Journal of Econometrics ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach☆

Yonghong An

Department of Economics, Texas A&M University, 3051 Allen Building, College Station, TX 77843, United States

ARTICLE INFO

Article history:
Available online xxxx

JEL classification:
C14
D44

Keywords:
First-price auctions
Measurement error models
Non-equilibrium beliefs
Asymmetric value distributions
Nonparametric identification and estimation

ABSTRACT

This paper studies identification and estimation of two models for first-price auctions: (1) bidders' beliefs about their opponents' bidding behavior are not in equilibrium but follow "level-k" thinking, and (2) bidders' values are asymmetrically distributed. Exploiting the nonparametric methodology developed for measurement error models (e.g., Hu, 2008), we show that both models can be identified by a unified methodology. The proposed methodology is applied to US Forest Service timber auction data and the estimation results suggest that bidders hold heterogeneous and non-equilibrium beliefs.

1. Introduction

Recent empirical studies on auctions have documented asymmetry among bidders, e.g., Athey et al. (2011) have established that in timber auctions there are asymmetric bidders who draw values from different distributions. Experimental studies of first-price auctions (e.g., see Crawford and Iriberri, 2007) demonstrate that bidders may hold non-equilibrium and asymmetric beliefs about their opponents' bidding strategies. The *ex ante* asymmetry of bidders lead to deviation of their bidding strategies from those in standard symmetric auctions, and the deviation has been employed to explain bidders' overbidding behavior in experiments. In presence of bidders' asymmetry, standard methods in auctions, e.g., Guerre et al. (2000) do not apply to robust analyses of the data. However, little work has been done to rigorously identify and estimate first-price auctions with asymmetric bidders, especially when they hold non-equilibrium and heterogeneous beliefs.

This paper shows that the recent developments in the econometric literature on nonlinear measurement error (e.g., Hu, 2008; Hu and Schennach, 2008) can be applied to identify first-price auctions with bidders holding non-equilibrium and heterogeneous beliefs about their opponents or having asymmetric value distributions. In the former auction model, bidders' beliefs follow a "level-k" thinking model (level-k auctions a.k.a. Crawford and Iriberri, 2007).[1] Bidders in level-k auctions hold heterogeneous beliefs (each belief is called a "type"), which have a hierarchy structure based on a nonstrategic type-$L0$, about their opponents' bidding strategies, and they bid to maximize their payoff by best responding to their own beliefs. Then heterogeneous bidding strategies display due to the heterogeneous beliefs across bidders.

The two models are first proved to be fitted into a measurement error model as in Hu (2008), then a unified methodology is proposed to nonparametrically identify both models. From the observed bids, we identify (1) the number of possible types among the bidders, what the types are (including Nash-type), and the bidding strategy of each individual type, (2) the proportion of each type of bidders, (3) the distribution of bidders' private values, and in level-k auctions (4) the specification of nonstrategic type-$L0$ on which bidders' recursive reasoning is based. The main assumption of identification is that bids from independent auctions contested by bidders with their identity being observed, and each bidder has three bids from three independent auctions. The results of identification provide a testable implication of level-k auctions against the model with asymmetric value distributions: the value

☆ This paper is a substantially revised version of the first chapter of my dissertation "Nonparametric Identification and Estimation of Level-k Auctions". I am deeply indebted to Yingyao Hu for his guidance and encouragement on this project. I benefited greatly from the comments of Edi Karni, Matthew Shum and Richard Spady. I thank Arie Beresteanu, Pedro Carneiro, Ben Gillen, Matt Jackson, Przemek Jeziorski, Elena Krasnokutskaya, Robert Moffitt, Mitsukuni Nishida, and Tiemen Woutersen for their helpful comments. I am grateful to Ben Gillen for explaining the USFS timber auction data. The usual disclaimer applies.

E-mail address: y.an@taum.edu.

[1] The "level-k" thinking was introduced by Stahl and Wilson (1994) and Nagel (1995) and obtained strong experimental support. Using experimental data, Crawford and Iriberri (2007) show that level-k auctions provide a unified explanation of the winner's curse in common-value auctions and overbidding in independent-private-value (IPV) auctions.

http://dx.doi.org/10.1016/j.jeconom.2017.06.014
0304-4076/

distributions for all types of bidders are identical if data are generated by level-k auctions. The methodology is applied to the United States Forest Service (USFS) timber auction data. The empirical results show that bidders display three-type non-equilibrium beliefs which can be explained by the level-k auction model. Moreover, we find strong support of the level-k thinking from the data by testing the implications of the model.

The main contribution of this paper is to illustrate that models of nonclassical measurement errors can be applied to analyze first-price auctions with bidders holding non-equilibrium and heterogeneous beliefs. The first-price auctions fit into a model of misclassification error where bidders' unobserved type is the latent true value and the observed bids are measurements of the true value. Thus, the developments in measurement errors have been applied to identify first-price auctions with bidders' heterogeneity, e.g., Li and Vuong (1998), Li et al. (2000), An et al. (2010) and Krasnokutskaya (2011). More recently, Hu et al. (2013) propose a general result of identification for first-price auctions with non-separable unobserved heterogeneity, which can be asymmetric value distributions we described above. Nevertheless, their approach does not apply to level-k auctions, where the two essential assumptions for identification in Hu et al. (2013) violate: First, the heterogeneity is unobserved to the econometrician but common knowledge among bidders. In level-k auctions, the unobserved heterogeneity is bidders' thinking level, which is private information to both other bidders and the econometrician. Second, bidders' value distribution is monotone in the unobserved heterogeneity, in the sense of first-order stochastic dominance (FOSD). In level-k auctions, bidders' values are assumed to be independent of their thinking levels.

Our nonparametric analysis of models with non-equilibrium behavior is novel in the literature. The existing literature on such models mainly focuses on experimental data with underlying value distribution being known to the econometrician.[2] By contrast, we identify the model's main components nonparametrically. Gillen (2009) also considers the identification and estimation of level-k auctions using a standard approach as in Guerre et al. (2000), and his results rely on the assumption that the number of the unobserved types, what the types are, and specifications of type-$L0$ to are all known to researchers, whereas all these objectives are nonparametrically identified and estimated in our paper. In an online appendix to our paper (An, 2015), we provide an additional comparison with Gillen (2009).

The empirical evidence in this paper demonstrates that non-equilibrium and heterogeneous beliefs of bidders on their opponents' bidding behavior could be an important source of asymmetry in auctions. Empirical results in Athey et al. (2011) show that bidders may be asymmetric in value distributions, while no existing evidence on the asymmetric beliefs of bidders from field data. Since our results show auctions with asymmetric value distributions and level-k auctions can be identified by a unified approach, the empirical results in this paper open a window for further research on comparison or testing different asymmetric auction models. Furthermore, the measurement error approach in this paper can be used to analyze various behavioral models that incorporate agents' non-equilibrium behavior, e.g., cognitive hierarchy models in Camerer et al. (2004) which also fit into the framework of Hu (2008).

The remainder of the paper proceeds as follows. Section 2 presents the models of first-price auctions with level-k thinking and asymmetric value distributions within an independent private value (IPV) paradigm. Section 3 provides nonparametric identification of both models. Section 4 conducts a Monte Carlo study of level-k auctions and addresses some practical implementation issues of the methodology. Section 5 applies the methodology to USFS data and tests the validity of level-k auction models. Section 6 concludes. The Appendix contains proofs. A comparison with Gillen (2009) as well as some technical details omitted in the paper are provided as supporting information in an online Appendix.

2. The model

In this section, we present the model of level-k auctions following Crawford and Iriberri (2007) as well as the model with asymmetric value distributions. Then provide simple theoretical analyses of the models. The auctions are assumed to be IPV first-price sealed-bid ones with non-binding reserve price.

2.1. The model of level-k auctions

A single and indivisible object is sold through a first-price sealed-bid auction among $I(\geq 2)$ risk-neutral bidders. The private information of each bidder consists of two independent components: value v and level of sophistication τ ("type" hereafter). The bidders' private values v are distributed according to a commonly known cumulative distribution function $F(\cdot)$, which is absolutely continuous on $[\underline{v}, \overline{v}] \subset \mathbb{R}_+$ with a strictly positive density $f(\cdot)$.

Each bidder is randomly assigned a type by nature, and it is invariant across all the auctions, i.e., the possibility that bidders improve their types by learning is assumed away. Bidders of different types bid heterogeneously: the bidders of lowest type (denoted by $L0$) bid non-strategically without considering their opponents, higher-type bidders believe that their opponents are all *one* type lower than they are, and they best respond to their opponents accordingly.

The iterated bidding behavior of bidders implies that beliefs of all higher-type bidders are recursively based on the assumption of lowest type-$L0$. Consequently, different specifications of bidding strategy for the lowest type characterize different bidding structure of bidders in level-k auctions. To capture the description above, we formulate the type τ as a random variable, $\tau \in \mathcal{K} \equiv \{Lk : k = 1, 2, \ldots, K, \infty\}$, where k is the level in the hierarchical structures. Since a level-k model does not necessarily converge to Bayesian-Nash equilibrium (BNE) as $k \to \infty$ (this is a similar result to Battigalli and Siniscalchi, 2003), we may include a type $L\infty$ indicating the type of bidders who play BNE. Considering that bidders in field are more experienced, we exclude type-$L0$ in this paper and assume bidders of type-$L0$ only exist in higher-type bidders' minds. In the literature on the level-k auctions, the set $\mathcal{K}$ including the specification of type $L0$ is assumed to be known to the econometrician. In this paper, however, the number of types in the set $\mathcal{K}$ as well as the specification of type-$L0$ are both identified from data.

A bidder i who is of type Lk believes that all other $(I-1)$ bidders are of type Lj, $j = k-1$. Given her value v_i, bidder i bids by maximizing the expected utility $(v_i - b_i)\Pr(\max_{l\neq i} b_l \leq b_i)$. The belief of this bidder (type) is associated with the winning probability as follows.

$$\Pr(\max_{l\neq i} b_l \leq b_i) = \Pr(s_{Lj}(\max_{l\neq i} v_l) \leq b_i) = F^{I-1}\left(s_{Lj}^{-1}(b_i)\right). \tag{1}$$

Consequently, a standard approach solves the optimal bidding strategy for a type-Lk bidder as.[3]

$$b_i = s_{Lk}(v_i, I, F) = v_i - \frac{1}{I-1}\frac{F\left(s_{Lj}^{-1}(b_i)\right)}{f\left(s_{Lj}^{-1}(b_i)\right)}\left(\frac{ds_{Lj}^{-1}(t)}{dt}|_{t=b_i}\right)^{-1}. \tag{2}$$

[2] For example, see Costa-Gomes et al. (2001), Camerer et al. (2004), Costa-Gomes and Crawford (2006), Georganas (2011), Georganas et al. (2015), and Ivanov et al. (2010).

[3] For a bidder of any type, if her value is $\underline{v}$ then her winning probability is zero. In this case we assume she bids $\underline{v}$, i.e., $s_{Lk}(\underline{v}) = \underline{v}$ for all $Lk \in \mathcal{K}$.

Recall that in a symmetric BNE for a first-price IPV auction, a bidder's optimal bidding strategy is determined by the following differential equation:

$$s(v, I, F) = v - \frac{1}{I-1}\frac{F(v)}{f(v)}s'(v). \tag{3}$$

Comparison between the two optimal bidding strategies (2) and (3) illustrates how the heterogeneous beliefs of level-k bidders affect their bidding strategies to deviate from that of Bayesian-Nash bidders. Also, the heterogeneous bidding strategies imply the support for bidding strategy $s_{Lk}(\cdot)$ (for ease of notation, we suppress the arguments I and F) may be different for bidders of different types. The result is summarized in Lemma 1.

Lemma 1. *Let $s_{Lk}(\cdot)$ and $s_{Lm}(\cdot)$ denote bidding strategies for bidders of types Lk and Lm, respectively, where $Lk, Lm \in \mathcal{K}$. If both bidding strategies are strictly increasing, then $k > m \geq 1$ implies $s_{Lk}(\bar{v}) \leq s_{Lm}(\bar{v})$.*

The lemma above, together with the relationship $s_{L1}(\underline{v}) = \cdots = s_{Lk}(\underline{v}) = \underline{v}$, implies that the support of the bids is monotonically decreasing in types. Battigalli and Siniscalchi (2003) and Shneyerov and Wong (2011) discuss similar properties in first-price auctions.

Both the bidding strategies in (2) and Lemma 1 rely on a regularity condition that a bidding strategy $s_{Lk}(v)$ is strictly increasing in value v for any type Lk. To explore the restrictions the regularity condition imposes on the data, we take derivative of (2) with respect to v, then the monotonicity of $s_{Lk}(\cdot)$ requires:

$$\frac{db}{dv} = \left(1 + \frac{1}{I-1}\frac{d}{dv}\frac{F\left(s_{Lj}^{-1}(b)\right)}{f\left(s_{Lj}^{-1}(b)\right)\frac{ds_{Lj}^{-1}(b)}{dv}}\right)^{-1} > 0,$$
$$b \equiv s_{Lk}(v) \in [\underline{v}, \bar{v}]. \tag{4}$$

This regularity condition guarantees the existence of the strictly increasing bidding function $s_{Lk}(\cdot)$. A sufficient condition for (4) is:

Regularity Condition (RC). *For any private value $v \in [\underline{v}, \bar{v}]$, the distribution function of value $F(\cdot)$ satisfies*

$$1 + \frac{d}{dv}\frac{F\left(s_{Lk}^{-1}(v)\right)}{f\left(s_{Lk}^{-1}(v)\right)\frac{ds_{Lk}^{-1}(t)}{dt}|_{t=v}} > 0, \quad \text{for any} \quad Lk \in \mathcal{K}, v \in [\underline{v}, \bar{v}]. \tag{5}$$

This condition implies that the term of derivative is greater than -1 and it is easy to show that (5) suffices (4) given $I \geq 2$. Note that the regularity condition is imposed for all $Lk \in \mathcal{K}$. In fact, for some type Lk, the condition may impose no restriction on the value distribution $F(\cdot)$. For example, for the *random*-type bidders discussed in Crawford and Iriberri (2007), the regularity condition above imposes no restriction to bidders of random type-$L1$ because their bidding strategy does not depend on $F(\cdot)$. When valuations are log-normally distributed, our Monte Carlo experiments demonstrate that this regularity condition is satisfied.

2.2. First-price auctions with asymmetric value distributions

In first-price sealed-bid auctions, bidders may draw their values from different distributions. Such an asymmetry can be due to collusion and heterogeneity across bidders and has been documented in some empirical studies, e.g., Athey et al. (2011).[4] We present a model of first-price auctions with asymmetric value distributions (AVD hereafter) in this section and show that it is observably equivalent to level-k auctions in the next section.

Consider a first-price sealed-bid auction with $I(\geq 2)$ risk-neutral bidders whose private values $v_1, v_2, \ldots, v_I$, are randomly drawn from $F_1(\cdot), F_2(\cdot), \ldots, F_K(\cdot)$ $(K \leq I)$, which are CDFs with the same support $[\underline{v}, \bar{v}]$. The distributions are differentiable over $(\underline{v}, \bar{v}]$ and the density functions $f_1(\cdot), f_2(\cdot), \ldots, f_K(\cdot)$ are continuous and strictly positive over $(\underline{v}, \bar{v}]$. In addition to valuation, each bidder is one of the K $(K \leq I)$ types with type $k \in \mathcal{K} = \{1, 2, \ldots, K\}$ indicating bidders' values are distributed according to $F_k(\cdot)$. Let τ_i denote the type of ith bidder and $\tau_i \in \{1, 2, \ldots, K\}$, then $\tau_i = k$ means bidder i's value is a random draw from $F_k(\cdot)$. Bidder i's bidding strategy s_i is a mapping from her value and type to her bid, i.e., $s_{\tau_i}(\cdot, \cdot) : [\underline{v}, \bar{v}] \times \{1, 2, \ldots, K\} \to [\underline{v}, \bar{v}]$. In this setting, value distributions of bidders are common knowledge, whereas a bidder's value is private, i.e., τ_i is known to all the bidders and v_i is private information. Both type and value are unknown to the econometrician.

Lebrun (1999) proves the existence and uniqueness of a Bayesian Nash equilibrium (BNE) if all the K distributions have a mass point at $\underline{v}$. At this equilibrium, any two bidders who draw their values from the same distribution have the same bidding strategy. Moreover, if there exists a relation of stochastic dominance between two valuation distributions, the same relation of stochastic dominance extends to the bid distributions, i.e., $F_j(v) \leq F_i(v)$ for all $v \in [\underline{v}, \bar{v}]$ implies $F_j(s_j^{-1}(b)) \leq F_i(s_i^{-1}(b))$ for all $b \in [\underline{v}, \eta]$ with $\eta = s_1(\bar{v}) = \cdots = s_K(\bar{v})$.

Assumption 1. There exists a relation of stochastic dominance among the K value distributions $F_1(\cdot), \ldots, F_K(\cdot)$, i.e., $F_1(\cdot) \leq \cdots \leq F_K(\cdot)$. Moreover, the K distributions share the common support $[\underline{v}, \bar{v}]$ and have a mass point as $\underline{v}$.

3. Nonparametric identification

3.1. Data structure and econometric models

Suppose there are N bidders (N is the total number of bidders in a dataset, which is different from I, the number of bidders in an auction) participating in some auctions with homogeneous objects and the number of bidders in each auction does not vary. It is further assumed that the bidders' identity is observed, and let b_{it} denotes bidder i's bid in the tth auction, where $i \in \{1, 2, \ldots, N\}$, $t \in \{1, 2, \ldots, \bar{t}\}$ and $\bar{t} \geq 3$. In the population there are m types, where $2 \leq m \leq N < \infty$ and it is unknown to the econometrician.[5] The unknown probability of type-τ bidders is $p(\tau) > 0$ with $\sum_{\tau \in \mathcal{K}} p(\tau) = 1$. In level-$k$ auctions, each type of bidders employ a certain bidding strategy while in AVD each type corresponds to a value distribution. It is further assumed that in level-k auctions, the types are *consecutive* and starting from type $L1$ (our identification argument still holds if the type starts from Lk, $k \geq 2$. We consider this case later in this paper). The data structure is summarized as one of the conditions required for identification as follows.

Assumption 2 (*Data Structure*)**.** The econometrician observes bidders' identity and an i.i.d sample of bidders' bids. The sample contains three bids from three independent auctions for each bidder. The bids of bidder i $(= 1, 2, \ldots, N)$ are indexed by b_{it}, where $t = 1, 2, 3$ denotes the t-th auction bidder i participates in.

Assumption 2 introduces the panel data structure, which is essential for our identification. It states that there is no strategic dependence between the same bidder's actions across auctions.

[4] The asymmetry in Athey et al. (2011) is indicated by the characteristics of bidders, which is known to the researcher *ex ante*, whereas in our model the asymmetry is unknown to the researcher.

[5] There are only two cases where $m = 1$: (1) All bidders play a symmetric Bayesian-Nash equilibrium; (2) All bidders employ the same but non-equilibrium bidding strategy. We rule out these two uninteresting cases since bidders are symmetric.

Please cite this article in press as: An, Y., Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.06.014.

Thus dynamic optimization behaviors based on learning or budget constraints are ruled out. As will be shown, one of the bids acts as an instrumental variable for bidders' type. Therefore, two bids for each bidder are enough for identification if suitable instrumental variables for bidders' types are available.

In an AVD a bidder with her values being drawn from the same distribution across auctions may employ different bidding strategies if the type distribution of her opponents are different across auctions. For the purpose of identification for AVD, we assume in addition to Assumption 2 that if the data are generated from AVD, the type distribution is the same across auctions. For example, all the observed auctions have three bidders, the values of two of them are drawn from the first distribution, whereas the remaining one is from the second distribution. Such a restriction may be relaxed and our identification argument still holds if we only require that in each auction there are at least two types of bidders. Nevertheless, we main the assumption of the same type distribution across auctions for ease of exposition. It is also worth noting that under such assumption, the identity of bidder is not necessary for identifying AVD.

Given the data structure described above, the joint distribution of bids $G(b_1, b_2, b_3)$ can be directly identified. Our problem of identification is to uniquely determine the distributions of τ, $p(\tau)$ and of v, $F(\cdot)$ from $G(b_1, b_2, b_3)$, where $F(\cdot) \equiv (F_1(\cdot), F_2(\cdot), \ldots, F_K(\cdot))$ in AVD. The novelty of this paper is to show how a panel structure of data enables identification. To rationalize the requirement of the panel structure, suppose only one bid for each bidder is available, then the marginal distribution of bid $G(b)$ is identified from the data. The value distribution $F(\cdot)$ and probability of type $p(\cdot)$ are associated with $G(b)$ in both models as follows.

$$G(b) = \sum_{\tau \in \mathcal{K}} G(b|\tau)p(\tau) = \begin{cases} \sum_{\tau \in \mathcal{K}} F\left(s_\tau^{-1}(b)\right)p(\tau), & \text{level-}k \text{ auctions} \\ \sum_{\tau \in \mathcal{K}} F_\tau\left(s_\tau^{-1}(b)\right)p(\tau), & \text{AVD} \end{cases} \tag{6}$$

where $s_\tau(\cdot)$ is the bidding strategy of type τ and it may be different in two models. This equation introduces a finite mixture where observing $G(b)$ does not allow us to identify the components $F(\cdot)$ and the weights $p(\cdot)$. Even we have a panel data described in Assumption 2, the identification of $F(\cdot)$ for AVD requires a new method and the existing approaches such as Athey and Haile (2002) do not apply here. The main idea of identification in Athey and Haile (2002) is a bidder with observed identity participates in a large number of auctions such that her bid distribution can be recovered directly. The requirement on data imposed by this approach is unlikely to be satisfied. Our data structure in Assumption 2 only requires three bids for each bidder, which is insufficient for identification using Athey and Haile (2002).

As shown in (6), both level-k auctions and AVD are described by the finite mixture model and they are indistinguishable by only observing the joint distribution of bids $G(b_1, b_2, b_3)$. In fact, we show that the components and weights of both mixture models can be nonparametrically identified by a unified approach. Once the components, i.e., the bid distributions for different types $G(b|\tau)$ are identified, the value distribution $F(\cdot)$ for level-k auction model can be nonparametrically recovered by exploring the relationship of two consecutive types. The value distribution $F(\cdot)$ can be estimated for any of the $K(\geq 2)$ types and this provides with us a testable implication of level-k auctions against AVD since value distributions in AVD are distinct for any two types.

In the rest of this section, we propose a unified approach to identify both models. For ease of exposition, we focus on the identification of level-k auctions while conducting necessary discussions on AVD whenever necessary. The basic idea of the identifying strategy is to treat the type of bidders as unobserved heterogeneity, and to apply the results of Hu (2008) on misclassification. The procedure of identification is two-step: the first step identifies bid distributions for the m types by a matrix eigenvalue–eigenvector decomposition, and the second step identifies the value and type distributions based on the implications of the theoretical model for level-k auctions, as well as the identified bid distributions from the first step.

3.2. Identification

Let b_{i1}, b_{i2}, b_{i3} denote the randomly drawn three bids from the bids of bidder i, $i \in \{1, 2, \ldots, N\}$ and $g(b_{i1}, b_{i2}, b_{i3})$ be the joint density of the three bids. For ease of notation, we suppress the index of bidders and denote the three bids b_{i1}, b_{i2}, b_{i3} by b_1, b_2, b_3, respectively.

Due to the law of total probability, the joint density $g(b_1, b_2, b_3)$ can be expressed as[6]

$$\begin{aligned} g(b_1, b_2, b_3) &= \sum_{\tau \in \mathcal{K}} g(b_1, b_2, b_3, \tau) \\ &= \sum_{\tau \in \mathcal{K}} g(b_1|\tau, b_2, b_3)\, g(b_2|\tau, b_3)\, g(\tau, b_3). \end{aligned} \tag{7}$$

An instrumental assumption is conditional independence of bids on the type τ:

Assumption 3. A bidder's values in the three auctions are independent and the type is invariant.

In some widely used dataset, e.g., USFS timber auctions, it is common that a bidder is observed to participate several auctions. Most of the empirical studies on the timber auctions (e.g., see Baldwin et al., 1997; Haile, 2001; Campo et al., 2011) treat the data as a cross-sectional one, which imposes the independence of each bidder's values in different auctions. We may allow the bidders' values in different auctions to be correlated through some observed heterogeneity because in this case we can employ the procedure in Haile et al. (2003) to control the heterogeneity such that the "residual values" are independent. The type of a bidder may depend on her experience, education, etc. as shown in Goldfarb and Xiao (2011), thus it is less likely changing in the three auctions. The invariant type is also supported by some existing empirical evidence e.g., Ivanov et al. (2010) show that it is difficult for subjects to learn to act at a higher level of sophistication.

Assumption 3 does rule out serial correlation of the same bidder's type across auctions. In situations where this is a concern due to evolution of rationality levels (e.g., bidders may learn to adjust their levels of thinking across auctions), we may still achieve identification based on the result in Hu (2008) by imposing some restrictions to the learning process of bidders. An et al. (2015) consider such a setting in the framework of private provision of public goods. Nevertheless, we maintain Assumption 3 as a first-order approximation/simplification of the data-generating process in the empirical application to USFS timber auctions in Section 5.

[6] The boundedness of $g(b_1, b_2, b_3, \tau)$ is due to the finite support of the value density $f(v)$ and the validity of the regularity condition: For a given type $\tau \in \mathcal{K}$, the conditional density $g(b|\tau)$ can be established as

$$g(b|\tau) = f(s_\tau^{-1}(b))\frac{ds_\tau^{-1}(b)}{db}.$$

The first term on the right-hand side is bounded away from zero because of the assumed boundedness of $f(\cdot)$, and the second term is also positive due to the regularity condition. Consider that $g(b, \tau) = g(b|\tau)p(\tau)$, and $p(\tau) > 0$, thus $g(b, \tau) > 0$. Similarly, it is easy to argue that $g(b_1, b_2, b_3, \tau)$ is bounded away from infinity since both the value density $f(v)$ and the derivative of bidding strategy $s'_\tau(\cdot)$ are finite. The similar argument applies to AVD.

Under Assumption 3, $g(b_1|\tau, b_2, b_3) = g(b_1|\tau), g(b_2|\tau, b_3) = g(b_2|\tau)$, then (7) can be further simplified as:

$$g(b_1, b_2, b_3) = \sum_{\tau\in\mathcal{K}} g(b_1|\tau)\, g(b_2|\tau)\, g(\tau, b_3). \tag{8}$$

The continuity of bids enables us to discretize any two of the three bids b_1, b_2, b_3 according to the following method of discretization.

$$d = \begin{cases} 1 & \text{if } b \in [\underline{b}, b(1)], \\ 2 & \text{if } b \in (b(1), b(2)], \\ \dots \\ M & \text{if } b \in (b(M-1), \bar{b}], M \geq 2, \end{cases} \tag{9}$$

where the support of bid, $[\underline{b}, \bar{b}]$ is divided into M intervals by the $M-1$ cutoff points $b(1), b(2), \dots, b(M-1)$, $\underline{b} < b(1) < b(2) < \dots < b(M-1) < \bar{b}$, and $d \in \{1, 2, \dots, M\}$ is the discretized bid. Let us denote d_1 and d_3 the discretized bids of b_1 and b_3, respectively. Both d_1 and d_3 take values from $\{1, 2, \dots, M\}$ but the cutoff points for discretizing b_1 and b_3 can be different. Substituting the bids b_i by the discretized bids d_i $(i = 1, 3)$ in (8) yields

$$g(d_1, b_2, d_3) = \sum_{\tau\in\mathcal{K}} g(d_1|\tau)\, g(b_2|\tau)\, g(\tau, d_3).$$

Because type τ, and the two bids d_1, d_3 are all discrete, it is convenient to express the equation above in matrix form

$$B_{b_2,d_1,d_3} = B_{d_1|\tau} D_{b_2|\tau} B_{\tau,d_3}. \tag{10}$$

The matrices above are defined as follows.

$$\begin{aligned} \left(B_{b_2,d_1,d_3}\right)_{i,j} &\equiv g(b_2, d_1 = i, d_3 = j), \\ \left(B_{d_1|\tau}\right)_{i,k} &\equiv g(d_1 = i|\tau = Lk), \\ \left(B_{\tau,d_3}\right)_{k,j} &\equiv g(\tau = Lk, d_3 = j), \\ \left(B_{d_1,d_3}\right)_{i,j} &\equiv g(d_1 = i, d_3 = j), \\ D_{b_2|\tau} &\equiv \text{diag}\Big[g(b_2|L1) \;\; g(b_2|L2) \;\; \dots \;\; g(b_2|Lm)\Big], \end{aligned}$$

where $g(\cdot)$ denotes a probability density or a probability mass, $L1, L2, \dots, Lm \in \mathcal{K}$ are the m types of the N bidders, and $i, j \in \{1, 2, \dots, M\}$. For a fixed bid b_2, the dimension of the four matrices defined above B_{b_2,d_1,d_3}, $B_{d_1|\tau}$, $D_{b_2|\tau}$, and B_{τ,d_3} are $M \times M$, $M \times m$, $m \times m$, and $m \times M$, respectively. Our first step of identification is to recover the number of possible types m from the matrices defined above, which relies on the following assumption.

Assumption 4. For some known partition of bid support, the discretized bid distribution of any type is not a linear combination of those for the other types.

As will shown in the proof of Lemma 2, Assumption 4 imposes no restriction on the model when there are only two types of bidders. Considering that distribution functions of bids are highly nonlinear in most of the cases, this assumption is unlikely to be violated even when there are more than two types of bidders. Under this assumption, the number of types m is identified as the rank of the matrix B_{d_1,d_3} and the result is presented in Lemma 2.

Lemma 2. *Under Assumptions 2–4, Rank* $\left(B_{d_1,d_3}\right) = m$ *whenever* $M \geq m$.

Proof. See Appendix A. ■

In the literature on level-k behavioral models, it is an important practical issue to choose the number of types m but theoretical models usually do not provide much guidance. Lemma 2, however, provides a plausible approach to choose m from data because this assumption is empirically testable. To identify the number of types m, it is necessary test the rank of the matrix B_{d_1,d_3}. This can be done by using the method proposed in Robin and Smith (2000). The method is based on a sequential test of the null hypotheses H_r : $rank(B_{d_1,d_3}) = r$ against the alternative H_r': $\text{rank}(B_{d_1,d_3}) > r$ for $r = 0, 1, \dots, m-1$. The critical region for each step of testing is obtained from the result that the limiting distribution of the test statistic is a weighted average of χ^2-distributions. This procedure enables us to infer the rank of B_{d_1,d_3} if the significance level of each step depends on the sample size appropriately. Alternatively, we may also infer the rank of B_{d_1,d_3} by testing whether the determinant of B_{d_1,d_3} vanishes for different values of M. Under Lemma 2, the number of types m satisfies $m = \max\{M : det(B_{d_1,d_3}) \neq 0, M = 2, 3, \dots\}$, therefore m is equal to the largest value of M such that the null hypothesis $det(B_{d_1,d_3}) = 0$ is rejected. It is an important empirical issue to choose M and implement the discretization for the inference of the number of types, m, and more details will be addressed in estimation.

The result in Lemma 2 enables us to choose $M = m$ hereafter in this paper, thus the dimension of all the matrices defined in (10) is $m \times m$. Consequently, the three $m \times m$ matrices B_{d_1,d_3}, $B_{d_1|\tau}$, and B_{τ,d_3} are all full rank, and invertible. The invertibility of these matrices allows us to post-multiply both sides of (10) by $B^{-1}_{d_1,d_3} = B^{-1}_{\tau,d_3} B^{-1}_{d_1|\tau}$, and the operation yields the following equation:

$$B_{b_2,d_1,d_3} B^{-1}_{d_1,d_3} = B_{d_1|\tau} D_{b_2|\tau} B^{-1}_{d_1|\tau}. \tag{11}$$

The right-hand side of the equation above represents an eigenvalue–eigenvector decomposition of the matrix on the left-hand side, with $D_{b_2|\tau}$ being the diagonal matrix of eigenvalues, and $B_{d_1|\tau}$ being the corresponding matrix of eigenvectors. The left-hand side of the equation can be estimated from the data, therefore this equation can be used to identify both $B_{d_1|\tau}$ and $D_{b_2|\tau}$ simultaneously. It is worth noting that (11) is point-wise in b_2 and the eigenvalue–eigenvector decomposition holds for any b_2 on its support.

Remark. It is necessary to discretize the two bids b_1 and b_3 to accommodate the discreteness of the unobserved types of bidders. The two discretized bids d_1 and d_3 act as an observed contaminated measure, and an instrumental variable of bidders' type, respectively. This also justifies our requirement of three bids for each bidder. Nevertheless, only one bid needs to be discretized if a proxy or an instrument is available for bidders' type. Different choices of discretization may result in different matrices B_{d_1,d_3}, $B_{d_1|\tau}$, and B_{τ,d_3}. However, both Lemma 2 and (11) hold for all possible methods of discretization given Assumption 4 is imposed. The validity of Lemma 2 implies that the identified number of types is not affected by discretization. Our procedure of identification and estimation are based on the main identification equation (11), which is point-wise in b_2, a continuous variable. Therefore, we are still dealing with continuous bids as in the literature on nonparametric analysis of first-price auctions and discretization will have no impact on our estimation results. ■

Assumption 5. There exists a known realization of b_2, $\eta \in [\underline{v}, \bar{v}]$ such that $g(\eta|\tau = Lk) \neq g(\eta|\tau = Lj)$, whenever $Lk \neq Lj$, $Lk, Lj \in \mathcal{K}$.

Assumption 5 is a non-degeneracy condition and it can be easily satisfied. In fact, for both level-k and AVD models, the set $\{b : g(b|\tau = Lk) \neq g(b|\tau = Lj)\}$ has nonzero Lebesgue measure. First, in level-k auctions, $g(b|\tau = Lk) = F(s^{-1}_{Lk}(b)) = F(s^{-1}_{Lj}(b)) = g(b|\tau = Lj)$, $\forall b$ holds only if bidders of type Lk and Lj bid using the same strategy, and this is impossible by construction. In AVD, recall that in Lebrun (1999), the inverse bidding strategy $s^{-1}_k(\cdot)$, $k = 1, 2, \dots, K$ are characterized by

$$\frac{ds^{-1}_k(b)}{db} = \frac{F_k(s^{-1}_k(b))}{(I-1)f_k(s^{-1}_k(b))} \left\{ \sum_{l\neq k, l=1}^{K} \frac{1}{s^{-1}_l(b) - b} - \frac{(I-2)}{s^{-1}_k(b) - b} \right\}.$$

Applying the relationship between $F_k(\cdot)$ and $G_k(\cdot)$, we rewrite this equation as

$$1 = \frac{G_k(b)}{(I-1)g_k(b)} \left\{ \sum_{l \neq k, l=1}^{K} \frac{1}{s_l^{-1}(b) - b} - \frac{(I-2)}{s_k^{-1}(b) - b} \right\}. \tag{12}$$

It is readily to show that if $G_k(b) = G_j(b)$ for all b, then $s_k^{-1}(b) = s_j^{-1}(b)$, i.e., these two types are symmetric. We conclude that for bidders with different value distributions, the set $\{b : g(b|\tau = Lk) \neq g(b|\tau = Lj)\}$ has nonzero Lebesgue measure.

Proposition 1. *Suppose the condition RC and Assumptions 2–5 hold, then the number of types m, and the m bid distributions in level-k auctions are nonparametrically identified from the observed data.*

The proof of this proposition follows from the insight in Hu (2008) and Hu et al. (2013), and is included in Appendix A. To further identify the latent value distribution $F(\cdot)$, we present the following important proposition for level-k auctions.

Proposition 2. *Suppose there are I bidders participating in a first-price sealed-bid auction, and the bidders are of two consecutive types Lj, Lk with $k \geq 2$ and $j = k - 1$. The pseudo-private values for bidders of type Lk can be expressed as*

$$v \equiv \xi_{Lk}(b, G(b|Lj), I) = b + \frac{1}{I-1}\frac{G(b|Lj)}{g(b|Lj)}, b \in [\underline{v}, s_{Lk}(\bar{v})], \tag{13}$$

where $G(\cdot|Lj)$ and $g(\cdot|Lj)$ are CDF and density function of type-Lj bidders' bids, respectively.

Proof. See Appendix A. ■

Proposition 2 directly associates higher-type bidders' pseudo-values with lower-type bidders' bid distribution. A similar relationship between values and distribution of bids also holds for first-price auctions at Bayesian-Nash equilibrium, and it was first proposed in Guerre et al. (2000) (GPV hereafter). That paper employs the relationship to nonparametrically identify and estimate first-price auctions within an IPV paradigm. Intuitively, the existence of the similar relationship in both Bayesian-Nash equilibrium and the level-k model is due to the fact that in the level-k model, bidders of type Lk always best respond to bidders of type Lj, just as Bayesian-Nash bidders best respond to each other. However, bidders in level-k auctions are "asymmetric" in the sense that different types of bidders employ different bidding strategies, and bidders of lower-type never take into account the possibility that their opponents are of higher-types. Consequently, Proposition 2 does not provide a relationship between lower-type bidders' pseudo-values and higher-type bidders' bid distribution. The value-bid relationship described in (13) also has different implications in Bayesian-Nash equilibrium from the level-k model. As GPV show, the relationship in Bayesian-Nash equilibrium can be used to recover pseudovalues for each individual bidder from her bids. In level-k auctions, however, we can only recover value distribution for an individual type. The reason is that in level-k auctions only the distribution of types rather than the type of an individual bidder is identifiable.

Based on the iterative relationship represented in Proposition 2, the value distribution and all higher-type $Lk(k \geq 2)$ bidders' bidding strategies are identified.

Proposition 3. *Suppose $G(b|L1), G(b|L2), \ldots, G(b|Lm)$ are $m \geq 2$ identified bid distributions for type-L1, L2, ..., and Lm, respectively. If the values of m-types' bidders are i.i.d., then the inverse of bidding strategy for type-Lk bidders, $s_{Lk}^{-1}(b, F, I)$, can be identified as $s_{Lk}^{-1}(b, F, I) = \xi_{Lk}(b, G(b|Lj), I)$, where $\xi_{Lk}(b, G(b|Lj), I)$ is recovered from Proposition 2 for $k = 2, 3, \ldots, m$, and $j = k - 1$. Moreover, the value distribution $F(\cdot)$ can be uniquely determined as $F(v) = G(\xi_{Lk}(v, G(\cdot|Lj), I)|Lk)$ by bidding strategy and the corresponding bid distribution of any type-Lk, where $k = 2, \ldots, m$.*

Proof. The proof follows that of Theorem 1 in GPV. ■

In our analysis of the level-k auction model, bidders of type-$L0$ are assumed to only exist in higher-type bidders' mind, hence the bidding behavior of type-$L0$ is captured by type-$L1$ bidders' belief. This belief can be captured by a CDF $Z(\cdot)$, with its density $z(\cdot) > 0$ on the support of values, $[\underline{v}, \bar{v}]$. We show in the following that the CDF $Z(\cdot)$ is nonparametrically identified using the results in Proposition 3. Without loss of generality, we assume that there are two types of bidders: type $L1$ and $L2$. According to the theoretical model of level-k auctions, a type-$L1$ bidder i solves the maximization problem by best responding to her own belief:

$$\max_{b_i}(v_i - b_i)\Pr(\max_{l \neq i} b_l \leq b_i) = \max_{b_i}(v_i - b_i)Z^{I-1}(b_i),$$

with the solution

$$s_{L1}^{-1}(b) = b + \frac{1}{I-1}\frac{Z(b)}{z(b)}, b \in [\underline{v}, s_{L1}(\bar{v})], \tag{14}$$

where $s_{L1}^{-1}(\cdot)$ is the inverse function of type-$L1$ bidders' optimal bidding strategy. The inverse function $s_{L1}^{-1}(\cdot)$ associates type-$L1$ bidders' bid distribution $G(b|L1)$ with their value distribution directly, i.e., $G(b|L1) = F(s_{L1}^{-1}(b))$. According to Proposition 3, the value distribution $F(\cdot)$ is identified as $F(\cdot) = G(\xi_{L2}(\cdot, G(\cdot|L1), I)|L2)$ from the bid distribution $G(\cdot|L2)$ and bidding strategy $\xi_{L2}(\cdot)$. The identification of both type-$L1$ bidders' bid distribution $G(\cdot|L1)$ and the value distribution $F(\cdot)$, together with the regularity condition under which $s_{L1}(\cdot)$ is strictly increasing, implies that the inverse function $s_{L1}^{-1}(b)$ is nonparametrically identified, and can be expressed explicitly as

$$s_{L1}^{-1}(b) = F^{-1}\left(G(b|L1)\right), b \in [\underline{v}, s_{L1}(\bar{v})]. \tag{15}$$

According to (14), type-$L1$ bidders' belief about other bidders' bids, $Z(\cdot)$, is identified from the inverse bidding function $s_{L1}^{-1}(\cdot)$. In fact, $Z(\cdot)$ can be solved explicitly as:

$$Z(b) = \begin{cases} \exp\left\{-\frac{1}{I-1}\int_b^{\bar{b}} \frac{d\mu}{F^{-1}(G(\mu|L1)) - \mu}\right\}, & b \in (\underline{b}, \bar{b}], \\ 0, & b = \underline{b}. \end{cases} \tag{16}$$

Corollary 1. *Suppose the assumptions of Proposition 3 hold, and there are at least two types of bidders, then type-L1 bidders' belief, which is captured by a CDF of bids $Z(\cdot)$, is identified to be (16) from the observed bids.*

Remark. In the argument of identification above, bidders of type-$L0$ are assumed to bid randomly without any further restrictions. However, we may impose some sensible restrictions to type-$L0$'s behavior. A prominent one is that any expected utility maximizing bidder would never bid beyond her own value if she assumes that any bid could win the auction with non-zero probability (Battigalli and Siniscalchi 2003). Such a restriction may improve the identification result of type-$L0$ by putting a bound for $Z(\cdot)$. For example, Aradillas-Lopez and Tamer (2008) explore the bounds of parameters imposed by level-k rationality. The possibility is out of the scope of this paper and we will leave it for future research. ■

We have identified bidders' value distribution, and the bidding strategies for bidders of all types. It remains to show that the proportion of each type is identified. For this purpose, we start with the linear system $G(b) = \sum_{\tau \in \mathcal{K}} G(b|\tau)p(\tau)$ with $G(b)$ being the observed distribution of bids at tth auction and $G(b|\tau)$ being the identified bid distribution conditional on type τ. In this linear system, the probability vector $p \equiv (p(L1), \ldots, p(Lm))$ consists of $m - 1$ unknowns while b is continuous, thus the linear system is overdetermined. In this system, the number of linearly independent equations is always $m-1$ because of Assumption 4. Therefore, the solution of the linear system $p(L1), \ldots, p(Lm)$ exists and is unique.

The complete results of identification are summarized in Theorem 1.

Theorem 1. *Suppose that (1) observed bids are generated by level-k first-price auctions within an IPV paradigm. (2) The bidders of these auctions are of m ($m \geq 2$) consecutive types with the lowest type being L1, and all the m types are responding to a unique type-L0. (3) The condition RC and Assumptions 2–5 hold, then the number of types m, the probability vector $p \equiv (p(L1), \ldots, p(Lm))$ of type, the value distribution of bidders $F(\cdot)$, and the specification of type-L0, $Z(\cdot)$, are all nonparametrically identified from the observed bids.*

The application of the identifying procedure for level-k auctions to AVD is straightforward. First off, Proposition 1 holds under Assumptions 1–4, where the discussions on the assumptions can be easily extended to the AVD. The ordering of identified bid distributions for AVD relies on Assumption 1: according to the discussion above this assumption, the stochastic dominance of any two value distributions implies the stochastic dominance of the corresponding distributions of bids and this allows us to completely order the K bid distributions (denoted by $G_1(\cdot), \ldots, G_K(\cdot)$). To identify the K value distributions $F_1(\cdot), \ldots, F_K(\cdot)$ from $G_1(\cdot), \ldots, G_K(\cdot)$, the standard approach does not apply because the bidding strategies of different types are interdependent and there does not exist a relationship as in a symmetric auction proposed in Guerre et al. (2000). We utilize the characterization of equilibrium (equation (1)) in Lebrun (1999) to solve the value distributions. From (12) we obtain

$$s_k^{-1}(b) = b + (I-2)\left\{\sum_{l\neq k, l=1}^{K} \frac{1}{s_l^{-1}(b) - b} - \frac{(I-1)g_k(b)}{G_k(b)}\right\}^{-1},$$
$$k = 1, 2, \ldots, K. \tag{17}$$

With $G_k(b)$ and $g_k(b)$ being identified, (17) contains K equations with the same number of unknowns $s_k^{-1}(b)$, $k = 1, 2, \ldots, K$ for any given b. If the Jacobian matrix for (17) is full rank for any b, then we can uniquely determine $s_k^{-1}(\cdot)$, $k = 1, 2, \ldots, K$. Consequently, the results in Proposition 3 apply and the value distributions $F_k(\cdot)$, $k = 1, 2, \ldots, K$ are all nonparametrically identified.

3.3. Discussions on estimation and hypothesis testing

The methodology of identification provides a constructive two-step procedure to estimate a level-k auction model: the first step estimates the bid distribution function $G(\cdot|\tau)$, the density $g(\cdot|\tau)$, and type distribution $p(\cdot)$; the second step estimates the value distribution $F(\cdot)$ from the results of the first step and Proposition 2. The details are included in Appendix B. Note that the value distributions in AVD can be estimated using various methods, e.g., we parametrize the distributions then employ the systems method of moments estimator discussed in Section 6.10.3 of Cameron and Trivedi (2005).

The results of identification provide a testable implications for the model: For each type Lk, a value distribution $F(\cdot)$ can be estimated, and the $m-1$ estimates of value distribution are identical since the values of all the bidders are assumed to be i.i.d.. Therefore a prediction of level-k auction models is $F_{Lk}(\cdot) = F_{Lj}(\cdot)$ for any two types $Lk, Lj \in \mathcal{K}$, where $F_{Lk}(\cdot)$ and $F_{Lj}(\cdot)$ are value distribution functions estimated for type Lk and Lj, respectively. If we conduct the following hypothesis

$$H_0 : F_{Lk}(v) = F_{Lj}(v), \quad Lk, Lj \in \mathcal{K}, k, j = 2, 3, \ldots, m;\ v \in [\underline{v}, \overline{v}]$$

against the alternative $H_1 : F_{Lk}(v) \neq F_{Lj}(v)$ then level-k auction models are rejected if we reject H_0 under our assumptions of identification, otherwise we are in favor of level-k auction models. Of course, rejecting H_0 does not necessarily imply a rejection of level-k model, but at least one of the maintained assumptions. It is worth noting that the test requires $m-1 \geq 2$, i.e., there are at least three type of bidders in auctions.

Testing equality of two functions nonparametrically has been broadly studied in statistics literature, e.g., see Dette and Neumeyer (2001) and the references there. We follow the literature in constructing the test statistic and computing its critical region. For any two types $Lk, Lj \in \mathcal{K}$, the test statistic is computed as follows.

$$\widehat{S}_n = \frac{1}{n}\sum_{i=1}^{n}\left[\widehat{F}_{Lk}(v_i) - \widehat{F}_{Lj}(v_i)\right]^2, \tag{18}$$

where $v_i, i = 1, 2, \ldots, n$ are n design points chosen from the estimated support $[\widehat{\underline{b}}, \widehat{\overline{\xi}}]$. The asymptotic distribution of the test statistic $\widehat{S}_n$ is model-specific and very difficult to obtain. Even the distribution is available, the rate of convergence of $\widehat{S}_n$ is rather slow. Therefore, bootstrapped re-samples are used to compute the test statistic and conduct the hypothesis. To be specific, suppose that we generate T samples by bootstrap, and for each sample a test statistic $\widehat{S}$ can be obtained. Denote the sequence of statistic $\{\widehat{S}_{n,1}, \widehat{S}_{n,2}, \ldots, \widehat{S}_{n,T}\}$, then a bootstrap critical region of significance level α is then given as

$$\widehat{S}_n > \widehat{S}_{n,([(1-\alpha)T])}, \tag{19}$$

where [ ] represents the integer part, and $\{\widehat{S}_{n,(i)}\}_{i=1}^{T}$ is the sample $\{\widehat{S}_{n,i}\}_{i=1}^{T}$ arranged in increasing order of magnitude.

4. Nonparametric identification: Extensions

In this section, we extend the results of identification in Theorem 1 to the following two cases: (1) the lowest type is not necessarily $L1$ and unknown to the researcher; and (2) some bidders may play Nash equilibrium.

The identification results in Theorem 1 rely on the assumption that the lowest type of bidder in the data is known to be $L1$. If this assumption is relaxed and the lowest type is unknown to the researcher (it can be $L1$ or Lk, $k \geq 2$), we can still nonparametrically identify the model and the results are similar to that in Theorem 1. For the purpose of identification, it is necessary to maintain that (1) types are consecutive (e.g., $L8, L9$ and $L10$). (2) All bidders' reasoning starts from $L0$.

When the lowest type is unknown, the identification results in Propositions 1 and 2 still hold. Without loss of generality, we assume that from the first step of identification, we identify that $m = 2$ and the two types of bidders are Lj and Lk, where $k = j+1$. The identified CDF (pdf) for type Lk and Lj are denoted by $G(b|Lj)$ ($g(b|Lj)$) and $G(b|Lk)$ ($g(b|Lk)$), respectively. According to Proposition 3, the value distribution $F(\cdot)$ and the bidding strategy of type Lk can both be identified. Then employing the relationship $s_{Lj}^{-1}(b) = F^{-1}(G(b|Lj))$, the bidding strategy of type Lj, $s_{Lj}(\cdot)$ can be identified. Even though there is no lower type than type Lj in the data, the bidding strategy $s_{Lj}(\cdot)$ still satisfies

$$s_{Lj}^{-1}(b) = b + \frac{1}{I-1}\frac{G(b|Lj-1)}{g(b|Lj-1)}, b \in [\underline{v}, s_{Lj}(\overline{v})],$$

where $G(\cdot|Lj-1)$ and $g(\cdot|Lj-1)$ are CDF and pdf of bids for type-$L(j-1)$ which contain the information of the belief for bidders of type Lj. The results in Corollary 1 imply that the distribution function $G(\cdot|Lj-1)$ is identified. By repeating the argument above, we are able to identify $G(\cdot|Lj-2), G(\cdot|Lj-3), \ldots$. Suppose that $G(\cdot|Lj-l)$ is the *first* identified function such that $G(\cdot|Lj-l-1)$ is not a CDF anymore, while all other functions $G(\cdot|Lj-2), G(\cdot|Lj-3), \ldots, G(\cdot|Lj-l)$ are, then $G(\cdot|Lj-l)$ is corresponding to the CDF $Z(\cdot)$ in Corollary 1, and type-$(Lj-l+1)$ is just the lowest type $L1$.[7] Consequently, type Lj is identified. Therefore, upon the identification of the number of type, m, in the first step of identification, we are able to identify exactly what the m types are, given that all the types are consecutive. This result is summarized in the following corollary of Theorem 1. To state the result, we list an alternative assumption of (2) in Theorem 1 first:

Condition (2′). *The bidders are of m ($m \geq 2$) consecutive types, and all the m types are responding to a unique type-L0.*

Corollary 2. *Suppose that the condition* (2′) *and the conditions* (1) *and* (3) *in Theorem 1 hold, then the number of types m, the probability of each type, the value distribution of bidders $F(\cdot)$, and the specification of type-L0, $Z(\cdot)$, are all nonparametrically identified from the observed bids.*

Remark. Corollary 2 implies that the model is still identified without specifying the lowest type of bidders in the data. For example, we are able to distinguish the following two cases with $m = 3$: bidders are of type $\{L1, L2, L3\}$ and type $\{L8, L9, L10\}$. Intuitively, Corollary 2 is true because bidders of higher types always start their reasoning from $L0$ to $L1$, to $L2, \cdots$ Therefore bidding behavior of higher types (e.g., Lk) also reveals information of type $Ll(l = 0, 1, \ldots, k-1)$ even though type Ll does not exist in the data.

To analyze the model when some bidders play Nash equilibrium, we first consider identification of the model when bidders' are of type $\mathcal{K} = \{L1, \ldots, L(m-1), L\infty\}$ ($m \geq 2$) where $L\infty$ is Nash-type who plays BNE. Without loss of generality, we illustrate the identification procedure by assuming $\mathcal{K} = \{Lj, Lk, L\infty\}$, $k = j+1$. By employing the result in Proposition 1, both the number of types and the three bid distribution functions can be identified. Unfortunately, the three identified distributions cannot be completely ordered using Lemma 1. This is due to the fact that in a level-k auction model type-Lk's bidding strategy does not necessarily converge to that of a Nash-type as k goes to infinity and consequently the result in Lemma 1 does not apply to the Nash-type. Nevertheless, Lemma 1 enables us to obtain a partial order of three bid distributions: the bid distribution with the smallest support must correspond to type Lk or $L\infty$ (we assume the three supports are different. This is without loss of generality because any two supports equal with zero probability). We start from this distribution and take several steps to identify the model. First, we tentatively assume this distribution is for the Nash-type. Following GPV, we are able to recover a CDF F_∞.

$$F_\infty(v) = G(\xi_{L\infty}(v, G(\cdot|L\infty), I)|L\infty), \quad (20)$$

where $G(\cdot)$ is the bid distribution and the inverse bidding strategy $\xi_\infty(\cdot)$ is $\xi_{L\infty}(b, G(b|L\infty), I) = b + \frac{1}{I-1}\frac{G(b|L\infty)}{g(b|L\infty)}$ for all $b \in [\underline{v}, s_{L\infty}(\bar{v})]$. For the remaining two bid distributions, we first apply Lemma 1 to order them, then recover a value distribution $F_k(\cdot)$ using the results in Theorem 1. The maintained assumption of values being i.i.d. implies that the type we first chose is $L\infty$ only if $F_\infty(\cdot) = F_k(\cdot)$. Otherwise, the first chosen type must be Lk. By testing the two value distributions, we can achieve identification of the model. The belief $Z(\cdot)$ can be also identified following the procedure we discussed previously. Such a procedure can be readily applied to the case of more types.

[7] We may continue this procedure to obtain objectives $G(\cdot|Lj-l-1)$, $G(\cdot|Lj-l-2)$,... However, all bidders' reasoning in a level-k model starts from the belief of type $L1$, $Z(\cdot)$. Thus $G(\cdot|Lj-l-q)$, $q = 1, 2, \ldots$ would not be in the model. To simplify our exposition, we assume that none of the objectives $G(\cdot|Lj-l-q)$ are CDFs. Such restriction does rule out the possibility that $G(\cdot|Lj-l-q)$ may still be a CDF. However, it is unknown whether our method can be applied to identify such a richer model.

Proposition 4 (*Identification of Level-k Auctions with Bidders of Nash-type*). *Suppose the set of possible types $\mathcal{K} = \{Lk, \ldots, L(k+l), L\infty, l \geq 1\}$, then the level-k auction model is identified.*

The result above shows that our methodology still achieves identification when there exist bidders who play a BNE. The importance of the result is twofold: first, BNE is a maintained assumption in most of the empirical work in auctions. Hence it is crucial for our methodology to include BNE as a special case. Second, when $k \to \infty$ type-Lk does not necessarily converge to Nash-type, i.e., level-k auction model does not nest BNE as a special case. Therefore, we have to include it in the model explicitly.

5. Monte Carlo experiments

In this section, we present some Monte Carlo evidence for the methodology of estimating level-k auctions. Using simulated bids, we first estimate the bid distribution function $G(\cdot|\tau)$, then recover the type distribution $p(\cdot)$, and bidders' valuation distribution $F(\cdot)$.

Suppose N bidders participate in $3N$ homogeneous auctions with three bidders ($I = 3$) in each individual auction, and every bidder participates in three auctions. The bids are re-arranged such that the ith row of the data contains three bids of ith bidder b_{i1}, b_{i2}, b_{i3}. The true distribution $F(\cdot)$ of private values is truncated log-normal. The Monte Carlo experiment consists of $S = 400$ replications. For each replication, we first generate $3N$ random private values from the truncated log-normal distribution, then compute the corresponding bids for bidders of each type according to the theoretical model of level-k auctions. Using the two-step estimation procedure, we first estimate the m bid distributions from the computed bids. Then the proportion of types and distribution of values are estimated in the second step.

We consider two settings of level-k auction models. In the first setting, bidders are of two types: $L1$ and $L2$ (denoted by $L1-R$ and $L2-R$), with proportions being 0.55 and 0.45, respectively, and the type $L0$ is specified as "random", i.e., this nonstrategic type's bids are randomly distributed on the support of values. The parameters of the log-normal distribution $F(\cdot)$ are zero and one. The estimated distribution functions of values and bids are presented in Figs. 1–3 for sample size $N = 200$, $N = 500$, and $N = 5000$, respectively. In each figure, the left plot contains the estimated (labeled as avg.$G(b|\cdot)$) and "true" distributions (labeled as $G(b|\cdot)$) of bids for two types where the "true" distributions are empirical CDFs recovered from the sample assuming bidders' types are known. The right plot presents the estimated and "true" CDFs of bidders' value, where the estimated distribution of values are computed using the estimated CDF of bids for type $L1$, and labeled as "est-dist". We observe that the estimates track the true ones closely for the modest-sized datasets, especially when bids are high; when bids are low, bids of two types are close to each other. The two plots in Fig. 4 present the 90% confidence interval of bid distribution for two types ($N = 5000$). The type distribution $p(\tau)$ is estimated using least squares estimator and the results are presented in Table 1.

In the second setting, bidders are of three types: $L1-R$, $L2-R$, and $L1-T$, (corresponds to the truthful type $L0$ who bids her own value) with probability 0.5, 0.3, and 0.2, respectively. The mean and variance of the log-normal distribution $F(\cdot)$ are set to be five and three. Figs. 5 and 6 contain our estimation results for sample size $N = 500$ and $N = 5000$. Estimates of type distribution $p(\tau)$ is shown in Table 2. Not surprisingly we observe that for the same sample size the performance of the estimates for three types are inferior to that of two types. Nevertheless, for the data of modest size the estimates still perform well.

An observation from the estimated results is that the distribution of types is noisily estimated for the modest-sized data. The relative poor performance of the estimate is also observed in

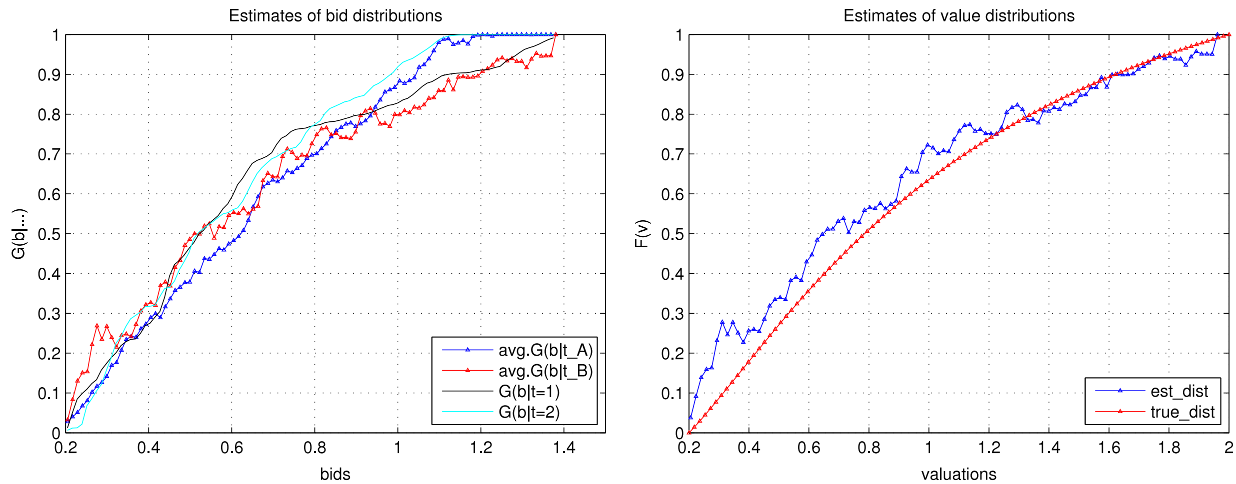

Fig. 1. Monte Carlo evidence: $m = 2, N = 200$.

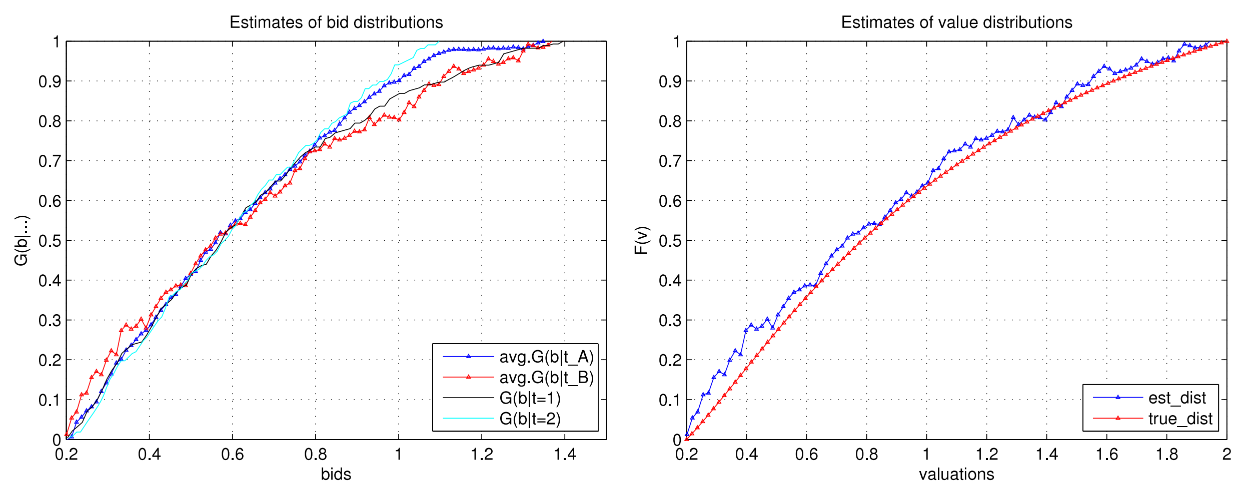

Fig. 2. Monte Carlo evidence: $m = 2, N = 500$.

Table 1
Estimated $p(\tau)$, two types.

		True value	Estimate
$N = 200$	$p(L1 - R)$	0.55	0.68
		–	(0.42)
	$p(L2 - R)$	0.45	0.32
		–	(0.42)
$N = 5000$	$p(L1 - R)$	0.55	0.56
		–	(0.35)
	$p(L2 - R)$	0.45	0.44
		–	(0.35)

Table 2
Estimated $p(\tau)$, three types.

		True value	Estimate
$N = 500$	$p(L1 - R)$	0.50	0.44
		–	(0.43)
	$p(L2 - R)$	0.30	0.40
		–	(0.40)
	$p(L1 - T)$	0.20	0.16
		–	(0.28)
$N = 5000$	$p(L1 - R)$	0.50	0.46
		–	(0.24)
	$p(L2 - R)$	0.30	0.27
		–	(0.39)
	$p(L1 - T)$	0.20	0.27
		–	(0.39)

the Monte Carlo study of Gillen (2009), where $p(\tau)$ is estimated parametrically. As a matter of fact, estimating the coefficients of a finite mixture model can be difficult even though the components are parametric. For example, the widely used expectation–maximization (EM) algorithm in estimating mixture models may suffer the unbounded likelihood functions, or depend severely on particular starting points, e.g., see Melnykov and Maitra (2010) for a survey on estimation of finite mixture models.

Before estimating the model, we estimated the number of possible types in both settings by testing the hypothesis that $\det(B_{d_1,d_3}) = 0$ for different value of M. For the two-type setting, we examine the rank condition by testing the null hypothesis $\det(B_{d_1,d_3}) = 0$ under different specifications of m, the number of types. Under the specification $m = 2$ the null hypothesis is rejected under the 5% significance level, but we fail to reject the hypothesis when $m = 3$ and $m = 4$. Analogously, for the three-type setting we are able to reject the null hypothesis for $m = 2$ and $m = 3$ but fail to reject it for $m = 4$ and $m = 5$ under the 5% significance level.[8] According to Condition 4, $m = \max\{M : det(B_{d_1,d_3}) \neq$

[8] To ensure the robustness of the estimated results, we tried several methods of discretization. The results show that the method of discretization does not affect the estimation much especially when the sample size is large.

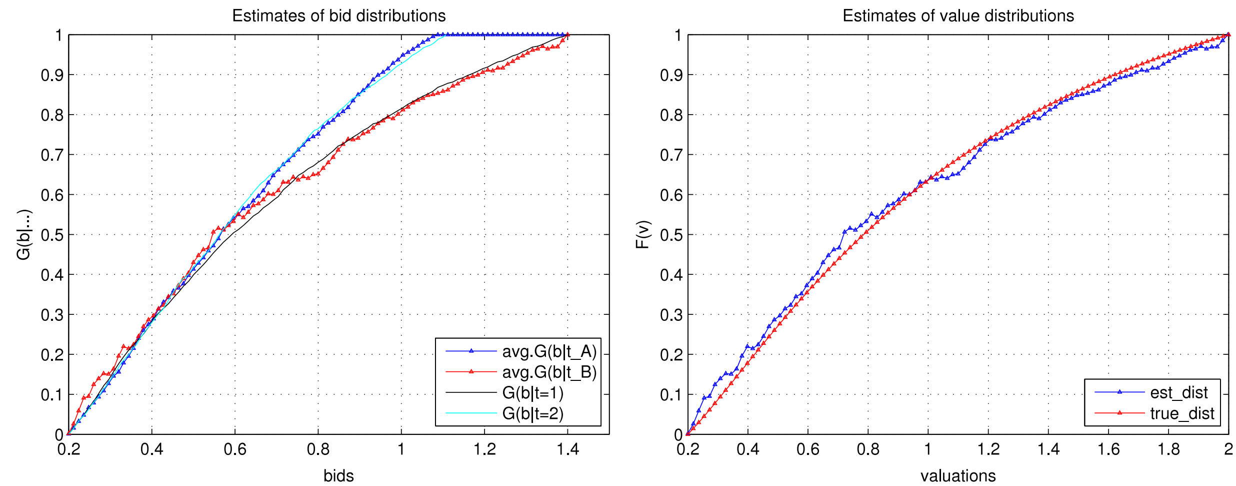

Fig. 3. Monte Carlo evidence: $m = 2$, $N = 5000$.

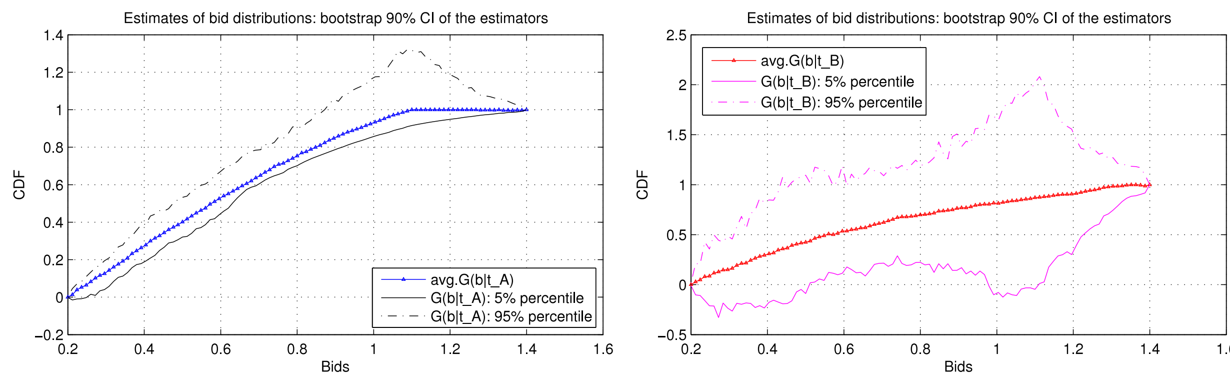

Fig. 4. Estimated bid distribution: bootstrap 90% confidence interval, $N = 5000$.

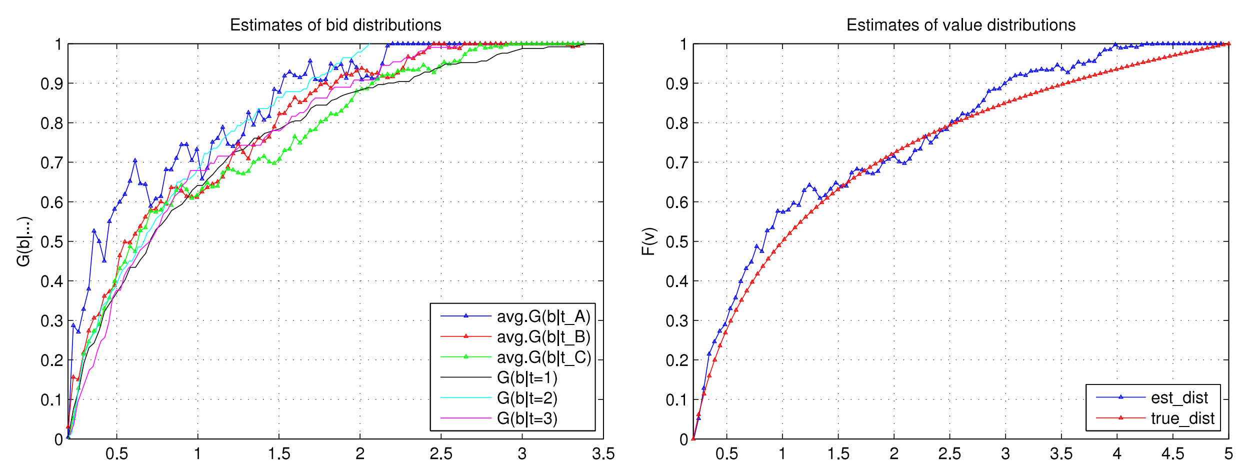

Fig. 5. Monte Carlo evidence: $m = 3$, $N = 500$.

$0, M = 2, 3, \ldots\}$, the testing results above imply that the number of possible types are two and three in the two settings.

6. Empirical application: USFS timber auctions

In this section, we apply the proposed estimation procedure to USFS timber auction data, and illustrate that bidders display non-equilibrium beliefs. Moreover, we recover the value distributions for each identified type and test whether the values distributions are identical across type. The testing results imply that the level-k auction model has support from field data. Our evidence does not rule out the possibility that the observed bidding behavior can also be explained by other alternative asymmetry of bidders, e.g., heterogeneous risk preference, asymmetric value distributions, etc. Rather the evidence sheds some light on our understanding on the effects of non-equilibrium beliefs on bidders' strategic bidding behavior.

Please cite this article in press as: An, Y., Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.06.014.

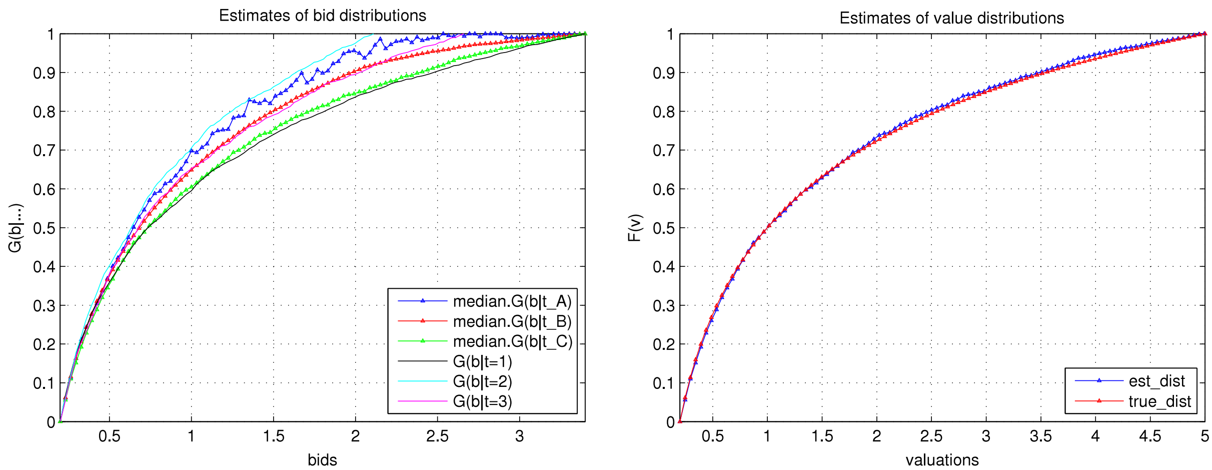

Fig. 6. Monte Carlo evidence: $m = 3, N = 5000$.

Table 3
Summary statistics 1 (5901 Auctions).

Variable	Mean	Standard deviation	Minimum	Maximum
Bids($/mbf)	52.45	45.91	0.13	534.58
Appraisal value	31.94	30.23	0.43	450.40
Volume (mbf)	2808	2958	10	56350
Number of bidders	3.76	1.86	2	12

6.1. The data

The USFS timber auction data have been extensively used for analysis of various auction models, and some of the analyses document the deviation of bidders' behavior from Bayesian-Nash equilibrium with risk-neutral bidders. For example, the empirical results in Baldwin (1995) suggest that bidders might be of decreasing absolute risk aversion. Athey and Levin (2001) argue that bidders' behavior is consistent with some amount of risk aversion. Recently, Dohmen et al. (2010) show empirically that risk aversion is related to cognitive ability while the heterogeneous bidding behavior of level-k bidders can be understood as the difference in cognitive ability among bidders. Therefore, the evidence of risk aversion from USFS auctions implies that non-equilibrium behavior of bidders might be appropriate to fit the data. In fact, in the literature on auction models, both level-k models and risk-aversion of bidders are employed to explain risk-neutral bidders' deviation from Bayesian-Nash equilibrium. However, it is not clear how we distinguish the two explanations both theoretically and empirically. We are investigating this research problem in an going project.

In all of the Forest Service auctions, we focus on those first-price sealed-bid ones. In order to adapt the private value paradigm, we only consider those auctions of the type "scaled sale". In a "scaled sale" auction, bidders bid on a per unit basis (thousand board-feet or mbf) and the payments are based on the winning bidder's unit prices and the *actual* volumes, which are measured by a third party at the time of harvest. Haile et al. (2003) empirically demonstrate that there is little evidence of common values for scaled sale auctions. Many other studies also assume private values for timber auctions, e.g., see Baldwin et al. (1997), Haile (2001) and Haile and Tamer (2003) among others. We further assume that the reserve price is non-binding for the auctions which is a widely acknowledged property in the literature, e.g., Haile (1996), Baldwin et al. (1997), Haile (2001), and Campo et al. (2011). The sample contains scaled sale sealed-bid first-price auctions conducted from 1982 to 1993 in all regions by eliminating both salvage and small-business set-aside sales. In each auction, we observe all the bidders' identity and their bids (dollar per mbf), the volume and the appraisal value of each tract. Table 3 presents some summary statistics on the number of bidders, bids per mbf, the appraisal value per mbf, and the volume in mbf. Bids and appraisal value are in 1982 real dollars. Different units of timber are used in different auctions, we convert all the units to mbf.

To apply the proposed estimation procedure to the data, it is necessary to address the heterogeneity of auctioned tracts across auctions. As argued in Haile (2001) and Campo et al. (2011), in all the observed features of auctions the appraisal values of the auctioned lots best capture the auction-specific heterogeneity. Hence we control the heterogeneity by running a first-stage linear regression of bids on appraisal value:

$$b_{it} = \beta_0 + \beta_1 X_t + \varepsilon_{it}, \quad \varepsilon_{it} \perp X_{it},$$

where b_{it} and X_t are the ith bid and the appraisal value of the tth auction. Then the "normalized" bid $\tilde{b}_{it}$ is constructed using the residual $\widehat{\varepsilon}_{it}$ and used for the estimation.[9] We further narrow down the sample to those auctions with three bidders. Of all the participants in those chosen auctions, $N = 462$ bidders participate in at least three auctions. We re-arrange the "normalized" bids of the 462 bidders such that $\tilde{b}_{it}$ denotes the tth bid of ith bidder, $t \geq 3$, $i = 1, \ldots, N$. Table 4 presents some summary statistics of bidders and their bids.

6.2. Estimation of type and bid distributions

The first step of the estimation is to estimate the number of possible types, m. For this purpose, we first bootstrap the data 1000 times and test the null hypothesis that the determinant of B_{d_1,d_3} vanishes, i.e., $\det(B_{d_1,d_3}) = 0$ under various specifications of the parameter m. Under the 5% significance level, we fail to

9 Such an approach has been used in literature to control for observed heterogeneity, e.g., see Haile et al. (2003) and Bajari et al. (2014) among others. The main idea is to express bids as $b = W\xi + \eta$, where W is a vector of characteristics and $\eta \perp W$. Thus the conditional CDF of bids $F_{b|W} = \Pr(W\xi + \eta \leq b) = F_\eta(b - W\xi)$. Normalized bids $b - W\xi$ are then constructed using the residuals when we regress bids on some observed auction characteristics.

Table 4
Summary statistics 2 ("Normalized" bids).

Variable	Mean	Standard deviation	Minimum	Maximum
# bidders in each auction	3	–	–	–
Identified bidders (N)	462	–	–	–
Auctions participated	4.99	4.56	3	65
Bids ($/mbf)	11.14	11.98	0.02	105.16

Table 5
Condition number of the matrix $B_{d_1,d3}$.

Possible types	Condition number				
	10th prctile	20th prctile	30th prctile	40th prctile	Median
2	13.3	15.1	17.3	20.6	24.3
3	10.7	13.0	15.3	18.7	22.7
4	19.1	23.4	29.1	35.3	44.3
5	27.1	33.6	40.1	49.5	62.3

reject the null hypothesis for the specification $m = 4$ and $m = 5$. Under the specifications $m = 2$ and $m = 3$, we reject the null hypothesis. To infer the parameter m from the testing results, we need to consider the fact that if the rank of the matrix B_{d_1,d_3} is m, then we will also reject the null hypothesis under the specification $m-1, m-2, \ldots, 2$. Therefore the testing evidence above does not support more than three types, and at the 5% significance level, the data support three types of bidders.

To reinforce the identification result of m, we further investigate the condition number of the matrix B_{d_1,d_3} under different specifications of the parameter m. Even though a deterministic relationship between condition number and determinant of a matrix does not exist, a larger condition number is a reasonable indicator of the matrix being closer to singular. Hence we bootstrap the data 1000 times and compute the median and some other percentiles of condition number for the matrix B_{d_1,d_3} under specifications $m = 2, 3, 4, 5$, and the results are presented in Table 5. The results for condition number also provide a reasonable implication that the number of possible types is three. After identifying the number of types, I proceed to estimate the three bid distributions. Fig. 7 presents the estimates of bid distribution for the three types. The two subplots (from left to right) are results of one-time estimation, and the median of the estimates based on 1000 bootstrapped samples, respectively. The consistent estimates in the two subplots imply that the one-time estimate is robust, and it is easy to identify (according to the monotonicity of the supports for the distributions) that the blue, red, and black curves are CDFs of bids for the lowest type, a higher type, and the highest type, respectively. Fig. 8 provides the 90% confidence interval of the three estimated bid distribution functions. Based on the estimated bid distributions, we proceed to estimate the distribution of types. Table 6 contains the estimated results, and the standard error is computed using bootstrap sampling (1000 times). For the robustness check, the medians of the estimates are also presented. Even though the estimate of type distribution is relatively noisy as in the Monte Carlo studies, we are able to get a consistent pattern of the distribution for different discretization methods: the higher type dominates, the highest type is substantial, and the lowest type is marginal.

6.3. *Estimation of value distribution, and hypothesis testing*

Upon estimating the number of types and the distribution of bids, we proceed to estimate value distribution. In the last step of estimation, the identified three types are not further specified. According to Proposition 2, a type-Lk bidder's pseudo value can be estimated from a type-$L(k-1)$ bidder's bid distribution, hence the assumption of consecutive types is enough for us to estimate the value distribution for the higher and the highest types using bid distribution of the lowest and the higher type, respectively. The estimates are presented in Fig. 9, where the left plot contains results of one-time estimation, and the right plot presents the median of the estimates based on 1000 time bootstrapped samples. In both plots, the red and the black lines are for the higher type and the highest type, respectively.[10]

Table 6
Estimated type distribution[a].

	Estimate
p(the lowest type)	0.07
	(0.22)
p(the higher type)	0.80
	(0.33)
p(the highest type)	0.13
	(0.28)

[a] Standard error is computed using bootstrap resampling (1000 times).

To further estimate the value distribution for the lowest type, it is necessary to specify the type by a two dimensional variable $\tau = (Lk, \omega)$, where Lk is the level and ω is the specification of type $L0$. In the literature on level-k auctions, two common specifications of ω are $\omega = R$ (*random* type, type-$L0$ bidders bid uniformly randomly on the support of values) and $\omega = T$ (*truthful* type, type-$L0$ bidders honestly bid their own values). For instance, Crawford and Iriberri (2007) illustrate that over-bidding in first-price sealed bid auctions can be explained by assuming bidders are random or/and truthful types. We specify bidders as *random* types and exclude the *truthful* types ($\omega = T$). It might be plausible to assume the existence of truthful types in analyzing experimental data because the subjects may not be experienced bidders, and it is likely for them to bid their own values. However, it is not reasonable to assume the existence of truthful types when we work on field data. The reason is straightforward: it is hard to imagine that a company participating in an auction bids its own value to get nothing no matter whether it wins or not. Furthermore, we specify $k = 1$ for the lowest type, consequently, the three types are specified as $\tau = (L1, R), (L2, R), (L3, R)$. Based on this specification, we estimate the value distribution of type $(L1, R)$, and the results represented by the blue lines in both plots of Fig. 9.

Using the estimated value distribution for three types, we test three pairwise hypotheses: $F_1(\cdot) = F_2(\cdot)$, $F_1(\cdot) = F_3(\cdot)$, and $F_2(\cdot) = F_3(\cdot)$, and the corresponding statistic are $\widehat{S}_{12}$, $\widehat{S}_{13}$, and $\widehat{S}_{23}$, respectively. Based on the 1000 times bootstrap resamples, $\widehat{S}_{12} = 0.0459(\text{std.} = 0.1105)$, $\widehat{S}_{13} = 0.0431(\text{std.} = 0.1539)$, and $\widehat{S}_{23} = 0.0072(\text{std.} = 0.0583)$, and the distribution for the three statistics are shown in Fig. 10. According to the results of these statistics, none of the three hypotheses are rejected even though the results are not statistically significant. To further understand

10 The value distributions are not directly estimated. Due to the small sample properties, the estimates in Fig. 9 are non-monotone.

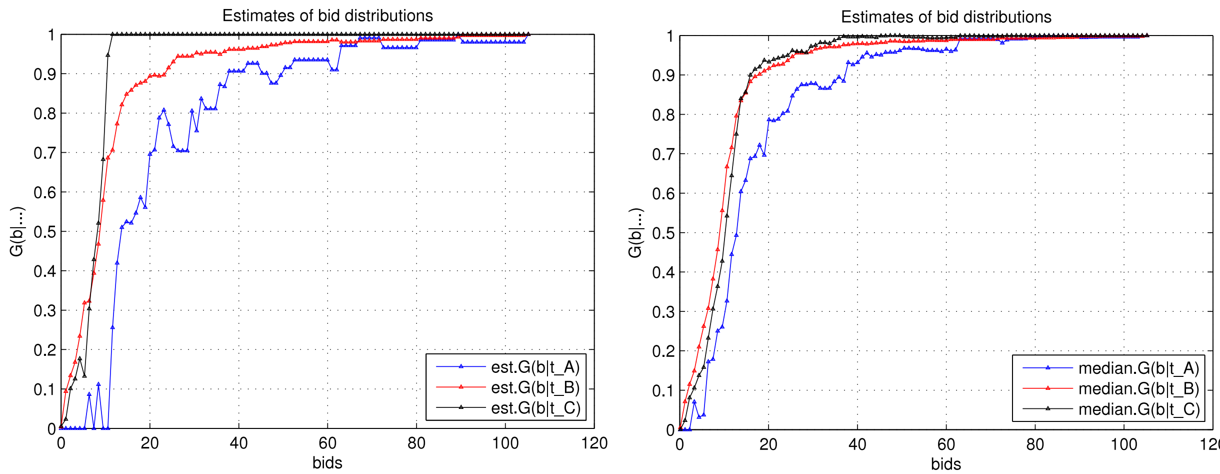

Fig. 7. USFS timber auctions: estimated bid distributions. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

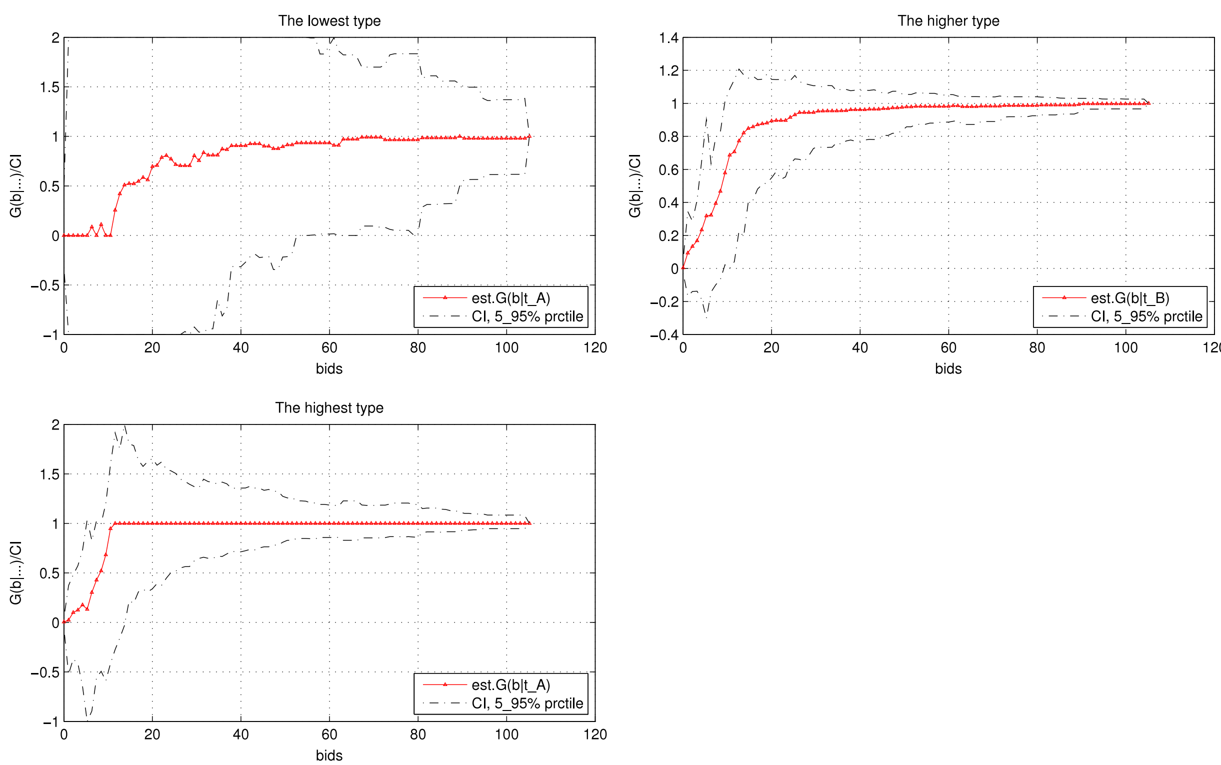

Fig. 8. USFS timber auctions, estimated bid distribution: bootstrap 90% CI.

the testing results, recall that the value distributions for type-($L2$, R) and type-($L3$, R) do not depend on the specification of type-($L1$, R). In other word, even type-($L1$, R) is mistakenly specified, the value distributions for the two higher types are still correct given these two higher types have the same ω, i.e., specification of type-$L0$. In this case, the estimated value distribution $F_1(\cdot)$ is not correct while the other two are still correct. From the testing results above we observe that test on $F_2(\cdot) = F_3(\cdot)$ is more statistically significant than the other two, and a natural explanation is that the specification of the lowest type as type-($L1$, R) may not correct, but both the two higher types have the same belief on type-$L0$ as the lowest type. Therefore, the testing results provide strong evidence that the observed bids may be rationalized by the level-k auction model. Furthermore, the results illustrate that bidders have high levels of sophistication because of the existence of three consecutive types. Experimental data heretofore seldom document level-3 behavior, but our results are consistent with the higher level of sophistication and motivation of bidders in the field.

7. Conclusion

We have developed a methodology to nonparametrically identify and estimate two models of first-price auctions: (1) bidders' beliefs are not at equilibrium but follow level-k thinking; (2) Bidders' values are asymmetrically distributed. The application of our method to USFS timber auctions demonstrates that bidders in field auction display non-equilibrium beliefs. The proposed methodology is generic and may be also applied to other games than auctions that employ level-k behavioral assumptions. The first-step of the identification procedure in particular does not fully depend on the

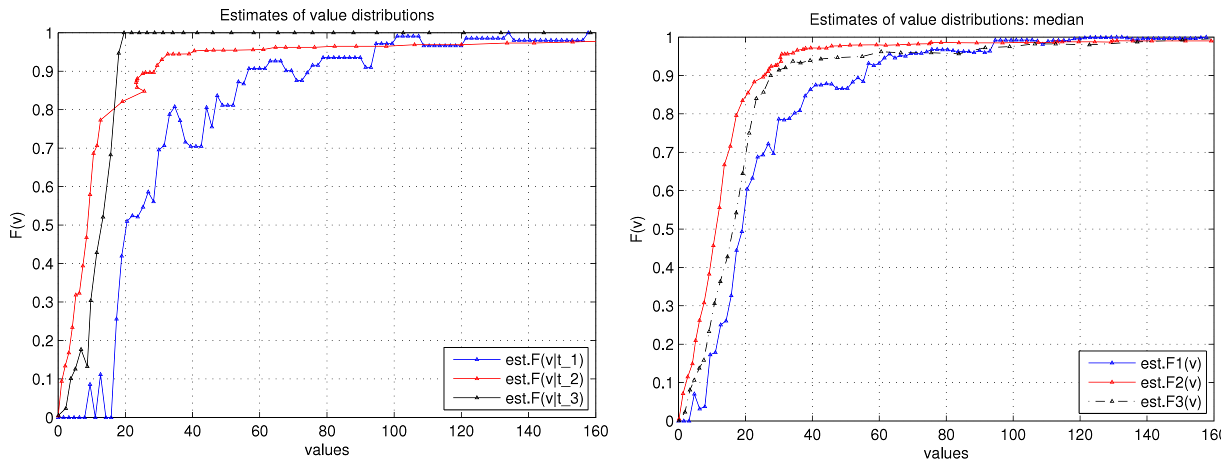

Fig. 9. Estimated value distributions.

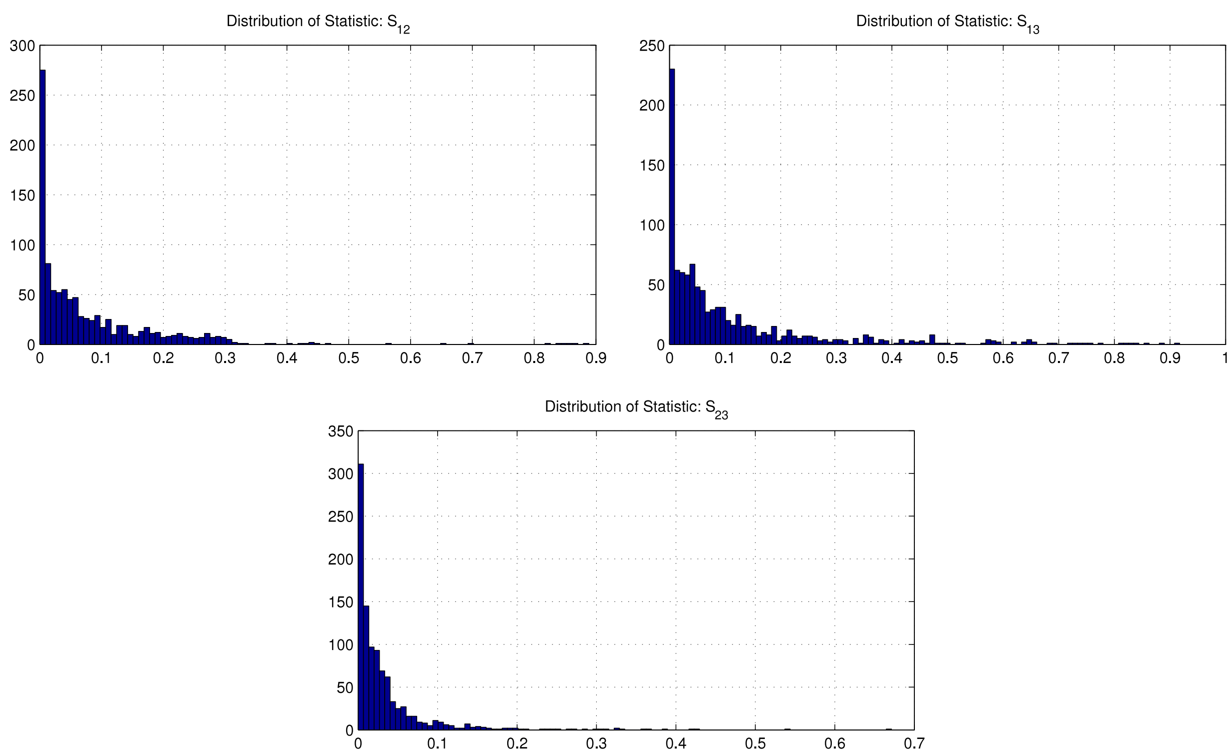

Fig. 10. Histograms of test statistic: "distance" between CDF of values.

structure of auctions. Given that we observe repeated behavior of players in a game, it is possible that we identify their strategies for different types of players. By applying the methodology to other games, the validity of level-k assumptions can be tested across the models. The methodology of identification and estimation proposed in this paper may be further applied to dynamic games. Relaxing assumptions of Nash equilibrium, and assuming bidders are of different levels of sophistication in dynamic games may help us understand the non-equilibrium behavior in dynamic settings. For instance, Aguirregabiria and Magesan (2009) have made the first step to relax the Nash equilibrium assumption in dynamics games. Based on rationalizability assumptions, they proposed methods to identify and estimate a class of dynamic games with incomplete information. The identification and estimation procedures proposed in this paper may be applied to dynamic games and permits us to conduct a nonparametric analysis of players' behavior in dynamic settings.

Appendix

The Appendix includes proofs. Appendix A provides proofs of lemmas and propositions. Appendix B present details of estimation.

Appendix A. Proofs of lemmas and propositions

Proof of Lemma 1. First consider the case where $m = k - 1$. The strict monotonicity of $s_{Lm}(\cdot)$ implies the existence of $s_{Lm}^{-1}(\cdot)$, hence $s_{Lm}^{-1}(s_{Lm}(\bar{v})) = \bar{v}$ holds. For a type-Lk bidder, her highest bid is $s_{Lk}(\bar{v})$. According to (1), the type-Lk bidder's winning probability always less than 1, hence a type-Lk bidder's bidding strategy s_{Lk} must be such that $s_{Lm}^{-1}(s_{Lk}(\bar{v})) \leq \bar{v} = s_{Lm}^{-1}(s_{Lm}(\bar{v}))$. Since $s_{Lm}^{-1}(\cdot)$ is also strictly increasing, we have $s_{Lk}(\bar{v}) \leq s_{Lm}(\bar{v})$.

Please cite this article in press as: An, Y., Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.06.014.

The argument above can be iteratively applied to the cases where $m, k \in \{1, 2, \ldots, K\}$ with $m < k$. As a result, the relationship $s_{Lk}(\bar{v}) \leq \cdots \leq s_{Lm} \leq \cdots \leq s_{L2} \leq s_{L1}(\bar{v})$ holds. This completes the proof. ■

Proof of Lemma 2. Based on Assumption 3, the joint distribution of two discretized bids d_1 and d_3, $g(d_1, d_3)$, can be expressed as follows.

$$g(d_1, d_3) = \sum_{\tau\in\mathcal{K}} g(d_1, d_3, \tau) = \sum_{\tau\in\mathcal{K}} g(d_1|\tau, d_3) g(\tau, d_3) = \sum_{\tau\in\mathcal{K}} g(d_1|\tau) g(\tau, d_3).$$

Rewriting it in matrix form, we obtain

$$B_{d_1,d_3} = B_{d_1|\tau} B_{\tau,d_3}. \tag{A.1}$$

The equation above holds for any integer M ($\geq$2), hence it is feasible to choose a large enough M such that $M \geq m$ even though m is and unknown but finite. Given $M \geq m$, we claim that the rank of both matrices $B_{d_1|\tau}$ and B_{τ,d_3} is equal to the number of types, m. To prove this claim, we assume that (without loss of generality under the assumption that bid distribution of any type is not a linear combination of those for other types) there are two types of bidders: $\tau = L1$ and $\tau = L2$, and both discretized bids d_1 and d_3 take values $1, 2, \ldots, M$, where $M \geq 2$. Then the matrix $B_{d_1|\tau}$ is

$$B_{d_1|\tau} = \begin{pmatrix} \Pr(d_1 = 1|\tau = L1) & \Pr(d_1 = 1|\tau = L2) \\ \Pr(d_1 = 2|\tau = L1) & \Pr(d_1 = 2|\tau = L2) \\ \ldots & \ldots \\ \Pr(d_1 = M|\tau = L1) & \Pr(d_1 = M|\tau = L2) \end{pmatrix}. \tag{A.2}$$

Because each element of the matrix $B_{d_1|\tau}$ is a probability mass, each column sum is one. Consequently, the only possibility that the two columns above are linearly dependent is that $\Pr(d_1 = i|\tau = L1) = \Pr(d_1 = i|\tau = L2)$, for all $i = 1, 2, \ldots, M$.[11] However, this only possibility implies that bid distributions for two types $L1$ and $L2$ are identical, and this contradicts the level-k auction model.[12] Therefore, all the columns of the matrix $B_{d_1|\tau}$ are linearly independent, and the rank of $B_{d_1|\tau}$ is equal to the number of types, m. Analogously, we can prove that all the rows of the matrix B_{τ,d_3} are linearly independent and the rank of this matrix is also equal to m. According to the following inequality regarding the rank of matrix B_{d_1,d_3}:

$$Rank(B_{d_1|\tau}) + Rank(B_{\tau,d_3}) - m \leq Rank(B_{d_1,d_3}) \leq \min\left\{Rank(B_{d_1|\tau}), Rank(B_{\tau,d_3})\right\},$$

We conclude that $Rank(B_{d_1,d_3}) = m$ whenever $M \geq m$. ■

Proof of Proposition 1. We first rewrite the key identification equation (11) for $b_2 = \eta$:

$$B_{\eta,d_1,d_3} B^{-1}_{d_1,d_3} = B_{d_1|\tau} D_{\eta|\tau} B^{-1}_{d_1|\tau}.$$

The eigenvalue–eigenvector decomposition above allows us to identify both the eigenvector matrix $B_{d_1|\tau}$ and the m eigenvalues $g(\eta|\tau = k)$, $k = 1, 2, \ldots, m$ up to unknown matches with τ. For the purpose of identification, the decomposition above needs to be unique. The uniqueness requires that (1) any two of the eigenvalues are distinct; and (2) the eigenvector matrix is normalized. The first requirement satisfies under Assumption 5. To normalize the eigenvector matrix $B_{d_1|\tau}$, we recall that each element in the matrix is a conditional probability, each column of the matrix sums up to one. Therefore, a plausible method of normalization is to divide each column by the corresponding column sum.

The next step is to recover $g(b_2|\tau)$ for all other realizations of b_2. Let $g(b_2, \tau)$ be the joint distribution function of bids b_2 and type τ, and it is a column vector (the dimension is $m \times 1$) with the kth component being $G(b_2, \tau = Lk)$. The joint distribution of b_2 and d_1, $G(b_2, d_1)$, can be defined analogously, and its ith component is $g(b_2, d_1 = i)$. We consider the following linear system:

$$G(b_2, d_1) = \sum_{\tau} G(b_2, d_1, \tau) = \sum_{\tau} G(d_1|\tau) G(b_2, \tau),$$

where the second equation is due to the conditional independence of d_1 and b_2: $G(d_1|\tau, b_2) = G(d_1|\tau)$. The linear system uniquely determines $G(b_2, \tau)$ for any given b_2 because in matrix form $G(d_1|\tau)$ is the invertible matrix $B_{d_1|\tau}$. Analogously, we can obtain another linear system characterizes the relationship between a vector of discretized bid $p(d_1)$ and the type distribution $p(\cdot)$: $p(d_1) = \sum_{\tau} G_{d_1|\tau} p(\tau)$, where both $p(d_1)$ and $p(\tau)$ are $m \times 1$ column vectors with the i-th components are $p(\tau = Li)$ and $p(d_1 = i)$, $i = 1, 2, \ldots, m$ respectively. The type probability $p(\tau)$ can be solved.

Combining the solution $G(b_2, \tau)$ with $p(\tau)$, we are able to obtain the conditional density for each τ, $g(b_2|\tau)$ up to unknown matches with τ. As pointed out in Hu (2008), the method to order the eigenvalues (eigenvectors) is model-specific and there may be multiple ordering conditions from which a researcher can choose. In level-k auctions with unique specification of type-$L0$, Lemma 1 states that the supports of bids are monotonically decreasing in types. Hence we use Lemma 1 to order the m "anonymous" densities $g(b_2|\tau)$ based on their supports.

The proof above readily applies to AVD where the order of m distribution functions can be ordered using the relationship of stochastic dominance implied by Assumption 1. ■

Proof of Proposition 2. For ease of notation, in this proof we use k and j to denote the type Lk and type Lj, respectively. Let $s_k(\cdot)$ and $s_j(\cdot)$ be the bidding strategies for type k and j, respectively. The strict monotonicity of $s_k(\cdot)$ implies that for any bid $b \in [\underline{v}, s_k(\bar{v})]$ there exists a unique value $v_1 \in [\underline{v}, \bar{v}]$ such that $s_k(v_1) = b$. Moreover, since $s_j(\cdot)$ is strictly increasing and the relationship $s_k(\bar{v}) \leq s_j(\bar{v})$ holds, then for each given $b = s_k(v_1)$ there exists unique value $v_2 \in [\underline{v}, \bar{v}]$ such that $s_k(v_1) = s_j(v_2) = b$, where the value v_2 may be greater than or less than the value v_1. Recall that the relationship between a bid b and the corresponding value v_1 for bidders of type k is

$$v_1 = b + \frac{1}{I-1} \frac{F(s_j^{-1}(s_k(v_1)))}{f(s_j^{-1}(s_k(v_1)))} \frac{1}{\frac{ds_j^{-1}(t)}{dt}|_{t=s_k(v_1)}},$$

where $s_k(v_1) = b$. The existence of a value v_2 implies that the equation above can be rewritten as follows.

$$\begin{aligned} v_1 &= b + \frac{1}{I-1} \frac{F(s_j^{-1}(s_k(v_1)))}{f(s_j^{-1}(s_k(v_1)))} \frac{1}{\frac{ds_j^{-1}(t)}{dt}|_{t=s_k(v_1)}} \\ &= b + \frac{1}{I-1} \frac{F(s_j^{-1}(s_j(v_2)))}{f(s_j^{-1}(s_j(v_2)))} \frac{1}{\frac{ds_j^{-1}(t)}{dt}|_{t=s_j(v_2)}} \\ &= b + \frac{1}{I-1} \frac{F(v_2)}{f(v_2)} \frac{1}{\frac{ds_j^{-1}(t)}{dt}|_{t=s_j(v_2)}}. \end{aligned} \tag{A.3}$$

[11] Suppose the two columns are linearly dependent and the coefficient is λ, i.e., $\Pr(d_1 = i|\tau = L1) = \lambda \Pr(d_1 = i|\tau = L2)$ for all $i = 1, 2, \ldots, M$. We add all the elements in the first column and obtain $\sum_{i=1}^M \Pr(d_1 = i|\tau = L1) = \lambda \sum_{i=1}^M \Pr(d_1 = i|\tau = L2) = 1$ since $\sum_{i=1}^M \Pr(d_1 = i|\tau = L2) = 1$, the only possible value for λ is one.

[12] In principle, two different continuous distributions may have the same discretization. However, we can always avoid this possibility by adopting a different methodology of discretization for two distributions.

According to $b = s_j(v_2)$, changing of variables yields

$$G(b|Lj) = F(v_2)$$
$$g(b|Lj) = f(v_2)\frac{ds_j^{-1}(t)}{dt}|_{t=s_j(v_2)}.$$

Combining these two equations above yields

$$\frac{G(b|Lj)}{g(b|Lj)} = \frac{F(v_2)}{f(v_2)}\left(\frac{ds_j^{-1}(t)}{dt}|_{t=s_j(v_2)}\right)^{-1}. \tag{A.4}$$

Substituting the right-hand side of (A.4) into (A.3), we obtain the following relationship between a type-k bidder's bid b, and her value v_1 as follows. v_1 and b can be rewritten as:

$$v_1 = b + \frac{1}{I-1}\frac{G(b|Lj)}{g(b|Lj)},$$

and the equation holds for any $b \in [\underline{v}, s_k(\bar{v})]$. This completes the proof. ■

Proof of Corollary 1. We prove the identification of $Z(\cdot)$ by seeking a contradiction: Suppose that $Z(\cdot)$ is not identified from $s_{L1}^{-1}(\cdot)$, then there exists at least another CDF $U(\cdot)$ with density $u(\cdot) > 0$ on the support $[\underline{v}, \bar{v}]$ such that $s_{L1}^{-1}(b) = b + U(b)/[(I-1)u(b)]$ holds. Therefore we obtain $U(b)/u(b) = Z(b)/z(b)$ for all $b \in [\underline{v}, s_{L1}(\bar{v})]$. Equivalently $u(b)/U(b) = z(b)/Z(b)$ except at $b = \underline{v}$ where $U(\underline{v}) = Z(\underline{v}) = 0$. Integrating both sides of the equation from any bid $\zeta \in (\underline{v}, s_{L1}(\bar{v}))$ to $s_{L1}(\bar{v})$, we have $\ln U(b) = \ln Z(b)$, and consequently $U(b) = Z(b)$ for $b \in [\zeta, s_{L1}(\bar{v})]$. Let the bid ζ converges to the lower bound of bid $\underline{b} = \underline{v}$, and consider that $U(\cdot) = Z(\cdot) = 0$ when $b = \underline{v}$, we conclude that $U(b) = Z(b)$ for all bid $b \in [\underline{v}, s_{L1}(\bar{v})]$, i.e., type-$L1$ bidders' belief is identified.

In fact, the distribution function $Z(\cdot)$ can be explicitly solved from (14) and (15) as follows.[13]

$$Z(b) = \begin{cases} \exp\{-\frac{1}{I-1}\int_b^{\bar{b}} \frac{d\mu}{F^{-1}(G(\mu|L1)) - \mu}\}, & b \in (\underline{b}, \bar{b}], \\ 0, & b = \underline{b}. \end{cases} \quad ■$$

Appendix B. Estimation of bid distribution and type distribution

The estimation procedure starts from discretization of bids and inference of number of types, m. There are uncountably many different methods of discretization that are consistent with (9) and the resulting matrices vary with the methods. However, given (10) holds, the methods of discretization do not affect our estimation asymptotically. In practice, those methods that lead to smaller condition number of the matrices are preferred because we need to take inverse of matrices in estimation. A related issue is to choose the number of discretization, $M \geq m$ when m is unknown. Since the number of types m is unknown to the researcher, it is impossible to provide a deterministic relationship between m and sample size. Therefore, there is a lack of theoretical guidance to choose the "large enough M". Nevertheless, m cannot exceed the number of bidders in the sample, which is finite. Hence one can always start from an M equals to the sample size and discrete two bids b_1, b_3. If $det(B_{d_1,d_3}) = 0$, then move to $M-1$. We can repeat the procedure and stop at M^* such that the $M^* \times M^*$ matrix B_{d_1,d_3} is full rank, then $m = M^*$. In practice, it is not necessary to start from a very large M because all the experimental evidence shows that m is not greater than three (Crawford and Iriberri 2007).

The estimation of bid distribution $G(b|\tau)$ is based on the main equation (11), in which the matrix $B_{d_1|\tau}$ is identical for all bids b_2. Therefore integrating the equation over b_2 helps maximize the rate of convergence in estimating the eigenvector matrix $B_{d_1|\tau}$. For this purpose, we take expectation over b_2 on both sides of (10), and define the two resulting matrices B_{Eb_2,d_1,d_3}, and $D_{Eb_2|\tau}$ as follows.

$$B_{Eb_2,d_1,d_3} \equiv [\mathbb{E}(b_2|d_1 = i, d_3 = j)\, g(d_1 = i, d_3 = j)]_{i,j},$$
$$D_{Eb_2|\tau} \equiv \text{diag}\Big[\mathbb{E}(b_2|\tau = L1) \quad \mathbb{E}(b_2|\tau = L2) \quad \dots \quad \mathbb{E}(b_2|\tau = Lm)\Big].$$

Substituting the matrices B_{b_2,d_1,d_3}, and $D_{b_2|\tau}$ in (10) by B_{Eb_2,d_1,d_3}, and $D_{Eb_2|\tau}$, respectively, we obtain

$$B_{Eb_2,d_1,d_3} = B_{d_1|\tau} D_{Eb_2|\tau} B_{\tau,d_3}.$$

Post-multiplying both sides of the equation above by $B_{d_1,d_3}^{-1} = B_{\tau,d_3}^{-1} B_{d_1|\tau}^{-1}$ yields an integrated version of the key equation (11)

$$B_{Eb_2,d_1,d_3} B_{d_1,d_3}^{-1} = B_{d_1|\tau} D_{Eb_2|\tau} B_{d_1|\tau}^{-1},$$

where the two matrices on the left-hand side are estimated as follows.

$$\widehat{B}_{Eb_2,d_1,d_3} = \left(\frac{1}{N}\sum_{n=1}^{N} b_{n2}\mathbf{1}(d_{n1} = i, d_{n3} = j)\right)_{i,j},$$
$$\widehat{B}_{d_1,d_3} = \left(\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}(d_{n1} = i, d_{n3} = j)\right)_{i,j},$$

where b_{n2} is the second bid of bidder n, and d_{n1}, d_{n3} are the first and the third discretized bids of bidder n, respectively.

According to the properties of eigenvalue–eigenvector decomposition, the eigenvector matrix $B_{d_1|\tau}$ can be expressed by a non-stochastic analytic function $\phi(\cdot)$ of the left-hand side (see Hu, 2008 for detailed discussions on the properties of the mapping $\phi(\cdot)$), and therefore $B_{d_1|\tau}$ may be estimated as

$$\widehat{B}_{d_1|\tau} := \phi\left(\widehat{B}_{Eb_2,d_1,d_3}\widehat{B}_{d_1,d_3}^{-1}\right). \tag{B.1}$$

Since the type τ is associated with the discretized bids through the matrix of eigenvector $\widehat{B}_{d_1|\tau}$, both bid distribution $G(\cdot|\tau)$ and the corresponding density $g(\cdot|\tau)$ can be estimated using matrix $\widehat{B}_{d_1|\tau}$. We exemplify the procedure by estimating the empirical cumulative distribution function $G(b|\tau)$ in the following.

To estimate $G(b|\tau)$, let $G(b_2, \tau)$ be the joint distribution function of bids b_2 and type τ, and it is a column vector (the dimension is $m \times 1$) with the kth component being $G(b_2, \tau = Lk)$. The joint distribution of b_2 and d_1, $G(b_2, d_1)$, can be defined analogously, and its ith component is $G(b_2, d_1 = i)$. The vector $G(b_2, \tau)$ is associated with $G(b_2, d_1)$ according to following equation $G(b_2, d_1) = B_{d_1|\tau} G(b_2, \tau)$. Then $G(b_2, \tau)$ is estimated as

$$\widehat{G}(b_2, \tau) = \widehat{B}_{d_1|\tau}^{-1}\widehat{G}(b_2, d_1),$$

where $\widehat{G}(b_2, d_1)$ is estimated from the sample as follows.

$$\widehat{G}(b_2, d_1 = i) = \frac{1}{N}\sum_{n=1}^{N}\mathbf{1}(b_{n2} < b_2, d_{n1} = i).$$

Analogously, we can obtain a relationship between a vector of discretized bid $p(d_1)$ and the type distribution $p(\cdot)$: $p(d_1) = B_{d_1|\tau} p(\tau)$, where both $p(d_1)$ and $p(\tau)$ are $m \times 1$ column vectors with the i-th components are $p(\tau = Li)$ and $p(d_1 = i)$, $i = 1, 2, \dots, m$ respectively. A natural estimator of $p(\tau)$ is

$$\widehat{p}(\tau) = \widehat{B}_{d_1|\tau}^{-1}\widehat{p}(d_1),$$

where the ith component of $p(d_1)$ is estimated as

$$\widehat{p}(d_1 = i) = \frac{1}{N}\sum_{n=1}^{N}\mathbf{1}(d_{n1} = i), \quad i = 1, 2, \dots, m.$$

[13] This differential equation presents a singularity at $b = \underline{b}$. The existence of a solution of such an equation is addressed in many papers, e.g., see Lebrun (1999).

Please cite this article in press as: An, Y., Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.06.014.

With the estimated joint distribution $\widehat{G}(b_2, \tau)$ and type distribution $\widehat{p}(\tau)$, the empirical CDF of bids for type-Lk is estimated as follows.

$$\widehat{G}(b_2|Lk) = \frac{\widehat{G}(b_2, Lk)}{\widehat{p}(Lk)} = \frac{e_k^T \widehat{B}_{d_1|\tau}^{-1} \widehat{G}(b_2, d_1)}{e_k^T \widehat{B}_{d_1|\tau}^{-1} \widehat{p}(d_1)}; k = 1, 2, \ldots, m, \quad \text{(B.2)}$$

where e_k is defined as a unit column vector with kth component is one and all other components are zeros. Analogously, the bid density $g(b|\tau)$ can be estimated as

$$\widehat{g}(b_2|Lk) = \frac{\widehat{g}(b_2, Lk)}{\widehat{p}(Lk)} = \frac{e_k^T \widehat{B}_{d_1|\tau}^{-1} \widehat{g}(b_2, d_1)}{e_k^T \widehat{B}_{d_1|\tau}^{-1} \widehat{p}(d_1)}, \quad \text{(B.3)}$$

where for every $d_1 \in \{1, 2, \ldots, m\}$, $\widehat{g}(b_2, d_1)$ is a kernel estimate of $g(b_2, d_1)$:

$$\widehat{g}\,(b_2, d_1 = i) = \frac{1}{Nh} \sum_{n=1}^{N} K\left(\frac{b_2 - b_{n2}}{h}\right) \mathbf{1}(d_1 = i),$$

with $K(\cdot)$ and h being a standard kernel function and bandwidth, respectively. Under mild conditions, $\widehat{G}(b|\tau)$ and $\widehat{g}(b|\tau)$ are both consistent and asymptotically normal. The additional assumptions and proofs are contained in the online Appendix.

In the estimation procedure above, the type distribution $p(\tau)$ is estimated as $\widehat{p}(\tau) = \widehat{B}_{d_1|\tau}^{-1}\widehat{p}(d_1)$. Even though this estimator $\widehat{p}(\tau)$ is uniformly consistent (the proof is in the appendix), it is not efficient because the N observations of bids are pooled and only m discretized values are used for the estimation. Hence we may employ an alternative estimator of $p(\tau)$ that uses all the N observations of bids in this subsection. The m components of the type distribution $p(\tau)$ are coefficients of the linear system $G(b) = \sum_{\tau \in \mathcal{K}} G(b|\tau)p(\tau), \quad b \in [\underline{b}, \bar{b}]$, with both $G(b)$ and $G(b|\tau)$ being recovered from data. The relationship above holds for all observed bids, and this permits us to define two column vectors $\overrightarrow{G}(b) = \left[G(b_1)\ \ G(b_2)\ \ \ldots\ \ G(b_N)\right]^T$, $\overrightarrow{p} = \left[p(\tau = L1)\ \ p(\tau = L2)\ \ \ldots\ \ p(\tau = Lm)\right]^T$, and a $N \times m$ matrix Σ is the entry at lth row and kth column being $G(b_i|Lk), i = 1, 2, \ldots, N$. Using $\overrightarrow{G}$, $\overrightarrow{p}$ and T, the linear system can be rewritten as $\overrightarrow{G} = \Sigma \overrightarrow{p}$. The least squares estimate of $\overrightarrow{p}$ is $\widehat{\overrightarrow{p}} = \left(\widehat{\Sigma}'\widehat{\Sigma}\right)^{-1}\widehat{\Sigma}'\widehat{\overrightarrow{G}}$, where $\left(\widehat{\Sigma}'\widehat{\Sigma}\right)^{-1}\widehat{\Sigma}'$ is the Moore–Penrose inverse of $\widehat{\Sigma}$. Entries of $\widehat{\Sigma}$ are obtained from the first step estimation and the elements of $\widehat{\overrightarrow{G}}$ are estimated from the sample as $\widehat{G}(b_i) = \frac{1}{N}\sum_{n=1}^{N} \mathbf{1}(b_{nt} < b_i)$.

B.1. Estimation of value distribution

We estimate the value distribution $F(\cdot)$ from estimated bid distribution $G(\cdot|\tau)$ using the relationship $G(\cdot|Lk) = F\left(\xi_{Lk}(\cdot, G(\cdot|Lj), I)\right)$, where $\xi_{Lk}(b, G(b|Lj), I)$ is the pseudovalue corresponding to bid b for bidders of type Lk. Proposition 2 provides a convenient approach to compute $\xi_{Lk}(\cdot)$

$$\xi_{Lk}(b, G(b|Lj), I) = b + \frac{1}{I-1}\frac{G(b|Lj)}{g(b|Lj)},$$

where $Lk, Lj \in \mathcal{K}, k \geq 2$ and $j = k-1$. Therefore a natural estimator of the pseudovalue $\xi_{Lk}(\cdot)$ is

$$\widetilde{\xi}_{Lk}(b, G(b|Lj), I) = b + \frac{1}{I-1}\frac{\widehat{G}(b|Lj)}{\widehat{g}(b|Lj)}, \quad b \in [\underline{b}_{Lk}, \bar{b}_{Lk}]. \quad \text{(B.4)}$$

The estimator $\widetilde{\xi}_{Lk}$ is similar to the estimator of pseudovalue in GPV except that in level-k auctions the unknown supports of bid b are generally different for different types, while all the bids are i.i.d. in GPV.

To estimate the m different unknown supports of bids, we first recall that all the supports share the same lower bound, which is the minimum bid $\underline{b}$. Therefore, the lower bound of bid, $\underline{b}$, can be estimated as

$$\widehat{\underline{b}} = \inf\left\{b_{it}, i = 1, 2, \ldots, N, t = 1, 2, 3\right\}. \quad \text{(B.5)}$$

The finite support of value distribution $F(\cdot)$ and the regularity condition (5) guarantee that the support of bid distribution $G(\cdot|Lk)$ is also finite. Hence, $\widehat{\underline{b}}$ is a sample minimum, i.e., $\widehat{\underline{b}} = \min\left\{b_{it}\right\}$. The vast literature on order statistics shows that for any conditional distribution function $G(b|Lk)$, $\widehat{\underline{b}}$ is a uniformly consistent estimator of $\underline{b}$. However, to further analyze the asymptotic properties of $\widehat{\underline{b}}$, we have to consider the fact that the sample $\{b_{it}, i = 1, \ldots, N; t = 1, 2, 3\}$ is drawn from the mixed distribution function $G(b) = \sum_{\tau \in \mathcal{K}} G(b|\tau)p(\tau)$. In the Appendix, we follow AL-Hussaini and El-Adll (2004) and Sreehari and Ravi (2010) to show that in the case of a mixture model, the normalized random variable $(\widehat{\underline{b}} - \underline{b})/w(n)$ is asymptotically distributed according to $H(b; 1)$, where $w(n)$ (n is the sample size) is a sequence related to $G(b|Lk)$ and $H(b; 1)$ is the Weibull distribution function with shape parameter 1.

The upper bounds of bids $\bar{b}_{Lk}$ for type-Lk differ across types, thus m upper bounds need to be estimated separately. Lemma 1 states that these upper bounds have the following relationship: $\bar{b}_{L1} \geq \bar{b}_{L2} \geq \cdots \geq \bar{b}_{Lm}$, which implies that the sample maximum $\widehat{\bar{b}} = \sup\left\{b_{it}, i = 1, 2, \ldots, N, t = 1, 2, 3\right\}$ can only be used to estimate the upper bound of bids for bidders of type-$L1$, i.e., $\bar{b}_{L1}$. The asymptotic properties of $\widehat{\bar{b}}$ are similar to that of $\widehat{\underline{b}}$, and it is shown in the Appendix that $(\widehat{\bar{b}} - \bar{b})/\varpi(n) \xrightarrow{d} H_2(b)$ where $H_2(b) \equiv e^b\mathbf{1}(b < 0) + \mathbf{1}(b \geq 0)$, and the sequence $\varpi(n)$ is dependent on $G(b|L1)$, or equivalently, dependent on the value distribution $F(\cdot)$ and the bidding strategy of type $L1$, $s_{L1}(\cdot)$. All the other $m-1$ upper bounds of bid distributions cannot be estimated directly from the sample. Instead, we estimate them according to the conditional bid distribution $G(b|\tau)$, which is estimated in the first-step:

$$\widehat{\bar{b}}_{Lk} = \inf\left\{b : \widehat{G}(b|Lk) = 1, Lk \in \mathcal{K}, k > 1\right\}.$$

The $m - 1$ estimators are uniformly consistent, and the proof is contained in the online Appendix.[14]

Due to the boundary effects of kernel estimator $\widehat{g}(\cdot)$, $\widehat{G}(\cdot)/\widehat{g}(\cdot)$ is an asymptotically biased estimate of $G(\cdot)/g(\cdot)$ at the boundaries. Thus the estimate $\widetilde{\xi}_{Lk}(b, G(b|Lj), I)$ introduced in (B.4) may be biased on the boundaries. To overcome this problem, we adopt the trimming method introduced in GPV described as follows.[15]

$$\widehat{\xi}_{Lk}(b, G(b|Lj), I) = \begin{cases} \widetilde{\xi}_{Lk}(b, G(b|Lj), I) & \text{if } b_{min} + \dfrac{\rho h}{2} \leq b_{it} \leq b_{max} - \dfrac{\rho h}{2}, \\ +\infty & \text{otherwise}, \end{cases} \quad \text{(B.6)}$$

where b_{min} and b_{max} are the minimum and the maximum of the observed bids, ρ and h are the length of the support of the kernel function $K(\cdot)$ and the bandwidth, respectively. We show in the online Appendix that so-defined pseudo private value $\widehat{\xi}_{Lk}$ converges uniformly to the true value.

Using the estimates of bid CDF $\widehat{G}(\cdot|Lk)$, bid density $\widehat{g}(\cdot|Lk)$, and the inverse of bidding strategy $\widehat{\xi}_{Lk}$, the value distribution $F(\cdot)$ can

[14] The estimators of boundaries above may be improved by more subtle methods as treated in the literature, e.g., Cuevas and Rodríguez-Casal (2004) and the reference there. In the present paper, consistent estimators for both the lower and the upper bound of bids are enough for me to illustrate the methodology of estimation, hence we will not further investigate other estimators.

[15] There is a vast literature on removing boundary effects of kernel estimation. The commonly used techniques are reflection of data (Silverman, 1996); transformation of data (Marron and Ruppert, 1994); pseudo-data methods (Cowling and Hall, 1996)); boundary kernel methods (Zhang and Karunamuni, 1998).

Please cite this article in press as: An, Y., Identification of first-price auctions with non-equilibrium beliefs: A measurement error approach. Journal of Econometrics (2017), http://dx.doi.org/10.1016/j.jeconom.2017.06.014.

be estimated by

$$\widehat{F}(\cdot) = \widehat{G}\big(\widehat{\xi}_{Lk}^{-1}(\cdot)|Lk\big), k = 2, \ldots, m, \tag{B.7}$$

where the support of $F(\cdot)$ is estimated as $[\widehat{\underline{b}}, \widehat{\overline{\xi}}]$.

The asymptotic properties of our estimator are analyzed in detail in the online Appendix, and we provide a brief summary here. It is worth noting that even though our estimators are conditional on the type τ, the asymptotic properties of CDFs and PDFs are similar to the existing results in the literature. The CDF $G(b|Lk)$ is estimated from (11) where the L.H.S. is constructed using a sample average. Considering that d_1, d_3 and τ are all discrete, we can achieve $\sqrt{N}$-rate for $\widehat{G}(b|Lk)$. The bid density $g(b|Lk)$ can be analogously estimated by kernel estimation. We show in Appendix that $\sqrt{Nh}[\widehat{g}(b|Lk) - g(b|Lk)]$ converges to a normal distribution, pointwise in b. These results permit us to show the convergence of the estimated pseudovalues $\widehat{\xi}_{Lk}$ and the value distribution $\widehat{F}(\cdot)$: $\sup|\widehat{\xi}_{Lk} - \xi_{Lk}| = \sup|\widehat{F}(\cdot) - F(\cdot)| = O(h^2 + \frac{1}{\sqrt{Nh}})$.

Appendix C. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jeconom.2017.06.014.

References

Aguirregabiria, V., Magesan, A., 2009. Estimation of Dynamic Discrete Games when Players Beliefs are not in Equilibrium. Manuscript, Department of Economics, Univeristy of Toronto.
AL-Hussaini, E., El-Adll, M., 2004. Asymptotic distribution of normalized maximum under finite mixture models. Statist. Probab. Lett. 70 (1), 109–117.
An, Y., 2015. Online appendix to "Identification of First-price Auctions with Non-Equilibrium Beliefs: A Measurement Error Approach". http://www.yonghongan.org/uploads/2/8/2/0/28204339/onlineappendix.pdf.
An, Y., Hu, Y., Liu, P., 2015. Estimating Private Provision of Public Goods with Heterogenous Participants: A Structural Analysis. Cemmap Working Papers. CWP 18/15.
An, Y., Hu, Y., Shum, M., 2010. Estimating first-price auctions with an unknown number of bidders: A misclassification approach. J. Econometrics 157, 328–341.
Aradillas-Lopez, A., Tamer, E., 2008. The identification power of equilibrium in simple games. J. Bus. Econom. Statist. 26 (3), 261–283.
Athey, S., Haile, P.A., 2002. Identification of standard auction models. Econometrica 70 (6), 2107–2140.
Athey, S., Levin, J., 2001. Information and competition in US forest service timber auctions. J. Polit. Econ. 109 (2), 375–417.
Athey, S., Levin, J., Seira, E., 2011. Comparing open and sealed bid auctions: Evidence from timber auctions. Quart. J. Econ. 126 (1).
Bajari, P., Houghton, S., Tadelis, S., 2014. Bidding for incomplete contracts: An empirical analysis of adaptation costs. Amer. Econom. Rev. 104, 1288–1319.
Baldwin, L., 1995. Risk Aversion in Forest Service Timber Auctions. Working Paper. RAND Corporation.
Baldwin, L., Marshall, R., Richard, J., 1997. Bidder collusion at forest service timber sales. J. Polit. Econ. 105 (4), 657–699.
Battigalli, P., Siniscalchi, M., 2003. Rationalizable bidding in first-price auctions. Games Econom. Behav. 45 (1), 38–72.
Camerer, C., Ho, T., Chong, J., 2004. A cognitive hierarchy model of games. Quart. J. Econ. 119 (3), 861–898.
Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.
Campo, S., Guerre, E., Perrigne, I., Vuong, Q., 2011. Semiparametric estimation of first-price auctions with risk-averse bidders. Rev. Econom. Stud. 78 (1), 112–147.
Costa-Gomes, M., Crawford, V., 2006. Cognition and behavior in two-person guessing games: An experimental study. Amer. Econom. Rev. 96 (5), 1737–1768.
Costa-Gomes, M., Crawford, V., Broseta, B., 2001. Cognition and behavior in normal-form games: an experimental study. Econometrica 69 (5), 1193–1235.
Cowling, A., Hall, P., 1996. On pseudodata methods for removing boundary effects in kernel density estimation. J. R. Stat. Soc. Ser. B Stat. Methodol. 58 (3), 551–563.
Crawford, V., Iriberri, N., 2007. Level-*k* auctions: Can a nonequilibrium model of strategic thinking explain the winner's curse and overbidding in private-value auctions? Econometrica 75 (6), 1721–1770.
Cuevas, A., Rodríguez-Casal, A., 2004. On boundary estimation. Adv. Appl. Probab. 36 (2), 340–354.
Dette, H., Neumeyer, N., 2001. Nonparametric analysis of covariance. Ann. Statist. 29 (5), 1361–1400.
Dohmen, T., Falk, A., Huffman, D., Sunde, U., 2010. Are risk aversion and impatience related to cognitive ability? Amer. Econom. Rev. 100 (3), 1238–1260.
Georganas, S., 2011. English auctions with resale: An experimental study. Games Econom. Behav. 73 (1), 147–166.
Georganas, S., Healy, P., Weber, R., 2015. On the persistence of strategic sophistication. J. Econom. Theory 159, 369–400.
Gillen, B., 2009. Identification and Estimation of Level-*k* Auctions. Manuscript, Department of Economics, UCSD.
Goldfarb, A., Xiao, M., 2011. Who thinks about the competition? Managerial ability and strategic entry in us local telephone markets. Amer. Econom. Rev. 101 (7), 3130–3161.
Guerre, E., Perrigne, I., Vuong, Q., 2000. Optimal nonparametric estimation of first-price auctions. Econometrica 68 (3), 525–574.
Haile, P., 1996. Auctions with Resale Markets: An Application to U.S. Forest Service Timber Sales. Working Paper, University of Wisconsin-Madison.
Haile, P., 2001. Auctions with resale markets: An application to U.S. forest service timber sales. Amer. Econom. Rev. 91 (3), 399–427.
Haile, P., Hong, H., Shum, M., 2003. Nonparametric Tests for Common Values in First-price Auctions. NBER Working Paper #10105.
Haile, P., Tamer, E., 2003. Inference with an incomplete model of english auctions. J. Polit. Econ. 111, 1–51.
Hu, Y., 2008. Identification and estimation of nonlinear models with misclassification error using instrumental variables: A general solution. J. Econometrics 144, 27–61.
Hu, Y., McAdams, D., Shum, M., 2013. Identification of first-price auctions with non-separable unobserved heterogeneity. J. Econometrics 174 (2), 186–193.
Hu, Y., Schennach, S., 2008. Instrumental variable treatment of nonclassical measurement error models. Econometrica 76, 195–216.
Ivanov, A., Levin, D., Niederle, M., 2010. Can relaxation of beliefs rationalize the winner's curse?: An experimental study. Economtrica 78 (4), 1435–1452.
Krasnokutskaya, E., 2011. Identification and estimation of auction models with unobserved heterogeneity. Rev. Econom. Stud. 78 (1), 293–327.
Lebrun, B., 1999. First price auctions in the asymmetric *N* bidder case. Internat. Econom. Rev. 40 (1), 125–142.
Li, T., Perrigne, I., Vuong, Q., 2000. Conditionally independent private information in ocs wildcat auctions. J. Econometrics 98, 129–161.
Li, T., Vuong, Q., 1998. Nonparametric estimation of the measurement error model using multiple indicators. J. Multivariate Anal. 65 (2), 139–165.
Marron, J., Ruppert, D., 1994. Transformations to reduce boundary bias in kernel density estimation. J. R. Stat. Soc. Ser. B Stat. Methodol. 56 (4), 653–671.
Melnykov, V., Maitra, R., 2010. Finite mixture models and model-based clustering. Statist. Surv. 4, 80–116.
Nagel, R., 1995. Unraveling in guessing games: An experimental study. Amer. Econom. Rev. 85 (5), 1313–1326.
Robin, J.-M., Smith, R.J., 2000. Tests of rank. Econom. Theory 16 (02), 151–175.
Shneyerov, A., Wong, A.C.L., 2011. Identification in first-price and dutch auctions when the number of potential bidders is unobservable. Games Econom. Behav. 72 (2), 574–582.
Silverman, B.W., 1996. Density Estimation for Statistics and Data Analysis. Chapman & Hall/CRC.
Sreehari, M., Ravi, S., 2010. On extremes of mixtures of distributions. Metrika 71 (1), 117–123.
Stahl, D., Wilson, P., 1994. Experimental evidence on players' models of other players. J. Econ. Behav. Org. 25 (3), 309–327.
Zhang, S., Karunamuni, R., 1998. On kernel density estimation near endpoints. J. Statist. Plann. Inference 70 (2), 301–316.