Unobserved heterogeneity in auctions

Philip A. Haile[†] and Yuichi Kitamura[‡]

[†]*Department of Economics and Cowles Foundation, Yale University, CEMMAP, and NBER, 37 Hillhouse Avenue, New Haven, CT 06511, USA.*
E-mail: philip.haile@yale.edu

[‡]*Department of Economics and Cowles Foundation, Yale University, and CEMMAP, 30 Hillhouse Avenue, New Haven, CT 06511, USA.*
E-mail: yuichi.kitamura@yale.edu

Summary: A common concern in the empirical study of auctions is the likely presence of auction-specific factors that are common knowledge among bidders but unobserved to the econometrician. Such unobserved heterogeneity confounds attempts to uncover the underlying structure of demand and information, typically a primary feature of interest in an auction market. Unobserved heterogeneity presents a particular challenge in first-price auctions, where identification arguments rely on the econometrician's ability to reconstruct from observables the conditional probabilities that entered each bidder's equilibrium optimization problem. When bidders condition on unobservables, it is not obvious that this is possible. Here we discuss several approaches to identification developed in recent work on first-price auctions with unobserved heterogeneity. Despite the special challenges of this setting, all of the approaches build on insights developed in other areas of econometrics, including those on control functions, measurement error, and mixture models. Because each strategy relies on different combinations of model restrictions, technical assumptions, and data requirements, their relative attractiveness will vary with the application. However, this varied menu of results suggests both a type of robustness of identifiability and the potential for expanding the frontier with additional work.

Keywords: *nonparametric identification*, *control function*, *measurement error*, *finite mixture*, *quasi-control function*, *first-price auctions*.

JEL codes: *C57*, *D44*, *C14*.

1. INTRODUCTION

The econometrics of auctions has been an active area of research over the last thirty years.[1] Auctions often provide applications in which an economic model can be tightly matched to actual market institutions, and where the equilibrium relationships obtained from a rich theoretical literature can often be 'inverted' to allow identification and estimation of model primitives. This makes auctions attractive for applied work aimed at combining theory and data to produce quantitative answers to questions about procurement practices, market design, and the roles of strategic behaviour and asymmetric information in determining market outcomes.

---

[1] Surveys of the literature can be found in, e.g., Hendricks and Paarsch (1995), Athey and Haile (2006, 2007), Hendricks and Porter (2007), and Hickman et al. (2012). See also the monograph of Paarsch and Hong (2006).

In this review we examine recent developments taking on an important challenge in this literature: unobserved heterogeneity. In many applications, one suspects that there is auction-specific information commonly known among bidders but unavailable to researchers. The presence of such unobserved heterogeneity can be important. The underlying information structure is at the heart of many questions concerning bidder behaviour, auction design, the division of surplus, etc., and the distinction between private information and information that is merely omitted from the analysis is essential. Accounting for unobserved heterogeneity is particularly important and challenging in first-price sealed bid auctions (henceforth 'first-price auctions') because standard identification results rely on the econometrician's ability to reconstruct from observables the conditional probabilities entering bidders' first-order conditions.[2] This strategy is threatened if bidders condition on common knowledge information unavailable to the researcher.

Further complicating matters is the problem of endogenous bidder entry. The effect of competition on bids and revenues can be theoretically ambiguous and is, therefore, itself an important empirical question. More broadly, a variety of identification and testing approaches in the auction literature exploit exogenous variation in the level of competition.[3] Once the potential for unobserved auction-level heterogeneity is acknowledged, one naturally suspects that such unobservables will affect bidders' participation decisions as well. Thus, for example, high levels of bidder participation may reflect latent auction-level factors that also alter bidders' valuations or information structure. In the case of first-price auctions, this means that ignoring unobserved heterogeneity can lead to double trouble: misspecification of bidders' first-order conditions and endogeneity of a key covariate.

Some of the errors that will result from ignoring unobserved heterogeneity in auctions are intuitive. For example, one will infer from bids too much within-auction correlation in bidders' private information, and too much cross-auction variation in this information. Recent work has demonstrated that ignoring unobserved heterogeneity can also lead to quantitatively important distortions of other less transparent forms, including (a) erroneous estimates of market power and information rents to bidders (e.g., Krasnokutskaya, 2011; Krasnokutskaya and Seim, 2011; Athey et al., 2011); (b) incorrect conclusions about optimal auction design (e.g., Krasnokutskaya, 2011; Roberts, 2013); and (c) wrong conclusions about the presence/significance of the winner's curse (e.g., Haile et al., 2003; Compiani et al., 2018). Thus, as the empirical auctions literature has matured, it has given increasing attention to strategies allowing identification even in the presence of unobserved heterogeneity.

Standard methods for handling unobserved heterogeneity in econometrics typically are not directly applicable in auction settings. For example, nonlinearity rules out reliance on first differences, and the (typically) small number of bids per auction rules out a fixed effects approach.[4] More fundamental is the fact that the observation-specific components of the 'error terms' in an auction model are equilibrium transformations of bidders' private information, i.e., of key primitives of interest. Assumptions (e.g., standard instrumental variables conditions) that merely restore independence (or conditional independence) of the error terms, therefore, generally do not enable identification of the relevant model primitives. Nonetheless, we will see below that the identification strategies developed for the auction context have connections to methods developed

---

[2] See Guerre et al. (2000) and the extensive literature that follows.

[3] See, e.g., Gilley and Karels (1981), Athey and Haile (2002), Haile et al. (2003), Guerre et al. (2009), Gillen (2010), and Campo et al. (2011).

[4] An exception is the case of certain large multi-unit auctions—e.g., Treasury auctions—where it is possible to treat each auction in isolation. See, e.g., Cassola et al. (2013) and Hortaçsu et al. (2018).

in other areas of econometrics. Some of these connections are more direct than others, and in some cases the new insights discussed here may prove useful in other types of models as well.

Because the primary threats of unobserved heterogeneity concern identification, we focus below exclusively on different approaches to nonparametric identification in the presence of unobserved heterogeneity. And because unobserved heterogeneity is particularly challenging in first-price auctions, we focus exclusively on this case. Each of the approaches we discuss requires assumptions beyond those of a standard baseline model.[5] Some rule out correlation between different bidders' information; some require an auxiliary equation and exclusion restriction; some restrict the support of the unobservable; some restrict the way that unobservables enter the auction model; and some rule out endogenous bidder entry. As a result, none of these approaches dominates another, and the most relevant result in practice will depend on the details of the application, the questions of interest, and the data available. Our goal here is to describe a range of alternative strategies currently available, focusing on the key insights permitting identification without reliance on parametric assumptions.

In the following section we set up an affiliated values auction model with unobserved heterogeneity and point out the key challenge that arises from the presence of unobservables in the equilibrium first-order conditions. Our discussion of solutions to this challenge then follows the history of the literature, beginning with two broad and well-established ideas. Section 3 discusses a control function strategy, where an auxiliary outcome (e.g., the number of bidders entering the auction or the value of the seller's reserve price) is exploited to allow one to indirectly condition on the unobservable. Section 4 then discusses identification obtained by adapting results from the literature on measurement error. Here, one starts from a model with independent bidder types so that observed correlation among bids can be attributed to the unobserved heterogeneity. Bids can then serve as multiple noisy measures of the unobserved state, allowing application of classic or modern results. We then move to two ideas developed in very recent work, both permitting correlated types and common values. Section 5 presents a 'quasi-control function' approach that relaxes the strict monotonicity requirement of the control function strategy, instead exploiting the way that bounds on the unobservable implied by endogenous entry outcomes shift with auction-level observables. Finally, in Section 6 we discuss identification obtained using a nonlinear finite mixture model. This approach restricts the unobserved heterogeneity to have finite support but allows a softening of the index structure required by several other strategies.

2. MODEL

2.1. Setup

Our baseline model is the standard symmetric affiliated values model of first-price sealed bid auctions (Milgrom and Weber, 1982).[6] For simplicity, we focus on the case with no binding reserve price. Bidders in auction t are risk-neutral and indexed by $i = 1, \ldots, N_t$. In addition to

---

[5] We emphasize, however, that we will not attempt to state the most general versions of the results possible. We focus on a version of each approach meant to convey the main insights most clearly.

[6] Milgrom and Weber (1982) considered bidding in a single auction, and by using this model we follow most of the literature in assuming that one observes data from independent auctions. Spatial dependence is discussed by, e.g., Hendricks et al. (2003) and Compiani et al. (2018). Work on the econometrics of dynamic auction models includes Jofre-Bonet and Pesendorfer (2003) and Balat (2011).

the number of bidders N_t, auction t is associated with characteristics

$$C_t = (X_t, Z_t, U_t).$$

Among the three components of C_t are two important distinctions. First, (X_t, Z_t) are observable to the econometrician, whereas $U_t \in \mathbb{R}$ is unobserved. Second, the variables (X_t, U_t) may affect bidder valuations, but Z_t does not—an exclusion restriction we will formalize below when we discuss results that require it.

The realization of (N_t, C_t) is common knowledge among bidders at auction t. Each bidder i has valuation for the good for sale given by a random variable $V_{it} \in \mathbb{R}$. However, beyond (N_t, C_t), each bidder i observes only a private signal $S_{it} \in \mathbb{R}$. Thus, a bidder may not know her own valuation. A bidder's expectation of her valuation given her information set is $E[V_{it}|S_{it}, C_t, N_t]$. More relevant for what follows, however, is a bidder's expected valuation conditional on (S_{it}, C_t, N_t) and an additional assumption that her equilibrium bid is pivotal (see, e.g., Milgrom and Weber, 1982):

$$w(s; n, c) \equiv E\left[V_{it} \,\middle|\, S_{it} = \max_{j \neq i} S_{jt} = s, N_t = n, C_t = c \right].$$

Following Compiani et al. (2018), we refer to $w(S_{it}; N_t, C_t)$ as bidder i's 'pivotal expected value' in auction t. We let $V_t = (V_{1t}, \ldots, V_{n_t t})$, $S_t = (S_{1t}, \ldots, S_{n_t t})$, and $S_{-i} = \{S_{jt}\}_{j \neq i}$.

The affiliated values model incorporates several important special cases often considered in applications. By specifying $S_{it} = V_{it}$, one obtains a private values model, and in that case $w(S_{it}; N_t, C_t) = V_{it}$. A further requirement of mutual independence among $(V_{1t}, \ldots, V_{n_t t})$ conditional on C_t will yield the 'independent private values' (IPV) model.[7] When $E[V_{it}|S_t]$ has nontrivial dependence on S_{-it}, one has a 'common values' model, also known as a model with 'interdependent values'. This is a broad class of models distinguished by the presence of a winner's curse. A special case, referred to as 'pure common values', arises when the value of the good is identical for all bidders.

Given $C_t = c$ and $N_t = n$, let $F_{SV}(S_t, V_t|n, c)$ denote the joint distribution of bidders' signals and valuations in auction t. We assume that this distribution is affiliated, exchangeable in the bidder indices i, admits a $\mathcal{C}^1$ density that is positive on $(\underline{s}, \overline{s})^n \times (\underline{v}, \overline{v})^n$, and is such that the conditional expectation $E[V_{it}|S_{it}, S_{-it}, N_t, X_t, U_t]$ exists and is strictly increasing in S_{it}.

We assume that observed bids reflect a symmetric Bayes Nash equilibrium in pure, strictly increasing, differentiable strategies.[8] Let

$$\beta(\cdot; N_t, C_t) : [\underline{s}, \overline{s}] \to \mathbb{R}$$

denote the equilibrium bidding strategy. Bidder i's equilibrium action (bid) in auction t can then be represented by the random variable

$$B_{it} = \beta(S_{it}; N_t, C_t).$$

---

[7] A relaxation of the IPV model specifies valuations as independent only after conditioning on a latent variable ω_t as well as the common knowledge auction-level information C_t. This is a special case of correlated private values known in the literature as the 'conditionally independent private values' model (see, e.g., Li et al., 2000). An important distinction between that model and models with unobserved heterogeneity is that in the latter U_t is observed by bidders, while in the former ω_t is unknown to bidders. These are different models with sometimes very different implications. For example, revenue equivalence holds under independence conditional on C_t, but not in the conditionally independent private values model. Thus, distinguishing between these information structures is an important motivation for exploring models that permit both dependent private information and unobserved heterogeneity.

[8] See Athey and Haile (2007) for a review of results on existence and uniqueness.

Let $B_t = (B_{1t}, \ldots, B_{n_t t})$. Let $M_{it} = \max_{j \neq i} B_{jt}$ denote the maximum bid among i's competitors at the auction. In effect, bidder i competes only against M_{it}: given $\{S_{it} = s, N_t = n, C_t = c\}$, her equilibrium bid solves

$$\max_b E\left[(V_{it} - b)\, 1\,\{M_{it} < b\}\,|\,S_{it} = s, N_t = n, C_t = c\right]. \tag{2.1}$$

2.2. *The identification challenge*

Our discussion of identification focuses primarily on one primitive of interest: the joint distribution $F_w(\cdot|n, c)$ of bidders' pivotal expected values $(w(S_{1t}; n, c), \ldots, w(S_{nt}; n, c))$. In private values models, this is identical to the joint distribution of valuations conditional on $N_t = n$, $C_t = c$. In more general models permitting common values, identification of $F_w(\cdot|n, c)$ is a form of partial identification that can allow one to address some important questions.[9] Of course, because C_t is not fully observed, one may also be interested in the distribution of U_t conditional on X_t, Z_t, N_t. Thus, we also discuss identification of this conditional distribution.

From the perspective of the econometrician, we assume that the observables consist of N_t, X_t, B_t, and in some cases excluded variables Z_t. Importantly, U_t is unobserved. To see why this poses a potential problem for identification, it is useful to review the pathbreaking insight of Guerre et al. (2000).[10] For simplicity, take the case of private values, where bidder i observes $S_{it} = V_{it} = v$, $N_t = n$, $C_t = c$. Let

$$G_{M|B}(m|b, n, c) = \Pr(M_{it} \leq m|\beta(S_{it}; n, c) = b, N_t = n, C_t = c),$$

and let

$$g_{M|B}(m|b, n, c) = \frac{\partial}{\partial m} G_{M|B}(m|b, n, c).$$

Bidder i's equilibrium bid b must then solve[11]

$$\max_{\tilde{b}} (v - \tilde{b})\, G_{M|B}(\tilde{b}|b, n, c). \tag{2.2}$$

This problem has first-order condition

$$v = b + \frac{G_{M|B}(b|b, n, c)}{g_{M|B}(b|b, n, c)}. \tag{2.3}$$

If there is no unobserved heterogeneity (U_t is degenerate), all terms on the right-hand side of (2.3) are observable. We then immediately have nonparametric identification of each bidder's valuation

---

[9] For example, this partial identification allows (a) testing for common values (e.g., Athey and Haile, 2002; Haile et al., 2003; Compiani et al., 2018); (b) comparisons of expected seller revenues across standard auction designs (e.g., Shneyerov, 2006); (c) assessment of the effect of competition on bidder market power (e.g., Compiani et al., 2018); (d) testing of the hypothesis of equilibrium bidding (e.g., Hendricks et al., 2003; Haile et al., 2003; Compiani et al., 2018); and (e) with additional data on realized bidder profits, quantification of the winner's curse (Hendricks et al., 2003). As discussed by, e.g., Laffont and Vuong (1996) and Athey and Haile (2007), in a common values model (even without unobserved heterogeneity) one is typically forced either to settle for such partial identification or to rely on a combination of additional restrictions and observables beyond bids and auction-level covariates. Hendricks et al. (2003) and Somaini (2018) offer two such approaches.

[10] See also Laffont and Vuong (1993), Li et al. (2000, 2002), Athey and Haile (2002), Hendricks et al. (2003), Haile et al. (2003), Campo et al. (2003), Hortaçsu and McAdams (2010), and Kastl (2011), among many other papers building on this insight.

[11] Observe that because equilibrium bids are strictly increasing in signals, conditioning on the value of a bidder's equilibrium bid is equivalent to conditioning on his signal. Hence, given the private values assumption, (2.1) and (2.2) are equivalent.

and, therefore, of the joint distribution $F_V(V_{1t}, \ldots, V_{nt}|n, c)$ (Li et al., 2002). With unobserved heterogeneity, however, bidders condition on information unavailable to the econometrician: C_t includes U_t. As a result, the right-hand side of (2.3) is no longer observed. Indeed, without additional information or structure, the functions $G_{M|B}$ and $g_{M|B}$ on the right-hand side are not identified.

This challenge extends to the more general affiliated values model, where the first-order condition takes the form (see, e.g., Haile et al., 2003 or Athey and Haile, 2007)

$$w(s_{it}; n_t, c_t) = b_{it} + \frac{G_{M|B}(b_{it}|b_{it}, n_t, c_t)}{g_{M|B}(b_{it}|b_{it}, n_t, c_t)}. \tag{2.4}$$

Absent unobserved heterogeneity, this equation implies identification of each $w(s_{it}; n_t, c_t)$ and, therefore, $F_w(\cdot | n_t, c_t)$. But when the econometrician is unable to observe all elements of c_t, this strategy breaks down. Additional data or structure will be needed to obtain identification of our primitives of interest.

2.3. Separability and 'homogenization'

One form of additional structure relied on by several of the approaches discussed below involves a separability restriction on the way the auction characteristics C_t affect bidder valuations. Suppose for example that

$$V_{it} = \Gamma(X_t, U_t) V_{it}^0, \tag{2.5}$$

where, conditional on N_t,

$$(V_{1t}^0, \ldots, V_{N_t t}^0, S_{1t}, \ldots, S_{N_t t}) \perp\!\!\!\perp (X_t, U_t). \tag{2.6}$$

Alone, (2.5) has no content; however, when combined with (2.6) it requires that, conditional on N_t, (X_t, U_t) affect the auction only through the multiplicative index $\Gamma(X_t, U_t)$. Note that under (2.5) and (2.6) we may take an arbitrary point x^0 and let

$$\Gamma(x^0, 0) = 1 \tag{2.7}$$

without loss of generality.

This type of structure proves useful because separability is preserved by equilibrium bidding. In particular, it is easily confirmed that[12]

$$\beta(S_{it}; n_t, x_t, u_t) = \Gamma(x_t, u_t) \beta^0(S_{it}; n_t), \tag{2.8}$$

where β^0 denotes the symmetric Bayes Nash equilibrium bidding strategy for a standardized auction t at which $\Gamma(X_t, U_t) = \Gamma(x^0, 0) = 1$. Letting $B_{it}^0 = \beta^0(S_{it}; n_t)$, we can rewrite (2.8) as

$$B_{it} = \Gamma(x_t, u_t) B_{it}^0. \tag{2.9}$$

Following Haile et al. (2003), we refer to the random variable B_{it}^0 as bidder i's 'homogenized' bid at auction t. Likewise, we refer to V_{it}^0 as i's 'homogenized valuation' and to $w^0(S_{it}; N_t) \equiv w(S_{it}; N_t)/\Gamma(X_t, U_t)$ as i's 'homogenized pivotal expected value'. One can easily confirm that

---

[12] See Haile et al. (2003), Athey and Haile (2007), or Krasnokutskaya (2011). We focus here on the case of multiplicative separability, although an analogous result holds by essentially the same argument under additive separability.

homogenized pivotal expected values must satisfy the first-order condition

$$w^0(s_{it}; n_t) = b_{it}^0 + \frac{G_{M^0|B^0}\left(b_{it}^0|b_{it}^0, n_t\right)}{g_{M^0|B^0}\left(b_{it}^0|b_{it}^0, n_t\right)}, \tag{2.10}$$

where

$$b_{it}^0 = b_{it}/\Gamma(x_t, u_t)$$

$$M_{it}^0 = M_{it}/\Gamma(x_t, u_t)$$

$$G_{M^0|B^0}\left(m^0|b^0, n\right) = \Pr\left(M_{it}^0 \le m|B_{it}^0 = b^0, N_t = n\right)$$

$$g_{M^0|B^0}\left(m^0|b^0, n\right) = \frac{\partial}{\partial m^0} G_{M^0|B^0}\left(m^0|b^0, n\right).$$

Equations (2.9) and (2.10) imply that, after rescaling bids appropriately, one can proceed as if the data reflected a sample of homogeneous auctions to recover estimates of the (homogenized) pivotal expected values $w^0(s_{it}; n_t)$. Of course, to do so one must first recover the scaling factors $\Gamma(x_t, u_t)$. Observe that knowledge of each $w^0(s_{it}; n_t)$ and $\Gamma(x_t, u_t)$ will imply identification of $F_w(\cdot \mid n, c)$.

3. CONTROL FUNCTION APPROACHES

The earliest and simplest approaches to unobserved heterogeneity in first-price auctions rely on a control function. The essence of the approach is to exploit an auxiliary observable outcome whose (conditional) variation mirrors that of the unobservable U_t, enabling the econometrician to indirectly condition on U_t.[13] This approach was first proposed by Campo et al. (2003) and further developed by Haile et al. (2003), both of whom used the number of bidders N_t as the auxiliary outcome (see also Guerre et al., 2009).

Suppose that the number of bidders at each auction t satisfies

$$N_t = \eta(X_t, Z_t, U_t), \tag{3.1}$$

where the function η is strictly increasing in its final argument. In addition, suppose that Z_t is independent of (S_t, V_t) conditional on X_t, U_t, N_t. As Campo et al. (2003) point out, the existence of a functional relationship between (X_t, Z_t, U_t) and N_t arises naturally when the meaningful decision to 'enter' an auction takes place before a bidder learns her signal and entry outcomes do not reflect randomization (e.g., mixed strategies). The monotonicity requirement captures the natural idea that unobservables can be ordered in a way such that auctions with 'better' unobservables attract more bidders. An important limitation of the strict monotonicity requirement, however, is that this restricts U_t to have finite support, with cardinality equal to that of $N_t|X_t, Z_t$.[14]

---

[13] This is more than typically required of a control function in a regression setting, where conditioning on a control variable need only deliver independence between regressors and the structural error in the outcome equation (e.g., Chesher, 2003; Imbens and Newey, 2009). Identification of the auction model is equivalent to identification of the bidding equation $B_{it} = \beta(S_{it}; N_t, X_t, U_t)$ ((2.4) is the inverse of this equation), where both S_{it} and U_t are structural errors. Knowledge of objects like average effects or the 'quantile structural function' generally are not sufficient to address the economic questions of interest in an auction setting. Olley and Pakes (1996) provide an early example of this kind of approach in a very different economic setting.

[14] In Section 5 we discuss an approach relaxing the strict monotonicity requirement.

The value of strict monotonicity for identification comes from the fact that conditioning on the observables (X_t, Z_t, N_t) indirectly fixes U_t. And when combined with the assumed conditional independence of Z_t, this implies that (taking $N_t = n$) the joint distribution of bids $(B_{1t}, \ldots, B_{nt})$ conditional on (X_t, Z_t, N_t) is identical to that conditional on (X_t, U_t, N_t). Thus we can rewrite the first-order condition (2.4) as

$$w(s_{it}; n_t, c_t) = b_{it} + \frac{G_{M|B}(b_{it}|b_{it}, n_t, x_t, z_t)}{g_{M|B}(b_{it}|b_{it}, n_t, x_t, z_t)}.$$

Now all terms on the right-hand side are observable, yielding identification of the realizations of $w_{it} = w(s_{it}; n_t, c_t)$ for all i and t, although the value of each c_t is not uniquely determined.

Although this result alone will be sufficient to answer some questions,[15] we can obtain a stronger result by requiring the instrument Z_t to provide the missing source of independent variation in X_t, U_t, and N_t.[16] Suppose, for example, that N_t is the only endogenous component of (N_t, X_t, Z_t), i.e.,

$$(X_t, Z_t) \perp\!\!\!\perp U_t.$$

This immediately yields identification of the function η, implying identification of the realization of each U_t as well. Since U_t is the only unobservable component of C_t, we now have identification of each realization $w(S_{it}; n, x, u)$ given any (n, x, u) in the support of (N_t, X_t, U_t). This implies identification of the conditional distribution $F_w(\,\cdot\,|N_t, X_t, U_t)$ on this support, as well as the distribution of U_t conditional on any subvector of (X_t, Z_t, N_t). Note that in this case, as long as Z_t has a nontrivial effect on N_t through (3.1), this instrument provides an exogenous source of variation in the level of competition. Such exogenous variation in competition has been exploited by Haile et al. (2003) and Guerre et al. (2009) to provide an approach for detecting the presence of common values and to identify a model with risk-averse bidders, respectively, while allowing for unobserved heterogeneity.

An attractive aspect of the control function approach is that it requires no restriction on functional form or on the information structure of the baseline auction model. This distinguishes it from all approaches discussed below. But the control function approach has limitations as well. We have already mentioned the implications of the strictly monotone relation (3.1) for the support of the unobservable. This structure also rules out 'selective entry' by bidders based on the realizations of their private signals (cf. Marmer et al., 2013; Gentry and Li, 2014; Bhattacharya et al., 2014; or Kong, 2017a). A further limitation is that reliance on a reduced form for bidder entry implies that some types of counterfactuals (those that would alter the reduced form) will not be identified, at least without additional work.[17]

One can avoid these limitations if there is a continuous outcome that can replace the number of bidders as the auxiliary outcome mirroring variation in U_t. Roberts (2013) points out that in many applications there is a natural candidate: the seller's reserve price.[18] If the seller observes the unobserved heterogeneity, it may be natural to assume that the reserve price is strictly increasing in U_t conditional on other observables. Focusing on the case of independent private values,

[15] For example, in a private values setting this reveals the division of surplus between seller and buyers.

[16] Thus far we have not required any variation in Z_t. Natural instruments may be measures of the number of potential bidders (e.g., Haile 2001; Hendricks et al., 2003; Haile et al. 2003), or of bidder entry costs (e.g., Kong, 2017a; Compiani et al. 2018).

[17] Of course, once the auction model primitives are identified using the reduced form entry model, it may become possible to identify a structural model of entry consistent with the reduced form.

[18] Adapting bidder first-order conditions to the case of a binding reserve price is straightforward. See, e.g., Milgrom and Weber (1982), Guerre et al. (2000), and Athey and Haile (2002).

Roberts (2013) shows that, given a standard regularity condition, this strict monotonicity follows if the seller observes U_t and sets the profit-maximizing reserve price.[19]

4. MEASUREMENT ERROR APPROACHES

The control function approach discussed above lets the researcher pin down, or fix, the value of U_t through an auxiliary observable. In effect, such a strategy makes the unobservable observable. This is not the only possible approach. In fact there is a close parallel between nonlinear measurement error models and econometric models with unobserved heterogeneity. Results from the measurement error literature can therefore offer useful identification strategies for auction models with unobserved heterogeneity. With these approaches one forgoes pinning down the realization of the unobserved heterogeneity for each observation and instead relies on a decomposition of observable distributions into components reflecting model primitives and those reflecting unobserved heterogeneity.

Krasnokutskaya (2011) was the first to apply a measurement error approach to auctions with unobserved heterogeneity. Building on Li (2000), her result relies on the celebrated lemma of Kotlarski (1967).[20] Kotlarski's Lemma is concerned with nonparametric deconvolution with repeated measurements subject to independent separable measurement error. Consider a triple of mutually independent random variables (Y^*, η_1, η_2) and let

$$Y_1 = Y^* + \eta_1$$
$$Y_2 = Y^* + \eta_2.$$

For normalization purposes let $E[\eta_1] = 0$. Assuming that the characteristic functions of Y_1 and Y_2 are non-vanishing,[21] Kotlarski (1967) shows that the distribution $F(Y^*, \eta_1, \eta_2) = F(Y^*)F(\eta_1)F(\eta_2)$ is identified from the joint distribution of (Y_1, Y_2). This result is constructive, which can be useful for nonparametric estimation.

Kotlarski's Lemma has wide applications in econometrics: it applies to measurement error problems, panel data, and, as Krasnokutskaya (2011) shows, auction models with unobserved heterogeneity. Krasnokutskaya (2011) considers an IPV model in which the unobserved heterogeneity enters multiplicatively, i.e.,

$$V_{it} = U_t V_{it}^0,$$

where

$$U_t \perp\!\!\!\perp V_{1t}^0 \perp\!\!\!\perp V_{2t}^0 \cdots \perp\!\!\!\perp V_{N_t t}^0.$$

Here V_{it}^0 can be regarded as i's homogenized valuation as before.[22] By the preservation of separability under equilibrium bidding, we then have

$$B_{it} = U_t B_{it}^0,$$

---

[19] Following Matzkin (2003), Roberts (2013) shows that the assumption of independence between (X_t, Z_t) and U_t may be relaxed under additional restrictions.

[20] Li et al. (2000) introduced Kotlarski's Lemma to the auction literature in a closely related model of conditionally independent private values without common knowledge unobservables.

[21] Evdokimov and White (2012) provide some results relaxing this assumption.

[22] Here we implicitly condition on any observable auction characteristics X_t.

with

$$U_t \perp\!\!\!\perp B_{1t}^0 \perp\!\!\!\perp B_{2t}^0.$$

By considering (any) two bidders at auction t, we have

$$\ln B_{1t} = \ln U_t + \ln B_{1t}^0$$
$$\ln B_{2t} = \ln U_t + \ln B_{2t}^0.$$

It is then immediate from Kotlarski's Lemma that the joint distribution of $(U_t, B_{1t}^0, B_{2t}^0)$ is nonparametrically identified from the observable joint distribution of (B_{1t}, B_{2t}). Using the first-order condition (2.10), one can then recover the marginal distributions of valuations for bidders 1 and 2. In the case of more than 2 bidders, a similar argument yields the distribution of valuations for other bidders. In the case of symmetric bidders, this result implies a form of overidentification, allowing falsification of the model.

Note that the data requirement for Krasnokutskaya's identification results is modest, and it does not require an instrument. Also, whereas some of the approaches we discuss require that the unobserved heterogeneity be discrete, here such an assumption is unnecessary. And although we have assumed symmetry for simplicity, this is not required. A critical requirement, however, is the independence of bidder types. In essence, therefore, this kind of approach is limited to IPV settings.[23] A separable structure and statistical independence between the unobserved heterogeneity and valuations are also critical; it is this structure that creates the equivalence to the classical measurement error setting.

An alternative approach that is also based on a result from the measurement error literature has been proposed by Hu et al. (2013). These authors build upon an identification result for nonlinear measurement error models due to Hu (2008). Consider a random vector (Y, W^*, W, Z), where Y denotes the dependent variable, W^* the unobserved independent variable, W a mismeasured indicator of W^*, and Z an instrument. Here W^* is assumed to be supported by a finite set, and W and Z are assumed to have the same finite support. The goal is to identify the joint distribution function $F_{Y,W^*,W,Z}$ of (Y, W^*, W, Z). Suppose

$$Y \perp\!\!\!\perp W \perp\!\!\!\perp Z | W^*, \tag{4.1}$$

which, among other things, allows for nonclassical measurement error. Also assume a full rank condition for the distribution of W^* given Z, which is essentially an instrument relevance condition. Then with additional conditions that guarantee uniqueness of the eigenvector–eigenvalue decomposition for an observable matrix, Hu (2008) shows that $F_{Y,W^*,W,Z}$ is identified.

Hu et al. (2013) show that the above result can be used to establish identification in auction models with unobserved heterogeneity when there are at least three bidders. Like that of Krasnokutskaya (2011), the identification strategy in Hu et al. (2013) is essentially limited to the IPV model, although here one can dispense with any separability requirement. In particular, Hu et al. (2013) assume

$$V_{1t} \perp\!\!\!\perp V_{2t} \perp\!\!\!\perp ... \perp\!\!\!\perp V_{N_t t} | U_t, \tag{4.2}$$

and that U_t is finitely supported, as in Hu (2008). Pick an arbitrary three bidders (without loss, the first three of the N_t bidders). The independence above implies that

$$B_{1t} \perp\!\!\!\perp B_{2t} \perp\!\!\!\perp B_{3t} | U_t. \tag{4.3}$$

---

[23] Common values models with independent types exist, but are not easily motivated in applications.

Hu et al. ([2013](#)) treat B_{1t}, B_{2t}, B_{3t} and U_t, respectively, as the dependent variable Y, the mismeasured indicator W, the instrument Z, and the true measurement W^* in Hu's measurement error model described above. To do this, they discretize B_{2t} and B_{3t}. Moreover, Hu et al. ([2013](#)) show that if $V_{it}|U_t$ is increasing in U_t in terms of first-order stochastic dominance (FOSD), then the rank condition and the uniqueness conditions for eigenvector–eigenvalue decomposition in Hu ([2008](#)) are satisfied.[24] Thus, by the result above, the joint distribution of $(B_{1t}, B_{2t}, B_{3t}, U_t)$ is identified (up to the discretization). This in turn allows one to recover the joint distribution of valuations and the unobserved heterogeneity through the first-order condition ([2.4](#)).

There have been at least two important extensions of this result in the recent literature. Gentry and Li ([2014](#)) have shown that it can be used to obtain identification in an IPV model with endogenous bidder entry under the additional assumption that higher values of the unobservable lead to higher entry probabilities. They allow for a binding reserve price and selective entry, the latter being ruled out by the control function and quasi-control function approaches discussed in Sections 3 and 5, respectively. Gentry and Li's result does not require an instrument for entry, but can incorporate instruments when a source of exogenous variation is required. Balat ([2011](#)) has shown that one need not rely on bids to provide the analogues of the 'measurements' and instrument in Hu ([2008](#)). He instead exploits the availability of multiple measures of participation—that at sequential stages (e.g., pre-qualification and bidding) or among multiple subgroups (e.g., large firms and small firms), each of which is assumed to respond (stochastically) to the auction-level unobservable. When such data are available, a significant advantage of this approach is that it can allow one to drop the assumption of independent bidder types.

Yet another identification strategy that could be described as a measurement error approach is due to D'Haultfoeuille and Fevrier ([2015](#)).[25] This paper first develops a nonparametric identification result for a model taking the form

$$Y \perp\!\!\!\perp W \perp\!\!\!\perp Z | W^* \tag{4.4}$$

where the triple (Y, W, Z) is observable whereas W^* is latent. This setting is similar to ([4.1](#)) considered by Hu ([2008](#)), although here W^* is assumed to be continuously distributed and a different set of assumptions is imposed. In particular, D'Haultfoeuille and Fevrier ([2015](#)) impose strict monotonicity properties of the boundaries of supp(Y, W, Z) as functions of W^*. They show that one can then use the observed values of any two components of (Y, W, Z) to hold W^* fixed, revealing the distribution of the third component conditional on W^*. This yields identification of the conditional distributions of Y, W and Z given W^*. Because the unconditional distribution of each component is directly observed, this implies identification of the distribution of W^* as well.

This result naturally applies to measurement error problems, but D'Haultfoeuille and Fevrier ([2015](#)) also discuss its application to auction models. More specifically, consider an IPV model with valuations with multiplicative unobserved heterogeneity as in Krasnokutskaya ([2011](#)) and Hu et al. ([2013](#)). Then, under the conditional independence assumption ([4.2](#)), the bids for any three bidders once again satisfy ([4.3](#)), which is conformable to the model ([4.4](#)), with unobserved heterogeneity U_t continuously distributed. More importantly, from the form of the equilibrium

---

[24] This FOSD condition is implied by multiplicative (or additive) separability. Luo ([2018](#)) shows that the FOSD requirement can be weakened.

[25] Alternatively, we may consider the results by Hu ([2008](#)) and D'Haultfoeuille and Fevrier ([2015](#)) as identification strategies for nonparametric mixture models, with W^* in ([4.1](#)) and ([4.4](#)) interpreted as a mixing variable. We provide further identification results using mixture models in Section 6. The approach of D'Haultfoeuille and Fevrier ([2015](#)) also shares features of the control function approach, in that strict monotonicity conditions are used to indirectly condition on the unobservable.

bidding function, D'Haultfoeuille and Fevrier ([2015](#)) note that boundaries of the supports of B_{1t}, B_{2t}, B_{3t} conditional on U_t are all increasing in U_t, fulfilling the above boundary monotonicity requirement. Consequently, an application of the general identification result shows that the joint distribution of the bids and U_t is recovered, allowing identification of the joint distribution of the valuations and U_t.[26]

5. A 'QUASI-CONTROL FUNCTION' APPROACH

Compiani et al. ([2018](#)) have recently considered an approach that shares many features of the control function strategy (Section 3) while avoiding its strict monotonicity requirement and the resulting restrictions on the support of U_t. Compiani et al. ([2018](#)) consider an affiliated values model with the separable structure of assumptions (2.5) and (2.6). They assume, in addition, that the index function $\Gamma(X_t, U_t)$ scaling bidder valuations is weakly increasing in U_t.

Without loss, let U_t be uniform on [0,1]. Like Haile et al. ([2003](#)), Compiani et al. ([2018](#)) assume that the number of bidders participating in auction t can be represented by a reduced-form relation

$$N_t = \eta(X_t, Z_t, U_t),$$

with

$$U_t \perp\!\!\!\perp (X_t, Z_t)$$

and

$$F_{SV}(S_t, V_t | N_t, Z_t, X_t, U_t) = F_{SV}(S_t, V_t | N_t, X_t, U_t).$$

However, Compiani et al. ([2018](#)) require only weak monotonicity of η in U_t. Thus, Compiani et al. ([2018](#)) assume that unobservables making the good for sale (weakly) more valuable also lead to (weakly) higher levels of participation.

Compiani et al. ([2018](#)) motivate this structure with an example of a fully specified two-stage game of entry and bidding, where entering the auction involves costly acquisition of a signal S_{it}. For that example, they show that the representation of unobservables with a scalar U_t that is independent of X_t can be obtained as a result rather than an assumption, as can the weak monotonicity conditions.[27] Thus, these requirements may be less restrictive than they initially appear.

With weak monotonicity of η, the observed values of (N_t, X_t, Z_t) do not determine the realizations of U_t, although they do imply bounds. Compiani et al. ([2018](#)) show that point identification of the model can be obtained by exploiting the way that these bounds alter the support of equilibrium bids.[28]

---

[26] D'Haultfoeuille and Fevrier ([2015](#)) also consider another application of their results to auction models with U_t playing the role of an unobserved reserve price.

[27] As discussed by Compiani et al. ([2018](#)), this representation is obtained without loss with respect to many questions motivating estimation of an auction model. However, separate identification of the effects of X_t and U_t on the auction could be obtained only under additional restrictions.

[28] An interesting open question is whether useful partial identification results could be obtained using these bounds without the index structure relied upon by Compiani et al. ([2018](#)). Related ideas have been explored in other contexts by Manski and Tamer ([2002](#)) and Chesher ([2005](#)).

Their argument proceeds in three steps. First, consider the bounds on each realized u_t. Fix $(X_t, Z_t) = (x, z)$ and let the (conditional) support of N_t be

$$\left\{ \underline{n}(x, z), \underline{n}(x, z) + 1, \ldots, \overline{n}(x, z) \right\}.$$

The function $\eta(x, z, \cdot)$ is then characterized by a set of thresholds: $N_t = n$ if and only if $U_t \in [\tau_{n-1}(x, z), \tau_n(x, z)]$. Observed conditional entry probabilities therefore satisfy

$$\Pr(N_t = n | X_t = x, Z_t = z) = \tau_n(x, z) - \tau_{n-1}(x, z) \qquad n = \underline{n}(x, z), \ldots, \overline{n}(x, z). \quad (5.1)$$

Of course, $\tau_{\underline{n}(x,z)-1}(x, z) = 0$ and $\tau_{\overline{n}(x,z)}(x, z) = 1$, giving an initial value from which to solve (5.1) for all thresholds $\tau_n(x, z)$. By (5.1), these known thresholds bound the realization of each U_t; indeed, conditional on $(X_t, Z_t, N_t) = (x, z, n)$, U_t is uniform on $[\tau_{n-1}(x, z), \tau_n(x, z)]$.

Next, to pin down the function Γ, take logs of (2.9) to obtain

$$\ln B_{it} = \gamma(X_t, U_t) + \ln \beta^0(S_{it}, N_t), \quad (5.2)$$

where we have defined $\gamma(X_t, U_t) = \ln \Gamma(X_t, U_t)$. Now observe that

$$\sup\{\ln B_{it} | N_t = n, X_t = x, Z_t = z\} = \gamma(x, \tau_n(x, z)) + \ln \beta^0(\bar{s}; n)$$

while

$$\sup\{\ln B_{it} | N_t = n, X_t = \hat{x}, Z_t = \hat{z}\} = \gamma(\hat{x}, \tau_n(\hat{x}, \hat{z})) + \ln \beta^0(\bar{s}; n).$$

By differencing these equations, one can determine all first differences of the form

$$\gamma(\hat{x}, \tau_n(\hat{x}, \hat{z})) - \gamma(x, \tau_n(x, z)).$$

Additional first differences can be obtained by varying the value of N_t conditioned on, and differencing first differences with common terms. More first differences are obtained by applying similar argument using the infimum rather than the supremum. Given 'enough' of these first differences, the fact that we know the value of the function γ at one point (recall (2.7)) will determine γ over its entire domain. Compiani et al. (2018) provide sufficient conditions— roughly, that U_t and Z_t act as continuous substitutes in the 'production' of bidder entry, and that Z_t have variation sufficient to offset certain discrete variation in U_t. They also discuss the partial identification of γ obtained when the instrument induces more limited variation (even no variation) in bidder entry.

With γ known, identification of $F_w(\cdot | n, c)$ follows easily. Fix $X_t = x$, $N_t = n$ and recall (5.2). The random variable $\gamma(x, U_t)$ is independent of $\left(\ln B_{1t}^0, \ldots, \ln B_{nt}^0\right)$ and has (now) a known distribution. Thus, because the joint distribution of $(\ln B_{1t}, \ldots, \ln B_{nt})$ is observed, a standard deconvolution result implies identification of the joint distribution of the log homogenized bids $\left(\ln B_{1t}^0, \ldots, \ln B_{nt}^0\right)$. Since homogenized pivotal expected values must satisfy the first-order condition (2.10), their joint distribution is identified. Because we know the function γ, this also implies identification of the joint distribution $F_w(\cdot | n, c)$ for all n and c.

The approach of Compiani et al. (2018) shares with the control function approach of Haile et al. (2003) the potential disadvantage of relying on a reduced form for bidder entry outcomes. And like the strategy of Krasnokutskaya (2011) (and the finite mixture approach discussed in Section 6), it relies on a separability requirement limiting the way auction characteristics alter the environment. On the other hand, the approach deals explicitly with the endogeneity of bidder entry, requires no further restriction on the baseline auction model, and provides a strategy for isolating the exogenous variation in competition induced by the instrument.

6. A MIXTURE MODEL APPROACH

Mixtures have been widely used in order to incorporate unobserved heterogeneity in econometric models.[29] Nonparametric identifiability of mixture models is a challenging subject, although some approaches have been suggested in the literature. For example, measurement error models considered by Hu (2008) and D'Haultfoeuille and Fevrier (2015) can be interpreted as mixture models, and, as seen in Section 4, identification results in these models are potentially useful for auctions with unobserved heterogeneity, especially in the IPV model. Moreover, a recent paper by Kitamura and Laage (2018) provides sufficient conditions for identification of a finite mixture regression model. We demonstrate below that their result can be used to establish nonparametric identification of the affiliated values model while treating unobserved heterogeneity flexibly.

Kitamura and Laage (2018) consider a J-components mixture of nonparametric regressions

$$Y = \gamma(X, U) + \epsilon_U, \quad \Pr\{U = u\} = \lambda_u, \quad u = 1, ..., J, \qquad (6.1)$$

where the econometrician observes the outcome variable Y and the covariate X, whereas U and ε_U are unobserved. The functions $\{\gamma(\cdot, u)\}_{u=1}^{J}$, the probability weights $\{\lambda_u(\cdot)\}_{u=1}^{J}$, and the distributions of $\{\epsilon_u\}_{u=1}^{J}$ are unknown. The key assumptions for the identification results in Kitamura and Laage (2018) are: (i) $\epsilon_u \perp\!\!\!\perp X$ for every $u \in \{1, ..., J\}$, and (ii) there exists a segment where $\gamma(x, u)$, $u \in \{1, ..., J\}$ are non-parallel. With regularity conditions in terms of the characteristic functions and/or moment-generating functions of ε_u, $u \in \{1, ..., J\}$, Kitamura and Laage (2018) show that $\{\gamma(\cdot, u)\}_{u=1}^{J}$, $\{\lambda_u\}_{u=1}^{J}$ and the distribution functions F_{ϵ_u}, $u = 1, ..., J$ are all nonparametrically identified.

To apply this result to an auction model with unobserved heterogeneity, the multiplicative structure (2.5) is once again maintained, although this time we impose the following independence condition:

$$(V_{1t}^0, \ldots, V_{N_t t}^0, S_{1t}, \ldots, S_{N_t t}) \perp\!\!\!\perp X_t | N_t, U_t.$$

This assumption relaxes condition (2.6) by avoiding the requirement of independence between homogenized valuations/signals and U_t. Thus, for example, bidders' private information may interact with the unobservable, not just through an index of auction characteristics (X_t, U_t). These structures imply

$$w(S_{it}; N_t, X_t, U_t) = \Gamma(X_t, U_t) w^0(S_{it}; N_t, U_t), \qquad (6.2)$$

where $w^0(s; n, u) = E[V_{it}^0 | S_{it} = \max_{j \neq i} S_{jt} = s, N_t = n, U_t = u]$. As before, $w^0(s; n, u)$ is a homogenized pivotal expected value, although here it is allowed to depend on U_t. On the other hand, we assume that the support of U_t is $\{1, ..., J\}$, so that the finite mixture identification result in Kitamura and Laage (2018) applies.

Like the quasi-control function approach in Section 5 and the measurement error approach of Krasnokutskaya (2011), the mixture approach in this section imposes the multiplicative separability requirement. But note that it differs from the other two in that the separability is required only in terms of X_t, not U_t. The mixture approach, like that of Hu et al. (2013), also assumes a discrete and finite support for unobserved heterogeneity. On the other hand, it applies to general affiliated values models, as do the control function and quasi-control function approaches.

By the separability-preserving property of equilibrium bidding, from (6.2) we obtain

$$B_{it} = \Gamma(X_t, U_t) B_{it}^{0, U_t},$$

---

[29] See Compiani and Kitamura (2016) for a recent review.

where B_{it}^{0,U_t} is bidder i's homogenized bid (i.e., $B_{it}^{0,U_t} = \beta^0(S_{it}; N_t, U_t)$). Here the more flexible index structure is also inherited by equilibrium bids: B_{it}^{0,U_t} can depend on U_t even after homogenization. Taking logs of both sides and letting $b_{it} = \log B_{it}$, $\gamma(X_t, U_t) = \log \Gamma(X_t, U_t)$, and $b_{it}^{0,U_t} = \log B_{it}^{0,U_t}$, we have

$$b_{it} = \gamma(X_t, U_t) + b_{it}^{0,U_t}.$$

Note that our independence assumptions imply

$$(b_{1t}^{0,u}, \ldots, b_{nt}^{0,u}) \perp\!\!\!\perp X_t | N_t.$$

Finally, we assume

$$U_t \perp\!\!\!\perp X_t | N_t. \tag{6.3}$$

Note that this imposes independence of X_t from the unobserved heterogeneity. Thus, for example, this generally rules out entry outcomes N_t that depend on both X_t and U_t. Of course, if we introduce additional covariates (that can be correlated with U_t) and if the independence condition between X_t and U_t conditional on the extra covariates and the non-parallel condition with respect to X_t hold, then identification prevails once again.[30]

In any event, we are now in a position to apply the identification result of Kitamura and Laage (2018). For the remainder of this section we fix N_t at an arbitrary value n and suppress the index n except where it is necessary. Take any vector $c = (c_1, ..., c_n)'$ from $\mathbb{R}^n$ and construct the linear combination of log bids

$$\sum_{i=1}^{n} c_i b_{it} = \left(\sum_{i=1}^{n} c_i \right) \gamma(X_t, U_t) + \sum_{i=1}^{n} c_i b_{it}^{0,U_t}.$$

Rewrite this as

$$\tilde{b}_t^c = C\gamma(X_t, U_t) + \tilde{b}_t^{c,0,U_t}, \tag{6.4}$$

with $\tilde{b}_t^c \equiv \sum_{i=1}^{n} c_i b_{it}$, $C \equiv \sum_{i=1}^{n} c_i$, and $\tilde{b}_t^{c,0,U_t} \equiv \sum_{i=1}^{n} c_i b_{it}^{0,U_t}$. Note that

$$\tilde{b}_t^{c,0,u} \perp\!\!\!\perp X_t \tag{6.5}$$

holds for each $u \in \{1, ..., J\}$. Define

$$\lambda_u = \Pr\{U_t = u\}, \quad u \in \{1, ..., J\}. \tag{6.6}$$

Applying the result by Kitamura and Laage (2018) outlined above to (6.4), (6.5), and (6.6), we see that $C\gamma(\cdot, u)$, λ_u, and the distribution of $\tilde{b}_t^{c,0,u}$ are all identified for every $c \in \mathbb{R}^n$ and each $u \in \{1, ..., J\}$. But C is known, so $\gamma(\cdot, \cdot)$ is identified. As $c \in \mathbb{R}^n$ is arbitrary, the (marginal) distribution of every linear combination $\tilde{b}_t^{c,0,u}$ of $(\ln B_{1t}^{0,u}, ..., \ln B_{1n}^{0,u})$ is identified. Thus, by the Cramér–Wold device, the joint distribution of $(B_{1t}^{0,u}, ..., B_{nt}^{0,u})$ is identified. Since $\gamma(\cdot, \cdot)$ is already known, this implies the identification of the joint distribution of $(B_{1t}, ..., B_{nt})|N_t = n, X_t = x, U_t = u$. From the first-order condition for equilibrium bidding, the joint distribution of $(w(S_{1t}; n, x, u), ..., w(S_{nt}, ; n, x, u))$ is then uniquely determined.

---

[30] Moreover, Kitamura and Laage (2018) show that nonparametric identification can be obtained in the model (6.1) even without the independence restriction (6.3), at least for the case $J = 2$.

7. CONCLUSION

We have discussed several strategies for allowing unobserved heterogeneity in empirical models of first-price auctions without sacrificing nonparametric identifiability. Some of these methods are now well established in the empirical literature, while others have been developed only very recently. In all cases, some combination of new structure or new data must be added to the minimal setting in which one observes only bids and covariates, and where no structure is placed on the model beyond those needed to characterize equilibrium behaviour. The results we have discussed offer a range of alternatives, all of which build on insights developed for other types of econometric models.

Each of the identification strategies we have discussed offers advantages and disadvantages relative to others. And, while we have focused exclusively on identification, nonparametric/semiparametric estimators based on these identification results introduce additional trade-offs, suggesting that the most suitable approach in practice will vary with the application. We view these nonparametric identification results as relevant for empirical work employing parametric models as well. Even when practical concerns dictate the use of parametric assumptions for estimation, it is valuable to understand whether such assumptions are essential maintained hypotheses or merely choices of the finite-sample approximation method; without such an understanding, we cannot be precise about the foundation on which we build knowledge from data. Here, the availability of several alternative sufficient conditions for nonparametric identification suggests a form of robust identification that should encourage the use of models incorporating unobserved heterogeneity in practice. Indeed, applications using estimators based on these identification strategies so far indicate that accounting for unobserved heterogeneity can be important for the policy conclusions one reaches.[31]

REFERENCES

An, Y., Y. Huand and M. Shum (2010). Estimating first-price auctions with an unknown number of bidders: A misclassification approach. *Journal of Econometrics 157*, 328–41.

Athey, S.and and P. A. Haile (2002). Identification of standard auction models. *Econometrica 70*, 2107–40.

Athey, S.and and P. A. Haile (2006). Empirical models of auctions. In Blundell, R., W. Newey and T. Persson (Eds.), *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, *Volume 2*. Cambridge University Press, Cambridge.

Athey, S. and P. A. Haile (2007). Nonparametric approaches to auctions. In Heckman, J. J. and E. Leamer (Eds.), *Handbook of Econometrics*, *Volume 6A*, 3847–965. Elsevier, Amsterdam.

---

[31] See, e.g., Haile (2003), An et al. (2010), Krasnokutskaya (2011), Krasnokutskaya and Seim (2011), Roberts (2013), Athey et al. (2011), Balat (2011), Kong (2017b), and Compiani et al. (2018).

Athey, S., J. Levin and E. Seira (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *Quarterly Journal of Economics 126*, 207–57.

Balat, J. (2011). Highway procurement and the stimulus package: Identification and estimation of dynamic auctions with unobserved heterogeneity. Working paper, Yale University, New Haven, CT.

Bhattacharya, V., J. Roberts and A. Sweeting (2014). Regulating bidder participation in auctions. *RAND Journal of Economics 45*, 675–704.

Campo, S., E. Guerre, I. M. Perrigne and Q. Vuong (2011). Semiparametric estimation of first-price auctions with risk averse bidders. *Review of Economic Studies 78*, 112–47.

Campo, S., I. M. Perrigne and Q. Vuong (2003). Asymmetry in first-price auctions with affiliated private values. *Journal of Applied Econometrics 18*, 197–207.

Cassola, N., A. Hortaçsu and J. Kastl (2013). The 2007 subprime market crisis through the lens of European Central Bank auctions for shortterm funds. *Econometrica 81*, 1309–45.

Chesher, A. (2003). Identification in nonseparable models. *Econometrica 71*, 1405–41.

Chesher, A. (2005). Nonparametric identification under discrete variation. *Econometrica 73*, 1525–50.

Compiani, G., P. Haile and M. Sant'Anna (2018). Common values, unobserverd heterogeneity, and endogenous entry in U.S. offshore oil lease auctions. Technical report, Yale University, New Haven, CT, Cowles Foundation CFDP No. 2137.

Compiani, G. and Y. Kitamura (2016). Using mixtures in econometric models: A brief review and some new results. *Econometrics Journal 19*, C95–C127.

D'Haultfoeuille, X. and P. Fevrier (2015). Identification of mixture models using support variation. *Journal of Econometrics 189*, 70–82.

Evdokimov, K. and H. White (2012). Some extensions of a lemma of Kotlarski. *Econometric Theory 28*, 925–32.

Gentry, M. and T. Li (2014). Identification in auctions with selective entry. *Econometrica 82*, 314–44.

Gillen, B. (2010). Identification and estimation of level-k auctions. Technical report, Caltech, Pasadena, California.

Gilley, O. W. and G. V. Karels (1981). The competitive effect in bonus bidding: new evidence. *Bell Journal of Economics 12*, 637–48.

Guerre, E., I. M. Perrigne and Q. Vuong (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica 68*, 525–74.

Guerre, E., I. M. Perrigne and Q. Vuong (2009). Nonparametic identification of risk aversion in first-price auctions under exclusion restrictions. *Econometrica 77*, 1193–227.

Haile, P. A. (2001). Auctions with resale markets: An application to U.S. forest service timber sales. *American Economic Review 91*(3), 399–427.

Haile, P. A., H. Hong and M. Shum (2003). Nonparametric tests for common values in first-price sealed-bid auctions, New Haven, CT. Cowles Foundation Discussion Paper 1444.

Hendricks, K. and H. J. Paarsch (1995). A survey of recent empirical work concerning auctions. *Canadian Journal of Economics 28*, 403–26.

Hendricks, K., J. Pinkse and R. H. Porter (2003). Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions. *Review of Economic Studies 70*, 115–45.

Hendricks, K. and R. H. Porter (2007). An empirical perspective on auctions. In Armstrong, M. and R. Porter (Eds.), *Handbook of Industrial Organization*, *Volume 3*, 2073–2143, Elsevier, Amsterdam.

Hickman, B. R., T. P. Hubbard and Y. Saglam (2012). Stuctural econometric methods in auctions: A guide to the literature. *Journal of Econometric Methods 1*, 67–106.

Hortaçsu, A., J. Kastl and A. Zhang (2018). Bid shading and bidder surplus in the US treasury auction system. *American Economic Review 108*(1),147–69.

Hortaçsu, A. and D. McAdams (2010). Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish treasury auction market. *Journal of Political Economy 188*, 833–65.

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 (2013). Identification of first-price auctions with non-separable unobserved heterogeneity. *Journal of Econometrics 174*, 186–93.

Imbens, G. W. and W. K. Newey (2009). Identification and estimation of triangular simultaneous equations models without additivity. *Econometrica 77*, 1481–512.

Jofre-Bonet, M. and M. Pesendorfer (2003). Estimation of a dynamic auction game. *Econometrica 71*, 1443–89.

Kastl, J. (2011). Discrete bids and empirical inference in divisible good auctions. *Review of Economic Studies 78*, 974–1014.

Kitamura, Y. and L. Laage (2018). Nonparametric analysis of finite mixtures. Technical report, Yale University, New Haven, CT.

Kong, Y. (2017a). Selective entry in auctions: Estimation and evidence. Technical report, Rice University, Houston.

Kong, Y. (2017b). Sequential auctions with synergy and affiliation across auctions. Technical report, Rice University, Houston.

Kotlarski, I. (1967). On characterizing the gamma and the normal distribution. *Pacific Journal of Mathematics 20*, 69–76.

Krasnokutskaya, E. (2011). Identification and estimation in procurement auctions under unobserved auction heterogeneity. *Review of Economic Studies 28*, 293–327.

Krasnokutskaya, E. and K. Seim (2011). Bid preference programs and participation in highway procurement auctions. *American Economic Review 101*, 2653–86.

Laffont, J.-J. and Q. Vuong (1993). Structural econometric analysis of descending auctions. *European Economic Review 37*, 329–341.

Laffont, J.-J. and Q. Vuong (1996). Structural analysis of auction data. *American Economic Review, Papers and Proceedings 86*, 414–20.

Li, T., I. Perrigne and Q. Vuong (2000). Conditionally independent private information in ocs wildcat auctions. *Journal of Econometrics 98*, 129–61.

Li, T., I. Perrigne and Q. Vuong (2002). Structural estimation of the affiliated private value auction model. *RAND Journal of Economics 33*, 171–93.

Luo, Y. (2018). Unobserved heterogeneity in auctions under restricted stochastic dominance. Technical report, Univerisity of Toronto, Department of Econonomics, Working Paper No. 606.

Manski, C. F. and E. Tamer (2002). Inference on regressions with interval data on a regressor or outcome. *Econometrica 70*, 519–46.

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.

Matzkin, R. L. (2003). Nonparametric estimation of nonadditive random functions. *Econometrica 71*, 1339–75.

Milgrom, P. R. and R. J. Weber (1982). A theory of auctions and competitive bidding. *Econometrica 50*, 1089–1122.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*, 1263–97.

Paarsch, H. J. and H. Hong (2006). *An Introduction to the Structural Econometrics of Auction Data*. Cambridge, MA: MIT Press.

Roberts, J. W. (2013). Unobserved heterogeneity and reserve prices in auctions. *RAND Journal of Economics 44*, 712–32.

Shneyerov, A. (2006). An empirical study of auction revenue rankings: The case of municipal bonds. *RAND Journal of Economics 37*, 1005–22.

Somaini, P. (2018). Identification in auction models with interdependent costs. Technical report, Stanford University.

Managing editor Jaap Abbring handled this manuscript.