Identification in Auctions with Selective Entry*

Matthew Gentry† and Tong Li‡

First version: September 2010

This version: November 2012

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 accommodates a wide range of entry processes, in particular nesting the Levin and Smith (1994) and Samuelson (1985) models as special cases. To date, however, the model's identification properties have not been well understood. We establish identification results for the general AS model, using variation in factors affecting entry behavior (such as potential competition or entry costs) to construct identified bounds on model fundamentals. If available entry variation is continuous, the AS model may be point identified; otherwise, it will be partially identified. We derive constructive identification results in both cases, which can readily be refined to produce the sharp identified set. We also consider policy analysis in environments where only partial identification is possible, and derive identified bounds on expected seller revenue corresponding to a wide range of counterfactual policies while accounting for endogenous and arbitrarily selective entry. Finally, we establish that our core results extend to environments with asymmetric bidders and nonseperable auction-level unobserved heterogeneity.

*We are grateful to Co-Editor, Jean-Marc Robin, and three anonymous referees for their constructive comments that have greatly improved the paper. We also thank Yanqin Fan, Jeremy Fox, James Heckman, and participants in seminars at National University of Singapore, Singapore Management University, HKUST, Columbia University, UBC, University of Pennsylvania, Carnegie Mellon University, Penn State University, and 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, the European Meeting of the Econometric Society in Oslo, August 2011, and the Identification and Inference in Econometrics Conference at Vanderbilt University, March 2012, for helpful comments and discussions. Support from the National Science Foundation (SES-0922109) is gratefully acknowledged. Previous versions of the paper were circulated under the title "Partial Identification in Auctions with Selective Entry."

†Department of Economics, London School of Economics. Email: m.l.gentry@lse.ac.uk.

‡Department of Economics, Vanderbilt University. Email: tong.li@vanderbilt.edu.

Introduction

Endogenous participation plays a major role in real-world auction markets. Empirical auction studies routinely find that large proportions 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 (2010a; 2010b)), in online auction markets (Bajari and Hortacsu (2003)), and in other procurement settings (Krasnokutskaya and Seim (2011)). Further, it is well-known that 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. These observations in turn reinforce the growing emphasis on accounting for entry in empirical applications.

While the economic importance of endogenous participation in auction markets is well understood, there is still no clear consensus regarding how to account for entry in structural analysis. To ensure identification, most studies incorporating entry do so via one of two polar entry paradigms: that of Samuelson (1985) (the *S model*), in which potential bidders observe own values exactly prior to entry, or that of Levin and Smith (1994) (the *LS model*), in which potential bidders have no information on own values prior to entry. These models are econometrically appealing because they simplify selection, but involve stark restrictions which (improperly enforced) can substantially distort structural results.[1] Consequently, several recent studies have begun to explore structural analysis based on a more general

[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 (2012). 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.

entry framework we call the *Affiliated-Signal (AS) model.* First proposed in Ye (2007), this framework embodies the following basic structure: 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. The AS model naturally generalizes the S and LS models to accommodate endogenous and arbitrarily selective entry; applications include Marmer, Shneyerov, and Xu (2007), who propose nonparametric specification tests of the AS, S, and LS models, and Roberts and Sweeting (2010a; 2010b), who estimate a parametric variant of the AS model using data on California timber auctions.[2] To date, however, identification in the AS model has not been well understood. This uncertainty in turn both limits the AS model's use in applications and motivates our current investigation.

This paper explores identification in auctions with endogenous and selective entry, seeking to characterize information nonparametrically available under the general AS model. For clarity, we focus discussion on the special case where the post-entry mechanism is a second-price auction, but our fundamental insights extend immediately to first-price, ascending, and Dutch auctions, and indeed to any standard auction (in the sense of Riley and Samuelson (1981)) for which an appropriate value recovery rule is known. Within this general class of mechanisms, we establish the following four results. First, we map observed variation in entry behavior (induced by factors such as potential competition or entry costs) into identified bounds on AS fundamentals, where continuous entry variation permits point identification and discrete entry variation yields partial identification. Second, we explore pointwise sharpness of these identified bounds, first deriving a test to verify whether given bounds are sharp, then using this condition as the basis for an algorithm to construct the pointwise sharp identified set. Third, we translate our bounds on fundamentals into bounds on expected seller revenue corresponding to a wide range of counterfactual auction mechanisms, which account for endogenous and selective AS entry. Finally, we extend our core

[2]Notably, both applications find the imperfect selection permitted by the AS model to be economically important.

results to environments with asymmetric bidders and nonseparable unobserved auction-level heterogeneity. We thereby characterize information nonparametrically available in a general class of auction mechanisms under endogenous and arbitrarily selective entry, illustrate the capacity of this information to support a rich set of counterfactual policy analyses, and establish that these conclusions apply even in more complicated econometric environments.[3] To our knowledge, these represent the first identification results applicable outside the polar S and LS cases, and provide a formal theory of identification to complement the large and growing empirical literature on auctions with entry.[4]

While our focus on selective entry is distinctive, our work builds on a large and influential literature on nonparametric identification in auctions. The possibility of such identification was first demonstrated by Guerre, Perrigne, and Vuong (2000) for first-price auctions with independent private values, and has been extended to many other environments in more recent studies; see, e.g., Li, Perrigne, and Vuong (2000) and Li, Perrigne, and Vuong (2002) for first-price auctions with conditionally independent and affiliated private values, Athey and Haile (2002) for other standard auction formats, Krasnokutskaya (2011) and Hu, McAdams, and Shum (2011) for auctions with unobserved heterogeneity, and Athey and Haile (2005) for a comprehensive survey of the literature. The primary focus of this literature is recovery of model primitives such as the distribution of values from observables such as the distribution of bids. While these results represent an essential point of departure for our current study, recovery of values *per se* is not our primary interest. Rather, we seek to derive restrictions on

[3]It should be emphasized, 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).

[4]To our knowledge, the only other study touching on nonparametric identification in the AS model is the previously-cited work by Marmer, Shneyerov, and Xu (2007), which proposes nonparametric specification tests for the S, LS, and AS models. These tests depend on quantiles of the *ex post* distribution of values among entrants, so a key first step is to show the quantities required are identified. This analysis parallels our discussion of directly identified objects in Section 2.1 below. Beyond this point, however, our investigations diverge: Marmer, Shneyerov, and Xu (2007) use directly identified objects to test competing entry specifications, whereas we use them to derive bounds on AS fundamentals. Note that these results are natural complements in applications: Marmer, Shneyerov, and Xu (2007) indicate when the general AS model is required, while we characterize what can be learned in cases where it is required.

ex ante fundamentals taking as given *ex post* quantities already known to be identified, where the key identification challenge is that *ex post* quantities represent an unknown selection of model fundamentals. As noted above, therefore, our core identification results apply to any standard auction such that an appropriate value recovery rule is known.[5]

Our results also contribute to two fields of inquiry in econometrics more broadly. The first of these is the literature on selection, which since at least the work of Heckman (1976) has been a central subfield of econometrics. This literature is too broad to survey in detail here; for current purposes, the key distinction is that we consider selection within an auction game, which naturally directs econometric analysis. The second is the literature on partial identification, also a substantial subfield in econometrics.[6] As noted above, our core results imply point identification of fundamentals where entry variation is continuous, but only identified bounds where entry is discrete. This result parallels the typical finding that discrete regressors induce partial identification in nonparametric regression contexts; see., e.g., Chesher (2005), Magnac and Maurin (2008), and Chesher and Smolinski (2010) among others. Among studies which considers partial identification in auctions specifically, our work is most similar in spirit to Haile and Tamer (2003), who relax assumptions on bidding behavior to obtain bounds on model fundamentals and counterfactual revenue in ascending auctions, and Tang (2011), who provides bounds for counterfactual revenue in auctions with affiliated values. Relative to both, however, we consider a very different problem (auctions with entry) and relax a different set of assumptions (those governing the nature of selection). We thus contribute both to the literature on auctions with entry specifically and to discussion on broader problems of interest in econometrics.

The rest of this paper is organized as follows. Section 1 describes the AS model, and

[5]Under standard assumptions, this class of mechanisms includes first-price, second-price, English and Dutch auctions; see Athey and Haile (2005) for further details. An interesting extension we do not consider in detail is to environments where the *ex post* value distribution is only partially identified; our core intuition seems likely to extend, but formalization becomes somewhat more involved.

[6]This literature was pioneered by Charles Manski; see Manski (2003) for a summary of early contributions. Recent contributions include Manski and Tamer (2002), Tamer (2003), Molinari (2008), Fan and Park (2009), Chesher and Rosen (2011), and Komarova (2012) to name just a few.

outlines key features of the entry equilibrium. Section 2 presents our core identification, partial identification, and sharpness results. Section 3 explores policy analysis under partial identification in the AS model, deriving bounds on expected seller revenue corresponding to a wide range of counterfactual mechanisms. Section 4 extends our core identification results to environments with unobserved heterogeneity and asymmetric bidders. Finally, Section 5 concludes. Supplemental material is provided in three appendices: Appendix A illustrates our key results through a range of numeric examples, Appendix B describes how to apply our basic results under other auction mechanisms, and Appendix C gives formal proofs of all results.

1 The AS model: Setup and equilibrium

Our goal in this study is to explore the ability of auction models to yield policy-relevant insights without strong *a priori* assumptions on the nature of selection. We formalize this investigation in the context of the *AS model*, a framework which allows bidders to select into entry based on preliminary signals, but which imposes minimal assumptions on the nature of this selection. This section formally defines the AS model and derives its key equilibrium predictions. These in turn provide the groundwork for our subsequent identification analysis. We maintain the following notational conventions throughout: stars (e.g. s^*) denote equilibrium quantities, hats (e.g. $\hat{s}$) denote quantities whose identification follows trivially from standard results in the literature, and bars (e.g. $\bar{s}$) denote fixed values.

1.1 Model setup

Consider allocation of an indivisible good among N potential bidders via a two-stage auction game, where bidders have independent private values for the good being sold. Timing of this game is as follows. First, in Stage 1, each potential bidder i observes a private signal S_i of her (unknown) private value V_i, and all potential bidders simultaneously choose whether to enter the auction. Entry involves payment of an entry cost c, which may be interpreted as any combination of opportunity, value discovery, and bid preparation costs. Next, in Stage 2, the n bidders who chose to enter in Stage 1 learn their actual valuations v_i and submit bids for the object being sold. Finally, auction outcomes are determined by the rules of the auction mechanism, which are common knowledge to all participants. For current purposes, we assume that bidders observe the number of potential bidders N prior to entry, but (as in most sealed-bid procurement settings) do not observe the number of entrants n until the auction concludes. Allowing n to be observed prior to bidding would slightly change some details of the derivation, but would not alter any of our core results.[7]

Let $F(v, s)$ denote the joint distribution linking Stage 2 values V_i to Stage 1 signals S_i. We impose the following structure on this distribution:

Assumption 1. *Each bidder i draws value-signal pairs (V_i, S_i) from a joint distribution $F(v, s)$ satisfying the following properties:*

1. *The random variable V_i has positive support on a bounded interval $\mathcal{V}$, and the joint distribution function $F(v, s)$ is continuous in (v, s).*[8]

2. *For each bidder i, the conditional distribution of V_i is stochastically ordered in S_i: $s' \geq s$ implies $F(v|s') \leq F(v|s)$.*

3. *The random pairs (V_i, S_i) are independent across bidders: $(V_i, S_i) \perp (V_j, S_j)$ for all $j \neq i$.*

4. *WLOG, we normalize first-stage signals S_i to have a uniform marginal distribution on $[0, 1]$: $S_i \sim U[0, 1]$.*

[7] As noted above, our choice to focus on unknown n is motivated by our belief that this best reflects the institutional structure typical of sealed-bid lettings, where measures of potential competition such as planholders and active firms are plausibly common knowledge, but the set of bids received is revealed only after the auction concludes. In circumstances where known n is considered a preferable assumption, the main change in the argument would be to allow bidding strategies and bid distributions to depend on n.

[8] Note that we impose continuity on $F(v, s)$ but not on $F(v|s)$. This permits us to nest the perfectly-selective S case, under which $F(v, s) = \min\{F_v(v), s\}$ is continuous (for continuous $F_v(\cdot)$) but $F(v|s) = \mathbf{1}[v \geq F_v^{-1}(s)]$ is not.

Condition 1.1 is standard, and Condition 1.2 formalizes the sense in which higher signals are "good news." Condition 1.3 generalizes the standard Independent Private Values framework to incorporate Stage 1 signals. Finally, Condition 1.4 is without loss of generality since monotone transformations preserve information. Note that in place of Condition 1.2, prior work has typically imposed the stronger condition of *affiliation* between V_i and S_i in the sense of Milgrom and Weber (1982); see, e.g., Ye (2007) and Marmer, Shneyerov, and Xu (2007). This in turn motivates our "Affiliated-Signal" label. As stochastic ordering alone is sufficient for all results, however, we prefer to employ this slightly weaker restriction.

To ensure a well-behaved equilibrium, we impose one additional regularity condition on the conditional distribution $F(\cdot|\cdot)$:

Assumption 2. *The integral $\int_{\mathcal{V}} F(y|S)dy$ exists and is continuous in S at each $S \in [0,1]$.*

Note that this is a weaker restriction than continuity of $F(y|s)$ in (y, s). This extra generality matters here for the following reason: Assumption 2 formally nests the perfectly selective S model, whereas the standard continuity assumption can only approach it as a limit.[9] Assumption 2 thus permits us to derive results which subsume either polar case.

Most prior work on identification in auctions has considered recovery of private values V_i from observed bids B_i; see Athey and Haile (2002) for a comprehensive survey. This is not our focus here. Rather, starting from objects already known to be identified, we seek to derive bounds on underlying AS fundamentals. To highlight this distinction, we frame our discussion in terms of the simplest possible mechanism: a second-price sealed-bid auction. We stress, however, that this is only for expositional convenience: all core results extend readily to any auction in the class considered by Riley and Samuelson (1981) (any

[9] In particular, the S model assumes signals are perfectly informative, which given our normalization of uniform S implies $F(y|s) = \mathbf{1}[y \geq F_v^{-1}(s)]$. We then have:

$$\int_{\mathcal{V}} F(y|s)dy \equiv \int_{F_v^{-1}(s)}^{\bar{v}} 1dy.$$

Assumption 1.1 implies that $F_v(\cdot)$ is continuous and strictly increasing, so the inverse $F_v^{-1}(s)$ exists and is continuous. Hence the S model satisfies Assumption 2.

RS auction), such that the *ex post* distribution of values among entrants is identified from observation of Stage 2 bids.[10] This will of course involve minor differences in detail, which we discuss more fully in Appendix B.

1.2 Equilibrium

We seek to characterize entry and bidding behavior in a symmetric pure strategy Bayesian Nash equilibrium of the AS model.[11] Temporarily suppose that the Stage 1 entry decision involves an *entry threshold* $\bar{s} \in [0,1]$ such that bidder i chooses to enter if and only if $s_i \geq \bar{s}$; the proof establishes that any equilibrium in the class considered has a payoff-equivalent threshold representation. The (selected) distribution of values among entrants at threshold $\bar{s}$ is then

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

When the Stage 2 mechanism is a second-price auction, it is a weakly dominant strategy for entrants to bid values. An entrant with value v thus outbids any given *potential* rival when that rival either does not enter (probability $\bar{s}$) or enters and draws a value below v (probability $(1-\bar{s})F^*(v;\bar{s})$). Let $F^*_{1:N-1}(v;\bar{s})$ denote the probability that an entrant with value v wins against $N-1$ potential rivals who enter according to $\bar{s}$. By independence, we then have:

$$F^*_{1:N-1}(v;\bar{s}) \equiv \left[\bar{s} + (1-\bar{s})F^*(v;\bar{s})\right]^{N-1}. \tag{2}$$

[10] As discussed in Riley and Samuelson (1981), the class of RS auctions includes any anonymous mechanism such that only the high bidder wins, under which there is a unique symmetric, monotonic bidding strategy. Stage 2 identification results are known for a variety of RS auctions, of which the most important are the first-price, second-price, English, and Dutch bidding rules. See Athey and Haile (2005) for further details.

[11] As usual, there may exist other, asymmetric equilibria. Note that these could in principle be nested by our discussion of bidder asymmetry in Section 4.2: simply introduce a set of trivial bidder types. For the moment, we focus on the symmetric case.

By standard arguments in mechanism design, the expected Stage 2 profit of an entrant drawing value v_i against $N-1$ potential rivals at entry threshold $\bar{s}$ is then[12]

$$\pi(v_i;\bar{s},N) = \int_0^{v_i} F^*_{1:N-1}(y;\bar{s})dy.$$

Now consider the Stage 1 entry decision of a bidder with signal s_i. Facing $N-1$ rivals who enter according to $\bar{s}$, this bidder's expected profit from competing in Stage 2 is

$$\begin{aligned}\Pi(s_i;\bar{s},N) &\equiv E_V\left[\pi(V_i;\bar{s},N)|s_i\right] \\ &= \int_0^{\bar{v}} [1-F(y|s_i)]\cdot F^*_{1:N-1}(y;\bar{s})dy. \qquad (3)\end{aligned}$$

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

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

Now consider any candidate *equilibrium* threshold s^*. By standard arguments, if s^* implies nontrivial entry, a bidder drawing signal $S_i = s^*$ must be indifferent to entry when facing $N-1$ potential rivals who also enter according to s^*:

$$\Pi(s^*;s^*,N) \equiv c.$$

Since $\Pi(s_i;\bar{s},N)$ is increasing in s_i and strictly increasing in $\bar{s}$, this breakeven condition uniquely determines the equilibrium threshold s^*. We formalize this intuition via the following proposition:

Proposition 1. *Under Assumptions 1 and 2, any symmetric pure strategy Bayesian Nash equilibrium has a payoff-equivalent equilibrium in which Stage 1 entry decisions involve a signal threshold s^* such that bidder i enters if and only if $S_i \geq s^*$. This threshold is uniquely*

[12]See, e.g., Krishna (2009) for details.

determined as follows.

- *If* $\Pi(0; 0, N) > c$*, then* $s^* = 0$ *and all potential bidders always enter.*

- *If* $\Pi(1; 1, N) < c$*, then* $s^* = 1$ *and no potential bidder ever enters.*

- *Otherwise,* s^* *satisfies the breakeven condition*

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

  where $\Pi(s_i; s^*, N)$ *is defined as in Equation (3).*

Further, considered as a function of (c, N)*, the equilibrium threshold* $s^*(c, N)$ *satisfies the following monotonicity properties:*

- *For any* $N \geq 1$*,* $s^*(c, N)$ *is continuous and weakly increasing in* c*, with strict monotonicity whenever* $s^*(c, N) \in (0, 1)$*.*

- *For any* $c \geq 0$*,* $N' > N$ *implies* $s^*(c, N') \geq s^*(c, N)$*. If in addition* $s^*(c, N) \in (0, 1)$*, then* $s^*(c, N') > s^*(c, N)$ *and* $s^*(c, N') \in (0, 1)$*.*

While our discussion assumed a second-price auction in Stage 2, this result in fact applies under a wide range of Stage 2 mechanisms. In a companion paper (Gentry and Li (2012)), we show that the characterizations of equilibrium profit $\Pi(s_i; \bar{s}, N)$ and the equilibrium entry threshold s^* in Proposition 1 extend to any mechanism in the class considered by Riley and Samuelson (1981): that is, any anonymous auction with a unique symmetric equilibrium such that bidding strategies are monotonic.[13] This class of mechanisms in turn contains the standard first-price, second-price, English, and Dutch rules, which together encompass the vast majority of auctions seen in practice. In Appendix B, we state a version of Proposition 1 which applies to other mechanisms; for completeness, the proof in Appendix C establishes this more general result.

[13]See formal statement and additional discussion in Appendix B. For completeness, the proof of Proposition 1 in Appendix C establishes the result for the general case.

2 Identification in the AS model

We consider identification based on a large sample of auctions from an AS process $\mathcal{L}$, where for each auction ℓ the following variables are observed: number of potential bidders N_ℓ, number of entrants n_ℓ, and a vector of submitted bids $\mathbf{b}_\ell$. 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. Optionally, the econometrician may also observe a vector of auction-level instruments $\mathbf{Z}$, which are assumed to shift entry behavior without affecting the fundamental distribution $F(v, s)$. In the symmetric case considered here, natural candidates for inclusion in $\mathbf{Z}$ would be entry fees or other factors affecting auction-level entry costs. For the moment, we formalize this notion by permitting entry costs to vary deterministically across auctions ℓ as a function of $\mathbf{Z}$: $c_\ell = c(\mathbf{z}_\ell)$, where c_ℓ is the cost realization prevailing in auction ℓ.[14] In environments with asymmetric bidders, $\mathbf{Z}$ could also include the types of bidder i's rivals; we consider this extension in Section 4.2 below.[15] As usual, all results generalize immediately to any further set of auction-level covariates $\mathbf{X}_\ell$: simply repeat all arguments conditional on realization $\mathbf{x}$.

To identify the joint distribution $F(v, s)$, we must be able to attribute at least some observable variation to changes in the dimension S; i.e. to observables affecting equilibrium entry behavior through channels other than $F(v, s)$. In the current context, this requires either variation in N_ℓ on a nontrivial set $\mathcal{N}$, or variation in $\mathbf{Z}_\ell$ on a nontrivial set $\mathcal{Z}$. For the moment, we assume this variation is *excludable* in the following sense:

Assumption 3. *For all $N \in \mathcal{N}$ and $\mathbf{z} \in \mathcal{Z}$, $F(v, s|N, \mathbf{z}) = F(v, s)$ and $c(N, \mathbf{z}) = c(\mathbf{z})$, with all equalities conditioned also on $\mathbf{x}$ if further covariates are observed.*

[14] Our discussion of unobserved heterogeneity in Section 4.1 relaxes this restriction to permit entry costs to vary stochastically across auctions, where cost realizations are observed to bidders but not the econometrician.

[15] The asymmetric case introduces the additional complication of potential multiple equilibria, which leads us to discuss it separately as an extension.

Excludable variation in N_ℓ is frequently cited as a basis for auction-related hypothesis testing: for instance, Haile, Hong, and Shum (2003) use such variation to construct a test for common values, and MSX use it to construct tests for competing entry specifications. Excludable variation in $\mathbf{z}_\ell$ directly extends the long tradition of instrumental variables in econometrics.[16] We therefore adopt Assumption 3 as a baseline. In Section 4.1, however, we establish that our results in fact extend to the much more general case of excludability conditional on the realization of an nonseparable unobserved auction-level random variable U.

Finally, for the moment, we assume that $\mathcal{L}$ is an AS process involving no unobserved heterogeneity:

Assumption 4. *Outcomes $(N_\ell, n_\ell, \mathbf{b}_\ell)$ represent repeated draws from symmetric equilibrium play under AS fundamentals $F(v, s)$ and $c(\cdot)$ which are either* invariant across auctions *or* invariant conditional on covariates $\mathbf{x}$ *if these are observed.*

While this restriction is quite typical in applications, it is also strong in the sense that it implicitly requires the econometrician to observe the same auction-level information as potential bidders. Consequently, while we maintain Assumption 4 to establish intuition, we also consider identification under nonseparable unobserved heterogeneity, interpreted as an auction-level factor U known to bidders but not the econometrician. As noted above, Section 4.1 establishes that our baseline results in fact extend to this much more general case.

2.1 Directly identified objects

For each $(N, \mathbf{z}) \in \mathcal{L}$, a large sample from process $\mathcal{L}$ will directly identify two statistical objects. First, given an arbitrary entry threshold $\bar{s}$, the probability that any particular bidder enters is simply $1 - \bar{s}$. We can thus identify the entry threshold $\hat{s}_N(\mathbf{z})$ prevailing at

[16] 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 (2010b) for nonparametric identification of multinomial choice demand models, to name only a few.

each $(N, \mathbf{z})$ directly from observed entry decisions:

$$\hat{s}_N(\mathbf{z}) \equiv 1 - \frac{E[n|N, z]}{N}.$$

Second, in a second-price auction, it is a weakly dominant strategy to bid values. The distribution of values $\hat{F}(v|N, \mathbf{z})$ prevailing at each $(N, \mathbf{z})$ is thus identical to the corresponding conditional distribution of bids $\hat{G}(b|N, \mathbf{z})$:

$$\hat{F}(v|N, \mathbf{z}) \equiv \hat{G}(v|N, \mathbf{z}).$$

It is straightforward to establish identification of $\hat{F}(v|N, \mathbf{z})$ under other common auction types using well-known results in the literature. Our fundamental results extend to any of these. See Appendix B for further discussion.

Next relate the objects above to their equilibrium counterparts. In equilibrium, the probability that bidder i enters is $1 - s_N^*(c)$, and the distribution of values conditional on entry is $F^*(v; s_N^*(c)) \equiv F(v|S_i \geq s_N^*(c))$. It follows that

$$s_N^*(c(\mathbf{z})) \equiv \hat{s}_N(\mathbf{z}) \tag{5}$$

and

$$F^*\left[v; s_N^*(c(\mathbf{z}))\right] \equiv \hat{F}(v|N, \mathbf{z}). \tag{6}$$

Finally, consider variation across $(N, \mathbf{z}) \in \mathcal{L}$. Let $\mathcal{S}$ denote the set of entry thresholds identified by process $\mathcal{L}$:

$$\mathcal{S} \equiv \{s \in [0, 1] | s = \hat{s}_N(\mathbf{z}) \text{ for some } (N, \mathbf{z}) \in \mathcal{L}\}. \tag{7}$$

We can then restate the observations above as follows: the objects directly identified by process $\mathcal{L}$ are the threshold set $\mathcal{S}$ itself and the *ex post* distribution $F^*(v; \hat{s})$ for each $\hat{s} \in$

$\mathcal{S}$. The threshold set $\mathcal{S}$ thus summarizes the information generated by entry and bidding behavior under AS process $\mathcal{L}$.

2.2 Complete variation: point identification

We begin with the ideal case: suppose the econometrician observes a set of instruments $\mathbf{Z}$ (with support on some set $\mathcal{Z}$) which induce an identified set $\mathcal{S}$ having *nonempty interior*. Since N is discrete by construction, this supposition requires at least one element of $\mathbf{Z}$ to be continuous. As above, elements in $\mathbf{Z}$ must be *excludable* in the sense that shift entry behavior without affecting $F(v, s)$; we here focus on cost shifters for simplicity.

Choose any $\hat{s} \in \text{int}(\mathcal{S})$, and consider the conditional distribution $F(\cdot|\hat{s})$. By Equation (1), we have

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

which in turn implies

$$F(v|\hat{s}) = -\frac{\partial}{\partial \hat{s}}[(1-\hat{s})F^*(v; \hat{s})]. \tag{8}$$

But $\hat{s}$ and $F^*(v; \hat{s})$ are identified for all $\hat{s} \in \mathcal{S}$. Consequently, if $\hat{s} \in \text{int}(\mathcal{S})$, we can obtain the RHS derivative exactly by taking limits of identified quantities. It follows that $F(\cdot|\hat{s})$ is point-identified for all $\hat{s} \in \text{int}(\mathcal{S})$.[17]

In turn, point identification of $F(\cdot|s)$ permits point identification of $c(\mathbf{z})$. To see this, choose any $\mathbf{z} \in \mathcal{Z}$ which induces nontrivial entry: that is, such that $\hat{s}_N(\mathbf{z}) \in (0, 1)$ for some $N \in \mathcal{N}$. By Proposition 1 and identity (6), $c(\mathbf{z})$ must then satisfy the breakeven condition

$$c(\mathbf{z}) \equiv \int_0^{\bar{v}} [1 - F(y|\hat{s}_N(\mathbf{z}))] \cdot \left[\hat{s}_N(\mathbf{z}) + (1-\hat{s}_N(\mathbf{z}))\hat{F}(y|N, \mathbf{z})\right]^{N-1} dy.$$

Thus identification of $F(\cdot|S)$ at $S = \hat{s}_N(\mathbf{z})$ implies identification of $c(\mathbf{Z})$ at $\mathbf{Z} = \mathbf{z}$. It follows

[17] At the cost of somewhat more cumbersome notation, the argument extends immediately to the closure of the set.

that $c(\mathbf{z})$ is point-identified at any $\mathbf{z} \in \mathcal{Z}$ such that $\hat{s}_N(\mathbf{z}) \in \text{int}(\mathcal{S})$ for some N; that is, for any realization $\mathbf{z}$ inducing local variation in entry.

Combining these observations leads to the AS-model equivalent of a full-support condition: the AS model is point-identified (almost) everywhere if we observe data generated at (almost) every $\hat{s} \in (0,1)$. Our next proposition formalizes this intuition.

Proposition 2. *Suppose the econometrician observes instruments $\mathbf{Z}$ satisfying Assumption 4, which have positive support on a set $\mathcal{Z} \subset R^k$. Then the following statements hold:*

1. *If $\hat{s} \in int(\mathcal{S})$, then $F(\cdot|S)$ is point-identified at $S = \hat{s}$.*

2. *If $\hat{s}_N(\mathbf{z}) \in int(\mathcal{S})$ for some $N \in \mathcal{L}$, then $c(\mathbf{Z})$ is point-identified at $\mathbf{Z} = \mathbf{z}$.*

3. *If $cl(\mathcal{S}) = [0,1]$, then $F(v|s)$ is* fully identified: *that is, point-identified almost everywhere.*

One could also state sufficient conditions for full identification in terms of model fundamentals: for instance, if $c(\mathcal{Z})$ spans the interval $[0, \bar{v}]$, then any threshold $\hat{s} \in [0,1]$ will be an equilibrium for some $\mathbf{z}$.[18] In practice, however, such conditions are difficult to check, whereas restrictions on $\mathcal{S}$ can be verified directly. We therefore prefer to state results in terms of the identified set $\mathcal{S}$.

2.3 Incomplete variation: partial identification

While the case of complete entry variation is useful as an ideal, there are many applications of interest where it will not hold. In particular, most auction studies either assume that bidders are symmetric or consider a discrete set of bidder types. While such applications typically control for a set of auction-level covariates $\mathbf{X}$, these are often difficult to exclude *a priori* from the distribution $F(v, s|\mathbf{x})$. Consequently, all informative entry variation will

[18]Note that when $c = 0$ it is always a best response to enter (since $\Pi(s_i; \bar{s}, N) \geq 0$), and when $c = \bar{v}$ it is always a best response not to enter (since $\Pi(s_i; \bar{s}, N) \leq \bar{v}$ with certainty). Appendix C establishes continuity of $\Pi(s_i; \bar{s}, N)$ in its first two arguments, which in turn implies the full-support condition required.

arise through changes in potential competition, and the set of thresholds $\mathcal{S}(\mathbf{x})$ prevailing at each $\mathbf{x}$ will necessarily be discrete. Furthermore, even in applications with excludable covariates, elements of $\mathbf{Z}$ may be either discrete or otherwise insufficient to induce complete entry variation. This section develops identification results applicable to such more general cases, deriving identified bounds on AS fundamentals under arbitrary entry variation which nest the conclusions of Proposition 2 as a special case. As above, we abstract from further discussion of nonexcludable covariates $\mathbf{X}$, but all results extend immediately conditional on $\mathbf{x}$.

More precisely, our goal in this section is as follows: taking the identified threshold set $\mathcal{S}$ as given, derive a map from observed entry variation as given into identified bounds on AS fundamentals. This map naturally extends the logic of Proposition 2: where local variation in $\hat{s}$ is available, calculate the derivative (8) exactly; everywhere else, approximate via small finite differences $\Delta\hat{s}$:

$$F(v|\hat{s}) \approx -\frac{\Delta[(1-\hat{s})F^*(v;\hat{s})]}{\Delta\hat{s}}. \tag{9}$$

We now formalize this intuition. Define *nearest-neighbor functions* $t^+(s)$ and $t^-(s)$ as follows:

$$t^+(s) = \begin{cases} \inf\{t \in \mathcal{S}|t > s\} & \textit{if } \max\{\mathcal{S}\} > s; \\ 1 & \textit{otherwise.} \end{cases}$$

$$t^-(s) = \begin{cases} \sup\{t \in \mathcal{S}|t < s\} & \textit{if } \min\{\mathcal{S}\} < s \\ 0 & \textit{otherwise.} \end{cases}$$

By construction, $t^+(s)$ and $t^-(s)$ are then the nearest upper and lower neighbors of s in $\mathcal{S}$ (or uninformative values where such neighbors are missing). For each $\hat{s} \in \mathcal{S}$, evaluating (9) at these neighbors will produce two natural approximations to $F(v|\hat{s})$:

$$\check{F}^{+}(v|\hat{s}) = \begin{cases} \lim_{t \uparrow t^{-}(\hat{s})} \left\{ \frac{(1-t)F^{*}(v;t) - (1-\hat{s})F^{*}(v;\hat{s})}{\hat{s}-t} \right\} & \text{if } t^{-}(\hat{s}) \in \mathcal{S}; \\ 1 & \text{otherwise.} \end{cases} \tag{10}$$

$$\check{F}^{-}(v|\hat{s}) = \begin{cases} \lim_{t \downarrow t^{+}(\hat{s})} \left\{ \frac{(1-\hat{s})F^{*}(v;\hat{s}) - (1-t)F^{*}(v;t)}{t-\hat{s}} \right\} & \text{if } t^{+}(\hat{s}) \in \mathcal{S}; \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

By definition, $\check{F}^{+}(v|\hat{s})$ and $\check{F}^{-}(v|\hat{s})$ are identified for all $\hat{s} \in \mathcal{S}$, and by stochastic ordering it can be shown that they also bound $F(\cdot|\hat{s})$:

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

with equality if $\hat{s} \in \text{int}(\mathcal{S})$. Thus $\check{F}^{+}(v|\hat{s})$ and $\check{F}^{-}(v|\hat{s})$ provide a basis for partial identification of $F(\cdot|s)$, a result we formalize in the next proposition.

Proposition 3 (Bounds on $F(v|s)$)**.** *For any $s \in [0,1]$, define $F^{+}(v|s)$ and $F^{-}(v|s)$ as follows:*

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

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

Then $F^{+}(v|s)$ 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 $s \in int(\mathcal{S})$.

It only remains to translate identified bounds on $F(\cdot|s)$ into identified bounds on $c(\cdot)$.

As above, at any $(N, \mathbf{z})$ with nontrivial entry, the corresponding entry threshold $\hat{s}_N(\mathbf{z})$ must satisfy the breakeven condition

$$c(\mathbf{z}) \equiv \int_0^{\bar{v}} [1 - F(y|\hat{s}_N(\mathbf{z}))] \cdot \left[\hat{s}_N(\mathbf{z}) + (1 - \hat{s}_N(\mathbf{z}))\hat{F}(y|N, \mathbf{z})\right]^{N-1} dy. \tag{12}$$

The RHS integral is identified up to $F(\cdot|\hat{s}_N(\mathbf{z}))$ and is decreasing in $F(y|\hat{s}_N(\mathbf{z}))$ at each y. Thus substituting the identified bounds $F^+(\cdot|\hat{s}_N(\mathbf{z}))$ and $F^-(\cdot|\hat{s}_N(\mathbf{z}))$ into (12) yields identified bounds on $c(\cdot)$ at $\mathbf{z}$:

$$c_N^+(\mathbf{z}) = \int_0^{\bar{v}} \left[1 - F^-(y|\hat{s}_N(\mathbf{z}))\right] \cdot \left[\hat{s}_N(\mathbf{z}) + (1 - \hat{s}_N(\mathbf{z}))\hat{F}(y; N, \mathbf{z})\right]^{N-1} dy. \tag{13}$$

$$c_N^-(\mathbf{z}) = \int_0^{\bar{v}} \left[1 - F^+(y|\hat{s}_N(\mathbf{z}))\right] \cdot \left[\hat{s}_N(\mathbf{z}) + (1 - \hat{s}_N(\mathbf{z}))\hat{F}(y; N, \mathbf{z})\right]^{N-1} dy \tag{14}$$

As usual, pooling these restrictions across N will generate tighter bounds on $c(\mathbf{z})$, which in turn yields the following proposition:

Proposition 4. *Choose any $(N, \mathbf{z}) \in \mathcal{L}$ with nontrivial entry, define $c_N^+(\mathbf{z})$ and $c_N^-(\mathbf{z})$ as above, and construct $c^+(\mathbf{z})$ and $c^-(\mathbf{z})$ as follows:*

$$\begin{aligned} c^+(\mathbf{z}) &= \min_{N \in \mathcal{N}} c_N^+(\mathbf{z}) \\ c^-(\mathbf{z}) &= \max_{N \in \mathcal{N}} c_N^-(\mathbf{z}). \end{aligned}$$

Then $c^+(\mathbf{z})$ and $c^-(\mathbf{z})$ are identified and $c^+(\mathbf{z}) \geq c(\mathbf{z}) \geq c^-(\mathbf{z})$, with equality if $\bar{s}_N(z) \in int(\mathcal{S})$ for some $N \in \mathcal{N}$.[19]

Thus, to recapitulate: given an identified set $\mathcal{S}$, we obtain identified bounds on the AS fundamentals $F(v|s)$ and $c(\mathbf{z})$. These bounds collapse to equalities on the interior of $\mathcal{S}$, thereby extending the special case of complete entry variation in Proposition 2 to the general case of an arbitrary entry set. Propositions 3 and 4 thus define a map from observed entry

[19] Obviously, at any $(N, \mathbf{z})$ such that entry is trivial ($\hat{s}_N(\mathbf{z}) = 0$), only the upper bound would apply.

variation into identified sets in a general class of auctions with endogenous and selective entry.

2.4 Sharp identification

While the bounds in Proposition 3 represent natural, intuitive, and directly estimable approximations to $F(v|s)$, they do not exploit all information in the model. In particular, pooling bounds on $c(\cdot)$ across N generates a set of cross-equation restrictions, which might in principle lead to sharper bounds on $F(v|s)$. This section characterizes the sharp identified set in the AS model. The argument proceeds as follows: we first derive a necessary and sufficient condition to determine whether any given bounds are sharp, then discuss how to use this condition to construct the sharp identified set. We thus characterize in an important sense what *can* be known in environments with entry; that is, what can be inferred from an economic model of entry *per se*, without parametric restrictions on the nature of selection. For clarity, we focus discussion on the baseline case where all entry variation is induced by N, but all results extend readily to incorporate excludable instruments $\mathbf{Z}$.[20]

Let $\tilde{c}$ denote any scalar and $\tilde{F}(V|S)$ denote any conditional distribution having support on $\mathcal{V} \times [0,1]$, and define a *candidate model* for process $\mathcal{L}$ as follows:

Definition 1. A *candidate model* for process $\mathcal{L}$ is a pair $\{\tilde{F}(\cdot|\cdot), \tilde{c}\}$ such that $\tilde{F}(\cdot|\cdot)$ implies a joint distribution satisfying Assumption 1, and $\{\tilde{F}(\cdot|\cdot), \tilde{c}\}$ rationalizes observed outcomes as an equilibrium under Proposition 1.

We now unpack this definition. For each $N \in \mathcal{N}$, define a functional transformation

[20]Obviously, if the instruments $\mathbf{Z}$ induce complete entry variation, the AS model will be fully identified and our bounds will be trivially sharp. Otherwise, the question of sharpness will remain relevant. As a general rule, however, more informative entry variation will produce bounds which are more likely to be sharp, which motivates our focus on the least informative N-only case.

$\hat{T}_N(\cdot)$, a function $\hat{\lambda}_N(\cdot)$, and a constant $\hat{\kappa}_N$ as follows:

$$\begin{aligned}
\hat{T}_N(F) \quad &\equiv - \int F(y) \cdot \left[\hat{s}_N + (1 - \hat{s}_N) \cdot \hat{F}(y|N)\right]^{N-1} dy. && (15)\\
\hat{\lambda}_N(y) \quad &\equiv \quad (1 - \hat{s}_N) \cdot \hat{F}(y|N) && (16)\\
\hat{\kappa}_N \quad &\equiv \int \left[\hat{s}_N + (1 - \hat{s}_N)\hat{F}(y|N)\right]^{N-1} dy. && (17)
\end{aligned}$$

By construction, these objects are identified for all $N \in \mathcal{L}$, and by equilibrium we have

$$\begin{aligned}
c \quad &\equiv \int [1 - F(y|\hat{s}_N)] \cdot \left[\hat{s}_N + (1 - \hat{s}_N) \cdot \hat{F}(y|N)\right]^{N-1} dy \\
&\equiv \quad \hat{\kappa}_N + \hat{T}_N \left[F(y|\hat{s}_N)\right]
\end{aligned}$$

for each $N \in \mathcal{N}$. We can thus restate the restrictions generated by the AS model as follows:

Lemma 1 (Candidate model)**.** $\{\tilde{F}(\cdot|\cdot), \tilde{c}\}$ *is a candidate model for process* $\mathcal{L}$ *if and only if the following conditions hold:*

1. Ordered distribution: *for each* $s \in [0, 1]$, $\tilde{F}(\cdot|s)$ *is a distribution, and* $\tilde{F}(\cdot|s)$ *is stochastically ordered in* s*:*

$$s' \geq s \text{ implies } \tilde{F}(\cdot|s') \leq \tilde{F}(\cdot|s). \tag{18}$$

2. Selection: *for each* $y \in [\underline{v}, \bar{v}]$ *and for all* $N \in \mathcal{L}$,

$$\hat{\lambda}_N(y) = \int_{\hat{s}_N}^{1} \tilde{F}(y|s)ds. \tag{19}$$

3. Equilibrium: *for each* $N \in \mathcal{L}$,

$$\hat{T}_N \left[\tilde{F}(\cdot|\hat{s}_N)\right] + \hat{\kappa}_N = \tilde{c}. \tag{20}$$

We next formally define the terms *pointwise identified set* and *pointwise sharp identified set*:

Definition 2 (Pointwise identified and pointwise sharp)**.** A set $\mathcal{F} \subset [0,1] \times \mathcal{V} \times \mathcal{S}$ is *pointwise identified* if every candidate distribution $\tilde{F}(\cdot|\cdot)$ is a *selection from* $\mathcal{F}$: that is, if for each candidate model $\{\tilde{F}(\cdot|\cdot), \tilde{c}\}$, the tuple $(\tilde{F}(y|s), y, s) \in \mathcal{F}$ for all $y \in \mathcal{V}$ and $s \in \mathcal{S}$. The pointwise identified set $\mathcal{F}$ is *pointwise sharp* if for every tuple $(\Gamma, v, s) \in \mathcal{F}$, there exists a candidate model $\{\tilde{F}(\cdot|\cdot), \tilde{c}\}$ such that $\Gamma = \tilde{F}(y|s)$.

It is clear that the set implied by Proposition 3 is a pointwise identified set, as any candidate model belongs to this set. Let $\mathcal{F}$ be either this set or any pointwise identified subset of it, and consider establishing sharpness at a point $(\bar{\Gamma}, \bar{y}, \bar{s})$ on the upper envelope of $\mathcal{F}$. For there to exist a candidate model $\{\tilde{F}(\cdot|\cdot), \tilde{c}\}$ passing through $(\bar{\Gamma}, \bar{y}, \bar{s})$, the maximum expected profit across all distributions feasible at $(\bar{\Gamma}, \bar{y}, \bar{s})$ must be at least as large as c^-. But the maximum expected profit will occur at the minimal feasible distribution, which can be immediately ascertained from $\mathcal{F}$. Extending this intuition leads to a formal test for pointwise sharpness of $\mathcal{F}$:

Proposition 5. *Let $\mathcal{F}$ be a pointwise identified set, c^+ and c^- be as in Proposition 4, and $F^+(y|s) = \max\{\Gamma | (\Gamma, y, s) \in \mathcal{F}\}$ and $F^-(y|s) = \min\{\Gamma | (\Gamma, y, s) \in \mathcal{F}\}$ be the bounds on $F(\cdot|\cdot)$ implied by $\mathcal{F}$. For each point $(\bar{y}, \bar{s}_k) \in \mathcal{V} \times \mathcal{S}$, define an* upper test function $\ddot{F}^+_{\bar{y}k}(\cdot)$ *and a* lower test function $\ddot{F}^-_{\bar{y}k}(\cdot)$ *as follows:*

$$\ddot{F}^+_{\bar{y}k}(y) \equiv \begin{cases} F^-(y|\bar{s}_k) & \text{if } y < \bar{y} \\ \max\{F^-(y|\bar{s}_k), F^+(\bar{y}|\bar{s}_k)\} & \text{if } y \geq \bar{y}; \end{cases}$$

$$\ddot{F}^-_{\bar{y}k}(y) \equiv \begin{cases} \min\{F^+(y|\bar{s}_k), F^-(\bar{y}|\bar{s}_k)\} & \text{if } y < \bar{y} \\ F^+(y|\bar{s}_k) & \text{if } y \geq \bar{y}. \end{cases}$$

Then $\mathcal{F}$ is pointwise sharp if and only if for each $\bar{y} \in \mathcal{V}$ and $\bar{s}_k \in \mathcal{S}$,

$$\hat{\kappa}_k + \hat{T}_k \left[\ddot{F}^+_{\bar{y}k}(\cdot)\right] \;\; \geq \;\; c^- \tag{21}$$

and

$$\hat{\kappa}_k + \hat{T}_k \left[\ddot{F}^{-}_{\bar{y}k}(\cdot) \right] \;\; \leq \;\; c^+. \tag{22}$$

Note that by construction $\ddot{F}^{+}_{\bar{y}k}(y)$ and $\ddot{F}^{-}_{\bar{y}k}(y)$ represent the *minimal feasible distribution attaining* $F^+(\cdot|\cdot)$ and the *maximal feasible distribution attaining* $F^-(\cdot|\cdot)$ at $(\bar{y}, \bar{s}_k)$, respectively. Proposition 5 thereby implement the intuition described above.

A crucial point in Proposition 5 is that the conditions given are both necessary and sufficient: that is, Proposition 5 is both a test for and a characterization of the pointwise sharp identified set. This fact in turn suggests an algorithm for constructing the pointwise sharp identified set: beginning from an initial identified set $\mathcal{F}_0$ implied by Proposition 3, test sharpness at each $(\bar{y}, \bar{s}) \in \mathcal{V} \times \mathcal{S}$. At each $(\bar{y}, \bar{s})$ where at least one bound is infeasible, discard all points failing the corresponding condition, and iterate until no further improvement is possible. The next proposition formalizes details of this algorithm:

Proposition 6. *Let $\mathcal{F}_0$ be the pointwise identified set implied by Proposition 3, and consider the following iterative algorithm (where $\mathcal{F}_j$ is the pointwise identified set at the jth iteration):*

1. *For each $\bar{s} \in \mathcal{S}$, obtain the upper and lower bounds on $F(\cdot|\bar{s})$ implied by $\mathcal{F}_j$:*

$$\begin{aligned} F^+_j(y|s) &= \max\{\Gamma|(\Gamma, y, s) \in \mathcal{F}_j\} \\ F^-_j(y|s) &= \min\{\Gamma|(\Gamma, y, s) \in \mathcal{F}_j\}. \end{aligned}$$

2. *For each $k \in \mathcal{N}$ and $\bar{y} \in \mathcal{V}$, evaluate $\hat{\kappa}_k + \hat{T}_k \left[\ddot{F}^{+}_{\bar{y}k}(\cdot) \right] \geq c^-$. If true, keep $F^+_{j+1}(\bar{y}|\hat{s}_k) = F^+_j(\bar{y}|\hat{s}_k)$. Otherwise, find the unique constant $\Gamma^+_{\bar{y}k}$ such that*

$$c^- = \hat{\kappa}_k + \hat{T}_k \left[\max\{F^-_j(Y|\hat{s}_k), \mathbf{1}[Y \geq \bar{y}] \cdot \Gamma^+_{\bar{y}k}\} \right],$$

and update $F^+_{j+1}(\bar{y}|\hat{s}_k) = \Gamma^+_{\bar{y}k}$.

3. *For each $k \in \mathcal{N}$ and $\bar{y} \in \mathcal{V}$, evaluate $\hat{\kappa}_k + \hat{T}_k \left[\ddot{F}^-_{\bar{y}k}(\cdot)\right] \leq c^+$. If true, keep $F^-_{j+1}(\bar{y}|\hat{s}_k) = F^-_j(\bar{y}|\hat{s}_k)$. Otherwise, find the unique constant $\Gamma^-_{\bar{y}k}$ such that*

$$c^+ = \hat{\kappa}_k + \hat{T}_k \left[\min\{F^+_j(Y|\hat{s}_k), \mathbf{1}[Y \leq \bar{y}] \cdot \Gamma^-_{\bar{y}k}\}\right],$$

and update $F^-_{j+1}(\bar{y}|\hat{s}_k) = \Gamma^-_{\bar{y}k}$.

4. *Obtain a new identified set $\mathcal{F}_{j+1}$ corresponding to these updated bounds:*

$$\mathcal{F}_{j+1} = \left\{(\Gamma, y, s)|\Gamma \in \left[F^-_{j+1}(y|s), F^+_{j+1}(y|s)\right]\right\}.$$

The resulting sequence $\{\mathcal{F}_j\}$ converges to a fixed point, which is the pointwise sharp identified set.

The key to Proposition 5 is that order of rejection is irrelevant: refinement of (say) the upper bound F^+ at point $(\bar{y}, \hat{s}_k)$ depends on values of F^+ at no other point, and a tighter lower bound $F^-(\cdot|\cdot)$ can only change the decision at $F^+(\bar{y}|\hat{s}_k)$ from "accept" to "reject." Hence ordering of the algorithm will not affect the set of points eventually rejected. Note also that while Proposition 5 allows for more general cases, convergence in our numerical examples (Appendix B) was always achieved in a single step. This is not surprising since the algorithm discards *all* points inconsistent with identified set $\mathcal{F}_{j-1}$ at iteration j.

Finally, Propositions 5 and 6 imply that improvement in existing bounds on $F(v|s)$ is possible only by comparing feasible expected profits with the bounds c^+ and c^- in Proposition 4. This in turn implies that even refinement of $\mathcal{F}$ can yield no further information on c:

Corollary 1. *The bounds c^+ and c^- defined in Proposition 4 are sharp.*

Appendix A applies the results in Propositions 5 and 6 to a set of numerical examples, comparing initial identified bounds to the sharp identified set for a variety of selectivity parameters and competition structures. The clear and encouraging message of these comparisons is that scope for improvement on Proposition 3 is minimal: in almost all cases, our initial bounds were sharp everywhere they were informative. Exceptions to this rule were of necessity quite artificial: large gaps early in an otherwise consecutive sequence of competition levels, in which case improvement is sometimes possible over the gap. Thus while our results indicate when and how the bounds in Proposition 3 can be refined, simulation evidence suggests that such refinement will seldom be required in practice.

2.5 Bounds in the S and LS cases

To conclude this section, we explore how our proposed bounds behave applied to data generated by the S and LS polar cases. As noted above, the S model corresponds 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 set $v_i \equiv F_v^{-1}(s_i)$. In this case, $F(v|\hat{s})$ is degenerate for each $\hat{s} \in \mathcal{S}$, with all mass at $\hat{v} = F_v^{-1}(\hat{s})$. Meanwhile, the bounds $F^+(v|\hat{s})$ and $F^-(v|\hat{s})$ in Proposition 3 will in general be well-defined distributions. Hence $F^+(v|\hat{s})$ and $F^-(v|\hat{s})$ do not collapse to $F(v|\hat{s})$, and consequently derived bounds on $c(\cdot)$ also do not collapse.

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

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

By definition of $F^+(v|s)$ and $F^-(v|s)$ in Proposition 3, this equality in turn implies that informative bounds on $F(v|s)$ will collapse to $F_v(v)$ for any s. As long as $\mathcal{N}$ contains at least three elements, at least one $s \in \mathcal{S}(\mathbf{z})$ will have two informative bounds at each $\mathbf{z} \in \mathcal{Z}$, and bounds at this s will point-identify both $F_v(v)$ and $c(\mathbf{z})$. At least in terms of identification, therefore, estimation based on the general AS model entails only minor losses relative to estimation based on the more restrictive LS polar case.

3 Policy analysis under partial identification

Counterfactual policy analysis is a leading motive for structural estimation. Such analysis is straightforward when model fundamentals are point identified, but becomes more challenging when they are only partially identified. This is particularly true in the AS model: policy choices will affect outcomes directly, through the Stage 1 entry threshold s^* and through the selected Stage 2 distribution $F^*(\cdot; s)$ and valid counterfactual analysis must account for all three effects. This section addresses this problem directly, using the objects identified above to construct bounds on counterfactual revenue which apply to a wide range of counterfactual policies under endogenous and arbitrarily selective entry. We thus establish that nonparametric analysis based on the AS model can yield practically relevant policy insights even in the presence of partial identification.

More precisely, the problem addressed in this section can be stated as follows: given a counterfactual mechanism M satisfying certain regularity properties, construct bounds on the expected equilibrium revenue $R^*_M(N)$ that would obtain if the auctions in $\mathcal{L}$ were instead run according to M at competition N. We address this problem within a general class of counterfactual mechanisms we call *RS auctions* (after the work of Riley and Samuelson (1981)), which includes standard first-price, second-price, English, and Dutch auctions:

Definition 3. 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. Entry and bidding realizations affect the auctioneer's award decision only through the highest bid.[21]

[21] This restriction rules out mechanisms where the auctioneer makes award decisions conditional on, e.g., realized number of entrants. Note that the statement permits award decisions to be random conditional on the highest bid, or to depend on auction-level characteristics or covariates.

4. For any 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 Samuelson (1981), the key determinant of equilibrium outcomes in any RS auction M is the *award rule* $\alpha_M(\cdot) : \mathbb{R} \to [0,1]$ in the direct equivalent to M, where $\alpha_M(Y)$ represents the probability that M results in sale when the maximum reported value among bidders is Y. For instance, a public reserve price would correspond to the award rule $\alpha_M(y) = \mathbf{1}[y \geq r]$, and a secret reserve price to the award rule $\alpha_M(y) = F_r(\beta(y))$. For current purposes, we impose the following additional regularity conditions on the mechanism M:

Assumption 5. *Let $\mathbf{R}_{-i} \in \{0 \cup \mathcal{V}\}^{N-1}$ denote values reported by rivals of bidder i (where type zero indicates no entry). The counterfactual mechanism M is an RS auction with a direct equivalent such that:*

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

2. *Expected Stage 2 profit of a low-type bidder (entrant with value $\underline{\mathrm{v}}$) is weakly decreasing in each element of $\mathbf{R}_{-i}$.*[22]

3. *WLOG, M is specified such that the expected payment of an entrant reporting value $v_i \leq \underline{\mathrm{v}}$ is*

$$\mathbf{1}[n=1]\left\{\alpha(v_i)v_i - \int_0^{v_i} \alpha(y)dy\right\} + \rho_M(\mathbf{R}_{-i}), \tag{23}$$

where $\rho_M(\mathbf{w}_{-i})$ is a symmetric nondecreasing function.

Conditions 1 and 2 of Assumption 5 are standard, and satisfied in the vast majority of mechanisms seen in practice. Condition 3 looks more restrictive, but in fact is without loss of generality since it applies only to *out-of-equilibrium* reports $v_i \leq \underline{v}$. This latter fact is important in ensuring generality of our results, so we state it formally as a lemma:

[22] Since a low-type bidder only wins when no other bidder enters, the main function of this condition is to rule out hypothetical mechanisms where the auctioneer provides subsidies for high realizations of rival types.

Lemma 2. *Any RS auction satisfying Conditions 5.1 and 5.2 is payoff- and performance-equivalent to some RS auction also satisfying Condition 5.3.*

While perhaps not obvious from the initial statement, our primary motive for adopting the normalization in Assumption 5.3 is that it intuitively nests the leading special cases of interest. We illustrate this fact via two simple examples:

Example 1. Consider a first-price auction with a reserve price $r \leq \underline{v}$. This structure can be nested under Assumption 5.3 by setting $\alpha(y) = \mathbf{1}[y \geq r]$ and $\rho(\cdot) = 0$.

Example 2. Consider a second-price auction with a secret reserve price $r \sim F_r(\cdot)$ and an entry fee $e > 0$. This structure can be nested under Assumption 5.3 by setting $\alpha(y) = F_r(y)$ and $\rho(\cdot) = e$.

For the moment, we further assume $\rho_M(\cdot) \geq 0$; roughly speaking, that the mechanism does not involve transfers to non-winning bidders. In the simple examples above, this is equivalent to requiring a non-negative entry fee:

Assumption 6. $\rho_M(\mathbf{R}_{-i}) \geq 0$ *for all rival reports* $\mathbf{R}_{-i}$.

As noted above, this assumption will be satisfied in most real-world auction mechanisms. It is not required to characterize expected revenue, but greatly simplifies analysis of monotonicity.

Having thus framed the problem, our argument proceeds in three steps. We first characterize expected revenue generated by mechanism M under *arbitrary* (potentially out-of-equilibrium) competition structure $(N, \bar{s})$:

Lemma 3. *Expected seller revenue under RS auction* M *at competition structure* $(N, \bar{s})$ *is given by*

$$R_M(N, \bar{s}) = \int_{v_0}^{\bar{v}} \{\alpha_M(y)[y - \lambda_M(y; N, \bar{s})] + [1 - \alpha_M(y)]v_0\}\, dF^*_{1:N}(y; \bar{s}) + N(1 - \bar{s})E[\rho_M | N, \bar{s}] \tag{24}$$

where

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

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

1. *$R_M(N,\bar{s})$ is identified for any $\bar{s} \in \mathcal{S}$.*

2. *$R_M(N,\bar{s})$ is decreasing in $\bar{s}$ for all N.*

We next consider counterfactual entry. When entry is endogenous, the equilibrium threshold s^*_M will depend on M, and when $F(v|s)$ and $c(\cdot)$ are not point identified, we cannot determine this dependence exactly. However, we can use the bounds on fundamentals derived above to *bound* the counterfactual entry behavior induced by M:

Lemma 4. *Let $c^+(\mathbf{z})$ and $c^-(\mathbf{z})$ be identified bounds on $c(\mathbf{z})$, $F^+(\cdot|s)$ and $F^-(\cdot|s)$ be identified bounds on $F(\cdot|s)$, and $s^*_M(N,\mathbf{z})$ be the (unknown) equilibrium entry threshold induced by counterfactual mechanism M at $(N,\mathbf{z})$. Define $s^+_M(N,\mathbf{z})$ and $s^-_M(N,\mathbf{z})$ as follows:*

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

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

where $\Pi_M(N,s|F)$ denotes expected profit of an entrant drawing from distribution F under mechanism M given competition (N,s):

$$\Pi_M(N,s|F) = \int_{v_0}^{\bar{v}} \alpha_M(y)[1-F(y)]F^*_{1:N-1}(y;s)dy - E[\rho_M|s].$$

Then $s_M^+(N,\mathbf{z})$ and $s_M^-(N,\mathbf{z})$ are identified and

$$s_M^+(N,\mathbf{z}) \geq s_M^*(N,\mathbf{z}) \geq s_M^-(N,\mathbf{z}),$$

with equality if $s_M^*(N,\mathbf{z}) \in int(\mathcal{S})$.

It only remains to combine these results. Lemma 3 establishes that conditional counterfactual revenue $R_M(N,\bar{s})$ is *identified* for all $\bar{s} \in \mathcal{S}$ and *decreasing in* $\bar{s}$ for all N, while Lemma 4 provides identified bounds $s_M^+(N,\mathbf{z}) \in \mathcal{S}$ and $s_M^-(N,\mathbf{z}) \in \mathcal{S}$ on the counterfactual entry threshold $s_M^*(N,\mathbf{z})$ induced by M. Substituting $s_M^+(N,\mathbf{z})$ and $s_M^-(N,\mathbf{z})$ into $R_M(N,\cdot)$ thus produces identified bounds on true counterfactual revenue $R_M^*(N,\mathbf{z})$ under M:

Proposition 7. *Choose any $(N,\mathbf{z})$, define $s_M^+(N,\mathbf{z})$ and $s_M^-(N,\mathbf{z})$ as in Lemma 4, and let $R_M^*(N,\mathbf{z})$ be (unknown) expected revenue under mechanism M at $(N,\mathbf{z})$. Define $R_M^+(N,\mathbf{z})$ and $R_M^-(N,\mathbf{z})$ as follows:*

$$R_M^-(N,\mathbf{z}) = \begin{cases} R_M\left(N, s_M^+(N,\mathbf{z})\right) & \textit{if } s_M^+(N,\mathbf{z}) \in \mathcal{S}(\mathcal{L}) \\ 0 & \textit{otherwise} \end{cases}$$

$$R_M^+(N,\mathbf{z}) = \begin{cases} R_M\left(N, s_M^-(N,\mathbf{z})\right) & \textit{if } s_M^-(N,\mathbf{z}) \in \mathcal{S}(\mathcal{L}) \\ \check{R}_M(N,0) & \textit{otherwise,} \end{cases}$$

where

$$\check{R}_M(N,0) = \int_{v_0}^{\bar{v}} \left\{ \alpha_M(y)[y - \int_{v_0}^{y} \frac{\alpha_M(t)}{\alpha_M(y)} \cdot \frac{F^*(t;\min\mathcal{S})^{N-1}}{F^*(y;\min\mathcal{S})^{N-1}}]dt + [1-\alpha_M(y)]v_0 \right\} dF_{1:N}^*(y;\min\mathcal{S}) + NE^*[\rho_M;\min\mathcal{S}]$$

is a semi-informative upper bound applicable when $s_M^-(N,\mathbf{z}) \equiv 0$.[23]

Then $R_M^+(N,\mathbf{z})$ and $R_M^-(N,\mathbf{z})$ are identified and $R_M^+(N,\mathbf{z}) \geq R_M(N,\mathbf{z}) \geq R_M^-(N,\mathbf{z})$, with equality if $s_M^*(N,\mathbf{z}) \in int(\mathcal{S})$.

[23]In particular, $\check{R}_M(0;N)$ 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 $\check{R}_M(0;N) \geq R_M(0;N) \geq R_M(s_M;N)$. Further, since $\min\mathcal{S} \in \mathcal{S}$, $\check{R}(0;N)$ is identified.

We thus obtain identified bounds on expected revenue applicable to a wide range of counterfactual mechanisms under endogenous and arbitrarily selective entry. Given appropriate choice of M, these bounds specialize immediately to policy instruments such as reserve prices and entry fees.[24] Appendix 2 illustrates our counterfactual bounds in a range of numerical examples, which together suggest their information content is encouragingly good. We thus establish that the AS model can support a rich variety of counterfactual and policy analyses under minimal assumptions on the nature of entry.

4 Extensions

4.1 Unobserved heterogeneity

Our discussion thus far has maintained Assumption 4, under which auction-level fundamentals $\{F_\ell(\cdot|\cdot), c_\ell\}$ are presumed to depend (at most) on observables $\mathbf{Z}$ and $\mathbf{X}$. In practice, however, bidder behavior will often depend on auction-level factors not observed by the econometrician, and in such cases Assumption 4 may fail to hold. Building on recent developments in the econometrics literature, this section extends our core identification results to a set of environments with unobserved heterogeneity, where this heterogeneity is not constrained to enter via any particular functional form. We thus establish the relevance of our results in a much broader class of models than have been heretofore considered.

In particular, consider introduction of an auction-level random variable U, which is observable to all potential bidders but not to the econometrician. This random variable is assumed to auction-level fundamentals $\{F_\ell(\cdot|\cdot), c_\ell\}$ through its realizations u_ℓ, leading to the following relaxation of Assumption 4:

Assumption 7. *For each auction ℓ generated by process $\mathcal{L}$, $F_\ell(v,s) = F(v,s|u_\ell)$ and $c_\ell = c(\mathbf{z}_\ell|u_\ell)$, both potentially conditional on further covariates $\mathbf{X}$.*

[24] In the special case of a public reserve price, our revenue bounds can also be translated into bounds on the seller's optimal reserve price following Haile and Tamer (2003), though in practice such bounds tend to be wide.

Let $F_u(\cdot)$ denote the distribution of the unobserved random variable U, and $\mathcal{U} \subset \mathbf{R}$ denote its support. We impose the following additional structure on this new economic environment:

Assumption 8. *The random variables* $\{(V_1, S_1), ..., (V_N, S_N); U\}$ *satisfy the following properties:*

1. Conditional independence: *for all potential bidders* i *and* j, $(V_i, S_i) \perp (V_j, S_j)|U$.
2. Stochastic ordering: *for all* (v, s), $u' \geq u$ *implies* $F(v|s, u') \leq F(v|s, u)$.
3. Entry ordering: *for all* N, $u' \geq u$ *implies* $s^*_N(u') \leq s^*_N(u)$.

Condition 8.1 extends the IPV model to accommodate unobserved heterogeneity, Condition 8.2 ensures that higher realizations of U are "good news," and Condition 8.3 specifies that entry is more likely at higher realizations u. Interestingly, while the last condition corresponds closely to intuition, it is not a consequence of the first two; there exist fundamentals satisfying stochastic ordering for which $s^*_N(u)$ can be increasing. But decreasing $s^*_N(u)$ is clearly the economic leading case, and in particular can be shown to hold if V is either additively or multiplicatively separable in U. Thus Assumption 8 represents a natural generalization of Assumption 1 to environments with unobserved heterogeneity.

As above, we assume the number of potential bidders N_ℓ, the number of entrants n_ℓ, and the vector of submitted bids $\mathbf{b}_\ell$ are observed for each auction. Let $s^*_N(\mathbf{z}; u)$ and $G^*(b|N, \mathbf{z}; u)$ be the equilibrium entry threshold and equilibrium distribution of bids corresponding to realization $u \in \mathcal{U}$ at $(N, \mathbf{z})$. If U were observed, these objects could be estimated directly for each u, and identification would follow as in Section 2. In practice, however, entry and bid decisions will directly identify only *averages* of $s^*_N(\mathbf{z}; U)$ and $G^*(b|N, \mathbf{z}; u)$ across U:

$$\begin{aligned}
\hat{s}_N &= E_U\left[s^*_N(\mathbf{z}; U)\right] \\
\hat{G}(b|N) &= E_U\left[G^*\left(b|N, \mathbf{z}; U\right)\right].
\end{aligned}$$

The identification problem induced by unobserved heterogeneity can thus be stated as follows: recover $s^*_N(\mathbf{z}; U)$, $G^*(b|N, \mathbf{z}; U)$ and $F_u(U)$ from quantities identified without seeing u. Our solution to this problem builds on the work of Hu, McAdams, and Shum (2011), who propose several novel results on identification in auctions with nonseparable unobserved heterogeneity.[25] As in Section 2.4, we assume for clarity that all entry variation is induced by N, but all results extend to excludable instruments $\mathbf{Z}$.

4.1.1 Identification under unobserved heterogeneity

For the moment, we focus on the case where U is discrete; that is, where $\mathcal{U}$ is a finite set (we discuss the continuous case briefly below). Let K denote the cardinality of $\mathcal{U}$, and normalize $\mathcal{U} = \{1, ..., K\}$ without loss of generality. Following Hu, McAdams, and Shum (2011), the argument will invoke *discretization* of the bid space, formally defined as follows:

Definition 4. A *discretization* is any one-to-one, onto mapping $\mathcal{D} : [0, \bar{v}] \to \mathcal{U}$; that is, any partition of the bid space into K intervals. Let $d_i \equiv \mathcal{D}(b_i)$ denote the interval corresponding to bid b_i, and $D_i \equiv \mathcal{D}(B_i)$ denote the corresponding random variable.

Define a *realized bid* W_i for each potential bidder i as follows: $W_i = B_i$ if i enters and $W_i = 0$ otherwise. By construction, the resulting random variables $(W_1, ..., W_N)$ follow a mixed joint distribution, with mass points on the zero axes representing probabilities of non-entry. Let $\hat{G}_w(W_1, ..., W_N|N)$ denote the joint distribution of $(W_1, ..., W_N)$ at N, and $\hat{g}_w(W_1, ..., W_N|N)$ denote the corresponding mixed joint density (with the convention that points on any zero axis represent probability masses).

As in Hu, McAdams, and Shum (2011), recovery of the equilibrium conditional distribution $G^*_w(W_i|N; u)$ from the identified joint distribution $\hat{G}_w(W_1, ..., W_N|N)$ requires an auxiliary distributional assumption. For any two bidders j and k and any discretization $\mathcal{D}$,

[25] See Hu (2008) for further development of the underlying identification argument; Krasnokutskaya (2011) considers identification in the special case of auctions with additively (or multiplicatively) separable unobserved heterogeneity.

let $\hat{M}_{d_j,d_k}(N)$ be the matrix characterizing the (identified) joint distribution of the discretized variables $D_j \equiv \mathcal{D}(W_j)$ and $D_k \equiv \mathcal{D}(W_k)$ at N:

$$\hat{M}_{d_j,d_k}(N) = \left[\ \Pr\{D_j = j', D_k = k'|N\} \ \right]_{\{j',k'\}\in\mathcal{U}\times\mathcal{U}}$$

We then impose the following full-rank condition on $\hat{M}_{d_j,d_k}(N)$:

Assumption 9. *For each $N \in \mathcal{N}$ such that $N \geq 3$, there exists a discretization $\mathcal{D}$ such that $\hat{M}_{d_j,d_k}(N)$ has rank K.*

As noted in Hu, McAdams, and Shum (2011), this condition follows automatically when the maximum realized bid is increasing in u. Assumption 9 is therefore without loss of generality in a first-price auction under Assumption 8 above, or in a second-price auction when the upper limit of $\mathcal{V}$ is increasing in u. Further, since $\hat{M}_{d_j,d_k}(N)$ is identified for any $\mathcal{D}$, the underlying rank condition can be readily verified in applications. Hence Assumption 9 is unlikely to be a major practical constraint.

We can now apply Lemma 1 in Hu, McAdams, and Shum (2011) to establish identification of the equilibrium distribution $G^*_w(W_i|N;U)$ and the heterogeneity distribution $F_u(U|N)$ prevailing at competition level N:

Lemma (Hu, McAdams, and Shum (2011))**.** *Suppose U is finite, Assumptions 7-9 are satisfied, and $N \geq 3$. Then $G^*_w(W_i|N;U)$ and $F_u(U|N)$ are identified from $\hat{G}_w(W_1, ..., W_N|N)$.*

Finally, note that identification of $G^*_w(W_i|N,u)$ at any u implies identification of $s^*_N(u)$ and $G^*(b|N;u)$ for that u:

$$G^*_w(b|N;u) \equiv s^*_N(u) + [1 - s^*_N(u)]\, G^*(b|N;u),$$

so

$$s^*_N(u) \equiv G^*_w(0|N;u)$$

and

$$G^*(b|N;u) = \frac{G^*_w(b|N;u) - G^*_w(0|N;u)}{1 - G^*_w(0|N;u)}.$$

In turn, given identification of $s^*_N(u)$ and $G^*(b|N;u)$ at each u, we can apply results from Section 2 to obtain identified bounds on $F(v|s;u)$ and $c(u)$. Thus all results above can be extended to accommodate nonclassical, nonseperable unobserved heterogeneity.

4.1.2 Further discussion

We conclude this subsection with three further comments. First, the discussion above deliberately permitted U to enter both $F(v|s;\cdot)$ and $c(\cdot)$. The entry ordering assumption 8.3 imposed an equilibrium restriction on entry induced by U, so this relationship is not entirely arbitrary. In practice, however, it is still more general than typically seen; most applications allow U to enter $F(v|s;\cdot)$ or $c(\cdot)$ but not both. The results above of course specialize to either special case.

Second, while only noted in passing above, the results here yield identification of the heterogeneity distribution $F_u(\cdot|N)$ *at each* N. We can therefore relax the excludability restrictions in Assumption 3 to excludability of N *conditional on* U: e.g.

$$F(v,s|N,U) = F(v,s|U).$$

So long as $\Pr(N|U)$ has nondegenerate support at each U, all results above immediately extend. This represents a substantial generalization of prior approaches in the literature, and may be of considerable value in applications where endogeneity of N is a concern.

Finally, though the discussion above took U to be discrete, the underlying logic extends readily to the continuous case. As in Hu, McAdams, and Shum (2011), the cost of this extension is a stronger nonparametric completeness condition, which replaces Assumption 9 above:

Assumption (Nonparametric completeness, continuous U)**.** *For all bidders i, j, functions*

$h(\cdot)$ having bounded conditional expectation, and observed $N \geq 3$,

1. *$E[h(W_k)|w_i; N] = 0$ for all $w_i \in \{0 \times \mathcal{V}\}$ implies $h(w_k) = 0$ for all $w_k \in \{0 \times \mathcal{V}\}$.*

2. *$E[h(U)|w_i; N] = 0$ for all $w_i \in \{0 \times \mathcal{V}\}$ implies $h(u) = 0$ for all $u \in \mathcal{U}$.*

These restrictions correspond closely to those required in other nonparametric identification settings (e.g. Newey and Powell (2003) and Berry and Haile (2010a,b)), and essentially guarantee that the distributions $F_u(\cdot|N)$ and $G_w^*(W_i|N; U)$ decomposing $\hat{G}_w(W_1, ..., W_N|N)$ into a product of marginals are unique. See Hu and Shiu (2012) for further discussion and sufficient conditions. Granted this additional condition, however, the argument is otherwise very similar. We can thus generalize our core results to a broad class of environments with nonclassical and nonseperable unobserved heterogeneity, obtaining partial identification of model fundamentals and policy outcomes accounting for endogenous and arbitrary selective entry.

4.2 Bidder asymmetry

A second potential extension of our results is to environments with asymmetric bidders. In particular, suppose process $\mathcal{L}$ involves a set $\mathcal{T}$ of potential bidder types, which are observable both to potential bidders and to the econometrician. Let $F_{\tau_i}(v, s)$ and $c_{\tau_i}(\cdot)$ denote the signal-value distribution and entry cost function of a bidder with type $\tau_i \in \mathcal{T}$, and $\tau_\ell \in \mathcal{T}^{N_\ell}$ be the vector of type realizations in auction ℓ. The type space $\mathcal{T}$ can be either discrete or continuous; we require only that continuous types affect model fundamentals continuously.

Further suppose a single type-symmetric equilibrium is played at each type vector τ, and let $s^*(\tau)$ be the corresponding vector of type-specific equilibrium entry thresholds. A repeated sample from auctions with type vector τ will then permit identification of the type-specific threshold $s^*_{\tau_i}(\tau)$ and *ex post* bid distribution $G^*_{\tau_i}(\cdot; s^*_{\tau_i}(\tau))$ for each type τ_i at τ, and pooling across other type vectors $\tilde{\tau}$ containing τ_i will generate a set of type-specific identified thresholds $\mathcal{S}_{\tau_i}$. Identification of type-specific fundamentals can then proceed as above.

The key additional complication under bidder asymmetry is that there may exist multiple type-symmetric equilibria at any type vector τ. If more than one such equilibrium is played with positive probability, observed entry and bidding decisions will represent an unknown mixture distribution, and identification may break down. This problem is the focus of much ongoing research, and a full treatment is beyond the scope of this paper. As typical in applications, therefore, we simply assume that the equilibrium played is either directly observed or depends deterministically on τ. Granted this restriction, however, asymmetry may in fact substantially improve identification; in particular, a continuous typeset $\mathcal{T}$ will typically induce a continuous identified set $\mathcal{S}_{\tau_i}$, thus permitting point identification as in Section 2.2.

5 Conclusion

In this paper, we explore a general approach to identification in auctions with entry based on a framework we call the AS model. In the process, we make three core contributions to the related literature. First, we derive a general map from observed entry variation to identified bounds on model fundamentals, where complete variation yields point identification and incomplete variation yields partial identification. This map is directly implementable, applies to a broad class of auction mechanisms under arbitrarily selective entry, and can readily be refined into the sharp identified set. Second, we translate identified bounds on fundamentals into bounds on expected revenue under a wide range of counterfactual mechanisms, again accounting for endogenous and selective entry. We therefore demonstrate that the AS model can yield practically relevant policy insights even in applications with incomplete entry variation. Finally, we discuss extensions to environments with asymmetric bidders and nonseperable unobserved heterogeneity, thus establishing the relevance of our findings within an even broader class of economic environments. We thus both characterize what can be learned without parametric restrictions on selection and illustrate this information's capacity to yield insights on auctions with entry.

While our main focus thus far has been nonparametric identification, our results also provide a straightforward basis for nonparametric estimation because they are constructive. In particular, all objects established as directly identified in Section 2.1 have standard nonparametric estimators, and plugging these in to subsequent results in turn produces consistent estimators for our sharp nonparametric bounds. Especially in the partially-identified case, however, deriving asymptotic properties for these nonparametric estimators turns out to be quite challenging, both due to the multiple steps involved and because properties of the initial steps will be mechanism-dependent. For current purposes, therefore, we therefore choose to focus on nonparametric identification rather than nonparametric inference.[26] In applications, particularly those involving high-dimensional covariates or unobserved heterogeneity, it may instead be desirable to adopt a parametric or semiparametric approach. For the former, see (e.g.) Roberts and Sweeting (2010a,b), who consider estimation of a parametric log-normal AS model incorporating rich normally-distributed unobserved heterogeneity. For the latter, one natural approach would be to combine nonparametric estimation of *ex post* identified quantities with a parametric copula $C_\theta(\cdot)$ for the joint distribution $F(v, s)$. Variation in the entry threshold $\hat{s}$ will then induce a set of equality restrictions which must be jointly satisfied by $(\theta_0, F_0(\cdot))$ at each $v \in \mathcal{V}$, and under standard regularity conditions point identification follows so long as the dimensionality of θ is less than the cardinality of $\mathcal{S}$.[27] In both cases, our results provide insight on the role of parametric restrictions as data

[26] In this respect, we follow many other notable recent studies: e.g. Athey and Haile (2002) on auctions, Chesher (2005) on models with discrete covariates, Berry and Haile (2010b,a) on multinomial choice models, Chesher and Smolinski (2010) on sharp identified sets for discrete variable IV models, and Chesher and Rosen (2011) on sharp identified sets for binary response models. These studies each provide valuable insight on information nonparametrically available within their respective applications, but do not explore details of nonparametric inference.

[27] In particular, adding and subtracting $\int_0^s F(v|t)dt$ to the RHS of Equation 1 and simplifying produces the following identity:

$$F^*(v; s) = \frac{F(v, 1) - F(v, s)}{1 - s}.$$

Rearranged and combined with the copula restriction $F(v, s) \equiv C_\theta(F(v), s)$, this in turn implies

$$(1 - \hat{s})F^*(v; \hat{s}) \equiv F(v, 1) + C_\theta(F(v, 1), \hat{s}) \text{ for each } \hat{s} \in \mathcal{S}.$$

smoothers *vis-a-vis* their role in driving analysis. Given our finding that fully nonparametric analysis has substantial capacity to yield policy-relevant conclusions on auctions with entry, we interpret our results quite positively in this respect.

Finally, while our analysis thus far has been framed in the context of auctions, many of our key insights could apply to environments with selection more generally. In particular, our core partial identification results in Section 2 may be relevant in any structural model where agents select into a binary (or ordered) decision based on an imperfect signal. Potential examples include selection into geographic markets in industrial organization, or selection into treatment in labor applications. The key restriction is that our results require Stage 2 "values" to be identified, with the corresponding advantage of producing pointwise sharp bounds on both joint and marginal distribution of fundamentals. We are exploring these connections in ongoing research.

Treating $F(v,1)$ as a parameter for each $v \in \mathcal{V}$ and solving for $(\theta, F(v,1))$ will yield a set of candidate models at each $v \in \mathcal{V}$. The true model $(\theta_0, F_0(\cdot))$ will be a selection from this set satisfying the further restriction that $F_0(v,1)$ is a distribution across v. So long as this solution is unique, the model will be semiparametrically identified.

Appendix A: Numerical examples

Sections 2 and 3 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 exercise illustrates the kind of information that could in principle be recovered using the methods developed above.

Our parametric specification is as follows. We model the joint distribution $F(v, s)$ using a Gaussian copula $C_\rho(F_v, s)$, where the marginal distribution $F_v(\cdot) \sim N(\mu = 100, \sigma = 10)$ and the entry cost is $c = 2$. These numeric values are chosen to be qualitatively similar to estimates in the literature.[28] The correlation parameter ρ measures the degree of dependence between s and v, with $\rho = 0$ generating the no-selection case and $\rho \to 1$ approaching the perfect-selection case. In what follows, we present results for $\rho = 0.2$, $\rho = 0.75$, and $\rho = 0.95$, representing minimally, moderately, and highly selective DGPs respectively. Except where noted otherwise, we assume potential competition N varies exogenously on the set $\mathcal{N} = \{2, 3, ..., 16\}$.

A.1: Bounds on fundamentals

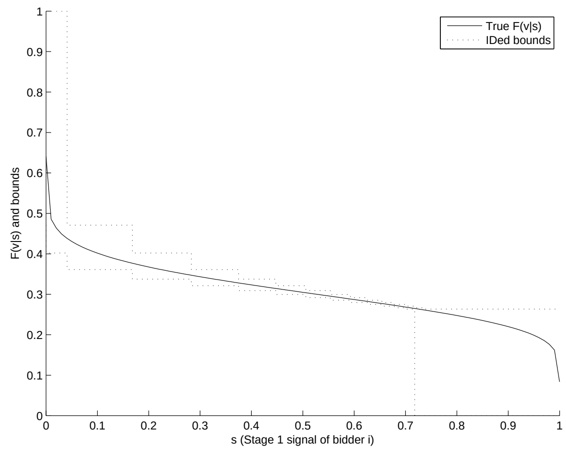
We first illustrate the bounds on model fundamentals derived in Section 2.3. Given the parametric specifications above, it is straightforward to calculate the set of equilibrium entry thresholds $\mathcal{S} = \{s_2^*, ..., s_{16}^*\}$ satisfying the the breakeven condition (4) at each value of ρ considered. These and the corresponding *ex post* distributions $F^*(v; s_N^*)$ are the objects identified by a standard $(N, n, \mathbf{b})$ sample, and the raw inputs into our fundamental bounds.

Given these identified threshold sets $\mathcal{S}$, we next use Proposition 3 to obtain identified bounds $F^+(\cdot|s)$ and $F^-(\cdot|s)$ on $F(\cdot|s)$. These bounds imply an identified set $\mathcal{F}$ in three-dimensional space, for which we provide two sets of graphical representations below. First, Figure 1 compares the S dimension of our bounds for two values of v across ρ. The "stairstep" nature of these bounds in s follows from lack of information on $F^*(v; s)$ at points outside the identified set $\mathcal{S}$. Meanwhile, Figure 2 illustrates the V dimension of our bounds at a selection of thresholds s_N^* in $\mathcal{S}$. Note the lack of informative upper (lower) bounds at the minimum (maximum) of $\mathcal{S}$, a point to which we return when discussing sharpness below.

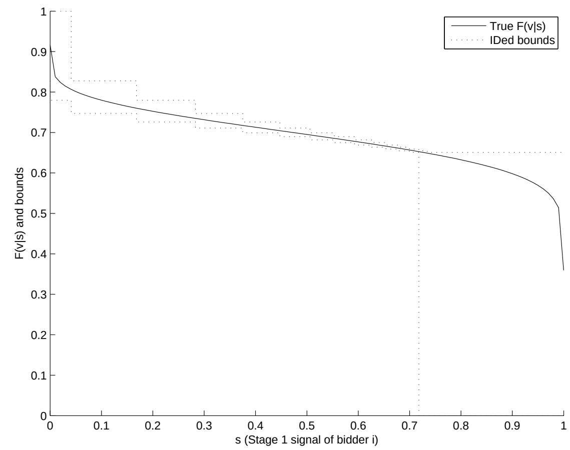
We note three further points on Figures 1 and 2. First, our bounds on $F(\cdot|s)$ become tighter as S becomes less informative. This is not surprising, since we know they yield point identification in the no-selection LS case (since then $F^*(v; s) \equiv F(v|s) \equiv F_v(v)$ by definition). Second, higher competition levels induce smaller steps in s_N^*, which in turn tends to yield tighter bounds on $F(\cdot|s)$. This pattern obviously depends somewhat on particular the DGP selected, but insofar as the marginal change in expected profit is decreasing in N is likely to be reasonably robust. Finally, as in Proposition 4, we can translate bounds on $F(v|s)$ into bounds on c. In the current examples these bounds are quite tight: $c \in [1.990, 2.010]$ when $\rho = 0.2$, $c \in [1.977, 2.024]$ when $\rho = 0.75$, and $c \in [1.985, 2.016]$ when $\rho = 0.95$, where as above true $c = 2$.

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

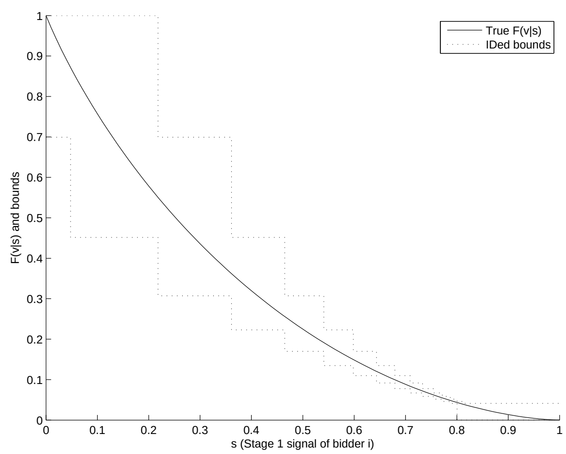
Figure 1: Bounds on $F(v|s)$ across S, $\mathcal{N} = \{2, ..., 16\}$



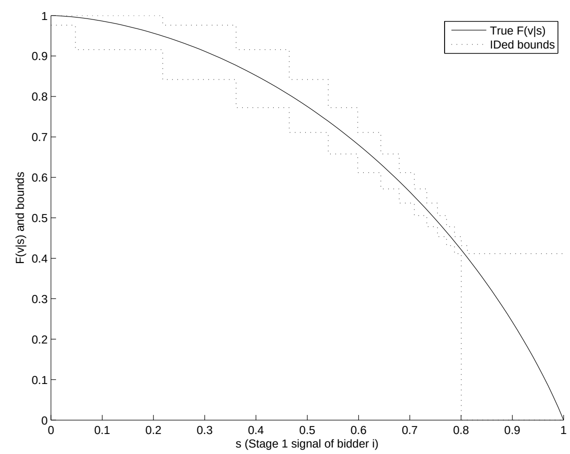
(a) Bounds at $v = 95$, $\rho = 0.2$



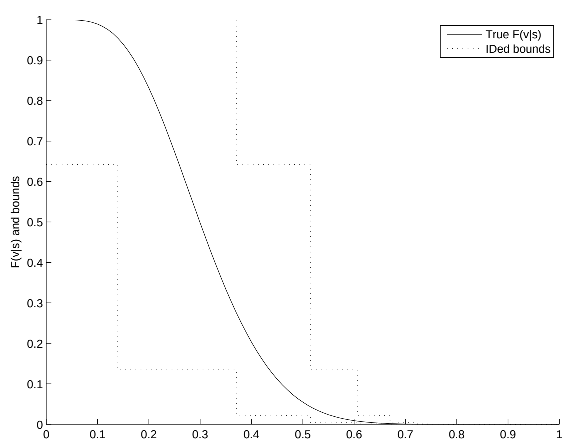
(b) Bounds at $v = 105$, $\rho = 0.2$



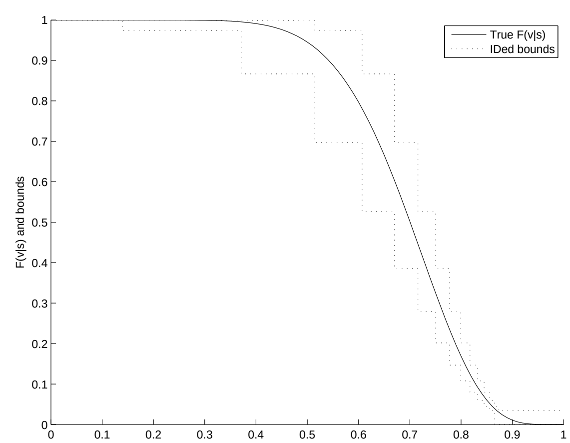
(c) Bounds at $v = 95$, $\rho = 0.75$



(d) Bounds at $v = 105$, $\rho = 0.75$

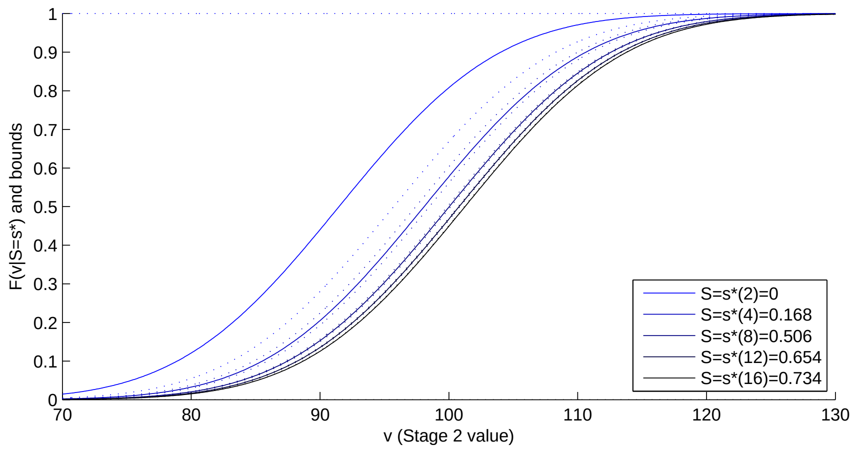


(e) Bounds at $v = 95$, $\rho = 0.95$

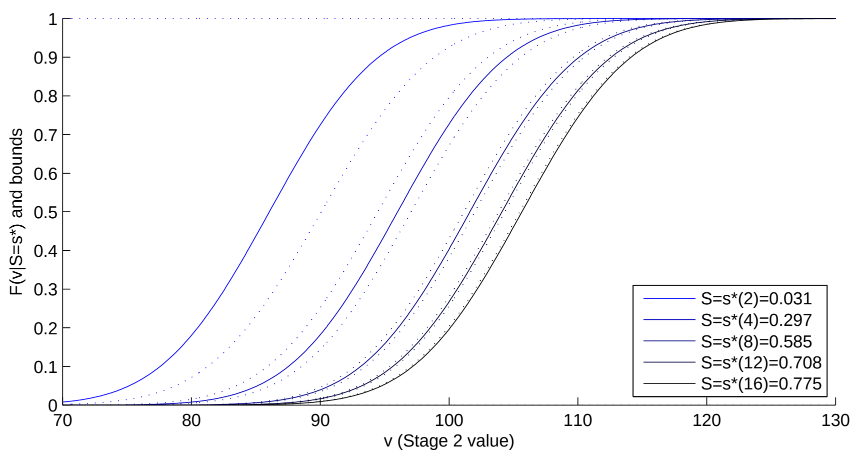


(f) Bounds at $v = 105$, $\rho = 0.95$

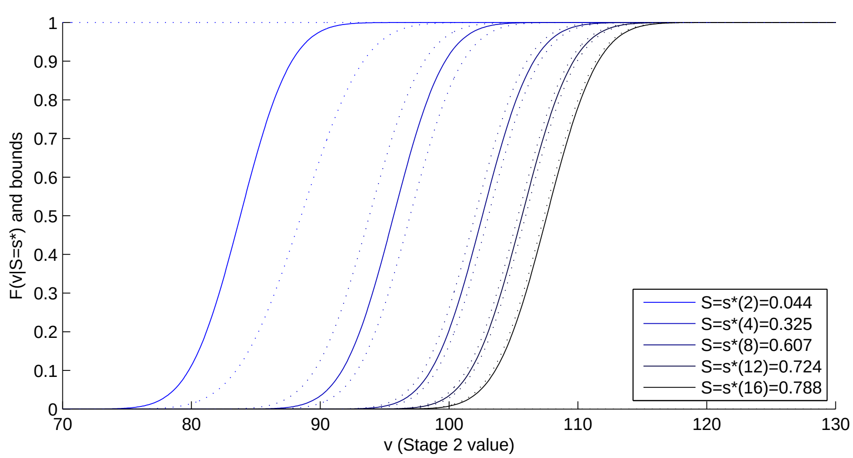
Figure 2: Bounds on $F(v|s)$ across $\mathcal{V}$, $\mathcal{N} = \{2, ..., 16\}$



(a) Bounds on $F(v|s_N^*)$ for selected N, $\rho = 0.2$



(b) Bounds on $F(v|s_N^*)$ for selected N, $\rho = 0.75$



(c) Bounds on $F(v|s_N^*)$ for selected N, $\rho = 0.95$

A.2 Sharp identified set

We next discuss refinement of bounds obtained as in the last subsection into the pointwise sharp identified set. Toward this end, we first apply the test of sharpness in Proposition 5 to the initial bounds above. At points where this test fails, we then apply the refinement in Proposition 6 to obtain the pointwise sharp identified bounds.

From the perspective of Proposition 3, the results of this procedure are quite encouraging: for all three baseline DGPs, our initial bounds are pointwise sharp everywhere they are informative. Refinement is possible at some points with uninformative initial bounds, and in all these cases the procedure in Proposition 6 yields the pointwise sharp identified set in a single iteration from its starting point. Not surprisingly, the largest improvements are achieved where the initial bounds are widest; Figure 4a plots initial and final bounds for $\rho = 0.95$ as an illustrative example. Thus while the results in Section 2.4 indicate when and how the bounds in Proposition 3 can be refined into the sharp identified set, the numerical results presented here strongly suggest that this extra step will typically be unnecessary in applications.

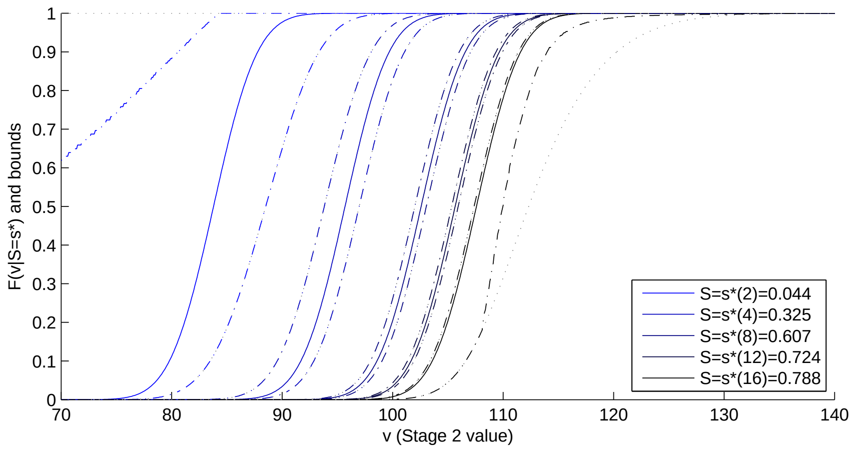
We conclude this section with a counterexample: a sample DGP in which initially informative bounds can be further refined. Not surprisingly given the discussion above, this example is also of necessity artificial: bounds on c contain relatively little information on $F(v|s)$, so to induce refinement we need both tight bounds on c and a large gap in identified entry thresholds. The discussion above suggests such a pattern is most likely to obtain when the DGP involves a large gap in N at low competition levels, so for clarity we consider the extreme case $\mathcal{N} = \{2, 15, 16\}$. Figure 4b plots the results, which as expected show substantial gains in the sharp identified set. In practice, however, it is difficult to envision applications where such large gaps in competition arise naturally, and even adding a single intermediary point ($N = 9$, Figure 4c) is sufficient to dissipate most gains. Hence in applications Proposition 3 alone likely to be sufficient.

A.3 Bounds on counterfactuals

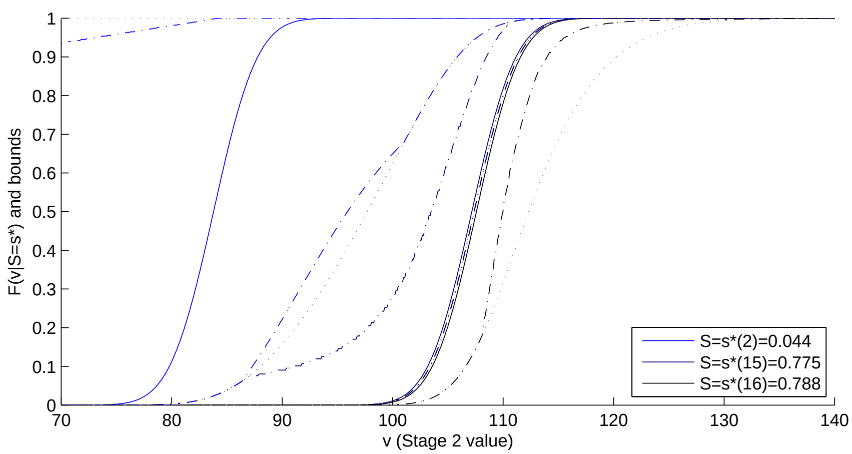
Finally, following Section 3, we translate the bounds on $F(v|s)$ and c obtained above into identified bounds on counterfactual seller revenue. While results in Section 3 are defined more broadly, for simplicity we here focus on the special case of a public reserve price r; we also assume the no-sale outcome yields value $v_0 = 40$ to the seller. As in Section 3, we first use Lemma 4 to bound the counterfactual entry equilibrium $s^*_r(N)$ corresponding to each candidate reserve price r, then use Proposition 7 to map these candidate thresholds into bounds on counterfactual revenue $R^*_r(N)$. These bounds will obviously depend on both the underlying DGP and the particular N in question; we illustrate results at $N = 4$ and $N = 9$ for each value of ρ considered. Bounds on counterfactual entry are plotted in Figure 4, bounds on counterfactual revenue are plotted in Figure 5.

On balance, the results in Figures 4 and 5 suggest that our counterfactual bounds can be quite informative: while richer identified sets will obviously improve precision, even variation in N alone is sufficient to generate economically meaningful restrictions on the objects of interest. In particular, using only in-sample variation, the revenue bounds in Figure 5 yield a remarkably accurate impression of the potential revenue implications of a binding reserve price, in a fashion which accounts fully for endogenous and selective bidder response. Thus the results in Sections 2 and 3 can support robust policy analysis even in the presence of partial identification induced by endogenous and arbitrarily selective entry.

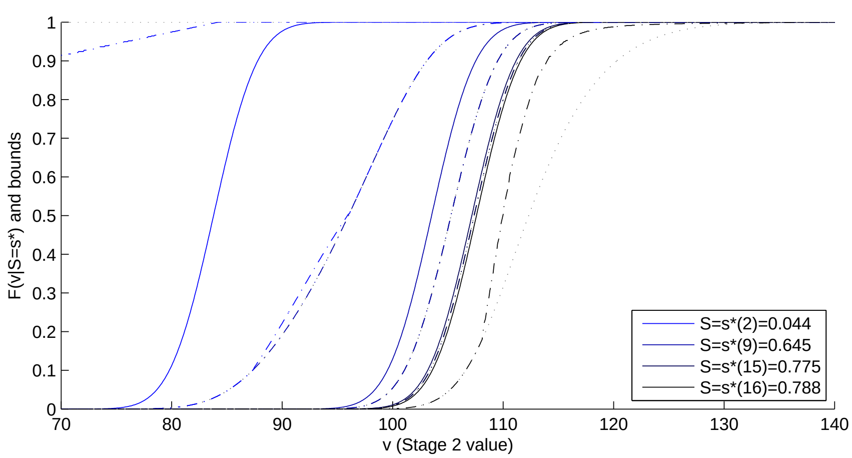
Figure 3: Pointwise sharp bounds on $F(v|s)$, $\rho = 0.95$, various $\mathcal{N}$



(a) Sharp bounds on $F(v|s_N^*)$ at selected N, $\mathcal{N} = \{2, 3, ..., 16\}$

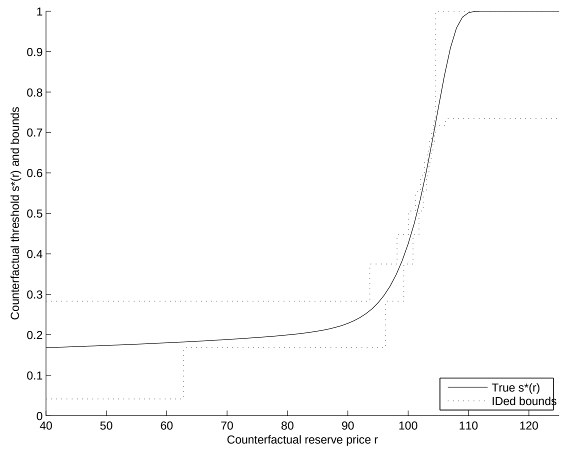


(b) Sharp bounds on $F(v|s_N^*)$ at selected N, $\mathcal{N} = \{2, 15, 16\}$

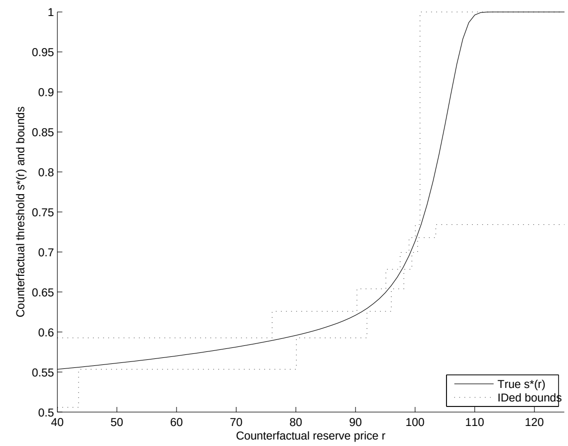


(c) Sharp bounds on $F(v|s_N^*)$ at selected N, $\mathcal{N} = \{2, 9, 15, 16\}$

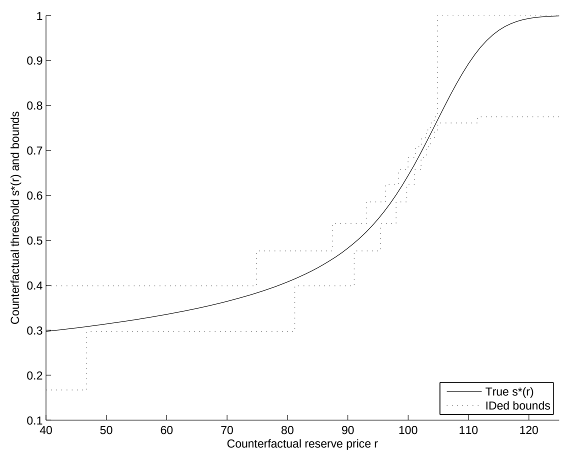
Figure 4: Bounds on counterfactual entry threshold $s_r^*(N)$, $\mathcal{N} = \{2, ..., 16\}$



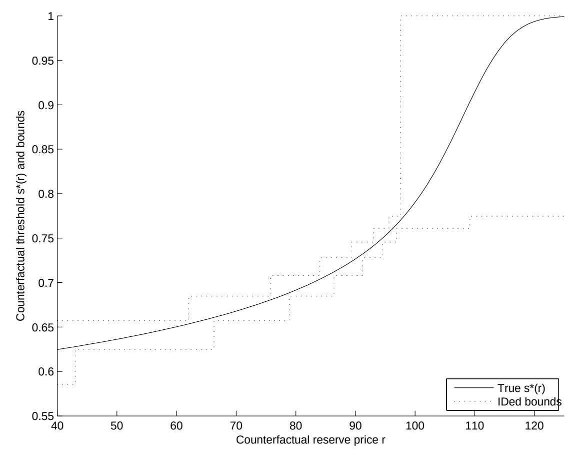
(a) Bounds on $s_r^*(N)$ at $N = 4$, $\rho = 0.2$



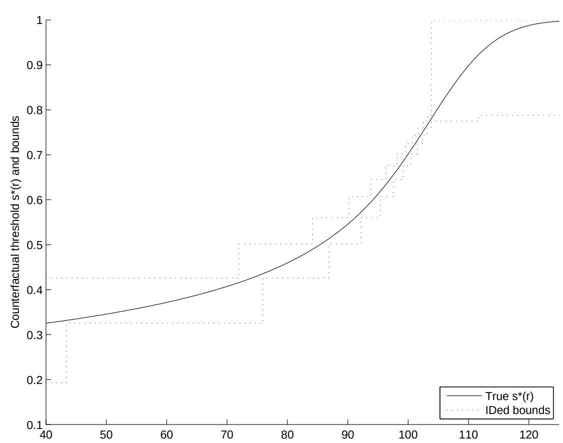
(b) Bounds on $s_r^*(N)$ at $N = 9$, $\rho = 0.2$



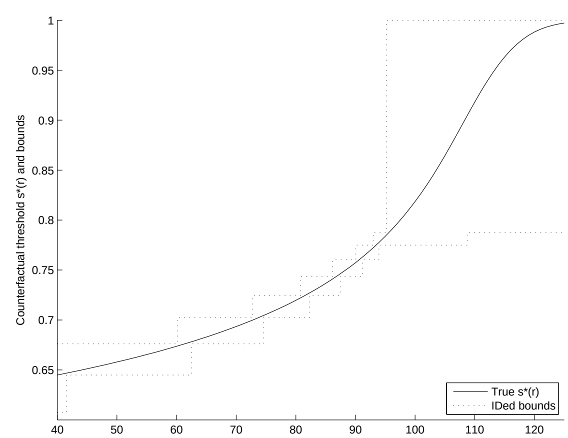
(c) Bounds on $s_r^*(N)$ at $N = 4$, $\rho = 0.75$



(d) Bounds on $s_r^*(N)$ at $N = 9$, $\rho = 0.75$

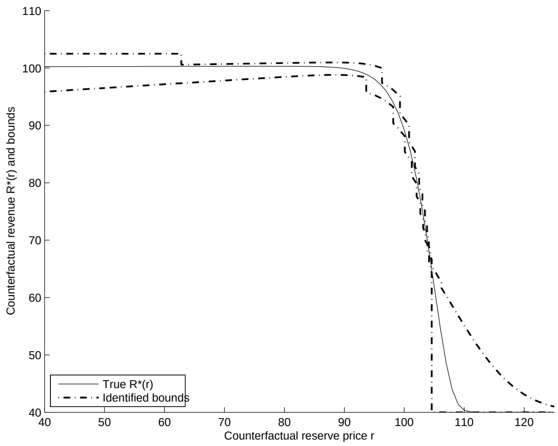


(e) Bounds on $s_r^*(N)$ at $N = 4$, $\rho = 0.95$

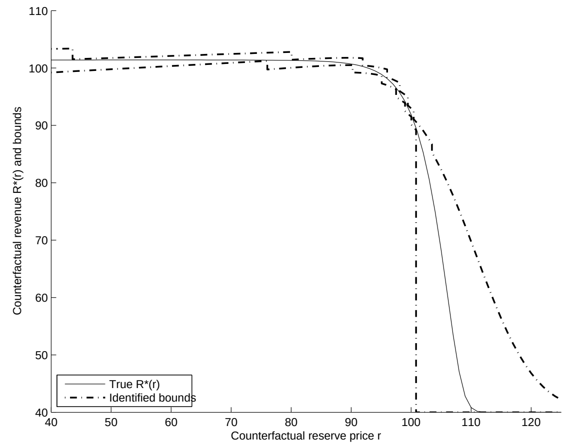


(f) Bounds on $s_r^*(N)$ at $N = 9$, $\rho = 0.75$

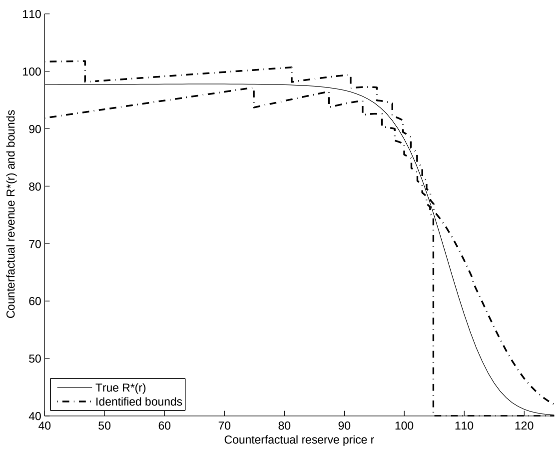
Figure 5: Bounds on counterfactual revenue $R_r^*(N)$, $\mathcal{N} = \{2, ..., 16\}$



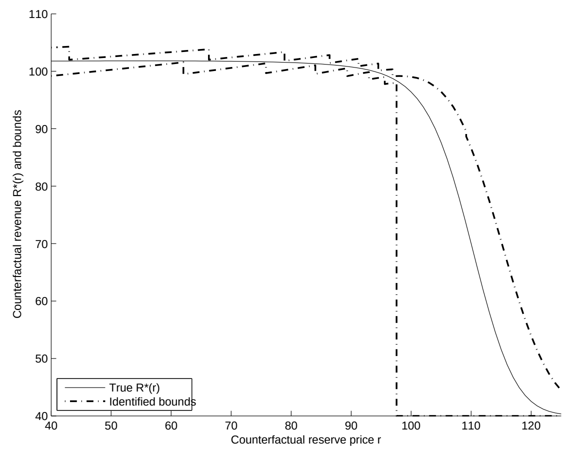
(a) Bounds on $R_r^*(N)$ at $N = 4$, $\rho = 0.2$



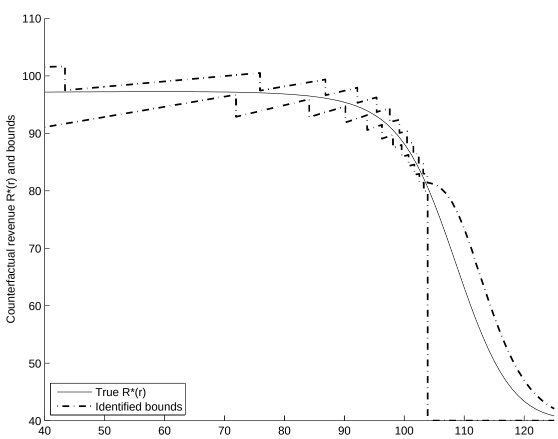
(b) Bounds on $R_r^*(N)$ at $N = 9$, $\rho = 0.2$



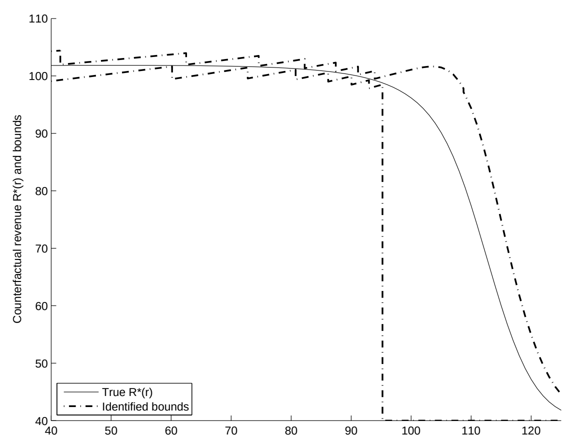
(c) Bounds on $R_r^*(N)$ at $N = 4$, $\rho = 0.75$



(d) Bounds on $R_r^*(N)$ at $N = 9$, $\rho = 0.75$



(e) Bounds on $R_r^*(N)$ at $N = 4$, $\rho = 0.95$



(f) Bounds on $R_r^*(N)$ at $N = 9$, $\rho = 0.75$

Appendix B: Other mechanisms

For expositional convenience, our discussion thus far has focused on the case where the Stage 2 auction is run according to simple second-price sealed bid rules. As noted throughout the text, however, our underlying arguments extend to most other standard auctions. In this appendix, we briefly sketch details of this more general argument.

As in Section 3, we consider extension of our underlying results to the general class of mechanisms we call RS auctions, formally described in Definition 3. As above, we define the *award rule* $\alpha_M(y)$ corresponding to mechanism M as the equilibrium probability that M results in sale when the maximum realized value among entrants is $Y = y$; this award rule can be deterministic as with a public reserve price ($\alpha(y) = \mathbf{1}[y \geq r]$), or stochastic as in the case of a secret reserve price ($\alpha(y) = F_r(\beta(y))$). We maintain the regularity conditions in Assumption 5: namely, that a low-type bidder weakly prefers less potential competition, and that the award rule $\alpha(y)$ is weakly increasing in y.[29] Further, as discussed in the introduction, our main focus in this paper is to obtain identified bounds on AS fundamentals taking as given quantities whose identification is well-established in the literature. For current purposes, therefore, we impose the following preliminary identification assumption:

Assumption (Stage 2 identification)**.** *At each $(N, \mathbf{z}) \in \mathcal{L}$, the mechanism M is such that under the hypothesis of equilibrium play, the distribution of values corresponding to bids submitted is nonparametrically identified.*

Mechanisms under which this assumption is known to hold include first-price, second-price, ascending, and Dutch auctions; we outline arguments applicable to each of these mechanisms below. See Athey and Haile (2005) for a comprehensive survey of the literature.

As above, the first step in our extended identification argument is to characterize equilibrium entry behavior within the class of mechanisms considered. Toward this end, we extend Proposition 1 to accommodate a general RS auction M:

Proposition (Equilibrium in RS auctions)**.** *Let M be an RS auction satisfying Assumption 5, and $\alpha_M(\cdot)$ and $\rho_M(\cdot)$ be the award rule and low-type payment function induced by M. Then when competing against $N-1$ potential rivals who enter according to $\bar{s}$, an entrant with value v_i in Stage 2 expects profit*

$$\pi_M(v_i; \bar{s}, N) = \int_0^v \alpha_M(y) \cdot F^*_{1:N-1}(y; \bar{s}) dy - E[\rho_M | N, \bar{s}]$$

and a potential entrant with signal s_i in Stage 1 expects Stage 2 profit

$$\Pi_M(s_i; \bar{s}, N) = \int_0^v [1 - F(y|s_i)] \cdot \alpha_M(y) \cdot F^*_{1:N-1}(y; \bar{s}) dy - E[\rho_M | N, \bar{s}].$$

Further, any symmetric pure strategy equilibrium has an equivalent equilibrium in which Stage 1 entry decisions can be characterized by a signal threshold s^* such that bidder i enters if and only if $S_i \geq s^*$. This threshold is uniquely determined as follows:

[29] We also impose a normalization on the out-of-equilibrium award rule $\alpha(y)$, but as noted in Lemma 2 this is without loss of generality.

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

$$\Pi_M(s^*; s^*, N) \equiv c.$$

Finally, considered as a function of (c, N), the equilibrium threshold $s^*(c, N)$ is increasing in both arguments, strictly if $s^*(c, N) \in (0, 1)$.

This proposition establishes that the key features of equilibrium entry behavior which our identification argument exploits (a breakeven threshold increasing in entry costs and potential competition) in fact generalize to any RS auction satisfying Assumption 5. Given the auxiliary assumption of Stage 2 identification, it is then straightforward to extend our identification argument to any mechanism in the class considered: at each $(N, \mathbf{z}) \in \mathcal{L}$, observed entry frequencies will identify the equilibrium threshold $s^*(N, \mathbf{z})$, and observed bids will identify the *ex post* distribution $F^*(v; s^*(N, \mathbf{z}))$. Partial identification of model fundamentals can then proceed as in Section 2.

It only remains to discuss conditions under which the auxiliary assumption of Stage 2 identification is reasonable. Not surprisingly, both the equilibrium bid function $\beta_M(\cdot; N, \mathbf{z})$ and the set of bids observed will depend on idiosyncratic features of the mechanism in question, and a general treatment of Stage 2 identification is beyond the scope of this paper. For current purposes, therefore, we simply discuss feasibility under the four standard auction rules: first-price, second-price, English, and Dutch. See Athey and Haile (2002, 2005) for further details.

First-price auctions

Suppose the mechanism M is a first-price auction: the high bidder wins and pays the amount bid. In this case, an entrant with value v_i facing $N-1$ rivals who enter according to s^* chooses bid b_i^* to solve the following optimization problem:

$$\max_b (v_i - b) \cdot G_M^*(b; N, s^*),$$

where $G_M^*(b; N, s^*)$ is the equilibrium probability that bid b results in award under mechanism M at competition (N, s^*).[30] The optimal b_i^* must satisfy the first-order condition

$$(v_i - b_i^*) g_M^*(b_i^*; N, s^*) - G_M^*(b_i^*; N, s^*) = 0, \tag{25}$$

and by hypothesis of equilibrium this solution will be unique. As noted by Guerre, Perrigne, and Vuong (2000), we can therefore invert Equation (25) to obtain the inverse bidding function $\xi_M(\cdot; N, \mathbf{z})$ corresponding to observables $(N, \mathbf{z})$:

$$v_i \equiv b_i^* + \frac{G_M^*(b_i^*; N, s_N^*(\mathbf{z}))}{g_M^*(b_i^*; N, s_N^*(\mathbf{z}))} \equiv \xi_M(b_i^*; N, \mathbf{z}).$$

[30] Note that in principle this probability could depend on both rival bids and mechanism-dependent factors such as reserve prices.

Under the hypothesis of equilibrium behavior, $G^*_M(b; N, s^*_N(\mathbf{z}))$ is the empirical frequency with which bid b results in award under M at observables $(N, \mathbf{z})$, and $g^*_M(b; N, s^*_N(\mathbf{z}))$ is the derivative of this quantity with respect to b. Hence both $G^*_M(\cdot; N, s^*_N(\mathbf{z}))$ and $g^*_M(\cdot; N, s^*_N(\mathbf{z}))$ are identified for any $(N, \mathbf{z})$, which in turn permits identification of the inverse bidding function $\xi_M(\cdot; N, \mathbf{z})$. Identification of $F^*(v; N, \mathbf{z}) \equiv G^*_M\left(\xi^{-1}_M(v; N, \mathbf{z}); N, s^*_N(\mathbf{z})\right)$ follows immediately.

Second-price auctions

As is well-known, in a second-price auction it is a dominant strategy to bid values. Hence identification of $\xi_M(\cdot)$ is trivial: $\xi_M(b; N, \mathbf{z}) \equiv b$ for all $(N, \mathbf{z})$. Identification can then proceed as in text.

English auctions

Theoretical work on English (ascending) auctions tends to focus on the special case of an ascending button auction: the auction proceeds by means of an continuously ascending price, bidders drop out as the auction proceeds, and the last bidder standing wins the auction at the currently posted price. As is well known, when bidders have private values this auction format is strategically equivalent to a second-price auction: it is a dominant strategy for each bidder to remain in until the posted price exceeds their willingness to pay. Hence posted prices at which bidders drop out can be interpreted as reflecting valuations of the exiting bidders.

Relative to the cases considered above, however, an English auction involves two additional complications. First, by construction, the highest bidder's willingness-to-pay is never revealed. Hence data on bids can reveal (at most) valuations for $n-1$ losing entrants; and nonparametric analysis must therefore correct for the fact that only a set of order statistics are observed. Fortunately, there exist well-known results for achieving this; see the relevant sections in, e.g., Athey and Haile (2005) for details. Second, in many real-world applications, the data-generating process is not a true button auction but a sequence of discrete ascending bids. In such cases the second-highest bid need not be identical to the second-highest value, leading to *partial identification* of the corresponding value distribution; see Haile and Tamer (2003) for further discussion. While we do not formally consider this partially identified case, in principle many of our fundamental insights should extend; in particular, replacing ex post distributions with appropriate upper and lower bounds in, e.g., Proposition 3 will leave all relevant inequalities unchanged. We leave this as an extension for future work.

Dutch auctions

The Dutch auction mechanism essentially represents the opposite of an English button auction: the auctioneer posts a continuously declining price, with the object awarded to the first bidder indicating willingness to buy. As is well known, the Dutch auction format is strategically equivalent to the first-price sealed bid auction format, with the highest bid satisfying the same first-order condition as in Guerre, Perrigne, and Vuong (2000). Analytically, the main additional complication in a Dutch auction is that only the highest bid is observed; this is analogous but opposite to the problem noted in button auctions above. Again, however, in an IPV environment it is possible to use standard results on order statistics to map the distribution of the *maximum* entrant value into the ex post distribution of values among entrants. Partial identification of AS fundamentals can then proceed as in Section 2.

Appendix C: Proofs

Proof of Proposition 1. For completeness, we establish the general version of Proposition 1 given in Appendix B, which characterizes entry and bidding behavior in the symmetric pure strategy Bayesian Nash Equilibrium of any RS auction satisfying Assumption 5. As in text, let $F^*_{1:N-1}(v;\bar{s})$ be the probability that the maximum rival value is below v:

$$\begin{aligned} F^*_{1:N-1}(v;\bar{s}) &\equiv [\Pr(S_j \leq \bar{s}) + \Pr(V_j \leq v \cap S_j \geq \bar{s})]^{N-1} \\ &= [\Pr(S_j \leq \bar{s}) + \Pr(V_j \leq v) - \Pr(V_j \leq v \cap S_j \leq \bar{s})]^{N-1} \\ &= [\bar{s} + F_v(v) - F(v,\bar{s})]^{N-1} \\ &= [\bar{s} + F(v,1) - F(v,\bar{s})]^{N-1} \end{aligned}$$

Note the following properties of $F^*_{1:N-1}(v;\bar{s})$:

- $F(v,s)$ is continuous in (v,s) by Assumption 1.1, so $F^*_{1:N-1}(v;\bar{s})$ is *continuous in* $(v,\bar{s})$ for any N.
- $F^*_{1:N-1}(v;\bar{s})$ is *increasing in* $\bar{s}$ for all (v,N):

  $$\begin{aligned} \bar{s} - F(y,\bar{s}) &= \bar{s} - \int_0^{\bar{s}} \int_0^{v} dF(y|t) \cdot 1dt \\ &= \bar{s} - \int_0^{\bar{s}} F(y|t)dt, \end{aligned}$$

  so $\partial\{\bar{s} - F(y,\bar{s})\}/\partial\bar{s} = 1 - F(y|\bar{s}) \geq 0$.

By the Revelation Principle (see Krishna (2009)), any mechanism with an equilibrium in pure strategies has an equivalent direct mechanism such that participants truthfully report types. We can therefore restrict attention WLOG to direct mechanisms with truthful equilibria. In particular, let M be an arbitrary direct mechanism involving allocation rule $Q(\mathbf{v};E)$ and payment rule $P(\mathbf{v};E)$, where $E \equiv (\bar{s},N)$ is an entry structure and $\mathbf{v}$ is a vector of (realized) bidder values, 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 expected allocation and payment functions facing bidder i given entry E.

Now consider an arbitrary bidder with value $v_i \in V$, and permit bidders to report any signal z in $Z = [0,\bar{v}]$. For truth-telling to be an equilibrium, we must have

$$\pi(v_i;E) \equiv q(v_i;E)\cdot v_i - p(v_i;E) \geq \max_{z\in Z}\{q(z;E)\cdot v_i - p(z;E)\}.$$

It follows that $\pi(v_i;E)$ is the maximum of a family of affine functions, which in turn implies that $\pi(\cdot;E)$ is a convex function on $\mathcal{V}$.

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.

Now consider RS auctions specifically. Let $\alpha(\cdot)$ be the probability award is made when the highest reported value is y. By Definition 3, the probability of allocation to an entering bidder with value y is

$$
\begin{aligned}
q(y;E) &= \alpha(y)\cdot \Pr(S_j \geq \bar{s} \cap V_j \leq y\,\forall j) \\
&= \alpha(y)\cdot \prod_{j\neq i} \Pr(S_j \geq \bar{s} \cap V_j \leq y) \\
&= \alpha(y)\cdot F^*_{1:N-1}(y;\bar{s}),
\end{aligned}
$$

and under Assumption 5 low-type profits $\pi_0(E)$ are

$$
\pi_0(E) \equiv \bar{s}^{N-1}\alpha(\underline{v})\underline{v} - \int_0^{\underline{v}} \alpha(y)dy - E^*[\rho|\bar{s},N],
$$

where the low-type payment function $\rho(\cdot)$ is defined as in Assumption 5 and $E^*[\rho|\bar{s},N]$ is the expectation of $\rho(\cdot)$ conditional on rivals entering according to $\bar{s}$. The expected Stage 2 profit of an entrant with value v facing $N-1$ potential rivals who enter according to threshold $\bar{s}$ is therefore

$$
\pi(v;\bar{s},N) = \int_0^{v} \alpha(y)\cdot F^*_{1:N-1}(y;\bar{s})dy - E[\rho|N,\bar{s}]. \tag{26}
$$

Next consider *ex ante* expected profit of an entrant with signal s_i:

$$
\begin{aligned}
\Pi(s_i;\bar{s},N) &\equiv E[\pi(V;\bar{s},N)|S_i=s_i] \\
&= \int_{\underline{v}}^{\bar{v}} \int_0^{v} \alpha(y)\cdot F^*_{1:N-1}(y;\bar{s})dy\,dF(v|s_i) - E[\rho|N,\bar{s}] \\
&= \int_0^{v} [1-F(y|s_i)]\cdot\alpha(y)\cdot F^*_{1:N-1}(y;\bar{s})dy - E[\rho|N,\bar{s}],
\end{aligned}
$$

where the third equality follows by changing order of integration. Note the following properties of $\Pi(s_i;\bar{s},N)$:

- $F(y|s_i)$ is decreasing in s_i for all y by stochastic ordering, so $\Pi(s_i;\bar{s},N)$ is increasing in s_i. $F^*_{1:N-1}(y;\bar{s})$ is increasing in $\bar{s}$ for all y from above, so $\Pi(s_i;\bar{s},N)$ is increasing in $\bar{s}$. $F^*_{1:N-1}(y;\bar{s})$ is decreasing in N for all y, so $\Pi(s_i;\bar{s},N)$ is decreasing in N.

- $\Pi(s;\bar{s},N)$ is *continuous* in $(s;\bar{s})$. For any pairs $(s,\bar{s})$ and $(s',\bar{s}')$,

$$
\begin{aligned}
\left|\alpha(y)F^*_{1:N-1}(y;\bar{s}) - \alpha(y)F^*_{1:N-1}(y;\bar{s}')\right| &\leq \alpha(y)\left|F^*_{1:N-1}(y;\bar{s}) - F^*_{1:N-1}(y;\bar{s}')\right| \\
&\leq \left|F^*_{1:N-1}(y;\bar{s}) - F^*_{1:N-1}(y;\bar{s}')\right|,
\end{aligned}
$$

  and

$$
\begin{aligned}
\left|F(y|s)\cdot F^*_{1:N-1}(y;\bar{s}) - F(y|s')\cdot F^*_{1:N-1}(y;\bar{s}')\right| &\leq \left|F(y|s)\cdot F^*_{1:N-1}(y;\bar{s}) - F(y|s)\cdot F^*_{1:N-1}(y;\bar{s}')\right| \\
&\quad + \left|F(y|s)\cdot F^*_{1:N-1}(y;\bar{s}') - F(y|s')\cdot F^*_{1:N-1}(y;\bar{s}')\right| \\
&= F(y|s)\left|F^*_{1:N-1}(y;\bar{s}) - F^*_{1:N-1}(y;\bar{s}')\right| \\
&\quad + F^*_{1:N-1}(y;\bar{s}')\left|F(y|s) - F(y|s')\right| \\
&\leq \left|F^*_{1:N-1}(y;\bar{s}) - F^*_{1:N-1}(y;\bar{s}')\right| \\
&\quad + \left|F(y|s) - F(y|s')\right|.
\end{aligned}
$$

It follows that

$$
\begin{aligned}
\left|\Pi(s;\bar{s},N)-\Pi(s';\bar{s}',N)\right| &\leq \int_0^{\bar{v}} \alpha(y)\left|F^*_{1:N-1}(y;\bar{s})-F^*_{1:N-1}(y;\bar{s}')\right| dy \\
&\quad + \int_0^{\bar{v}} \alpha(y)\cdot\left|F(y|s)\cdot F^*_{1:N-1}(y;\bar{s})-F(y|s')\cdot F^*_{1:N-1}(y;\bar{s}')\right| dy \\
&\leq 2\int_0^{\bar{v}} \left|F^*_{1:N-1}(y;\bar{s})-F^*_{1:N-1}(y;\bar{s}')\right| dy \\
&\quad + \int_0^{\bar{v}} \left|F(y|s)-F(y|s')\right| dy \\
&= 2\int_0^{\bar{v}} \left|F^*_{1:N-1}(y;\bar{s})-F^*_{1:N-1}(y;\bar{s}')\right| dy \\
&\quad + \left|\int_0^{\bar{v}} F(y|s)dy-\int_0^{\bar{v}} F(y|s')dy\right|,
\end{aligned}
$$

where the last equality follows by monotonicity of $F(y|s)$ and $F^*_{1:N-1}(y;\bar{s})$ in $\bar{s}$. As $|s-s'|\to 0$, the first term converges to zero by continuity of $F^*_{1:N-1}(y;\bar{s})$ in $\bar{s}$, and the second term converges to zero by continuity of $\int_{\mathcal{V}} F(y|s)ds$ in s under Assumption 2. Hence $|\Pi(s;\bar{s},N)-\Pi(s';\bar{s}',N)|\to 0$ as $|s-s'|\to 0$, so $\Pi(s;\bar{s},N)$ is continuous in $(s;\bar{s})$.

Now consider a symmetric equilibrium in threshold strategies. If $\Pi(0;0,N)\geq c$, entry at Stage 1 is a dominant strategy. If $\Pi(1;1,N)\leq c$, remaining out is a dominant strategy. Otherwise, we seek an *interior equilibrium* $s^*\in(0,1)$ such that a marginal entrant (potential bidder with signal s^*) is exactly indifferent to entry:

$$\Pi(s^*;s^*,N)\equiv c. \tag{27}$$

First establish existence of s^*. $\Pi(s_i;\bar{s},N)$ is continuous and increasing in its first two arguments, and $\Pi(0;0,N)>c>\Pi(1;1,N)$ by hypothesis. Consequently at least one s^* solving 27 will exist. Since $\Pi(s_i;\bar{s},N)$ is increasing in s_i, all bidders with $S_i\geq s^*$ will (weakly) prefer to enter, and all bidders with $S_i<s^*$ will (weakly) prefer to remain out. Hence "enter if $S_i\geq s^*$" is a best response to itself, and s^* constitutes a threshold equilibrium.

Next establish uniqueness and monotonicity of s^*. Suppose there exists an interior solution $s^*\in(0,1)$ (else the solution is an endpoint, hence unique). By construction, we then have $\Pi(s^*;s^*,N)=c>0$, which can only obtain if the set $\mathcal{Y}^+\equiv\{y\in\mathcal{V}|[1-F(y|s^*)]\cdot\alpha(y)>0\}$ has positive measure. In particular, this implies $F(y|s^*)<1$ for each $y\in\mathcal{Y}^+$. By the discussion above, $F^*_{1:N-1}(y;s^*)$ will lie in the open interval $(0,1)$, will be strictly increasing in s^*, and will strictly decreasing in N at every y such that $F(y|s^*)<1$. Since $s^*\in(0,1)$ implies the set of such y has positive measure, the integral $\Pi(s^*;s^*,N)$ will be *strictly increasing* in its second argument and *strictly decreasing* in N. Hence any interior solution $s^*\in(0,1)$ will be *unique, strictly increasing in c,* and *strictly increasing in N*. Finally, note that $\bar{s}\to 1$ implies $F^*_{1:N-1}(y;\bar{s})\to 1$ for any y and N, which in turn implies $\Pi(\bar{s};1,N)=\Pi(\bar{s};1,N+1)$ for all $\bar{s}$ and N. Hence if $s^*_N\in(0,1)$ solves (27) N, we must have

$$c=\Pi(s^*_N;s^*_N,N)<\Pi(s^*_N;1,N)=\Pi(s^*_N;1,N+1)\leq\Pi(1;1,N).$$

Combining the above, we conclude that $s^*_N\in(0,1)$ implies $s^*_{N+1}\in(s^*_N,1)$.

It only remains to argue that any symmetric pure strategy equilibrium has an equivalent threshold equilibrium. To see this, choose any symmetric pure strategy equilibrium, and let $\mathcal{E}$ be the set of signals inducing entry in this equilibrium. Suppose that $\mathcal{E}$ is not a threshold set; that is, that there exists $\bar{s} > \min \mathcal{E}$ such that a bidder with signal $\bar{s}$ elects not to enter. Then from above $\Pi(\bar{s}; \mathcal{E}, N) \geq \Pi(\min(\mathcal{E}), \mathcal{E}, N) \geq c$. If either inequality is strict, then a bidder with $S_i = \bar{s}$ would strictly prefer entry, which contradicts equilibrium. Hence $\Pi(\bar{s}; \mathcal{E}, N) = \Pi(\min(\mathcal{E}), \mathcal{E}, N) = c$, which in turn implies $F(y|\bar{s}) = F(y|\min \mathcal{E})$ a.e. by monotonicity of $F(y|s)$. It follows that there exists a payoff-equivalent equilibrium where a bidder with signal $S_i = \bar{s}$ enters and a bidder with signal $S_i = \min(\mathcal{E})$ does not. Iterating this argument over sets of positive measure then establishes the claim.[31]

□

Proof of Proposition 2. Claims 1 and 2 established in text. Claim 3 follows since complete variation implies the condition in Claim 1 is satisfied everywhere.

□

Proof of Proposition 3. We establish claims for $F^+(v|s)$; the argument for $F^-(v|s)$ is analogous. First consider the candidate bound $\check{F}^+(v|\hat{s})$, defined for $\hat{s} \in \mathcal{S}$ as in Equation 10:

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

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

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

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

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

[31] The LS model provides a concrete illustration of this argument. As discussed in Levin and Smith (1994), equilibrium in this model involves random entry by bidders. The AS model interprets such randomization as entry conditional on the realization of an uninformative Stage 1 signal. In this case, any entry set $\mathcal{E}$ of measure $1 - s^*$ will constitute a symmetric pure strategy equilibrium. In practical terms, however, these equilibria are precisely equivalent to that induced by the threshold strategy s^*. As every similar example must involve a similar equivalence, our focus on threshold strategies involves no loss of generality.

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

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

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

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

  where the first inequality follows by hypothesis of stochastic ordering.

Now consider the bounds $F^+(v|s)$ defined in Proposition 3:

- Identification of $F^+(v|s)$ follows from (i) $t^-(s) \in \{0, \mathcal{S}\}$ 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 $s \neq \mathcal{S}$. Thus $F^+(v|s)$ depends only on identified quantities, hence is identified.

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

- Finally, to establish bounds, we consider cases:

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

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

$\square$

Proof of Proposition 4. Expected marginal profit $\Pi_N(s, s^*; F)$ is identified up to F at each N, and $\Pi_N(s, s; F)$ is stochastically ordered in F for all (s, s^*, N). Identification of $c^+(z)$ and $c^-(z)$ and the inequalities $c^+(z) \geq c(z) \geq c^-(z)$ thus 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 $F^{\pm}(y|\bar{s}_N(z)) = F(y|\bar{s}_N(z))$.

$\square$

Proof of Proposition 5. We first establish several basic properties of $\hat{T}_k(\cdot)$:

- $\hat{T}_k(\cdot)$ is *linear*: $\hat{T}_k(\alpha F) \equiv -\int \alpha F(y)\hat{F}_{1:k-1}(y;\hat{s}_k)dy = \alpha(-1)\int F(y)\hat{F}_{1:k-1}(y;\hat{s}_k)dy = \alpha\hat{T}_k(F)$.
- $\hat{T}_k(\cdot)$ is *additively separable:* $\hat{T}_k(F+G) \equiv -\int [F(y)+G(y)]\hat{F}_{1:k-1}(y;\hat{s}_k)dy = \hat{T}_k(F)+\hat{T}_k(G)$.
- $\hat{T}_k(\cdot)$ is *decreasing in* $F(\cdot)$: $F(y) \geq G(y)$ for all y implies $\hat{T}_k(F) \leq \hat{T}_k(G)$.

These properties in turn imply that the pointwise sharp identified set is *convex in* Γ: that is, for any $(\bar{y},\hat{s}_k) \in \mathcal{V}\times\mathcal{S}$, if $(\Gamma,\bar{y},\hat{s}_k) \in \hat{\mathcal{F}}$ and $(\Gamma',\bar{y},\hat{s}_k) \in \hat{\mathcal{F}}$, then $(t\Gamma+(1-t)\Gamma',\bar{y},\hat{s}_k) \in \hat{\mathcal{F}}$ for all $t \in [0,1]$. To see this, note that by definitions 1 and 2 there must exist candidate models $\{\tilde{F},\tilde{c}\}$ and $\{\tilde{F}',\tilde{c}'\}$ passing through $(\Gamma,\bar{y},\hat{s}_k)$ and $(\Gamma',\bar{y},\hat{s}_k)$ respectively. Let $\tilde{F}'' \equiv t\tilde{F}+(1-t)\tilde{F}'$ and $\tilde{c}'' \equiv t\tilde{c}+(1-t)\tilde{c}'$. Then by Condition 3 of Definition 1 we have

$$
\begin{aligned}
\hat{T}_k(\tilde{F}'') + \hat{\kappa}_k &= t[\hat{T}_k(\tilde{F}) + \hat{\kappa}_k] + (1-t)[\hat{T}_k(\tilde{F}') + \hat{\kappa}_k] \\
&= t\tilde{c} + (1-t)\tilde{c}' \\
&= \tilde{c}'' \in [\tilde{c},\tilde{c}'],
\end{aligned}
$$

and by Condition 2 of Definition 1 we have

$$
\begin{aligned}
\int_{\hat{s}_k}^1 \tilde{F}''(y|s)ds &= t\int_{\hat{s}_k}^1 \tilde{F}(y|s)ds + (1-t)\int_{\hat{s}_k}^1 \tilde{F}'(y|s)ds \\
&= t\hat{\lambda}_k(y) + (1-t)\hat{\lambda}_k(y) = \hat{\lambda}_k(y).
\end{aligned}
$$

$\tilde{F}''$ inherits the distributional properties of $\tilde{F}$ and $\tilde{F}''$ and $\tilde{F}''(\bar{y}|\hat{s}_k) \equiv t\Gamma+(1-t)\Gamma'$ by construction, so $\{\tilde{F}'',\tilde{c}''\}$ is a candidate model passing through $(t\Gamma+(1-t)\Gamma',\bar{y},\hat{s}_k)$.

Now turn to Proposition 5 itself. First consider the "only if" direction: $\mathcal{F}$ is pointwise sharp only if for each $\bar{y} \in \mathcal{V}$ and $\hat{s}_k \in \mathcal{S}$, $\hat{\kappa}_k + \hat{T}_k\left[\ddot{F}^+_{\bar{y}k}(\cdot)\right] \geq c^-$ and $\hat{\kappa}_k + \hat{T}_k\left[\ddot{F}^-_{\bar{y}k}(\cdot)\right] \leq c^+$. Suppose there exists $(\bar{y},\hat{s}_k)$ such that $\hat{\kappa}_k + \hat{T}_k\left[\ddot{F}^+_{\bar{y}k}(\cdot)\right] < c^-$. By construction, $\ddot{F}^+_{\bar{y}k}(\cdot)$ is the *smallest* distribution attaining the upper bound $F^+(\bar{y}|\hat{s}_k)$ which is also consistent with $F^-(\cdot|\hat{s}_k)$. Hence $\hat{\kappa}_k+\hat{T}_k\left[\ddot{F}^+_{\bar{y}k}(\cdot)\right] \geq \hat{\kappa}_k + \hat{T}_k[F]$ for any distribution F such that $F(\bar{y}) = F^+(\bar{y}|\hat{s}_k)$ and $F(y) \geq F^-(y|\hat{s}_k)$ for all y. It follows that there can exist no $F(\cdot)$ simultaneously satisfying Conditions 1, 2 and 3 of Definition 1, so $\mathcal{F}$ is not pointwise sharp. The argument for $\hat{\kappa}_k + \hat{T}_k\left[\ddot{F}^-_{\bar{y}k}(\cdot)\right] \leq c^+$ is analogous.

Next consider the "if" direction: $\mathcal{F}$ is pointwise sharp if for each $\bar{y} \in \mathcal{V}$ and $\hat{s}_k \in \mathcal{S}$, $\hat{\kappa}_k + \hat{T}_k\left[\ddot{F}^+_{\bar{y}k}(\cdot)\right] \geq c^-$ and $\hat{\kappa}_k + \hat{T}_k\left[\ddot{F}^-_{\bar{y}k}(\cdot)\right] \leq c^+$. The discussion above implies that the pointwise identified set is convex in Γ. It is thus sufficient to show existence of candidate models attaining the upper and lower envelopes of $\mathcal{F}$ at each point $(\bar{y},\hat{s}_k) \in \mathcal{V}\times\mathcal{S}$. As above, we illustrate the argument for the upper bound $F^+(\bar{y}|\hat{s}_k)$; the argument for $F^-(\bar{y}|\hat{s}_k)$ is analogous.

Suppose $\kappa_k + T_k\left[\ddot{F}_{\bar{y}k}(\cdot)\right] \geq c^-$, let $\tilde{c} \equiv \min\left\{c^+, \kappa_k + T_k\left[\ddot{F}_{\bar{y}k}(\cdot)\right]\right\}$, and note $\tilde{c} \in [c^-, c^+]$ by hypothesis. Define $\tilde{F}_{\bar{y}k}(\cdot)$ as follows:

$$\tilde{F}_{\bar{y}k}(\cdot) \equiv \gamma_{\bar{N}}\ddot{F}(\cdot|\bar{s}_{\bar{N}}) + (1-\gamma_{\bar{N}})F^+(\cdot|\bar{s}_{\bar{N}}),$$

where γ_k is a solution to

$$\kappa_k + \gamma_k T_k\left[\ddot{F}_{\bar{y}k}(\cdot)\right] + (1-\gamma_k)\cdot T_k\left[F^+(\cdot|\bar{s}_k)\right] \equiv \tilde{c}.$$

By construction, $\kappa_k + T_k[\ddot{F}_{\bar{y}k}(\cdot)] \geq \tilde{c}$ and $\kappa_k + T_k[F^+(\cdot|\bar{s}_k)] \equiv c_k^- \leq c^- \leq \tilde{c}$. Thus $\gamma_{\bar{N}}$ exists and lies in the unit interval. Hence $\tilde{F}_{\bar{y}k}(\cdot)$ exists and satisfies $F^-(\cdot|\bar{s}_k) \leq \ddot{F}_{\bar{y}k}(\cdot) \leq \tilde{F}_{\bar{y}k}(\cdot) \leq F^+(\cdot|\bar{s}_k)$. Further, by linearity and additive separability of $T_k(\cdot)$, $\kappa_k + T_k[\tilde{F}_k(\cdot)] \equiv \tilde{c}$.

Next construct candidate distributions $\tilde{F}_j(\cdot)$ for $j \neq k$. Toward this end, choose any $j \neq k$, and define $\tilde{F}_{\bar{y}j}(\cdot)$ as follows:

$$\tilde{F}_j(\cdot) \equiv \gamma_j F^-(\cdot|\bar{s}_j) + (1-\gamma_j)F^+(\cdot|\bar{s}_j),$$

where γ_j is a solution to

$$\gamma_j T_j\left[[F^-(\cdot|\bar{s}_j)\right] + (1-\gamma_j)\cdot T_j\left[F^+(\cdot|\bar{s}_j)\right] \equiv \tilde{c} - \kappa_j.$$

As above, $\kappa_j + T_j[F^-(\cdot|\bar{s}_j)] \equiv c_j^+ \geq c^-$ and $\kappa_j + T_j[F^+(\cdot|\bar{s}_j)] \equiv c_j^- \leq c^+$. Hence γ_j exists and lies on the unit interval. Further, $\tilde{F}_j(\cdot)$ exists, satisfies $F^-(\cdot|\bar{s}_j) \leq \tilde{F}_j(\cdot) \leq F^+(\cdot|\bar{s}_j)$, and by linearity $\kappa_j + T_j[\tilde{F}_j(\cdot)] = \tilde{c}$.

Now take $\tilde{c}$ as a candidate entry cost, and define a candidate conditional distribution $\tilde{F}(\cdot|s)$ on $[\bar{s}_{\underline{\text{N}}}, \bar{s}_{\bar{N}}]$ as follows:

$$\tilde{F}(y|s) = \begin{cases} \tilde{F}_{\bar{y}k}(\cdot) & \text{if } s = \hat{s}_k \\ \tilde{F}_j(\cdot) & \text{if } s = \hat{s}_j \text{ and } j \neq k \\ F^-\left[\cdot|\max\{\bar{s} \in \mathcal{S}|\bar{s} < s\}\right] & \text{otherwise.} \end{cases}$$

Note the following properties of $\tilde{F}(\cdot|s)$:

- For each $s \in [\hat{s}_{\underline{\text{N}}}, \hat{s}_{\bar{N}}]$, $\tilde{F}(\cdot|s)$ is a weighted average of distributions, hence a distribution.

- $\tilde{F}(\cdot|s)$ is monotonically decreasing in s on $[\hat{s}_{\underline{\text{N}}}, \hat{s}_{\bar{N}}]$. This follows since for all $j \in \{\underline{\text{N}}, \bar{N}-1\}$, $\tilde{F}(y|\hat{s}_j) \geq F^-(y|\hat{s}_j) = F^+(y|\hat{s}_{j+1}) \geq \tilde{F}(y|\hat{s}_{j+1})$ at identified thresholds, and $s \in (\hat{s}_j, \hat{s}_{j+1})$ implies $\tilde{F}(y|s) \geq F^-(y|\bar{s}_j) = F^+(y|\bar{s}_{j+1})$ between identified thresholds.

- For all $j \in \{\underline{\text{N}}, \bar{N}\}$, $\kappa_j + T_j[\tilde{F}_j(\cdot|\hat{s}_j)] \equiv \tilde{c}$.

Thus $\tilde{F}(\cdot|s)$ is a candidate conditional distribution satisfying Restrictions (18) and (20) in Definition 1.

It remains to show that $\tilde{F}(\cdot|s)$ satisfies Restriction (19). To see this, note that for $j \in \{\underline{N}, \bar{N}-1\}$,

$$\begin{aligned}
\int_{\hat{s}_j}^{\hat{s}_{j+1}} \tilde{F}(y|s)ds &= \int_{\hat{s}_j}^{\hat{s}_{j+1}} F^-(y|\hat{s}_j)ds \\
&= (\hat{s}_{j+1} - \hat{s}_j)F^-(y|\hat{s}_j) \\
&= (\hat{s}_{j+1} - \hat{s}_j)\frac{(1-\hat{s}_j)F^*(y;\hat{s}_j) - (1-\hat{s}_{j+1})F^*(y;\hat{s}_{j+1})}{(\bar{s}_{j+1} - \bar{s}_j)} \\
&= (1-\hat{s}_j)F^*(y;\hat{s}_j) - (1-\hat{s}_{j+1})F^*(y;\hat{s}_{j+1}).
\end{aligned}$$

Note further that for $j \leq \bar{N} - 2$,

$$\begin{aligned}
\int_{\hat{s}_j}^{\hat{s}_{j+2}} \tilde{F}(y|s)ds &= \int_{\hat{s}_j}^{\hat{s}_{j+1}} \tilde{F}(y|s)ds + \int_{\hat{s}_{j+1}}^{\hat{s}_{j+2}} \tilde{F}(y|s)ds \\
&= (1-\hat{s}_j)F^*(y;\hat{s}_j) - (1-\hat{s}_{j+1})F^*(y;\hat{s}_{j+1}) \\
&\quad +(1-\hat{s}_{j+1})F^*(y;\bar{s}_{j+1}) - (1-\hat{s}_{j+2})F^*(y;\hat{s}_{j+2}) \\
&= (1-\hat{s}_j)F^*(y;\hat{s}_j) - (1-\hat{s}_{j+2})F^*(y;\hat{s}_{j+2}).
\end{aligned}$$

By iteration, it follows that for $k \in \{j, ..., \bar{N}\}$,

$$\int_{\hat{s}_j}^{\hat{s}_k} \tilde{F}(y|s)ds = (1-\hat{s}_j)F^*(y;\hat{s}_j) - (1-\hat{s}_k)F^*(y;\hat{s}_k).$$

In particular,

$$\int_{\hat{s}_j}^{\hat{s}_{\bar{N}}} \tilde{F}(y|s)ds = (1-\hat{s}_j)F^*(y;\hat{s}_j) - (1-\hat{s}_{\bar{N}})F^*(y;\hat{s}_{\bar{N}}).$$

Now consider extending the candidate model $\tilde{F}(\cdot|s)$ to $s > \hat{s}_{\bar{N}}$ as follows: $\tilde{F}(\cdot|s) = F^*(\cdot;\hat{s}_{\bar{N}})$ for $s > \hat{s}_{\bar{N}}$. We know $\tilde{F}(\cdot|\hat{s}_{\bar{N}}) \geq F^-(\cdot|\hat{s}_{\bar{N}}) \equiv F^*(\cdot;\hat{s}_{\bar{N}})$, so $\tilde{F}(\cdot|s)$ satisfies Restriction (18). Further, by construction,

$$\int_{\hat{s}_{\bar{N}}}^{1} \tilde{F}(y|s)ds = (1-\hat{s}_{\bar{N}})F^*(y;\hat{s}_{\bar{N}});$$

so

$$\begin{aligned}
\int_{\hat{s}_j}^{1} \tilde{F}(y|s)ds &= (1-\hat{s}_j)F^*(y;\hat{s}_j) \\
&= \lambda_j(y)
\end{aligned}$$

for all $j \in \mathcal{N}$ and $y \in \mathcal{V}$. We conclude that there exists a candidate model $\tilde{F}(\cdot|\cdot)$ satisfying Restrictions (20), (19), and (18) such that $\tilde{F}(\bar{y}|\bar{s}_k) = F^+(\bar{y}, \bar{s}_k)$. Hence F^+ is pointwise sharp at $(\bar{y}, \bar{s}_k)$.

The argument for F^- makes use of the condition $\hat{\kappa}_k + \hat{T}_k[\ddot{F}^-_{\bar{y}k}(\cdot)] \leq c^+$ and has signs reversed, but is otherwise equivalent. Hence $\hat{\kappa}_k + \hat{T}_k[\ddot{F}^+_{\bar{y}k}(\cdot)] \geq c^-$ and $\hat{\kappa}_k + \hat{T}_k[\ddot{F}^-_{\bar{y}k}(\cdot)] \leq c^+$ for all $(\bar{y}, \hat{s}_k) \in \mathcal{V} \times \mathcal{S}$ implies $\mathcal{F}$ is pointwise sharp. This in turn establishes the proposition.

□

Proof of Proposition 6. Suppose $\mathcal{F}_0$ is an identified set, and consider refinement as in Proposition 6. The proof proceeds in three steps.

First, suppose $\mathcal{F}_j$ is a pointwise identified set, and consider points discarded in iteration j. By construction, these points are identically those where Proposition 5 fails; i.e. those which no candidate model at $\mathcal{F}_j$ can attain. It follows that $\mathcal{F}_{j+1}$ contains every candidate model contained in $\mathcal{F}_j$; i.e. all candidate models by definition of pointwise identified set. Hence if $\mathcal{F}_j$ is a pointwise identified set, then $\mathcal{F}_{j+1}$ is a pointwise identified set. We know $\mathcal{F}_0$ is pointwise identified, so we conclude $\mathcal{F}_j$ is pointwise identified at each j.

Second, the sequence of sets $\{\mathcal{F}_j\}_{j=0}^{\infty}$ converges to a bounded, nonempty limit $\hat{\mathcal{F}}$. First note that (by construction) $\{\mathcal{F}_j\}$ is a *contracting sequence*: $\mathcal{F}_{j+1} \subset \mathcal{F}_j$ for all j. Let χ be any point in the *limit superior* of the sequence $\{\mathcal{F}_j\}$: that is, any tuple $(\Gamma, \bar{y}, \hat{s}_k)$ such that there exist a sequence of points χ_{j_k} and a sub-sequence $\{\mathcal{F}_{j_k}\}$ of $\{\mathcal{F}_j\}$ with $\chi_{j_k} \to \chi$ and $\chi_{j_k} \in \mathcal{F}_{j_k}$ for all j_k. Now define a sequence χ_j for $j = \{0, ..., \infty\}$ as follows: $\chi_j = \chi_{j_{k'}}$, where $k' = \min\{k | j_k \geq j\}$. Since $\{\mathcal{F}_j\}$ is a contracting sequence, $\chi_j \in \mathcal{F}_j$ for all j, and since $\chi_{j_k} \to \chi$, so does χ_j. Thus χ is also in the *limit inferior* of $\{\mathcal{F}_j\}$, which in turn implies that $\{\mathcal{F}_j\}$ has a well-defined limit. Label this limit $\hat{\mathcal{F}}$. By construction, $\mathcal{F}_j$ is bounded within $\mathcal{V} \times \mathcal{S} \times [0, 1]$ for each j, so $\hat{\mathcal{F}}$ is bounded. Further, since $\mathcal{F}_j$ contains the true model pointwise at each j, $\hat{\mathcal{F}}$ must also contain the true model pointwise. Hence $\hat{\mathcal{F}}$ is nonempty.

Third, by construction, $\hat{\mathcal{F}}$ is a pointwise identified set such that the conditions of Proposition 5 hold (otherwise the algorithm would contract further, which is a contradiction). Hence $\hat{\mathcal{F}}$ is pointwise sharp.

$\square$

Proof of Corollary 1. Let $\mathcal{F}_0$ be the pointwise identified set from Proposition 3, and $\mathcal{F}_j$ be the jth iteration in Proposition 6. By construction, the upper bound on $\mathcal{F}_j$ is obtained by contracting the upper bound on $\mathcal{F}_{j-1}$ at points $(\bar{y}, \hat{s}_k)$ where $\hat{\kappa}_k + \hat{T}_k[\ddot{F}^+_{\bar{y}k}(\cdot)] < c^-$ at iteration $j-1$, where the new upper bound by definition satisfies $\hat{\kappa}_k + \hat{T}_k[\min\{\ddot{F}^+_{\bar{y}k}(Y), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_j(\bar{y}|\hat{s}_k)\}] \equiv c^-$. If no contraction point exists, then $F^+_j \equiv F^+_{j-1}$ by hypothesis. Alternatively, if at least one contraction point $(\bar{y}, \hat{s}_k)$ exists, then choose this point. By definition $\ddot{F}^+_{\bar{y}k}(Y) = \min\left\{F^-_{j-1}(Y|\hat{s}_k), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_{j-1}(\bar{y}|\hat{s}_k)\right\}$, so

$$
\begin{aligned}
\min\left\{\ddot{F}^+_{\bar{y}k}(Y), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_j(\bar{y}|\hat{s}_k)\right\} &= \min\left\{F^-_{j-1}(Y|\hat{s}_k), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_{j-1}(\bar{y}|\hat{s}_k), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_j(\bar{y}|\hat{s}_k)\right\} \\
&= \min\left\{F^-_{j-1}(Y|\hat{s}_k), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_j(\bar{y}|\hat{s}_k)\right\}
\end{aligned}
$$

since $F^+_j(\bar{y}|\hat{s}_k) \leq F^+_{j-1}(\bar{y}|\hat{s}_k)$ by contraction. But $F^+_j(Y|\hat{s}_k)$ is increasing in Y by definition and $F^-_{j-1}(Y|\hat{s}_k) \leq F^-_j(Y|\hat{s}_k) \leq F^+_j(Y|\hat{s}_k)$ by construction, so $\min\left\{F^-_{j-1}(Y|\hat{s}_k), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_j(\bar{y}|\hat{s}_k)\right\} \leq F^+_j(Y|\hat{s}_k)$ for all $(Y, \bar{y}, \hat{s}_k)$. It follows that

$$
\begin{aligned}
\hat{\kappa}_k + \hat{T}_k[F^+_j(Y|\hat{s}_k)] &\leq \hat{\kappa}_k + \hat{T}_k[\min\{\ddot{F}^+_{\bar{y}k}(Y), \mathbf{1}[Y \geq \bar{y}] \cdot F^+_j(\bar{y}|\hat{s}_k)\}] \\
&\equiv c^- \text{ by hypothesis.}
\end{aligned}
$$

Thus iteration on $\mathcal{F}_j$ can produce no new information on c.

In particular, at any N such that $\hat{\kappa}_N + \hat{T}_N[F_0^+(Y|\hat{s}_N)] \equiv c^-$, we must have $F_j^+(Y|\hat{s}_N) = F_0^+(Y|\hat{s}_N)$ for all j. Thus a candidate model with $\tilde{F}(Y|\hat{s}_N) = F_0^+(Y|\hat{s}_N)$ will be feasible at each iteration j. By definition of c^-, there exists at least one such N, which in turn implies existence of a candidate model with $\tilde{c} = c^-$. Hence the lower bound c^- is sharp. The argument for the upper bound c^+ is analogous.

$\square$

Proof of Lemma 2 (from Gentry and Li (2012)). Let M be any RS auction satisfying Assumptions 23.1 and 23.2, and suppose that M involves award rule $\alpha(\cdot)$ and induces low-type payment $p_0(\underline{v};\mathbf{w}_{-i})$. Define $\rho(\mathbf{w}_{-i})$ as follows:

$$\rho(\mathbf{w}_{-i}) = p_0(\underline{v};\mathbf{w}_{-i}) - \mathbf{1}[n=1]\left\{\alpha(\underline{v})\underline{v} - \int_0^{\underline{v}} \alpha(y)dy\right\}.$$

Now define $\hat{p}(w_i;\mathbf{w}_{-i})$ for $w_i \leq \underline{v}$ as in Assumption 23:

$$\hat{p}(w_i;\mathbf{w}_{-i}) \quad = \quad \mathbf{1}[n=1]\left\{\alpha(w_i)w_i - \int_0^{w_i} \alpha(y)dy\right\} + \hat{\rho}(\mathbf{w}_{-i}).$$

By construction, $\hat{p}(\cdot)$ induces the same low-type payments as $p_0(\cdot)$ under truthful revelation:

$$\begin{aligned}\hat{p}(\underline{v};\mathbf{w}_{-i}) &= \mathbf{1}[n=1]\left\{\alpha(\underline{v})\underline{v} - \int_0^{\underline{v}} \alpha(y)dy\right\} + \rho(\mathbf{w}_{-i}) \\ &\equiv p_0(\underline{v};\mathbf{w}_{-i}).\end{aligned}$$

It remains only to show that $\hat{p}(\cdot)$ induces truthful revelation from an entrant with type $\underline{v}$. The initial mechanism induces an equilibrium so $w_i > \underline{v}$ cannot be optimal. For $w_i \leq \underline{v}$, the new mechanism induces profit

$$\begin{aligned}\pi(w_i;\mathbf{w}_{-i}) &= \mathbf{1}[n=1]\alpha(w_i)\underline{v} - \hat{p}(w_i;\mathbf{w}_{-i}) \\ &= \mathbf{1}[n=1]\left[\alpha(w_i)[\underline{v}-w_i] + \int_0^{w_i} \alpha(y)dy\right] - \rho(\mathbf{w}_{-i}).\end{aligned}$$

The derivative of this function with respect to i's report w_i is

$$\pi'(w_i;\mathbf{w}_i) = \mathbf{1}[n=1]\alpha'(w_i)[\underline{v}-w_i] \geq 0\,\forall w_i \leq \underline{v}.$$

Hence a low-type bidder can do no better than report $w_i = \underline{v}$. Symmetry and monotonicity of $\rho(\cdot)$ follow from anonymity of the mechanism and Condition 2 of Assumption 5.

$\square$

Proof of Lemma 3 . For any $(s;N)$ pair, expected seller revenue under mechanism M is given by

$$R_M(s;N) = AV(s;N) - N\Pi^*(s;N),$$

where $AV(\cdot)$ 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 $AV(\cdot)$, let $Y_{1:N}$ be the maximum 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:

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

where $F^*_{1:N}(y;s) \equiv [s+(1-s)F^*(y;s)]^N$ is the distribution of $Y_{1:N}$ on $[v_0, \bar{v}]$ given entry threshold s, $f^*_{1:N}(y;s) \equiv N(1-s)F^*_{1:N-1}(y;s)f^*(y;s)$ is the corresponding density, and $\alpha(v_0) \equiv 0$ by Assumption 5.

To obtain $\Pi^*(s;N)$, we start from the result in Equation (3):

$$\begin{aligned} \pi(v;s,N) &= \pi_0(s,N) + \int_{\underline{v}}^{v} \alpha(t)\cdot F^*_{1:N-1}(t;s)dt \\ &= \int_{v_0}^{\underline{v}} \alpha(t)s^{N-1}dt + \int_{\underline{v}}^{v} \alpha(t)\cdot F^*_{1:N-1}(t;s)dt - E[\rho|N,s] \\ &= \int_{v_0}^{\underline{v}} \alpha(t)\cdot F^*_{1:N-1}(t;s)dt - E[\rho|N,s] \\ &= \lambda(v;s,N)\cdot \alpha(v)F^*_{1:N-1}(t;s) - E[\rho|N,s], \end{aligned}$$

where the second equation follows from Assumption 5 and

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

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

Integrating over the *ex post* density $f^*(y;s)$ and multiplying by probability of entry $(1-s)$ gives *ex ante* expected profit $\Pi^*(s;N)$:

$$\Pi^*(s;N) = (1-s)\int_{\underline{v}}^{\bar{v}} \lambda(y;s,N)\cdot\alpha(y)F^*_{1:N-1}(y;s)f^*(y;s)dy - (1-s)E[\rho|N,s]$$

and multiplying by N yields

$$
\begin{aligned}
N\Pi^*(s;N) &= \int_{\underline{v}}^{\bar{v}} \lambda(y;s,N)\alpha(y)\cdot N(1-s)F^*_{1:N-1}(y;s)f^*(y;s)dy - N(1-s)E[\rho|N,s] \\
&= \int_{\underline{v}}^{\bar{v}} \lambda(y;s,N)\alpha(y)\cdot f^*_{1:N}(y;s)dy - N(1-s)E[\rho|N,s] \\
&= \int_{\underline{v}}^{\bar{v}} \lambda(y;s,N)\alpha(y)\, dF^*_{1:N}(y;s)dy - N(1-s)E[\rho|N,s].
\end{aligned}
$$

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

$$
\begin{aligned}
R_M(s;N) &= \int_{v_0}^{\bar{v}} \{\alpha(y)y + [1-\alpha(y)]v_0\}\, f^*_{1:N}(y;s)dy \\
&\quad - \int_{\underline{v}}^{\bar{v}} \lambda(y;s,N)\alpha(y)\, dF^*_{1:N}(y;s) + N(1-s)E[\rho|N,s] \\
&= \int_{v_0}^{\bar{v}} \{\alpha(y)[y-\lambda(y;s,N)] + [1-\alpha(y)]v_0\}\, dF^*_{1:N}(y;s)dy + N(1-s)E[\rho|N,s].
\end{aligned}
$$

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

Identification of $R_M(s;N)$ for $s \in \mathcal{S}$ follows directly from Equation 24: $R_M(\cdot)$ depends only on mechanism components (α, ρ, v_0) (known by hypothesis) and distributions $F^*_{1:N-1}(\cdot;s)$ and $F^*_{1:N}(\cdot;s)$ (identified for $s \in \mathcal{S}$). Thus it only remains to show $R_M(s;N)$ is decreasing in s. Equation (24) implies that s affects seller revenue through (at most) three channels: the per-win profit function $\lambda_\alpha(v;s,N)$, the distribution $F^*_{1:N}(\cdot;s)$, and the residual term $N(1-s)E[\rho|N,s]$. We show that each of these partial effects is negative. For purposes of this derivation, let $F^*_w(y;s) \equiv s + (1-s)F^*(y;s)$ denote the *ex ante* distribution of Bidder i's value (so $F^*_{1:k}(y;s) \equiv F^*_w(y;s)^k$.

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

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

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 $\lambda(v;s,N)$ is increasing in s for all v, so the effect of s on R through $\lambda(v;s,N)$ is negative.

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

$$
\{\alpha(y)[y-\lambda(y;s,N)] + [1-\alpha(y)]v_0\} \tag{28}
$$

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 (28) is increasing in y.

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

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

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

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

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

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

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

Hence increasing y has two effects on the function (28): it increases $[y - \lambda(y; s, N)]$ and shifts weight from v_0 to $[y - \lambda(y; s, N)]$ (through $\alpha(y)$). Since $[y - \lambda(y; s, N)] \geq v_0$, both these effects are positive, so (28) 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 $F_{1:N}^*(y; s)$ is negative.

Finally, note that $\rho(\cdot) \geq 0$ and weakly increasing by Assumptions 5 and 6. Since the distribution of rival realizations $\mathbf{w}_{-i}$ is stochastically decreasing in s, an increase in s therefore implies a decrease in $(1 - s)NE[\rho|N, s]$.

Combining these observations implies that seller revenue $R_M(s; N)$ is decreasing in s for any N.

□

Proof of Lemma 4. We establish claims for $s_M^+(N, \mathbf{z})$; the argument for $s_M^-(N, \mathbf{z})$ is analogous. Suppose s^* is an equilibrium under mechanism M at $(N, \mathbf{z})$. Then Proposition 1 implies $\Pi_M(s^*, N; F) \equiv c(\mathbf{z})$. We know the function $\Pi_M(s, N; \tilde{F})$ is increasing in s and decreasing in $\tilde{F}$, so it follows that

$$c^+(\mathbf{z}) \geq c(\mathbf{z}) \equiv \Pi_M(s^*, N; F) \geq \Pi_M(s^*, N; F^+)$$

Hence if $\Pi_M(s', N; F^+) > c^+(\mathbf{z})$, then $s' > s_M^*(N, \mathbf{z})$. Taking the smallest such s' in $\mathcal{S}$ (or the uninformative bound 1 if no such s' exists) yields the identified bound $s_M^+(N, \mathbf{z})$ defined in Lemma 4, which establishes the claim.

$\square$

Proof of Proposition 7. Follows immediately by combining identification and monotonicity of $R(\bar{s}, N)$ in Lemma 3 with identified bounds on s^* in 4.

$\square$

Proof of Lemma on unobserved heterogeneity in Section 4.1. (Follows proofs of Lemma 1 and Theorem 1 in Hu, McAdams, and Shum (2011).) Fix $N \in \mathcal{L}$ such that $N \geq 3$, and let $\hat{G}(W_1, ..., W_N)$ and $\hat{g}(W_1, ..., W_N)$ denote the joint distribution and (mixed) joint density of realized bids at N (where as above values on the zero axes represent probability masses, and we temporarily suppress dependence on N). Let $\mathcal{D}$ be a discretization satisfying Assumption 9, and $D_i = \mathcal{D}(W_i)$ and $d_i = \mathcal{D}(w_i)$ denote the random variable and realized index induced by this discretization for bidder i. Let $\bar{w}_i \in \mathcal{W}$ denote a particular (fixed) value of W_i, and introduce the following probability arrays in matrix notation:

$$
\begin{aligned}
\hat{M}_{d_j \bar{w}_i | d_k} &= \left[\ \Pr\left\{ D_j = j', W_i = \bar{w}_i | D_k = k' \right\} \ \right]_{\{j',k'\} \in \mathcal{U} \times \mathcal{U}} \\
\hat{M}_{d_j | d_k} &= \left[\ \Pr\left\{ D_j = j' | D_k = k' \right\} \ \right]_{\{j',k'\} \in \mathcal{U} \times \mathcal{U}} \\
\hat{M}_{d_j , d_k} &= \left[\ \Pr\left\{ D_j = j', D_k = k' \right\} \ \right]_{\{j',k'\} \in \mathcal{U} \times \mathcal{U}} \\
M_{d_j | u} &= \left[\ \Pr\left\{ D_j = j' | U = u' \right\} \ \right]_{\{j',u'\} \in \mathcal{U} \times \mathcal{U}} \\
M_{d_j , u} &= \left[\ \Pr\left\{ D_j = j', U = u' \right\} \ \right]_{\{j',u'\} \in \mathcal{U} \times \mathcal{U}} \\
M_{u | d_j} &= \left[\ \Pr\left\{ U = u' | D_j = j' \right\} \ \right]_{\{u',j'\} \in \mathcal{U} \times \mathcal{U}} \\
D_{\bar{w}_i | u} &= \operatorname{diag}\left\{ g(\bar{w}_i | U = 1), ... g(\bar{w}_i | U = K) \right\}. \\
\hat{D}_{d_j} &= \operatorname{diag}\left\{ \Pr(D_j = 1), ..., \Pr(D_j = K) \right\}.
\end{aligned}
$$

These definitions in turn imply the following useful identities: for each $\bar{w}_i \in \mathcal{W}$,

$$
\begin{aligned}
\hat{M}_{d_j \bar{w}_i | d_k} &= M_{d_j | u} D_{\bar{w}_i | u} M_{u | d_k} && (29) \\
\hat{M}_{d_j | d_k} &= M_{d_j | u} M_{u | d_k} && (30) \\
\hat{M}_{d_j , d_k} &= M_{d_j | u} M_{d_k , u}^T && (31) \\
\hat{M}_{d_j , d_k} &= \hat{M}_{d_j | d_k} \hat{D}_{d_k}. && (32)
\end{aligned}
$$

Next observe that full rank for $\hat{M}_{d_j,d_k}$ implies full rank for all other matrices in Equations 29-32. First, since $\hat{M}_{d_j,d_k} = M_{d_j|u} M_{d_k,u}^T$ by Equation 31, we know $\operatorname{rank}(\hat{M}_{d_j,d_k}) = \min\{\operatorname{rank}(M_{d_j|u}), \operatorname{rank}(M_{d_k,u}^T)\}$, which in turn implies $\operatorname{rank}(M_{d_j|u}) = \operatorname{rank}(M_{d_k,u}) = K$ since all matrices are $K \times K$. Similarly, by Equation (32), $\operatorname{rank}(M_{d_j|d_k}) = \operatorname{rank}(D_{d_j}) = K$, and by Equation (30), $\operatorname{rank}(M_{u|d_k}) = K$. Thus if $\hat{M}_{d_j,d_k}$ has full rank, then all other matrices in Equations (30)-(32) also have full rank.

Now proceed as follows. Rearrange (30) to obtain

$$M_{u|d_k} = M_{d_j|u}^{-1}\hat{M}_{d_j|d_k}, \tag{33}$$

and postmultiply both sides of (29) by $\hat{M}_{d_j|d_k}^{-1}$ to get

$$\begin{aligned} \hat{M}_{d_j\bar{w}_i|d_k}\hat{M}_{d_j|d_k}^{-1} &= M_{d_j|u}D_{\bar{w}_i|u}M_{u|d_k}\hat{M}_{d_j|d_k}^{-1} \\ &= M_{d_j|u}D_{\bar{w}_i|u}M_{u|d_k}\hat{M}_{d_j|d_k}^{-1}. \end{aligned} \tag{34}$$

Substituting (33) into (34), we thus obtain:

$$\begin{aligned} \hat{M}_{d_j\bar{w}_i|d_k}\hat{M}_{d_j|d_k}^{-1} &= M_{d_j|u}D_{\bar{w}_i|u}M_{d_j|u}^{-1}\hat{M}_{d_j|d_k}\hat{M}_{d_j|d_k}^{-1} \\ &= M_{d_j|u}D_{\bar{w}_i|u}M_{d_j|u}^{-1}. \end{aligned} \tag{35}$$

The LHS of this equation is identified, and the RHS implies that the LHS has an eigenvalue-eigenvector decomposition. The eigenvalues of this decomposition will be the diagonal elements of $D_{\bar{w}_i|u}$, and the eigenvectors will be the rows of $M_{d_j|u}$. Iterating this argument across $\bar{w}_i \in \mathcal{W}$ will trace out the K conditional densities $g(w_i|\cdot)$, and stochastic ordering of the corresponding distributions $G(W_i|\cdot)$ in u implies a unique map from these densities to the elements of $\mathcal{U}$. Equation (33) then implies identification of $M_{u|d_k}$, hence identification of $\mathbf{F}_{u|N}$ through $\hat{\Pr}\{d_k|N\}$.

□

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 (2010a): "Identification in Differentiated Products Markets Using Market Level Data," Working Paper, Yale University.

——— (2010b): "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.

Chesher, A., and A. Rosen (2011): "Semiparametric Structural Models of Binary Response: Shape Restrictions and Partial Identification," CeMMAP Working Paper CWP30/11.

Chesher, A., and K. Smolinski (2010): "Sharp Identified Sets for Discrete Variable IV Models," CeMMAP Working Paper CWP11/10.

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 (2012): "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, Yale University.

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

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

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

Hu, Y. (2008): "Identification and Estimation of Nonlinear Models with Misclassification Error using Instrumental Variables: A General Solution," *Journal of Econometrics*, 144, 27–61.

Hu, Y., D. McAdams, and M. Shum (2011): "Identification of First-Price Auctions with Non-seperable Unobserved Heterogeneity," Working Paper, Johns Hopkins University.

Hu, Y., and J. Shiu (2012): "Nonparametric Identification Using Instrumental Variables: Sufficient Conditions for Completeness," Working Paper, Johns Hopkins University.

Komarova, T. (2012): "Partial Identification in Asymmetric Auctions in the Absence of Independence," *Econometrics Journal*, forthcoming.

Krasnokutskaya, E. (2011): "Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity," *Review of Economic Studies*, 78, 293–327.

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

Krishna, V. (2009): *Auction Theory*. Academic Press.

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

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 (2010a): "Affiliation and Entry in First-Price Auctions with Heterogeneous Bidders," Working Paper, Vanderbilt University.

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

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.

Li, T., and X. Zheng (2012): "Information Acquisition and/or Bid Preparation: A Structural Analysis of Entry and Bidding in Timber Sale Auctions," *Journal of Econometrics*, 168, 29–46.

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. (2004): *Putting Auction Theory to Work*. Cambridge Univ Pr.

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.

Newey, W. K., and J. L. Powell (2003): "Instrumental Variables Estimation in Non-parametric Models," *Econometrica*, 71, 1565–1578.

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.