdot \tilde{\nu}$ and to see whether $\tilde{\gamma}$ satisfies the inequality restrictions defining $\Gamma^{(n)}$. If so, then $\tilde{\gamma} \in \tilde{\Gamma}^{(n)}$ and we set $\gamma^1 = \tilde{\gamma}$. If not, then we reject $\tilde{\gamma}$, draw a new innovation $\tilde{\nu}$, and repeat until $\tilde{\gamma} \in \tilde{\Gamma}^{(n)}$ is obtained. At this point we set $\gamma^1 = \tilde{\gamma}$ and proceed to the next iteration. Repeating these steps S times yields a Markov chain $\{\gamma^s\}_{s=0}^{S}$ exploring $\tilde{\Gamma}^{(n)}$.

As usual with Markov Chain type algorithms, the procedure above induces substantial autocorrelation in successive elements of $\{\gamma^s\}_{s=0}^{S}$. To avoid redundant Step 2 calculations, in practice we begin with a very large number of draws ($S = 40,000,000$) then subsample every $l = 200$th element. Since each iteration is virtually costless, this can be implemented reasonably quickly even for very large S.

Step 2 Given a subsample $\{\gamma^{s_l}\}_{l=1}^{L}$ of parameters in $\tilde{\Gamma}^{(n)}$, we then proceed to Step 2: evaluating Conditions 2 and 3 of Theorem 2 for each $\gamma^{s_l} \in \{\gamma^{s_l}\}_{l=1}^{L}$. Toward this end, we first use spline interpolation to invert $\tilde{v}_k(\cdot; \gamma^{s_l})$ for $\tilde{F}_k(\cdot; \gamma^{s_l})$ at each $k \in \mathcal{K}$. For each $k = 1, ..., K_1$, we then define $\phi_k(\cdot; \gamma^{s_l}) : [0, \tilde{v}_1(1, \gamma^{s_l})] \to \mathbb{R}$ as follows:

$$\phi_k(y; \gamma^{s_l}) = \frac{(1 - s_k)\tilde{F}_k(y; \gamma^{s_l}) - (1 - s_{k+1})\tilde{F}_{k+1}(y; \gamma^{s_l})}{(s_{k+1} - s_k)}.$$

Condition 2 requires ϕ_k to be increasing in y:

$$\phi_k(y'; \gamma^{s_l}) \geq \phi_k(y; \gamma^{s_l}) \quad \text{for all } y, y' \text{ such that } y' \geq y. \tag{35}$$

Meanwhile, Condition 3 requires ϕ_k to be decreasing in k:

$$1 \geq \phi_1(y; \gamma^{s_l}) \geq \phi_2(y; \gamma^{s_l}) \geq \cdots \geq \phi_{K-1}(y; \gamma^{s_l}) \geq \tilde{F}_K(y; \gamma^{s_l}). \tag{36}$$

To quantify violations of Condition 2, we first specify a grid Y of points in $[0, \tilde{v}_1(1; \gamma^{s_l})]$. We then sum up negative variation in $\phi_k(\cdot; \gamma^{s_l})$ across $y \in Y$ and $k \in \mathcal{K}$ to obtain

$$\mathbf{V}^2(\gamma^{s_l}) = \sum_{k=1}^{K-1} \sum_{g=1}^{|Y|-1} |\phi_k(y; \gamma^{s_l}) - \phi_k(y'; \gamma^{s_l})| \cdot \mathbb{I}[\phi_k(y'; \gamma^{s_l}) < \phi_k(y; \gamma^{s_l})].$$

Meanwhile, to quantify departures from Condition 3, we first compute pointwise violations of the inequalities (36) above at each $y \in Y$: i.e. we compute

$$\begin{aligned} V^3(y; \gamma^{s_l}) \equiv\ & |1 - \phi_1(y; \gamma^{s_l})| \cdot \mathbb{I}[\phi_1(y; \gamma^{s_l}) < 1] \\ & + \sum_{k=1}^{K-1} |\phi_{k+1}(y; \gamma^{s_l}) - \phi_k(y; \gamma^{s_l})| \cdot \mathbb{I}[\phi_k(y; \gamma^{s_l}) < \phi_{k+1}(y; \gamma^{s_l})] \\ & \qquad + |\tilde{F}_K(y; \gamma^{s_l}) - \phi_{K-1}(y; \gamma^{s_l})| \cdot \mathbb{I}[\phi_{K-1}(y; \gamma^{s_l}) < \tilde{F}_K(y; \gamma^{s_l})]. \end{aligned}$$

We then sum up these violations across $y \in Y$ to obtain the final criterion

$$\mathbf{V}^3(\gamma^{s_l}) \equiv \sum_{y \in Y} V^3(y).$$

Given tolerances $\epsilon_2, \epsilon_3 > 0$ specified by the researcher, we keep γ^{s_l} in the feasible set if $\mathbf{V}^2(\gamma^{s_l}) < \epsilon_2$ and $\mathbf{V}^3(\gamma^{s_l}) < \epsilon_3$. Otherwise we interpret γ^{s_l} as violating at least one of Conditions 2 and 3 and therefore drop it from the feasible set.

Simulation results Building on this algorithm, we construct approximations to the sharp nonparametric identified sets for several variants of the model specified in Section 7. As in Section 7, bidders have CRRA utility with $w_0 = c_0$, valuations are drawn from a $N(5, \sigma_0)$ distribution truncated on $[0, 10]$, and dependence between s_i and v_i is characterized by a Gumbel copula with parameter θ_0. For purposes of this exercise, we set $c_0 = 0.2$, $\sigma_0 = 2$, and vary ρ_0 and θ_0, assuming that N varies exogenously on $\mathcal{N} = [3, 4, 5, 6, 7, 8]$. In simulating feasible parameters, we draw a Markov chain of length $S = 40,000,000$ from $\tilde{\Gamma}^{(1)}$ based on step size $\kappa = 0.001$, subsampling every 200th element. We then compute $\mathbf{V}^2(\gamma^{s_l})$ and $\mathbf{V}^3(\gamma^{s_l})$ for each element of the resulting length-$100,000$ chain, reporting results for various values of ϵ_1, ϵ_2. Results reported are based on approximation of λ_0^{-1} within a shape-constrained Bernstein sieve space of order 24; preliminary test with a sieve space of order 32 yielded very similar results at greater computation cost.

Results of this procedure are summarized in Figures 1-4 for four data generating processes: limited selection with moderate risk aversion ($\theta_0 = 1.5$, $\rho_0 = 0.3$), moderate selection with moderate risk aversion ($\theta_0 = 1.5$, $\rho_0 = 0.1$), limited selection with limited risk aversion ($\theta_0 = 1.1$, $\rho_0 = 0.3$), and moderate selection with limited risk aversion ($\theta_0 = 1.1$, $\rho_0 = 0.1$). In each case we plot three sets of pointwise bounds derived from the analysis above. We first plot pointwise maxima and minima for λ_0^{-1} over the entire subsample $\{\gamma^{s_l}\}_{l=1}^{L}$; i.e. bounds exploiting only shape restrictions on λ_0^{-1} plus the invariant support restriction in Condition 1 of Theorem 2. We then plot two further sets of pointwise bounds incorporating restrictions derived from Conditions 2 and 3: first plotting pointwise maxima and minima for λ_0^{-1} over the subset of $\{\gamma^{s_l}\}_{l=1}^{L}$ for which both $\mathbf{V}^2(\gamma^{s_l})$ and $\mathbf{V}^3(\gamma^{s_l})$ are less than 0.5, then plotting pointwise maxima and minima over the subset of $\{\gamma^{s_l}\}_{l=1}^{L}$ for which both $\mathbf{V}^2(\gamma^{s_l})$ and $\mathbf{V}^3(\gamma^{s_l})$ are less than 0.1. Note that both thresholds 0.5 and 0.1 are large relative to the numerical error in $\mathbf{V}^2(\gamma^{s_l})$, $\mathbf{V}^3(\gamma^{s_l})$ we expect in simulations. Hence numerical bounds in all cases are likely to be somewhat conservative.

Several important patterns are suggested by this analysis. First, pointwise bounds are reasonably tight in all cases, with upper bounds on λ_0^{-1} particularly sharp. Since we are here approximating λ_0^{-1} within a very flexible sieve space, this suggests that observed bid data in fact contains substantial information on underlying (nonparametric) primitives. Second, data generating processes involving less selection support substantially tighter bounds on λ_0^{-1}. This pattern (more selection leads to wider bounds on primitives) is quite consistent with GL (2014)'s findings in the risk neutral case; intuitively, it follows since more selection induces more variation in $v_1, ..., v_K$ and thereby gives the econometrician more "degrees of freedom" in attempting to find a rationalizing model. Finally, even looking *only* at invariant support (i.e. ignoring Conditions 2 and 3 of 2 altogether), identified sets are clearly inconsistent with risk neutrality even when the true data generating process involves very little risk aversion ($\rho_0 = 0.1$). This suggests that building on the restrictions formalized in Theorem 2 it may be possible to derive nonparametric or robust semiparametric tests for risk aversion applicable even under endogenous and arbitrarily selective entry.

[Figure 1 about here.]

[Figure 2 about here.]

[Figure 3 about here.]

[Figure 4 about here.]

Appendix C: Estimation with parametric C_0 but nonparametric U_0

As an extension of the simple CRRA-based estimator explored in the main text, we also consider estimation of a semiparametric AS-RA model as in Section 5.3: C_0 restricted within a parametric family but λ_0^{-1} and F_0 left fully nonparametric. We implement estimation in this case via two-step sieve minimum distance. In particular, we first approximate λ^{-1} within the sieve space of shape constrained Bernstein polynomials $\mathcal{P}^{(n)}$ defined in Appendix B. We then minimize a criterion function derived from the compatibility condition (17) with respect to θ and the Bernstein sieve parameters, given estimates of s_k, b_k, and R_k derived as in Section 7.1.

Two-step Bernstein polynomial sieve estimator

Extending the arguments in Section 6 to accommodate a nonparametric function λ_0^{-1}, it is straightforward to show that true primitives θ_0, λ_0^{-1} must satisfy the identification criterion:

$$Q(\lambda_0, \theta_0) = \sum_{k=1}^{K} \int_0^1 \left(\Delta\tilde{b}_{k,\theta_0}(a) + \Delta\lambda_0^{-1}(\tilde{R}_{k,\theta_0}(a)) \right)^2 da, \tag{37}$$

where $\Delta\lambda_0^{-1}(\tilde{R}_{k,\theta_0}(a))$ is defined by

$$\Delta\lambda_0^{-1}(\tilde{R}_{k,\theta_0}(a)) \equiv \lambda_0(\tilde{R}_{k,\theta}(a)) - \frac{1}{K}\sum_{l=1}^{K}\lambda_0(\tilde{R}_{l,\theta}(a))$$

and as in Section 6 we define for each $k = 1, ..., K$:

$$\tilde{b}_{k,\theta}(a) \equiv \hat{b}_k(\hat{h}_\theta^k(a)),$$
$$\tilde{R}_{k,\theta}(a) \equiv \hat{R}_k(\hat{h}_\theta^k(a)).$$

Extending the notation in Appendix B, let $\tilde{\mathbf{p}}_{k,\theta}^{(n)}(a) \equiv [p_{J_n,j}(\tilde{R}_{k,\theta}(a))]_{j=0}^{J_n}$ denote the vector of order-J_n Bernstein basis polynomials evaluated at $\tilde{R}_{k,\theta}(a)$. Let $\gamma \in \Gamma^{(n)}$ be any coefficient vector satisfying shape constraints in Definition 7. Plugging in the Bernstein sieve approximation to λ^{-1} in (37) at γ then yields the following sieve estimation criterion:

$$Q(\gamma, \theta) = \sum_{k=1}^{K}\int_0^1 \left(\Delta\tilde{b}_{k,\theta}(a) + \Delta\tilde{\mathbf{p}}_{k,\theta}^{(n)}(a)\cdot\gamma\right)^2 da, \tag{38}$$

where we first substitute the Bernstein sieve approximation $\tilde{\mathbf{p}}_{k,\theta}(a))\cdot\gamma$ for $\lambda^{-1}(\tilde{R}_{k,\theta}(a))$ and then exploit the fact that $\tilde{\mathbf{p}}_{k,\theta}(a))\cdot\gamma$ is additively separable in γ. As above, in practice we discretize (38) on a finite grid $A^{(n)}$ to obtain the final minimum distance criterion:

$$\tilde{Q}(\gamma, \theta) = \sum_{k\in\mathcal{K}}\sum_{a\in A^{(n)}} \left(\Delta\tilde{b}_{k,\theta}(a) + \Delta\tilde{\mathbf{p}}_{k,\theta}^{(n)}(a)\cdot\gamma\right)^2. \tag{39}$$

Note that in contrast to the CRRA case described above, consistency here requires $A^{(n)}$ to become dense in $[0, 1]$ as $J_n \to \infty$.

As above, the discretized criterion (39) leads to a significant computational advantage: taking θ as given, minimization of (39) with respect to γ is a simple constrained quadratic programming problem. To exploit this structure, we again implement estimation via a nested minimization algorithm. First, in the inner loop, we obtain $\gamma(\theta) \equiv \arg\min_\gamma \tilde{Q}(\gamma, \theta)$ via constrained OLS regression of $\Delta\tilde{b}_{k,\theta}(a)$ on $-\Delta\tilde{\mathbf{p}}_{k,\theta}^{(n)}(a)$ subject to the constraint $\gamma \in \Gamma^{(n)}$. Then, in the outer loop, we search over θ to locate

$$\hat{\theta} = \arg\min_\theta\{\tilde{Q}(\gamma(\theta), \theta)\}.$$

We thus obtain a fast, stable algorithm in which for any θ the unique global solution $\gamma(\theta)$ to the inner-loop minimization can be obtained almost immediately using standard quadratic programming methods. Since θ is typically low-dimensional, it is therefore simple to combine grid search and gradient-based search in the outer loop to verify that estimates $(\hat{\gamma}, \hat{\theta})$ correspond to a unique global minimum of the objective (39).

Monte Carlo performance

We apply essentially the same Monte Carlo design to study the performance of our two-step sieve polynomial estimator, with only three minor changes in estimation procedure. First, since nonparametric estimation of λ^{-1} requires A to become dense in $[0, 1]$, we take A to be a uniform grid on $[0, 1]$. Second, boundary performance is relatively more important in two-step sieve estimation than in two-step CRRA estimation. Since over-smoothing near boundaries can significantly degrade such performance, we cap bandwidth for quantiles close to the boundaries according to the following rule:

$$h(a) = \min\{\mathrm{bwc} \cdot S^{\frac{1}{3}}, h_0 + a, h_0 + 1 - a\},$$

where h_0 is a given constant (we take $h_0 = .05$). Finally, as theory predicts $b_k(0) = b'_k(0) = 0$, we enforce these constraints directly in local polynomial estimation.

Figure 1 summarizes estimates of $\lambda^{-1}(x)$ resulting from our two-step Bernstein sieve polynomial estimation procedure with scale factor $S = 4000$ and bandwidth constant $bwc = 5$. Two features of our proposed estimator are evident in Figure 1. First, levels of $\lambda^{-1}(x)$ are very imprecisely estimated. This is due primarily to substantial error in estimates of $\tilde{R}_k(a)$ near $a = 0$, which generates significant noise (and corresponding bias) in sieve estimates of $\lambda^{-1}(x)$ for x near zero. Second, while levels of λ^{-1} are very imprecise, the *slope* $d\lambda^{-1}/dx$ is in fact well estimated on the interior of the support of x. This is significant because under the CRRA hypothesis the slope of λ^{-1} should be approximately constant across any range of x considered. In large samples, the Bernstein sieve estimator thus provides a means to assess the suitability of our baseline CRRA structure.

Building on this observation, we now shift focus to estimates for the derivative $d\lambda^{-1}(x)/dx$ implied by Figure 1. Figure 2 plots estimates $d\lambda^{-1}(x)/dx$ for the 10th-90th quantiles of R across 100 Monte Carlo repetitions, and Figure 3 plots pointwise quantiles of these estimates. While (as above) point estimates are slightly biased down, both figures generally confirm that $d\lambda^{-1}/dx$ is well-estimated on the interior of the support of R, with larger bandwidths leading to more precise estimates. While more work would be required to devise a formal test of the CRRA specification on the basis of this information, Figures 1-3 suggest that such an approach would be feasible in principle.

[Figure 5 about here.]

[Figure 6 about here.]

[Figure 7 about here.]

References

Ackerberg, D., K. Hirano, and Q. Shahriar (2006): "The Buy-It-Now Option, Risk Aversion and Impatience in an Empirical Model of eBay Bidding," Unpublished Manuscript, University of Arizona.

Athey, S., and P. Haile (2005): "Nonparametric Approaches to Auctions," *Handbook of Econometrics*, Vol. 6.

Athey, S., and J. Levin (2001): "Information and Competition in U.S. Forest Service Timber Auctions," *Journal of Political Economy*, 109, 375–413.

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 –257.

Bajari, P., and A. Hortacsu (2003): "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions," *The RAND Journal of Economics*, 34(2), 329–355.

Bajari, P., and A. Hortacsu (2005): *Journal of Political Economy*113, 703–741.

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

Bhattacharya, V., J. W. Roberts, and A. Sweeting (2014): "Regulating Bidder Participation in Auctions," Unpublished Manuscript, Massachusetts Institute of Technology.

Bhattacharya, V., and A. Sweeting (2014): "Selective Entry and Auction Design," Unpublished Manuscript, Massachusetts Institute of Technology.

Campo, S., E. Guerre, I. Perrigne, and Q. Vuong (2011): "Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders," *Review of Economic Studies*, 78(1), 112–147.

Fan, Y., T. Li, and M. Pesendorfer (2015): "New Local Polynomial Quantile Estimation of First-Price Auctions," Working Paper, Vanderbilt University.

Fang, H., and X. Tang (2014): "Inference of Bidders' Risk Attitutes in Ascending Auctions with Endogenous Entry," *Journal of Econometrics*, 180(2), 198–216.

Gentry, M., and T. Li (2012a): "Auctions with Selective Entry," Unpublished manuscript, London School of Economics.

——— (2012b): "Identification in Auctions with Selective Entry," Working Paper, Vanderbilt University and London School of Economics. Stable URL: http://ssrn.com/abstract=2268597.

——— (2014): "Identification in Auctions with Selective Entry," *Econometrica*, 82(1), 315–344.

Gentry, M., and C. Stroup (2014): "Entry and Competition in Takeover Auctions," Unpublished Manuscript, London School of Economics.

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

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.

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

Heckman, J. J. (1976): "The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic Social Measurement*, 5, 475–492.

Hendricks, K., J. Pinkse, and R. Porter (2003): "Empirical Implications of Equilibrium Bidding in First Price, Symmetric, Common Value Auctions," *Review of Economic Studies*, 70(1), 115–145.

Hickman, B., and T. Hubbard (2014): "Replacing Sample Trimming with Boundary Correction in Nonparametric Estimation of First-Price Auctions," *Journal of Applied Econometrics*, Forthcoming.

Krasnokutskaya, E., and K. Seim (2011): "Bid Preference Programs and Participation in Highway Procurement Auctions," *American Economic Review*, 101, 2653–2686.

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

Li, H., and G. Tan (2000): "Hidden Reserve Prices with Risk Averse Bidders," Unpublished manuscript, University of British Columbia.

Li, T., J. Lu, and L. Zhao (2014): "Auctions with Selective Entry and Risk Averse Bidders: Theory and Evidence," *RAND Journal of Economics*, forthcoming.

Li, T., and B. Zhang (2010): "Testing for Affiliation in First-Price Auctions Using Entry Behavior," *International Economic Review*, 51(3), 837–850.

Li, T., and B. Zhang (2014): "Affliation and Entry in First-Price Auctions with Heterogeneous Bidders: An Analysis of Merger Effects," *American Economic Journal: Microeconomics*, Forthcoming.

Li, T., and X. Zheng (2009): "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions," *Review of Economic Studies*, 76(4), 1397–1429.

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, 871–896.

Marmer, V., A. Shneyerov, and P. Xu (2013): "What Model for Entry in First-Price Auctions? A Nonparametric Approach," *Journal of Econometrics*, 176(1), 46–58.

Maskin, E., and J. Riley (1984): "Optimal auctions with risk-averse bidders," *Econometrica*, 52(6), 1473–1518.

Matthews, S. (1987): "Comparing Auctions for Risk Averse Buyers: A Buyer's Point of View," *Econometrica*, 55(3), 633–646.

Roberts, J., and A. Sweeting (2013): "When Should Sellers Use Auctions?," *American Economic Review*, forthcoming.

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

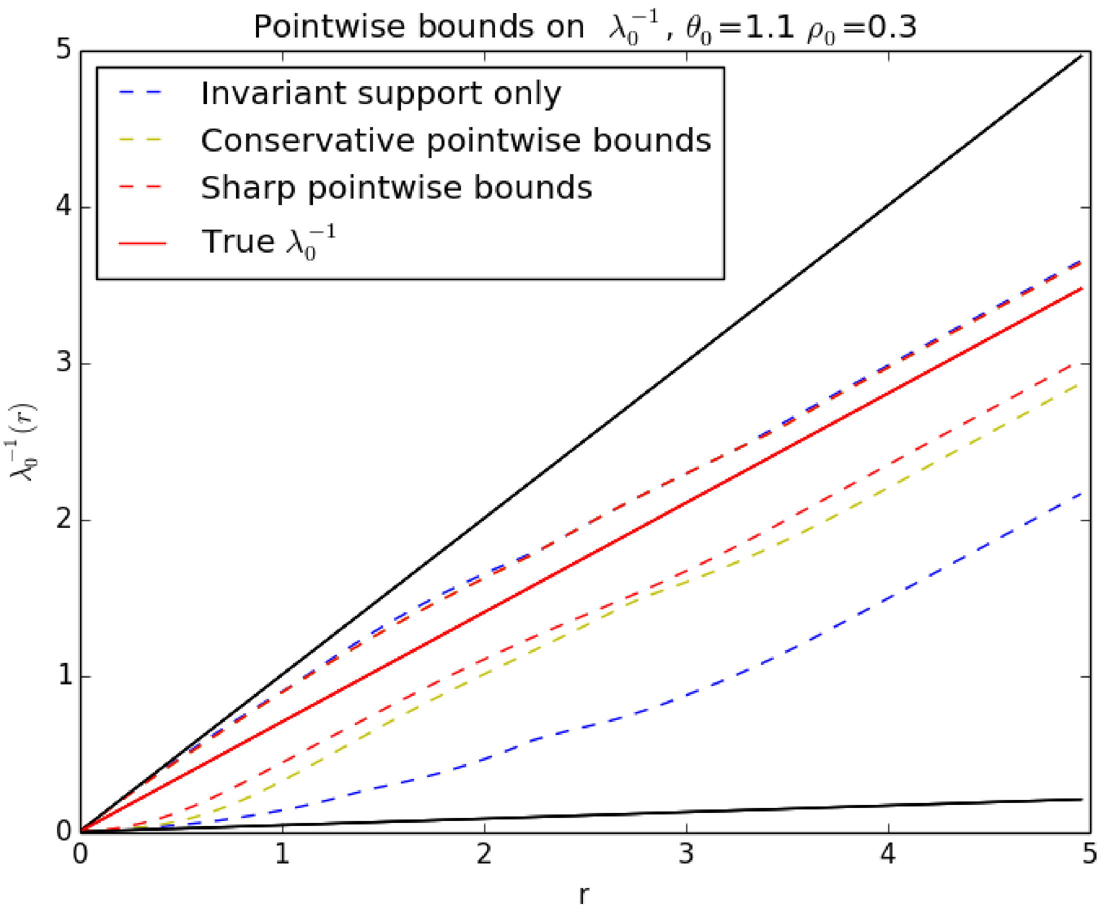
Smith, J., and D. Levin (1996): "Ranking Auctions with Risk Averse Bidders," *Journal of Economic Theory*, 68, 549–561.

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

Zhu, Y., and S. Grundl (2014): "Nonparametric Tests in Moment Equality Models with an Application to Infer Risk Aversion in First-Price Auctions," Unpublished Manuscript, University of Wisconsin.

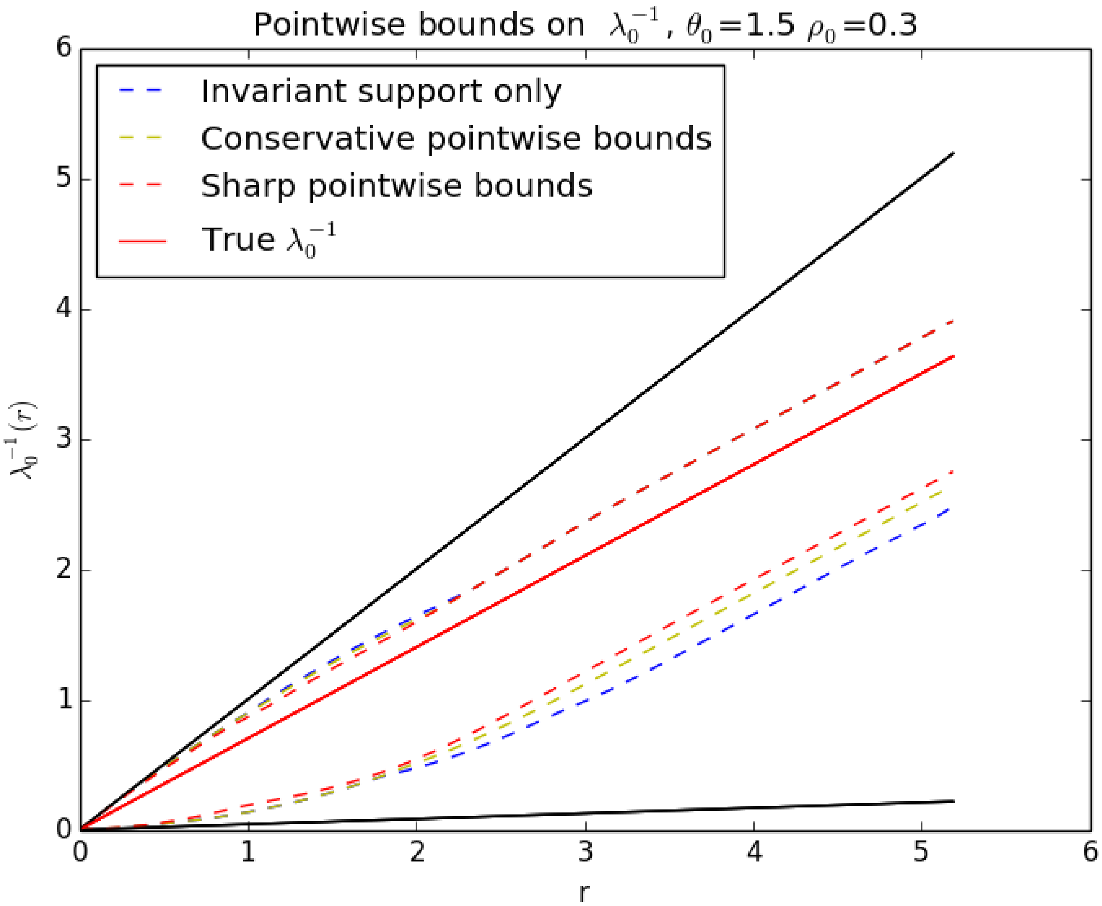
Zincenko, F. (2014): "Nonparametric Estimation of First-Price Auctions with Risk Averse Bidders," Unpublished Manuscript, University of Pittsburgh.

Figure 1: Bernstein sieve approximation to bid-stage identified set for $\lambda_0^{-1}(r)$



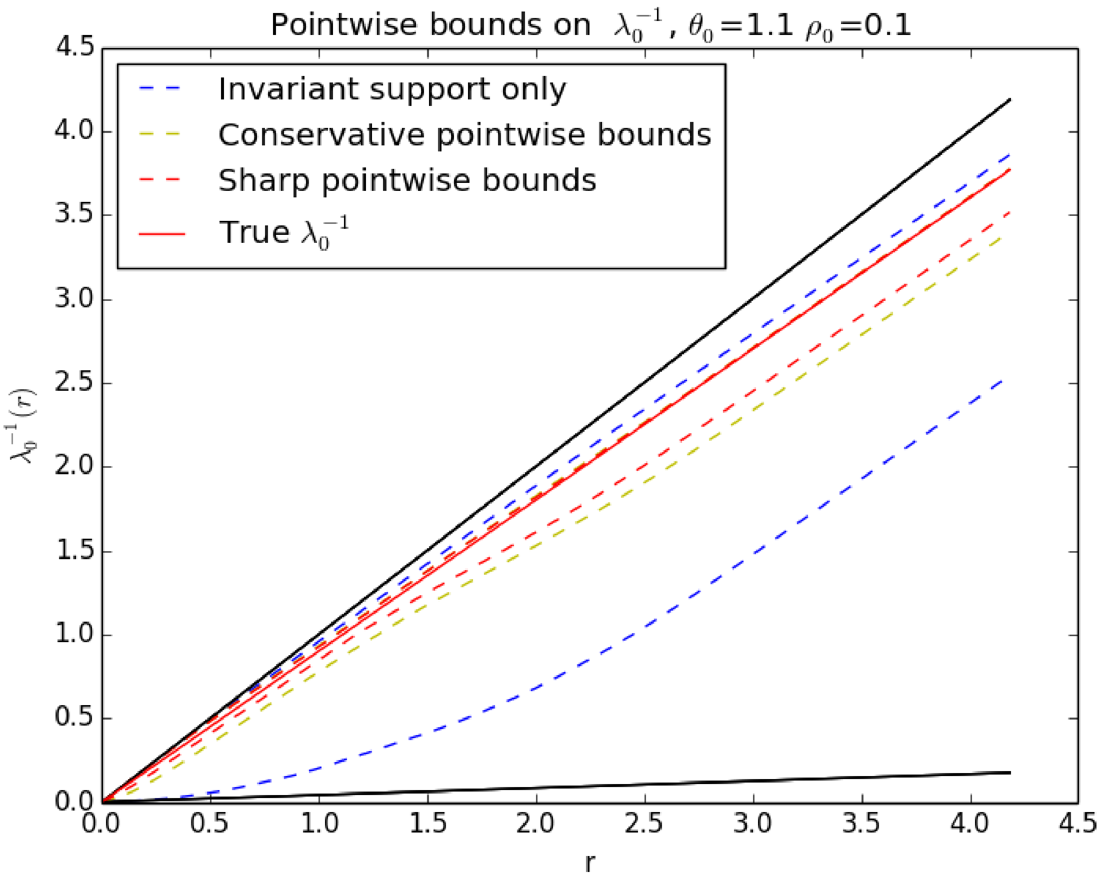
Approximation based on a shape-constrained Bernstein sieve of order 24, with feasible region $\tilde{\Gamma}^{(n)}$ for γ_0 simulated using length-200,000 subsample of length-40,000,000 Markov chain $\{\gamma^s\}_{s=1}^{S}$. "Invariant support only" denotes pointwise bounds obtained enforcing Condition 1 of Theorem 2 only. "Sharp pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.1$, $\mathbf{V}^3 < 0.1$. "Conservative pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.5$, $\mathbf{V}^3 < 0.5$. Pointwise upper bounds overlap.

Figure 2: Bernstein sieve approximation to bid-stage identified set for $\lambda_0^{-1}(r)$



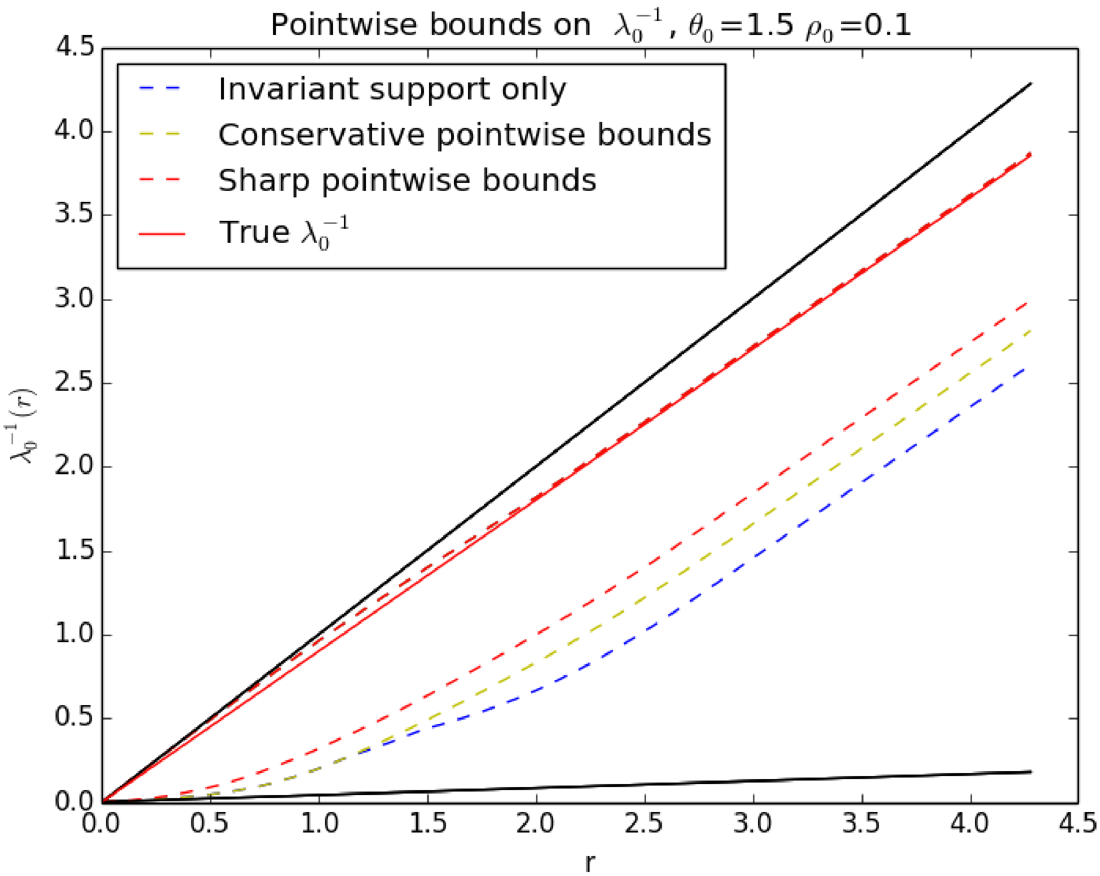
Approximation based on a shape-constrained Bernstein sieve of order 24, with feasible region $\tilde{\Gamma}^{(n)}$ for γ_0 simulated using length-200,000 subsample of length-40,000,000 Markov chain $\{\gamma^s\}_{s=1}^S$. "Invariant support only" denotes pointwise bounds obtained enforcing Condition 1 of Theorem 2 plus shape restrictions. "Sharp pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.1$, $\mathbf{V}^3 < 0.1$. "Conservative pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.5$, $\mathbf{V}^3 < 0.5$. Pointwise upper bounds overlap.

Figure 3: Bernstein sieve approximation to bid-stage identified set for $\lambda_0^{-1}(r)$



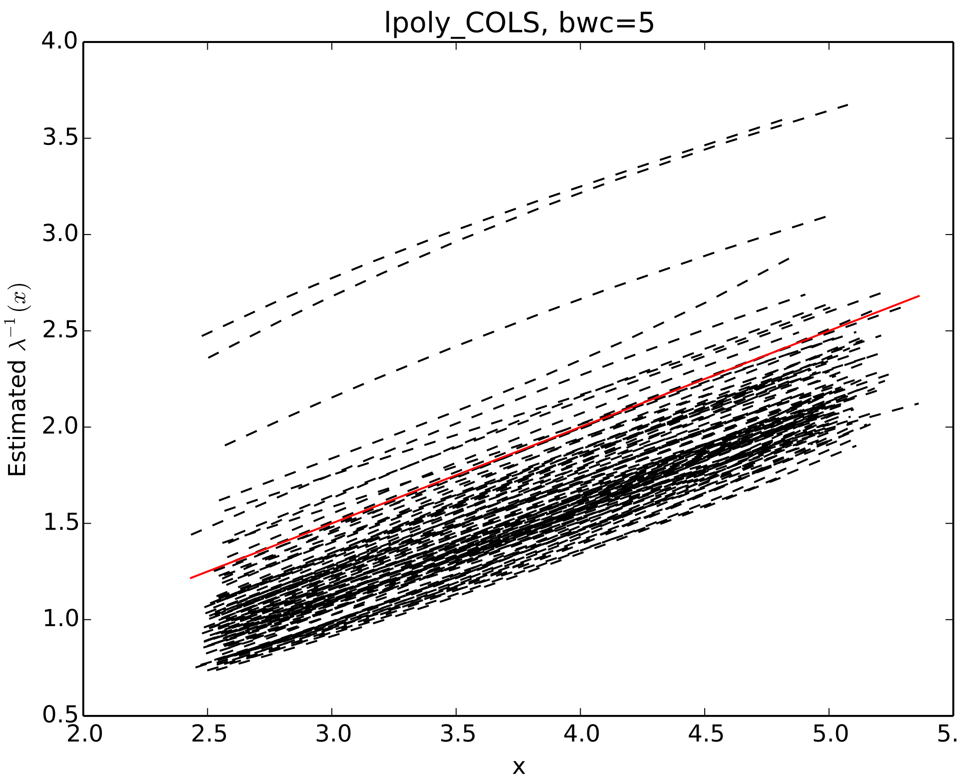
Approximation based on a shape-constrained Bernstein sieve of order 24, with feasible region $\tilde{\Gamma}^{(n)}$ for γ_0 simulated using length-200,000 subsample of length-40,000,000 Markov chain $\{\gamma^s\}_{s=1}^S$. "Invariant support only" denotes pointwise bounds obtained enforcing Condition 1 of Theorem 2 plus shape restrictions. "Sharp pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.1$, $\mathbf{V}^3 < 0.1$. "Conservative pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.5$, $\mathbf{V}^3 < 0.5$. Sharp and conservative pointwise upper bounds overlap.

Figure 4: Approximate bid-stage identified set for $\lambda_0^{-1}(r)$, $\theta_0 = 1.5$, $\rho_0 = 0.1$



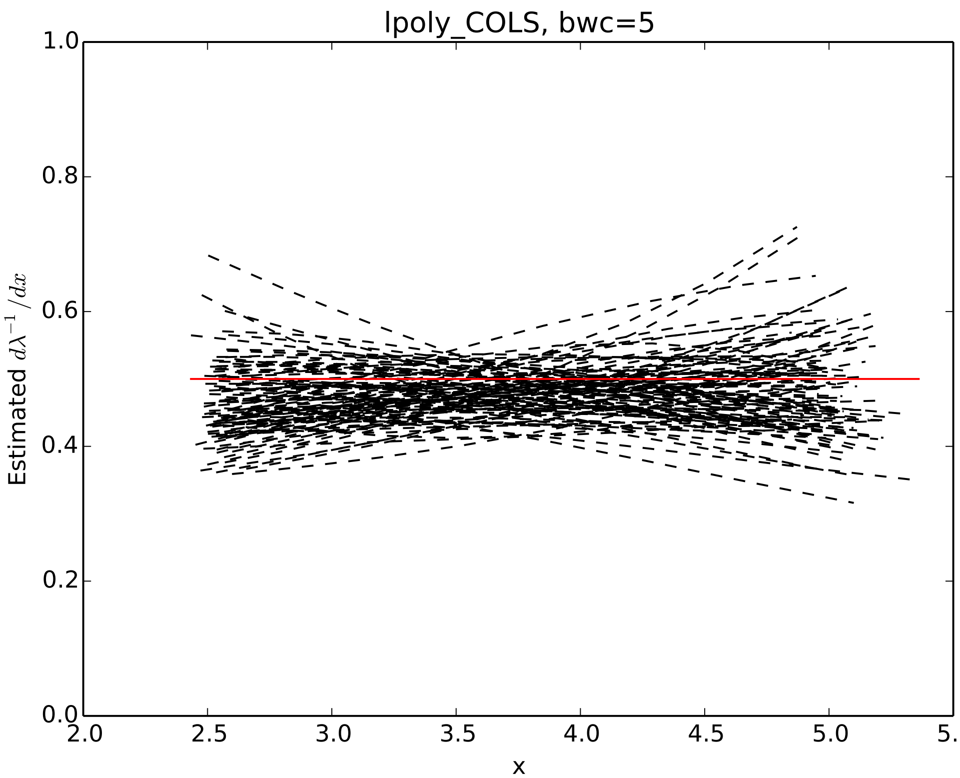
Approximation based on a shape-constrained Bernstein sieve of order 24, with feasible region $\tilde{\Gamma}^{(n)}$ simulated using size-200,000 subsample of length-40,000,000 Markov chain $\{\gamma^s\}_{s=1}^S$. "Invariant support only" denotes pointwise bounds obtained enforcing Condition 1 of Theorem 2 plus shape restrictions. "Sharp pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.1$, $\mathbf{V}^3 < 0.1$. "Conservative pointwise bounds" simulated from elements of Monte Carlo chain for which $\mathbf{V}^2 < 0.5$, $\mathbf{V}^3 < 0.5$. Pointwise upper bounds overlap.

Figure 5: Estimates of $\lambda^{-1}(x)$ (interior deciles of R)



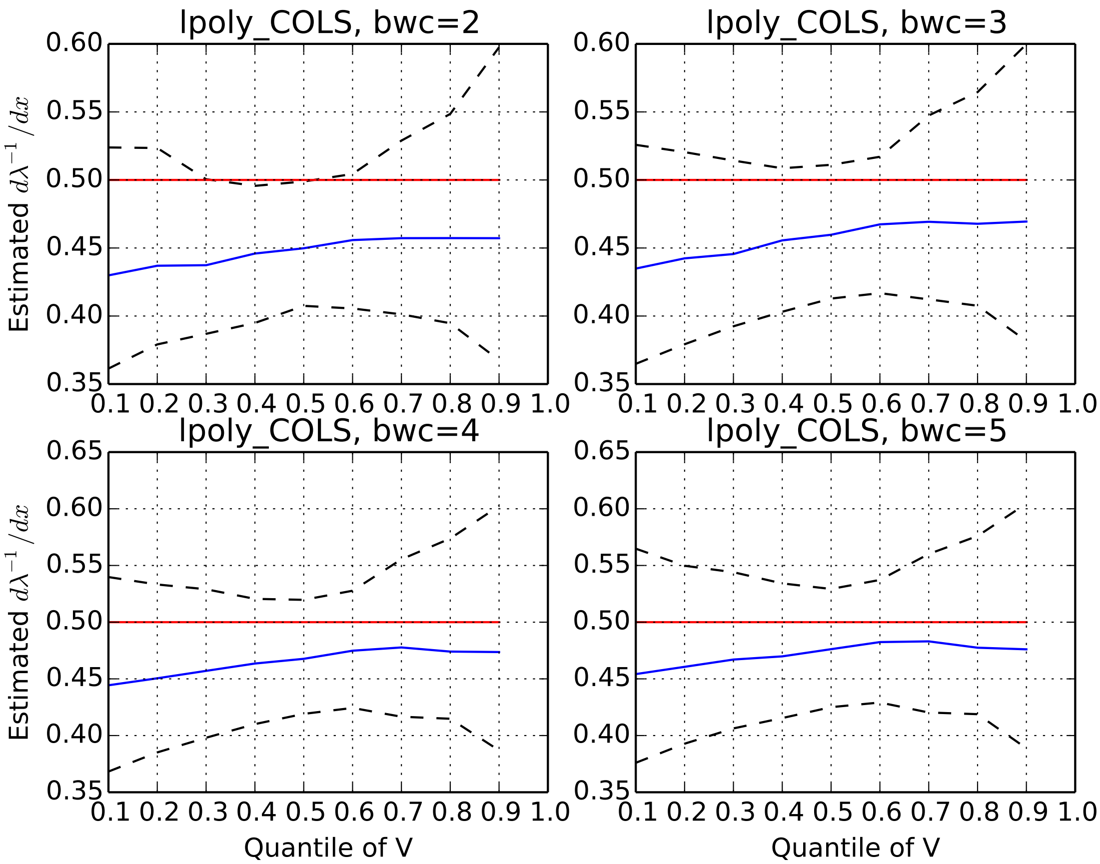
Estimates based on 100 Monte Carlo samples from AS-CRRA data generating process with $\rho_0 = 0.5$, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, $c_0 = w_0 = 0.2$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 4000$ bids observed at each observation. First step local polynomial bandwidths are given by $h = \text{bwc} \cdot S^{-1/3}$, with bwc $= 5.0$ for estimates reported.

Figure 6: Estimates of $d\lambda^{-1}/dx$ (interior deciles of R)



Estimates based on 100 Monte Carlo samples from AS-CRRA data generating process with $\rho_0 = 0.5$, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, $c_0 = w_0 = 0.2$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 4000$ bids observed at each observation. First step local polynomial bandwidths are given by $h = \text{bwc} \cdot S^{-1/3}$, with bwc $= 5.0$ for estimates reported.

Figure 7: Quantiles of $d\lambda^{-1}/dx$ (interior deciles of R)



Dashed lines represent 10th-90th quantiles of $d\lambda^{-1}/dx$ over 100 Monte Carlo simulations; blue line is median; red line is $d\lambda_0^{-1}/dx$. Underlying data generating process is AS-CRRA with $\rho_0 = 0.5$, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, $c_0 = w_0 = 0.2$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 4000$ bids observed at each observation. First step local polynomial bandwidths are given by $h = \text{bwc} \cdot S^{-1/3}$, with bwc varying on $\{2.0, 3.0, 4.0, 5.0\}$ as indicated.

Table 1: Two-step OLS-CRRA estimator, $S = 500$, $\rho_0 = 0.5$, $\theta_0 = 1.5$

$\hat{\rho}$	mean	std	25%	50%	75%
bwc = 1.0	0.399221	0.257272	0.218931	0.411732	0.543550
bwc = 1.5	0.424039	0.249043	0.285707	0.440248	0.552921
bwc = 2.0	0.417082	0.251915	0.270722	0.435488	0.553536
bwc = 3.0	0.439930	0.255433	0.307828	0.429347	0.568354
bwc = 4.0	0.494217	0.261150	0.341435	0.476652	0.596515
$\hat{\theta}$	mean	std	25%	50%	75%
bwc = 1.0	1.535023	0.346495	1.367761	1.441261	1.629569
bwc = 1.5	1.513945	0.379520	1.363254	1.438873	1.554221
bwc = 2.0	1.520083	0.398367	1.356968	1.425464	1.528418
bwc = 3.0	1.459115	0.315543	1.347753	1.419355	1.493907
bwc = 4.0	1.379589	0.211879	1.340931	1.397429	1.464333

Estimates based on 100 Monte Carlo samples from AS-CRRA process with $\rho_0 = 0.5$, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 500$ bids observed at each competition level. First step local polynomial bandwidths computed via $\text{bwc} \cdot S^{-1/3}$, for values of *bwc* reported above.

Table 2: Two-step OLS-CRRA estimator, $S = 1000$, $\rho_0 = 0.5$, $\theta_0 = 1.5$

$\hat{\rho}$	mean	std	25%	50%	75%
bwc = 1.0	0.420992	0.175029	0.357833	0.439752	0.528544
bwc = 1.5	0.414720	0.181355	0.356295	0.458151	0.530549
bwc = 2.0	0.425663	0.167566	0.372587	0.448925	0.532161
bwc = 3.0	0.472407	0.147228	0.420716	0.472703	0.551467
bwc = 4.0	0.491436	0.138030	0.436531	0.492518	0.563187
$\hat{\theta}$	mean	std	25%	50%	75%
bwc = 1.0	1.568788	0.296468	1.408262	1.483247	1.627447
bwc = 1.5	1.549957	0.271925	1.407975	1.465911	1.607106
bwc = 2.0	1.532709	0.273605	1.400348	1.457619	1.570694
bwc = 3.0	1.455302	0.196691	1.393846	1.428791	1.493555
bwc = 4.0	1.417398	0.108818	1.382068	1.408845	1.454900

Estimates based on 100 Monte Carlo samples from AS-CRRA process with $\rho_0 = 0.5$, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 1000$ bids observed at each competition level. First step local polynomial bandwidths computed via $\text{bwc} \cdot S^{-1/3}$, for values of *bwc* reported above.

Table 3: Two-step OLS-CRRA estimator, $S = 2000$, $\rho_0 = 0.5$, $\theta_0 = 1.5$

$\hat{\rho}$	mean	std	25%	50%	75%
bwc = 1.0	0.471312	0.118759	0.435467	0.497566	0.553754
bwc = 1.5	0.478903	0.105939	0.444104	0.497610	0.550294
bwc = 2.0	0.471831	0.120922	0.437695	0.500650	0.546182
bwc = 3.0	0.483431	0.088783	0.440036	0.498591	0.544441
bwc = 4.0	0.499125	0.075337	0.452781	0.501195	0.556878

$\hat{\theta}$	mean	std	25%	50%	75%
bwc = 1.0	1.512006	0.129052	1.432392	1.488613	1.547362
bwc = 1.5	1.493737	0.109186	1.427603	1.469686	1.529703
bwc = 2.0	1.496589	0.131696	1.427018	1.459374	1.513609
bwc = 2.5	1.472719	0.082343	1.415344	1.451792	1.510353
bwc = 3.0	1.450352	0.058761	1.408881	1.437910	1.489030

Estimates based on 100 Monte Carlo samples from AS-CRRA process with $\rho_0 = 0.5$, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 2000$ bids observed at each competition level. First step local polynomial bandwidths computed via bwc $\cdot\, S^{-1/3}$, for values of *bwc* reported above.

Table 4: Two-step OLS-CRRA estimator, $S = 2000$, $\rho_0 = 0.0$, $\theta_0 = 1.5$

$\hat{\rho}$	mean	std	25%	50%	75%
bwc = 1.0	0.020129	0.036026	0	0	0.024512
bwc = 1.5	0.015771	0.031992	0	0	0.003280
bwc = 2.0	0.013178	0.028219	0	0	0.001045
bwc = 3.0	0.008931	0.022009	0	0	0.000000
bwc = 4.0	0.006521	0.018608	0	0	0.000000
$\hat{\theta}$	mean	std	25%	50%	75%
bwc = 1.0	1.504600	0.070124	1.457686	1.491877	1.528829
bwc = 1.5	1.493492	0.057861	1.455038	1.485242	1.518443
bwc = 2.0	1.485448	0.049356	1.452146	1.479397	1.512814
bwc = 3.0	1.473774	0.043599	1.443437	1.470796	1.496075
bwc = 4.0	1.463839	0.039724	1.436120	1.459060	1.485859

Estimates based on 100 Monte Carlo samples from AS-CRRA process with risk-neutral bidders, $\theta_0 = 1.5$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 1000$ bids observed at each competition level. First step local polynomial bandwidths computed via $\text{bwc} \cdot S^{-1/3}$, for values of *bwc* reported above.

Table 5: Two-step OLS-CRRA estimator, $S = 2000$, $\rho_0 = 0.5$, $\theta_0 = 1.0$

$\hat{\rho}$	mean	std	25%	50%	75%
bwc = 1.0	0.448991	0.078023	0.421457	0.455907	0.497214
bwc = 1.5	0.444433	0.065119	0.413911	0.446102	0.488955
bwc = 2.0	0.432176	0.066427	0.400208	0.429393	0.476924
bwc = 3.0	0.400118	0.071315	0.357837	0.394831	0.450712
bwc = 4.0	0.367880	0.078002	0.317557	0.364526	0.423907
$\hat{\theta}$	mean	std	25%	50%	75%
bwc = 1.0	1.016308	0.071549	1.000000	1.004317	1.016048
bwc = 1.5	1.009654	0.012810	1.000000	1.005333	1.016150
bwc = 2.0	1.009850	0.012851	1.000000	1.005471	1.016427
bwc = 3.0	1.011076	0.013207	1.000000	1.005319	1.017767
bwc = 4.0	1.012490	0.013765	1.000067	1.007402	1.020937

Estimates based on 100 Monte Carlo samples from AS-CRRA process with no selection, $\rho_0 = 0.5$ $\mu_0 = 5.0$, $\sigma_0 = 2.0$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 2000$ bids observed at each competition level. First step local polynomial bandwidths computed via $\text{bwc} \cdot S^{-1/3}$, for values of *bwc* reported above.

Table 6: Two-step OLS-CRRA estimator, $S = 2000$, $\rho_0 = 0.2$, $\theta_0 = 1.2$

$\hat{\rho}$	mean	std	25%	50%	75%
bwc = 1.0	0.143641	0.099563	0.048830	0.151904	0.218030
bwc = 1.5	0.127124	0.089961	0.044346	0.125522	0.191729
bwc = 2.0	0.108466	0.083150	0.031770	0.108805	0.169897
bwc = 3.0	0.080438	0.073766	0.000000	0.072119	0.132281
bwc = 4.0	0.068299	0.069315	0.000000	0.052697	0.113191
$\hat{\theta}$	mean	std	25%	50%	75%
bwc = 1.0	1.213277	0.056411	1.181536	1.206031	1.223302
bwc = 1.5	1.209346	0.041990	1.181824	1.206144	1.221634
bwc = 2.0	1.207091	0.032267	1.184517	1.207244	1.223182
bwc = 3.0	1.206266	0.029824	1.183567	1.205641	1.222456
bwc = 4.0	1.205981	0.028782	1.184788	1.206167	1.221968

Estimates based on 100 Monte Carlo samples from AS-CRRA process with $\rho_0 = 0.2$, $\theta_0 = 1.2$, $\mu_0 = 5.0$, $\sigma_0 = 2.0$, and $\mathcal{N} = [2, 4, 6, 8]$, assuming approximately $S = 2000$ bids observed at each competition level. First step local polynomial bandwidths computed via bwc $\cdot\, S^{-1/3}$, for values of *bwc* reported above.