Identification of English Auctions When Losing Entrants Are Not Observed

Yunmi Kong[a]

September 14, 2022

Abstract

A typical feature of English auctions and negotiated-price markets modeled as English auctions is that only the transaction price and identity of the winner are observed; auction entrants other than the winner are commonly not recorded in the data. Meanwhile, existing identification results for independent private values, including Athey and Haile (2002), require that the set of entrants be observed. This paper fills the gap by establishing nonparametric identification of asymmetric bidders' value distributions when losing entrants are not observed, under a general class of entry models in which each potential entrant enters the auction with some bidder-specific probability.

Keywords: English auctions, negotiated price markets, entry, unknown number of bidders, nonparametric identification, asymmetric bidders

---

I thank the Editor and referees for thoughtful and constructive comments. This paper contains some content that was previously circulated under the title "Selective Entry in Auctions: Estimation and Evidence".

[a]Department of Economics, Rice University, 6100 Main St, Houston, Texas 77005, USA.
E-mail: yunmi.kong.01@gmail.com

https://doi.org/10.1016/j.ijindorg.2022.102875

1 Introduction

English auctions are one of the most commonly observed auction formats, used to sell objects ranging from fish, timber, and tobacco to art, cars, and wine. In addition, markets with negotiated prices often resemble English auctions and are modeled as such in the economics literature;[1] examples include Woodward and Hall (2012), Allen, Clark, and Houde (2014, 2019), and Cuesta and Sepúlveda (2019) on mortgage and consumer loan markets.

A typical feature of English auction and negotiated-price market data is that only the transaction price and identity of the winning bidder are observed. To establish identification from these observables, most papers reference Athey and Haile (2002). In their Theorem 2, Athey and Haile (2002) establish that value distributions of asymmetric bidders under the independent private values paradigm are nonparametrically identified from auction prices and winner identities alone. An assumption underlying this result, as stated in the paper, is that the "econometrician always observes the number of bidders" n, so that the auction price can be modeled accurately as an $(n-1)$:n order statistic, i.e. the second best among n bidders' valuations.[2] In addition, if bidders are asymmetric, one must observe not only the number but "the set of bidders who participate in each auction".

However, many candidate applications fail to satisfy this condition. In the example of a consumer loan, the econometrician may observe the final interest rate and identity of the lending bank as well as the set of all N banks in the consumer's neighborhood ("potential entrants") but will typically not observe the $n \leq N$ banks from which the consumer actually solicited quotes ("entrants"). To assume $n = N$—that consumers always solicit quotes from every bank in the neighborhood—would be a misspecification and bias estimates of banks' value distributions. Due to their oral nature, even English auctions organized by federal and state governments often lack records of n. This leaves the econometrician ignorant of the true order statistic that the observed auction price represents. Further complicating matters, the unobserved entrants may be different in every auction instance, may be stochastic, and may be

[1] Some models of negotiated price markets, such as sequential search models, are not suited for representation by an English auction. I proceed taking as given the researcher's choice to model their particular application as an English auction.

[2] Under the independent private values paradigm, the English auction is outcome-equivalent to a second-price sealed-bid auction. See Milgrom and Weber (1982).

determined by an endogenous entry process. As a result, there is a gap between the observed data and available identification theorems. Intuitively, an unobserved number of entrants poses a challenge to identification because it is difficult to distinguish whether a high auction price is caused by a large number of entrants or high value distributions.

This paper establishes nonparametric identification of asymmetric bidders' value distributions when the number/set of entrants in an English auction are unobserved. The assumed observables are the auction price, identity of the winning bidder, and set of potential entrants. Requiring knowledge only of *potential* entrants grants a significant advantage for empirical work because data on potential entrants are more readily available than data on entrants to each auction, as in the example above and as I elaborate further below. More precisely, I establish identification under a popular class of entry models in which each potential entrant enters the auction with some bidder-specific probability, which is unknown to the econometrician. I show that the entry probabilities are identified separately from the value distributions. The literature shows that knowledge of these entry probabilities often enables subsequent identification of entry model primitives, as I later illustrate. I emphasize, however, that the focus of this paper is on addressing the differential challenge that arises from the nature of English auction data, rather than on studying entry models per se,[3] and the identification proof itself is agnostic about the specific entry model generating the entry probabilities. Discussion of applicability to specific entry models, including models of selective entry, follows in Section 3. As for the auction paradigm, this paper focuses on independent private values; for English auctions under the common value paradigm, Athey and Haile (2002) establish a strong negative result in which the model is not identified even when all bids are observed and the set of entrants is known.

Concretely, I consider an English auction with N potential entrants, in which each potential entrant i independently enters the auction with probability p_i, and the resulting $n \leq N$ entrants each draw a value from the distribution $F_i(\cdot)$ and bid in the auction. I explain that, upon assigning appropriate placeholder outcomes to auctions

[3] There is an extensive literature on identifying entry models when the set of entrants to each auction is known. Methodological studies and empirical applications include Li and Zheng (2009), Athey et al. (2011), Krasnokutskaya and Seim (2011), Li and Zheng (2012), Marmer et al. (2013), Roberts and Sweeting (2013), Bhattacharya et al. (2014), Gentry and Li (2014), Li and Zhang (2015), and others.

with one or zero entrants, the distribution of auction prices and winner identities generated by this auction model is equivalent to that generated by a translated model in which all N potential entrants 'enter' the auction but now draw their value from an adjusted distribution given by $(1 - p_i) + p_i F_i(\cdot)$. The atom of size $(1 - p_i)$ at the infimum of this distribution represents potential entrant i's probability of not entering the auction. The advantage of the translated model is that there are now N value draws from the adjusted distributions, where N is known, rather than $n \leq N$ value draws from $F(\cdot)$, where n is unknown. After reframing the English auction model in this manner, I apply a result from Nowik (1990), from the literature on competing risks, which shows identification of just such atomic distributions given observables that are equivalent to those available in English auctions. To clarify the distinction from the proof in Athey and Haile (2002), their Theorem 2 applies a result from Meilijson (1981), which is from the same literature as Nowik (1990) but assumes distributions are non-atomic and thus cannot be directly applied to the atomic distributions that arise as above upon translating auctions with probabilistic entry.

In addition to Athey and Haile (2002), closely related work includes Adams (2007) and Komarova (2013). Komarova (2013) provides an alternative to Meilijson (1981)'s proof and a more extensive discussion when it comes to identifying bidders' value distributions in English auctions. A subsection of that paper considers the case of a stochastic number of entrants where the set of entrants is unobserved but the probability distribution of the set of entrants is known by the econometrician. The subsection is intended as an illustration and does not provide a general proof of identification. In the case of symmetric bidders, Adams (2007) proves identification when the observed auction price equals the $(n-1){:}n$ order statistic and the probability distribution of n is known by the econometrician or there is an instrument that shifts that distribution of n in an infinitely continuous manner.

When two or more $(n-m){:}n$ order statistics are observable per auction, such as on eBay, Song (2004) and Freyberger and Larsen (2022) identify the value distribution when neither n nor its distribution are known. There is also a literature studying identification in first-price auctions with missing information on bidders. Guerre and Luo (2022) prove identification from winning bids with unknown numbers of bidders by using the monotonic increase of the upper bound of first-price auction bids in the number of bidders. Meanwhile, An, Hu, and Shum (2010) and Shneyerov and Wong

(2011) study first-price auctions where the number of observed entrants differs from the competition that bidders perceive and base their bidding strategies on.

My paper fills the need for an identification result in English auctions when the econometrician does not know the set of entrants nor the probability distribution thereof. This identification result does not require symmetry of the bidders, instruments, or observability of multiple bids per auction.

2 Identification

Consider an English auction for a single, indivisible item. There is a set of N potential entrants. Following the literature on auctions with entry, I define potential entrants as the universe of bidders that have ex ante a non-zero probability of entering the auction.[4] For example, a common empirical implementation is to count all firms active in the relevant industry, geographic area, and timeframe as potential entrants. This information is readily available compared to records of the specific set of firms that entered each auction. Each potential entrant $i \in \{1, ..., N\}$ enters the auction with probability $0 < p_i < 1$, resulting in $n \leq N$ entrants to the auction. Competitors' entry decisions are unknown at the time of one's own decision, and the entry decision of each potential entrant is independent of the others'.[5] This type of entry equilibrium can be generated by various entry models which I delineate in Section 3. For purposes of the main identification result, it does not matter which entry model is generating p_i so long as there is a fixed p_i for each i.

Upon entering the auction, each entrant learns her private value v_i for the auction item, which is independently distributed across bidders according to $F_i(\cdot)$ with support $[\underline{v}, \bar{v}]$. As the entrant learns her value only after deciding to enter, there is no selection on values here, following Levin and Smith (1994). After first demonstrating identification for this baseline scenario, I explain the extension to selective entry in Section 3. Following Milgrom and Weber (1982)'s model of English auctions, I treat the English auction here as equivalent to a second-price sealed-bid auction. It is a weakly dominant strategy for each entrant to bid their value v_i, and this is unaffected by whether they know the number of entrants n. Then the entrant with the highest

[4] By ex ante, I mean in expectation prior to the realization of any bidder-auction specific random variables.

[5] This rules out models in which potential entrants condition their entry decisions on the decisions of others or entry is based on signals that are correlated even after conditioning on covariates.

v wins the auction and pays a price equal to the second highest v among n entrants. Thus, the auction price is the order statistic $v_{(n-1):n}$; in particular, it is not $v_{(N-1):N}$, as only $n \leq N$ bidders entered and formed valuations of the auction item.

If there are no entrants, the item is not sold. If there is only one entrant, the sole entrant wins the auction at a public, non-binding reserve price, $r \leq \underline{v}$. Section 3 discusses the case of a binding reserve price. Note that as $v_{(n-1):n} \geq \underline{v}$ for $n \geq 2$, there can be auction prices t that equal r or $\underline{v}$ but there cannot be any $t \in (r, \underline{v})$ if $r < \underline{v}$. Thus, the cumulative distribution function of observed t will be flat in $(r, \underline{v})$ if $r < \underline{v}$, and $\underline{v}$ is identified as the supremum of t in that flat region.

The question is whether the value distributions $F_i(\cdot)$ are identified from data typically observed in applications modeled as English auctions. These data include the auction price, identity of the winning bidder, and set of potential entrants. In particular, the set and number of entrants, bids other than the auction price, and the entry probabilities p_i are not observed. Intuitively, it is not obvious whether strong $F_i(\cdot)$ are distinguishable from large p_i, as both would manifest as high auction prices in the observed data. Incorrectly assuming that $p_i = 1$ for all i—i.e. that the auction price equals $v_{(N-1):N}$—would lead to negative bias in value distribution estimates.

To approach this identification problem, I begin by demonstrating that this auction with probabilistic entry can be translated into an auction where all N potential entrants 'enter' but draw their values from adjusted distributions that now have an atom at the infimum.

Proposition 1. *The following two models of second-price sealed-bid auctions generate identical distributions of 'auction prices' t and winner identities.*

1. *Each potential entrant $i \in \{1, ..., N\}$ enters the auction with probability p_i and draws her value from distribution $F_i(\cdot)$, where $F_i(\cdot)$ has support $[\underline{v}, \bar{v}]$. When there are no entrants or the auction price equals the reserve price, assign a placeholder 'auction price' of $t = \underline{v}$.*

2. *Each potential entrant $i \in \{1, ..., N\}$ enters the auction with probability 1 and draws her value from distribution $H_i(\cdot) \equiv [(1 - p_i) + p_i F_i(\cdot)]$, where $H_i(\cdot)$ has support $[\underline{v}, \bar{v}]$. Always define the 'auction price' t to equal the second highest of drawn values. If all N bidders bid $\underline{v}$, the winner identity is $\emptyset$.*

The appendix provides a proof for the general case. Here, I use the simple example of $N = 3$ to illustrate the proposition, denoting the set of potential bidders as $\{i, j, k\}$.

Under model 1 of the proposition, consider the event that potential bidder i wins the auction with auction price $t \in (\underline{v}, \bar{v}]$. Since $t > \underline{v}$, there must have been at least one other entrant besides i. There are three possible combinations of entrants that could have generated this event: $\{i, j\}$, $\{i, k\}$, and $\{i, j, k\}$, which occur with probability $p_i p_j (1 - p_k)$, $p_i(1 - p_j)p_k$, and $p_i p_j p_k$, respectively. Conditional on entrant set $\{i, j\}$, the probability density of the event is $[1 - F_i(t)] f_j(t)$; conditional on $\{i, k\}$, it is $[1 - F_i(t)] f_k(t)$; and conditional on $\{i, j, k\}$, it is $[1 - F_i(t)]\{f_j(t)F_k(t) + f_k(t)F_j(t)\}$. Therefore, summing across the possible entrant combinations, the probability density of the event is

$$\begin{aligned}
&p_i p_j (1 - p_k)[1 - F_i(t)] f_j(t) + p_i(1 - p_j)p_k[1 - F_i(t)] f_k(t) \\
&\quad + p_i p_j p_k [1 - F_i(t)]\{f_j(t)F_k(t) + f_k(t)F_j(t)\} \\
= \;& p_i[1 - F_i(t)]\{p_j f_j[(1 - p_k) + p_k F_k(t)] + p_k f_k[(1 - p_j) + p_j F_j(t)]\} \\
= \;& [1 - H_i(t)]\{h_j(t)H_k(t) + h_k(t)H_j(t)\}.
\end{aligned}$$

The last expression, where $h_i(\cdot) \equiv p_i f_i(\cdot)$ denotes the first derivative of $H_i(\cdot)$, is the probability density of the stated event under model 2 of the proposition, confirming the result.

The intuition for Proposition 1 is as follows. Under model 1 of the proposition, potential bidder i winning with auction price t implies that i entered and bid higher than t, a second-highest bidder $j \neq i$ entered and bid exactly t, and all other potential bidders $k \neq i, j$ either did not enter or entered and bid less than t. The probability of the first component is $p_i[1 - F_i(t)]$, the probability density of the second component is $p_j f_j(t)$, and the probability of the third component for each k is $(1 - p_k) + p_k F_k(t)$. By independence across bidders, the probability density of the event is then $p_i[1 - F_i(t)] \sum_{j \neq i} \left(p_j f_j(t) \prod_{k \neq i,j} [(1 - p_k) + p_k F_k(t)] \right)$. By definition of $H_i(\cdot)$, this equals $(1 - H_i(t)) \sum_{j \neq i} \left(h_j(t) \prod_{k \neq i,j} H_k(t) \right)$, which is the probability density of the same event under model 2 of the proposition.

The standard identification approach is to assume knowledge of the $n \leq N$ entrants to each auction and then identify $F_i(\cdot)$. Here, to deal with unobserved n, I approach identification from the perspective of the second model in Proposition 1, assuming knowledge only of the N *potential* entrants and then identifying $H_i(\cdot)$. Identification of $H_i(\cdot)$ implies identification of both p_i and $F_i(\cdot)$.

Given Proposition 1, I proceed to establish identification of $H_i(\cdot)$ by applying a result from Nowik (1990), from the literature on competing risks. Preceding that, I restate the relevant portion of Nowik's result. The original problem studied by Nowik (1990) is that of a machine made up of N components, each of which have a lifetime v_i with distribution $H_i(\cdot)$. The machine fails when k out of the N components fail. The goal is to identify $H_i(\cdot)$ from the distribution of the observed lifetime of the machine t and the set of components that failed by time t, the "fatal set". A "life-supporting" component is defined as a component that must fail in order for the machine to fail. If there is an atom at the infimum of $H_i(\cdot)$, its size represents the probability that component i is 'dead on arrival'.

Nowik (1990). *Assume that the components' lifetimes $v_i \sim H_i(\cdot)$, $i \in \{1, ..., N\}$, are independent. Assume, further, that $H_i(\cdot)$, $i \in \{1, ..., N\}$, are mutually absolutely continuous and that each possesses a single positive atom at the (common) essential infimum. Then a necessary and sufficient condition for identifiability of all distributions $H_i(\cdot)$ is that there is at most one life-supporting component.*

The English auction in which not all potential entrants enter translates into a competing risks problem where machine components can be dead on arrival. Using the framework of the second model in Proposition 1, the "components of the system" correspond to the N potential entrants. The auction price corresponds to the lifetime of the machine t. The auction stops or "dies" when t exceeds $N-1$ out of N private values each drawn from the respective $H_i(\cdot)$. Therefore, the "fatal set" corresponds to all potential entrants other than the auction winner. If all components of the system are dead on arrival, no component remains alive at time t, which corresponds to the auction winner's identity being $\emptyset$. Importantly, Nowik (1990)'s identification result does not require knowledge of which components were live on arrival—the set of entrants—but only of which components had failed by the time the machine died—the N potential entrants minus the auction winner. Meilijson (1981)'s proof, in assuming non-atomic distributions, does not allow for the possibility of components being dead on arrival and requires that the set of entrants be known. Nowik (1990)'s proof addresses this possibility and requires only that the set of N *potential* entrants be known. Formally,

Theorem 1. *In the independent private values model, assume that $F_i(\cdot)$ are continuous with no constant sections and same support, $0 < p_i < 1$ for all $i \in \{1, ..., N\}$,* *and entry decisions are independent across bidders. Then $F_i(\cdot)$ and p_i for all i are identified from the auction price, the identity of the winner, and the set of potential entrants. If bidders are symmetric, $F(\cdot)$ and p are identified from the auction price and number of potential entrants.*

Proof. By Proposition 1, I reinterpret the auction as one in which all N potential entrants 'enter', but draw values from $H_i(\cdot) \equiv [(1-p_i) + p_i F_i(\cdot)]$ instead of $F_i(\cdot)$. Each $H_i(\cdot)$ has an atom of size $(1-p_i)$ at the infimum of its support, representing the probability (in the original model) that potential entrant i does not enter.

Then the conditions of Nowik (1990)'s theorem are satisfied here: the values v_i are independent across bidders, $H_i(\cdot)$ is without constant sections and possesses an atom at the infimum of its support, and there are no "life-supporting" bidders, since any bidder has a positive probability of winning. The "components of the system" are the set of all N potential entrants, and the "fatal set" is the set of all potential entrants other than the auction winner. Hence the $H_i(\cdot)$ are identified by Nowik (1990), and p_i and $F_i(\cdot)$ are identified as an immediate implication. □

Intuition for the separate identification of p_i is as follows. As discussed in the proof of Proposition 1, we have that $\Pr(\text{no winner}) = \prod_{j=1}^{N}(1-p_j)$ and $\Pr(\text{bidder } i \text{ wins at price } r) = p_i \prod_{j\neq i}(1-p_j)$.[6] Then

$$\frac{\Pr(\text{no winner})}{\Pr(\text{bidder } i \text{ wins at price } r) + \Pr(\text{no winner})} \tag{1}$$

$$= \frac{\prod_{j=1}^{N}(1-p_j)}{p_i \prod_{j\neq i}(1-p_j) + \prod_{j=1}^{N}(1-p_j)} = \frac{1-p_i}{p_i + 1 - p_i} = 1 - p_i.$$

So the p_i are identified as a function of observed probabilities.

3 Application to specific entry models

This section discusses the application of Theorem 1 to specific entry models from the literature. First, consider English auctions with no entry cost but with a public reserve price r, which is binding as $\underline{v} < r < \bar{v}$. Let bidder i's value be an independent draw from $J_i(\cdot)$ on support $[\underline{v}, \bar{v}]$. In this case, the entry probability 'p_i' in Theorem 1

[6] Note that, as this is a second-price auction, winning with an auction price of r does not mean the winning bidder bid r but that the second highest bidder either did not bid or bid exactly r.

corresponds to $1-J_i(r)$ with $0 < 1-J_i(r) < 1$. The '$F_i(\cdot)$' in Theorem 1 corresponds to the value distribution conditional on entry, $[J_i(\cdot) - J_i(r)]/[1 - J_i(r)]$, with support $[r, \bar{v}]$. Thus, Theorem 1 allows identification of $J_i(\cdot)$ on $[r, \bar{v}]$. Identification of $J_i(\cdot)$ below the lowest reserve price is impossible even if all bids are observed and the set of entrants is known.

Second, consider the popular Levin and Smith (1994) model of endogenous entry adopted by Li and Zheng (2009), Athey, Levin, and Seira (2011), and Krasnokutskaya and Seim (2011), among others. In this model, potential entrants have no information about their value prior to entering the auction, as in the baseline scenario of Section 2. There are two versions of this model that have been applied empirically. One is a mixed strategy entry model (e.g., see Athey, Levin, and Seira (2011)) in which each potential entrant has a constant entry cost and is indifferent between entering and not entering the auction in equilibrium. The other is a pure strategy version (e.g., see Krasnokutskaya and Seim (2011)) in which each potential entrant draws an entry cost from a distribution, independently of competitors, and enters the auction if this entry cost is less than her expected gain from entry. In general, the entry equilibrium need not be unique when potential entrants are asymmetric; applications typically verify uniqueness in their specific case or assume that only one equilibrium is being played within the data. For both variations on Levin and Smith (1994), Theorem 1 applies as long as each potential entrant's ex ante probability of entry p_i satisfies $0 < p_i < 1$. However, the entry equilibrium may involve $p_i = 1$ for some i, as in Athey, Levin, and Seira (2011). In this case, it is no longer possible to identify p_i from our English auction data because, e.g., the fraction in (1) is no longer defined.

Third, identification extends to entry models with imperfect selection as represented by Marmer, Shneyerov, and Xu (2013) and Gentry and Li (2014). In a selective entry model, each potential entrant i observes a signal s_i of her value v_i prior to entry and enters only if her signal is sufficiently favorable to offset a fixed entry cost, i.e. if $s_i > \bar{s}_i$ where $\bar{s}_i$ is called an entry threshold. Both the signals and values are drawn independently across potential entrants. Without loss of generality, signals are normalized to have a uniform marginal distribution on [0,1], so that $1 - \bar{s}_i$ equals the probability that potential entrant i enters the auction. The distribution of bidder i's v_i conditional on her signal s_i is $F_i(\cdot|s_i)$ with support $[\underline{v}, \bar{v}]$, where higher signal s_i leads to stochastically dominant $F_i(\cdot|s_i)$. Thus, the value distribution conditional on entry is different from the distribution unconditional on entry, and entry is "selective".

Specifically, the distribution of bidder i's values conditional on entry is

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

Defining $F_i(\cdot) \equiv \int_0^1 F_i(\cdot|s)ds$, a sufficient condition for potential entrant i's probability of entry to satisfy $0 < 1-\bar{s}_i$ is that there exists a signal $s_i \in (0,1)$ such that $\int \left(\int_{\underline{v}}^{v}(v-t)d\prod_{j\neq i}F_j(t)\right)dF_i(v|s_i) > c$, i.e., the expected profit of entry given s_i would be greater than the entry cost even if all $N-1$ potential competitors always entered regardless of signal. A sufficient condition for $1-\bar{s}_i < 1$ is that there exists a signal $s_i \in (0,1)$ such that $\int(v-r)dF_i(v|s_i) < c$, i.e., the expected profit of entry given s_i would be less than the entry cost even if no other bidders entered the auction. As these are sufficient rather than necessary conditions, $0 < \bar{s}_i < 1$ can well occur even if these conditions are not met.

We can translate this auction model in the manner shown in Proposition 1, now replacing the entry probability 'p_i' with $1-\bar{s}_i$ and replacing the entrant's value distribution '$F_i(\cdot)$' with $F_i^*(\cdot;\bar{s}_i)$. That is, the distribution of auction prices and winner identities generated by this English auction model is equivalent to that generated by a model in which all N potential entrants 'enter' the auction but now draw a value from the adjusted distribution $H_i(\cdot) = [\bar{s}_i+(1-\bar{s}_i)F_i^*(\cdot;\bar{s}_i)]$. Then, by Theorem 1, the entry threshold $\bar{s}_i$ and conditional distribution $F_i^*(\cdot;\bar{s}_i)$ are identified from observations of the auction price, identity of the winning bidder, and set of potential entrants. Once $\bar{s}_i$ and $F_i^*(\cdot;\bar{s}_i)$ are identified in this way, we can then employ the methods of Gentry and Li (2014) directly to identify remaining primitives of the selective entry model. For example, the unconditional joint distribution of values and signals for each bidder i, $F_i(v,s)$, is additionally identified provided an instrument that exogenously shifts the entry threshold $\bar{s}_i$. Gentry and Li (2014) provide a detailed exposition on this step, which I do not repeat here. Meanwhile, if the entry probability $1-\bar{s}_i = 1$ for some i, it is no longer possible to apply Theorem 1, analogously to the $p_i = 1$ scenario of the Levin and Smith (1994) model.

Fourth, consider the perfectly selective entry model of Samuelson (1985) in which each potential entrant is perfectly informed about her value prior to her entry decision and enters the auction if $v_i \geq \tilde{v}_i$, where $\tilde{v}_i$ is her entry threshold. In the symmetric case, Theorem 1 allows for identification of $F(\cdot)$ on $[\tilde{v},\bar{v}]$. In the asymmetric case, the entry thresholds may differ among potential entrants, leading to different value supports conditional on entry. This violates the mutually absolutely continuous condition of Nowik (1990), so Theorem 1 does not apply.[7]

Finally, consider consumer search for a loan as in Allen, Clark, and Houde (2014), which they model as an English auction. Consumers have a marginal cost of search effort that is publicly observed by all parties. Allen, Clark, and Houde (2014) do not specify why it is publicly observed; it could be because it is a function of observable characteristics. Consumers with lower search costs choose higher search effort. Consumers do not choose the number of quotes directly, but higher search effort stochastically increases the realized number of quotes.[8] Lenders are ex ante identical, and their idiosyncratic cost shocks are drawn independently only after being included in the set from which the consumer obtains quotes, leading to non-selective entry similar to Levin and Smith (1994). Thus, we could model each potential lender as being included in the quote set with probability p, which is decreasing in consumer search cost, a function of observed covariates. Then Theorem 1 would apply.

4 Concluding remarks

This paper studies identification of entry probabilities p_i and value distributions $F_i(\cdot)$ given the constraints of English auction data. When it comes to identification of the entry model primitives that generate p_i, the literature shows that knowledge of p_i can be used to subsequently identify primitives of the specific entry model involved. For instance, in the symmetric Levin and Smith (1994) model, entry into the auction incurs an entry cost c, representing expenditures such as the cost of evaluating the auction item, developing a valuation, and (physically) participating in the auction. At equilibrium values of p, each potential entrant must be indifferent between entering or not. If we let $\Pi(p)$ denote a potential entrant's ex ante expected gain from entering the auction given p, equilibrium of the entry model requires the zero-profit condition

$$\Pi(p) - c = 0. \tag{2}$$

The ex ante expected gain of entry $\Pi(p)$ is easily computed given $F(\cdot)$ and p, so the

[7]If the respective value supports of each potential entrant conditional on entry were known, repeated application of Theorem 1 on segments of the full support could perhaps be shown to restore identification in some cases, but this is left to future research.

[8]Allen, Clark, and Houde (2014) restrict the realized number of quotes to be either 2 or N. We could extend this to allow any number of quotes according to a binomial distribution.

entry cost c is identified from the equilibrium condition above. Empirical applications that identify entry costs from entry probabilities using this approach include Li and Zheng (2009), Athey, Levin, and Seira (2011), Krasnokutskaya and Seim (2011), and Li and Zheng (2012), among others. Selective entry models such as Gentry and Li (2014) also identify bidders' entry costs through a zero-profit condition resembling (2). Similarly, consumer search applications also rely on entry probabilities and resulting entrant distributions to identify search costs; see e.g. Salz (2022) and Cuesta and Sepúlveda (2019).

When it comes to estimating the value distributions identified according to Section 2, one could approximate $F_i(\cdot)$ with sieves and use a sieve extremum estimator. For example, Kong (2020) uses a sieve maximum likelihood estimator for this purpose in an English auction application with the data limitations posited in this paper. See Komarova (2017) for properties of sieve extremum estimators in k-out-of-n systems and Chen (2007) for general properties of sieve estimators.

Appendix

Proof of Proposition 1

Proof. Let $[N] \equiv \{1, ..., N\}$ and let $\mathcal{P}(X)$ denote all the subsets of set X. Define $\mathcal{S}_{c\setminus\{i,j\}} \equiv \{A \in \mathcal{P}([N] \setminus \{i,j\}) \mid |A| = c\}$, the set of all subsets of $[N]$ of cardinality c that exclude elements i, j. In model 1, the probability density of auction price $t \in (r, \bar{v}]$ and winning bidder i is equal to

$$p_i[1 - F_i(t)] \sum_{j \neq i} \left[p_j f_j(t) \sum_{c=0}^{N-2} \left\{ \sum_{A \in \mathcal{S}_{c\setminus\{i,j\}}} \left(\prod_{\ell \notin A} (1 - p_\ell) \prod_{m \in A} p_m F_m(t) \right) \right\} \right]. \tag{A.1}$$

This expression sums across the probability densities that arise under each of the possible realizations of the set of entrants. The first summation operator sums across the possible identities of the second highest bidder j, which must exist since $t > r$. The second summation operator sums across the possible number c of entrants excluding i and j, which can be as low as zero and as high as $N - 2$, corresponding to the scenario that all potential entrants enter. The third summation operator sums across the possible sets of entrant identities, excluding i and j, with cardinality c.

Label each element of $[N]\setminus\{i,j\}$ as $k_1, ..., k_{N-2}$. The set of all subsets of $[N]\setminus\{i,j\}$,

$\bigcup_{c=0}^{N-2} \mathcal{S}_{c\setminus\{i,j\}}$, can be divided into subsets that exclude k_1 as an element and those that include it. In particular,

$$\bigcup_{c=0}^{N-2} \mathcal{S}_{c\setminus\{i,j\}} = \left\{ \bigcup_{c=0}^{N-3} \mathcal{S}_{c\setminus\{i,j,k_1\}} \right\} \bigcup \left\{ A \cup \{k_1\} \mid A \in \bigcup_{c=0}^{N-3} \mathcal{S}_{c\setminus\{i,j,k_1\}} \right\}.$$

A simple analogy is that the subsets of $\{1,2\}$ can be divided into those that exclude 1 and those that include it as $\mathcal{P}(\{1,2\}) = \{\emptyset, \{2\}\} \cup \{\{1\}, \{1,2\}\} = \{\emptyset, \{2\}\} \cup \{A \cup \{1\} \mid A \in \{\emptyset, \{2\}\}\}$. Using this, and given independence across potential entrants, (A.1) can be rewritten as

$$p_i[1 - F_i(t)] \sum_{j \neq i} \Bigg[p_j f_j(t) \big[(1 - p_{k_1}) + p_{k_1} F_{k_1}(t) \big] \times \sum_{c=0}^{N-3} \Bigg\{ \sum_{A \in \mathcal{S}_{c\setminus\{i,j,k_1\}}} \Big(\prod_{\ell \notin A} (1 - p_\ell) \prod_{m \in A} p_m F_m(t) \Big) \Bigg\} \Bigg], \tag{A.2}$$

where the expression $[(1 - p_{k_1}) + p_{k_1} F_{k_1}(t)]$ sums across the scenarios in which k_1 is excluded from the set of entrants and the scenarios in which it is included, and the second line of (A.2) sums across all the possible entry scenarios of potential entrants other than i, j, and k_1. Next, $\bigcup_{c=0}^{N-3} \mathcal{S}_{c\setminus\{i,j,k_1\}}$ can analogously be divided into subsets that exclude k_2 as an element and those that include it as

$$\bigcup_{c=0}^{N-3} \mathcal{S}_{c\setminus\{i,j,k_1\}} = \left\{ \bigcup_{c=0}^{N-4} \mathcal{S}_{c\setminus\{i,j,k_1,k_2\}} \right\} \bigcup \left\{ A \cup \{k_2\} \mid A \in \bigcup_{c=0}^{N-4} \mathcal{S}_{c\setminus\{i,j,k_1,k_2\}} \right\}.$$

Using this, (A.2) can be rewritten as

$$p_i[1 - F_i(t)] \sum_{j \neq i} \Bigg[p_j f_j(t) \Big(\prod_{k=k_1}^{k_2} \big[(1 - p_k) + p_k F_k(t) \big] \Big) \times \sum_{c=0}^{N-4} \Bigg\{ \sum_{A \in \mathcal{S}_{c\setminus\{i,j,k_1,k_2\}}} \Big(\prod_{\ell \notin A} (1 - p_\ell) \prod_{m \in A} p_m F_m(t) \Big) \Bigg\} \Bigg].$$

Repeated application of this procedure for the remaining elements $k_3, ..., k_{N-2}$ results in

$$p_i[1-F_i(t)]\sum_{j\neq i}\left[p_j f_j(t)\prod_{k\neq i,j}\big[(1-p_k)+p_kF_k(t)\big]\right]. \tag{A.3}$$

Define the derivative of $H_i(\cdot)$ as $h_i(\cdot) = p_i f_i(\cdot)$, where $f_i(\cdot)$ is the density of $F_i(\cdot)$. Simple algebra gives $p_i[1 - F_i(t)] = 1 - H_i(t)$. So (A.3) equals

$$(1-H_i(t))\sum_{j\neq i}\left(h_j(t)\prod_{k\neq i,j}H_k(t)\right).$$

But this expression equals the probability density of auction price $t \in (r, \bar{v}]$ and winning bidder i under model 2.

Finally, consider the special events of the auction price (or the placeholder 'auction price' referenced in model 1 of the proposition) equaling $\underline{v}$ or there being no winner (i.e., the auction item failing to sell). In model 1, the probability of 'auction price' $\underline{v}$ and winning bidder i is the probability that only potential bidder i enters, $p_i \prod_{j\neq i}(1-p_j)$, and the probability of no winner equals that of no potential bidders entering, $\prod_{j=1}^{N}(1-p_j)$. In model 2, the probability of 'auction price' $\underline{v}$ and winning bidder i is the probability that bidder i bids higher than $\underline{v}$ while all other bidders bid $\underline{v}$, $p_i \prod_{j\neq i}(1-p_j)$, and the probability of no winner is the probability that all bidders bid $\underline{v}$, $\prod_{j=1}^{N}(1-p_j)$. This completes the proof. □

References

Adams, C. P. (2007): "Estimating demand from eBay prices," *International Journal of Industrial Organization*, 25, 1213–1232.

Allen, J., R. Clark, and J.-F. Houde (2014): "The Effect of Mergers in Search Markets: Evidence from the Canadian Mortgage Industry," *American Economic Review*, 104, 3365–96.

——— (2019): "Search Frictions and Market Power in Negotiated-Price Markets," *Journal of Political Economy*, 127, 1550–1598.

An, Y., Y. Hu, and M. Shum (2010): "Estimating First-Price Auctions with an Unknown Number of Bidders: A Misclassification Approach," *Journal of Econometrics*, 157, 328–341.

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

Athey, S., J. Levin, and E. Seira (2011): "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions," *Quarterly Journal of Economics*, 126, 207–257.

Bhattacharya, V., J. W. Roberts, and A. Sweeting (2014): "Regulating Bidder Participation in Auctions," *RAND Journal of Economics*, 45, 675–704.

Chen, X. (2007): "Large Sample Sieve Estimation of Semi-Nonparametric Models," in *Handbook of Econometrics*, ed. by J. J. Heckman and E. E. Leamer, Elsevier, vol. 6B, 5549–5632.

Cuesta, J. I. and A. Sepúlveda (2019): "Price Regulation in Credit Markets: A Trade-off between Consumer Protection and Credit Access," https://sites.google.com/site/ignaciocuesta/.

Freyberger, J. and B. J. Larsen (2022): "Identification in ascending auctions, with an application to digital rights management," *Quantitative Economics*, 13, 505–543.

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

Guerre, E. and Y. Luo (2022): "Nonparametric Identification of First-Price Auction with Unobserved Competition: A Density Discontinuity Framework," https://arxiv.org/abs/1908.05476.

Komarova, T. (2013): "A New Approach to Identifying Generalized Competing Risks Models with Application to Second-Price Auctions," *Quantitative Economics*, 4, 269–328.

——— (2017): "Extremum Sieve Estimation in k-out-of-n Systems," *Communications in Statistics - Theory and Methods*, 46, 4915–4931.

Kong, Y. (2020): "Not Knowing the Competition: Evidence and Implications for Auction Design," *The RAND Journal of Economics*, 51, 840–867.

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

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

Li, T. and B. Zhang (2015): "Affiliation and Entry in First-Price Auctions with Heterogeneous Bidders: An Analysis of Merger Effects," *American Economic Journal: Microeconomics*, 7, 188–214.

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, 1397–1429.

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

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

Meilijson, I. (1981): "Estimation of the Lifetime Distribution of the Parts from the Autopsy Statistics of the Machine," *Journal of Applied Probability*, 18, 829–838.

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

Nowik, S. (1990): "Identifiability Problems in Coherent Systems," *Journal of Applied Probability*, 27, 862–872.

Roberts, J. W. and A. Sweeting (2013): "When Should Sellers Use Auctions?" *American Economic Review*, 103, 1830–1861.

Salz, T. (2022): "Intermediation and Competition in Search Markets: An Empirical Case Study," *Journal of Political Economy*, 130, 310–345.

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

Shneyerov, A. and A. C. L. Wong (2011): "Identification in first-price and Dutch auctions when the number of potential bidders is unobservable," *Games and Economic Behavior*, 72, 574–582.

Song, U. (2004): "Nonparametric Estimation of an eBay Auction Model with an Unknown Number of Bidders," http://papers.economics.ubc.ca/legacypapers/dp0514.pdf.

Woodward, S. E. and R. E. Hall (2012): "Diagnosing Consumer Confusion and Sub-Optimal Shopping Effort: Theory and Mortgage-Market Evidence," *American Economic Review*, 102, 3249–76.