Partial Identification in Auctions with Selective Entry*

Matthew Gentry† and Tong Li‡

First version: September 2010
This version: August 2011

Abstract

This paper considers nonparametric identification of a two-stage entry and bidding model for auctions which we call the Affiliated-Signal (AS) model. This model assumes that potential bidders have private values, observe imperfect signals of their true values prior to entry, and choose whether to undertake a costly entry process. The AS model is a theoretically appealing candidate for the structural analysis of auctions with entry: it formally nests the Levin and Smith (1994) model and approaches the Samuelson (1985) model as a limit. Unfortunately, since pre-entry signals are not observed, the AS model is non-parametrically non-identified. In this paper, we explore a partial identification approach to structural analysis using the AS model. In particular, we show how exogenous variation in entry behavior (induced by variation in factors such as potential competition or entry costs) can be used to construct non-parametric bounds on the fundamentals of the AS model under a general class of second-stage bidding mechanisms, and translate these bounds into bounds on counterfactual seller revenue corresponding to a wide variety of potential award rules. Finally, we state conditions under which the AS model is exactly identified.

*Earlier versions of this paper have been presented at the EC^2 Conference on Identification in Econometrics in Toulouse, December 2010, the North American Summer Meeting of the Econometric Society in St. Louis, June 2011, and the European Meeting of the Econometric Society in Oslo, August 2011. We thank participants of these conferences for comments and Yanqin Fan for helpful suggestions. Support from National Science Foundation (SES-0922109) is gratefully acknowledged.

†Department of Economics, Vanderbilt University. Email: matthew.l.gentry@vanderbilt.edu.
‡Department of Economics, Vanderbilt University. Email: tong.li@vanderbilt.edu.

1 Introduction

Endogenous participation clearly matters in real-world auction markets. Studies of auctions in a variety of economic contexts routinely find that large fractions of eligible bidders elect not to submit bids. For example, Hendricks, Pinske, and Porter (2003) report an overall participation rate of less than 25 percent in US Minerals Management Service "wildcat auctions" held from 1954-1970. Li and Zheng (2009) find that only about 28 percent of planholders in Texas Department of Transportation mowing contracts actually submit bids. Similar results have been reported for timber auctions (Athey, Levin, and Seira (2011), Li and Zhang (2009; 2010)), in online auction markets (Bajari and Hortacsu (2003)), and in other procurement settings (Krasnokutskaya and Seim (2009)). Such endogenous participation can overturn core predictions of classical auction theory: for instance, Levin and Smith (1994) show that the possibility of entry can lead to a zero optimal reserve price, and Li and Zheng (2009) show that it can cause a seller to prefer *less* potential competition. Hence properly accounting for entry is practically important in applied research.

Though the importance of endogenous participation is well-established in the auction literature, there is still no clear consensus regarding how best to account for entry in structural analysis. To ensure identification, most applied work on auctions with entry has been based on (variants of) two polar models: that of Samuelson (1985) (the *S model*) and that of Levin and Smith (1994) (the *LS model*). These models both turn on the intuition that entry involves costs (to learn about the object being sold, prepare bids, etc), but they involve radically different assumptions about what potential bidders know when making entry decisions. In particular, in the S model, potential bidders are assumed to know their values exactly prior to entry, whereas in the LS model they are assumed to have no specific information *ex ante*. In turn, such strong informational assumptions translate into stark restrictions on the underlying DGP: the S model implies that entrants are perfectly selected, while the LS model implies that entrants are a random sample from the population. Consequently, inappropriately enforcing either polar model can substantially distort structural analysis, with significant consequences for counterfactual policy results.[1]

In this paper, we consider one potential solution to these problems: structural analysis based on a framework we call the *Affiliated-Signal (AS) model*. This model was first proposed in Ye (2007), and has since been explored by several other authors, most notably Marmer, Shneyerov, and Xu (2007) (henceforth MSX), who propose

[1]For instance, consider an independent private values (IPV) setting where the seller's value is normalized to zero. Then the optimal reserve price is zero in the LS model (see Levin and Smith (1994)), but can be positive when entry involves selection, as shown in Li and Zheng (2007). Structural estimation based on an incorrect LS specification would force a researcher to the conclusion that the optimal reserve price is zero. Marmer, Shneyerov, and Xu (2007) and Roberts and Sweeting (2010b) discuss parameter bias and policy implications resulting from improper entry specifications; Roberts and Sweeting (2010b) and Gentry (2010) provide simulation evidence on the potential magnitudes of the biases involved.

nonparametric tests of the AS, S, and LS models, and Roberts and Sweeting (2010a; 2010b), who apply a parametric variant of the model to California timber auctions. The basic structure of the AS model is as follows: potential bidders have private values, observe imperfect signals of their true values prior to entry, choose whether to undertake a costly entry process, then (conditional on entry) learn their exact values and submit bids. As noted by MSX, the AS model provides an ideal theoretical bridge between the S and LS polar cases: it formally nests the LS model and approaches the S model as a limit. Thus structural estimation based on the AS model would represent a substantial improvement over structural estimation based on either the S or LS extreme cases.

Unfortunately, as noted in MSX, the general AS model is nonparametrically non-identified: since preliminary signals are unobservable, entry decisions depend on a conditional distribution that cannot be recovered from the data.[2] Consequently, attempts to apply the AS model must adopt one of two approaches: impose parametric restrictions such that the model is identified, or try to learn as much as possible from the data without such restrictions. Roberts and Sweeting (2010a; 2010b) explore the first approach, using a joint log-normal distribution (with unobserved auction heterogeneity) to model dependence between signals and values. In this work, we explore the second: since the AS model is not exactly identified, we seek to obtain identified *bounds* on objects of interest. We thus adopt a variant of the partial identification approach pioneered by Manski, which has motivated a large and growing literature in econometrics.[3]

In particular, this paper establishes three core results on identification in auctions with entry. First, using exogenous variation in entry behavior (induced by variation in, e.g., potential competition or entry costs), we derive identified bounds on fundamentals under endogenous and arbitrarily selective entry in a general class of auction mechanisms considered in Riley and Samuelson (1981). Second, we translate these bounds on fundamentals into bounds on seller revenue corresponding to a wide range of counterfactual award rules (again accounting for endogenous and selective entry). Finally, we explore sharpness properties of the underlying bounds and state conditions under which all bounds collapse to exact identification. To our knowledge, these

[2]Nonparametric identification of auction models focuses on using observables such as bids to identify model primitives such as the distribution of values. This literature was established by Guerre, Perrigne, and Vuong (2000), who address nonparametric identification and estimation in first-price IPV auctions, and has focused primarily on auction models without entry. See, e.g., Li, Perrigne, and Vuong (2002) for the affiliated private value (APV) model, Li, Perrigne, and Vuong (2000) for the conditionally independent private information model, Krasnokutskaya (2009) for an asymmetric auction with unobserved auction heterogeneity, Hortacsu (2002) for treasury bond auctions, and Athey and Haile (2002) for other standard auction models/formats. Athey and Haile (2005) provide a comprehensive survey of the literature.

[3]See Manski (2003) for a summary of the early partial-identification literature; recent additions to the literature include Manski and Tamer (2002), Magnac and Maurin (2008), Molinari (2008), Fan and Park (2009), and Tamer (2003), to name only a few.

are the first such results in the identification literature.[4]

Within the existing literature on partial identification in auctions, our work is most similar in spirit to that of Haile and Tamer (2003), who derive identified bounds on fundamentals in ascending auctions under weak behavioral assumptions, then translate their results into bounds on the seller's optimal reserve price. However, we focus on a very different problem (endogenous entry, not considered in Haile and Tamer), and relax a different set of assumptions (those governing the nature of selection). Another related paper is Tang (2011), who provides bounds for counterfactual revenue in affiliated values (AV) auction settings. Again, however, our work is set in a different context (auctions with selective entry) and focuses on a much different set of problems. We thus contribute both to the literature on partial identification and to the literature on the econometrics of auctions with entry.[5]

The plan of the paper is as follows. Section 2 describes the IPV AS model and outlines the entry and bidding equilibria corresponding to a general class of auction mechanisms. Sections 3 and 4 explore identification of the AS entry model within this general class of mechanisms and present our core partial-identification results. Section 5 translates these core partial-identification results into bounds on seller revenue corresponding to a wide range of counterfactual award rules. Finally, Section 6 concludes. Detailed proofs and a numerical example are included as appendices.

2 The AS model

We consider allocation of a single indivisible good among N potential bidders via a two-stage auction mechanism, where bidders have indepentent private values for the good being sold. First, in Stage 1, each of N potential bidders observes a private signal s_i (to be formalized below), and all potential bidders simultaneously choose whether to enter the auction at cost c. Then, in Stage 2, the n bidders who chose to enter in Stage 1 learn their true values v_i and submit bids for the object being sold. Finally, auction outcomes are determined according to the rules of the Stage 2 mechanism, which are common knowledge to all participants. Consistent with institutional features common to many official procurement lettings, we assume that bidders observe the number of potential bidders N prior to entry, but do not observe

[4]It should be noted, however, that our main focus in this work is nonparametric *identification*, not nonparametric inference. Consequently, while we derive nonparametric bounds on model primitives and other quantities of policy interest, we do not develop asymptotic distribution theory for these bounds. In this respect, we follow several prior studies, e.g. Athey and Haile (2002), Haile and Tamer (2003) and Manski and Tamer (2002).

[5]Though much of the early structural auction literature focused on the no-entry case, auctions with entry have also received substantial attention in recent years. See, e.g., Athey, Levin, and Seira (2011), Bajari and Hortacsu (2003), Li (2005), Li and Zhang (2009; 2010), Li and Zheng (2007; 2009), Hendricks, Pinske, and Porter (2003), Krasnokutskaya and Seim (2009), MSX, and Roberts and Sweeting (2010b; 2010a) among others.

the number of actual entrants n prior to bidding.[6]

2.1 Model setup

Most existing results on identification in auctions concern recovery of Stage 2 value distributions from Stage 2 bids; we refer readers to Athey and Haile (2005) for details. In contrast, this work focuses on identification of the Stage 1 entry model *given* objects already established as Stage 2 identified using appropriate rule-specific techniques. To emphasize this distinction, we frame our analysis in terms of a general class of mechanisms we call *RS auctions* (after the structure explored by Riley and Samuelson (1981)):

Definition 1. A *RS auction* is any bidding mechanism having the following properties:

1. Mechanism rules are anonymous.
2. If award is made, it is to the bidder submitting the highest bid.
3. The probability of award depends only on the highest bid.
4. For any nondegenerate distribution of rival values, there exists a unique symmetric bidding equilibrium such that bids submitted are strictly increasing in bidder values.

As noted by Riley and Sameulson (1981), this class of mechanisms includes all four standard auctions (first-price, Vickery, English ascending, and Dutch), plus many less common auction types. Further, and more important for current purposes, the class of RS auctions also contains almost all bidding rules for which Stage 2 identification results are known. Hence RS auctions represent a natural focal point for our current investigation.

We formalize the remaining assumptions of the AS entry model as follows.

Assumption 1. *The seller and all potential bidders are risk-neutral.*

Assumption 2. *All bidders are* ex ante *symmetric, and draw (unknown) signal-value pairs* (V, S) *independently from a continuous joint distribution having density* $f(v, s)$.[7]

(i) *The marginal density of second-stage values* ($f(V)$) *has positive support on a bounded interval* $[\underline{v}, \bar{v}]$.

[6] Allowing bidders to observe n prior to bidding would slightly change the details of the derivation, but would not substantially alter any of our core results.

[7] We discuss how this assumption can be relaxed as an extension.

(ii) *WLOG, we normalize first-stage signals (S) to have a uniform marginal distribution on* $[0, 1]$*:* $s_i \sim U[0, 1]$.

Assumption 3. *For each bidder i, the random variables S_i and V_i are* affiliated *in the sense of Milgrom and Weber (1982).*[8]

Assumption 4. *Information structure:*

(i) *The second-stage auction mechanism M, the number of potential bidders N, and all other model fundamentals are common knowledge, but the number of actual bidders n is not revealed until the auction concludes.*

(ii) *Each bidder i observes own signal s_i prior to entry, but does not learn own value v_i until after entry.*

Finally, as noted above, we focus on RS auctions. By Definition 1, we can characterize allocations resulting from such auctions by an *award rule* $\alpha(\cdot)$, where $\alpha(y)$ represents the probability of award when the maximum value among entrants is y. This function $\alpha(\cdot)$ will be the main focus of our counterfactual analysis. For current purposes, we add two further regularity conditions on the mechanism M:

Assumption 5. *The second-stage auction mechanism M is a Riley-Samuelson auction such that:*

(i) *The award rule $\alpha(y)$ is weakly increasing in the maximum entrant value y.*

(ii) *A low-type bidder (entrant with value* $\underline{\mathrm{v}}$*) weakly prefers less Stage 2 competition.*

While one can construct theoretical mechanisms violating these conditions, they are satisfied by almost all auctions used in practice. Hence we maintain them through the rest of the paper.

2.2 Stage 2: Bidding equilibrium

As usual in the literature, we focus on the class of symmetric subgame perfect Bayesian Nash equilibria. Since bidders are *ex ante* symmetric by hypothesis, so any such symmetric equilibrium must involve a common Stage 1 signal threshold $\bar{s} \in [0, 1]$ such

[8] Loosely speaking, that two random variables are affiliated means that a large value of one variable makes the value of the other variable more likely to be large than small. Formally, let $T \equiv (S, V)$ with a density $f_T(T)$. Let T and T' be any two values of T. It is said that the elements of T are affiliated if $f_T(T \vee T')f_T(T \wedge T') \geq f_T(T)f_T(T')$, where $T \vee T'$ denotes the component-wise maximum of T and T', and $T \wedge T'$ denotes the component-wise minimum of T and T'. See Milgrom and Weber (1982) for details.

that bidder i chooses to enter if and only if $s_i \geq \bar{s}$.[9] Then the (selected) distribution of values among bidders entering at equilibrium $\bar{s}$ will be given by

$$F^*(v;\bar{s}) \equiv \frac{1}{1-\bar{s}} \int_{\bar{s}}^{1} F(v|t)\, dt. \tag{1}$$

Let $w_j \equiv \mathbf{1}[s_j \geq \bar{s}] \cdot v_j$ be potential bidder j's *realized value* at $\bar{s}$ (so that $w_j \equiv 0$ if bidder i stays out and v_j otherwise). Then the *ex ante* probability that bidder j will draw a realized value below v given threshold $\bar{s}$ is

$$F_w^*(v;\bar{s}) = \bar{s} + (1-\bar{s})F^*(v;\bar{s}).$$

Further, in any symmetric, monotonic Stage 2 bidding equilibrium, $b_i^* > b_j^*$ if and only if $w_i > w_j$. Hence $F_w^*(v_i;\bar{s})$ can be interpreted as the probability that entering bidder i with value v_i will outbid any given potential rival j at entry equilibrium $\bar{s}$.

Equilibrium bidding behavior will obviously depend on the specific rules of the Stage 2 mechanism in question. However, via standard arguments in mechanism design, we can characterize expected symmetric equilibrium *profit* for any RS mechanism. This is formally shown in a companion paper (Gentry and Li (2011b)); for conciseness, we simply state the relevant proposition here.

Proposition 1. *In any symmetric Stage 2 equilibrium of any RS mechanism, a bidder with value v facing competition structure $(N,\bar{s})$ earns expected profit*

$$\pi(v;\bar{s},N) = \pi_0(\bar{s},N) + \int_{\underline{\mathrm{V}}}^{v} \alpha(y) \cdot F_w^*(y;\bar{s})^{N-1}\, dy, \tag{2}$$

where $\pi_0(\bar{s},N)$ is a mechanism-specific intercept that does not depend on v.

Thus, conditional on competition, the award probablity function $\alpha(\cdot)$ determines bidder profits up to an additive constant.

2.3 Stage 1: Entry equilibrium

Given this symmetric Stage 2 equilibrium profit function $\pi(v;\bar{s},N)$, we can characterize the symmetric Stage 1 entry threshold $\bar{s}$ as a function of model fundamentals and competition N. We seek to obtain a symmetric threshold function $\bar{s}(c,N) : \{\mathbb{R}^+,\mathbb{N}\} \to [0,1]$ such that "enter if $s_i \geq \bar{s}(c,N)$" is the unique Stage 1 equilibrium given the Stage 2 continuation play as in Proposition 1 above.

[9] Intuitively, any symmetric subgame perfect BNE must (by definition) involve a symmetric entry equilibrium, and given any symmetric bidding equilibrium the only possible symmetric entry equilibrium in a game with informative signals takes the threshold form. This latter claim is shown formally in the appendix.

Toward this end, consider the Stage 1 decision faced by potential bidder i (with Stage 1 signal s_i) facing N potential rivals who enter according to threshold $\bar{s}$. Then bidder i's *ex ante* expected Stage 2 profit $\Pi(s_i;\bar{s},N)$ is given by

$$\begin{aligned}
\Pi(s_i;\bar{s},N) &= E_v[\pi(v;\bar{s},N)|S=s_i] \\
&= \pi_0(\bar{s},N) + \int_{\underline{v}}^{\bar{v}} f(v|s_i) \int_{\underline{v}}^{v} \alpha(y)\cdot F_w^*(y;\bar{s})^{N-1} dy\, dv \\
&= \pi_0(\bar{s},N) + \int_{\underline{v}}^{\bar{v}} \alpha(y)\cdot[1-F(y|s_i)]\cdot F_w^*(y;\bar{s})^{N-1} dy,
\end{aligned}$$

where the second line follows from Proposition 1 and the third follows from integration by parts. The key properties of this *ex ante* profit function are outlined in Gentry and Li (2011b); we restate these as a lemma below.

Lemma 1. Ex ante *expected Stage 2 profit for a bidder with Stage 1 signal s_i facing competition structure $(\bar{s},N)$ is*

$$\Pi(s_i;\bar{s},N) = \pi_0(\bar{s},N) + \int_{\underline{\mathrm{V}}}^{\bar{v}} \alpha(y)\cdot[1-F(y|s_i)]\cdot F_w^*(y;\bar{s})^{N-1} dy. \tag{3}$$

This function is weakly increasing in s_i for all $(\bar{s},N)$, strictly decreasing in $\bar{s}$ for all (s_i,N), and strictly decreasing in N for all s_i and any $\bar{s}<1$.

Bidder i will choose to enter whenever expected net profit from entry is positive; i.e. whenever

$$\Pi(s_i;\bar{s},N) \geq c.$$

For threshold $\bar{s}\in(0,1)$ to constitute a Stage 1 entry equilibrium, it must satisfy a standard breakeven condition: a bidder drawing signal $S_i=\bar{s}$ must be indifferent to entry when facing N potential bidders who also enter according to $\bar{s}$. Proposition 2 formally outlines the properties of this equilibrium.

Proposition 2. *A symmetric entry equilibrium in the AS model is characterized by a signal threshold $\bar{s}$ such that only bidders with $s_i\geq\bar{s}$ choose to enter. This signal threshold is determined as follows.*

- *If $\Pi(0;0,N)>c$, then $\bar{s}=0$ and all potential bidders always enter.*
- *If $\Pi(1;1,N)<c$, then $\bar{s}=0$ and no potential bidder ever enters.*
- *Otherwise, the signal threshold $\bar{s}$ satisfies the breakeven condition*

$$\Pi(\bar{s};\bar{s},N) \equiv c, \tag{4}$$

where $\Pi(s_i;\bar{s},N)$ is defined as in Lemma 1.

Further, considered as a function of (c, N), the equilibrium threshold $\bar{s}(c, N)$ satisfies the following monotonicity properties:

- *For any $N \geq 1$, $\bar{s}(c, N)$ is continuous and weakly increasing in c, with strict monotonicity whenever $\bar{s}(c, N) \in (0, 1)$.*
- *For any $c \geq 0$, $N' > N$ implies $\bar{s}(c, N') \geq \bar{s}(c, N)$. If in addition $\bar{s}(c, N) \in (0, 1)$, then $\bar{s}(c, N') > \bar{s}(c, N)$ and $\bar{s}(c, N') \in (0, 1)$.*

Taken together, Propositions 1 and 2 characterize the unique symmetric Bayesian Nash equilibrium of the general AS model under any Stage 2 RS auction rules. Proposition 2 in particular will be crucial in establishing (partial) identification of AS-model fundamentals.

3 Econometrics: Identification and nonidentification

Full nonparametric identification in the IPV AS model involves recovery of two fundamentals: the joint signal-value distribution $F(v, s)$ and the entry cost c. Unfortunately, the rich informational structure permitted by the AS model comes at a significant practical cost: the AS model is nonparametrically nonidentified in general. Intuitively, this is because the general AS model imposes only weak restrictions on the mechanism by which bidders are selected. Hence there may be no one-to-one map from characteristics of entering bidders back to fundamentals of the model.

Our key insight is that exogenous variation in the entry threshold $\bar{s}$ generates usable information on the fundamentals of the AS model. The precision of this information depends on the nature of variation in $\bar{s}$: the AS model will be *partially identified* in DGPs where equilbrium $\bar{s}$ takes a discrete set of values, but may be *exactly identified* when equilibrium $\bar{s}$ takes a continuum of values. Our core contribution in this paper is to derive natural identified bounds on AS-model fundamentals that correspond to any set of identified entry thresholds.

Throughout this section, we assume that the econometrician has access to a large sample of auctions from some AS auction process $\mathcal{L}$. For each auction ℓ, the following variables are observed: number of potential bidders N_ℓ, number of actual bidders n_ℓ, and a vector of submitted bids $\mathbf{b}_\ell$. This structure is standard in the literature on auctions with entry: in applications, N_l is typically proxied by variables such as number of planholders (e.g. Li and Zheng (2009)) or number of bidders in related auctions (e.g. Roberts and Sweeting (2010a)) and n_ℓ is taken to be the number of bids submitted. Our identification argument requires three additional restrictions on the DGP $\mathcal{L}$, which are formalized below.

First, our primary goal is to translate existing results on identification in auctions without entry into partial identification results applicable to auctions with arbitrarily selective entry. Consequently, in this section we restrict attention to RS auction mechanisms which are *Stage 2 identified:*

Definition (Stage-2 identified)**.** RS mechanism M is *Stage-2 identified* if, for any marginal distribution $F^*(v)$ and any $n > 1$, a sample of observed bids $\mathbf{b}_\ell$ generated by n bidders competing (without entry) under M based on draws from $F^*(v)$ would permit consistent nonparametric estimation of $F^*(v)$.

Assumption 6 (Stage 2 identification)**.** *Process $\mathcal{L}$ involves an auction mechanism M that is Stage 2 identified.*

This focus on Stage 2 identified mechanisms is natural given our objectives: in general, absent Stage 2 identification, the question of Stage 1 identification is not likely to be interesting.[10]

Second, as noted above, our identification results turn on the presence of variation in the entry threshold $\bar{s}$: intuitively, in order for objects identified in Stage 2 to convey meaningful information on the joint distribution $F(v, s|x)$, at least some variation in entry behavior must be induced by variation in factors other than $F(v, s|x)$. We therefore introduce a key exclusion restriction:

Assumption 7 (Exogenous entry variation)**.** *Process $\mathcal{L}$ involves either (i) exogenous variation in N_ℓ for fixed (x, z), or (ii) exogenous variation in z_ℓ for fixed (N, x), or both.*

Exogenous variation in number of potential bidders has been exploited in prior work for testing purposes: for instance, Haile, Hong, and Shum (2003) use variation in N_ℓ to construct a test for common values, and MSX use variation in N_ℓ to test competing entry models. Exogenous variation in cost shifters z_ℓ directly extends a long tradition of instrumental variables in econometrics.[11] Both are sources of exogenous variation in entry behavior, which we in turn exploit as a source of (partially) identifying information on the AS model.

Finally, for current purposes, we assume that process $\mathcal{L}$ involves no unobserved auction-level heterogeneity:

Assumption 8. *For each auction ℓ generated by process $\mathcal{L}$, the distribution $F_\ell(v, s)$ depends (at most) on an observable $J \times 1$ vector x_ℓ, the entry cost c_ℓ depends (at most) on an observable $K \times 1$ vector z_ℓ, and this dependence is continuous (and monotonic for $c(\cdot)$) in all continuous covariates. That is, for all ℓ, $F_\ell(v, s) = F(v, s|x_\ell)$ and $c_\ell(\cdot) = c(z_\ell)$, where elements of x and z may overlap.*

[10] An interesting question which we do not directly address is how to apply our method to cases where the Stage 2 distribution is partially identified (e.g. Haile and Tamer (2003)). This should be a relatively straightforward extension of the method, but would substantially complicate notation and discussion.

[11] Recent work using instrumental variables to address identification of nonparametric models includes Chesher (2005) for nonparametric identification of models with discrete endogenous variables and Berry and Haile (2010) for nonparametric identification of multinomial choice demand models, to name only a few.

This assumption is obviously strong, and warrants further discussion. Krasnokutskaya (2009) has established identification in auctions with unobserved heterogeneity but without entry. Even in this case, unobserved auction-level heterogeneity introduces significant challenges, and identification turns on strong additional shape restrictions. Introducing endogenous and selective entry would complicate Krasnokutskaya's identification problem exponentially, since *ex post* value distributions would then depend on both unobserved auction-level heterogeneity (via an unobserved value shifter) and unobserved bidder-level heterogeneity (via unobserved Stage 1 signals). Further, while auction-level heterogeneity can be modeled simply (additively or multiplicatively), entry-induced heterogeneity will typically affect observed distributions via the complex nonlinear relationship described in Proposition 2. Consequently, even strong shape restrictions as in Krasnokutskaya (2009) are unlikely to permit simultaneous identification of both auction- and entry-level unobserved heterogeneity.

Given an additional *separation restriction*, however, many of the identification results outlined below will extend even to the case of unobserved auction-level heterogeneity. In particular, suppose that auction-level heterogeneity is realized only in Stage 2 – a structure which we find plausible in applications where c is interpreted as a learning cost.[12] Then auction-level heterogeneity will affect bidding but not entry, which in turn will permit distinction between entry-related and heterogeneity-related effects on the *ex post* value distribution. We will return to the details of this argument as an extension, but for now simply note that the possibility exists; for expositional clarity, our main results will be framed under Assumption 8.

3.1 Identified objects

In principle, given a large enough sample from a process $\mathcal{L}$ satisfying Assumptions 1-8 above, we can always identify at least two classes of statistical object from the data. First, using standard nonparametric regression methods, we can directly identify the equilibrium entry threshold $\bar{s}_x(z, N)$ corresponding to each $(x, z, N) \in \mathcal{L}$:

$$\bar{s}_x(z, N) \equiv \frac{1}{N} E[n_\ell | x, z, N].$$

The set of thresholds $\bar{s}$ thus identified will play a crucial role in our subsequent analysis. Consequently, we introduce some additional notation for this identified set: for any x, let

$$\mathcal{S}_x(\mathcal{L}) \equiv \{s \in [0, 1] | s = \bar{s}_x(z, N) \text{ for some } (z, N) \in \mathcal{L}(x)\}.$$

[12] For instance, in highway construction contract auctions, many features likely to be associated with unobserved heterogeneity (e.g. specifics of the work to be performed, soil type, etc) will only be revealed to bidders via detailed study of the plans and location for the project in question. We interpret "entry" in this context as the decision to undertake this detailed study. Hence the assumption that unobserved heterogeneity is revealed only after entry would seem highly consistent with the underlying economic motivation of the model.

$\mathcal{S}_x(\mathcal{L})$ then characterizes exogenous variation in entry behavior at fundamentals x: i.e. all variation available to (partially) identify the *ex ante* joint distribution $F(v, s|x)$. For notational compactness, we state all subsequent identification results in terms of properties of $S_x(\mathcal{L})$.

Second, by hypothesis, the mechanism M is Stage 2 identified. We can thus recover the value distribution $F^*(\cdot|x, z, N)$ that corresponds to bids submitted at each $(x, z, N) \in \mathcal{L}$. Given the structure of the AS model, we interpret this as the distribution of values among *entrants* at (x, z, N): i.e. the distribution of values conditional on drawing $S_i \geq \bar{s}_x(z, N)$ by Proposition 2. The class of distributions directly identified by bids in process $\mathcal{L}$ are thus the *ex post* distributions

$$F^*(v; \bar{s}|x) \equiv F(v|S_i \geq \bar{s}, x)$$

for each identified entry threshold $\bar{s} \in \mathcal{S}_x(\mathcal{L})$.

In what follows, we abstract from variation in distribution-related covariates x. This is purely for notational convenience; all derivations below can be repeated conditional on x. Hence all identification results stated below immediately generalize to the case of arbitrarily many covariates, though estimation in such cases would obviously be subject to standard curse-of-dimensionality concerns.

3.2 Nonidentified fundamentals

By definition, identification in the IPV AS model involves recovery of two fundamentals: the joint signal-value distribution $F(v, s)$ and the entry cost c. Meanwhile, as discussed in the last subsection, the class of distributions directly identified are the *ex post* distributions conditional on entry: $F^*(v; \bar{s}) \equiv F(v|S_i \geq \bar{s})$ for all $\bar{s} \in \mathcal{S}(\mathcal{L})$. In general, this class of distributions is related to the true joint distribution $F(v, s)$ by the identity

$$F(v, s) = F^*(v; 0) - (1 - \bar{s})F^*(v; \bar{s}). \tag{5}$$

This relationship (derived in the Appendix) immediately implies our first core identification result:

Lemma 2. *Under Assumptions 1-7, $F(v, s)$ is fully identified (at x) if and only if $cl(\mathcal{S}(\mathcal{L})) = [0, 1]$.*

This lemma establishes the AS-model equivalent of a full-support condition: if we observe data generated at every possible entry threshold $\bar{s} \in [0, 1)$, we can fully recover $F(v, s)$. Unfortunately, in many applications of interest, there may be no plausible excludable cost shifter z_ℓ. In such cases, all informative variation in $\bar{s}$ will be driven by variation in N_ℓ, so $\mathcal{S}(\mathcal{L})$ will necessarily be a finite set and the joint distribution $F(v, s)$ will not be point-identified.

In turn, non-identification of $F(v, s)$ implies non-identification of c. As outlined in Lemma 1, each (nontrivial) entry equilibrium $\bar{s} \in \mathcal{S}(\mathcal{L})$ will involve a breakeven condition of the form

$$E_v[\pi(v;\bar{s},N)|S_i=\bar{s}] \equiv c.$$

If the conditional distribution $F(v|\bar{s})$ were known, the expectation on the left could be calculated, and c would be identified. By definition, however, $F(v|\bar{s}) \equiv \partial F(v,\bar{s})/\partial\bar{s}$ depends on local properties of $F(v,s)$, and such local properties will not be identified when $\mathcal{S}_x(\mathcal{L})$ is a finite set. Hence in many cases c also will not be identified.

Having thus outlined the fundamental nonidentification problem in the general AS model, it is interesting to return to the question of identification in the LS and S polar cases. The LS model assumes that potential bidders have no specific information about their values prior to entry, which is equivalent to setting $s_i \perp v_i$ in the AS model. We then have $F^*(v;s) = F(v|s) = F_v(v)$ for any s, so the LS model is identified. On the other hand, the S model assumes that potential bidders know their values exactly prior to entry, which is heuristically equivalent to assuming that v_i is a deterministic function of s_i.[13] The identified distribution $F^*(v;s)$ will then be a truncation of the true distribution $F_v(v)$, so we can use the known functional relationship between a true distribution and its truncated counterpart to recover $F_v(v)$ above the truncation point. Thus, in essence, the LS model obtains identification by assuming away selection, while the S model obtains identification by assuming perfect selection. Meanwhile, the AS model permits much more general selection behavior, but this flexibility comes at the cost of exact identification.

4 Bounds on fundamentals in the AS model

Subsection 3.1 outlined the objects directly identified by a sample from AS process $\mathcal{L}$, and Subsection 3.2 outlined why these objects may not point-identify the fundamentals of interest in the general AS model. This section focuses on a more positive question: what do the objects we *can* identify tell us about the underlying auction process? In the spirit of Haile and Tamer (2003), our answer to this question is based on the principle of partial identification: since exact identification is often impossible, we instead seek to obtain the best available *bounds* on the fundamentals of interest. As above, the objects of interest are the joint signal-value distribution $F(v,s)$ and the entry cost c.

4.1 Bounds on distributions: $F(v|s)$ and $F(v,s)$

As noted in Subsection 3.1, a sample from process $\mathcal{L}$ will directly identify two classes of object: a set of equilibrium entry thresholds $\mathcal{S}$, and an *ex post* distribution $F^*(v;\bar{s})$ for each $\bar{s} \in \mathcal{S}$.

[13] In particular, to preserve the normalization $S \sim U[0,1]$, we would set $s_i = F_v^{-1}(v_i)$.

To establish the main results in this section, we will need some additional notation. Define functions $s^+(t)$ and $s^-(t)$ as follows:

$$\begin{aligned} s^+(t) &= \inf\{\bar{s} \in \mathcal{S} \cup \{1\} | \bar{s} > t\} \\ s^-(t) &= \sup\{\bar{s} \in \mathcal{S} \cup \{0\} | \bar{s} < t\}. \end{aligned}$$

The construction here is somewhat involved, but the underlying idea is actually very simple: given any $t \in [0, 1]$, return the nearest upper and lower neighbors of t in the identified set $\mathcal{S}$. The specific forms chosen are intended to compactly summarize three important special cases: if $\mathcal{S}$ is a set of discrete points, then $s^+(t)$ returns the nearest identified threshold strictly greater than t; if $\mathcal{S}$ contains an interval and $t \in \text{int}(\mathcal{S})$, then $s^+(t) \equiv t$; and if $t \geq \max\{\mathcal{S}\}$, then $s^+(t) \equiv 1$ (and conversely for $s^-(t)$).

The next lemma establishes bounds on $F(v|\bar{s})$ for each threshold $\bar{s}$ in the identified set $\mathcal{S}$.

Lemma 3. *Choose any $\bar{s} \in \mathcal{S}$, let $s^+(\cdot)$ and $s^-(\cdot)$ be as above, and define $\check{F}^+(v|\bar{s})$ and $\check{F}^-(v|\bar{s})$ as follows.*

$$\check{F}^+(v|\bar{s}) = \begin{cases} \lim_{t \uparrow s^-(\bar{s})} \left\{ \frac{(1-t)F^*(v;t) - (1-\bar{s})F^*(v;\bar{s})}{\bar{s}-t} \right\} & \text{if } s^-(\bar{s}) \in \mathcal{S}; \\ 1 & \text{otherwise.} \end{cases}$$

$$\check{F}^-(v|\bar{s}) = \begin{cases} \lim_{t \downarrow s^+(\bar{s})} \left\{ \frac{(1-\bar{s})F^*(v;\bar{s}) - (1-t)F^*(v;t)}{t-\bar{s}} \right\} & \text{if } s^+(\bar{s}) \in \mathcal{S}; \\ 0 & \text{otherwise.} \end{cases}$$

Then $\check{F}^+(v|\bar{s})$ and $\check{F}^-(v|\bar{s})$ are identified for all v, represent distributions over $[\underline{v}, \bar{v}]$, and bound the conditional distribution $F(v|\bar{s})$:

$$\check{F}^+(v|\bar{s}) \geq F(v|\bar{s}) \geq \check{F}^-(v|\bar{s}) \, \forall v,$$

with equality whenever $\bar{s} \in int(\mathcal{S})$.

To explore the intuition behind this lemma, consider a heuristic derivation of $\check{F}^+(v|\bar{s})$ and $\check{F}^-(v|\bar{s})$. By definition, the distribution of values identified by bids submitted at threshold $\bar{s}$ is $F^*(v; \bar{s}) \equiv F(v|S_i \geq \bar{s})$. Given the normalization $S_i \sim U[0, 1]$, this observed distribution can be written in terms of the conditional distribution $F(v|s)$ as follows:

$$F^*(v; s) = \frac{1}{1-\bar{s}} \int_{\bar{s}}^1 F(v|t) dt.$$

Rearranging this identity gives $(1 - \bar{s})F^*(v; s) = \int_{\bar{s}}^1 F(v|t)dt$, which immediately implies that

$$F(v|\bar{s}) = -\frac{\partial}{\partial \bar{s}}[(1 - \bar{s})F^*(v; \bar{s})].$$

We can then approximate the RHS derivative via a small finite difference $\Delta\bar{s}$:

$$F(v|\bar{s}) \approx -\frac{\Delta[(1-\bar{s})F^*(v;\bar{s})]}{\Delta\bar{s}}$$

In practice, the smallest differences $\Delta\bar{s}$ we can recover will be $\Delta\bar{s} = [s^+(\bar{s}) - \bar{s}]$ and $\Delta\bar{s} = [\bar{s} - s^-(\bar{s})]$. These approximations yield $\check{F}^+(v|\bar{s})$ and $\check{F}^-(v|\bar{s})$ defined above. $\check{F}^+(v|\bar{s}) \geq F(v|s) \geq \check{F}^-(v|\bar{s})$ follows from affiliation, and if $\bar{s} \in \text{int}(\mathcal{S})$ we can let $\Delta\bar{s} \to 0$ and the corresponding approximations become exact. Hence $\check{F}^+(v|\bar{s})$ and $\check{F}^-(v|\bar{s})$ represent natural bounds on the true conditional distribution $F(v|\bar{s})$, and the statement in Lemma 3 follows.

As defined in Lemma 3, however, $\check{F}^+(v|\bar{s})$ and $\check{F}^-(v|\bar{s})$ bound $F(v|\bar{s})$ only for $\bar{s} \in \mathcal{S}$. The next proposition extends these functions into bounds on $F(v|s)$ for any $s \in [0,1]$.

Proposition 3. *Choose any $v \in [\underline{v}, \bar{v}]$, and define functions $F^+(v|\cdot)$ and $F^-(v|\cdot)$ as follows:*

$$F^+(v|t) = \begin{cases} \check{F}^+(v|t) & \textit{if } t \in \mathcal{S}; \\ \check{F}^+[v|s^-(t)] & \textit{if } t \notin \mathcal{S}. \end{cases}$$

$$F^-(v|t) = \begin{cases} \check{F}^-(v|t) & \text{if } t \in \mathcal{S}; \\ \check{F}^-[v|s^+(t)] & \text{if } t \notin \mathcal{S}. \end{cases}$$

Then for any $s \in [0,1]$, $F^+(v|s)$ and $F^-(v|s)$ are identified, represent distributions over $[\underline{v}, \bar{v}]$, and bound $F(v|s)$:

$$F^+(v|s) \geq F(v|s) \geq F^-(v|s),$$

with equality whenever $s \in int(\mathcal{S})$.

This proposition naturally extends the logic above: given any $s \in [0,1]$, find the nearest identified neighbors $\bar{s} \in \mathcal{S}$. By Lemma 3, we can obtain identified bounds on $F(v|\cdot)$ at these neighbors, and by affiliation these bounds will also apply to $F(v|s)$. When local variation in $\bar{s}$ is available, the derivative approximations become exact, and exact identification follows. For clarity, we restate this latter fact as a corollary:

Corollary 1. *$F(v|\bar{s})$ is exactly identified for any $\bar{s} \in int(\mathcal{S}(\mathcal{L}))$.*

Finally, since the conditional distribution $F(v|s)$ is directly related to the joint density $F(v,s)$, the identified bounds $F^+(v|s)$ and $F^-(v|s)$ immediately imply identified bounds on $F(v,s)$:

Corollary 2. *Define $F^+(v,s)$ and $F^-(v,s)$ as follows:*

$$F^+(v,s) = \int_0^s F^+(v|t)dt$$

$$F^-(v,s) = \int_0^s F^-(v|t)dt.$$

Then $F^+(v,s)$ *and* $F^-(v,s)$ *are identified and* $F^+(v,s) \geq F(v,s) \geq F^-(v,s)$.[14]

We can thus construct nonparametric bounds on $F(v|s)$ and $F(v,s)$ under any Stage 2 identified RS auction mechanism.

4.2 Bounds on $c(z)$

We translate bounds on $F(v|\bar{s})$ into bounds on $c(z)$ using the characterization of Stage 1 equilibrium given in Proposition 2. In particular, given any z and any nontrivial equilibrium threshold $\bar{s}_N(z)$, we know $c(z)$ must satisfy a breakeven condition of the form

$$\begin{aligned} c(z) &\equiv E_v[\pi(v;\bar{s}_N(z),N)|S_i = \bar{s}_N(z)] \\ &= \pi_0(\bar{s}(z,N),N) + \int_{\underline{v}}^{\bar{v}} \alpha(y)\cdot[1 - F(y|\bar{s}(z,N))]\cdot F_w^*(y;\bar{s}(z,N))^{N-1}dy. \quad (6) \end{aligned}$$

The only unknown on the RHS of (6) is $F(v|s)$: the distribution $F_w^*(\cdot;\bar{s}) \equiv$ is identified directly from observed entry and bidding decisions for any $\bar{s} \in \mathcal{S}$ and knowledge of the Stage 2 mechanism implies knowledge of $\pi_0(\cdot)$ and $\alpha(\cdot)$. Further, by inspection, the RHS integral is decreasing in $F(y|s)$. Consequently identified bounds on $F(y|s)$ immediately imply identified bounds on $c(z)$ for any N:

Proposition 4. *Choose any* $(N,z) \in \mathcal{L}$, *let* $\bar{s}_N(z)$ *be the (identified) entry threshold at* (N,z), *and define* $c_N^+(z)$ *and* $c_N^-(z)$ *as follows:*

$$c_N^+(z) = \pi_0(\bar{s}_N(z),N) + \int_{\underline{\mathrm{V}}}^{\bar{v}} \alpha(y)\cdot[1 - F^-(y|\bar{s}_N(z))]\cdot F_w^*(y;\bar{s}_N(z))^{N-1}dy.$$

$$c_N^-(z) = \pi_0(\bar{s}_N(z),N) + \int_{\underline{\mathrm{V}}}^{\bar{v}} \alpha(y)\cdot[1 - F^+(y|\bar{s}_N(z))]\cdot F_w^*(y;\bar{s}_N(z))^{N-1}dy.$$

Then $c_N^+(z)$ *and* $c_N^-(z)$ *are identified and* $c_N^+(z) \geq c(z) \geq c_N^-(z)$, *with equality if* $\bar{s}_N(z) \in int(\mathcal{S}(\mathcal{L}))$.

[14] If desired, we could extend this proposition to incorporate the Frechet-Hoeffding bounds (see Nelsen (1999)):

$$F_v(v) + s - 1 \leq F(v,s) \leq \min\{F_v(v), s\}.$$

In practice, $F_v(v)$ may be unknown, but can be bounded as follows:

$$F^*(v;\min\mathcal{S}) \leq F_v(v) \leq \min\mathcal{S} + (1-\min\mathcal{S})F^*(v;\min\mathcal{S}).$$

Combining these results yields alternative bounds on $F(v,s)$, which in some cases might be tighter than those above.

In turn, pooling these bounds across N will imply sharper bounds on c:

Corollary 3. *For any z, we can pool identified bounds across N to obtain sharper bounds on $c(z)$:*

$$c^+(z) \equiv \min_{N \in \mathcal{N}} c_N^+(z) \geq c(z) \geq \max_{N \in \mathcal{N}} c_N^-(z) \equiv c^-(z).$$

Finally, from Corollary 1 above, we know that $F(v|\bar{s})$ is exactly identified whenever $\bar{s} \in \text{int}(\mathcal{S}(\mathcal{L}))$. Equation (6) then implies an analogous condition for exact identification of entry costs:

Corollary 4. *For any z such that $\bar{s}_N(z) \in int(\mathcal{S}(\mathcal{L}))$ for some N, $c(z)$ is exactly identified.*

4.3 Full identification

Taken together, Corollaries 1, 2, and 4 reinforce the conclusion of Lemma 2: the AS model is exactly identified (almost) everywhere if and only if we observe data generated at (almost) every $\bar{s} \in [0, 1]$. The conditions under which this will occur will depend on the nature of the underlying fundamentals, but (roughly) will require an excluded instrument z that induces sufficient variation in entry cost $c(z)$. The next proposition formalizes this intuition.

Proposition 5. *Suppose the econometrician observes a cost shifter z satisfying Assumptions 8 and 7 above, which has positive support on a set $Z \subset R^k$. Then the following statements hold:*

1. *If $z \in int(Z)$, then $c(z)$ is identified and $F(v|\bar{s})$ is locally identified at each $\bar{s} \in \mathcal{S}(\mathcal{L}(z))$.*

2. *If $Z = R^k$ and the range of $c(\cdot)$ is unbounded in R^+, then $F(v|s)$, $F(v, s)$, and $c(\cdot)$ are fully identified.*

The sufficient Condition 2 can probably be relaxed somewhat in many applications: as noted above, the fundamental property needed for full identification is $\text{cl}(\mathcal{S}(\mathcal{L})) = [0, 1]$. Condition 2 merely ensures that this property will hold absent further restrictions on N and $F(v, s)$.

4.4 Sharp bounds

One key message of Proposition 5 is that full identification in the general AS model depends on strong conditions: we require both an excludable cost shifter z and a sufficiently variable cost function $c(\cdot)$, either or both of which can easily fail in practice. In such cases, nonparametric analysis must fall back on bounds like those we derive in subsections 4.1 and 4.2. As defined above, however, these bounds may not be sharp: they represent natural, intuitive, and directly estimable approximations to the fundamentals of interest, but may not fully exhaust all variation in the data. This subsection formally characterize sharp bounds in the AS model.

Begin by defining a *candidate model* corresponding to process $\mathcal{L}$ as follows:

Definition 2 (Candidate model)**.** A *candidate model* for process $\mathcal{L}$ is any pair $\{\tilde{F}(\cdot|\cdot), \tilde{c}(\cdot)\}$ satisfying the following conditions for all $z \in \mathcal{L}$:

1. **Distribution:** for all $s \in [0,1]$, $\tilde{F}(\cdot|s)$ defines a distribution over $[\underline{v}, \bar{v}]$.

2. **Selection:** $\tilde{F}(\cdot|\cdot)$ implies the set of distributions identified by sub-process $\mathcal{L}(z)$*:* $(1-s)F^*(v;s) = \int_s^1 F(v|t)dt$ *for all* v*, for all* $s \in \mathcal{S}(z)$*.*

3. **Affiliation:** $\tilde{F}(\cdot|\cdot)$ implies a joint distribution $\tilde{F}(v,s) \equiv \int_0^s \tilde{F}(v|t)dt$ satisfying affiliation.

4. **Entry:** $\Pi^*(s, N_s(z); \tilde{F}) \equiv \tilde{c}(z)$ for all $s \in \mathcal{S}(z)$, where $N_s(z)$ denotes the competition level N corresponding to s under $\mathcal{L}(z)$ and

$$\Pi^*(s, N, \tilde{F}) \equiv \int_V \pi(v; s, N) d\tilde{F}(v|s). \tag{7}$$

Taken together, the conditions in Definition 2 exhaust the restrictions generated by the AS model. We thus use Definition 2 to illustrate the two main conclusions of this section: why the bounds derived above may not be sharp, and how (in principle) sharp bounds might be obtained.

To see why the bounds in Propositions 3 and 4 may not be sharp, note that $F^+(v|s)$ and $F^-(v|s)$ directly exploit only the distribution and selection conditions of Definition 2.[15] Hence it is concievable that there could exist a (v,s) pair such that *no* candidate $\tilde{F}(\cdot|\cdot)$ attaining (say) the upper bound $F^+(v|s)$ at (v,s) could simultaneously satisfy the entry condition (4) for all other $s \in \mathcal{S}(z)$. In this case, the bounds given in sections 4.1 and 4.2 would not be sharp.

Extending this intuition, we can conceptually characterize sharp bounds in the AS model as follows. Let $\mathcal{M}(\mathcal{L})$ be the set of all candidate models $\{\tilde{F}(\cdot|\cdot), \tilde{c}(\cdot)\}$ satisfying Definition 2. By assumption, the true model $\{F(v|s), c(\cdot)\}$ satisfies Definition 2, so $\mathcal{M}(\mathcal{L})$ is nonempty. Further, since Definition 2 can be evaluated for each $\{\tilde{F}(\cdot|\cdot), \tilde{c}(\cdot)\}$ given objects identified by $\mathcal{L}$, the set $\mathcal{M}(\mathcal{L})$ is (in principle) identified. Then sharp bounds on AS fundamentals will be given by the upper and lower envelopes of $\mathcal{M}(\mathcal{L})$:

[15] More precisely, we incorporate a slightly weaker "differences" version of the selection condition (2):

$$(1-s)F^*(v;s) - (1-s')F^*(v;s') = \int_s^{s'} F^+(v|t)dt$$

for $s' > s$, with F^- substituted for F^+ when $s' < s$.

Lemma 4. *Let $\mathcal{M}(\mathcal{L})$ be the set of all candidate models at $\mathcal{L}$. Then $\mathcal{M}(\mathcal{L})$ is the sharp identified set at $\mathcal{L}$, and implies sharp bounds $\{\tilde{c}^+, \tilde{c}^-\}$ and $\{\tilde{F}^+, \tilde{F}^-\}$ as follows:*

$$
\begin{aligned}
\tilde{c}^+(z) &= \sup\{\tilde{c}(z) \text{ s.t. } \tilde{c}(\cdot) \in \mathcal{M}(\mathcal{L})\} \\
\tilde{c}^-(z) &= \inf\{\tilde{c}(z) \text{ s.t. } \tilde{c}(\cdot) \in \mathcal{M}(\mathcal{L})\}
\end{aligned}
$$

and for each (v, s)

$$
\begin{aligned}
\tilde{F}^+(v|s) &= \sup\{\tilde{F}(v|s) \text{ s.t. } \tilde{F}(v|s) \in \mathcal{M}(\mathcal{L})\} \\
\tilde{F}^-(v|s) &= \inf\{\tilde{F}(v|s) \text{ s.t. } \tilde{F}(v|s) \in \mathcal{M}(\mathcal{L})\}.
\end{aligned}
$$

This statement follows immediately from construction of $\mathcal{M}(\mathcal{L})$; it is intended to illustrate how the bounds derived in sections 4.1 and 4.2 differ from the conceptually identified sharp bounds. Direct implementation of Lemma 4 would involve evaluating Definition 2 at each member of an infinite functional set, and hence would be impossible in practice. Indirect implementation could perhaps be attempted using sieve or other functional approximation methods, but this is well outside the scope of the current investigation.

Finally, note that our discussion in this section parallels that of Ciliberto and Tamer (2009) on sharpness in oligopoly entry models. In particular, at any $t \in \mathcal{S}$, our bounds on $F(v|t)$ exploits information generated by the nearest-neighbors $s^+(t)$ and $s^-(t)$, but not that potentially generated by more distant entry equilibria. Meanwhile, Ciliberto and Tamer (2009) exploit zero-one bounds on probability magnitudes, but not the condition that probabilities must sum to one. In both cases the condition omitted is a cross-equation restriction which, though potentially informative, would very difficult to implement.[16] Hence in what follows we focus on the directly identified (and directly implementable) bounds in sections 4.1 and 4.2, but note that in principle sharper bounds might exist.

4.5 Bounds in the S and LS special cases

To conclude this section, we explore how our proposed bounds would behave when applied to data generated by the S and LS polar cases. As noted above, the S model is roughly equivalent to the limit case in which Stage 2 values are a deterministic function of Stage 1 signals. In particular, to preserve the normalization $S_i \sim U[0, 1]$, we would set $v_i = F_v^{-1}(s_i)$. In this case, the true conditional distribution $F(v|\bar{s})$ is degenerate, with all mass at $v^* \equiv F_v^{-1}(\bar{s})$. Meanwhile, the bounds $F^+(v|\bar{s})$ and $F^-(v|\bar{s})$ will in general be well-defined distributions. Hence $F^+(v|\bar{s})$ and $F^-(v|\bar{s})$ do not collapse to $F(v|\bar{s})$, and consequently the (derived) bounds on c and $F(v, s)$ also do not collapse.

[16] In particular, in our case, sharp bounds on the distribution function $F(\cdot|s)$ can be characterized only through the integral condition (7), which does not have a tractable functional inverse.

Results for the LS special case are more favorable. Recall that the LS model can be formally nested in the AS model by assuming that Stage 2 values are independent of Stage 1 signals. For any (v, s) independence implies

$$F^*(v; s) = F(v|s) = F_v(v).$$

In turn, this equality implies that $F^+(v|s)$ and $F^-(v|s)$ collapse to the true marginal distribution $F_v(v)$ for non-extremum $s \in \mathcal{S}(\mathcal{L})$. Consequently, $c(\cdot)$ is identified, the lower bound $F^-(v, s)$ collapses to $F(v, s)$ for $s < \max \mathcal{S}$, and differences $F^+(v, s') - F^+(v, s)$ collapse to true values $F(v, s') - F(v, s)$ for $s, s' > \min \mathcal{S}$.[17] Thus, at least in terms of identification, estimation based on the general AS model entails only marginal losses relative to estimation based on the more restrictive LS special case.

5 Bounds on counterfactual revenue

Counterfactual policy simulations are a leading motivation for structural analysis. In the context of RS auctions, one key policy variable is the seller's award rule, which specifies under what circumstances (and with what probabilities) any particular auction will result in a sale. The best-known example of such an award rule is a public reserve price, which represents a particularly interesting application of our results: structural analysis based on the IPV LS model will always find a zero optimal reserve price (see Levin and Smith (1994)) and structural analysis based on the IPV S model will typically find a positive reserve price (see Li and Zheng (2007)), but the AS model does not predetermine policy outcomes. However, several other types of award rules (such as secret reserve prices) are also frequently used in practice. Hence we frame our analysis in terms of a general award rule $\alpha(\cdot)$, but note that all results can trivially be specialized to the (much simpler) case of a public reserve price.

In particular, we start from the set of entry thresholds $S(\mathcal{L})$ identified by process $\mathcal{L}$, and seek to derive bounds on expected seller revenue R_α corresponding to counterfactual award rule $\alpha(\cdot)$. This problem is complicated considerably by the very general nature of the AS entry model: in the presence of endogenous and selective entry, the award rule $\alpha(\cdot)$ will affect seller revenue directly, through the Stage 1 entry threshold $\bar{s}$, and through the selected Stage 2 distribution $F^*(\cdot; \bar{s})$. Valid counterfactual revenue bounds need to account for all three effects. As typical in the literature, we assume that the no-sale outcome yields value $v_0 \leq \underline{v}$ to the seller.

The argument in this section proceeds as follows. First, using theoretical results in our companion paper (Gentry and Li (2011b)), we characterize seller revenue $R_\alpha(N, \bar{s})$ for *any* $\bar{s}$. Second, using the breakeven condition (6) and bounds on fundamentals established in Section 4, we establish bounds on the counterfactual *entry threshold* $\bar{s}_\alpha$ characterizing equilibrium play under award rule $\alpha(\cdot)$. Finally, using the fact that

[17] Since we have no definite upper bound on $F(v|s)$ for $s < \min \mathcal{S}$, the upper bound $F^+(v, s)$ may not collapse.

$R_\alpha(N, \bar{s})$ is decreasing in $\bar{s}$, we translate bounds on $\bar{s}_\alpha$ into bounds on true expected revenue R_α for any (N, z). We thus obtain identified bounds on counterfactual revenue corresponding to a wide range of award rules in a general class of auctions with endogenous and arbitrarily selective entry.

For current purposes, we restrict attention to counterfactual mechanisms satisfying two additional properties:

Assumption 9 (Counterfactual award rule)**.** *The counterfactual mechanism is such that*

1. *The seller never awards inefficiently:* $\alpha(y) = 0$ *for* $y \leq v_0$*.*
2. *Low-type profits* $\pi_0^\alpha(s, N)$ *take the form*

$$\pi_0^\alpha(s, N) = s^{N-1} \int_{v_0}^{\bar{v}} \alpha(y) dy - \rho \tag{8}$$

for some constant $\rho \geq 0$*.*

Condition 1 is standard in the literature: it pins down seller preferences over sale outcomes. Condition 2 looks slightly more obscure, but is satisfied by the vast majority of mechanisms used in practice. For instance, in a standard first-price auction, Condition 2 supports any public reserve price $r \in [v_0, \bar{v}]$, any secret reserve price, and a wide range of other stochastic award rules. Thus, though apparently restrictive, Assumption 9 actually involves minimal loss of generality. The constant ρ can be interpreted as an entry fee.

Given Assumption 9, we can proceed with the argument sketched above. First, building on results in Gentry and Li (2011b), we characterize seller revenue under award rule $\alpha(\cdot)$ at arbitrary entry threshold $\bar{s}$:

Lemma 5. *Under Assumptions 1-5 and 9, expected seller revenue corresponding to award rule* $\alpha(\cdot)$ *at competition structure* $(\bar{s}, N)$ *is given by*

$$R_\alpha(\bar{s}; N) = \int_{v_0}^{\bar{v}} \{\alpha(y)[y - \lambda_\alpha(y; \bar{s}, N)] + [1 - \alpha(y)]v_0\}\, dG^*_{1:N}(y; \bar{s}) + N(1 - \bar{s})\rho \tag{9}$$

where

$$\lambda_\alpha(v; s, N) \equiv \begin{cases} 0 & \textit{if } \alpha(v) = 0; \\ \int_{v_0}^{v} \frac{\alpha(t)}{\alpha(v)} \cdot \frac{F_w^*(t;s)^{N-1}}{F_w^*(v;s)^{N-1}} dt & \textit{otherwise.} \end{cases}$$

Further, considered as a function of $\bar{s}$*,* $R_\alpha(\bar{s}; N)$ *satisfies the following properties:*

1. $R_\alpha(\bar{s}; N)$ *is decreasing in* $\bar{s}$ *for all* $\alpha(\cdot)$ *and* N*.*
2. $R_\alpha(\bar{s}; N)$ *is identified for any* $\bar{s} \in \mathcal{S}(\mathcal{L})$*.*

When entry is endogenous, the equilibrium threshold $\bar{s}_\alpha$ will depend on $\alpha(\cdot)$, and when $F(v|s)$ and $c(\cdot)$ are not identified, we typically will not be able to pin down this dependence exactly. However, we can use the bounds on fundamentals derived in Section 4 to *bound* the relationship between $\bar{s}_\alpha$ and $\alpha(\cdot)$:

Lemma 6. *Choose any (z, N) and let $c^+(z)$ and $c^-(z)$ be identified bounds on $c(z)$, $F^+(v|\bar{s})$ and $F^-(v|\bar{s})$ be identified bounds on $F(v|s)$, and $\bar{s}_\alpha(z, N)$ be the (unknown) equilibrium entry threshold under counterfactual award rule $\alpha(\cdot)$. Define $s^+_\alpha(z, N)$ and $s^-_\alpha(z, N)$ as follows:*

$$s^+_\alpha(z, N) = \begin{cases} \inf\{s \in \mathcal{S} | \Pi_\alpha(s, N; F^+) > c^+(z)\} & \text{if } \exists \text{such } s; \\ 1 & \text{otherwise} \end{cases}$$

$$s^-_\alpha(z, N) = \begin{cases} \sup\{s \in \mathcal{S} | \Pi_\alpha(s, N; F^-) < c^-(z)\} & \text{if } \exists \text{such } s; \\ 0 & \text{otherwise} \end{cases}$$

where $\Pi_\alpha(\cdot; \tilde{F})$ denotes expected breakeven profit at $\alpha(\cdot)$ implied by conditional distribution $\tilde{F}$:

$$\Pi_\alpha(s, N; \tilde{F}) = \int_{v_0}^{\bar{v}} [1 - \tilde{F}(y|s)]\alpha(y) F^*_w(y; s)^{N-1} dy - \rho.$$

Then $s^+_\alpha(z, N)$ and $s^-_\alpha(z, N)$ are identified and $s^+_\alpha(z, N) \geq \bar{s}_\alpha(z, N) \geq s^-_\alpha(z, N)$, with equality if $s_\alpha(z, N) \in int(\mathcal{S}(\mathcal{L}))$.

Finally, taken together, Lemmas 5 and 6 imply identified bounds on true expected seller revenue $R_\alpha(z, N)$ corresponding to counterfactual award rule $\alpha(\cdot)$:

Proposition 6. *Choose any (z, N), define $s^+_\alpha(z, N)$ and $s^-_\alpha(z, N)$ as in Lemma 6, and let $R_\alpha(z, N)$ be (unknown) expected revenue under $\alpha(\cdot)$ at (z, N). Define $R^+_\alpha(z, N)$ and $R^-_\alpha(z, N)$ as follows:*

$$R^-_\alpha(z, N) = \begin{cases} R_\alpha(s^+_\alpha(z, N); N) & \text{if } s^+_\alpha(z, N) \in \mathcal{S}(\mathcal{L}) \\ 0 & \text{otherwise} \end{cases}$$

$$R^+_\alpha(z, N) = \begin{cases} R_\alpha(s^-_\alpha(z, N); N) & \text{if } s^-_\alpha(z, N) \in \mathcal{S}(\mathcal{L}) \\ \check{R}_\alpha(0; N) & \text{otherwise,} \end{cases}$$

where

$$\check{R}_\alpha(0; N) = \int_{v_0}^{\bar{v}} \left\{ \alpha(y)[y - \int_{v_0}^{y} \frac{\alpha(t)}{\alpha(y)} \cdot \frac{F^*_w(t; \min \mathcal{S})^{N-1}}{F^*_w(y; \min \mathcal{S})^{N-1}}]dt + [1 - \alpha(y)]v_0 \right\} dG^*_{1:N}(y; \min \mathcal{S}) + N\rho$$

is a semi-informative upper bound applicable when $s^-_\alpha(z, N) \equiv 0$.[18]

Then $R^+_\alpha(z, N)$ and $R^-_\alpha(z, N)$ are identified and $R^+_\alpha(z, N) \geq R_\alpha(z, N) \geq R^-_\alpha(z, N)$, with equality if $s_\alpha(z, N) \in int(\mathcal{S}(\mathcal{L}))$.

[18] In particular, $\check{R}_\alpha(0; N)$ is the revenue that would result if all potential bidders always enter but draw values from distribution $F^*(\cdot; \min \mathcal{S})$. Since $\min \mathcal{S} \geq 0$, we know $F^*(v; \min \mathcal{S}) \leq F^*(v; 0)$, so $\check{R}_\alpha(0; N) \geq R_\alpha(0; N) \geq R_\alpha(s_\alpha; N)$. Further, since $\min \mathcal{S} \in \mathcal{S}$, $\check{R}(0; N)$ is identified.

The intuition behind this result is straightforward: Lemma 5 establishes that conditional revenue is decreasing in $\bar{s}$ for any $\alpha(\cdot)$, Lemma 6 establishes bounds s_α^+ and s_α^- on the counterfactual entry threshold $\bar{s}_\alpha$, and we know $R_\alpha(\bar{s}, N)$ is identified for any $\bar{s} \in \mathcal{S}(\mathcal{L})$. Hence when $s_\alpha^+ \in \mathcal{S}(\mathcal{L})$, $R_\alpha(s_\alpha^+; N)$ gives an identified lower bound on $R_\alpha(\bar{s}_\alpha; N)$ (and conversely for s_α^-). When either $s_\alpha^+ \notin \mathcal{S}(\mathcal{L})$ or $s_\alpha^- \notin \mathcal{S}(\mathcal{L})$, the corresponding $R_\alpha(s_\alpha^\pm; N)$ is not identified, so we fall back on less informative bounds 0 and $\check{R}_\alpha$ respectively. We thus obtain identified bounds on expected revenue under a wide range of counterfactual award rules $\alpha(\cdot)$.

Revenue bounds applicable to the special case of a public reserve price can easily be obtained by setting $\alpha(y) \equiv \mathbf{1}[y \geq r]$ in Proposition 6. In some applications, however, researchers may wish to characterize not just expected seller revenue but also the seller's *optimal reserve price* (ORP). Consequently, following Haile and Tamer (2003), we translate the revenue bounds above into bounds on the seller's ORP:

Corollary 5 (Bounds on Seller's ORP)**.** *Let $R^+(r; z, N)$ and $R^-(r; z, N)$ be revenue bounds derived from Proposition 6 (with $\alpha_r(y) \equiv \mathbf{1}[y \geq r]$), $R_-^* \equiv \sup_r R^-(r; z, N)$ be the maximum value attained by the lower revenue bound, and $r^*(z, N) = \arg\max R(r; z, N)$ be the seller's true optimal reserve price at (z, N). Define $r_+^*(z, N)$ and $r_-^*(z, N)$ as follows:*

$$\begin{aligned} r_+^*(z, N) &= \sup\{r | R^+(r; z, N) \geq R_-^*\} \\ r_-^*(z, N) &= \inf\{r | R^+(r; z, N) \geq R_-^*\}. \end{aligned}$$

Then r_+^* and r_-^* are identified and $r_+^*(z, N) \geq r^*(z, N) \geq r_-^*(z, N)$, with equality if $s_{r^*}(z, N) \in int(\mathcal{S}(\mathcal{L}))$.

Thus, to summarize: we obtain identified bounds on expected revenue under a wide range of counterfactual award rules $\alpha(\cdot)$, which apply to the general class of RS auctions and account for endogenous and arbitrarily selective entry. To our knowledge, these are the first such results reported in the literature. Further, in the special case of a public reserve price, these revenue bounds can be translated into bounds on the seller's optimal reserve price following Haile and Tamer (2003). We thus establish that the general AS model can support a rich variety of counterfactual and policy analyses under relatively weak assumptions on the nature of entry and selection.

6 Conclusion

Though entry is clearly an important feature of many real-world auction markets, there is still no clear consensus in the literature regarding how to account for entry in the structural analysis of auction data. Most applied research seeking to incorporate entry does so via strong informational assumptions: bidders either know nothing about their values prior to entry (LS model) or know their values exactly prior to entry (S model). Recognizing the limitations of these polar-case approaches, several recent papers explore a more general framework in which bidders observe signals affiliated with their true values prior to entry, which we call the AS model. This more general AS framework is theoretically attractive: it imposes few *a priori* restrictions on information structure, serves as a natural bridge between the S and LS models, and does not constrain policy parameters such as the optimal reserve price. Unfortunately, however, the AS model is nonparametrically non-identified. Thus the few existing applications of the AS model obtain identification via parametric functional-form restrictions.

In this paper, we explore an alternative approach: rather than obtain identification via parametric restrictions, we exploit exogenous variation in entry behavior to construct bounds on quantities of interest in the general AS model. In the process, we make three core contributions to the related literature. First, we develop nonparametric bounds on AS model fundamentals applicable to a general class of auctions with endogenous and arbitrarily selective entry. Second, we translate these bounds on fundamentals into bounds on expected revenue corresponding to a wide range of counterfactual award rules, again accounting for endogenous and selective entry. Finally, we outline conditions under which all bounds collapse to exact identification. To our knowledge, these are the first formal (partial) identification results applicable to the general AS model, and represent the most general treatment of identification in auctions with entry to date.

For expositional clarity, our results thus far have been presented for the case of symmetric bidders. However, the underlying logic extends readily to environments with asymmetric bidders. In particular, suppose process $\mathcal{L}$ involves a set $\mathcal{T}$ of potential bidder types, where a bidder of type $\tau_i \in \mathcal{T}$ draws from affiliated joint distribution $F_{\tau_i}(v, s|\cdot)$ and has entry cost $c_{\tau_i}(\cdot)$.[19] For any any competition set τ and any type-symmetric, monotonic Stage 2 equilibrium, there will exist at least one type-symmetric Stage 1 equilibrium $\bar{s}(\tau) = \{\bar{s}_1(\tau), ..., \bar{s}_N(\tau)\}$ such that a bidder with type τ_i will optimally enter whenever $s_i \geq \bar{s}_i(\tau)$ and rivals enter according to $\bar{s}_{-i}(\tau)$. Given an equilibrium selection rule, observed entry decisions will permit identification of the threshold set $\bar{s}(\tau)$ for any competition set τ, and pooling across auctions involving type τ_i will generate a set of type-specific identified thresholds $\mathcal{S}_{\tau_i}(\mathcal{L})$. Results in Section 4 can then be applied to obtain bounds on $F_{\tau_i}(\cdot)$ and $c_{\tau_i}(\cdot)$ for each type τ_i, and partial identification follows. The key additional complication in the asymmetric case is the potential presence of multiple Stage 1 equilibria, and the consequent need to specify an equilibrium selection rule. Insofar as bidder asymmetries induce greater variation in identified threshold sets $\mathcal{S}_{\tau_i}(\mathcal{L})$, however, they may actually improve identification.[20]

[19] These types can be either discrete or continuous; we require only that continuous types affect model fundamentals continuously.

[20] In particular, if bidder types are continuous, the set of identified threshold sets $\mathcal{S}_{\tau_i}(\mathcal{L})$ should also be continuous for any τ_i, and the possibility of exact identification follows by Corollaries 1 and 4.

A second potentially important extension of our results is to environments with unobserved heterogeneity. As noted in our discussion of Assumption 8, this extension requires two additional restrictions. First, following Krasnokutskaya (2009), we assume auction-level unobserved heterogeneity u enters the value distribution either additively or multiplicatively: that is, $v_i = u\epsilon_i$ or $v_i = u + \epsilon_i$, where ϵ_i is IID across bidders.[21] Second, we assume the realization of u is revealed to bidders only after Stage 1 entry is complete. This second assumption seems plausible in applications where c is interpreted as a cost of learning, and permits us to separate heterogeneity induced by the unobserved latent variable u from that induced by unobserved signals s_i.[22] The identification argument can then proceed as follows. First, following Subsection 3.1, we know we can identify the set of equilibrium entry thresholds $\mathcal{S}(\mathcal{L})$ corresponding to process $\mathcal{L}$. For each $\bar{s} \in \mathcal{S}(\mathcal{L})$, the observed bid distribution will depend on two components: the ex-post selected distribution $F^*(\epsilon_i; \bar{s})$ and the distribution of unobserved heterogeneity $F_u(u)$. Second, following Krasnokutskaya (2009), each of these components can be recovered using deconvolution methods on an appropriate sample of observed bids. It follows that the ex-post distribution $F^*(\epsilon_i; \bar{s})$ is identified for any $\bar{s} \in \mathcal{S}(\mathcal{L})$. Finally, we can apply results in Section 4 to the identified ex post distribution $F^*(\epsilon_i; \bar{s})$ to obtain bounds on remaining model fundamentals, and partial identification of the overall model follows immediately.

[21] Krasnokutskaya (2009) assumes u enters multiplicatively, but her argument can easily be adapted to the additive case as well.

[22] For instance, in highway construction auctions, many features likely to generate project-level heterogeneity (e.g. exact location, specifics of work to be performed, soil type, etc) are observable only to bidders obtaining detailed project plans (the outcome typically taken to represent entry). Hence the assumption that u is revealed following entry seems reasonable. Without this separating restriction, identification would require us to successfully disentangle distribution effects directly attributable to heterogeneity (u) from those attributable to changes in endogenous selection (including those induced by shifts in u). In a fully nonparametric model with endogenous and arbitrarily selective entry, this is likely to prove impossible in practice.

Appendix 1: Numerical example

Sections 4 and 5 develop identified bounds for model fundamentals and seller revenue in auctions with arbitrarily selective entry. In this appendix, we explore a simple numeric example designed to illustrate what these theoretical identified bounds might look like in practice. Consistent with our emphasis in the rest of the paper, this example focuses on *identification*, not estimation: the figures that follow illustrate the bounds that would obtain in an infinite auction sample. Nevertheless, this simple exercise should help to indicate what kind of information could in principle be recovered using the methods developed above.

Details of this example are as follows. We model the joint distribution $F(v, s)$ using a Gaussian copula $C_\rho(F_v, F_s)$, where the marginal distribution $F_v(\cdot) \sim N(\mu = 100, \sigma = 10)$ and as above we normalize $F_s(\cdot) \sim U[0, 1]$. The correlation parameter ρ measures the degree of affiliation between s and v, with $\rho = 0$ generating the no-information LS case and $\rho \to 1$ approaching the full-information S case. In what follows, we take $\rho = 0.75$ unless noted otherwise; other values of ρ yield quite similar results. Entry involves cost $c = 2$, and we assume potential competition varies exogenously on the set $N = \{4, 5, ..., 16\}$. These parameter values are chosen to be quanlitatively similar to existing findings in the literature.[23]

Given this parametric specification, it is straightforward to calculate the set of equilibrium entry thresholds $\mathcal{S} = \{\bar{s}_4, ..., \bar{s}_{16}\}$ satisfying the the breakeven condition (4). From subsection 3.1, we know these thresholds (and the corresponding value distributions $F^*(v; \bar{s})$ are the objects identified by a standard $(N, n, \mathbf{b})$ sample. We can then use results in Sections 4 and 5 to obtain identified bounds on quantities of interest as follows.

First, Proposition 3 implies identified bounds on $F(v|s)$ for any (v, s). Figures 1 and 2 illustrate these bounds (across s) for two values of v.

Second, based on Proposition 4, we can translate identified bounds on $F(v|s)$ into identified bounds on c. Applying this proposition to our numeric example and pooling results across N yields identified bounds $c^+ = 2.026$ and $c^- = 1.971$, where (as noted above) true $c = 2$.

Third, following Section 5, we can translate these bounds on $F(v|s)$ and c into identified bounds on counterfactual seller revenue. Results in Section 5 are framed in terms of an arbitrary award rule $\alpha(\cdot)$, but for simplicity we here restrict attention to the special case of a counterfactual public reserve price r. As in Section 5, the first step in this process is to obtain identified bounds on the counterfactual entry threshold $\bar{s}_r$ corresponding to each candidate reserve price r. These bounds are illustrated for $N = 6$ and $N = 9$ in Figures 3 and 4 below.

Finally, using Proposition 6, we can translate bounds on $\bar{s}_r$ to bounds on counterfactual revenue R_r at any (N, r). These bounds are illustrated for $N = 6$ and $N = 9$ in Figures 5 and 6 below. In both cases, the seller's value is assumed to be $v_0 = 60$.

[23] See, e.g., Roberts and Sweeting (2010a; 2010b) and Li and Zheng (2009) for examples.

If so desired, we could also adapt the argument of Haile and Tamer (2003) to translate bound on counterfactual revenue to bounds on the optimal reserve price r as in Corrolary 5. In this example, the implied bounds on optimal r would be rather wide: an upper bound of roughly 90 when $N = 6$, and of roughly 100 when $N = 9$ (with uninformative lower bound $v_0 = 60$ in each case). However, in both cases the identified bounds on counterfactual *revenue* are surprisingly tight.

Figure 1: IDed bounds on $F(v|s)$, $v = 95$

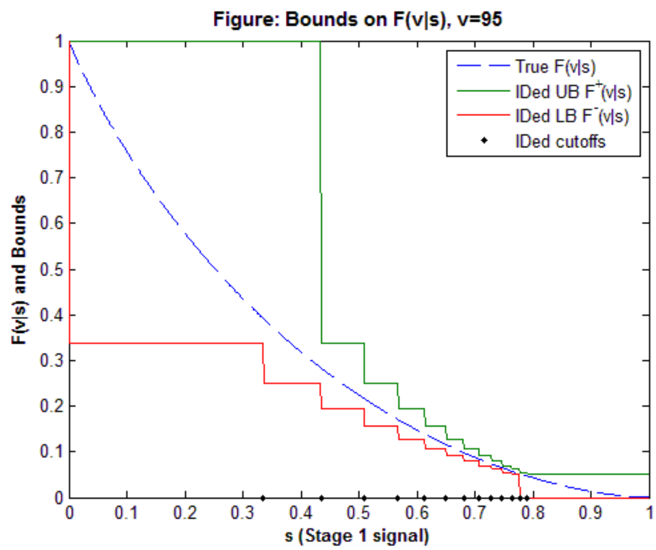


Figure 2: IDed bounds on $F(v|s)$, $v = 105$

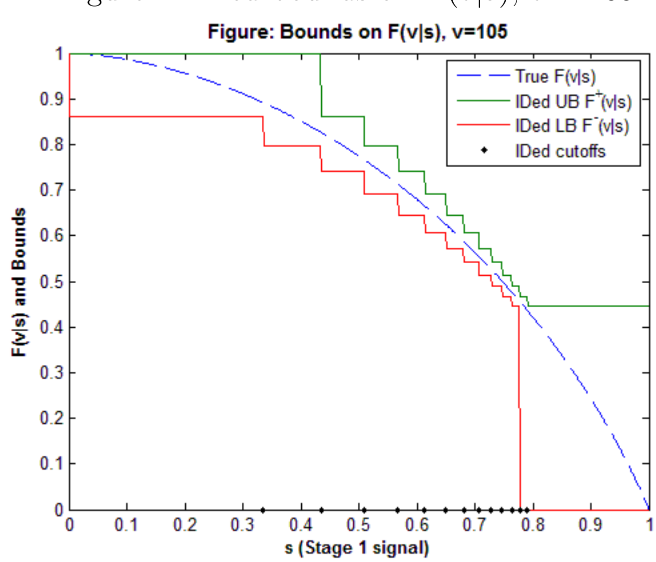


Figure 3: IDed bounds on $\bar{s}_r$ at $N = 6$, $\rho = 0.75$

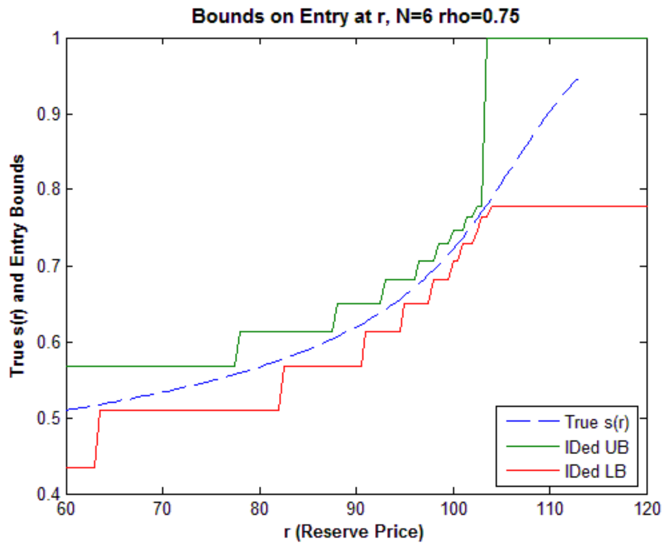


Figure 4: IDed bounds on $\bar{s}_r$ at $N = 9$, $\rho = 0.75$

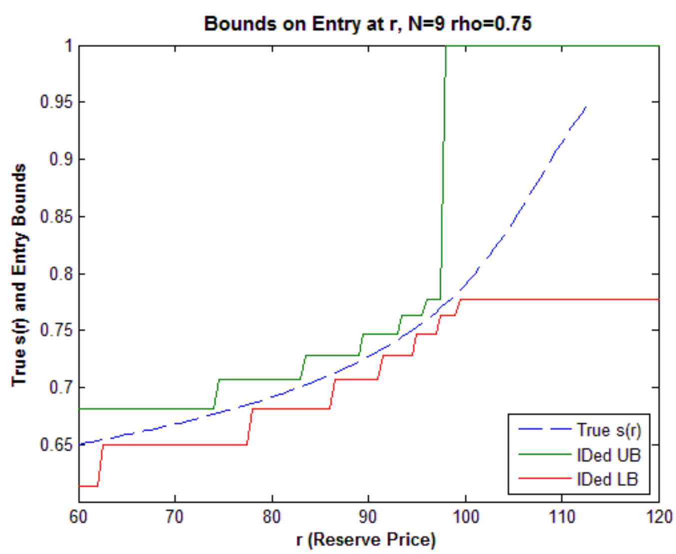


Figure 5: IDed bounds on CF revenue at $N = 6$, $\rho = 0.75$

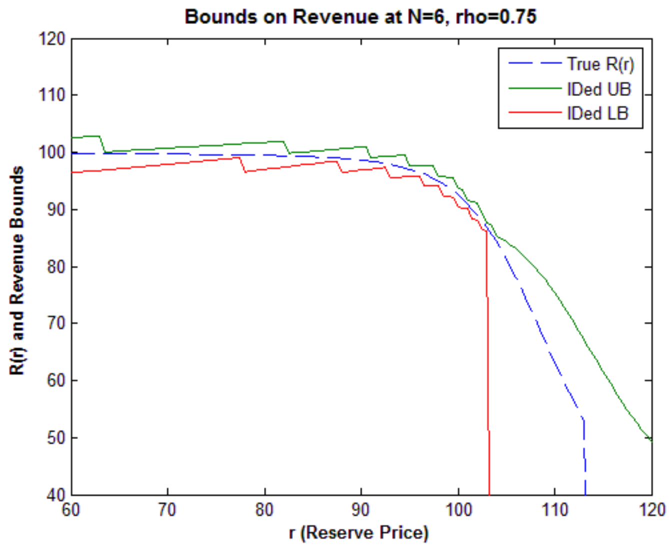
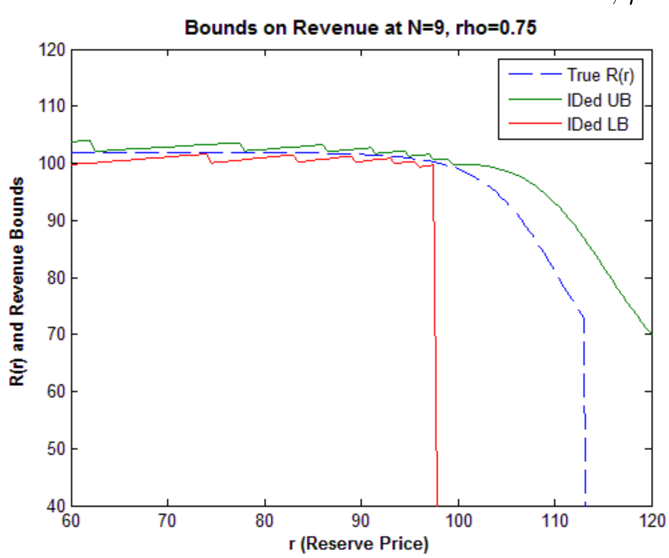


Figure 6: IDed bounds on CF revenue at $N = 9$, $\rho = 0.75$



Appendix 2: Proofs

Proof of Proposition 1 (following Krisha 2002). By the Revelation Principle, any symmetric equilibrium in any mechanism is payoff-equivalent to the truthful equilibrium in some symmetric direct mechanism. Thus WLOG we restrict attention to direct mechanisms in which entrants truthfully report values. In particular, let M be an arbitrary direct mechanism involving allocation rule $Q(\mathbf{v};E)$ and payment rule $P(\mathbf{v};E)$, where $\mathbf{v}$ is a vector of (realized) bidder values and $E \equiv (\bar{s},N)$ is an entry structure, and let $q(v_i;E) \equiv \int_{V_{-i}} Q(v_i,\mathbf{v}_{-i},E)f(\mathbf{v}_{-i}|E)d\mathbf{v}_{-i}$ and $p(v_i;E) \equiv \int_{V_{-i}} P(v_i,\mathbf{v}_{-i},E)f(\mathbf{v}_{-i}|E)d\mathbf{v}_{-i}$ be the corresponding (expected) allocation and payment functions facing bidder i. For truth-telling to be an equilibrium, we must have

$$q(v_i;E)\cdot v_i - p(v_i;E) = \max_z\{q(z;E)\cdot v_i - p(z;E)\},\ \forall v_i \in [\underline{v},\bar{v}].$$

By the Integral Form Envelope Theorem (see Milgrom (2004)), this restriction in turn implies that any incentive-compatible direct mechanism must yield equilibrium bidder profit $\pi(\cdot;E)$ of the form

$$\pi(v;E) = \pi_0(E) + \int_{\underline{v}}^{v} q(y;E)\,dy,$$

where $\pi_0(E)$ is the (mechanism-determined) profit of the lowest entering bidder.

We now return to RS auctions. By Definition 1, the probability of allocation to an entering bidder with value y is

$$\begin{aligned} q(y;E) &= \alpha(y)\cdot \Pr(W_j \le y\,\forall j) \\ &= \alpha(y)\cdot \prod_{j\neq i}\Pr(W_j \le y) \\ &= \alpha(y)\cdot F_w^*(y;\bar{s})^{N-1}. \end{aligned}$$

Equation 2 of Proposition 1 follows immediately. □

Proof of Lemma 1. By definition, a bidder with signal s_i facing competition $\bar{s},N$ will earn expected profit

$$\begin{aligned} \Pi(s_i;\bar{s},N) &\equiv E_v[\pi(v;\bar{s},N)|S=s_i] \\ &= \int_{\underline{v}}^{\bar{v}} \pi(v;\bar{s},N)f(v|s_i)dv \\ &= \pi_0(\bar{s},N) + \int_{\underline{v}}^{\bar{v}}\int_{\underline{v}}^{v} \alpha(y)\cdot F_w^*(y;\bar{s})^{N-1}dydv \\ &= \pi_0(\bar{s},N) + \int_{\underline{v}}^{\bar{v}} \alpha(y)\cdot[1-F(y|s_i)]\cdot F_w^*(y;\bar{s})^{N-1}dy, \end{aligned}$$

where line 3 follows from Proposition 1 and line 4 follows from changing order of integration.

Affiliation implies that $F(y|s_i)$ is decreasing in s_i , so $\Pi(s_i;\bar{s},N)$ is increasing in s_i (see Milgrom and Weber (1982)).

$F^*_w(y;\bar{s}) = \bar{s} + \int_{\bar{s}}^1 F(v|t)dt$, so $\partial F^*_w(y;\bar{s})/\partial\bar{s} = 1 - F(v|\bar{s}) \geq 0$ and $\Pi(s_i;\bar{s},N)$ is increasing in $\bar{s}$.

Finally, $F^*_w(y;\bar{s}) \in [0,1]$, so $F^*_w(y;\bar{s})^{N-1}$ is decreasing in N and $\Pi(s_i;\bar{s},N)$ is decreasing in N (strictly for $F^*_w(y;\bar{s}) < 1$). $\square$

Proof of Proposition 2. Assume symmetric continuation play as in Proposition 1. By definition, in a symmetric entry equilibrium characterized by threshold $\bar{s}$, "enter if and only if $s_i \geq \bar{s}$" must be a best response to itself. Since $\Pi(s_i,\bar{s},N)$ is increasing in s_i, it will be a best respose to enter for all $s_i > s'$ if it is also a best response to enter for $s_i = s'$. We can thus establish the first part of the proposition case by case:

- If $\Pi(0;0,N) > c$, then $\Pi(s_i;0,N) > c$ for all s_i and it is always a best response to enter. Hence $\Pi(0;0,N) > c$ implies $\bar{s} = 0$ is a symmetric equilibrium.

- If $\Pi(1;1,N) < c$, then $\Pi(s_i;0,N) < c$ for all s_i and it is never a best response to enter. Hence $\Pi(1;1,N) < c$ implies $\bar{s} = 1$ is a symmetric equilibrium.

- Otherwise, "enter if and only if $s_i \geq \bar{s}$" will be a best response to itself when the following breakeven condition is satisfied:

$$\Pi(\bar{s},\bar{s},N) \equiv c.$$

  By Lemma 1, $\Pi(\bar{s},\bar{s},N)$ is strictly increasing in both $\bar{s}$ arguments. Further, since all component distributions are assumed continuous, $\Pi(\bar{s},\bar{s},N)$ is continuous in $\bar{s}$. Hence if $\Pi(0;0,N) \leq c \leq \Pi(1;1,N)$, there will exist a unique $\bar{s}$ such that $\Pi(\bar{s},\bar{s},N) \equiv c$.

The cases above are mutually exclusive and exhaustive. Hence there will exist a unique threshold equilibrium $\bar{s}$. We now argue that no other symmetric pure strategy entry equilibrium can exist if signals are informative. To see this, let $S \subset [0,1]$ be an arbitrary signal set, and suppose "enter if $s_i \in S$" is a symmetric entry equilibrium. Then ex ante profit corresponding to signal s_i will be

$$\Pi(s_i;S,N) = \pi_0(S,N) + \int_V \alpha(y)[1 - F(y|s_i)]F^*(y;S)^{N-1}dy.$$

Now suppose $s \in S$, and consider $s' > s$. Then by affiliation we must have

$$\Pi(s';S,N) \geq \Pi(s;S,N) \geq c,$$

If $F(y|s') < F(y|s)$ for some y and some $s \in S$, this inequality will be strict, so "enter at s'" will be the unique best response to S at s' and we must have $s' \in$

S. Hence if signals are strictly informative, any symmetric equilibrium must be a threshold equilibrium. The only exception occurs when signals are uninformative ($F(y|s') = F(y|s)$) for some s, s'. In this case, only the *measure* of S will matter for characterizing equilibrium, but focusing on the threshold equilibrium still involves no loss of generality.

Finally, we establish the stated monotonicity results:

- We know $\Pi(\bar{s}, \bar{s}, N)$ is strictly increasing in $\bar{s}$ for $\bar{s} \in (0,1)$. Hence $c' > c$ implies $\Pi(\bar{s}', \bar{s}', N) > \Pi(\bar{s}, \bar{s}, N)$ and consequently $\bar{s}' > \bar{s}$. However, in the boundary case $\bar{s} = 0$ (or $\bar{s} = 1$), we could have $c < c' < \Pi(\bar{s}, \bar{s}, N')$ (or vice versa), which would imply $\bar{s}' = \bar{s} = 0$ (or $\bar{s}' = \bar{s} = 1$). Hence $\bar{s}(c, N)$ is increasing in c, strictly when $\bar{s} \in (0,1)$.

- We know $\Pi(\bar{s}, \bar{s}, N)$ is strictly increasing in $\bar{s}$ and strictly decreasing in N for $\bar{s} \in (0,1)$. Suppose s' is equilibrium at N' and s is equilibrium at N. Then if $\bar{s} \in (0,1)$ we must have $\Pi(\bar{s}', \bar{s}', N') = c = \Pi(\bar{s}, \bar{s}, N)$, which implies $\bar{s}' > \bar{s}$. However, in the boundary cases $\bar{s} = 0$ (or $\bar{s} = 1$), we could have $\Pi(\bar{s}', \bar{s}', N') > \Pi(\bar{s}, \bar{s}, N) > c$ (or vice versa), which would imply $\bar{s}' = \bar{s} = 0$ (or $\bar{s}' = \bar{s} = 1$). Hence $\bar{s}(c, N)$ is increasing in N, strictly when $\bar{s} \in (0,1)$.

□

Proof of Lemma 2. By definition,

$$\begin{aligned}
F(v,s) &= \int_{\underline{v}}^{v} \int_{0}^{s} f(v,s) ds dv \\
&= \int_{\underline{v}}^{v} \int_{0}^{1} f(v,s) ds dv - \int_{\underline{v}}^{v} \int_{s}^{1} f(v,s) ds dv \\
&= F_v(v) - (1-s) \left[\frac{1}{1-s} \int_{s}^{1} \int_{\underline{v}}^{v} f(v|s) dv ds \right] \\
&= F_v(v) - (1-s) \left[\frac{1}{1-s} \int_{s}^{1} F(v|s) ds \right] \\
&\equiv F^*(v;0) - (1-s) F^*(v;s).
\end{aligned}$$

Comparing the first and last lines of this identity establishes the "if" result: observing data at (almost) all $s \in [0,1]$ permits exact identification of $F(v,s)$. The "only if" result follows from noting that $\mathrm{cl}(\mathcal{S}_x(\mathcal{L})) \neq [0,1]$ implies $\exists$ set $\mathcal{S}^C(\mathcal{L})$ of positive measure such that $F^*(v;s)$ is *not* identified (in general) for $s \in \mathcal{S}^C(\mathcal{L})$. Hence we have no direct information about $F(v,s)$ on $\mathcal{S}^C(\mathcal{L})$. Hence absent further assumptions $F(v,s)$ is not fully identified. □

Proof of Lemma 3. We establish claims for $\check{F}^{+}(v|\bar{s})$; the argument for $\check{F}^{-}(v|\bar{s})$ is analogous.

By construction, if $s^-(\bar{s}) \notin \mathcal{S}$ then $s^-(\bar{s}) \equiv 0$, and if $s^-(\bar{s}) = 0$ and $0 \notin \mathcal{S}$ then $\check{F}^+(v|\bar{s}) \equiv 1 \geq F(v|\bar{s})$. Hence we focus on the case $s^-(\bar{s}) \in \mathcal{S}$.

When $s^-(\bar{s}) \in \mathcal{S}$, there are two possible subcases:

- $\bar{s} = s^-(\bar{s})$: By construction of $s^-(\bar{s})$, this occurs when $\bar{s} \in \mathrm{int}(\mathcal{S})$, which implies that there exists an open neighborhood of identified thresholds $t \in \mathcal{S}$ around $\bar{s}$. Consequently, we can identify the function $(1-t)F^*(v;t)$ at points arbitrarily close to $\bar{s}$, and the limit defining $\check{F}^+(v|\bar{s})$ converges to the corresponding derivative:

$$\lim_{t \uparrow s^-(\bar{s})} \left\{ \frac{(1-t)F^*(v;t) - (1-\bar{s})F^*(v;\bar{s})}{\bar{s}-t} \right\} = -\frac{\partial}{\partial \bar{s}}(1-\bar{s})F^*(v;\bar{s}) \equiv F(v|s).$$

  Hence $\check{F}^+(v|\bar{s}) = F(v|s)$, so $\check{F}^+(v|\bar{s})$ is a distribution and $F(v|s)$ is exactly identified.

- $\bar{s} > s^-(\bar{s})$: By construction, $s^-(\bar{s})$ is then the nearest lower neighbor of $\bar{s}$ in $\mathcal{S}$ (but separated by an open interval). In this case,

$$\begin{aligned}
\lim_{t \uparrow s^-(\bar{s})} \left\{ \frac{(1-t)F^*(v;t) - (1-\bar{s})F^*(v;\bar{s})}{\bar{s}-t} \right\} &= \frac{[1-s^-(\bar{s})]F^*(v;s^-(\bar{s})) - (1-\bar{s})F^*(v;\bar{s})}{\bar{s}-s^-(\bar{s})} \\
&= \frac{1}{\bar{s}-s^-(\bar{s})} \left\{ F^*(v;s^-(\bar{s})) - F^*(v;\bar{s}) \right\} \\
&= \frac{1}{\bar{s}-s^-(\bar{s})} \left\{ \int_{s^-(\bar{s})}^{1} F(v|t)dt - \int_{\bar{s}}^{1} F(v|t)dt \right\} \\
&= \frac{1}{\bar{s}-s^-(\bar{s})} \int_{s^-(\bar{s})}^{\bar{s}} F(v|t)dt \\
&= F(v|S_i \in [s^-(\bar{s}), \bar{s}]).
\end{aligned}$$

  Line 1 implies $\check{F}^+(v|\bar{s})$ is identified (since it depends only on identified components), Line 5 implies that $\check{F}^+(v|\bar{s})$ is a distribution, and Line 4 implies that $\check{F}^+(v|\bar{s})$ bounds $F(v|s)$:

$$\begin{aligned}
\frac{1}{\bar{s}-s^-(\bar{s})} \int_{s^-(\bar{s})}^{\bar{s}} F(v|t)dt &\geq \frac{1}{\bar{s}-s^-(\bar{s})} \int_{s^-(\bar{s})}^{\bar{s}} F(v|\bar{s})dt \\
&= \frac{\bar{s}-s^-(\bar{s})}{\bar{s}-s^-(\bar{s})} F(v|\bar{s}) = F(v|\bar{s}),
\end{aligned}$$

  where the first inequality follows since affiliation implies $F(v|s') \leq F(v|s)$ for $s' \geq s$ (see Milgrom and Weber (1982)).

Taken together, the cases above establish all claims in Lemma 3. □

Proof of Proposition 3. We establish claims for $F^+(v|s)$; the argument for $F^-(v|s)$ is analogous.

Identification of $F^+(v|t)$ follows from (i) $s^-(t) \in \{\mathcal{S}, 1\}$ by construction, (ii) identification of $\check{F}^+(v|s)$ for $s \in \mathcal{S}$, and (iii) $\check{F}^+(v|s) \equiv 1$ for $s = 0$ if $s \neq \mathcal{S}$. Hence $F^+(v|t)$ depends only on objects recoverable from process $\mathcal{L}$.

The distribution and exact identification properties of $F^+(v|t)$ are inherited directly from the corresponding properties of $\check{F}^+(v|t)$.

Finally, to establish bounds, we consider cases:

- If $t \in \mathcal{S}$, then $F^+(v|t) \equiv \check{F}^+(v|t) \geq F(v|t)$.
- Otherwise, $F^+(v|t) \equiv \check{F}^+(v|s^-(t)) \geq F(v|s^-(t)) \geq F(v|t)$, where the last inequality follows by the stochastic-dominance property of affiliation.

Taken together, the cases above establish all claims in Proposition 3. □

Proof of 4. Identification of $c^+(z)$ and $c^-(z)$ and the inequalities $c^+(z) \geq c(z) \geq c^-(z)$ follow immediately from identification of $F^-(y|\bar{s}_N(z))$ and $F^+(y|\bar{s}_N(z))$ and $F^+(y|\bar{s}_N(z)) \geq F(y|\bar{s}_N(z) \geq F^-(y|\bar{s}_N(z))$, with exact equality obtaining when $F^{\pm}(y|\bar{s}_N(z)) = F(y|\bar{s}_N(z))$. □

Proof of Proposition 5. To establish Statement 1, suppose $z \in \text{int}(Z)$ and $\bar{s} \in (0,1)$. Then there exists an open ϵ-ball $B_\epsilon(z) \subset Z$ around z. By Assumption 8, $c(\cdot)$ is continuous and monotonic in continuous components of z, and hence maps open sets to open sets. Thus $\mathcal{L}$ involves an open ϵ-ball $B_\epsilon(c(z)) \subset R^+$ of costs around $c(z)$. Finally, by Proposition 2, the equilibrium threshold $\bar{s}(\cdot, N)$ is continuous and monotonic in $c(\cdot)$ for $\bar{s} \in (0,1)$. Hence $\bar{s}(\cdot, N)$ also maps open sets to open sets, so $\bar{s}(c(z), N) \in \text{int}(\mathcal{S}(\mathcal{L}))$. Exact local identification then follows from Propositions 3 and 4.

To establish Statement 2, suppose $Z = R^k$ and the range of $c(\cdot)$ is unbounded in R^+. Then $z \in \text{int}(Z)$ by definition, and for an appropriate choice of z we can produce any $c(z) \in [0, \infty)$. Hence we can sustain any $s \in [0,1]$ as an equilibrium for some z, so $\mathcal{S}(\mathcal{L}) = [0,1]$ and full identification follows from Lemma 2. □

Proof of revenue characterization in Lemma 5 (from Gentry-Li 2011b) . For any $(s;N)$ pair, expected seller revenue at allocation rule α is given by

$$R_\alpha(\bar{s}; N) = AV_\alpha(s; N) - N\Pi^*_\alpha(s; N),$$

where $AV_\alpha(\cdot)$ is *ex ante* expected allocation value of the object being auctioned and $\Pi^*(s;N)$ is expected *ex ante* equilibrium profit for an arbitrary bidder.

To obtain $AV_\alpha(\cdot)$, let $Y_{1:N}$ be max realized value among N potential bidders. Then net value created is$Y_{1:N}$ if sale, v_0 if no sale. Conditional on $Y_{1:N}$, expected allocation value is thus

$$\alpha(Y_{1:N})Y_{1:N} + [1 - \alpha(Y_{1:N})]v_0.$$

Integrating with respect to $Y_{1:N}$, we obtain *ex ante* expected allocation value:

$$\begin{aligned}
AV_\alpha(s;N) &= s^N v_0 + \int_{\underline{v}}^{\bar{v}} \left\{\alpha(y)y + [1-\alpha(y)]v_0\right\} g^*_{1:N}(y;s)dy \\
&= \int_{v_0}^{\bar{v}} \left\{\alpha(y)y + [1-\alpha(y)]v_0\right\} dG^*_{1:N}(y;s),
\end{aligned}$$

where $g^*_{1:N}(y;s) \equiv NF^*_w(y;s)^{N-1}f^*_w(y;s)$ is the density of $Y_{1:N}$ on $[\underline{v}, \bar{v}]$ given entry threshold s, $G^*_{1:N}(y;s) = F^*_w(y;s)^N$ is the corresponding distribution on $[v_0, \bar{v}]$, and $\alpha(v_0) \equiv 0$ by Assumption 9.

To obtain $\Pi^*(s;N)$, we start from the result in Proposition 1:

$$\begin{aligned}
\pi_\alpha(v;s,N) &= \pi_0^\alpha(s,N) + \int_{\underline{v}}^{v} \alpha(t)\cdot F^*_w(t;s)^{N-1}dt \\
&= \int_{v_0}^{\underline{v}} \alpha(t)s^{N-1}dt + \int_{\underline{v}}^{v} \alpha(t)\cdot F^*_w(t;s)^{N-1}dt - \rho \\
&= \int_{v_0}^{\underline{v}} \alpha(t)\cdot F^*_w(t;s)^{N-1}dt - \rho \\
&= \lambda_\alpha(v;s,N)\cdot \alpha(v)F_w(v;s)^{N-1} - \rho,
\end{aligned}$$

where the second equation follows from Assumption 9 and

$$\lambda_\alpha(v;s,N) \equiv \begin{cases} 0 & \text{if } \alpha(v) = 0; \\ \int_{v_0}^{v} \frac{\alpha(t)}{\alpha(v)} \cdot \frac{F^*_w(t;s)^{N-1}}{F^*_w(v;s)^{N-1}}dt & \text{otherwise.} \end{cases}$$

gives the average incremental profit (above $-\rho$) a bidder of type v recieves per win.

Integrating over the distribution $F^*_w(y;s)$ then gives $\Pi^*(s;N)$

$$\Pi^*(s;N) = \int_{\underline{v}}^{\bar{v}} \lambda_\alpha(y;s,N)\cdot\alpha(y)F^*_w(y;s)^{N-1}f^*_w(y;s)dy - (1-s)\rho$$

and multiplying by N yields

$$\begin{aligned}
N\Pi^*(s;N) &= \int_{\underline{v}}^{\bar{v}} \lambda_\alpha(y;s,N)\alpha(y)\cdot NF^*_w(y;s)^{N-1}f^*_w(y;s)dy - N(1-s)\rho \\
&= \int_{\underline{v}}^{\bar{v}} \lambda_\alpha(y;s,N)\alpha(y)\cdot g^*_{1:N}(y;s)dy - N(1-s)\rho \\
&= \int_{\underline{v}}^{\bar{v}} \lambda_\alpha(y;s,N)\alpha(y)\, dG^*_{1:N}(y;s)dy - N(1-s)\rho.
\end{aligned}$$

Combining the results above gives a final expression for seller revenue:

$$\begin{aligned} R_\alpha(s;N) &= \int_{v_0}^{\bar{v}} \{\alpha(y)y + [1-\alpha(y)]v_0\}\, g^*_{1:N}(y;s)dy \\ &\quad - \int_{\underline{v}}^{\bar{v}} \lambda_\alpha(y;s,N)\alpha(y)\, dG^*_{1:N}(y;s) + N(1-s)\rho \\ &= \int_{v_0}^{\bar{v}} \{\alpha(y)[y-\lambda_\alpha(y;s,N)] + [1-\alpha(y)]v_0\}\, dG^*_{1:N}(y;s)dy + N(1-s)\rho. \end{aligned}$$

where the second equality follows because $\int_{v_0}^{\underline{v}} \lambda_\alpha(y;s,N)\, dG^*_{1:N}(y;s) = 0$: $\lambda_\alpha(v_0;s,N) \equiv 0$ and $g^*_{1:N}(y;s) \equiv 0$ for $y \in (v_0, \underline{v})$. □

Proof of Lemma 5. Identification of $R_\alpha(s;N)$ for $s \in \mathcal{S}$ follows directly from Equation 9: $R_\alpha(\cdot)$ depends only on mechanism components (α, ρ, v_0) (known by hypothesis) and distributions $F^*_w(\cdot;s)$ and $G^*_{1:N}(\cdot;s)$ (identified for $s \in \mathcal{S}$). Thus it only remains to show $R_\alpha(s;N)$ is decreasing in s. Equation (9) implies that s affects seller revenue through (at most) three channels: the per-win profit function $\lambda_\alpha(v;s,N)$, the distribution $G^*_{1:N}(\cdot;s)$, and the residual term $N(1-s)\rho$. We show that each of these partial effects is negative.

First, consider effects through the per-win profit function $\lambda_\alpha(v;s,N)$. Note that

$$\frac{\partial}{\partial s}\lambda_\alpha(v;s,N) = \int_{v_0}^{v} \frac{\alpha(t)}{\alpha(v)} \cdot \frac{\partial}{\partial s}\left\{\frac{F^*_w(t;s)^{N-1}}{F^*_w(v;s)^{N-1}}\right\} dt.$$

By algebra,

$$\begin{aligned} \frac{\partial}{\partial s}\left\{\frac{F^*_w(t;s)^{N-1}}{F^*_w(v;s)^{N-1}}\right\} &= \frac{(N-1)F^*_w(t;s)^{N-2}\frac{\partial}{\partial s}F^*_w(t;s)}{F^*_w(v;s)^{N-1}} - \frac{(N-1)F^*_w(t;s)^{N-1}\frac{\partial}{\partial s}F^*_w(v;s)}{F^*_w(v;s)^{N}} \\ &= (N-1)\frac{F^*_w(t;s)^{N-2}}{F^*_w(v;s)^{N-1}}\left\{[1-F(t|s)] - \frac{F^*_w(t;s)}{F^*_w(v;s)}[1-F(v|s)]\right\} \\ &\geq 0\, \forall t \leq v, \end{aligned}$$

since $t \leq v$ means $F^*_w(t;s) \leq F^*_w(v;s)$ and $F(t|s) \leq F(v|s)\ \forall s$. Thus $\lambda_\alpha(v;s,N)$ is increasing in s for all v, so the effect of s on R through $\lambda_\alpha(v;s,N)$ is negative.

Next, consider effects through the distribution $G^*_{1:N}(\cdot;s)$. It is easy to show that $G^*_{1:N}(v;s)$ is increasing in s for any v, hence $s' \geq s$ means $G^*_{1:N}(\cdot;s)$ first-order stochastically dominates $G^*_{1:N}(\cdot;s')$. Thus if the integrand

$$\{\alpha(y)[y-\lambda_\alpha(y;s,N)] + [1-\alpha(y)]v_0\} \tag{10}$$

is increasing in y, an increase in s will involve taking the expectation of an increasing function with respect to a stochastically dominated distribution, which must imply a decrease in revenue. It is therefore is sufficient to show that the integrand (10) is increasing in y.

- First, note that $[y - \lambda_\alpha(y; s, N)]$ is increasing in y:

$$\begin{aligned} \frac{\partial}{\partial y}[y - \lambda_\alpha(y; s, N)] &\equiv 1 - \frac{\partial}{\partial y} \int_{v_0}^{y} \frac{\alpha(t)}{\alpha(y)} \cdot \frac{F_w^*(t; s)^{N-1}}{F_w^*(y; s)^{N-1}} dt \\ &= 1 - \frac{\partial}{\partial y} \frac{1}{\alpha(y) F_w^*(y; s)^{N-1}} + 1 \\ &= -\frac{\partial}{\partial y} \frac{1}{\alpha(y) F_w^*(y; s)^{N-1}} \geq 0 \end{aligned}$$

  since $\alpha(y) F_w^*(y; s)^{N-1}$ is increasing in y by construction.

- Second, note that $[y - \lambda_\alpha(y; s, N)] \geq v_0$ for $y \geq v_0$:

$$\begin{aligned} [y - \lambda_\alpha(y; s, N)] &\equiv [y - \lambda_\alpha(y; s, N)]|_{v_0} + \int_{v_0}^{y} \frac{\partial}{\partial t}[t - \lambda_\alpha(t; s, N)] dt \\ &= v_0 + \int_{v_0}^{y} \frac{\partial}{\partial t}[t - \lambda_\alpha(t; s, N)] dt \\ &\geq v_0 \end{aligned}$$

  since we know $\frac{\partial}{\partial y}[y - \lambda_\alpha(y; s, N)] \geq 0$.

- Finally, note that (by construction) $\alpha(y)$ is increasing in y.

Hence increasing y has two effects on the function (10): it increases $[y - \lambda_\alpha(y; s, N)]$ and shifts weight from v_0 to $[y - \lambda_\alpha(y; s, N)]$ (through $\alpha(y)$). Since $[y - \lambda_\alpha(y; s, N)] \geq v_0$, both these effects are positive, so (10) is increasing in y. It follows that increasing s leads to taking an expectation of an increasing function with respect to a stochastically dominated distribution. Hence the effect of s on R through the distribution $G_{1:N}^*(y; s)$ is negative.

Finally, note that $\rho_\alpha \geq 0$ by construction. Hence an increase in s implies a decrease in $(1 - s) N \rho$.

Combining these observations, we conclude that seller revenue $R_\alpha(s; N)$ is decreasing in s for any N. □

Proof of Lemma 6. We establish claims for $s_\alpha^+(z, N)$; the argument for $s_\alpha^-(z, N)$ is analogous. Suppose $\bar{s}$ is an equilibrium at (z, N, α). Then Proposition 2 implies $\Pi_\alpha(\bar{s}, N; F) \equiv c(z)$. Since $\Pi_\alpha(s, N; \tilde{F})$ is increasing in s and decreasing in F, it follows that

$$c^+(z) \geq c(z) \equiv \Pi_\alpha(\bar{s}, N; F) \geq \Pi_\alpha(\bar{s}, N; F^+)$$

Hence $\Pi_\alpha(s', N; F^+) > c^+(z)$ implies $s' \neq \bar{s}_\alpha(z, N)$ and (in particular, by monotonicity of $\Pi_\alpha(\cdot)$ in s) $s' > \bar{s}_\alpha(z, N)$. Taking the smallest such s' identified by $\mathcal{L}$ (or the uninformative bound 1 if no such s' exists) yields $s_\alpha^+(z, N)$ defined above. □

Proof of Corollary 5. See Haile and Tamer (2003) for proof that $r_+^*(z, N)$ and $r_-^*(z, N)$ bound $r^*(z, N)$. The final equality statement follows from exact identification of counterfactual revenue when $s_{r^*}(z, N) \in \text{int}(\mathcal{S}(\mathcal{L}))$. □

References

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

Athey, S., and P. A. Haile (2002): "Identification of Standard Auction Models," *Econometrica*, 70(6), 2107–2140.

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.

Berry, S., and P. Haile (2010): "Nonparametric Identification of Multinomial Choice Demand Models with Heterogeneous Consumers," Working Paper, Yale University.

Chesher, A. (2005): "Nonparametric Identification Under Discrete Variation," *Econometrica*, 73, 1525–1550.

Ciliberto, F., and E. Tamer (2009): "Market Structure and Multiple Equilibria in Airline Markets," *Econometrica*, 77(6), 1791–1828.

Fan, Y., and S. Park (2009): "Partial Identification of the Distribution of Treatment Effects and its Confidence Sets," *Advances in Econometrics*, 25, 3–70.

Gentry, M., and T. Li (2011): "A Revenue Equivalence Result for Auctions with Selective Entry," Working Paper, Vanderbilt University.

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

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

Haile, P. A., and E. Tamer (2003): "Inference with an Incomplete Model of English Auctions," *The Journal of Political Economy*, 111(1), 1–51.

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

Hortacsu, A. (2002): “Mechanism Choice and Strategic Bidding in Divisible Good Auctions: An Empirical Analysis of the Turkish Treasury Auction Market,” Working Paper, University of Chicago.

Krasnokutskaya, E. (2009): “Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity,” *Review of Economic Studies*, forthcoming.

Krasnokutskaya, E., and K. Seim (2009): “Bid Preference Programs and Participation in Highway Procurement Auctions,” *American Economic Review*, forthcoming.

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

Li, T. (2005): “Econometrics of First-Price Auctions with Entry and Binding Reservation Prices,” *Journal of Econometrics*, 126(1), 173 – 200.

Li, T., I. Perrigne, and Q. Vuong (2000): “Conditionally Independent Private Information in OCS Wildcat Auctions,” *Journal of Econometrics*, 98(1), 129–162.

Li, T., I. Perrigne, and Q. Vuong (2002): “Structural Estimation of the Affiliated Private Value Auction Model,” *The RAND Journal of Economics*, 33(2), 171–193.

Li, T., and B. Zhang (2009): “Affiliation and Entry in First-Price Auctions with Heterogeneous Bidders,” Working Paper, Vanderbilt University.

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 X. Zheng (2007): “Information Acquisition and/or Bid Preparation: A Structural Analysis of Entry and Bidding in Timber Sale Auctions,” *Journal of Econometrics*, 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.

Magnac, T., and E. Maurin (2008): “Partial Identification in Monotone Binary Models: Discrete Regressors and Interval Data,” *Review of Economic Studies*, 75(3), 835–864.

Manski, C. (2003): *Partial Identification of Probability Distributions*. Berlin: Springer- Verlag.

Manski, C. F., and E. Tamer (2002): “Inference on Regressions with Interval Data on a Regressor or Outcome,” *Econometrica*, 70(2), 519–546.

Marmer, V., A. Shneyerov, and P. Xu (2007): “What Model for Entry in First-Price Auctions? A Nonparametric Approach,” Working Paper, University of British Columbia.

Milgrom, P. R., and R. J. Weber (1982): “A Theory of Auctions and Competitive Bidding,” *Econometrica*, 50(5), 1089–1122.

Molinari, F. (2008): “Partial Identification of Probability Distributions with Misclassified Data,” *Journal of Econometrics*, 144(1), 81–117.

Nelsen, R. B. (1999): *An Introduction to Copulas*, Lecture Notes in Statistics. Springer-Verlag New York.

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

Roberts, J., and A. Sweeting (2010a): “Competition versus Auction Design,” Working Paper, Duke University.

Roberts, J., and A. Sweeting (2010b): “Entry and Selection in Auctions,” Working Paper, Duke University.

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

Tamer, E. (2003): “Incomplete Simultaneous Discrete Response Model with Multiple Equilibria,” *The Review of Economic Studies*, 70(1), 147 –165.

Tang, X. (2011): “Bounds on Revenue Distributions in Counterfactual Auctions with Reserve Prices,” *RAND Journal of Economics*, 42(1), 175–203.

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