Chapter 60

# NONPARAMETRIC APPROACHES TO AUCTIONS *

SUSAN ATHEY

*Harvard University, USA*

*NBER, USA*

*e-mail:* *athey@fas.harvard.edu*

PHILIP A. HAILE

*Yale University, USA*

*NBER, USA*

*e-mail:* *philip.haile@yale.edu*

## Contents



* We have benefited from the input of Jaap Abbring, Estelle Cantillon, Ali Hortaçsu, Elena Krasnokutskaya, Jonathan Levin, Harry Paarsch, Isabelle Perrigne, Martin Pesendorfer, Joris Pinkse, Rob Porter, Unjy Song, Elie Tamer, an anonymous referee, and students at Stanford, University College London, and Yale. Ying Fan, Dan Quint and Gustavo Soares provided capable research assistance. Research support from the National Science Foundation (grants SES-0112047 and SES-0351500) and the Alfred P. Sloan Foundation is gratefully acknowledged. Any conclusions, findings, or opinions are those of the authors and do not necessarily reflect the views of any funding organization.

*Handbook of Econometrics, Volume 6A*

*DOI: 10.1016/S1573-4412(07)06060-6*

## Abstract

This chapter discusses structural econometric approaches to auctions. Remarkably, much of what can be learned from auction data can be learned without restrictions beyond those derived from the relevant economic model. This enables us to take a nonparametric perspective in discussing how the structure of auction models can be combined with observables to uncover (or test hypotheses about) primitives of interest in auction markets. We focus on first-price sealed-bid and ascending auctions, including extensions to Dutch auctions, Internet auctions, multi-unit auctions, and multi-object auctions. We consider a wide range of underlying structures of bidder demand and information, as well as a variety of types of data one may encounter in applications. We discuss identification and testable restrictions of these models and present a variety of estimation approaches.



## Keywords

auctions, identification, estimation, testing

*JEL classification*: C5, C14, D44

## 1. Introduction

Auctions provide opportunities for economists to examine field data from markets that can involve rich strategic interaction and asymmetric information while nonetheless being simple enough that the salient forces can be convincingly captured by a tractable economic model. The primitives of any strategic model include the set of players, the information structure, the rules of play, and players' objectives. In auction markets, one can often describe these key elements with an unusually high degree of confidence. Consequently, auctions have been at the center of efforts to combine economic theory with econometric analysis to understand behavior and inform policy.

Early work by Hendricks and Porter (1988) and others played an important role in demonstrating the empirical relevance of private information and the ability of strategic models to predict behavior. More recently, there has been a great deal of attention to econometric approaches to auctions that incorporate restrictions from economic theory as assumptions of an econometric model.[1] The goal of this structural approach is to address questions that can only be answered with knowledge concerning the distribution functions that characterize the underlying demand and information structure. Structural empirical work on auctions has examined, for example, the division of rents in auctions of public resources, whether reserve prices in government auctions are adequate, the effects of mergers on procurement costs, whether changes in auction rules would produce greater revenues, whether bundling of procurement contracts is efficient, the value of seller reputations, the effect of information acquisition costs on bidder participation and profits, whether bidders' private information introduces adverse selection, and whether firms act as if they are risk averse.

Many of these questions have important implications well beyond the scope of auctions themselves. In all of economics there is a tradeoff between the assumptions one relies on and the questions one can address. Because an auction is a market institution that is particularly easy to capture with a theoretical model, one may have more confidence than usual that imposing significant structure from economic theory in interpreting data can be useful. Combined with the fact that private information and strategic behavior are paramount in auctions, this suggests that auctions may enable economists to get at questions of importance to many other types of markets.

Remarkably, much of what can be learned from auction data can be learned without restrictions beyond those derived from economic theory. In particular, identification often does not depend on unverifiable parametric distributional assumptions. This is important: although economics can determine or at least shape the specification of many components of an empirical model, it rarely provides guidance on distribution functions governing unobservables.

[1] A seminal paper in this literature is Paarsch (1992a), which builds on insights in Smiley (1979) and Thiel (1988).

Our focus in this chapter is on structural econometric approaches to auctions, with an emphasis on nonparametric identification. This focus should not be confused with a presumption that nonparametric *estimation* methods are always preferred. Approximation methods are virtually always needed for estimation in finite samples, and parametric estimators will be most appropriate in some applications. However, as emphasized at least since Koopmans (1945), the question of what the observables and the assumed underlying structure are capable of revealing (i.e., the identification question) is fundamentally distinct from the choice of statistical methods used in practice for estimation. When identification holds nonparametrically, one can be confident that estimates have valid interpretations as finite sample approximations and are not merely artifacts of unverifiable maintained assumptions. Equally important for our purpose, because a discussion of nonparametric identification makes clear how the structure of a model and the observables enable (or, in some important cases, fail to enable) estimation, it also provides an ideal perspective for discussing recent developments in empirical approaches to auctions. Our goals are to describe key insights from a wide range of recent work in this area in a unified framework, to present several new results, and to point out areas ripe for exploration.

We focus on two auction formats that are dominant in practice: first-price sealed-bid auctions and ascending (or "English") auctions. First-price auctions are particularly prevalent in government procurement – a common source of data in applied work. Our discussion of first-price auctions will include the closely related Dutch auction. Ascending auctions, in several variations, are the most frequently observed in practice. They are widely used in sales of antiques, art, timber, and in Internet auctions.[2] As we will see below, each type of auction presents different econometric challenges. We will also examine extensions to other environments, including multi-unit and multi-object auctions. We consider a wide range of underlying structures of bidder demand and information, as well as a range of types of data one may encounter in applications. We discuss identification, testable restrictions, and a variety of parametric, semi-parametric, and nonparametric estimation approaches. Much of the recent innovation in the literature has been on the identification question. In many cases this is because standard statistical methods can be applied for estimation and testing once identification results are obtained. This is not always the case, however, and in some cases the development of methods for estimation has lagged development of identification results. Here and elsewhere, our discussion will point to a number of opportunities for additional work.

[2] Among the auction forms commonly discussed in the theoretical literature, we exclude the second-price sealed-bid ("Vickrey") auction, which is closely related to the ascending auction but uncommon in practice. Some Internet auctions, like those on the eBay site, use a system of proxy bidding that has the flavor of a second-price sealed bid auction, although in practice bidders usually have the ability to observe and respond to the bids of at least some of their opponents, as in an ascending auction (see Lucking-Reiley (2000) for more stylized facts about Internet auctions). Alternative models of Internet auctions are offered by Bajari and Hortaçsu (2003a), Ockenfels and Roth (2006) and Song (2003). We will discuss a structural empirical model based on the last of these in Section 6.3.4.

Before proceeding, we first make precise what is meant by identification in this context. Let $\mathbb{G}$ denote the set of all joint distributions over a specified set of observable random variables. Define a *model* as a pair $(\mathbb{F}, \Gamma)$, where $\mathbb{F}$ is a set of joint distributions over a specified set of latent random variables and $\Gamma$ is a collection of mappings $\gamma : \mathbb{F} \to \mathbb{G}$. In this chapter, $\mathbb{F}$ will typically be the set of joint distributions of bidder valuations and information ("types") satisfying certain statistical properties (e.g., independence, symmetry, etc.), while $\Gamma$ will consist of a single mapping from the true distribution of types to a distribution of bids implied by the assumption of Bayesian Nash equilibrium. Implicit in the specification of a model is the assumption that it contains the true structure $(\mathcal{F}, \gamma)$ generating the observables. A model is said to be identified (or identifiable) if the observables uniquely determine the true structure within $(\mathbb{F}, \Gamma)$.

DEFINITION 1.1. A model $(\mathbb{F}, \Gamma)$ is *identified* if for every $(\mathcal{F}, \widetilde{\mathcal{F}}) \in \mathbb{F}^2$ and $(\gamma, \tilde{\gamma}) \in \Gamma^2$, $\gamma(\mathcal{F}) = \tilde{\gamma}(\widetilde{\mathcal{F}})$ implies $(\mathcal{F}, \gamma) = (\widetilde{\mathcal{F}}, \tilde{\gamma})$.

In some cases, useful inferences can be made even when a model is not identified. In *partially identified* models one may be able to identify some components of interest, or place bounds on components of interest [see, e.g., Manski (1995)]. A separate question is whether a model places refutable restrictions on observables; i.e., whether the model is testable. A model is testable if some joint distributions in $\mathbb{G}$ cannot be generated by the model.

DEFINITION 1.2. A model $(\mathbb{F}, \Gamma)$ is *testable* if $\bigcup_{\gamma \in \Gamma,\, \mathcal{F} \in \mathbb{F}} \gamma(\mathcal{F})$ is a strict subset of $\mathbb{G}$.

With these definitions in hand, we can preview some of the themes that emerge in what follows. First, a remarkable number of positive nonparametric identification results can be obtained by exploiting the relationships between observables and the primitives of interest that are implied by economic theory. Richer statistical structures (e.g., arbitrary correlation) for bidders' information and/or more limited sets of observables (e.g., only the winning bid) create greater challenges, but even here a number of positive results can be obtained. There are limits to the positive results, however. For example, identification of models with common values, risk aversion, or unobserved heterogeneity can be obtained only with strong *a priori* restrictions. This is particularly the case in ascending auctions, where theory provides less guidance on the appropriate interpretation of the observed bids.

A second major theme is the need to make modeling choices and the importance of testing these choices when possible. Often a particular set of assumptions (e.g., independent private values, risk neutral bidders) is postulated for particular application based on characteristics of the relevant market. Modeling choices can have important implications for the conclusions one reaches. Ideally, researchers would like to combine economic justifications for modeling choices with statistical evidence supporting these choices and/or an analysis of the range of outcomes possible under alternative assumptions. We discuss a number of results that clarify when this will be possible. In many cases some assumptions can be tested while maintaining others. In general this possibility depends on the auction format (e.g., ascending versus first-price) and the data configuration (e.g., whether all bids are observed or just the winning bid, or whether particular types of exogenous variation are present and observable). For example, an assumption of private values is testable under some data configurations but not others (Section 8). Another example arises when the econometrician must make modeling choice regarding the source of observed correlation among bids: this may result from correlation of bidders' private information or from common knowledge among bidders of auction-specific factors affecting all bidders' valuations. These alternative models often have different implications for counterfactuals, but it is difficult to distinguish between them empirically (Section 6.1.2). Other examples include choices of how participation is modeled (Section 6.3), and of bidders' risk preferences (Section 6.4). Typically, some modeling choices will be testable (particularly in data configurations that include some type of exogenous variation in the environment – e.g., in the number of bidders or bidder covariates), while others will not.

Our focus in this chapter unavoidably leads us to ignore many interesting and important issues given attention in the empirical auctions literature. Fortunately, there are now several excellent surveys, each with a somewhat different focus, that provide useful complements to our chapter. Laffont (1997) and Hendricks and Paarsch (1995) provide early surveys reviewing empirical studies of the implications of equilibrium bidding in auctions as well as approaches to estimation of the primitives of auction models. Perrigne and Vuong (1999) survey methods for structural analysis of first-price auctions, including a synthesis of their own extensive contributions (with several coauthors) to nonparametric identification and estimation of these models. Hong and Shum (2000) provide an introduction to parametric structural approaches. Kagel (1995) surveys the extensive work on auctions in the experimental economics literature. Finally, Hendricks and Porter (in press) provide a recent and extensive review of the large empirical literature on auctions, covering a wide range of economic questions and econometric approaches.[3]

The structure of the chapter can be described as follows. Section 2 describes the underlying theoretical framework for our initial focus and provides the characterizations of equilibrium bidding behavior that underlie the econometric approaches that follow.[4] In Sections 3 and 4 we then discuss first-price and ascending auctions in the simplest and most widely considered case of private values, assuming that there is no binding

[3] Reiss and Wolak (Chapter 64 in this volume) include a discussion of auctions among several examples of the structural empirical approach in industrial organization. See also the recent monograph by Hong and Paarsch (2006).

[4] For additional detail, Krishna (2002) provides an excellent synthesis of a large theoretical literature on auctions. McAfee and McMillan (1987) provide a shorter introduction to much of the relevant theory. Milgrom and Weber (1982) is a central paper in the early theoretical literature that covers many of the models we consider. Milgrom (2004) treats some of the newer literature on combinatorial auctions.

reserve price and that the data available consist of bids from independent auctions of identical goods. These results provide many of the key building blocks for considering richer private values specifications, specification testing, endogenous participation, risk aversion, as well as other types of data in Sections 5 and 6. In Section 7 we take up the case of common values models, where identification is more difficult and often fails. This provides one motivation for a discussion of testing for common values in Section 8. We conclude with two sections on important topics that have been the subject of very recent work. Section 9 addresses dynamics. Section 10 discusses work in progress on multi-unit and multi-object auctions.

## 2. Theoretical framework

### 2.1. Demand and information structures

Throughout we denote random variables in upper case and their realizations in lower case. We use boldface to indicate vectors. To emphasize the distinction between latent variables and observables, we adopt the convention of denoting the cumulative distribution function (CDF) of a latent random variable $\mathbf{Y}$ by $F_{\mathbf{Y}}(\cdot)$ and the CDF of an observable random variable $\mathbf{Y}$ by $G_{\mathbf{Y}}(\cdot)$. Much of the discussion will involve order statistics. We let $Y^{(k:n)}$ denote the $k$th order statistic from the sample $(Y_1, \dots Y_n)$, with $F_Y^{(k:n)}(\cdot)$ denoting the corresponding marginal CDF. We follow the standard convention of indexing order statistics lowest to highest so that, e.g., $Y^{(n:n)}$ is the maximum.

For most of the chapter, the underlying theoretical framework involves the sale of a single indivisible good to one of $n \in \{\underline{n}, \dots, \overline{n}\}$ risk neutral bidders, with $\overline{n} \geqslant \underline{n} \geqslant 2$.[5] Later, when we consider auctions with reserve prices or participation costs, these $n$ bidders will be referred to as "potential bidders." We consider risk aversion, sequential auctions, multi-unit auctions, and multi-object auctions separately below. We let $\mathcal{N}$ denote the set of bidders, although when bidders are symmetric, $n = |\mathcal{N}|$ will be a sufficient statistic. We let $\mathcal{N}_{-i}$ denote the set of competitors faced by bidder $i$. The utility bidder $i \in \{1, \dots, \overline{n}\}$ would receive from the good is $U_i$, which we assume to have common support (denoted $\operatorname{supp} F_{U_i}(\cdot)$ or $\operatorname{supp} U_i$) for all $i$. Often $U_i$ is referred to as $i$'s "valuation." We let $\mathbf{U} = (U_1, \dots, U_n)$.

Bidder $i$'s private information (his "type") consists of a scalar signal $X_i$. We let $\mathbf{X} = (X_1, \dots, X_n)$, $\underline{x}_i = \inf \operatorname{supp} X_i$, and $\overline{x}_i = \sup \operatorname{supp} X_i$. Signals are informative in the sense that the expectation

$$E[U_i \mid X_i = x_i,\ \mathbf{X}_{-i} = \mathbf{x}_{-i}]$$

strictly increases in $x_i$ for all realizations $\mathbf{x}_{-i}$ of $i$'s opponents' signals. Note that because signals play a purely informational role and any monotonic transformation $\theta(X_i)$

[5] Translation to procurement settings, where bidders compete to sell, is straightforward.

contains the same information as $X_i$ itself, the marginal distribution of $X_i$ is irrelevant; i.e., without a normalization on $X_i$, the theoretical model is over-parameterized. It is therefore desirable (and without loss of generality) to impose a normalization such as[6]

$$X_i = E[U_i \mid X_i].$$

We will see below that different normalizations will sometimes turn out be more convenient.

Except where otherwise stated, we assume that the set of bidders and the joint distribution $F_{\mathbf{X},\mathbf{U}}(\cdot; \mathcal{N})$ of bidders' signals and valuations are common knowledge. While these are standard assumptions in the theoretical literature on auctions, in a few cases (e.g., an ascending auction with private values) these assumptions are inconsequential. In a first-price auction, these assumptions can be relaxed somewhat; for example, we consider the possibility that $\mathcal{N}$ is unknown in Section 6.3.3.

This framework is a generalization of that studied in Milgrom and Weber's (1982) influential theoretical exploration of auctions and nests a wide range of special cases, each involving different assumptions about bidders' private information. One key distinction is that between *private values* (PV) and *common values* (CV) models.

DEFINITION 2.1. Bidders have *private values* if $E[U_i \mid X_1 = x_1, \dots, X_n = x_n] = E[U_i \mid X_1 = x_1]$ for all $x_1, \dots, x_n$ and all $i$; bidders have *common values* if $E[U_i \mid X_1 = x_1, \dots, X_n = x_n]$ strictly increases in $x_j$ for all $i$, $j$, and $x_j$.[7]

In private values models, bidders do not have private information about the valuations of their opponents. For the settings we will consider, this is equivalent to assuming bidders know their own valuations ($X_i = U_i$). In a common values model, by contrast, each bidder $i$ would update her beliefs about her valuation $U_i$ if she learned an opponent's signal $X_j$ in addition to her own signal $X_i$. Even in a private values auction a bidder would like to know her competitors' private information for strategic reasons. However, in a common values auction, knowledge of opponents' signals would alter her expectation of her own valuation. This is the characteristic of common values auctions that leads to the "winner's curse." Roughly speaking, winning a common values auction reveals (in equilibrium) to the winner that her signal was more optimistic than those of her opponents. Rational bidders anticipate this information when forming expectations

6 It is important to avoid confusing this extra degree of freedom in the usual specification of the theoretical model with issues concerning econometric identification. Since the marginal distribution of $X_i$ is irrelevant in the theoretical model, it is not a primitive whose identification should even be considered.

7 Alternatively, one might define private and common values in terms of the conditional distributions $F_{U_i}(U_i \mid X_1, \dots, X_n)$ and $F_{U_i}(U_i \mid X_i)$. For our purposes a definition in terms of conditional expectations is adequate. Note that for simplicity of exposition our definition of common values rules out cases where the winner's curse arises for some realizations of types but not others.

of the utility they would receive by winning.[8] Note that common values models incorporate a wide range of structures in which information about the value of the good is dispersed among bidders, not just the special case in which the value of the object is identical for all bidders (defined as *pure common values* below).[9]

A second way in which this general framework can be specialized is through restrictions on the joint distribution of signals. Common assumptions are independence or affiliation.[10] Note that dependence (or affiliation) of signals is neither necessary nor sufficient for common values. Finally, a common restriction in the literature is *symmetry*, i.e., that the joint distribution $F_{\mathbf{X},\mathbf{U}}(X_1, \ldots, X_n, U_1, \ldots, U_n; \mathcal{N})$ is exchangeable in the bidder indices. For clarity, we will often explicitly refer to models as "symmetric" or "asymmetric." Combining these types of restrictions leads to a number of special cases that have been considered in the literature, including:

- *Independent Private Values (IPV)*: private values with $U_i$ independent;
- *Symmetric Independent Private Values*: private values with $U_i$ i.i.d.;
- *Affiliated Private Values (APV):* private values with $(U_1, \ldots, U_n)$ affiliated;
- *Pure Common Values*: common values with $U_i = U_0\ \forall i$;
- *Mineral Rights*: pure common values with signals i.i.d. conditional on $U_0$.

Finally, for a few results we will make an additional assumption of *exogenous variation in the number of bidders*, which holds when variation in the set of bidders is independent of the joint distribution of bidders' valuations and signals.

DEFINITION 2.2. A bidding environment has *exogenous variation in the number of bidders* if $\bar{n} > \underline{n}$ and, for all $\mathcal{N}, \mathcal{N}'$ such that $\mathcal{N} \subset \mathcal{N}' \subseteq \{1, \ldots, \bar{n}\}$, $F_{\mathbf{X},\mathbf{U}}(\cdot; \mathcal{N})$ is identical to the marginal distribution of $\{(U_i, X_i)\}_{i\in\mathcal{N}}$ obtained from $F_{\mathbf{X},\mathbf{U}}(\cdot; \mathcal{N}')$.

### 2.2. *Equilibrium bidding*

We restrict attention to econometric approaches that exploit the structure of equilibrium bidding to obtain identification or testable restrictions. Hence we must first provide the

[8] Note that the presence of the winner's curse does not imply that winners regret winning; rather, the winner's curse refers to the "bad news" [Milgrom (1981)] about the object's value contained the information that one has won the auction. Rational bidders anticipate this.

[9] While our terminology follows, e.g., Klemperer (1999), Athey and Haile (2002), and Haile, Hong and Shum (2003), there is some variation in the terminology used in the auction literature. Early on, the term "common values" was sometimes used in the way we use it but sometimes used to refer to the special case we call "pure common values." Similarly, "affiliated values" was sometimes used for the class of models we call "common values," despite the fact that purely private values can be affiliated (see below). Recently some authors [e.g., Krishna (2002)] have adopted the term "interdependent values" to refer to the broad class of models we refer to as common values models.

[10] The random variables $\mathbf{Y} = (Y_1, \ldots, Y_n)$ with joint density $f_{\mathbf{Y}}(\cdot)$ are affiliated if for all $\mathbf{y}$ and $\mathbf{y}'$, $f_{\mathbf{Y}}(\mathbf{y}\vee\mathbf{y}')f_{\mathbf{Y}}(\mathbf{y}\wedge\mathbf{y}') \geqslant f_{\mathbf{Y}}(\mathbf{y})f_{\mathbf{Y}}(\mathbf{y}')$, where $\vee$ denotes the component-wise maximum, and $\wedge$ the component-wise minimum. See Milgrom and Weber (1982) for additional discussion. Note that affiliation allows independence as a special case.

necessary characterizations of equilibrium. Following the literature, we generally restrict attention to (perfect) Bayesian Nash equilibria in weakly undominated strategies. We focus on equilibrium in pure bidding strategies $\beta_i(\cdot; \mathcal{N})$, $i = 1, \ldots, n$, mapping each bidder's signal (and, implicitly, any public information) into a bid. When bidders are *ex ante* symmetric we further restrict attention to symmetric equilibria, so that $\beta_i(\cdot) = \beta(\cdot)\ \forall i$. Below we discuss conditions under which there are other equilibria in first-price auctions. We will denote a bidder $i$'s equilibrium bid by $B_i$, with $\mathbf{B} = \{B_1, \ldots, B_n\}$. We let $\underline{b}_i = \inf[\operatorname{supp}[B_i]]$ and $\overline{b}_i = \sup[\operatorname{supp}[B_i]]$.

### 2.2.1. First-price auctions

In a first-price sealed-bid auction bidders submit bids simultaneously, and the good is awarded to the high bidder at a price equal to his bid. If there is a reserve price, $r$, the seller has committed to consider only bids of at least $r$. For first-price auctions we make the following additional assumptions:

ASSUMPTION 2.1 *(First-price auction assumptions).*
- (i) For all $i$, $U_i$ has compact, convex support denoted $\operatorname{supp} F_{U_i}(\cdot) = [\underline{u}, \overline{u}]$.
- (ii) The signals $\mathbf{X}$ are affiliated, with $\operatorname{supp} F_{\mathbf{X}}(\cdot) = \times_{i=1}^{n} \operatorname{supp} F_{X_i}(\cdot)$.
- (iii) $F_{\mathbf{X}}(\cdot)$ has an associated joint density $f_{\mathbf{X}}(\cdot)$ that is strictly positive on the interior of $\operatorname{supp} F_{\mathbf{X}}(\cdot)$.

The following result summarizes existence and uniqueness results for this model. This will enable us to then proceed to the key characterization results used for empirical work.

THEOREM 2.1. *Consider the first-price auction.*
- (i) (Existence in strictly increasing strategies) *An equilibrium exists in pure, nondecreasing strategies, where for each $i$,* $\operatorname{supp}[B_i] \subseteq \operatorname{supp}[\max_{j\in\mathcal{N}\setminus i} B_j]$. *In addition, a pure strategy equilibrium in strictly increasing strategies exists in all models except in the CV model with asymmetric bidders and signals that are not independent*; *in the latter case, strategies are strictly increasing except that at most one bidder may bid* $\inf[\operatorname{supp}[B^{(n:n)}]]$ *with positive probability.*[11]
- (ii) (Uniqueness) *In a PV model with either* (a) *independence* (*IPV*), *or* (b) *symmetry, if $f_{\mathbf{X}}(\cdot)$ is continuously differentiable there is a unique equilibrium. This equilibrium is in pure, strictly increasing, and differentiable strategies.*[12]

[11] See Athey (2001) and Reny and Zamir (2004) for existence of equilibrium in nondecreasing strategies, and Milgrom and Weber (1982), McAdams (2007) and Lizzeri and Persico (2000) for the characterization. McAdams (2007) argues that in any monotone equilibrium, strategies are strictly increasing except that at most one bidder may, with positive probability, choose the lowest bid that wins with strictly positive probability, if such a bid exists. In PV auctions it is possible to rule out mass points at the reserve price, if it binds, or at the bottom of the value distribution if the reserve price does not bind.

[12] In the IPV case, all equilibria are in monotone strategies; see Lebrun (1999), Bajari (2001), and Maskin and Riley (2003) for uniqueness results. Milgrom and Weber (1982) show existence of the equilibrium for

(iii) (Uniqueness in monotone class for symmetric models) *If we restrict attention to pure strategy equilibria in nondecreasing strategies, then when bidders are symmetric and $f_{\mathbf{X}}(\cdot)$ is continuously differentiable there is a unique equilibrium, which is in symmetric, strictly increasing, and differentiable strategies.*[13]

All of our positive identification results for common values models rely on symmetry, so in our discussions of CV auctions we will proceed under the assumption that strategies are strictly increasing. For the first-price auction models for which uniqueness has not been established, we will also assume that all observations in a given data set are derived from the same equilibrium.

As shown by Milgrom and Weber (1982), a bidder $i$ participates if and only if his signal exceeds a threshold value

$$x_i^*(\mathcal{N}) = \inf\Big\{x_i\colon\ E\Big[U_i \;\Big|\; X_i = x_i,\ \max_{j\in\mathcal{N}_{-i}} B_j \leqslant r\Big] \geqslant r\Big\}. \tag{2.1}$$

When there is no reserve price, let $x_i^*(\mathcal{N}) = \underline{x}_i$. Here, expectations over others' bids represent equilibrium expectations. A bidder $i$ who has observed signal $X_i = x_i > x_i^*(\mathcal{N})$ solves

$$\max_{\tilde{b}}\Big(E\Big[U_i \;\Big|\; X_i = x_i,\ \max_{j\in\mathcal{N}_{-i}} B_j \leqslant \tilde{b}\Big] - \tilde{b}\Big)\Pr\Big(\max_{j\in\mathcal{N}_{-i}} B_j \leqslant \tilde{b} \;\Big|\; X_i = x_i\Big), \tag{2.2}$$

where we adopt the convention that $B_j < r$ for any bidder $j$ who does not participate.

Define

$$\tilde{v}_i(x_i, m_i; \mathcal{N}) = E\Big[U_i \;\Big|\; X_i = x_i,\ \max_{j\in\mathcal{N}_{-i}} B_j = m_i\Big].$$

This is bidder $i$'s expectation of his valuation conditional on his own signal and the highest competing bid. This highest competing bid is informative because, in equilibrium, bids are strictly increasing in signals. In particular, if we let

$$v_i(x_i, y_i; \mathcal{N}) = E\Big[U_i \;\Big|\; X_i = x_i,\ \max_{j\in\mathcal{N}_{-i}} B_j = \beta_i(y_i; \mathcal{N})\Big] \tag{2.3}$$

APV auctions. McAdams (2007) shows that for a nonmonotone equilibrium to exist, both independence of signals and private values must be relaxed. He shows that with private values or independent signals, all equilibria are outcome-equivalent to a monotone equilibrium; i.e., bidding strategies are identical to those in a monotone equilibrium except possibly for subsets of types whose equilibrium bids win with probability zero. McAdams (2004b) shows that if bidders are symmetric, there is a unique equilibrium within the monotone class. So together, these results imply that for the symmetric PV model, there is a unique equilibrium.

[13] See footnote 12 for a discussion of when nonmonotone equilibria can exist. McAdams (2004b) proves uniqueness within the monotone class. For characterizations, see Milgrom and Weber (1982). See also Lizzeri and Persico (2000), who show that when the density of the value distribution is $C^1$, in two-bidder first-price auctions with a binding reserve price, among monotone pure strategy equilibria there exists a unique equilibrium in strictly increasing, differentiable strategies, except that one bidder may choose the reserve price with positive probability.

then

$$v_i(x_i, y_i; \mathcal{N}) = \tilde{v}_i\big(x_i, \beta_i(y_i; \mathcal{N}); \mathcal{N}\big).$$

The expectation $v_i(x_i, x_i; \mathcal{N})$ will play an important role below. This expectation is taken conditioning both on $i$'s own private information and on the event that $i$'s equilibrium bid is "pivotal," i.e., that infinitesimal deviations from his equilibrium bid would change the outcome of the auction.

Let

$$G_{M_i|B_i}(m_i|b_i; \mathcal{N}) = \Pr\Big(\max_{j\neq i} B_j \leqslant m_i \;\Big|\; B_i = b_i, \mathcal{N}\Big)$$

denote the distribution of the maximum equilibrium bid among $i$'s opponents conditional on $i$'s own equilibrium bid and the set of bidders $\mathcal{N}$. Let $g_{M_i|B_i}(m_i|b_i; \mathcal{N})$ denote the corresponding conditional density, which exists and is positive for all $b_i$ and almost every $m_i$ in the support of $B_i$ under the assumptions outlined above. Note that with strictly increasing equilibrium bidding, conditioning on $\{B_i = b\}$ is equivalent to conditioning on $\{X_i = \beta_i^{-1}(b; \mathcal{N})\}$. Bidder $i$'s bidding Problem (2.2) can then be rewritten

$$\max_{\tilde{b}} \int_{-\infty}^{\tilde{b}} \big[\tilde{v}_i(x_i, m_i; \mathcal{N}) - \tilde{b}\big] g_{M_i|B_i}\big(m_i|\beta_i(x_i; \mathcal{N}); \mathcal{N}\big)\, \mathrm{d}m_i.$$

This objective function is differentiable almost everywhere. Differentiating with respect to $\tilde{b}$, we see that for almost every signal $x_i$ of bidder $i$, a necessary condition for $b_i$ to be an optimal bid (i.e., for $\beta_i(x_i; \mathcal{N}) = b_i$) is

$$v_i(x_i, x_i; \mathcal{N}) = b_i + \frac{G_{M_i|B_i}(b_i|b_i; \mathcal{N})}{g_{M_i|B_i}(b_i|b_i; \mathcal{N})} \equiv \xi_i(b_i; \mathcal{N}). \tag{2.4}$$

Equation (2.4) characterizes an equilibrium bid as equal to the bidder's expectation of his valuation (conditional on being pivotal) less a strategic "markdown" $\frac{G_{M_i|B_i}(b_i|b_i;\mathcal{N})}{g_{M_i|B_i}(b_i|b_i;\mathcal{N})}$.[14] This first-order condition does not always lead to an analytic solution for equilibrium bidding strategies. With *ex ante* symmetric bidders, however, we can write

$$v_i(x, x; \mathcal{N}) = v(x, x; n) = E\Big[U_i \;\Big|\; X_i = \max_{j\neq i} X_j = x\Big]$$

and $x_i^*(\mathcal{N}) = x^*(n)\ \forall i$. In that case, Milgrom and Weber (1982) have shown that the equilibrium bid function has the form

$$\beta(x; n) = rL\big(x^*(n)|x; n\big) + \int_{x^*(n)}^{x} v(t, t; n)\, \mathrm{d}L(t|x; n) \tag{2.5}$$

[14] This is analogous to the markdown of an oligopsonist, which bases its price on the equilibrium elasticity of its residual supply curve; in the auction model, $G_{M_i|B_i}(b_i|b_i; \mathcal{N})$ plays the role of the residual supply curve.

for $x \geqslant x^*$, where

$$L(t|x;n) \equiv \exp\left(-\int_t^x \frac{f_1(z|z;n)}{F_1(z|z;n)}\,\mathrm{d}z\right)$$

and $F_1(\cdot|x;n)$ is the distribution of the maximum signal among a bidder's opponents conditional on the number of bidders and on his own signal being $x$.

Before proceeding, we pause to make two observations about the support of the equilibrium bid distribution.[15]

THEOREM 2.2. *In the IPV model of the first-price auction,* $\operatorname{supp}[B_i]$ *is the same for all* $i$.

PROOF. With independence, the inverse bid function for bidder $i$ can be written

$$\xi_i(b_i;\mathcal{N}) = b_i + \frac{1}{\sum_{k\in\mathcal{N}\setminus i} \frac{g_{B_k}(b_i)}{G_{B_k}(b_i)}}.$$

If there are two bidders, the result is immediate given that the value distributions have the same support. Now suppose $n > 2$ and $\bar{b}_i < \bar{b}_j$. We know that $\xi_j(b_j;\mathcal{N})$ must be continuous at $\bar{b}_i$: otherwise (given strictly monotone strategies) we would contradict our assumption that valuations are drawn from a convex set. Then, note that

$$\bar{u} = \xi_i(\bar{b}_i;\mathcal{N}) = \bar{b}_i + \frac{1}{\sum_{k\in\mathcal{N}\setminus i} \frac{g_{B_k}(\bar{b}_i)}{G_{B_k}(\bar{b}_i)}} < \bar{b}_i + \frac{1}{\sum_{k\in\mathcal{N}\setminus\{i,j\}} \frac{g_{B_k}(\bar{b}_i)}{G_{B_k}(\bar{b}_i)}} = \xi_j(\bar{b}_i;\mathcal{N}).$$

But $\xi_j(\bar{b}_i;\mathcal{N}) > \bar{u}$ contradicts the assumption that $U_i$ has the same support for all $i$. Given the properties established in Theorem 2.1, standard arguments then show that $\operatorname{supp}[B_i] = [\max\{r,\underline{u}\},\bar{b}]\ \forall i$. $\square$

Outside of the IPV model, it is not known in general whether bid distributions have the same support for all bidders when bidders are asymmetric. We do know that if we relax the assumption that valuations have common support, the bids may or may not have the same support.[16]

Note that the theory also implies that the upper bound of the bid distribution is closely related to features of the distribution of valuations. In the symmetric IPV model,

$$U_i = B_i + \frac{G_B(B_i;n)}{(n-1)g_B(B_i;n)}$$

[15] See Lebrun (1999) for an alternative proof.

[16] In Section 5.1 we give an example where valuations have different supports but bids have identical supports. To see an example where bid distributions have different supports, suppose that there are three bidders. $F_{U_1}(u_1) = \frac{8}{5}u_1 - \frac{16}{25}u_1^2$ for $u_1 \in [0,5/4]$, while for $i \in \{2,3\}$, $F_{U_i}(u_i) = \frac{1}{100}(4 + 2u_i - \sqrt{2}\sqrt{8 - 7u_i + 2u_i^2})^2$ for $u_i \in [0,3/2]$ and $F_{U_i}(u_i) = \frac{1}{9}u_i^2$ for $u_i \in [3/2,3]$. For this example, $G_{B_1}(b_1) = 2b_1 - b_1^2$ for $b_1 \in [0,1]$, while for $i \in \{2,3\}$, $G_{B_i}(b_i) = b_i^2/4$ for $b_i \in [0,2]$.

so

$$E[U_i] = E[B_i] + \frac{1}{n-1}\int_{\underline{b}}^{\bar{b}} G_B(b;n)\,\mathrm{d}b = \frac{n-2}{n-1}E[B_i] + \frac{1}{n-1}\bar{b}.$$

Thus, the mean valuation is a linear function of the mean bid and $\bar{b}$. When $n = 2$, this yields $E[U_i] = \bar{b}$. The average "markdown" for a bidder in the symmetric IPV model is given by

$$E[U_i - B_i] = \frac{1}{n-1}\big(\bar{b} - E[B_i]\big).$$

Although it seems that these kinds of relationships might be useful, they have not to our knowledge been explored in the econometric analysis of auctions.

### 2.2.2. *Ascending auctions*

The standard model of an ascending auction is the so-called "clock auction" or "button auction" model of Milgrom and Weber (1982), where the price rises continuously and exogenously. Bidders indicate their willingness to continue bidding continuously as well, for example by raising their hands or depressing a button as the price rises. As the auction proceeds, bidders exit observably and irreversibly (by lowering their hands, releasing their buttons, etc.) until only one bidder remains. This final bidder obtains the good at the price at which his last opponent exited; i.e., the auction ends at a price equal to the second highest exit price ("bid") $b^{(n-1:n)}$.

The participation rule for an ascending auction is identical to that for a first-price auction. An equilibrium bidding strategy specifies a price at which to exit, conditional on one's own signal and on any information revealed by previous exits by opponents. With strictly increasing bidding strategies, the price at which a bidder exits reveals his signal to others. So in a common values auction, an exit causes the remaining bidders to update their beliefs about their valuations; hence, the prices at which bidders plan to exit change as the auction proceeds. In a private values auction there is no such updating, and each bidder has a weakly dominant strategy to bid up to his valuation, i.e.,

$$\beta_i(x_i;\mathcal{N}) = E[U_i \mid X_i = x_i] = x_i \equiv u_i. \tag{2.6}$$

In common values auctions there are multiple equilibria, even with *ex ante* symmetric bidders and restriction to symmetric strictly increasing weakly undominated strategies [Bikhchandani, Haile and Riley (2002)]. In any such equilibrium, however, if $i$ is one of the last two bidders to exit, his exit price $b_i$ is

$$E\big[U_i \bigm| X_i = x_i,\ X_j = x_i\ \forall j \notin \{i \cup \mathcal{E}_i\},\ X_k = x_k\ \forall k \in \mathcal{E}_i\big], \tag{2.7}$$

where $\mathcal{E}_i$ denotes the set of bidders who exit before $i$. Milgrom and Weber (1982) originally identified the equilibrium in which all bidders follow (2.7), which reduces to the weakly dominant strategy (2.6) in the case of private values.

While the Milgrom–Weber model yields a trivial relation between a bidder's valuation and his bid in a private values auction, we will see that even in this case identification can present challenges, due to the fact that the auction ends before the winner bids (exits). Furthermore, in many applications the Milgrom–Weber model may represent too great an abstraction from actual practice, for example if prices are called out by bidders rather than by the auctioneer, or if bidders are free to make a bid at any point in the auction, regardless of their activity (or lack thereof) earlier in the auction. In Section 4.3 we will discuss an econometric approach that relaxes the structure of the button auction model.

## 3. First-price auctions with private values: Basic results

### 3.1. Identification

We begin by considering the case of private values auctions, assuming that bidders' valuations at each auction are draws from the same joint distribution $F_{\mathbf{U}}(\cdot)$. The primitive of interest in a PV auction is this joint distribution: it completely characterizes bidder demand and information. With knowledge of $F_{\mathbf{U}}(\cdot)$ one can, for example, simulate outcomes under alternative market mechanisms, assess efficiency and the division of surplus, and determine an optimal reserve price. The simple idea underlying the structural approach to PV auctions is to use the distribution of bids observed in a sample of auctions along with the equilibrium mapping between valuations and bids (the observables) to learn about $F_{\mathbf{U}}(\cdot)$.

Even when a closed form solution like (2.5) is available, however, it is not immediately clear how one would proceed to use this equilibrium characterization for a first-price auction to obtain identification. Even in the simplest symmetric IPV model, the equilibrium bid function takes the form (recall that $x_i = u_i$)

$$\beta(u; n) = \frac{\int_{-\infty}^{u} t f_U^{(n-1:n-1)}(t)\,\mathrm{d}t}{F_U^{(n-1:n-1)}(u)},$$

which depends on the unknown distribution $F_U(\cdot)$ of valuations, i.e., on the object one would like to estimate.

Several approaches were initially taken to address this problem within the symmetric IPV model. Following Smiley (1979) and Paarsch (1992a), early work focused on parametric specifications of $F_U(\cdot)$ admitting simple closed form equilibrium bid functions that made it feasible to derive likelihoods or moment conditions.[17] Laffont, Ossard and Vuong (1995) proposed an approach combining parametric assumptions with a simulation based estimator that is made feasible in the symmetric IPV framework by the revenue equivalence theorem [e.g., Myerson (1981)]. Bajari (1997) proposed a Bayesian

[17] Smiley (1979) considered only common values models.

approach applicable in the more difficult case of asymmetric independent private values. The role of the parametric distributional assumptions in these empirical approaches was not initially clear.

An important breakthrough due to Guerre, Perrigne and Vuong (2000) came from the simple but powerful observation that equilibrium is attained when each player is acting optimally against the distribution of behavior by opponents.[18] When bids are observable, both the distribution of opponent behavior and the optimal (equilibrium) action of each bidder are observable, enabling identification of the latent joint distribution of bidder valuations under fairly weak restrictions. In particular, the first-order condition (2.4) can be written

$$u_i = b_i + \frac{G_{M_i|B_i}(b_i|b_i;\mathcal{N})}{g_{M_i|B_i}(b_i|b_i;\mathcal{N})}. \tag{3.1}$$

Thus, each bidder's latent private value can be expressed as a functional of his equilibrium bid and the joint distribution of the competing equilibrium bids he faces.[19] In fact, the function $\xi_i(b_i,\mathcal{N}) \equiv b_i + \frac{G_{M_i|B_i}(b_i|b_i;\mathcal{N})}{g_{M_i|B_i}(b_i|b_i;\mathcal{N})}$ is the inverse of bidder $i$'s equilibrium bid function, the mapping needed to infer valuations from bids. Since the joint distribution of bids is observable, identification of each private value $u_i$ (and, therefore, of the joint distribution $F_{\mathbf{U}}(\cdot)$) follows directly from (3.1). Formally,

$$F_{\mathbf{U}}(\mathbf{u}) = G_{\mathbf{B}}\big(\xi_1^{-1}(u_1,\mathcal{N}),\dots,\xi_n^{-1}(u_n,\mathcal{N})\big). \tag{3.2}$$

This proves the following identification result, combining results from Guerre, Perrigne and Vuong (2000), Li, Perrigne and Vuong (2002), and Campo, Perrigne and Vuong (2003).

THEOREM 3.1.
(i) *Suppose all bids are observed in first-price sealed-bid auctions. Then the symmetric affiliated private values model is identified.*
(ii) *Suppose all bids and bidder identities are observed in first-price sealed-bid auctions. Then the asymmetric affiliated private values model is identified.*

### *3.2. Estimation*

For purposes of estimation, suppose one observes bids from independent auctions $t = 1,\dots,T$. We will add an auction index $t$ to the notation above as necessary. For

[18] This approach was first described in print by Laffont and Vuong (1993), who attribute the idea to an early draft of Guerre, Perrigne and Vuong (2000).

[19] Note that in general this kind of approach relies on there being a unique equilibrium or on an assumption that the equilibrium selected is the same across observations. Otherwise the observed distribution of opponent bids would be a mixture of those in each equilibrium, and would not match the distribution characterizing a bidder's beliefs in a given auction.

example, $b_{it}$ will denote the realized bid of bidder $i$ at auction $t$. Let $T_{\mathcal{N}}$ denote the number of auctions in which $\mathcal{N}$ is the set of bidders. We let $T_n = \sum_{\mathcal{N}:\,|\mathcal{N}|=n} T_{\mathcal{N}}$. We assume that for all $n = \underline{n} \ldots \bar{n}$, $T_n \to \infty$ and $T \to \infty$. When we consider asymmetric settings, we consider only sets $\mathcal{N}$ for which $T_{\mathcal{N}} \to \infty$.

A two-step estimation procedure can be employed, closely following the identification result in Theorem 3.1. In the first step, estimates of each $\frac{G_{M_i|B_i}(b_{it}|b_{it};\mathcal{N})}{g_{M_i|B_i}(b_{it}|b_{it};\mathcal{N})}$ are obtained from the observed bids. These estimates are then used with Equation (3.1) to construct estimates of each latent valuation $u_i$. This pseudo-sample of valuations (often referred to as a sample of "pseudo-values") is then treated as a sample from the true distribution $F_{\mathbf{U}}(\cdot)$, subject to first-stage estimation error.

In principle each step could be parametric or nonparametric. As noted by Perrigne and Vuong (1999), a challenge in a fully parametric method is the need for internal consistency between the parametric families chosen for the distributions of bids and of valuations, since these are related by the equilibrium bid function. This issue would be avoided if only one of the two steps were treated parametrically. Jofre-Bonet and Pesendorfer (2003) and Athey, Levin and Seira (2004) follow this approach, motivated by a desire to include covariates in a parsimonious way.[20] Fully parametric methods based on maximum likelihood or moment conditions (rather than the two-step "indirect" approach discussed here) have been explored by, e.g., Paarsch (1992a, 1992b), Donald and Paarsch (1993, 1996), and Laffont, Ossard and Vuong (1995). In practice the applicability of these methods has been limited to distributional families leading to simple closed forms for equilibrium bid functions and/or to the symmetric independent private values setting. As first explored by Donald and Paarsch (1993), a violation of a standard regularity condition for maximum likelihood estimation arises in a first-price auction, leading to nonstandard asymptotic distributions [see also Donald and Paarsch (1996), Chernozhukov and Hong (2003), and Hirano and Porter (2003)].

Below we describe the fully nonparametric estimators that have thus far been proposed in the literature.[21]

### 3.2.1. *Symmetric bidders*

Consider first the case of symmetric bidders, where $G_{M_i|B_i}(b|b;\mathcal{N})$ can be written $G_{M|B}(b|b;n)$ $\forall i$. Following Li, Perrigne and Vuong (2002), let

$$G_{M,B}(m,b;n) \equiv G_{M|B}(m|b;n)g_B(b;n)$$

[20] Note, however, that theory predicts that bid distributions should have compact support. To be consistent with theory, an upper bound on the support of the bid distributions should be incorporated in estimation.

[21] Thus far, the literature has focused on kernel estimators. One possible alternative is sieve estimation [e.g., Chen (2007)]. As we discuss below, such an approach might have a practical advantage in environments with observed auction heterogeneity.

and

$$g_{M,B}(m; b; n) \equiv g_{M|B}(m|b; n)g_B(b; n)$$

where $g_B(\cdot)$ is the marginal density of a bidder's equilibrium bid, given the number of bidders $n$. Note that here we depart from our usual notational convention, since $G_{M,B}(\cdot)$ is not the joint distribution of $(M, B)$ but its derivative with respect to its second argument. Let

$$\widehat{G}_{M,B}(b, b; n) = \frac{1}{nT_n h_G} \sum_{t=1}^{T} \sum_{i=1}^{n} K\left(\frac{b - b_{it}}{h_G}\right) \mathbf{1}\{m_{it} < b, n_t = n\}, \tag{3.3}$$

$$\hat{g}_{M,B}(b, b; n) = \frac{1}{nT_n h_g^2} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbf{1}\{n_t = n\} K\left(\frac{b - b_{it}}{h_g}, \frac{b - m_{it}}{h_g}\right), \tag{3.4}$$

where $M_{it}$ denotes the maximum of $i$'s opponents' bids at auction $t$, $K(\cdot)$ is a kernel, and $h_G$ and $h_g$ are appropriately chosen bandwidth sequences. Under standard conditions, $\widehat{G}_{M,B}(b, b; n)$ and $\hat{g}_{M,B}(b, b; n)$ are consistent estimators of $G_{M,B}(b, b; n)$ and $g_{M,B}(b; b; n)$. Noting that

$$\frac{G_{M,B}(b, b; n)}{g_{M,B}(b, b; n)} = \frac{G_{M|B}(b|b; n)}{g_{M|B}(b|b; n)}$$

we see that $\frac{\widehat{G}_{M,B}(b,b;n)}{\hat{g}_{M,B}(b,b;n)}$ is a consistent estimator of $\frac{G_{M|B}(b|b;n)}{g_{M|B}(b|b;n)}$. Equation (3.1) then implies that

$$\hat{u}_{it} \equiv b_{it} + \frac{\widehat{G}_{M,B}(b_{it}, b_{it}; n)}{\hat{g}_{M,B}(b_{it}, b_{it}; n)}$$

is a consistent estimate of the latent valuation $u_{it}$ that generated the observed bid $b_{it}$.

Naively treating each $\hat{u}_{it}$ as a draw from $F_{\mathbf{U}}(\cdot)$ might suggest a kernel density estimator of the form

$$\hat{f}_{\mathbf{U}}(u_1, \ldots, u_n) = \frac{1}{T_n h_f^n} \sum_{t=1}^{T} K_f\left(\frac{u_1 - \hat{u}_{1t}}{h_f}, \ldots, \frac{u_n - \hat{u}_{nt}}{h_f}\right) \mathbf{1}\{n_t = n\},$$

where $K_f(\cdot)$ is a multivariate kernel and $h_f$ is a bandwidth. Li, Perrigne and Vuong (2002, Proposition 2) show that with bandwidths $h_G$, $h_g$, and $h_f$ that vanish at appropriate rates, under standard smoothness conditions $\hat{f}_{\mathbf{U}}(\cdot)$ is in fact a uniformly consistent estimator of $f_{\mathbf{U}}(\cdot)$ on any inner compact subset of its support. The restriction to the region of support away from the boundaries follows from the usual problem of asymptotic bias at the boundaries with kernel estimates.

Li, Perrigne and Vuong (2002, pp. 180–181) suggest triweight kernels (using products of univariate kernels for the multivariate kernels) and a standard rule of trimming the pseudo-values associated with bids within one bandwidth of either boundary of the bid data. The most important practical question is the choice of bandwidth. Guerre, Perrigne and Vuong (2000) and Li, Perrigne and Vuong (2002) suggest following Silverman's (1986) "rule of thumb." To our knowledge, data driven bandwidth selection procedures have not been explored. Guerre, Perrigne and Vuong (2000) also point out that the assumption of exchangeability can be imposed by averaging $\hat{f}_{\mathbf{U}}(u_1, \ldots, u_n)$ over all permutations of the bidder indices. When there is exogenous variation in the number of bidders, it may be useful to further exploit this restriction by optimally combining information from auctions with different numbers of bidders. As we discuss in more detail below, the overidentifying exchangeability restriction or exogenous variation in participation can also serve as a basis for specification testing.

An important but largely unresolved question is the asymptotic distribution of the estimator $\hat{f}_{\mathbf{U}}(\cdot)$. The challenge is to appropriately account for the estimation error arising from the first-stage estimation of the markdown component of the equilibrium bid functions [Guerre, Perrigne and Vuong (2000)]. Of course, one is often interested in confidence intervals on an estimate of some functional of $f_{\mathbf{U}}(\cdot)$, rather than on $\hat{f}_{\mathbf{U}}(\cdot)$ itself. For example, the goal of the empirical exercise may be to determine optimal selling procedures, to assess efficiency, or to describe how valuations are affected by various factors. For the symmetric case, Haile, Hong and Shum (2003) have shown that the estimates $\hat{u}_{it}$ themselves have asymptotic normal distributions, as do all fixed quantiles (and many other functionals) of their empirical distribution. In practice, a bootstrap procedure has sometimes been applied for inference on these functionals of $F_{\mathbf{U}}(\cdot)$ or others expected to have a normal limiting distribution [e.g., Hendricks, Pinkse and Porter (2003), Haile, Hong and Shum (2003), Krasnokutskaya (2004)]. Outside the IPV model, a block bootstrap is used, reflecting the assumption that auctions are independent, whereas bids may be correlated within an auction. In particular, to construct one bootstrap sample of bids for a given value of $n$, auction indices $s$ are sampled with replacement from the set $\{t\colon\ n_t = n\}$. All bids from each selected auction $s$ are then included in the bootstrap sample. Haile, Hong and Shum (2003) have also explored the use of subsampling.

In the special case of (symmetric) independent private values, the joint distribution $F_{\mathbf{U}}(\cdot)$ is a product of identical marginal distributions, $F_U(\cdot)$, and the first order condition (3.1) simplifies to

$$u = b + \frac{G_B(b; n)}{(n-1) g_B(b; n)}, \tag{3.5}$$

where $G_B(\cdot; n)$ is the marginal distribution of equilibrium bids in auctions with $n$ bidders, and $g_B(\cdot; n)$ is the associated density. Because $G_B(\cdot; n)$ and $g_B(\cdot; n)$ are univariate functions, this simplifies estimation. Let

$$\widehat{G}_B(b; n) = \frac{1}{nT_n} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbf{1}\{b_{it} \leqslant b,\ n_t = n\},$$

$$\hat{g}_B(b; n) = \frac{1}{nT_n h_g} \sum_{t=1}^{T} \sum_{i=1}^{n} K\left(\frac{b - b_{it}}{h_g}\right) \mathbf{1}\{n_t = n\},$$

$$\hat{u}_{it} = b_{it} + \frac{\widehat{G}_B(b_{it}; n_t)}{(n_t - 1)\hat{g}_B(b_{it}; n_t)},$$

where $K(\cdot)$ is a kernel (satisfying standard conditions) and $h_g$ is an appropriately chosen bandwidth sequence.[22] Guerre, Perrigne and Vuong (2000) show that with appropriately chosen bandwidth sequence $h_f$, one then obtains a uniformly consistent estimator of $f_U(\cdot)$ from the kernel density estimator

$$\hat{f}_U(u) = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{n_t}\sum_{i=1}^{n_t}\frac{1}{h_f}K\left(\frac{u - \hat{u}_{it}}{h_f}\right).$$

### 3.2.2. Asymmetric bidders

Extending the approach above to the case of asymmetric bidders is straightforward, but more data intensive. With symmetric bidders, estimation of the distribution of opposing bids (and the markdown term $\frac{G_{M,B}(b|b;n)}{g_{M,B}(b|b;n)}$ this distribution implies) is performed separately for each value of $n$. This reflects the fact that variation in $n$ changes the distribution of the maximum opposing bid and, therefore, the equilibrium bidding strategy that is inverted to recover private values from the observed bids. With asymmetric bidders, variation in the *identities* of opposing bidders can have a similar effect, even when the *number* of opponents is held constant. Depending on the nature of bidder asymmetries, different approaches will be taken, although the general principle is clear: to estimate the markdown $\frac{G_{M_i|B_i}(b_i|b_i;\mathcal{N}_t)}{g_{M_i|B_i}(b_i|b_i;\mathcal{N}_t)}$ for a bidder $i$ in auction $t$, the relevant sample is the set of auctions $s$ in which $G_{M_i|B_i}(\cdot|\cdot;\mathcal{N}_s) = G_{M_i|B_i}(\cdot|\cdot;\mathcal{N}_t)$.

In the most general case, each bidder is allowed to draw her valuation from a different distribution and each set of bidders $\mathcal{N}$ is treated separately. Again let $M_{it}$ denote the maximum bid among $i$'s opponents at auction $t$. Letting $T_{\mathcal{N}_t}$ denote the number of auctions in which the set of bidders is $\mathcal{N}_t \ni i$, one could let

$$\widehat{G}_{M,B}(b_{it}, b_{it}; \mathcal{N}_t) = \frac{1}{T_{\mathcal{N}_t} h_G}\sum_{s=1}^{T} K\left(\frac{b_{it} - b_{is}}{h_G}\right)\mathbf{1}\{m_{is} < b_{it},\ \mathcal{N}_s = \mathcal{N}_t\},$$

$$\hat{g}_{M,B}(b_{it}, b_{it}; \mathcal{N}_t) = \frac{1}{T_{\mathcal{N}_t} h_g^2}\sum_{s=1}^{T}\mathbf{1}\{\mathcal{N}_s = \mathcal{N}_t\}K\left(\frac{b_{it} - b_{is}}{h_g}\right)K\left(\frac{b_{it} - m_{is}}{h_g}\right)$$

[22] See Guerre, Perrigne and Vuong (2000) for details. They also propose kernel smoothing over the different values of $n$ in estimating each $G_B(\cdot; n)$ and $g(\cdot; n)$ rather than the pure "binning" approach described here. Asymptotically there is no difference and, since $N$ is discrete, kernel smoothing is a generalization. In finite sample, kernel smoothing that is not equivalent to binning will utilize bids from auctions with $n' \neq n$ bidders to estimate the markdown $\frac{G_B(b;n)}{(n-1)g_B(b;n)}$ in (3.5). Whether this is desirable will depend on the data available, although we are not aware of a careful analysis of this question.

and

$$\hat{u}_{it} = b_{it} + \frac{\widehat{G}_{M,B}(b_{it}, b_{it}; \mathcal{N}_t)}{\hat{g}_{M,B}(b_{it}, b_{it}; \mathcal{N}_t)}$$

to obtain consistent estimators under standard conditions. In practice, however, this approach may require a great deal of data, since many observations will be needed for each set $\mathcal{N}$ considered.

In some cases, one may be able to categorize bidders into a smaller set of heterogeneous classes, assuming exchangeability within each class. This structure can lead to significant practical advantages, as it allows use of substantially more data for each estimated pseudo-value. For example, Campo, Perrigne and Vuong (2003) studied "wildcat" auctions for mineral extraction rights on the US outer-continental shelf, where bids may come from "solo" bidders (a single firm) or "joint" bidders (more than one firm, legally bidding as one).[23] This leads them to consider the case of two classes of bidders, I and II. The first-order condition for a class-I bidder in an auction in which the set of bidders is $\mathcal{N}$ can be written

$$u_{\mathrm{I}} = b_{\mathrm{I}} + \frac{G^{\mathrm{I}}_{M,B}(b_{\mathrm{I}}, b_{\mathrm{I}}; \mathcal{N})}{g^{\mathrm{I}}_{M,B}(b_{\mathrm{I}}, b_{\mathrm{I}}; \mathcal{N})}. \tag{3.6}$$

Define the relation $=^{\mathrm{I,II}}$ such that $\mathcal{N}_t =^{\mathrm{I,II}} \mathcal{N}$ holds iff $\mathcal{N}_t$ and $\mathcal{N}$ have the same number of bidders, $n^{\mathrm{I}}$ and $n^{\mathrm{II}}$, from each class. Let $T^{\mathrm{I,II}}_{\mathcal{N}} = \sum_{t=1}^{T} \mathbf{1}\{\mathcal{N}_t =^{\mathrm{I,II}} \mathcal{N}\}$. Now $G^{\mathrm{I}}_{M,B}(b_{\mathrm{I}}, b_{\mathrm{I}}; \mathcal{N})$ can be estimated by

$$\begin{aligned}&\widehat{G}^{\mathrm{I}}_{M,B}(b, b; \mathcal{N})\\ &\quad = \frac{1}{T^{\mathrm{I,II}}_{\mathcal{N}} \times h_G \times n^{\mathrm{I}}} \sum_{s=1}^{T} \sum_{i=1}^{|\mathcal{N}|} K\left(\frac{b - b_{is}}{h_G}\right) \mathbf{1}\{m_{is} < b,\ \mathcal{N}_s =^{\mathrm{I,II}} \mathcal{N},\ i \in \text{class I}\}.\end{aligned}$$

Analogous adjustments are made to an estimator for $g^{\mathrm{I}}_{M,B}(b, b; \mathcal{N})$ and to the first-order condition for a class-II bidder [see Campo, Perrigne and Vuong (2003, pp. 186–187) for details]. Note in particular that in estimating $\frac{G^{\mathrm{I}}_{M,B}(b_{\mathrm{I}}, b_{\mathrm{I}}; \mathcal{N})}{g^{\mathrm{I}}_{M,B}(b_{\mathrm{I}}, b_{\mathrm{I}}; \mathcal{N})}$ one can use data from all auctions $t$ with $\mathcal{N}_t =^{\mathrm{I,II}} \mathcal{N}$. Furthermore, the sample of bids is cut less finely across bidders than in the completely general case.

Note that we have treated asymmetries as resulting from differences in the distributions from which bidders draw unobservables. In some cases, it may be more natural that asymmetries arise instead from observable covariates $Z_i$ that are idiosyncratic to each bidder – e.g., distance to a construction site [e.g., Bajari (1997), Flambard and

[23] Athey, Levin and Seira (2004) provide another example, treating loggers and sawmills as two different classes of bidders at timber auctions.

Perrigne (2006)]. Conditional on having the same value of the covariates, bidder valuations may still be exchangeable. Without further restriction, this is similar to the case in which bidders fall into discrete categories; indeed, it is exactly the same if the covariates are discrete. In the case of continuous covariates, standard smoothing techniques would lead to similar approaches for estimating the joint distribution $F_{\mathbf{X},\mathbf{U}}(\cdot|Z_1, \ldots, Z_n)$. In Section 6.2.1 we will see how the presence of bidder-specific covariates can actually aid identification in some cases.

### 3.3. Incomplete bid data and Dutch auctions

#### 3.3.1. Independent private values

The results above exploited the assumed observability of all bids from each auction. In some applications, however, not all bids are available. For example, for some auctions only the transaction price $B^{(n:n)}$ is recorded. One example is a Dutch auction, where the auctioneer starts with a very high price and lowers it continuously until one bidder is willing to take the good at the current price. Although a Dutch auction is seemingly different from a first-price sealed-bid auction, the two formats are strategically equivalent (assuming the same information is observable prior to bidding).[24] Since a Dutch auction ends as soon as the winner makes his bid, only the winning bid can be observed. We will see that in some cases the winning bid is sufficient for identification. In other environments, only a partial set of bids may be available. For example, in a procurement setting, the buyer might retain information regarding the best losing bid in case the auction winner defaults. Viewed somewhat differently, identification results for the case of incomplete bid data can clarify how much information one would need to collect to create a useful data set.

In an asymmetric IPV first-price (or Dutch) auction, identification of each marginal distribution $G_{B_i}(\cdot)$ from observation of the winning bid and winner's identity is formally equivalent to identification of the well known competing risks model with independent nonidentically distributed risks.[25] For that model, nonparametric identification was shown by Berman (1963). Since knowledge of each $G_{B_i}(\cdot)$ completely determines the distribution of

$$B_i + \frac{G_{M_i|B_i}(B_i|B_i;\mathcal{N})}{g_{M_i|B_i}(B_i|B_i;\mathcal{N})}$$

[24] Brendstrup and Paarsch (2003) point out that in a Dutch auction the set of actual opponents may be observable before the bidding decision is made. With a binding reserve price that creates a distinction between the potential bidders and actual bidders (see Section 6.3 for definitions), this destroys strategic equivalence. The basic approach for first-price auctions can still be applied, however, if the distribution of the actual bidders' valuations (which reflects truncation at the reserve price) is the object of interest. See the related discussion in Section 6.3.1.

[25] The data generating process mapping bids to observables is formally identical to that in a complementary risks model, where failure of all components triggers the observable system failure, and one observes the identity of the last component to fail. This is isomorphic to the competing risks model.

identification of the marginal distributions $F_{U_i}(\cdot)$ then follows. This gives the following result from Athey and Haile (2002).

THEOREM 3.2. *Suppose that the transaction price and the number of bidders* (*and, if bidders are asymmetric, the set $\mathcal{N}$ and identity of the winner*) *are observed in first-price auctions with independent private values. Then $F_{\mathbf{U}}(\cdot)$ is identified.*

To gain some intuition, consider the symmetric case, where the observable transaction price $B^{(n:n)}$ has distribution $G_B^{(n:n)}(b)$, $G_{B_i}(\cdot)$ can be written as $G_B(\cdot)$, and

$$\frac{G_B^{(n:n)}(b)}{g_B^{(n:n)}(b)} = \frac{G_B(b)^n}{n g_B(b) G_B(b)^{n-1}} = \frac{n-1}{n}\left(\frac{G_B(b)}{(n-1)g_B(b)}\right). \tag{3.7}$$

As first observed by Guerre, Perrigne and Vuong (1995), identification then follows from (3.5).

In the asymmetric case, the derivation of Berman's (1963) equation (2) [see also Prakasa-Rao (1992, Theorem 7.3.1 and Remarks 7.3.1)] yields the relation (fixing $\mathcal{N}$)

$$G_{B_i}(b_i) = \exp\left\{\int_{-\infty}^{b_i}\left(\sum_{j=1}^{n} G_i^w(s)\right)^{-1} \mathrm{d}G_j^w(s)\right\}, \tag{3.8}$$

where $G_i^w(b_i) = \Pr(B_i \leqslant b_i, B_i \geqslant B_j\ \forall j)$. Since each $G_i^w(b_i)$ is observable, each $G_{B_i}(b_i)$ is identified. The marginal distributions $G_{B_i}(\cdot)$ uniquely determine the underlying distributions $F_{U_i}(\cdot)$ through the first-order condition (3.1) as in the case in which all bids are observed.

An immediate implication of Theorem 3.2 is identification from the transaction price in a Dutch auction.

COROLLARY 3.1. *Suppose that the transaction price and the number of bidders* (*and, if bidders are asymmetric, the set $\mathcal{N}$ and the identity of the winner*) *are observed in Dutch auctions with independent private values. Then $F_{\mathbf{U}}(\cdot)$ is identified.*

As suggested by Laffont, Ossard and Vuong (1995), the requirement that $n$ be observable by the econometrician may fail in some Dutch auctions, where one might expect only the transaction price (i.e., the only bid made in the auction) to be recorded. It should be clear that without knowledge of $n$, knowledge of $G_B^{(n:n)}(\cdot)$ is insufficient to determine even $G_B(\cdot)$.[26]

Extending the estimation approaches described in the preceding sections to cases in which only the transaction price (winning bid) is observed is straightforward in the symmetric case, where (3.7) can be used. In the asymmetric case, Brendstrup and Paarsch

[26] Laffont, Ossard and Vuong (1995) suggest an approach for estimating $n$ when it is unknown but fixed. They assume that identification follows from a parametric distributional assumption.

(2003) have recently proposed substituting the empirical distribution function and a kernel density estimator for, respectively, $G_i^w(s)$ and $\frac{\mathrm{d}G_i^w(s)}{\mathrm{d}s}$ in Equation (3.8). The close relation of the model to the competing risks model suggests that other nonparametric estimators such as the Nelson–Aalan or Kaplan–Meier estimators [e.g., David and Moeschberger (1978), Andersen et al. (1991)] might also be used to estimate each $G_{B_i}(\cdot)$. Once these distributions are estimated, one might then simulate bids from these estimated distributions in order to estimate pseudo-values using the first-order condition (3.1).

### *3.3.2. Affiliated private values*

Next we consider what can be learned from the top two bids in first-price auctions in a richer private values environment. Intuitively, the top two bids contain much of the critical information for a first-price auction. First, these are the only two bids necessary to determine the distribution of the maximum opposing bid for each bidder, suggesting that at least some information about the markdown components of the equilibrium bid functions could be learned. Second, in equilibrium, the top two bids are monotonic transformations of the top two signals. As the following result, adapted from Athey and Haile (2002) shows, this is sufficient to enable partial identification in a symmetric first-price sealed-bid auction.

THEOREM 3.3. *Assume that the two highest bids are observed in first-price auctions. If bidders are asymmetric, assume that the set $\mathcal{N}$ and the identity of the winner are also observed. Then*

(i) *the equilibrium bid functions $\beta_i(\cdot;\mathcal{N})$ are identified for all $i = 1,\dots,n$;*
(ii) *with symmetric private values, the joint distribution of $U^{(n-1:n)}$ and $U^{(n:n)}$ is identified.*

PROOF. Part (ii) follows immediately from part (i), since in the symmetric private values case the two highest bids are made by the bidders with the two highest valuations. To prove part (i) for the more general asymmetric case, consider bidder 1 without loss of generality. Let $I^{(n:n)}$ denote the identity of the winning bidder. For almost all such $b_1 \in \operatorname{supp}[G_{B_1}(\cdot)]$ (using Bayes' rule, and canceling common terms)

$$\begin{aligned}
&\frac{\Pr(\max_{j\neq 1} B_j \leqslant b_1 \mid B_1 = b_1;\mathcal{N})}{\frac{\partial}{\partial m}\Pr(\max_{j\neq 1} B_j \leqslant m \mid B_1 = b_1;\mathcal{N})|_{m=b_1}} \\
&\quad= \frac{\frac{\partial}{\partial y}\Pr(\max_{j\neq 1} B_j \leqslant b_1,\ B_1 \leqslant y;\mathcal{N})|_{y=b_1}}{\frac{\partial^2}{\partial m\partial y}\Pr(\max_{j\neq 1} B_j \leqslant m,\ B_1 \leqslant y;\mathcal{N})|_{m=y=b_1}} \\
&\quad= \frac{\frac{\partial}{\partial y}G_{\mathbf{B}}(y, b_1,\dots,b_1;\mathcal{N})|_{y=b_1}}{\sum_{j\neq 1}\frac{\partial^2}{\partial y\partial s_j}G_{\mathbf{B}}(y, s_2,\dots,s_n;\mathcal{N})|_{y=s_2=\cdots=s_n=b_1}}
\end{aligned}$$

$$= \frac{\frac{\partial}{\partial y}\Pr(B^{(n:n)} \leqslant y,\ I^{(n:n)} = 1, \mathcal{N})|_{y=b_1}}{\frac{\partial^2}{\partial m \partial y}\Pr(B^{(n-1:n)} \leqslant m,\ B^{(n:n)} \leqslant y,\ I^{(n:n)} = 1, \mathcal{N})|_{m=y=b_1}}.$$

Since the last expression is the ratio of two observable functions, the right-hand side of (2.4) is identified almost everywhere, which determines bidder 1's (inverse) equilibrium bid function. □

Estimation approaches based on this partial identification result have not yet been explored. Note that estimates of each $\beta_i(\cdot; \mathcal{N})$ are themselves of interest, since these characterize the wedge between bids and valuations that determine the division of surplus and can lead to inefficiencies. In the symmetric case, knowledge of $\beta(\cdot; n)$ and the joint distribution of $(U^{(n-1:n)}, U^{(n:n)})$ would enable evaluation of rent extraction by the seller, the effects of introducing a reserve price, and the outcomes under a number of alternative selling mechanisms. As discussed in Section 8, this partial identification result can also be sufficient to enable discrimination between private and common values environments.

Observing the top two bids, however, is insufficient to identify the full joint distribution $F_{\mathbf{U}}(\cdot)$. In fact, Athey and Haile (2002) have shown that observation of *all* bids is needed, even in a symmetric setting.

THEOREM 3.4. *In the symmetric private values model, suppose that* $(B^{(n:n)}, B^{(n-1:n)})$ *are observed in first-price auctions but some* $B^{(j:n)}$, $j < n-1$ *is unobserved. Then* $F_{\mathbf{U}}(\cdot)$ *is not identified.*

PROOF. Let the point $(u_1, u_2, \ldots, u_n)$ be on the interior of the support of $F_{\mathbf{U}}(\cdot)$, with $u_1 < \cdots < u_n$. Starting with the true joint density $f_{\mathbf{U}}(\cdot)$, define a new joint density $\tilde{f}_{\mathbf{U}}(\cdot)$ by shifting mass $\delta$ from a neighborhood of $(u_1, \ldots, u_j, \ldots, u_n)$ (and each permutation) to a neighborhood of the point $(u_1, \ldots, u_j + \epsilon, \ldots, u_n)$ (and each permutation).[27] For small $\epsilon$ and $\delta$, this change preserves exchangeability. Since the distribution of $\max_{k \neq i} B_k$ is unaffected for any $i$ by this change, equilibrium bidding strategies (given by (3.1)) remain the same for all bidders. Furthermore, the only order statistic affected in moving from $\tilde{f}_{\mathbf{U}}(\cdot)$ to $f_{\mathbf{U}}(\cdot)$ is $U^{(j:n)}$. Since $B^{(j:n)} = \beta(U^{(j:n)}; n)$ is unobserved, the distribution of observables is unchanged. □

Intuitively, even under exchangeability, $F_{\mathbf{U}}(\cdot)$ is an $n$-dimensional joint distribution. Identifying this distribution with data of dimension $n-1$ or lower would require additional restrictions.

[27] Athey and Haile (2002, Theorem 4) describe this in more detail.

## 4. Ascending auctions with private values: Basic results

### 4.1. Identification

With private values, equilibrium bidding strategies in Milgrom and Weber's (1982) model of the ascending auction are particularly simple: it is a weakly dominant strategy for each bidder to exit the auction at his valuation. Hence, unlike the first-price auction, here there is no need to estimate inverse bid functions in order to relate the observed bids to the underlying valuations. This does not make identification trivial, however. The reason is the fact that the auction ends as soon as only one bidder remains. Because the auction stops at the second highest bid, the only information revealed about the winner's valuation is the censoring point $B^{(n-1:n)}$. This partial observability of bids is the main challenge to identification.

When valuations are independent, Athey and Haile (2002) have shown that identification does hold, even if one observes only the transaction price (and the identity of the winner, if bidders are asymmetric). This is easier to see when bidders are symmetric. In that case valuations are i.i.d. draws from the marginal distribution $F_U(\cdot)$. The transaction price is the order statistic $U^{(n-1:n)}$, which has distribution $F_U^{(n-1:n)}(\cdot)$. The distribution of an order statistic from an i.i.d. sample of size $n$ from an arbitrary distribution $F(\cdot)$ has the distribution [see, e.g., Arnold, Balakrishnan and Nagaraja (1992)]

$$F^{(i:n)}(s) = \frac{n!}{(n-i)!(i-1)!} \int_0^{F(s)} t^{i-1}(1-t)^{n-i}\,\mathrm{d}t \quad \forall s. \tag{4.1}$$

Since the right-hand side of (4.1) is strictly increasing in $F(s) \in [0,1]$, $F^{(i:n)}(s)$ uniquely determines $F(s)$ for every $s$.

When bidders have asymmetric independent private values, the identification argument is more subtle. Athey and Haile (2002) point out that the asymmetric ascending auction model is isomorphic to a model considered in the statistics literature on reliability, where Meilijson (1981) has provided a proof. To get some intuition, fix $\mathcal{N}$ with $|\mathcal{N}| = 3$ and define

$$\begin{aligned}
\widetilde{G}_3(t) &\equiv \Pr(\text{price} \leqslant t,\ 3 \text{ is the winner}) \\
&= \Pr(B_1 \leqslant t;\ B_2 \leqslant t;\ B_3 \geqslant t) + \Pr(B_1 \leqslant B_3;\ B_2 \leqslant B_3;\ B_3 \leqslant t) \\
&= F_{U_1}(t)F_{U_2}(t)\bigl(1 - F_{U_3}(t)\bigr) + \int_{\underline{u}}^{t} F_{U_1}(x)F_{U_2}(x)\,\mathrm{d}F_{U_3}(x) \\
&= \int_{\underline{u}}^{t} \bigl(F_{U_1}(x)F_{U_2}(x)\bigr)'\bigl(1 - F_{U_3}(x)\bigr)\,\mathrm{d}x,
\end{aligned}$$

where $(F_{U_1}F_{U_2})'$ is the first derivative of $F_{U_1}F_{U_2}$ and the last equality follows from integration by parts. Differentiating both sides, we obtain

$$\tilde{g}_3(t) = \bigl(F_{U_1}(t)F_{U_2}(t)\bigr)'\bigl(1 - F_{U_3}(t)\bigr),$$

which implies that

$$\begin{aligned}
\big(F_{U_1}(t)F_{U_2}(t)\big)' &= \frac{\tilde{g}_3(t)}{1-F_{U_3}(t)},\\
F_{U_1}(t)F_{U_2}(t) &= \int_{\underline{u}}^{t} \frac{\tilde{g}_3(x)}{1-F_{U_3}(x)}\,\mathrm{d}x,\\
\log F_{U_1}(t) + \log F_{U_2}(t) &= \log \int_{\underline{u}}^{t} \frac{\tilde{g}_3(x)}{1-F_{U_3}(x)}\,\mathrm{d}x.
\end{aligned}$$

Rewrite this as

$$\begin{bmatrix} 1 & 1 & 0\\ 1 & 0 & 1\\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \log F_{U_1}(x)\\ \log F_{U_2}(x)\\ \log F_{U_3}(x) \end{bmatrix} = \log \begin{bmatrix} \int_{\underline{u}}^{t} \frac{\tilde{g}_3(x)}{1-F_{U_3}(x)}\,\mathrm{d}x\\ \int_{\underline{u}}^{t} \frac{\tilde{g}_2(x)}{1-F_{U_2}(x)}\,\mathrm{d}x\\ \int_{\underline{u}}^{t} \frac{\tilde{g}_1(x)}{1-F_{U_1}(x)}\,\mathrm{d}x \end{bmatrix}. \tag{4.2}$$

This is a $3 \times 3$ system of operator equations defining how the three observable marginal distributions are related to the three marginal distributions $F_{U_i}(\cdot)$ of interest. Meilijson (1981) showed that this system has a unique solution.

We summarize these results in the following theorem.

THEOREM 4.1. *In an ascending auction with symmetric independent private values, $F_{\mathbf{U}}(\cdot)$ is identified when the transaction price and the number of bidders are observable. In an ascending auction with asymmetric independent private values, $F_{\mathbf{U}}(\cdot)$ is identified when the transaction price, the identity of the winning bidder, and the set $\mathcal{N}$ are observable.*

One attractive feature of this result is that it implies that one need not use bids other than the transaction price to estimate $F_U(\cdot)$. This is valuable because in many applications one may have little confidence in the interpretation of losing bids implied by Milgrom and Weber's (1982) button auction model. With independence, one is free to ignore losing bids altogether, relying only on the assumption that the transaction price equals the second highest valuation.

Athey and Haile (2002) give a much more negative result when the independence assumption is dropped, even with symmetric bidders. The proof mirrors that of Theorem 3.4.

THEOREM 4.2. *In a symmetric private values model, the joint distribution $F_{\mathbf{U}}(\cdot)$ is not identified from the observable bids in an ascending auction.*

### *4.2. Estimation*

Here we make the same sampling assumptions made in the discussion of estimation for first-price sealed bid auctions (see Section 3.2). In an ascending auction, typically one treats the highest price offered by each bidder as his "bid," i.e., his exit price in the button auction model.[28] Using these data, several parametric estimation approaches for the symmetric IPV model have been explored in the literature. Maximum likelihood, nonlinear least squares, and GMM are among the methods considered. Due to the simplicity of the equilibrium bid function, likelihood functions or moment conditions are easily derived from the probability density function of the winning bid alone [e.g., Paarsch (1992b), Donald and Paarsch (1996), Baldwin, Marshall and Richard (1997), Haile (2001)], or of the $n-1$ losing bids [e.g., Donald and Paarsch (1996), Paarsch (1997)]. In the former case, the likelihood of a winning bid $b$ is $f_U^{(n-1:n)}(b)$. For the latter case, the likelihood for the losing bids in a given auction is

$$n!\left[1 - F_U\left(b^{(n-1:n-1)}\right)\right] \prod_{j<n} f_U\left(b^{(j:n)}\right).$$

To our knowledge, nonparametric estimation of the symmetric IPV model has been performed only in simulations [Haile and Tamer (2003)], although this is actually simpler than nonparametric estimation in the case of a first-price auction. Following Haile and Tamer (2003), for $H \in [0, 1]$ define the strictly increasing differentiable function $\phi(H; i, n)$ implicitly as the solution to

$$H = \frac{n!}{(n-i)!(i-1)!} \int_0^{\phi} s^{i-1}(1-s)^{n-i}\, \mathrm{d}s \tag{4.3}$$

so that by (4.1)

$$F_U(u) = \phi\left(F_U^{(i:n)}(u); i, n\right) \quad \forall u, i \leqslant n. \tag{4.4}$$

In particular,

$$F_U(u) = \phi\left(F_U^{(n-1:n)}(u); n-1, n\right) \quad \forall u. \tag{4.5}$$

Since the winning bid is $B^{(n-1:n)} = U^{(n-1:n)}$, one can construct an estimator of $F_U(u)$ by substituting the empirical distribution

$$\widehat{G}_B^{(n-1:n)}(u) = \frac{1}{T_n} \sum_{t=1}^{T} \mathbf{1}\left\{n_t = n,\ B^{(n_t-1:n_t)} \leqslant u\right\}$$

for $F_U^{(n-1:n)}(u)$ inside the function $\phi(\cdot)$ on the right-hand side of (4.5). Since $G_B^{(n-1:n)}(\cdot) = F_U^{(n-1:n)}(\cdot)$, by standard arguments $\widehat{G}_B^{(n-1:n)}(u)$ converges uniformly to $F_U^{(n-1:n)}(u)$ almost surely and has a normal asymptotic approximation. Convergence of

[28] See, e.g., the surveys of Paarsch (1994) and Hendricks and Paarsch (1995). A source of ambiguity arises when one observes $n$ such bids, with $B^{(n:n)}$ significantly higher than $B^{(n-1:n)}$. In such cases, one may question the applicability of the button auction model. For now we assume the button auction structure and treat the distributions $G_B^{(n:n)}(\cdot)$ and $G_B^{(n-1:n)}(\cdot)$ as identical.

$\phi(\widehat{F}_U^{(n-1:n)}(u); n-1, n)$ to $F_U(u)$ then follows, with an asymptotic normal distribution obtainable by the delta method.[29] In practice, one can use the relation

$$F_U^{(n-1:n)}(u) = \sum_{j=n-1}^{n} \binom{n}{j} F_U(u)^j \big(1 - F_U(u)\big)^{n-j} \tag{4.6}$$

instead of the equivalent but more computationally demanding (4.5) when solving for each $\widehat{F}_U(u)$. Monotonicity of the relation between $F_U^{(n-1:n)}(u)$ and $F_U(u)$ makes numerical solution particularly simple.

Note that when there is exogenous variation in the number of bidders, there will be as many different estimators $\phi(\widehat{F}_U^{(n-1:n)}(u); n-1, n)$ of $F_U(u)$ available as there are different values of $n$ in the data. If one observes losing bids beyond the transaction price and assumes these are generated by the button auction model, additional estimators will be available, based on Equation (4.4) with $i < n-1$. An efficient estimator would take an optimally weighted average of these different estimators, imposing the constraint that the estimated CDF be monotone.

In the case of asymmetric bidders, no simple relation like (4.6) is available. However, a likelihood approach provides several possible estimation strategies. The likelihood for the observable event $\{i$ wins at price $p\}$ is

$$\big(1 - F_{U_i}(p)\big) \sum_{j \neq i} f_{U_j}(p) \prod_{k \neq i, j} F_{U_k}(p).$$

Hence, if we let $I_t$ denote the winner of auction $t$, the likelihood function has the form

$$\mathcal{L} = \prod_t \big(1 - F_{U_{I_t}}(p)\big) \sum_{j \neq I_t} f_{U_j}(p) \prod_{k \neq I_t, j} F_{U_k}(p). \tag{4.7}$$

Parametric or nonparametric MLE might then be applied. Brendstrup and Paarsch (2004) have recently proposed a "semi-nonparametric" [Gallant and Nychka (1987)] estimation approach based on this likelihood.[30]

### 4.3. *An alternative, incomplete model of ascending auctions*

In some cases an auction institution closely matching the structure of the button auction model is observed in practice. Bidders may, for example, raise their hands or other objects to indicate their participation continuously as the auctioneer raises the price [see, e.g., Zulehner's (2003) description of cattle auctions]. When the auction is conducted in a less structured oral format, however, one may question the applicability of the button

[29] In fact, the convergence is uniform. These results follow from those given in Haile and Tamer (2002, Appendix A; 2003, Theorem 3).

[30] They also consider auctions in which multiple units are sold sequentially, focusing on bids in the (single-unit, asymmetric) auction of the final unit.

auction model as an empirical structure. Of particular concern is the fact that there is no need for a bidder to continuously indicate whether she is "in" or "out" as the auction proceeds. Nontrivial minimum bid increments are often used, and a bidder is free to "jump bid" or to remain silent for most of the auction and bid only when it looks like the auction is about to end (if others do not bid the price past her valuation first). Such behaviors are common in practice and raise the possibility that bidders will fail to reveal their full willingness to pay, or even fail to bid altogether. Several theoretical extensions of the standard model have been proposed, mainly focusing on the case of common values [Avery (1998), Harstad and Rothkopf (2000), Izmalkov (2003)]. Until recently, however, all empirical models of the ascending auction relied on significant abstractions for tractability of the underlying theoretical model.

As an alternative to relying on the structure of the button auction or another stylized model, Haile and Tamer (2003) have proposed an empirical approach to ascending auctions with symmetric independent private values using two simple assumptions to govern the interpretation of the observed bids:

ASSUMPTION 4.1. Bidders do not bid more than they are willing to pay.

ASSUMPTION 4.2. Bidders do not allow an opponent to win at a price they are willing to beat.

These assumptions allow bidding as in the dominant strategy equilibrium of the button auction model but do not require it. In particular, bids need not be equal to valuations or even monotonic in valuations, and the price need not equal the second highest valuation. These assumptions define an "incomplete" model of an ascending auction: they place some restrictions on the relation between valuations and bids, but do not fully characterize behavior. While this incomplete model is insufficient to identify the distribution of valuations from the distribution of bids, it does provide partial identification; in particular, one may still obtain informative *bounds* on the distribution of valuations.

### *4.3.1. Bounding the distribution of bidder valuations*

To obtain an upper bound on the distribution function $F_U(\cdot)$, observe that Assumption 4.1 is equivalent to assuming $b_i \leqslant u_i$ for all $i$. In an $n$-bidder auction, it is easy to confirm that this implies $b^{(i:n)} \leqslant u^{(i:n)}$ $\forall i$, which then gives

$$G_B^{(i:n)}(u) \geqslant F_U^{(i:n)}(u) \quad \forall i, n, u. \tag{4.8}$$

Recalling the definition (4.3) and Equation (4.4), we know that

$$F_U(u) = \phi\big(F_U^{(i:n)}(u); i, n\big) \quad \forall i, n, u. \tag{4.9}$$

Since the function $\phi(\cdot; i, n) : [0, 1] \to [0, 1]$ is strictly increasing, (4.8) and (4.9) together give

$$\phi\big(G_B^{(i:n)}(u); i, n\big) \geqslant F_U(u) \quad \forall i, n, u.$$

For each $u$, this yields $\sum_{n=\underline{n}}^{\overline{n}} n$ upper bounds on $F_U(u)$. The most informative bound (i.e., the smallest upper bound) is obtained by taking the minimum at each value of $u$:

$$F_U^+(u) = \min_{i,n} \phi\big(G_B^{(i:n)}(u); i, n\big). \tag{4.10}$$

A similar approach can be taken to obtain a lower bound on $F_U(\cdot)$. Letting $\Delta \geqslant 0$ denote the minimum bid increment in the auction, Assumption 4.2 implies that all losing bidders have valuations less than $b^{(n:n)} + \Delta$, implying $u^{(n-1:n)} \leqslant b^{(n:n)} + \Delta$. If $G_\Delta^{(n:n)}(\cdot)$ denotes the distribution of $B^{(n:n)} + \Delta$, this gives

$$G_\Delta^{(n:n)}(u) \leqslant F_U^{(n-1:n)}(u) \quad \forall n, u.$$

Applying the monotonic transformation $\phi(\cdot; n-1, n)$ to both sides gives

$$\phi\big(G_\Delta^{(n:n)}(u); n-1, n\big) \leqslant F_U(u) \quad \forall n, u.$$

This yields multiple lower bounds on $F_U(u)$ (one for each value of $n$). The most informative bound can be constructed by taking the pointwise maximum:

$$F_U^-(u) = \max_{n} \phi\big(G_\Delta^{(n:n)}(u); n-1, n\big). \tag{4.11}$$

We summarize these results in the following theorem.

THEOREM 4.3. *$F_U^-(u) \leqslant F_U(u) \leqslant F_U^+(u)$ for all $u$.*

Note that in principle this approach can be followed even when the transaction price is the only bid available from each auction – the only modification required is that the minimum in (4.10) would be taken over $n$ only, fixing $i = n$. However, an essential requirement of the approach is that the number of bidders, $n$, be observable to the econometrician. This is also essential for the methods discussed above for both sealed-bid and button auction models, but the assumption may be more suspect in an ascending auction in which the button auction structure is inappropriate. The number $n$ will be observed if all bidders make some bid during the auction, or if bidders must pre-qualify, register, or otherwise identify themselves in order to be eligible to bid. This is the case in the timber auctions studied by Haile and Tamer (2003) and many other public sector auctions.[31]

In general, the informativeness of the bounds $F_U^+(u)$ and $F_U^-(u)$ depends on the deviation of the true data generating process from that implied by the button auction model. In fact, if the restriction $B^{(n:n)} = B^{(n-1:n)}$ implied by the button auction model is consistent with the data, the bounds $F_U^-(\cdot)$ and $F_U^+(\cdot)$ collapse to the true distribution $F_U(\cdot)$, providing point identification. By contrast, imposing the full structure of the button auction model when this is not the true data generating process need not result in

[31] Song (2004) has recently considered identification and estimation for ascending auctions (and others) when $n$ is not observed. We will discuss one such case in Section 6.3.4 below.

an estimate of $F_U(\cdot)$ that lies within the bounds, regardless of sample size.[32] While this should not be surprising – imposing false restrictions should be expected to yield misleading estimates – it is a useful reminder that imposing structure in order to obtain point identification is not equivalent to selecting a point estimate within bounds obtained from a less restrictive but incomplete model.

Estimation of the bounds is achieved by substituting the empirical distributions

$$\widehat{G}_B^{(i:n)}(b) = \frac{1}{T_n} \sum_{t=1}^{T} \mathbf{1}\{n_t = n,\ b^{(i:n_t)} \leqslant b\}$$

and

$$\widehat{G}_\Delta^{(n:n)}(b) = \frac{1}{T_n} \sum_{t=1}^{T} \mathbf{1}\{n_t = n,\ b^{(n_t:n_t)} + \Delta_t \leqslant b\}$$

for the corresponding CDFs in (4.10) and (4.11). Since the empirical distribution functions are uniformly consistent and asymptotically normally distributed estimators of their population analogs, differentiability of $\phi(\cdot; i, n)$ ensures that each $\phi(G_B^{(i:n)}(u); i, n)$ and $\phi(G_\Delta^{(n:n)}(u); n-1, n)$ are consistent and asymptotically normally distributed as well. Continuity of the min and max functions then ensures consistency of the estimates of the estimators

$$\widehat{F}_U^+(u) = \min_{i,n} \phi\big(\widehat{G}_B^{(i:n)}(u); i, n\big),$$

$$\widehat{F}_U^-(u) = \max_{n} \phi\big(\widehat{G}_\Delta^{(n:n)}(u); n-1, n\big).$$

These estimators have nonnormal asymptotic distributions, due to the max and min. However, Haile and Tamer (2002) show that the bootstrap (see Section 3.2.1) may be used for inference. A more difficult problem is that, while these estimators are consistent, in practice the max and min can lead to severe finite sample bias, potentially even leading to estimated upper and lower bounds that cross. Intuitively, taking the minimum (maximum) of several consistent estimators makes it likely that an estimator with downward (upward) sampling error is selected. One solution, discussed in greater detail by Haile and Tamer (2003), is to define bounds in finite samples based on smooth approximations to the max and min functions in the definitions of $\widehat{F}_U^+(u)$ and $\widehat{F}_U^-(u)$ above. This amounts to using weighted averages instead of the max or min.

### 4.3.2. Bounding the optimal reserve price

Unlike point estimates of $F_U(\cdot)$, it is not immediately clear whether bounds on $F_U(\cdot)$ would be useful.[33] For example the key policy choice for the seller in the symmetric

[32] See Haile and Tamer (2003) for additional discussion and simulation results.

[33] Haile and Tamer (2003) demonstrate an additional use of the bounds by showing how to incorporate auction covariates nonparametrically. Building on Manski and Tamer (2002), the resulting bounds on conditional

IPV environment is the reserve price [Myerson (1981)]. When $F_U(\cdot)$ is continuously differentiable, the optimal reserve price $r^*$ is defined by the equation[34]

$$r^* = c_0 + \frac{1 - F_U(r^*)}{f_U(r^*)}, \tag{4.12}$$

where $c_0$ is the seller's valuation (or marginal cost) of the good. However, nondegenerate bounds on $F_U(\cdot)$ place no restriction on its derivative $f_U(\cdot)$ at any given point. Hence, just as a monopolist's price need not shift in the same direction as demand, $r^*$ need not lie between the reserve prices that would be optimal if $F_U^+(\cdot)$ or $F_U^-(\cdot)$ were the true distribution of valuations. Note that the same problem arises any time one wishes to construct confidence bands on the optimal reserve price from confidence bands on nonparametric point estimates of $F_U(u)$, e.g., using the method described in Section 4.2.

Observe, however, that when the seller's own valuation for the good is $c_0$, $r^*$ solves $\max_r \pi(r)$, where[35]

$$\pi(r) = (r - c_0)\bigl(1 - F_U(r)\bigr).$$

Since $F_U(r)$ must lie between $F^-(r)$ and $F^+(r)$, $\pi(r)$ must lie between

$$\pi_1(r) = (r - c_0)\bigl(1 - F_U^+(r)\bigr)$$

and

$$\pi_2(r) = (r - c_0)\bigl(1 - F_U^-(r)\bigr).$$

Figure 1 illustrates. Under the additional assumption that $\pi(r)$ is strictly quasi-concave in $r$ (which ensures a unique solution to (4.12)) we can use the bounding "profit" functions $\pi_1(\cdot)$ and $\pi_2(\cdot)$ to place bounds on $r^*$. Let $r_1^* \in \arg\sup \pi_1(r)$, $r_2^* \in \arg\sup \pi_2(r)$, and $\pi_1^* = \pi_1(r_1^*)$. We obtain the trivial result $r^* = r_1^*$ when $\pi_2(r_1^*) = \pi_2(r_2^*) = \pi_1^*$, or when $\pi_2(r_1^*) = \pi_1^*$ and either $\pi_1(\cdot)$ or $\pi_2(\cdot)$ has slope zero at $r_1^*$. For these trivial cases let $r^- = r^+ = r_1^*$. For all other cases define

$$r^- = \sup\bigl\{r < r_1^*\colon \pi_2(r) \leqslant \pi_1^*\bigr\},$$
$$r^+ = \inf\bigl\{r > r_1^*\colon \pi_2(r) \leqslant \pi_1^*\bigr\}.$$

Haile and Tamer (2003) prove the following result.[36]

distributions can then be used to estimate bounds on parameters of a semiparametric model describing how valuations shift with auction characteristics.

[34] This is easily derived for a second-price sealed-bid or button auction, where a reserve price of $r$ implies expected revenue $rnF_U(r)^{n-1}(1 - F_U(r)) + \int_t^\infty un(n-1) f_U(u)F_U(u)(1 - F_U(u))\,du$. Myerson (1981) shows that, under a regularity condition, a standard auction with an optimal reserve price is optimal among all possible selling mechanisms. Haile and Tamer (2003) show that $r^*$ is also optimal in their incomplete model ascending auction as long as Assumptions 4.1 and 4.2 are interpreted as a partial characterization of equilibrium behavior in some true but unspecified auction mechanism.

[35] Note that $\pi(r)$ is not the expected profit of the seller when $n > 1$. The usefulness of this function is the fact that its maximum is attained at the same value of $r$ that maximizes the seller's expected profit.

[36] Bounds are said to be *sharp* if they exhaust all information available from the data and *a priori* assumptions.

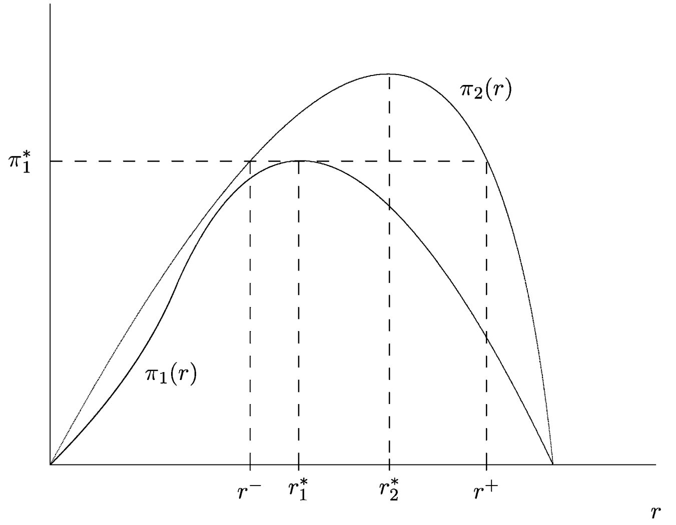


Figure 1.

THEOREM 4.4. *Suppose $\pi(r)$ is continuously differentiable and strictly quasi-concave in $r$. Then $r^* \in [r^-, r^+]$. Given the bounds $F^+(\cdot)$ and $F^-(\cdot)$ on $F_U(\cdot)$, the bounds $r^-$ and $r^+$ on $r$ are sharp.*

Intuition for the result can be seen in Figure 1. We know that the true function $\pi(\cdot)$ lies between $\pi_1(\cdot)$ and $\pi_2(\cdot)$ and must, therefore, reach a peak of at least $\pi_1^*$. Such a peak cannot be reached outside the interval $[r^-, r^+]$. However, prices arbitrarily close to either of these endpoints could be the true optimum $r^*$.

For estimation, assume for simplicity that $\pi_2(r)$ has nonzero slope at $r = r^-$ and $r = r^+$.[37] Let

$$
\begin{aligned}
\hat{\pi}_1(r) &= (r - c_0)\big(1 - \widehat{F}_U^+(r)\big),\\
\hat{\pi}_2(r) &= (r - c_0)\big(1 - \widehat{F}_U^-(r)\big),\\
\hat{\pi}_1^* &= \sup_r \hat{\pi}_1(r),\\
\hat{r}_1^* &= \arg\sup_r \hat{\pi}_1(r)
\end{aligned}
$$

and define the correspondence $\pi_2^c(\cdot)$ by

$$\big\{\pi \in \pi_2^c(r)\big\} \quad \Longleftrightarrow \quad \Big\{\lim_{r' \downarrow r} \hat{\pi}_2(r') \leqslant \pi \leqslant \lim_{r' \uparrow r} \hat{\pi}_2(r')\Big\}.$$

[37] Haile and Tamer (2003) provide estimators that do not require this assumption.

This defines a smooth sample analog of $\pi_2(\cdot)$ that can be used to define consistent estimators of $r^-$ and $r^+$:

$$\hat{r}^- = \sup\{r < \hat{r}_1^*: \pi = \hat{\pi}_1^* \text{ for some } \pi \in \pi_2^c(r)\},$$
$$\hat{r}^+ = \inf\{r > \hat{r}_1^*: \pi = \hat{\pi}_1^* \text{ for some } \pi \in \pi_2^c(r)\}.$$

### *4.3.3. Asymmetric and affiliated private values*

In principle, the applicability of Assumptions 4.1 and 4.2 is not limited to environments with symmetric independent private values. Haile and Tamer (2001) have explored extensions to models of asymmetric and/or affiliated private values. While it is encouraging that any restrictions at all on the joint distribution $F_{\mathbf{U}}(\cdot)$ can be obtained without the assumption of independence that was required for identification in the button auction model, in practice the bounds one can obtain without the independence assumption are likely to be quite wide. Intuitively, when one observes only bounds on realizations of random variables, it is difficult to learn much about their correlation structure. Of course, without knowledge of the correlation structure, a number of important positive and normative questions cannot be answered.[38] Thus, while the bounds approach provides a way of addressing concerns about the appropriateness of the standard button auction model, it may provide little help in environments in which the button auction model itself is unidentified.

## 5. Specification testing

Identification of the models discussed above relies on behavioral assumptions and on assumptions about the underlying demand and information structure. Obviously, then, the choice of model is important. In some environments there are overidentifying restrictions that can be used to test some assumptions while maintaining others. Several testing approaches have been described in the literature to date, although so far there has been little attention to development of formal statistical tests.

[38] Optimal auction design with correlated valuations is much more complex than in the IPV case, requiring precise information about the underlying correlation structure [cf., Crémer and McLean (1988) and McAfee and Reny (1992)]. Quint (2004) has shown that even the simpler question of the optimal reserve price cannot be addressed with a bounds approach in such an environment. In particular, for any reserve price $r \geqslant u_0$ and any distribution of bids, there exists an underlying joint distribution of valuations consistent with these bids and Assumptions 1 and 2 such that $r$ is the optimal reserve price. Hence, no restriction on the optimal reserve price can be obtained from nondegenerate bounds on the joint distribution of valuations.

### 5.1. Theoretical restrictions in first-price auction models

We first consider restrictions imposed by equilibrium bidding in first-price auctions.[39] Recall that a model is testable if there exists some joint distribution of observables that cannot be rationalized by the model. It is then natural to ask what set of distributions *can* be rationalized. Here we provide two results for the affiliated private values (APV) framework.[40] The first gives necessary conditions for a distribution of bids to be rationalized by equilibrium behavior, while the second gives necessary and sufficient conditions in two special cases: symmetric affiliated private values, or independent private values (IPV).

THEOREM 5.1. *Consider the APV first-price auction with fixed $\mathcal{N}$. Necessary conditions for $G_{\mathbf{B}}(\cdot;\mathcal{N})$ to be rationalized by equilibrium bidding are*

(a) $(B_1,\ldots,B_n)$ *are affiliated*;

(b) *for each $i$, $\xi_i(\cdot,\mathcal{N})$ is continuous and strictly increasing on* $\operatorname{supp}[B_i]$;

(c) $\operatorname{supp}[\xi_1(B_1,\mathcal{N})]\times\cdots\times\operatorname{supp}[\xi_n(B_n,\mathcal{N})]$ *is a convex, compact set, and on this set the joint distribution $G_{\mathbf{B}}(\xi_1^{-1}(u_1,\mathcal{N}),\ldots,\xi_n^{-1}(u_n,\mathcal{N}))$ is absolutely continuous* (*with respect to the Lebesgue measure*) *as a function of* $\mathbf{u}$, *with a strictly positive density*;

(d) $\underline{b}_i=\underline{b}$ *for all $i$, and $\xi_i(\underline{b},\mathcal{N})=\underline{b}$ for all $i\in\mathcal{N}$*;

(e) $\xi_i(\bar{b}_i,\mathcal{N})=\xi_j(\bar{b}_j,\mathcal{N})$ *for all $i,j\in\mathcal{N}$; and*

(f) *for each $i$,* $\operatorname{supp}[B_i]\subseteq\operatorname{supp}[\max_{j\in\mathcal{N}\setminus i}B_j]$, *and* $\operatorname{supp}[\max_{j\in\mathcal{N}\setminus i}B_j]$ *is convex.*

PROOF. Given strictly increasing bidding strategies and affiliated private values with an atomless type distribution, affiliation of bids and strict monotonicity of $\xi_i(\cdot,\mathcal{N})$ on $\operatorname{supp}[B_i]$ follow directly. Continuity of $\xi_i(\cdot,\mathcal{N})$ follows from strict monotonicity of the bidding strategies together with the assumption that $\operatorname{supp}F_{U_i}(\cdot)$ is convex. Since Assumption 2.1 requires that $F_{\mathbf{U}}(\cdot)$ have a strictly positive joint density on a compact convex set, the relationship between $F_{\mathbf{U}}(\cdot)$ and $G_{\mathbf{B}}(\cdot)$ given by (3.2) implies that (c) must hold. The assumption that $\operatorname{supp}F_{U_i}(\cdot)$ does not vary with $i$, together with the equilibrium conditions $\max(r,\underline{u})=\underline{b}_i$ and $\beta_i(\max(r,\underline{u}))=\max(r,\underline{u})$ for each $i$, imply (d) and (e). The necessity of $\operatorname{supp}[B_i]\subseteq\operatorname{supp}[\max_{j\in\mathcal{N}\setminus i}B_j]$ follows from (d) and the fact that when bidding against opponents who use strictly increasing strategies it is never optimal for bidder $i$ to bid more than the minimum necessary to win with a particular probability. The same logic implies that $\operatorname{supp}[\max_{j\in\mathcal{N}\setminus i}B_j]$ is convex, since

[39] Recall that we maintain Assumption 2.1, restricting the primitives of the model, and that we focus on equilibria in strictly increasing strategies. Theorem 2.1 guarantees that such an equilibrium exists for the APV model. If bidders are symmetric, Theorem 2.1 implies that there is a unique equilibrium in the class of equilibria in nondecreasing strategies. When bidders are asymmetric, we do not have a uniqueness result.

[40] Guerre, Perrigne and Vuong (2000) gave a similar result for the symmetric independent private values model. See also Li, Perrigne and Vuong (2002).

no bidder $j \in \mathcal{N}\backslash i$ would find it optimal to place a bid at the upper boundary of a gap in this support. $\square$

Note that we do not provide sufficient conditions for $G_{\mathbf{B}}(\cdot)$ to be rationalized in the APV model with asymmetric bidders, since a full equilibrium characterization is not available for that case. However, in the special cases of IPV or symmetric bidders when valuations have a continuously differentiable density, Theorem 2.1 implies that there is a unique equilibrium, which has strictly increasing and differentiable strategies and the same support for the equilibrium bids of all bidders. When bidders are symmetric, the unique equilibrium is symmetric. In these settings we have necessary and sufficient conditions for a bidding distribution to be rationalized. The statement of the conditions of Theorem 5.1 can then also be simplified somewhat, exploiting differentiability of strategies.

THEOREM 5.2. *Consider the APV first-price auction with fixed $\mathcal{N}$. Assume that $f_{\mathbf{U}}(\cdot)$ is continuously differentiable and suppose, further, that either*

(i) *$(U_1, \ldots, U_n)$ are mutually independent or*

(ii) *bidders are symmetric.*

*Necessary and sufficient conditions for $G_{\mathbf{B}}(\cdot; \mathcal{N})$ to be rationalized by equilibrium bidding are*:

(a) *$(B_1, \ldots, B_n)$ are affiliated, and in case* (i) *they are independent, while in case* (ii) *they are exchangeable*;

(b) *for each $i$, $\xi_i(\cdot, \mathcal{N})$ is differentiable and strictly increasing on* $\text{supp}[B_i]$;

(c) *$G_{\mathbf{B}}(\xi_1^{-1}(u_1, \mathcal{N}), \ldots, \xi_n^{-1}(u_n, \mathcal{N}))$ is absolutely continuous* (*with respect to the Lebesgue measure*) *as a function of* $\mathbf{u}$, *with a positive continuously differentiable density on* $\text{supp}[\xi_1(B_1, \mathcal{N})] \times \cdots \times \text{supp}[\xi_n(B_n, \mathcal{N})]$ *and zero density elsewhere*;

(d) *$\xi_i(\underline{b}, \mathcal{N}) = \underline{b}$ for all $i \in \mathcal{N}$*;

(e) *$\xi_i(\bar{b}_i, \mathcal{N}) = \xi_j(\bar{b}_j, \mathcal{N})$ for all $i, j \in \mathcal{N}$; and*

(f) $\text{supp}[B_i]$ *is convex, compact, and the same for all $i$.*

PROOF. Given strictly increasing, differentiable bidding strategies and the conditions on the $F_{\mathbf{U}}(\cdot)$, affiliation of bids and the relevant independence and symmetry conditions follow directly. For condition (f), equal supports is necessary by Theorem 2.2 in case (i) and by symmetry in case (ii). Convex, compact support follows because bidding strategies are strictly increasing, continuous functions of random variables with convex, compact support. Condition (b) is necessary because differentiability of $\xi_i(\cdot, \mathcal{N})$ is equivalent to differentiability of $\xi_i^{-1}(\cdot, \mathcal{N})$ (since $\xi_i(\cdot, \mathcal{N})$ is strictly increasing), with the latter equal to the equilibrium bidding strategy under the assumptions of the model. Conditions (d) and (e) are necessary following the arguments in Theorem 5.1. Recall that Assumption 2.1 requires that $F_{\mathbf{U}}(\cdot)$ have a strictly positive joint density on a compact, convex set, and that we have assumed that it has a differentiable density. The set $\text{supp}[\xi_1(B_1, \mathcal{N})] \times \cdots \times \text{supp}[\xi_n(B_n, \mathcal{N})]$ is the support of valuations implied by the model, and it is convex and compact by (f) and differentiability of $\xi_i(\cdot)$. The relationship between $F_{\mathbf{U}}(\cdot)$ and $G_{\mathbf{B}}(\cdot)$ specified by (3.2) then implies (c).

To see that the stated conditions are sufficient for $G_{\mathbf{B}}(\cdot;\mathcal{N})$ to be rationalized, observe that they ensure that $\xi_i^{-1}(\cdot)$ is well defined, differentiable, and strictly increasing on $\operatorname{supp}[\xi_i(B_i)]$ for each $i$, and that in symmetric models it is the same for each $i$, so that the expression for $F_{\mathbf{U}}(\cdot)$ in (3.2) is well defined and satisfies the relevant affiliation, independence, symmetry, and differentiability conditions. The conditions guarantee that the implied $F_{\mathbf{U}}(\cdot)$ has a support that is convex and compact; that it has a strictly positive, continuously differentiable density on this support; and that the support is the same for all bidders. The definition of $\xi_i(\cdot)$ implies that if each bidder $i$ uses the bidding strategy $\xi_i^{-1}(\cdot)$, his first-order condition for optimality is satisfied. Under independence or symmetry, bidder payoffs satisfy a single crossing property: for any fixed monotone strategies by opposing bidders, a higher realized valuation $u_i$ leads to a higher marginal return to increasing one's bid. Standard results from the literature on auctions and mechanism design [see, e.g., Fudenberg and Tirole (1991)] imply when the single crossing property holds, local optimality of a bid (i.e., first-order conditions hold) together with monotonicity of the bidding strategy are necessary and sufficient for global optimality of the strategy. Thus the strategies $\{\xi_i^{-1}(\cdot)\}_{i\in\mathcal{N}}$ form an equilibrium. □

The importance of a result providing sufficient conditions for a bid distribution to be rationalized by equilibrium behavior should not be underappreciated. Without such a result, one would have no way of ensuring that the interpretation of bids based on the first-order conditions is valid. In particular, for an observed bid $b_i$ and an implied valuation $u_i = \xi_i(b_i)$, there would be no guarantee that $b_i$ was actually an equilibrium bid for a bidder with valuation $u_i$. Verifying that the observed bids actually can be rationalized by equilibrium behavior is analogous to verifying second-order conditions for optimality: only when such sufficient conditions are verified can we be sure that the mappings (forward or inverse) provided by the first-order conditions relate valuations to optimal (best-response) bids. Theorem 5.2 can then be used in two ways. First, in an application one can attempt to verify that sufficient conditions for bid data to be consistent with the assumptions of the model are satisfied. Second, the necessary conditions suggest specification tests, which we discuss further in the following section.

We note that it is possible to generalize the overall empirical approach to the case where condition (e) above fails by relaxing the assumption that $\operatorname{supp}[U_i]$ is the same for all bidders $i$. The latter assumption is typically maintained in the literature [see, e.g. Campo, Perrigne and Vuong (2003)]. In independent private values models, it ensures that $\operatorname{supp}[B_i]$ is the same for all bidders $i$ [Lebrun (1999)] and that the equilibrium is unique [Lebrun (1999), Bajari (2001)]. However, plausible specifications of primitives would lead to distributions of bids that violate condition (e), so it may be useful to relax that assumption in practice.

Let us briefly consider some examples. Consider maintaining the assumption that $\inf[\operatorname{supp}[U_i]]$ is the same for all $i$, but allow $\bar{u}_i = \sup[\operatorname{supp}[U_i]]$ to vary with $i$. For affiliated private values models, there exists an equilibrium in nondecreasing strategies where $\text{supp}[B_i] \subseteq \text{supp}[\max_{j \neq i} B_j]$, and these supports are convex and compact; in addition, in any equilibrium (mixed or pure) strategies must be strictly increasing (i.e., separating) [Maskin and Riley (2003), McAdams (2007)]. Thus, the distributions of valuations can be identified using (3.1). For example, suppose $n = 2$, $U_1$ is uniformly distributed on $[0, 3/2]$, and $U_2$ is independent of $U_1$, with distribution $F_{U_2}(u_2) = (1/4)u_2^2$ on the support $[0, 2]$. Then, it can be shown that $\text{supp}[B_i] = [0, 1]$ for $i \in \{1, 2\}$, and that for $b \in [0, 1]$, $G_{B_1}(b) = b$ and $G_{B_2}(b) = b^2$. With these bid distributions, $g_{M_1}(1; \mathcal{N}) = 1$ while $g_{M_2}(1; \mathcal{N}) = 2$, violating the boundary condition (e). However, in this example, the $F_{U_i}(\cdot)$ each can be identified if we expand the set of permissible distributions of valuations to allow supports that vary across bidders.

If both $\inf[\text{supp}[U_i]]$ and $\sup[\text{supp}[U_i]]$ vary with $i$, then it is possible that some bidders never win in equilibrium. For example, if there are two bidders in an IPV auction, and $\text{supp}[U_1] = [0, 1]$ while $\text{supp}[U_2] = [100, 101]$, in equilibrium $B_2 = 1$ with probability 1, while $B_1 \leqslant 1$ Maskin and Riley (2000a). Clearly, very little can be said about the distribution of $U_2$ in this case. Despite the possibility of degenerate equilibria like this, Maskin and Riley (2000b) show that the distribution of winning bids, $G_B^{(n:n)}(\cdot)$, is continuous on its support. This implies that a mass point in $G_B^{(n:n)}(\cdot)$ such as the one in the latter example can only occur either (i) at the bottom of the support if the support of winning bids is nondegenerate, or (ii) if the support of winning bids is degenerate. Thus, outside of cases (i) and (ii), the equilibrium must be in strictly increasing strategies on $\text{supp}[B^{(n:n)}]$, so that it will be possible to recover the distribution of bidders' valuations on the pre-image of the interior of $\text{supp}[B^{(n:n)}]$ using (3.1). This would lead to a partial identification result.

For the remainder of the chapter, we follow the existing literature and maintain the assumption that valuation distributions have the same support, while noting that many of the results generalize.

### 5.2. *Testing monotonicity of bid functions and boundary conditions*

Here we consider two possible types of tests based on Theorem 5.1. Guerre, Perrigne and Vuong (2000) have suggested a specification test based on the observation that the right-hand side of bidder $i$'s first-order condition (2.4), i.e.,

$$\xi_i(b_i, \mathcal{N}) \equiv b_i + \frac{G_{M_i|B_i}(b_i|b_i; \mathcal{N})}{g_{M_i|B_i}(b_i|b_i; \mathcal{N})}$$

is the inverse of his equilibrium bidding strategy. This is true for private and common value auctions.[41] Since bidding strategies must be strictly increasing, so must $\xi_i(\cdot, \mathcal{N})$.

[41] One way to see this in the common values case is to note that one possible normalization of signals sets $X_i = v_i(X_i, X_i; \mathcal{N})$, so that the first order condition may still be directly interpreted as giving bidder $i$'s inverse bidding strategy.

Although testing monotonicity of $\xi_i(\cdot, \mathcal{N})$ is conceptually straightforward, no formal statistical test has been developed for this problem. Existing tests of monotonicity in the statistics literature[42] are not directly applicable due to the fact that realizations of the random variables $\xi_i(B_i, \mathcal{N})$ are estimated rather than observed directly. In applications, researchers often find few (if any) violations of strict monotonicity [e.g., Hendricks, Pinkse and Porter (2003), Haile, Hong and Shum (2003)], in which case no formal test would reject. Nonetheless, formal tests could be valuable.

Two things about such a test should be noted, however. First, the alternative hypothesis is simply that some component of the specification is incorrect. A failure of monotonicity may indicate the presence of unobserved heterogeneity, risk aversion, nonequilibrium bidding behavior, or violation of some other maintained assumption. In general, testing one assumption will require maintaining others, so many of the other specification tests discussed below will share this limitation. Second, no test of this hypothesis will be consistent against all violations of the maintained assumptions. In particular, one can easily construct examples in which one or more maintained assumptions are violated, but monotonicity of $\xi_i(\cdot, \mathcal{N})$ still holds.

Another potential specification test is based on the boundary condition (e) from Theorem 5.1. This restriction can be simplified in the case of the IPV model. Let the common support of the bid distribution be denoted $\operatorname{supp}[B_i] = [\underline{b}, \overline{b}]$. Then, the boundary condition requires

$$g_{M_i}(\overline{b}; \mathcal{N}) = g_{M_j}(\overline{b}; \mathcal{N}) \quad [\forall i, j \in \mathcal{N}] \tag{5.1}$$

which is a testable restriction.

### *5.3. Multi-sample tests with exogenous variation in participation*

Athey and Haile (2002) discuss a different principle for specification testing that can be used in both first-price and ascending auctions whenever (a) there is exogenous variation in the number of bidders, and (b) the underlying model is identified with a fixed number of bidders. For simplicity, consider the case of symmetric bidders, although the same principle applies to asymmetric settings. Let $\widehat{F}_{\mathbf{U}}(\mathbf{u}; n)$ denote a consistent estimator of $F_{\mathbf{U}}(\mathbf{u})$ obtained using data from $n$-bidder auctions. With exogenous variation in the number of bidders, for $n' \neq n$, $\widehat{F}_{\mathbf{U}}(\mathbf{u}; n)$ should equal $\widehat{F}_{\mathbf{U}}(\mathbf{u}; n')$ up to sampling error. Hence a test of the null hypothesis of equal distributions provides a specification test.

While testing equality of distributions is a standard problem [e.g., McFadden (1989)], complications arise both in ascending and first-price auctions. In an ascending auction, the complication is the fact that identification relies on mappings between distributions of order statistics and the underlying marginal distributions (Theorem 4.1). Hence, asymptotic distributions of test statistics must account for this transformation of the data.

[42] See, e.g., Bowman, Jones and Gijbels (1998), Gijbels et al. (2000), or Hall and Heckman (2000).

In a first-price auction, the complications are more challenging, arising from the fact that valuations are estimated rather than observed directly. This first-stage nonparametric estimation of $F_{\mathbf{U}}(\cdot)$ introduces nontrivial complications to the asymptotic theory needed for inference. Haile, Hong and Shum (2003) develop several formal tests applicable when all bids are observed, based on comparisons of different estimates of the marginal distribution $F_U(\cdot)$ obtained from auctions with different number of bidders.[43] In models identified with partially observed bids, similar tests may be applicable, although this has not yet been explored.

### 5.4. Multi-sample tests with multiple order statistics

In IPV settings, a variation on the type of testing approach above may be available without exogenous variation in participation. In an IPV auction each marginal distribution $F_{U_i}(\cdot)$ is identified from observation of the transaction price (and bidder identities if the environment is asymmetric) in both ascending and first-price auctions. Athey and Haile (2002) have shown that observation of any other order statistic $B^{(j:n)}$ can be substituted for observability of the transaction price – in a symmetric environment, for example, this follows from (4.1). When two or more order statistics (e.g., the top two bids) are observed, the estimates of $F_{U_i}(\cdot)$ implied by each of these should be identical up to sampling error.

### 5.5. Direct tests of exchangeability or independence

There are other potential approaches to specification testing when bidders are assumed to be symmetric or types are assumed independent. With symmetric bidders, the joint distribution of bidder valuations is exchangeable and each bid $B_i = \beta(U_i; n)$. Hence, the joint distribution of bids must also be exchangeable. When bidder identities are observed, there are several ways to approach testing such a hypothesis. One is to test exchangeability of the bids (or subsets of bids) directly. Nonparametric tests from the statistics literature may be directly applicable. For example, Romano (1988, 1989) suggests tests based on the supremum distance between the values of a multivariate CDF evaluated at permutations of its arguments.

One implication of exchangeability is equality of marginal distributions. For example, in a symmetric model, any subset of bidders should have bids governed by the same marginal distribution as those of another subset of bidders. A standard Kolmogorov–Smirnov test of equal distributions could then be applied.

Alternative tests may be useful when covariates are available and additional structure is assumed. Suppose, for example, that valuations are assumed to have the structure

$$U_{it} = h(Z_{1t}, Z_{2i}, \mathbf{Z}_{2(-i)}, A_{it})$$

[43] They focus on tests of the private values hypothesis. However, their tests, which are based on comparisons of the empirical distributions of pseudo-values for auctions with different numbers of bidders, could be directly applied.

where $Z_{1t}$ is an auction-specific covariate, $Z_{2i}$ is a bidder-specific covariate, $\mathbf{Z}_{2(-i)}$ denotes the bidder-specific covariates of $i$'s opponents, and $A_{it}$ is a private idiosyncratic factor. The restriction to scalar covariates is only for expositional simplicity. Assume further that the conditional distribution function $F_{\mathbf{A}}(A_1, \ldots, A_n | Z_1, Z_{21}, \ldots, Z_{2n})$ is exchangeable in the indices $(1, \ldots, n)$. Loosely speaking, with this structure, all bidder valuations are affected in the same way by covariates. In particular, the distribution of bidder $i$'s valuation conditional on $(Z_1, Z_{2i}, \mathbf{Z}_{2(-i)})$ is the same for all $i$. Since bids are equal to valuations in an ascending auction, this can be tested, for example, by examining coefficient estimates in a regression of bids on covariates (auction- and bidder-specific) interacted with bidder dummies (or indicators for different "classes" of bidders).

In a first-price auction, the structure above implies that the distribution of $\max_{j \neq i} B_j$ is the same for all $i$ conditional on $(Z_1, Z_{2i}, \mathbf{Z}_{2(-i)})$. Hence, the distribution of $i$'s bids should depend only on $(Z_1, Z_{2i}, \mathbf{Z}_{2(-i)})$, not on the index $i$ itself. This may again be evaluated in a regression. Bajari and Ye (2003) apply these regression-based approaches in their analysis of highway construction contracts [see also Porter and Zona (1993, 1999)].

Note that similar restrictions will hold in a common values model, where it is the joint distribution of the random variables

$$v_i(X_i, X_i, \mathcal{N})$$

that must be exchangeable. As we will see below, this distribution will often be identified in a common values model, even though $F_{\mathbf{U},\mathbf{X}}(\cdot)$ is not identified. Hence, specification testing may be possible even for under-identified models.

Another direct approach to specification testing is applicable in first-price auctions in the widely used independent private values model (symmetric or asymmetric). Since each $B_i$ is a measurable function of $U_i$, bids must also be independent. In a first-price sealed-bid auction in which all bids are observed, one can directly test this restriction using standard nonparametric tests [Guerre, Perrigne and Vuong (2000)]. For example, Romano (1988, 1989) suggests tests based on the supremum distance between an estimated joint distribution and the joint distribution obtained as the product of the estimated underlying marginal distributions.[44] In practice, it is typical to assume that valuations, and thus bids, are independent conditional on a set of auction-specific and perhaps bidder-specific covariates. Su and White (2003) propose a testing approach that may then be applicable. An alternative is to test for correlation of residuals from a regression of bids on bidder-specific or auction-specific covariates. Bajari and Ye (2003) do this in their analysis of highway construction procurement auctions. In an ascending auction, the problem of partially observed bids appears to make direct testing impossible (recall, however, the indirect tests discussed in Section 5.4).

44 Other tests of the hypothesis that bids are uncorrelated (an implication of independence) could also be applied. See, e.g., Chapter 8 of Hollander and Wolfe (1999).

## 6. Extensions of the basic results

### 6.1. Auction heterogeneity

#### 6.1.1. Observed auction heterogeneity

In practice, one rarely has access to data from auctions of identical objects. For example, the goods for sale at each auction often differ in observable characteristics, and we may expect distributions of valuations to shift with these observables. All of the identification results above hold in the presence of auction-specific covariates. In particular, the previous discussion can be reinterpreted as being conditioned on a given realization of the covariate values. To make this concrete, let $\mathbf{Z}$ be a vector of auction covariates. We extend the notation defined above to condition on $\mathbf{Z}$ by defining $\beta_i(\cdot;\mathcal{N},\mathbf{Z})$, $F_{\mathbf{U}}(\cdot|\mathbf{Z})$, $G_{M_i|B_i}(b|b;\mathcal{N},\mathbf{Z})$ and $g_{M_i|B_i}(b|b;\mathcal{N},\mathbf{Z})$, etc. Assuming all auction-specific heterogeneity is captured by $\mathbf{Z}$, in a first-price auction the first-order condition for bidder $i$ at auction $t$ becomes

$$u_{it} = b_{it} + \frac{G_{M_i|B_i}(b_{it}|b_{it};\mathcal{N},\mathbf{z}_t)}{g_{M_i|B_i}(b_{it}|b_{it};\mathcal{N},\mathbf{z}_t)} \tag{6.1}$$

which uniquely determines $F_{\mathbf{U}}(\cdot|\mathbf{z}_t)$ in the affiliated private values model when all bids and bidder identities are observable. In an ascending auction with private values that are independent conditional on $\mathbf{Z}_t$, one can use the conditional distribution of transaction prices $F_U^{(n-1:n)}(\cdot|\mathbf{z}_t)$ for any given value of $\mathbf{z}_t$ to uniquely determine $F_{\mathbf{U}}(\cdot|\mathbf{z}_t)$ through Equation (4.4).

The nonparametric estimation methods discussed above can also be extended, for example by using standard kernel smoothing over covariates. Guerre, Perrigne and Vuong (2000) discuss details of such an approach for the case of a first-price auction with symmetric independent private values, and this approach is easily extended to the other models. This type of approach has been applied to ascending auctions by Haile and Tamer (2003).

Unless the dimensionality of the covariates is fairly small relative to the sample size, however, a fully general nonparametric estimation approach may not be practical. One alternative suggested by Haile, Hong and Shum (2003) exploits the observation that additive (or multiplicative) separability is preserved by equilibrium bidding.[45] In particular, suppose that in an auction with characteristics $\mathbf{z}_t$ valuations are given by

$$u_{it} = \Gamma(\mathbf{z}_t) + a_{it} \tag{6.2}$$

for some (possibly unknown) function $\Gamma(\cdot)$, with the bidder-specific private information $A_{it}$ independent of $\mathbf{Z}_t$. Then, if we let $\mathbf{z}_0$ be such that[46]

[45] This approach has been applied by Krasnokutskaya (2004), Bajari, Houghton and Tadelis (2004), and Shneyerov (2005).

[46] We assume for simplicity that such a $\mathbf{z}_0$ exists. If it does not, the argument extends but with more cumbersome notation.

$$\Gamma(\mathbf{z}_0) = 0 \tag{6.3}$$

equilibrium bidding also follows the additively separable structure (an analogous result applies in the case of multiplicative separability)

$$\beta_i(u_i; \mathcal{N}, \mathbf{z}) = \Gamma(\mathbf{z}) + \beta_i(u_i; \mathcal{N}, \mathbf{z}_0). \tag{6.4}$$

Proving this is trivial in an ascending auction, where the bid function is the identity function. For a first-price sealed-bid auction, let

$$\breve{\beta}_i(a_i, \mathbf{z}; \mathcal{N}) \equiv \beta_i\big(a_i + \Gamma(\mathbf{z}); \mathcal{N}, \mathbf{z}\big)$$

so that under (6.2) a bidder's first-order condition can be written

$$a_{it} + \Gamma(\mathbf{z}_t) = \breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N}) + \frac{G_{M_i|B_i}(\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N})|\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N}); \mathcal{N}, \mathbf{z}_t)}{g_{M_i|B_i}(\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N})|\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N}); \mathcal{N}, \mathbf{z}_t)}. \tag{6.5}$$

Note that the events $\{\breve{\beta}_i(A_i, \mathbf{z}; \mathcal{N}) = \breve{\beta}_i(a_i, \mathbf{z}; \mathcal{N})\}$ and $\{\breve{\beta}_i(A_i, \mathbf{z}_0; \mathcal{N}) = \breve{\beta}_i(a_i, \mathbf{z}_0; \mathcal{N})\}$ are equivalent for any $\mathbf{z}$. Under (6.4), the events $\{\breve{\beta}_j(A_j, \mathbf{z}; \mathcal{N}) = \breve{\beta}_i(a_i, \mathbf{z}; \mathcal{N})\}$ and $\{\breve{\beta}_j(A_j, \mathbf{z}_0; \mathcal{N}) = \breve{\beta}_i(a_i, \mathbf{z}_0; \mathcal{N})\}$ are also equivalent for $j \neq i$, so the expression

$$\frac{G_{M_i|B_i}(\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N})|\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N}); \mathcal{N}, \mathbf{z}_t)}{g_{M_i|B_i}(\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N})|\breve{\beta}_i(a_{it}, \mathbf{z}_t; \mathcal{N}); \mathcal{N}, \mathbf{z}_t)}$$

on the right-hand side of (6.5) is invariant to $\mathbf{z}_t$. Hence, (6.4) guarantees that (6.5) is satisfied for all $\mathbf{z}_t$ whenever it is for $\mathbf{z}_t = \mathbf{z}_0$.

This preservation of additive separability is useful because it implies that the effects of covariates on valuations can be controlled for using a regression of bids on covariates. In particular, we can write

$$b_{it} = \alpha(\mathcal{N}_t) + \Gamma(\mathbf{z}_t) + \epsilon_{it}, \tag{6.6}$$

where $\alpha(\mathcal{N}_t)$ is an intercept specific to auctions in which the set of bidders is $\mathcal{N}_t$ (in a symmetric environment, this can be $\alpha(n_t)$) and $\epsilon_{it} \equiv \beta_i(u_{it}; \mathcal{N}_t, \mathbf{z}_0) - \alpha(\mathcal{N}_t)$ has mean zero conditional on $\mathbf{z}_t$. Both $\alpha(\mathcal{N}_t)$ and $\Gamma(\mathbf{z}_t)$ are then identified from observation of bids, $\mathcal{N}_t$, and $\mathbf{z}_t$; indeed, they can be estimated consistently using standard regression techniques.

Let $\widehat{\Gamma}(\mathbf{z}_t)$ denote a consistent estimate of $\Gamma(\mathbf{z}_t)$. Then $b_{it} - \widehat{\Gamma}(\mathbf{z}_t)$ provides a consistent estimate of $\beta_i(u_{it}; \mathcal{N}_t, \mathbf{z}_0)$, i.e., the bid $i$ *would have submitted* in auction $t$ if $\mathbf{Z}_t$ were equal to $\mathbf{z}_0$. Of course, a sample of bids from auctions with the same value of $\mathbf{Z}$ of is exactly what we would like to have. Estimation of (6.6) provides an approach for "homogenization" of the bid data by replacing each $b_{it}$ with

$$b_{it}^h = b_{it} - \widehat{\Gamma}(\mathbf{z}_t).$$

These homogenized bids can then be used to consistently estimate the underlying distribution of valuations $F_{\mathbf{U}}(\cdot; \mathbf{z}_0)$ using the methods described in the previous sections;

i.e., with $\Gamma(\cdot)$ known, $F_{\mathbf{U}}(\cdot; \mathbf{z}_0)$ is identified through (6.1). Finally, since (6.3) and (6.2) imply

$$\Pr(U_{1t} \leqslant u_1, \ldots, U_{nt} \leqslant u_n) = F_{\mathbf{U}}\big(u_1 - \Gamma(\mathbf{z}_{it}), \ldots, u_n - \Gamma(\mathbf{z}_{it}); \mathbf{z}_0\big),$$

$F_{\mathbf{U}}(\cdot; \mathbf{z})$ is then identified for all $\mathbf{z}$ in the support of the auction covariates. As usual, in a first-price auction, equilibrium bidding implies that the distribution of the mean-zero $\epsilon_{it}$ will vary with $\mathcal{N}_t$. So the "second stage" of estimating the joint distribution $F_{\mathbf{U}}(\cdot; \mathbf{z}_0)$ must be done separately for each $\mathcal{N}_t$.

The number of observations available for the first-stage regression of bids on covariates is $\sum_n nT_n$, which is often quite large. Hence, a nonparametric or flexible parametric specification of $\Gamma(\cdot)$ will be feasible in data sets of reasonable size. Assuming that $\Gamma(\cdot)$ is known up to a finite parameter vector has an advantage for some purposes in that estimates from the first stage will converge at the parametric rate, leaving the asymptotic distribution of nonparametric estimators applied to the homogenized sample unaffected. Note that the function $\Gamma(\cdot)$, which characterizes the effects of covariates on valuations, is sometimes of direct interest itself. Equation (6.4) implies that one can estimate this primitive directly with a regression of bids on covariates. Bajari, Houghton and Tadelis (2004), for example, have recently exploited this observation to investigate the importance of renegotiation costs in procurement auctions.

This approach preserves the fully nonparametric specification of the idiosyncratic component of bidders' private values and allows direct inference (through the first-stage estimates) on the way observables affect valuations. However, it places a strong restriction on the way observables enter. An alternative nonparametric approach is to use series or sieves [e.g., Chen (2007)], approximating the bid distribution with a sequence of parametric models. In a given data set this will amount to assuming a flexible parametric model, and one might also take such an approach directly. For example, in an ascending auction with symmetric independent private values, one might specify the conditional distribution $F_U^{(n-1:n)}(u|\mathbf{z})$ as a finite mixture of parametric distributions. Letting $H(\cdot; \boldsymbol{\gamma})$ be a parameterized distribution function, the distribution of the transaction price could be specified as

$$F_U^{(n-1:n)}(u|\mathbf{z}, \boldsymbol{\theta}, J) = \frac{1}{\sum_{j=1}^{J} \omega(\mathbf{z}; \boldsymbol{\theta}_j)} \sum_{j=1}^{J} \omega(\mathbf{z}; \boldsymbol{\theta}_j) H\big(u; \boldsymbol{\gamma}(\mathbf{z}; \boldsymbol{\theta}_j)\big) \tag{6.7}$$

given parametric specifications of the functions $\boldsymbol{\gamma}(\cdot)$ and $\omega(\cdot)$. Given an estimate $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_J)$, Equation (4.9) implies that

$$\widehat{F}_U(u|\mathbf{z}) \equiv \phi\big(F_U^{(n-1:n)}(u|\mathbf{z}; \hat{\boldsymbol{\theta}}, J); n-1, n\big) \tag{6.8}$$

would provide a consistent estimator of $F_U(u|\mathbf{z})$ under (6.7).

In a first-price sealed-bid auction, a similar approach might be applied. For a given set of bidders $\mathcal{N}$, the conditional distribution $G_{M_i|B_i}(m_i|b_i; \mathcal{N}, \mathbf{z})$ could be assumed to have the parametric form

$$G_{M_i|B_i}(m_i|b_i;\mathcal{N},\mathbf{z},\boldsymbol{\theta},J)=\sum_{j=1}^{J}\frac{\omega(b_i,\mathbf{z};\boldsymbol{\theta}_j)}{\sum_{j'=1}^{J}\omega(b_i,\mathbf{z};\boldsymbol{\theta}_{j'})}H\big(m;\boldsymbol{\gamma}(b_i,\mathbf{z};\boldsymbol{\theta}_j)\big) \tag{6.9}$$

providing a flexible parametric first step of the two-step estimation procedure discussed in Section 3.2. This kind of specification allows the distribution of bids to vary with auction covariates in richer ways than is allowed by the "homogenization" approach described above. This flexibility comes at the price of placing *a priori* structure on the distributions of bids and valuations. Of course, some approximation must always be used in a finite sample, and a finite mixture may perform well in practice. Note that here the effects of covariates on valuations, which are often of primary interest, would be obtained indirectly, through (6.8) or through (6.1) and (6.9).

### 6.1.2. *Unobserved auction heterogeneity*

In many applications one may suspect that there are factors affecting bidder valuations that are common knowledge among bidders but unobserved by the econometrician. For example, suppose valuations are given by the equation

$$U_i = V_0 + A_i. \tag{6.10}$$

Even if the idiosyncratic components $A_1, \dots, A_n$ are i.i.d., the valuations $U_1, \dots, U_n$ will be correlated unconditional on $V_0$ (they will be affiliated if the densities $f_{A_i}(\cdot)$ are log-concave). When bidders' information consists only of their valuations $u_i$, not the individual components $v_0$ and $a_i$, this merely provides one motivation for an affiliated private values model. When bidders observe both $v_0$ and $a_i$, however, the situation can be more complicated.

As noted above, information regarding an auction that is common knowledge among the bidders creates no problem for the characterization of equilibrium bidding strategies – the theory can be thought of as holding for each value of the public information. However, for empirical work, difficulties can arise when the econometrician is unable to condition on all the information that is public to bidders.

There are at least three issues that arise in the presence of unobserved heterogeneity. The first is whether unobserved heterogeneity is empirically distinguishable from other structures that introduce correlation among bids. In an ascending auction, equilibrium bids satisfy $B_i = U_i$ regardless of whether bidders observe only their own valuations or also factors shifting all bidders' valuations. Hence it will be impossible to distinguish an environment with unobserved heterogeneity from an environment with correlated private values but no unobserved heterogeneity. In a first-price auction, the situation is similar. As long as the conditions of Theorem 5.2 hold, the data can be rationalized by equilibrium bidding. However, unobserved heterogeneity can account for some or all of the observed correlation (if any) among bids.

This observational equivalence can be important. If (6.10) holds, for example, an assumption about whether each bidder $i$ observes only $U_i$ or also $V_0$ can have significant implications for bidding strategies (and, therefore, the appropriate interpretation of bids) in a first-price auction. Thus, one must rely on an assumption regarding which model is appropriate.[47] In her application to highway procurement auctions, Krasnokutskaya (2004) compares the estimated bid function under an assumption of affiliated private values (with no unobserved heterogeneity) to the average (over the unobserved heterogeneity) bid function under the assumption of independent private values with unobserved heterogeneity. She finds that the estimated average bid function under unobserved heterogeneity is steeper than the estimated bid function under affiliated private values, and that estimated average markups are substantially higher when one ignores unobserved heterogeneity. Hence, the modeling choice can have important implications.

The second issue is whether the joint distribution $F_{\mathbf{U}}(\cdot)$, is identified under the assumption of unobserved heterogeneity. We will see that in a first-price sealed-bid auction, identification can be obtained through additional structure, e.g., on the statistical and functional relationships between the unobserved heterogeneity and bidder valuations or on the effects of unobserved heterogeneity on observed outcomes other than bids. In an ascending auction, the available identification results require additional sources of variation in the data, such as bidder-specific covariates.

The third issue is whether identification of $F_{\mathbf{U}}(\cdot)$ is adequate for the economic questions one wishes to answer. In the presence of unobserved heterogeneity, knowledge of this distribution is sufficient to answer some important questions but not others – in particular, not those concerning outcomes that depend on bidders' beliefs about opponents' valuations, since these beliefs vary with the realization of the factor that is unobservable to the econometrician. In ascending or second-price auctions (or any mechanism with a dominant strategy equilibrium), $F_{\mathbf{U}}(\cdot)$ is the only primitive relevant for predicting equilibrium outcomes, designing the auction rules, or performing counterfactual simulations. However, if we wish to consider policy questions concerning first-price auctions or other mechanisms in which beliefs play a more significant role, it will be necessary to know the joint distribution of bidders' private information and the unobserved heterogeneity (e.g., the distribution $F_{\mathbf{A},V_0}(\cdot)$ if one assumes (6.10)), not just $F_{\mathbf{U}}(\cdot)$. Below we will discuss conditions under which this joint distribution is identified.

[47] Although the literature has not yet considered approaches for distinguishing between the two models, it may be possible to develop a test based on exogenous variation in participation. In particular, it is possible to estimate the primitives of each model for a fixed set of potential bidders $\mathcal{N}$. Then, these primitives can be used to make "out of sample" predictions about bid distributions for other sets of potential bidders (e.g., a subset of the original set $\mathcal{N}' \subset \mathcal{N}$). We conjecture that in general the specific bid distributions predicted by the two models for the set of bidders $\mathcal{N}'$ will differ across the two models. However, to our knowledge this has not been formally analyzed. Note that a test of this restriction would rely on the assumption that participation does not vary with the unobserved heterogeneity. This assumption may be strong in practice; it may be satisfied, however, if bidders pay a cost to acquire a signal and the unobserved heterogeneity is not observed by bidders until they bear the cost of investigating the auction. Instrumental variables approaches like that explored in Haile, Hong and Shum (2003) may also be useful.

*6.1.2.1. First-price sealed-bid auctions* To demonstrate the problem of unobserved heterogeneity in first-price auctions, we begin with a very general case. In a private values first-price sealed-bid auction, suppose that information $\mathbf{w}_t$ is common knowledge among the bidders at auction $t$. Following the discussion in Section 6.1.1, the first-order condition relating bids to the underlying valuations is

$$u_{it} = b_{it} + \frac{G_{M_i|B_i}(b_{it}|b_{it}; \mathcal{N}, \mathbf{w}_t)}{g_{M_i|B_i}(b_{it}|b_{it}; \mathcal{N}, \mathbf{w}_t)}. \tag{6.11}$$

If the econometrician does not observe $\mathbf{w}_t$, the conditional distribution $G_{M_i|B_i}(b_{it}|b_{it}; \mathcal{N}, \mathbf{w}_t)$ is not identified. This creates a serious challenge to any attempt to uncover the markdown $\frac{G_{M_i|B_i}(b_{it}|b_{it};\mathcal{N},\mathbf{w}_t)}{g_{M_i|B_i}(b_{it}|b_{it};\mathcal{N},\mathbf{w}_t)}$. Indeed, because this markdown is a nonlinear function of $\mathbf{w}_t$, even the average markdown is not identified in general [Hendricks, Pinkse and Porter (2003)].

Identification requires additional structure, and several possibilities have been explored in the literature. All begin by assuming that the unobservable is a scalar, which we will denote by $W$. Some early work took parametric approaches to disentangling the common shock $W$ from idiosyncratic factors, but more recently nonparametric identification results have been derived, exploiting additional data and/or assumptions about the way common shocks affect outcomes.

One approach, first proposed by Campo, Perrigne and Vuong (2003), is to exploit observables that are sufficient for the unobserved factor. This can be natural when there is an observable endogenous variable besides bids that responds to the unobservable $W$.[48] Both Campo, Perrigne and Vuong (2003) and Haile, Hong and Shum (2003) have used this approach by positing a model in which the number of bidders in auction $t$ can be represented as a function of observables $\mathbf{Z}_t$ and the unobservable $W_t$:

$$N_t = \alpha(\mathbf{Z}_t, W_t). \tag{6.12}$$

If $\alpha(\mathbf{z}, \cdot)$ is a strictly increasing function for all $\mathbf{z}$, then the joint distribution of $(X_1, \ldots, X_n, U_1, \ldots, U_n)$ conditional on $(N_t, \mathbf{Z}_t)$ is identical to that conditional on $(N_t, \mathbf{Z}_t, W_t)$. Then

$$\begin{aligned}
v(x, x; n, w, z) &\equiv E[U_i \mid X_i = x,\ N_t = n,\ W_t = w,\ \mathbf{Z}_t = \mathbf{z}] \\
&= E[U_i \mid X_i = x,\ N_t = n,\ \mathbf{Z}_t = \mathbf{z}] \\
&\equiv v(x, x; n, \mathbf{z})
\end{aligned}$$

and identification follows from the first-order condition

[48] A similar idea was used to address the problem of identification with unobserved heterogeneity in a very different environment by Olley and Pakes (1996).

$$v(x_{it}, x_{it}; n_t, \mathbf{z}_t) = b_{it} + \frac{\Pr(\max_{j\neq i} B_{jt} \leqslant b_{it} \mid B_{it} = b_{it},\ \mathbf{Z}_t = \mathbf{z}_t,\ N_t = n_t)}{\frac{\partial}{\partial m} \Pr(\max_{j\neq i} B_{jt} \leqslant m \mid B_{it} = b_{it},\ \mathbf{Z}_t = \mathbf{z}_t,\ N_t = n_t)|_{m=b_{it}}},$$

where the right-hand side is a known function of observables.

The assumption of strict monotonicity of $N$ in $W$ is strong although it is clear that there must be an invertible relation between $W$ and the observables for this kind of approach. With weak monotonicity, conditioning on $(N_t, \mathbf{Z}_t)$ would limit the realization of $W_t$ to some set $\mathcal{W}(N_t, \mathbf{Z}_t)$, and in some applications this might be sufficient to use the first-order condition above as a useful approximation.

The economic interpretation of (6.12) can be important when taking this kind of approach. For example, to predict outcomes under alternative selling mechanisms, one must consider whether changing mechanisms would alter the relation between bidder participation and $\mathbf{Z}$ [see, e.g., Athey, Levin and Seira (2004)]. If so, one would need a fully specified economic model of participation and bidding. However, a reduced form may be adequate for some questions and applications – for example, when (6.12) describes the determination of matches between auctions and potential bidders based on unobserved characteristics of the object offered for sale, or when the economic questions of interest do not depend on counterfactual predictions regarding participation.

Other approaches to handling unobserved heterogeneity in a first-price auction are closely related to ideas from the econometrics literatures on measurement error with repeated measures [Li and Vuong (1998), Li (2002), Schennach (2004)] and duration models with unobserved heterogeneity and multiple spells [see, e.g., Lancaster (1990)]. These literatures consider multiple observations for each of many units, with observations within each unit reflecting both a common (unobserved) shock as well as idiosyncratic shocks. In the auction setting, the auction plays the role of the unit, with the individual bids being the observations within unit.

Consider a simplified model of unobserved heterogeneity in which bidder valuations take the additively separable form in (6.10), and $(A_1, \ldots, A_n, V_0)$ are mutually independent with compact support. This is a special case of a *conditionally independent private values* model (itself a special case of affiliated private values, so long as each $f_{A_i}(\cdot)$ is log-concave).[49] Li, Perrigne and Vuong (2000) considered this structure under the assumption that bidders observe only their valuations $U_i$. They showed that, in that case, the joint distribution $F_{\mathbf{A}, V_0}(\cdot)$ is nonparametrically identified up to a location normalization. While there is no unobserved heterogeneity in their model, their approach

[49] In a general specification of conditionally independent private values, one would assume only $F_{\mathbf{A},V_0}(a_1, \ldots, a_n, v_0) = F_{V_0}(v_0) \prod_{i=1}^{n} F_{A_i}(a_i|v_0)$. With this more general specification, the linearity assumed in (6.10) would be without loss of generality, since whatever the distributions of $U_i|V_0$, one can let $A_i = U_i - V_0$. The de Finetti Theorem [e.g., Chow and Teicher (1997)] tells us that any infinite sequence of exchangeable random variables can be represented by this more general conditionally independent structure. However, finite exchangeable sequences, like those arising in symmetric auctions with a finite number of potential bidders, need not have such a representation. Athey and Haile (2000, Proposition 4) explore limitations of the flexibility of the more restrictive conditionally independent structure considered here.

turns out to be a useful starting point. To see the idea behind their result, recall that observation of all bids and bidder identities is sufficient to identify the joint distribution $F_{\mathbf{U}}(\cdot)$ in a first-price auction with affiliated private values. Once $F_{\mathbf{U}}(\cdot)$ is known, a result from the literature on measurement error can be applied to separately identify the component distributions $F_{\mathbf{A}}(\cdot)$ and $F_{V_0}(\cdot)$ up to a location normalization. Li, Perrigne and Vuong (2000) develop consistent nonparametric estimators for this environment using empirical characteristic functions.[50]

Krasnokutskaya (2004) shows that a very similar approach can be applied in the case of unobserved heterogeneity – i.e., when valuations take the additively separable form in (6.10) and $v_0$ is observed by bidders but not the econometrician.[51] In essence, she reverses the two steps of Li, Perrigne, and Vuong's (2000) approach: she first uses a deconvolution technique to remove the effects of unobserved heterogeneity from bids, then recovers the idiosyncratic factors $a_i$ through the first-order condition for a hypothetical auction with no unobserved heterogeneity. In this sense, the approach is similar to the "homogenization" approach for incorporating observable auction heterogeneity, discussed in Section 6.1.1.

For the first step, recall from Section 6.1.1 that the additive separability in (6.10) is preserved by equilibrium bidding.[52] So if $\beta_i(u_{it}; \mathcal{N}_t, v_{0t})$ denotes bidder $i$'s equilibrium bid given $u_{it}$, $\mathcal{N}_t$, and $v_{0t}$, then

$$\beta_i(u_{it}; \mathcal{N}_t, v_{0t}) = \beta_i(u_{it} - v_{0t}; \mathcal{N}_t, 0) + v_{0t}. \tag{6.13}$$

If one observes all bids from each auction, the following result from Kotlarski (1966) implies identification of the joint distribution of $(\beta_1(A_1; \mathcal{N}, 0), \ldots, \beta_n(A_n; \mathcal{N}, 0), V_0)$ up to a location normalization.[53]

LEMMA 6.1. *Let $Y_1$, $Y_2$, and $Y_3$ be mutually independent random variables with nonvanishing characteristic functions $\phi_1(\cdot)$, $\phi_2(\cdot)$, and $\phi_3(\cdot)$, respectively. Let $Q_1 = Y_1 + Y_3$, $Q_2 = Y_2 + Y_3$. Then*

(i) *the joint distribution of $(Q_1, Q_2)$ completely determines the distributions of $Y_1$, $Y_2$, and $Y_3$ up to location*;

(ii) *if $\psi(\cdot,\cdot)$ denotes the characteristic function of $(Q_1, Q_2)$, and $\psi_i(\cdot,\cdot)$ is its derivative with respect to its $i$th argument, then under the normalization $E[Y_1] = 0$, $\phi_3(t) = \exp\{\int_0^t \frac{\psi_1(0,s)}{\psi(0,s)}\,\mathrm{d}s\}$, $\phi_1(t) = \frac{\psi(t,0)}{\phi_3(t)}$, and $\phi_2(t) = \frac{\psi(0,t)}{\phi_3(t)}$.*

[50] See also the discussion of a closely related special case of the mineral rights model in Section 7.2.1 below.

[51] An alternative is discussed in Section 8.2 below.

[52] Krasnokutskaya (2004) focuses on the case of multiplicative rather than additive separability. The analysis is equivalent with a logarithmic transformation. As she points out, a more general model allowing unobserved heterogeneity affecting both the location and scale of private values is also identifiable, since one can apply the deconvolution step (Lemma 6.1) to the bids twice – once in logs and once in levels.

[53] To see the connection to the original measurement error framework, observe that with an appropriate location normalization, under (6.10) each $A_i$ can be interpreted as an independent mean-zero measurement error on $v_0$.

A proof can be found in Prakasa-Rao (1992, Theorem 2.1.1 and Remark 2.1.11).[54] A key to the result is the fact that the characteristic function of the sum of independent random variables is the product of the characteristic functions of the component variables. With multiple observations involving one component in common, this separability can be exploited to isolate the characteristic functions (and, thereby, the distributions) of the individual components. Identification is up to a location normalization, since adding a constant to $Y_3$ and subtracting the same constant from $Y_1$ and $Y_2$ has no effect on observables.

Lemma 6.1 can be used to relate characteristic functions of the observed bids to those of the "homogenized" bids $\beta_1(A_1; \mathcal{N}, 0), \ldots, \beta_n(A_n; \mathcal{N}, 0)$ and the unobserved factor $V_0$. Identification of the distribution of each $A_i$ then follows from the first-order condition for a hypothetical auction in which $v_{0t} = 0$. In particular, if we let $B_i^0 = \beta_i(A_i; \mathcal{N}, 0) = B_i - V_0$,

$$A_i = B_i^0 + \frac{\Pr(\max_{j\neq i} \beta_j(A_j; \mathcal{N}, 0) \leqslant B_i^0)}{\frac{\partial}{\partial m} \Pr(\max_{j\neq i} \beta_j(A_j; \mathcal{N}, 0) \leqslant m)|_{m=B_i^0}} \equiv \tilde{\xi}_i\big(B_i^0; \mathcal{N}\big). \tag{6.14}$$

Note that unlike the case without unobserved heterogeneity, it is not possible to identify the valuations of bidders in a particular auction, because the realization of $V_0$ is unobserved. Despite this, because Lemma 6.1 implies that $G_{B_i^0}(\cdot)$ is identified, it follows that $\tilde{\xi}_i(\cdot; \mathcal{N})$ is also identified, so that the distribution of private information is given by

$$F_{A_i}(a_i) = G_{B_i^0}\big(\tilde{\xi}_i^{-1}(a_i; \mathcal{N})\big).$$

Nonparametric estimators can be developed by first substituting empirical characteristic functions for the population characteristic functions in part (ii) of Lemma 6.1, and then using simulation to construct pseudo-draws of the random variable on the right-hand-side of (6.14). We sketch the approach here. For simplicity, consider the special case in which there are two classes of bidders, with bidders in the same class drawing their valuations from the same marginal distribution (extension to more than two types is straightforward). Suppose one has a sample of $T$ auctions in which there are $n_1$ class-1 and $n_2$ class-2 bidders in each auction, and that all bids and bidder identities are observable. As above, estimation must be undertaken fixing the number of bidders of each type, which is equivalent here to fixing the set $\mathcal{N}$.

Let $c(j,t)$ denote the class of bidder $j$ in auction $t$. Impose the normalization $E[A_i] = 0$ for any class-1 bidder $i$. Let $G_{B^j}(\cdot)$ denote the marginal distribution of the equilibrium bid $B^j$ of a class-$j$ bidder, and let $G_{B^1,B^2}(b^1, b^2)$ denote the joint distribution of $(B^1, B^2)$. Similarly, let $B^{0,j} \equiv B^j - V_0$ denote the homogenized bid of a class-$j$ bidder. Note that the homogenized bids are independent. Let $\psi(\cdot,\cdot)$, $\phi_0(\cdot)$,

[54] See also Li and Vuong (1998, Lemma 2.1).

$\phi_{B^{0,1}}(\cdot)$ and $\phi_{B^{0,2}}(\cdot)$ denote the characteristic functions of $(B^1, B^2)$, $V_0$, $B^{0,1}$, and $B^{0,2}$, respectively.

Following Li and Vuong (1998) and Krasnokutskaya (2004) [see also Li, Perrigne and Vuong (2000)], define estimators

$$\hat{\psi}(\tau_1, \tau_2) = \frac{1}{Tn_1n_2} \sum_{t=1}^{T} \sum_{j:\, c(j,t)=1} \sum_{k:\, c(k,t)=2} \exp(\mathrm{i}\tau_1 b_{jt} + \mathrm{i}\tau_2 b_{kt}),$$

$$\hat{\psi}_1(\tau_1, \tau_2) = \frac{1}{Tn_1n_2} \sum_{t=1}^{T} \sum_{j:\, c(j,t)=1} \sum_{k:\, c(k,t)=2} \mathrm{i}b_{jt} \exp(\mathrm{i}\tau_1 b_{jt} + \mathrm{i}\tau_2 b_{kt}),$$

where, for each estimator, an average is taken over all possible pairs $(b^1, b^2)$. Let

$$\hat{\phi}_0(\tau) = \exp\left\{ \int_0^{\tau} \frac{\hat{\psi}_1(0, v)}{\hat{\psi}(0, v)} \,\mathrm{d}v \right\},$$

$$\hat{\phi}_{B^{0,1}}(\tau) = \frac{\hat{\psi}(\tau, 0)}{\hat{\phi}_0(\tau)},$$

$$\hat{\phi}_{B^{0,2}}(\tau) = \frac{\hat{\psi}(0, \tau)}{\hat{\phi}_0(\tau)}.$$

Given these estimated characteristic functions, one can obtain estimates of the marginal densities of $B^{0,1}$, $B^{0,2}$ and $V_0$ using the inverse Fourier transform. In particular, let

$$\hat{g}_{B^{0,i}}(b) = \frac{1}{2\pi} \int_{-\mu}^{\mu} \exp(-\mathrm{i}\tau b) \hat{\phi}_{B^{0,i}}(\tau) \,\mathrm{d}\tau, \tag{6.15}$$

$$\hat{f}_{V_0}(v) = \frac{1}{2\pi} \int_{-\mu}^{\mu} \exp(-\mathrm{i}\tau v) \hat{\phi}_{V_o}(\tau) \,\mathrm{d}\tau, \tag{6.16}$$

where $\mu$ is a trimming parameter.

As shown by Li and Vuong (1998), under certain smoothness conditions (6.15) and (6.16) provide uniformly consistent estimators of the density $f_{V_0}(\cdot)$ of $V_0$ and the densities of the homogenized bids $B_i^0$ for each bidder $i$. These densities can then be used to construct estimates of the right-hand-side of the first-order condition (rewriting (6.14))

$$A_{it} = B_{it}^0 + \frac{\prod_{j \neq i} G_{B^{0,c(j,t)}}(B_{it}^0)}{\sum_{j \neq i} g_{B^{0,c(j,t)}}(B_{it}^0) \prod_{k \neq i,j} G_{B^{0,c(k,t)}}(B_{it}^0)}, \tag{6.17}$$

where

$$G_{B^{0,j}}(b) = \int_{-\infty}^{b} g_{B^{0,j}}(s) \,\mathrm{d}s. \tag{6.18}$$

In contrast to other applications of the indirect approach to first-price auctions [e.g., Guerre, Perrigne and Vuong (2000)], however, here draws of the bids $B_{it}^0$ on the right-hand-side of (6.17) cannot be taken directly from the data. Instead, they must be simulated from the estimated densities $\hat{g}_{B_i^0}(b)$. Using simulated bids, (6.17) makes it possible to construct a pseudo-sample of draws of the idiosyncratic components $A_i$, which can be used to obtain estimates of their underlying densities $f_{A^j}(\cdot)$ using standard methods. Krasnokutskaya (2004) provides additional details and conditions under which this leads to uniformly consistent estimates of the marginal densities $f_{V_0}(\cdot)$ and $f_{A^j}(\cdot)$ for each bidder class $j$. She suggests the use of the bootstrap for inference.

Note that while the approach here is similar to that in Li, Perrigne and Vuong (2000), there are important distinctions. When $V_0$ is not observed by bidders, the joint distribution $F_{\mathbf{U}}(\cdot)$ is identified directly from the first-order condition and completely characterizes bidder demand and information. Since knowledge of $F_{\mathbf{U}}(\cdot)$ is sufficient for counterfactual simulations in a private values model with no unobserved heterogeneity, it is not clear under what circumstances one would need to separately identify $F_{\mathbf{A}}(\cdot)$ and $F_{V_0}(\cdot)$.[55] When $V_0$ is observed by the bidders, however, identification of the joint distribution $F_{\mathbf{U}}(\cdot)$ no longer follows directly from the first-order condition. Furthermore, even if $F_{\mathbf{U}}(\cdot)$ were identified, in this environment separate identification of $F_{\mathbf{A}}(\cdot)$ and $F_{V_0}(\cdot)$ is required for many counterfactuals.

The approach proposed by Krasnokutskaya (2004) is attractive in that it places no restriction on the distribution of the idiosyncratic factor $A_i$ or the distribution of $V_0$. It does restrict the way unobservables affect valuations. It may also require large samples – the slow convergence rates of deconvolution estimators is well known. Athey, Levin and Seira (2004) propose an alternative, trading flexibility in the specifications of $F_{V_0}(\cdot)$ and the $F_{A_i}(\cdot)$ for flexibility in how unobservable and observable auction characteristics affect valuations. They propose parametric estimation of the bid distributions and the distribution of auction heterogeneity. This is followed by estimation of the distribution of valuations based on (6.14) in a manner similar to Krasnokutskaya (2004). Mixtures of parametric models might be introduced to allow for more flexibility, as described at the end of Section 6.1.1. Although using a parametric first step is restrictive, it allows a parsimonious specification whereby the unobserved heterogeneity may affect some types of bidders differently than others, and where the distribution of the unobserved heterogeneity depends on auction characteristics. In principle, these features could be incorporated into Krasnokutskaya's (2004) approach by allowing auction characteristics to interact with $V_0$ and $A_i$ in (6.10), but in practice this may be not be feasible in data sets of moderate size.

*6.1.2.2. Ascending auctions* The challenges created by unobserved auction heterogeneity in an ascending auction are quite different. Because equilibrium is in weakly dominant strategies in the standard model of the ascending auction, unobserved heterogeneity does not affect the equilibrium mapping (the identity function) between

[55] A separate (and open) question, however, is whether imposing the structure of this model in estimation leads to more precise estimates in counterfactual simulations.

valuations and bids. For example, bidding in an environment with valuations characterized by (6.10) is the same regardless of whether bidders observe both $v_0$ and $a_i$ or only their sum. The main problem posed by such an environment is the fact that positive identification results for ascending auctions have been obtained primarily for environments with independent valuations, yet the presence of an unobserved factor like $v_0$ generally leads to a violation of independence.

In Section 6.2.1 we will show how additional data on bidder characteristics can be used to obtain identification of the joint distribution of valuations in an ascending auction without independence. This would not be sufficient for all economic questions of interest, however. As the preceding section makes clear, for example, separate identification of $F_{V_0}(\cdot)$ and each $F_{A_i}(\cdot)$ is needed even to simulate outcomes in a first-price sealed-bid auction. However, with an estimate of the joint distribution $F_{\mathbf{U}}(\cdot)$, it should be possible to use deconvolution techniques similar to those discussed above to separately estimate $F_{V_0}(\cdot)$ and each $F_{A_i}(\cdot)$ when $U_i = A_i + V_0$, under assumptions similar to those discussed above. This has not yet been investigated.

## 6.2. *Bidder heterogeneity*

### 6.2.1. *Observed bidder heterogeneity*

As discussed in prior sections, observable differences across bidders introduce asymmetry that can complicate the analysis of bidding data. However, when bidder-specific covariates are observable and vary across auctions, they can actually aid identification by enabling the distribution function for a single order statistic to reveal more information. This is particularly valuable in an ascending auction given the negative identification results above for environments without independence. In fact, with sufficiently rich variation in covariates, identification can be obtained with asymmetric dependent valuations, even when the transaction price is the only bid available (or the only bid assumed to have the unambiguous interpretation implied by the button auction model).

The idea behind this approach is familiar from other types of models, including the Roy model of labor supply [e.g., Heckman and Honoré (1990)] and competing risks models [e.g., Heckman and Honoré (1989)]. To see how this can work in the auction environment, suppose

$$U_i = g_i(W_i) + A_i,$$

where each $g_i(\cdot)$ is an unknown function, $W_i$ is a covariate reflecting characteristics of bidder $i$, and the private stochastic components $(A_1, \ldots, A_n)$ are drawn from an arbitrary joint distribution $F_{\mathbf{A}}(\cdot)$ and are independent of the matrix $\mathbf{W} = (\mathbf{W}_1, \ldots, \mathbf{W}_n)$. Suppose for the moment that each $g_i(\cdot)$ is known and that we could somehow observe

$u^{(n:n)}$.[56] Conditional on the vector $\mathbf{w}$, $U^{(n:n)}$ has cumulative distribution

$$\begin{aligned}F_U^{(n:n)}(u|\mathbf{w}) &= \Pr\bigl(U^{(n:n)} \leqslant u \mid \mathbf{w}\bigr)\\ &= F_{\mathbf{U}}(u, \ldots, u|\mathbf{w})\\ &= \Pr\bigl(g_i(w_i) + A_i \leqslant u\ \forall i\bigr)\\ &= F_{\mathbf{A}}\bigl(u - g_1(\mathbf{w}_1), \ldots, u - g_n(\mathbf{w}_n)\bigr).\end{aligned}$$

While the joint distribution $F_{\mathbf{U}}(\cdot|\mathbf{w})$ is observed only along the diagonal ($U_1 = \cdots = U_n$), sufficient variation in $(g_1(\mathbf{w}_1), \ldots, g_n(\mathbf{w}_n))$ would "trace out" the entire joint distribution $F_{\mathbf{A}}(\cdot)$. Furthermore, prior knowledge of the functions $g_i(\cdot)$ is not necessary with sufficient variation in covariates: at sufficiently large negative values of $g_j(w_j)\ \forall j \neq i$, bidder $i$ will have the largest valuation with probability arbitrarily close to one, so that variation in $\mathbf{w}_i$ and the point of evaluation $u$ would trace out the function $g_i(\cdot)$.

In practice we cannot observe $u^{(n:n)}$ in an ascending auction, and the distribution of an interior order statistic has a more complicated relation to the underlying joint distribution than does the maximum (or minimum, as in the case of competing risks). However, the following result shows that the fundamental idea behind this approach can be used to obtain identification in an ascending auction when only the transaction price is observable.[57]

THEOREM 6.1. *Assume*

(i) $U_i = g_i(\mathbf{W}_i) + A_i$, $i = 1, \ldots, n$.
(ii) $F_{\mathbf{A}}(\cdot)$ *has support* $\mathbb{R}^n$ *and a continuously differentiable density.*
(iii) $A_i$ *and* $\mathbf{W}_j$ *are independent for all* $i, j$.
(iv) $\operatorname{supp}(g_1(\mathbf{W}_1), \ldots, g_n(\mathbf{W}_n)) = \mathbb{R}^n$.
(v) *For all* $i$, $g_i(\cdot)$ *is continuously differentiable, with* $\lim_{\mathbf{w}_i \to (\infty, \ldots, \infty)} g_i(\mathbf{w}_i) = \infty$ *and* $\lim_{\mathbf{w}_i \to (-\infty, \ldots, -\infty)} g_i(\mathbf{w}_i) = -\infty$.

*Then* $F_{\mathbf{A}}(\cdot)$ *and each* $g_i(\cdot)$, $i = 1, \ldots, n$, *are identified up to a location normalization from observation of* $U^{(j:n)}$ *and* $\mathbf{W}$, *for any single value of* $j \in \{1, \ldots, n\}$.

PROOF. For simplicity let each $\mathbf{W}_i = W_i$ be a scalar. For $\mathcal{T} \subset \{1, \ldots, n\}$ define

$$\overline{F}_{\mathbf{A}}^{\mathcal{T}}(a_1, \ldots, a_n) \equiv \Pr(A_i > a_i\ \forall i \in \mathcal{T},\ A_j \leqslant a_j\ \forall j \notin \mathcal{T})$$

[56] This is the observable order statistic in the Roy model, where the wage in the chosen sector (the one offering the highest wage) is the only one observed, yet one is interested in the joint distribution of wage offers from all sectors.

[57] The result is a slight modification of Theorem 5 of Athey and Haile (2002), correcting a minor error in their proof.

and let $\overline{F}^{\mathcal{T}}_{\mathbf{A},A_i}(a_1,\dots,a_n) = \frac{\partial}{\partial a_i}\overline{F}^{\mathcal{T}}_{\mathbf{A}}(a_1,\dots,a_n)$. For arbitrary $u \in \mathbb{R}$, define $\mathbf{z} = (u - g_1(w_1),\dots,u - g_n(w_n))$. Then

$$\Pr\bigl(U^{(j:n)} \leqslant u \mid \mathbf{w}\bigr) = \sum_{\substack{\mathcal{T}\subset\{1,\dots,n\}\\|\mathcal{T}|=n-j}} \sum_{i\notin\mathcal{T}} \int_{-\infty}^{u} \overline{F}^{\mathcal{T}}_{\mathbf{A},A_i}\bigl(\tilde{u} - g_1(w_1),\dots,\tilde{u} - g_n(w_n)\bigr)\,\mathrm{d}\tilde{u},$$

where the summations are over the possible identities of the bidders with the $n-j$ highest bids, and the identity of the bidder $i$ with bid $B^{(j:n)}$. Differentiation yields

$$\begin{aligned}
&\frac{\partial}{\partial u}\frac{\partial^n}{\partial w_1\cdots\partial w_n}\Pr\bigl(U^{(j:n)} \leqslant u \mid \mathbf{w}\bigr)\\
&\quad= \sum_{\substack{\mathcal{T}\subset\{1,\dots,n\}\\|\mathcal{T}|=n-j}} \sum_{i\notin\mathcal{T}} (-1)^{n-j}\prod_{k=1}^{n}\bigl(-g_k'(w_k)\bigr)\frac{\partial}{\partial a_i}f_{\mathbf{A}}(\mathbf{a})\Big|_{\mathbf{a}=\mathbf{z}}\\
&\quad= \binom{n-1}{n-j}(-1)^{n-j}\prod_{k=1}^{n}\bigl(-g_k'(w_k)\bigr)\sum_{i=1}^{n}\frac{\partial}{\partial a_i}f_{\mathbf{A}}(\mathbf{a})\Big|_{\mathbf{a}=\mathbf{z}}
\end{aligned}$$

since there are $\binom{n-1}{n-j}$ subsets $\mathcal{T}$ of size $n-j$ that exclude $i$. Now observe that

$$\begin{aligned}
&\frac{\partial}{\partial u}\frac{\partial^n}{\partial w_1\cdots\partial w_n}F_{\mathbf{A}}\bigl(u - g_1(w_1),\dots,u - g_n(w_n)\bigr)\\
&\quad= \prod_{k=1}^{n}\bigl(-g_k'(w_k)\bigr)\sum_{i=1}^{n}\frac{\partial}{\partial a_i}f_{\mathbf{A}}(\mathbf{a})\Big|_{\mathbf{a}=\mathbf{z}}\\
&\quad= \frac{1}{\binom{n-1}{n-j}(-1)^{n-j}}\frac{\partial}{\partial u}\frac{\partial^n}{\partial w_1\cdots\partial w_n}\Pr\bigl(U^{(j:n)} \leqslant u \mid \mathbf{w}\bigr).
\end{aligned}$$

Hence, using the fundamental theorem of calculus,

$$\begin{aligned}
&\frac{\partial^n}{\partial w_1\cdots\partial w_n}F_{\mathbf{A}}\bigl(u - g_1(w_1),\dots,u - g_n(w_n)\bigr)\\
&\quad= \frac{1}{\binom{n-1}{n-j}(-1)^{n-j}}\frac{\partial^n}{\partial w_1\cdots\partial w_n}\Pr\bigl(U^{(j:n)} \leqslant u \mid \mathbf{w}\bigr).
\end{aligned}$$

Repeated application of the fundamental theorem of calculus shows that

$$\begin{aligned}
&\int_{w_1}^{\infty}\cdots\int_{w_n}^{\infty}\frac{\partial^n}{\partial\tilde{w}_1\cdots\partial\tilde{w}_n}F_{\mathbf{A}}\bigl(u - g_1(\tilde{w}_1),\dots,u - g_n(\tilde{w}_n)\bigr)\,\mathrm{d}\tilde{w}_n\dots\mathrm{d}\tilde{w}_1\\
&\quad= (-1)^n F_{\mathbf{A}}\bigl(u - g_1(w_1),\dots,u - g_n(w_n)\bigr)
\end{aligned}$$

so that

$$F_{\mathbf{A}}\big(u-g_1(w_1),\dots,u-g_n(w_n)\big) = \frac{(-1)^j}{\binom{n-1}{n-j}}\int_{w_1}^{\infty}\cdots\int_{w_n}^{\infty}\frac{\partial^n}{\partial\tilde{w}_1\cdots\partial\tilde{w}_n}\Pr\big(U^{(j:n)}\leqslant u\mid\tilde{\mathbf{w}}\big)\,\mathrm{d}\tilde{w}_n\ldots\mathrm{d}\tilde{w}_1. \tag{6.19}$$

Now note that $\lim_{\mathbf{w}_{-i}\to(-\infty,\dots,-\infty)} F_{\mathbf{A}}(u-g_1(w_1),\dots,u-g_n(w_n)) = F_{A_i}(u-g_i(w_i))$, where $F_{A_i}(\cdot)$ is the marginal distribution of $A_i$. For each $i$, then, variation in $u$ and $w_i$ identifies $g_i(\cdot)$ through Equation (6.19) by standard arguments. With knowledge of each $g_i(\cdot)$ we can then use (6.19) to uniquely determine $F_{\mathbf{A}}(\cdot)$ at any point $(a_1,\dots,a_n)$ through appropriate choices of $u$ and $\mathbf{w}$. □

Estimation based on this result has not yet been explored. For the competing risks model, however, Fermanian (2003) has recently proposed kernel methods that build directly on the closely related identification proof of Heckman and Honoré (1989).

#### 6.2.2. *Unobserved bidder heterogeneity*

We have already discussed several models with bidder heterogeneity that is either fixed across all auctions or captured by observable bidder-specific covariates. However, one can imagine situations in which asymmetries between bidders vary across auctions due to factors that are common knowledge to bidders but unobserved to the econometrician. For example, the match between the specifications of a procurement contract and each contractor's particular expertise might be common knowledge within the industry but unobservable to outsiders.

In the most general case, this type of environment requires a different marginal distribution $F_{U_{it}}(\cdot)$ for each bidder $i$'s valuation in each auction $t$. It should be clear that identification of such a model from bid data alone is impossible: the number of marginal distributions in the model is equal to the number of observations, even assuming one observes all bids from each auction.

Consider instead a more restrictive model

$$U_{it} = A_{it} + E_{it},$$

where all (i) $A_{it}$ are i.i.d. draws from a cumulative distribution $F_A(\cdot)$ with density $f_A(\cdot)$; (ii) $E_{it}$ and $A_{it}$ are mutually independent; (iii) $E_{it}$ is common knowledge among the bidders but unobserved to the econometrician; and (iv) each $E_{it}$ is an independent draw from a cumulative distribution $F_{E_i}(\cdot)$ with density $f_{E_i}(\cdot)$. From the econometrician's perspective, each bidder's valuation is then an independent draw from a density

$$f_{U_i}(\cdot) = f_A(\cdot) * f_{E_i}(\cdot),$$

where $*$ denotes convolution.

In an ascending auction, Theorem 4.1 implies that each $F_{U_i}(\cdot)$ is identified if one observes the transaction price, the set of bidders $\mathcal{N}$, and the winner's identity. This would be sufficient for some important questions and policy simulations, although not all. For example, it would not be sufficient to simulate outcomes under a first-price sealed-bid auction, since to do this one would need to know how much of the variation in valuations was common knowledge (through $E_i$) and how much was private information (through $A_i$). Separate identification of $F_A(\cdot)$ and $F_{E_i}(\cdot)$ for all $i$ is not possible from bid data, however. There are $n+1$ marginal distribution functions of interest. Yet even if one observed bids from all bidders (instead of the $n-1$ losing bids, as usually assumed), there are only $n$ marginal distributions of observable bids. Without additional restrictions, identification will not be possible.

In a first-price auction, the situation is further complicated by the nontrivial strategic behavior. In particular, even identification of each $F_{U_i}(\cdot)$ in the special case above is doubtful, since the markdown in each bidder's first-order condition

$$u_i = b_i + \frac{\Pr(\max_{j \neq i} B_j \leqslant b_i \mid B_i = b_i,\ E_1, \ldots, E_j)}{\frac{\partial}{\partial m} \Pr(\max_{j \neq i} B_j \leqslant b_i \mid B_i = b_i,\ E_1, \ldots, E_j)|_{m=b_i}}$$

involves expectations that are conditioned on the information $E_1, \ldots, E_j$ that is unobservable to the econometrician.[58] The problem here is closely related to that discussed in Section 6.1.2, although the dimensionality of the unobserved heterogeneity is higher, and the approaches thus far proposed to address unobserved heterogeneity do not appear to be applicable.

### 6.3. Endogenous participation

So far, we have focused on models in which any variation in the set of bidders is exogenous (the exception is the discussion of endogenous participation with unobserved heterogeneity in Section 6.1.2). In this section we consider several different models of how the set of bidders is determined, and we explore the consequences of these models for identification. Here it will be useful to draw a distinction between *potential* bidders and *actual* bidders. As before, we let $\mathcal{N}$ (with $|\mathcal{N}| = n$) denote the set of potential bidders – those who draw signals and decide whether to bid.[59] We let $\mathcal{A} \subseteq \mathcal{N}$ (with $|\mathcal{A}| = a$) denote the set of actual bidders, i.e., those who actually place a bid. Variation in both $\mathcal{N}$ and $\mathcal{A}$ is possible. Let $\widetilde{\mathcal{N}}$ be the random set whose realization is denoted by $\mathcal{N}$, and let $\widetilde{\mathcal{A}}$ be the random set whose realization is denoted by $\mathcal{A}$.

[58] Models similar to this have been explored in the related context of differentiated products oligopoly price competition [e.g., Berry, Levinsohn and Pakes (1995); see also Chapter 63 by Ackerberg et al. in this volume]. There, common knowledge differences in unobservable (to the econometrician) quality of products that differ across markets lead to asymmetries in the effective common knowledge marginal costs of supplying utility to a buyer choosing between firms. Identification in those models is obtained through a combination of parametric assumptions and restrictions from the demand side of the market. In the auction setting, the latter would be analogous to restrictions from the seller's (or auctioneer's) side of the market, for example using the assumption that the reserve price is set optimally. We are not aware of empirical approaches exploiting such information, although this is a direction worth exploring. See Einav (2004) for a related discussion.

[59] In the literature, sometimes agents with the option of acquiring a signal are referred to as potential bidders [e.g., Hendricks, Pinkse and Porter (2003)].

An example of why the set of potential bidders may vary is an environment in which obtaining a signal is costly. Firms may then decide whether to investigate a particular opportunity at random or based on some summary statistics about the auction (for example, the appraised value of the object). Fixing the set of potential bidders, the set of actual bidders may vary, for example, if there is a binding reserve price or if submitting a bid is costly. In such cases, typically only bidders with sufficiently favorable signals will bid. In addition, in an ascending auction that lacks a strict "activity rule" like that in the standard Milgrom–Weber model, the set of actual bidders can exclude even potential bidders with relatively high valuations, since others may push the price beyond these bidders' willingness to pay before they ever make a bid.[60]

In this section we will see that the consequences of endogenous variation in $\mathcal{A}$ and $\mathcal{N}$ for equilibrium and identification will depend on whether bidders' participation decisions are common knowledge among the bidders and whether these are observable by the econometrician. Often the number of actual bidders in an auction is observed by the econometrician; the set of potential bidders may or may not be observed. [61]

### *6.3.1. Binding reserve prices*

We first consider the case in which a reserve price may be binding. Recalling (2.1), in an $n$-bidder auction with reserve price $r$, only bidders with signals $x_i \geqslant x_i^*(r, \mathcal{N})$ participate (with $x_i^*(r, \mathcal{N}) = r$ in a private values auction). Ignoring this endogenous participation can result in misleading estimates due to the selection introduced by the participation decisions.[62] Throughout this section, we will assume $\mathcal{N}$ is observable, hold $\mathcal{N}$ fixed, and consider only bidders $i \in \mathcal{N}$.

*6.3.1.1. Ascending auctions* For ascending auctions we obtained positive identification results above primarily for models with independent private values (the exception is Theorem 6.1), so we will focus on such models here. Donald and Paarsch (1996)

[60] Auction-specific unobservables may affect either the number of potential bidders (e.g., if unobservables determine whether there is a suitable match between a specialized contractor and a contract offered by auction), or the number of actual bidders (e.g., if unobservables affect the profitability of an auction in an environment with costly signal acquisition). See Section 6.1.2 as well as Athey, Levin and Seira (2004), and Li and Zheng (2005).

[61] In the case that $\mathcal{N}$ is not observed but fixed in a sample, in most models of endogenous participation the common support assumption ensures that the union of identities of all actual bidders ever observed will converge to $\mathcal{N}$ as the sample of auctions grows [cf. Guerre, Perrigne and Vuong (2000)].

[62] A closely related model is that in which bidders must pay a fee to enter the auction [Samuelson (1985)] or, equivalently from the perspective of identification, preparing a bid is costly. This can lead to a participation rule very similar to that with a binding reserve price [Milgrom and Weber (1982)]. For first-price sealed-bid auctions, Haile, Hong and Shum (2003) discuss this case and provide results similar to those given in this section. Note that bid preparation costs are different from costs of acquiring a signal (discussed in Section 6.3.2), because in the former case a bidder places a bid if his signal is high enough, while in the latter case the participation decision must be made before bidders have obtained signals, and all bidders who acquire signals will bid (unless there is a binding reserve price).

and Paarsch (1997) were the first to incorporate reserve prices in structural models of ascending auctions in the IPV setting.[63] They observed that in a parametric framework one may account for the endogeneity of participation in one of two ways. First, if the number of potential bidders is observable, one may explicitly account (e.g., in a likelihood function) for the fact that the valuations (bids) of $(n - a)$ potential bidders were censored because these were below $r$. Alternatively, one can examine the bidding behavior of the actual bidders conditional on their decision to participate. This second approach is based on the fact that under independence each participating bidder $i$ has a valuation that is an independent draw from the distribution

$$F_{U_i}(u|r) = \frac{F_{U_i}(u) - F_{U_i}(r)}{1 - F_{U_i}(r)}. \tag{6.20}$$

This observation is useful for considering nonparametric identification as well. With this observation, Theorem 4.1 implies that each truncated distribution $F_{U_i}(\cdot|r)$ is nonparametrically identified.

COROLLARY 6.1. *In an ascending auction with symmetric independent private values, $F_U(\cdot|r)$ is identified when the transaction price and the number of actual bidders is observable. In the asymmetric independent private values model, for each $i \in \mathcal{N}$, $F_{U_i}(\cdot|r)$ is identified when the transaction price, the identity of the winning bidder, and the set $\mathcal{A}$ are observable.*

In many cases, this result alone will be sufficient to enable one to address interesting questions. In the symmetric case, for example, Haile and Tamer (2003) have shown that the truncated distribution $F_U(\cdot|r)$ can be sufficient to determine the optimal reserve price (recall Equation (4.12)). To state the result, let $F_{U|r}(\cdot)$ denote $F_U(\cdot|r)$, and let $c_0$ be the value the seller places on the good (or her marginal cost of providing it).

THEOREM 6.2. *Given any univariate CDF $\Phi(\cdot)$, let $\pi(r; \Phi) = (r - c_0)(1 - \Phi(r))$ and $p^*(\Phi) \in \arg\max_{p \in \operatorname{supp} \Phi(\cdot)} \pi(p; \Phi)$. Suppose $\pi(\cdot; F_U)$ is continuously differentiable and strictly quasi-concave. Then* (i) *if $r < p^*(F_U)$, $r^*(F_{U|r}) = p^*(F_U)$*; (ii) *if $r \geqslant p^*(F_U)$, $p^*(F_{U|r}) = r$.*

This result implies that in a symmetric IPV environment, the optimal reserve one would calculate by ignoring the endogenous participation is actually optimal, except when the actual reserve price results in truncation of the relevant region of support. This follows from the fact that the objective functions $\pi(\cdot; F_U)$ and $\pi(\cdot; F_{U|r})$ differ only by a multiplicative constant. The qualification concerning truncation is important but not surprising: if there are no data below the true optimal reserve price, this optimum cannot

[63] More recently, Donald, Paarsch and Robert (2006), and Bajari and Hortaçsu (2003a) have considered parametric models incorporating endogenous participation with reserve prices.

be detected. However, part (ii) of Theorem 6.2 ensures that when such truncation has occurred, the data will at least reveal this fact.

For some policy questions, including predicting revenues under a different mechanism or reserve price, the full (untruncated) distributions $F_{U_i}(\cdot)$ will be needed, even under the independent private values assumption. It should be clear that the value of $F_{U_i}(u)$ for $u$ lower than all observed reserve prices could not be determined except through a parametric assumption. However, if both $\mathcal{N}$ and $\mathcal{A}$ are observable, each $F_{U_i}(u)$ can be recovered for all $u \geqslant r$. In particular, since $F_{U_i}(r) = \Pr(i \notin \widetilde{\mathcal{A}})$, identification of $F_{U_i}(u)$ for all $u \geqslant r$ follows immediately from (6.20) and Corollary 6.1.

THEOREM 6.3. *In the symmetric independent private values model, $F_U(u)$ is identified for all $u \geqslant r$ when the transaction price and $|\widetilde{\mathcal{A}}|$ are observable. In the asymmetric independent private values model, each $F_{U_i}(\cdot)$ is identified when the transaction price, the identity of the winning bidder, and $\widetilde{\mathcal{A}}$ is observable.*

An estimate of $F_{U_i}(u)$ for $u \geqslant r$ will be sufficient for some policy questions, e.g., calculations of revenues with higher reserve prices or under some alternative mechanisms. Estimation of each $F_{U_i}(r) = \Pr(i \notin \widetilde{\mathcal{A}})$ based on a sample analog is straightforward. In the case of symmetry, a different approach to estimation of $F_U(\cdot)$ is available: observe that exchangeability implies [Haile, Hong and Shum (2003)]

$$\begin{aligned} F_U(r) &= \Pr(U_1 \leqslant r) \\ &= F_{\mathbf{U}}(r, \infty, \ldots, \infty; n) \\ &= \sum_{k=1}^{n} \frac{k}{n} \Pr\bigl(|\widetilde{\mathcal{A}}| = n - k\bigr). \end{aligned} \tag{6.21}$$

A sample analog of (6.21) places much weaker demands on the data than a sample analog of $\Pr(i \notin \widetilde{\mathcal{A}})$. Estimates of $F_U(\cdot|r)$ can be obtained from the winning bids as in Section 4.2, simply replacing $\mathcal{N}$ with $\mathcal{A}$. Combining such estimators to form

$$\widehat{F}_{U_i}(u) = \bigl[1 - \widehat{F}_{U_i}(r)\bigr]\widehat{F}_{U_i}(u|r) + \widehat{F}_{U_i}(r)$$

leads to a consistent estimator of $F_{U_i}(u)$.

Haile and Tamer (2003) point out that similar extensions apply to the bounds approach to ascending auctions discussed in Section 4.3. Their assumptions (see Section 4.3) imply that all bidders with valuations above the reserve price must participate, as in the standard model. Ignoring the endogenous participation and treating $\mathcal{A}$ as the set of potential bidders then leads to bounds on the CDF $F_U(u|r)$ for $u \geqslant r$. Combining these with an estimate of $F_U(r)$ obtained from the observable participation decision leads to bounds on $F_U(u)$ for $u \geqslant r$.

While we have treated the reserve price above as fixed, it should be clear that this is not necessary. As with other auction-specific covariates, the results above can be interpreted as holding for a given value of the reserve price. However, because economic theory places considerable structure on the effect of the reserve price on the distribution of participating bidders' valuations, in practice this structure should be utilized in estimation. For example, one would want to use data from all auctions with reserve prices below $s$ to estimate $F_{U_i}(u)$ for $u \geqslant s$. This requires a modified estimation approach that combines data drawn from different truncated distributions. Indeed, if the reserve price varies exogenously (e.g., as it would if it were set optimally by sellers with stochastic private values for the good that are independent of bidders' valuations), this variation can trace out much (or even all) of the distributions $F_{U_i}(\cdot)$. For example, if the support of the reserve price includes values below the lower boundary of the support of bidder valuations, then identification of the full distribution $F_U(\cdot)$ is immediate from the arguments above. The estimation problem in such cases is similar to that for other models with random truncation [e.g., Woodroofe (1985), Wang, Jewell and Tsai (1986)]. While this idea has been mentioned by Guerre, Perrigne and Vuong (2000), nonparametric estimators exploiting the presence of variation in reserve prices have not yet been investigated, either for ascending or first-price auctions.

*6.3.1.2. First-price auctions* Similar arguments apply to first-price auctions, although here we can consider a richer set of private values models. We will focus on the case in which the econometrician observes all of the bids as well as the realizations of the sets $\widetilde{\mathcal{A}}$ and $\widetilde{\mathcal{N}}$. In first-price auctions, it is necessary to make an assumption about whether the bidders observe the set $\widetilde{\mathcal{A}}$ before placing their bids. Since participation is determined by the realization of bidders' private information, it will often be most natural to assume that bidders do not know $\widetilde{\mathcal{A}}$ when choosing their bids. We will focus on this case.[64]

Since for any bidder $i$ making a bid in equilibrium

$$\begin{aligned} G_{M_i|B_i}(m_i|b_i;\mathcal{N}) &= \Pr\bigl(\widetilde{\mathcal{A}} = \{i\} \mid i \in \widetilde{\mathcal{A}},\ B_i = b_i, \mathcal{N}\bigr) \\ &\quad + \sum_{\mathcal{A}' \subset \mathcal{N}, i \in \mathcal{A}'} \Pr\Bigl(\widetilde{\mathcal{A}} = \mathcal{A}',\ \max_{k \in \mathcal{A}',\, k \neq i} B_k \leqslant m_i \Bigm| i \in \widetilde{\mathcal{A}}, B_i = b_i, \mathcal{N}\Bigr) \end{aligned}$$

the observables and the first-order condition (2.4) uniquely determine the valuation $u_{it}$ associated with the bid $b_{it}$ of each actual bidder. Letting $F_{\mathbf{U}}(\cdot|\mathcal{A}, r)$ denote the joint distribution of $\{U_i\colon\ U_i \geqslant r,\ i \in \mathcal{A}\}$, this gives the following result.

THEOREM 6.4. *For each $\mathcal{A} \subseteq \mathcal{N}$, the joint distribution $F_{\mathbf{U}}(\cdot|\mathcal{A}, r)$ is identified in a first-price auction from observation of the reserve price $r$, all bids, and the associated bidder identities. In a symmetric environment, it is sufficient to observe $r$ and all bids.*

Combined with the probabilities $\Pr(\widetilde{\mathcal{A}} = \mathcal{A} \mid \mathcal{N}, r)$ (for which identification is immediate when $\mathcal{A}$, $\mathcal{N}$, and $r$ are all observed), the joint distributions $F_{\mathbf{U}}(\cdot; \mathcal{A}, r)$ will be sufficient for a number of questions of interest, including predicting the effects of an

[64] In some auctions, bidders may be required to register or make a deposit in order to participate. If these actions are observable to bidders, $\mathcal{A}$ will be known at the time they choose their bids.

increase in the reserve price. As discussed above, however, in some cases one will need an estimate of the untruncated distribution of valuations. This does not appear to be possible in the case of correlated private values: there is simply no information available regarding the correlation of valuations below the reserve price. However, maintaining the assumption that $\mathcal{N}$ is observable, one can identify the marginal distributions of bidder valuations evaluated at values above $r$.[65]

THEOREM 6.5. *In a first-price auction with private values, $F_{U_i}(u_i)$ is identified for all $u_i \geqslant r$ from observation of all bids and the associated bidder identities. In a symmetric environment, it is sufficient to observe all bids.*

PROOF. For each $\mathcal{A}$ and each $i \in \mathcal{A}$, the joint distribution $F_{\mathbf{U}}(\cdot|\mathcal{A}, r)$ completely determines the conditional distribution $F_{U_i}(u_i|r) = \Pr(U_i \leqslant u_i \mid U_i \geqslant r)$. Further,

$$F_{U_i}(u_i|r) = \frac{F_{U_i}(u_i) - F_{U_i}(r)}{1 - F_{U_i}(r)} \tag{6.22}$$

for all $u_i \geqslant r$. $F_{U_i}(r)$ is identified from the observed participation decisions, as in the case of an ascending auction. The result then follows from (6.22). □

Note that in an independent private values auction, this provides identification of $F_{\mathbf{U}}(\mathbf{u})$ for $\mathbf{u}$ such that $u_i \geqslant r$ for all $i$. As with similar results in preceding sections, estimation is possible building directly on the identification result, substituting sample analogs for the probabilities $F_{U_i}(u_i|r)$ and $F_{U_i}(r)$ in (6.22).

### 6.3.2. *Costly signal acquisition and the identification of acquisition costs*

Levin and Smith (1994) have considered a model in which players ("firms") first choose whether to become potential bidders ("enter") by investing in signals of their valuations. Firms that invest observe private signals. The assumption of costly signals is natural in many environments, particularly in the procurement contexts that account for a large share of the data studied in the auctions literature. For example, acquiring a signal might require conducting/analyzing a seismic survey or reviewing detailed contract specifications. In this subsection, we discuss identification of both value distributions and the costs of signal acquisition.

Levin and Smith (1994) assume that the bidders observe the set of potential bidders before placing their bids; in Section 6.3.3 we discuss the alternative assumption that investments in signal acquisition are private information so that bidders place their bids without knowing which firms are potential bidders. Levin and Smith focus on symmetric

[65] Analogs of Theorems 6.4 and 6.5 were demonstrated for the case of symmetric independent private values by Guerre, Perrigne and Vuong (2000). Haile, Hong and Shum (2003) extended these results to symmetric affiliated private values and common values models.

equilibria of models with symmetric bidders. In equilibrium, firms acquiring a signal must expect to recover the cost of doing so on average. So when there are sufficiently many firms in the market, some must choose not to enter. In the unique symmetric equilibrium, entry is determined by mixed strategies, leading to exogenous variation in the set of potential bidders.[66] One caveat is that, as in virtually all entry games, when asymmetric equilibria are allowed, there will be multiple equilibria [see, e.g., Berry and Tamer (2005)].

To extend the econometric model to this setting, observe that the distribution of the set of potential bidders is determined by the mixing probabilities. Since firms make independent decisions about signal acquisition, the event $|\widetilde{\mathcal{N}}| = 1$ occurs with positive probability. This case was ruled out above because typically this is not an interesting case: if a bidder knows that $|\mathcal{N}| = 1$, she will simply bid the reserve price. However, for the purposes of this section and the next, we will allow $|\mathcal{N}| = 1$. If we assume that the reserve price $r$ is less than $\underline{u}_i$ for all firms $i$, the reserve price plays a role only when $|\mathcal{N}| = 1$, in which case the lone potential bidder bids the reserve. Hence when $r < \underline{u}_i$ for all $i$, the number of potential bidders is equal to the number of actual bidders, the model generates exogenous variation in the number of bidders, and the methods described above can be used to estimate primitive value distributions. When $r > \underline{u}_i$ there will also be also variation in the number of actual bidders for a given set of potential bidders, as in Section 6.3.1. There we assumed that the set of potential bidders was observable to the econometrician for some results. That may be unlikely in the presence of both a reserve price and costly signals, since the set of potential bidders varies across auctions. Li (2003) considers parametric estimation of a model based on Levin and Smith's model with $r > \underline{u}_i$.

In their study of US Forest Service timber auctions, Athey, Levin and Seira (2004) consider a variation of this model, allowing asymmetric bidders. They assume firms fall into two classes, "weak" and "strong" (generalizations to more than two types are also possible). Strong firms that choose to invest draw valuations from a distribution that stochastically dominates that of the "weak" firms. They restrict attention to type-symmetric equilibria, in which all members of a given class use the same strategies. Because firms are asymmetric, however, there may be multiple type-symmetric equilibria. Athey, Levin and Seira (2004) derive a restriction on primitives that guarantees a unique type-symmetric equilibrium, and this restriction can be verified empirically.

In any signal acquisition model that generates exogenous variation in $\widetilde{\mathcal{N}}$, if $\widetilde{\mathcal{N}}$ and all bids are observed (or in an IPV model if $\widetilde{\mathcal{N}}$ and the winning bid are observed), our prior results imply that (assuming $r < \underline{u}_i$) a bidder's *ex ante* gross expected profit $\Pi_i(\mathcal{N})$ from entering the auction is identified. In particular,

$$\Pi_i(\mathcal{N}) = E_{U_i}\big[\big(U_i - \beta_i(U_i;\mathcal{N})\big)G_{M_i|B_i}\big(\beta_i(U_i;\mathcal{N}) \mid \beta_i(U_i;\mathcal{N});\mathcal{N}\big)\big]$$

66 Hendricks, Pinkse and Porter (2003) have considered a variation on this model in a common values setting in which bidders choose whether to invest in a signal based on noisier (in a precise sense) preliminary estimates of their valuations. As they point out, their model can be interpreted as providing a purification of Levin and Smith's (1994) mixed strategy equilibrium.

with the right-hand side determined by the observed bid distribution and the first-order conditions for equilibrium bidding. Identification of $\Pi_i(\mathcal{N})$ requires no assumptions about the nature of the signal acquisition equilibrium (or equilibrium selection) beyond what is required to guarantee that variation in $\widetilde{\mathcal{N}}$ is exogenous. Estimates of $\Pi_i(\mathcal{N})$ can then be used to calculate all equilibria of an entry game for given entry costs. Thus, in an application, the existence of multiple equilibria in the entry game can be assessed empirically.

Athey, Levin and Seira (2004) show that in the unique type-symmetric equilibrium in their application, strong firms enter with probability one and weak firms are indifferent about entry. They further observe that for any firms that are indifferent about acquiring a signal (the weak firms in their application), the expected profit from entry must be zero. Thus entry costs are identified using $\Pi_i(\mathcal{N})$ and the distribution of $\widetilde{\mathcal{N}}$, which is directly observable. In particular, for any firm $i$ that is indifferent about acquiring a signal, signal acquisition costs must be equal to

$$\overline{\Pi}_i = \sum_{\mathcal{N}:\, i \in \mathcal{N}} \Pr(\widetilde{\mathcal{N}} = \mathcal{N} \mid i \in \widetilde{\mathcal{N}}) \Pi_i(\mathcal{N}).$$

Thus, in contrast to much of the empirical industrial organization literature on entry (where entry corresponds to signal acquisition in this model), which draws inferences solely from entry decisions,[67] the level of entry costs can be inferred. Hence it is possible to conduct counterfactual simulations about changes in these costs on the competitiveness of markets and bidder rents.

### 6.3.3. *Bidder uncertainty about the competition*

Throughout the preceding sections we maintained the assumption that bidders make their bids knowing the set of competitors they face. In the standard model of the ascending auction with private values, this is without loss of generality since the dominant strategy is not affected by the set of opponents. Furthermore, the assumption may be uncontroversial in an ascending auction; certainly if one believes bidders observe their opponents' exit prices (as in the standard model) it is natural to presume that bidders are aware of all competitors. In a sealed-bid auction, however, bidders need not gather to participate, making it less certain that bidders will know what competition they face. And in a first-price auction, a bidder's information about the competition is critical to the characterization of equilibrium bidding. In some procurement settings, firms may in fact know which of their competitors have the capability to compete for a given contract or even which firms have been invited to bid, but in other contexts this may not be public information.

Even if the set of firms who could in principle compete in an auction is common knowledge, in models where firms incur a cost to acquire a signal (see, e.g., Section 6.3.2) bidders may not know which other firms have actually invested in a signal

[67] See, e.g., Berry and Reiss (in press).

for a particular auction. There the investment choice is determined by randomization (in the case of a mixed strategy equilibrium) or as a function of private information (in a pure strategy equilibrium).

It is straightforward to modify theoretical models of costly signal acquisition to accommodate the case where bidders do not observe who has acquired a signal before bidding. McAfee, Quan and Vincent (2002) and Hendricks, Pinkse and Porter (2003) consider models with this feature for the case of first-price auctions. McAfee, Quan and Vincent (2002) show that a slightly stronger condition than affiliation of signals is required to ensure existence of a pure strategy Nash equilibrium in increasing strategies.[68] Li and Zheng (2005) also study such a model, highlighting an interesting testable theoretical possibility: bids may decrease when the number of firms increases, because each firm will enter with lower probability, and the resulting change in the distribution of potential bidders has ambiguous consequences for bidding strategies.

*6.3.3.1. Unknown potential competition* Now consider a first-price sealed-bid auction where $\widetilde{\mathcal{N}}$ is unobserved to both bidders and the econometrician. $\Pr(\widetilde{\mathcal{N}} = \mathcal{N})$ is identified as long as the set of bidders is observable at each auction. The distribution of the highest bid among $i$'s opponents is calculated taking the expectation over the set of potential bidders:

$$\begin{aligned} &G_{M_i|B_i}(m_i|b_i) \\ &\quad = \Pr\bigl(\widetilde{\mathcal{N}} = \{i\} \mid i \in \widetilde{\mathcal{N}}\bigr) \\ &\quad\quad + \sum_{\mathcal{N}:\, i\in\mathcal{N},\, |\mathcal{N}|>1} \Pr\Bigl(\max_{j\in\mathcal{N},\, j\neq i} B_j \leqslant m_i \Bigm| i \in \widetilde{\mathcal{N}}, B_i = b_i\Bigr) \Pr(\widetilde{\mathcal{N}} = \mathcal{N} \mid i \in \widetilde{\mathcal{N}}). \end{aligned} \tag{6.23}$$

Bidder $i$'s first-order condition is then given by

$$u_i = b_i + \frac{G_{M_i|B_i}(b_i|b_i)}{g_{M_i|B_i}(b_i|b_i)}. \tag{6.24}$$

This takes the usual form; however, here $G_{M_i|B_i}(\cdot)$ does not depend on $\mathcal{N}$. Equation (6.24) and observation of all bids then identifies the distribution of $U_i$, and straightforward extensions of the estimation techniques described above can be applied.

So far we have considered two assumptions that might be made when interpreting data from first-price auctions: (i) $\widetilde{\mathcal{N}}$ is observed by all bidders prior to bidding, or (ii) $\widetilde{\mathcal{N}}$

[68] In particular, they assume that there exists a nondecreasing function $h(\cdot)$ such that for each $i$, $U_i = h(X_i, V_0)$, where $(X_1, \ldots, X_n)$ are i.i.d. conditional on $V_0$. Private values, where $U_i = X_i$, is a special case. Each bidder bears a cost (constant across bidders) to learn the value of $X_i$. Bidders first invest in their signals and then place bids, but investment decisions are not publicly observable. They derive an equilibrium in which firms randomize in the signal acquisition decision. Then, for bidders who acquire a signal, bidding is in strictly monotone pure strategies. They show that a sufficient (but not necessary) condition for existence of a such an equilibrium is that $1 - \rho(1 - F_{X_i|V_0}(x_i|v_0))$ is log-supermodular in $(x_i, v_0)$, where $\rho$ is the entry probability in the mixed strategy equilibrium of the entry game.

is unobserved prior to bidding. In many settings, institutional detail may be available to guide the choice between these assumptions. When there is variation in $\mathcal{N}$, the data can also help guide this choice. If $\widetilde{\mathcal{N}}$ is observed by all bidders prior to bidding, then when $|\mathcal{N}| = 1$ the bidder must bid the reserve price. Thus, the hypothesis that $\widetilde{\mathcal{N}}$ is observable to bidders could be rejected if $\Pr(B^{(1:1)} = r \mid |\widetilde{\mathcal{N}}| = 1) < 1$. In addition, building on the discussion in Section 5, we note that both assumptions can have additional testable implications. If variation in $\widetilde{\mathcal{N}}$ is exogenous (as in the models of costly signal acquisition described above), it will be possible to estimate $F_{\mathbf{U}}(\cdot)$ separately for each $\mathcal{N}$ and compare the resulting estimates. If $\widetilde{\mathcal{N}}$ is observed by bidders before bidding, these estimates should be equal to each other (up to sampling error). On the other hand, if bidders have no information regarding the realization of $\widetilde{\mathcal{N}}$ when choosing their bids, then the distribution of $B_i$ itself should not vary with $\mathcal{N}$ (recall (6.24)).

*6.3.3.2. Noisy knowledge of the competition* Once we allow the possibility that bidders do not observe $\widetilde{\mathcal{N}}$ prior to bidding, it is natural to consider more carefully what bidders do know. In particular, it may be more reasonable to imagine that bidders have noisy signals of $\widetilde{\mathcal{N}}$ when choosing their bids. When the econometrician can condition on the same information available to bidders (excluding their signals of course), extending the methods is straightforward. Suppose, for example, that bidders form their beliefs about the set of competitors based on a public signal $\eta$ that is also observable to the econometrician. The signal, $\eta$, might contain information about how costly it will be to evaluate the object and acquire a signal, or information about the expected value of the object. In a model of costly signal acquisition, such factors will affect the entry probability of each bidder.

We can extend the methods above by treating $\eta$ as an auction-specific covariate to be conditioned on in bidders' first-order conditions. Note that the signal $\eta$ need not be a scalar and can include any information that may affect the set of potential bidders, including, e.g., characteristics of the good for sale or market conditions. Hendricks, Pinkse and Porter (2003) consider a simple example of this approach. They construct a binary signal $\eta_t = 1\{\Upsilon_t \geqslant \Upsilon^*\}$ of the number of potential bidders for tract $t$, where $\Upsilon_t$ is the number of firms ever to bid on an oil tract in a geographic neighborhood of the tract offered in auction $t$, and $\Upsilon^*$ is a specified threshold value.

In contrast, if bidders have signals (public or private) about factors that affect the number of competitors, but these signals are not observable to the econometrician, the problem of unobserved heterogeneity discussed in Section 6.1.2 arises. For example, in a model where acquiring a signal is costly, firms might observe an auction characteristic $v_0$ *before* making entry decisions. Another possibility is that firms observe auction characteristics that affect the cost of acquiring information about a particular object. Li and Zheng (2005) develop a model of a first-price auction with these features. They specify a semi-parametric model, leaving the distribution of unobserved heterogeneity unrestricted while assuming a functional form for the marginal distributions of valuations conditional on the heterogeneity. They estimate the model using Bayesian methods.

### 6.3.4. Internet auctions and unobserved participation

Internet auctions have recently attracted considerable attention from economists. In addition to providing a great deal of new data, Internet auctions introduce a number of new and interesting questions, including the role of seller reputations [see, e.g., the papers surveyed by Bajari and Hortaçsu (2003b, 2004)] and competition between sellers [e.g., Peters and Severinov (2006)].

Internet auctions are most often conducted in one of several variations on the standard ascending auction mechanism [Lucking-Reiley (2000)]. However, a challenge to structural analysis of bid data from Internet auctions is the fact that the number of bidders cannot be observed. Recall from Section 4 that a key assumption for the identification arguments in even the simplest ascending auction environments was observation of the number of bidders – either the number of potential bidders or the number who have valuations above the reserve price. An Internet auction typically takes place over several days (usually a week or more on eBay, for example), with bidders becoming aware of the auction at different times as they log onto the auction site while the auction is underway. A bidder who logs on to discover that the price has already risen past his valuation will not bid. Hence the number of submitted bids will not generally equal the number of bidders willing to pay the reserve price (if any).[69] The usual assumption that the transaction price is equal to the second-highest valuation is of little use if it is not known whether it is the second highest of two valuations or of ten, for example.

This problem has accounted for a substantial impediment to progress in addressing questions about the underlying demand structures at Internet auctions.[70] This includes even seemingly simple questions like how seller reputations affect bidders' willingness to pay, since this requires inference on the underlying distribution of bidder valuations.

Song (2003) has proposed a model capturing key departures of Internet auctions from the standard ascending auction model. Using this model, she derives conditions under which the identification of the distribution $F_U(\cdot)$ can be obtained in the symmetric independent private values paradigm without observing the number of bidders, or even assuming that this number is constant.

In her model, an auction takes place over an interval of time $[0, \tau]$. The distribution of $N$ can vary across auctions, and need not be known to bidders. In a given auction, each potential bidder $i$ draws a vector of "bidding opportunities" $(t_i^1, \dots, t_i^{\tau_i})$, with each $t_i^k \in [0, \tau]$. Taking $t_i^1 < \cdots < t_i^{\tau_i}$ without loss of generality, $t_i^1$ represents the time of $i$'s "arrival" at the auction, and $t_i^{\tau_i}$ represents $i$'s final bidding opportunity. No restriction is

[69] This problem can also arise in other applications, particularly in other ascending auctions with similar deviations from the button auction model, or in Dutch auctions, where only the winner makes a bid. Song (2004) explores identification and estimation in these and other auction models when the number of bidders is not observable to the econometrician.

[70] Bajari and Hortaçsu (2003a) avoid this problem with a common values model that admits an equilibrium in which all bidders willing to pay the reserve price will bid simultaneously at the end of the auction, as if in a second-price sealed bid auction. See Ockenfels and Roth (2006) for an alternative model of Internet auctions.

placed on the joint distribution of $(N, \{\tau_i\}, \{t_i^k\})$ except that (a) these are independent of bidders' valuations, and (b) each $t_i^{\tau_i}$ is continuously distributed on some interval $(t_i^0, \tau]$. In this model, bidders may "arrive" early or late, bid frequently or infrequently, and have different notions of what bidding at the "last minute" means.

At each bidding opportunity, a bidder may specify a "cutoff price" of any value above the current standing bid. Whenever a new cutoff price is submitted, the auctioneer raises the standing bid (denoted $s_t$) to the second-highest cutoff price, and the bidder with the highest cutoff price is named the standing high bidder. This matches the actual procedure on eBay, the most popular Internet auction site, for example. At time $\tau$, the standing high bidder wins the object at the standing bid (for simplicity we assume no reserve price and no minimum bid increment). Typically, the econometrician can observe the history of submitted cutoff prices (except the winner's), as well as the identity of the bidder who placed each bid. This information is publicly available for eBay auctions, for example.

There are many equilibria of this game. For example, all bidders can submit cutoff prices equal to their valuations at their first bidding opportunities; bidders may start with low cutoff prices and gradually raise them as the auction proceeds; or some/all bidders may wait until their final bidding opportunities to submit a cutoff price. In some of these equilibria (like the last example), some potential bidders will not bid, since at their planned bidding time the standing bid will already exceed their valuations. However, Song (2003) shows that in all equilibria the highest cutoff price submitted by bidder $i$ will be no larger than his valuation $u_i$, and it will equal his valuation if the standing bid at time $t_i^{\tau_i}$ was below $u_i$.[71]

Since the price can never rise above $u^{(n-1:n)}$, this ensures that the allocation is efficient and that the transaction price is $u^{(n-1:n)}$. Further, in some cases, the third-highest cutoff price submitted will be equal to $u^{(n-2:n)}$. To see this, let $b_i$ denote the highest cutoff price submitted by bidder $i$ (i.e., his "bid") and let $b^{(m-2:m)}$ denote the third-highest such bid (or $-\infty$ if there is no such bid). Here $m$ represents the number of *observed* bidders – those submitting cutoff prices at some point in the auction. Now suppose that at time $\tilde{t}$ the standing bid $s_{\tilde{t}}$ is no higher than $b^{(m-2:m)}$. In practice, whether this is true can be directly determined from the available bidding histories. In particular, recalling that two bids above $b$ are required for the standing bid to exceed $b$, this occurs whenever at least one of the two bidders making the highest bids (as of the end of the auction) makes no bid above $b^{(m-2:m)}$ prior to time $\tilde{t}$. In that case we have

$$s_{\tilde{t}} \leqslant b^{(m-2:m)} \leqslant u^{(m-2:m)} \leqslant u^{(n-2:n)}$$

implying that if the bidder with valuation $u^{(n-2:n)}$ has his final bidding opportunity at time $\tilde{t}$ or earlier, he will submit a cutoff price equal to his valuation. In that case,

[71] The first property is easily understood. The second follows from arguments similar to those used in analyzing a second-price auction.

$b^{(m-2:m)} = u^{(n-2:n)}$. While the final bidding time of this bidder is not known, by looking at auctions in which $s_{\tilde{t}} \leqslant b^{(m-2:m)}$ for $\tilde{t}$ sufficiently close to $\tau$, the probability that $b^{(m-2:m)} = u^{(n-2:n)}$ can be made arbitrarily close to one.[72]

This is useful because we typically cannot observe the cutoff price submitted by the auction winner.[73] However, by examining only the set of auctions in which

$$s_{\tilde{t}} \leqslant b^{(m-2:m)}, \quad \tilde{t} \in (\tau - \delta, \tau), \tag{6.25}$$

for small $\delta > 0$, we can treat the order statistics $(U^{(n-1:n)}, U^{(n-2:n)})$ as "observed." The following result, proved in Song (2003), implies that observation of $(U^{(n-1:n)}, U^{(n-2:n)})$ is sufficient to identify $F_U(\cdot)$, even though the realization of $N$ at each auction is unknown.

LEMMA 6.2. *Let $(Y^{(N:N)}, Y^{(N-1:N)}, Y^{(N-2:N)})$ denote random variables equal to the three highest of $N \geqslant 3$ independent draws from a univariate distribution $F_Y(\cdot)$, where $N$ is stochastic and unobserved. $F_Y(\cdot)$ is uniquely determined by the joint distribution of $(Y^{(N-1:N)}, Y^{(N-2:N)})$.*

PROOF. Given $Y^{(N-2:N)} = y'$, the pair $(Y^{(N-1:N)}, Y^{(N:N)})$ can be reinterpreted as the two order statistics for an i.i.d. sample of size two from the distribution

$$F_Y(\cdot|y') = \frac{F_Y(\cdot) - F_Y(y')}{1 - F_Y(y')}.$$

Although $Y^{(N:N)}$ is unobserved, Equation (4.1) implies that the observation of $Y^{(N-1:N)}$ alone is sufficient to identify the parent distribution $F_Y(\cdot|y')$ for this sample. Identification of $F_Y(\cdot)$ then follows from the fact that

$$\lim_{y' \downarrow \inf \operatorname{supp} F_Y^{(N-2:N)}(\cdot)} F_Y(\cdot|y') = F_Y(\cdot).$$

$\square$

Key to the applicability of this result is an assumption that auctions in which at least one of the two high bidders make late bids (i.e., where (6.25) holds) are representative of all auctions. If $\tau_i = 1\ \forall i$, this follows from the assumption that $\{\tau_i\}$ and $\{U_i\}$ are independent. In general, when $\tau_i > 1$, this requires the additional assumption that the equilibrium selection does not depend on $(u_1, \ldots, u_n)$.[74]

[72] In finite sample, of course, there will be a tradeoff between the bias of including auctions with $\tilde{t}$ far from $\tau$ and the reduction in the variance from doing so. Song (2003) suggests a data-driven approach for choosing the sample.

[73] If we could – for example, if such data were provided directly by eBay – a variation on the result below would still be applicable to address the problem that the number of potential bidders is unobserved.

[74] More precisely, the distribution of $U^{(n-1:n)}|U^{(n-2:n)}$ conditional on at least three bidders' being observed and (6.25) holding must be the same as the unconditional distribution of $U^{(n-1:n)}|U^{(n-2:n)}$.

Song (2003) proposes a semi-nonparametric estimator [Gallant and Nychka (1987)] applicable to the subset of auctions in which bids are observed from at least three distinct bidders. The likelihood function is constructed from the conditional density of $U^{(n-1:n)}$ given $U^{(n-2:n)}$, i.e.,

$$\frac{\partial}{\partial y} \Pr\bigl(U^{(n-1:n)} \leqslant y \bigm| U^{(n-2:n)} = x\bigr) = \frac{2(1 - F_U(y)) f_U(y)}{(1 - F_U(x))^2}$$

in which $n$ does not appear. Monte Carlo experiments suggest that the approach can perform well in sample sizes easily attainable from Internet auctions.

### 6.4. *Risk aversion*

Most of the empirical literature on auctions assumes risk neutrality of bidders. Risk neutrality is a natural assumption when the value of the object being sold is small relative to each bidder's wealth. Furthermore, in many applications bidders are firms, which economists usually assume to be profit maximizers. However, many auctions involve highly valuable goods (or contracts). And even when bidders represent firms, they may themselves be risk averse.[75] Risk aversion can have important implications for a wide range of policy questions, including the optimal reserve price and a seller's preference between the standard auction formats.[76]

Risk aversion also creates significant challenges for identification. In an ascending auction with private values, for example, risk aversion has no effect on equilibrium bidding in the standard model: bidding one's valuation is still a dominant strategy. While this implies that identification of $F_{\mathbf{U}}(\cdot)$ holds with risk aversion whenever it holds with risk neutrality, it also implies that there is no way to distinguish risk neutrality from risk aversion, i.e., no way to identify bidders' preferences. While the distribution $F_{\mathbf{U}}(\cdot)$ will be sufficient for some questions (for example, the effect of changing the reserve price) it will be inadequate for many others.

In a first-price auction, the implications of risk aversion for equilibrium bidding are nontrivial, since bidding involves a gamble. Bidding less aggressively leads to a lower chance of winning but higher profits conditional on winning. A more risk averse bidder will be less willing to accept a reduced probability of winning in order to obtain a higher profit when she wins. This suggests that there is at least hope for identification of preferences using data from first-price auctions. However, identifying risk preferences generally requires observation of choices from different menus of lotteries. Variations

[75] The incentives provided by the firms they work for may or may not "undo" such risk aversion. Athey and Levin (2001), Campo et al. (2002), and Perrigne (2003), for example, find evidence consistent with risk averse bidding behavior by firms at timber auctions.

[76] See, e.g., McAfee and McMillan (1987) and references therein. The theory of first-price auctions with risk averse bidders was initially developed by Maskin and Riley (1984). Campo et al. (2002) extend the analysis to the case in which there is no binding reserve price and establish additional smoothness properties used for identification and estimation.

in bidders' valuations do change the sets of lotteries available to them, but not in ways that are observable to the econometrician, since valuations are private information. This suggests that some observable exogenous variation will be needed to separately identify preferences and the distribution of valuations.[77] Below we will first explore possible approaches to identification in symmetric models, proceeding to consider models with asymmetric preferences in Section 6.4.2.

### 6.4.1. Symmetric preferences

We begin with a more formal illustration of the fundamental challenge for identification in models of first-price auctions with risk aversion. For simplicity, consider the case of symmetric independent private values, and assume that all bids are observable. Assume further that all bidders share the same continuously differentiable utility function $\omega(\cdot)$. Taking equilibrium behavior of her opponents as given, bidder $i$ solves the problem

$$\max_{\tilde{b}_i} \omega(u_i - \tilde{b}_i) \Pr\Big( \max_{j \in \mathcal{N}_{-i}} B_j \leqslant \tilde{b}_i \Big).$$

If we define

$$\lambda(s) \equiv \omega(s)/\omega'(s), \tag{6.26}$$

then first-order condition

$$\omega'(u_i - b_i) G_B(b_i) = (n-1)\omega(u_i - b_i) g_B(b_i)$$

can be rewritten usefully as

$$u_i = b_i + \lambda^{-1}\left( \frac{1}{n-1} \frac{G_B(b_i)}{g_B(b_i)} \right). \tag{6.27}$$

Now define the function

$$\xi(b_i, n, \lambda) = b_i + \lambda^{-1}\left( \frac{1}{n-1} \frac{G_B(b_i)}{g_B(b_i)} \right)$$

and let $\lambda_{\mathrm{I}}(\cdot)$ denote the identity function – i.e., the function $\lambda(\cdot)$ implied in the case of risk neutrality. Campo et al. (2002) show that as long as bids are independent (and additional regularity conditions are satisfied), an observed marginal bid distribution $G_B(\cdot)$ can be rationalized by equilibrium behavior if and only if there exists a utility function $\omega(\cdot)$ such that, for the associated $\lambda(\cdot)$, $\xi(\cdot; n, \lambda)$ is increasing (see Section 5.1). Hence, if bids are independent and $\xi(\cdot, n, \lambda_{\mathrm{I}})$ is increasing, it will be possible to find a distribution $F_U(\cdot)$ that rationalizes the observed bids within the symmetric risk neutral IPV

[77] However, if the distribution $F_{\mathbf{U}}(\cdot)$ is known or identified from other data – for example, in a laboratory setting or when one observes the same bidders participating in both first-price and ascending auctions – bid data from first-price auctions might then be used to estimate the utility function.

model. If $\xi(\cdot, n, \lambda_{\mathrm{I}})$ is decreasing at some point, the observed bids could not have been generated by equilibrium bidding by risk-neutral bidders, although there may exist another utility function $\omega(\cdot)$ with associated $\lambda(\cdot)$ such that $\xi(\cdot; n, \lambda)$ is increasing. Thus, allowing for risk aversion expands the set of observable bid distributions that can be rationalized by equilibrium bidding [Campo et al. (2002)].

Unfortunately, $\xi(\cdot, n, \lambda)$ need not violate the monotonicity restriction when the model is misspecified – in particular when the given function $\lambda(\cdot)$ does not correspond to that for the true preferences. When $\xi(\cdot, n, \lambda_{\mathrm{I}})$ is increasing, for example, the observed bids can be rationalized with risk neutrality, but they can also be rationalized with many different specifications of risk aversion. To suggest why, observe that if as long as a bidder is sufficiently risk averse, $\lambda^{-1}(\frac{1}{n-1}\frac{G_B(b)}{g_B(b)})$ does not vary much with $b$, ensuring that $\xi(b; n, \lambda)$ strictly increases in $b$. Consider the following example of a CRRA utility function (with zero initial wealth): $\omega(u) = u^{1-c}$, with $0 \leqslant c < 1$. Then $\lambda(s) = s/(1-c)$, and $\lambda^{-1}(z) = z(1-c)$. As $c$ approaches 1, $\xi(\cdot; n, \lambda)$ approaches the identity function. Intuitively, sufficiently risk averse bidders are not willing to risk losing the object by shading their bids, so they do not respond to the shape of the opposing bid distribution. Thus, even if $\frac{G_B(b)}{g_B(b)}$ is sharply decreasing in some places, there will be a critical level of risk aversion above which $\xi(\cdot; n, \lambda)$ is everywhere increasing.

Similarly, it is generally impossible to identify the degree of risk aversion from bid data in a fixed environment. Perhaps surprisingly,[78] however, this is true even with a strong functional form assumption on bidders' preferences. Again consider the CRRA example, and suppose that the data can be rationalized by a distribution $F_U(\cdot)$ and coefficient of relative risk aversion $c$. Then, for any $\tilde{c} \in (c, 1)$, if we let $\tilde{\omega}(s) = s^{1-\tilde{c}}$, we can find another distribution $F_{\widetilde{U}}(\cdot)$ that implies the same distribution of bids, but where $F_{\widetilde{U}}(\cdot)$ stochastically dominates $F_U(\cdot)$. In particular, to satisfy (6.27), we define $\widetilde{U}_i$ to be equal in distribution to

$$\begin{aligned}
\xi(B_i; n, \tilde{\lambda}) &= B_i + \frac{1-\tilde{c}}{n-1}\frac{G_B(B_i)}{g_B(B_i)} \\
&= \frac{\tilde{c}-c}{1-c}B_i + \frac{1-\tilde{c}}{1-c}\left(B_i + \frac{1-c}{n-1}\frac{G_B(B_i)}{g_B(B_i)}\right) \\
&= \frac{\tilde{c}-c}{1-c}B_i + \frac{1-\tilde{c}}{1-c}\xi(B_i; n, \lambda) \\
&= \frac{\tilde{c}-c}{1-c}B_i + \frac{1-\tilde{c}}{1-c}U_i.
\end{aligned}$$

It follows that whenever $\xi(b_i; n, \lambda)$ is increasing in $b_i$, so is $\xi(b_i; n, \tilde{\lambda})$. Hence, the data can be rationalized with risk aversion $\tilde{c}$. Campo et al. (2002) show that this argument holds for other parameterized families, as well as general utility functions.

[78] Recall that with risk neutrality the symmetric IPV model is overidentified when one observes all bids from each auction.

These results are quite negative. Only for bid distributions such that $\frac{G_B(b)}{g_B(b)}$ decreases sufficiently sharply in $b$ in places can risk aversion be distinguished from risk neutrality; it is impossible to distinguish among different parameterized functional forms for risk aversion; and there exists a large range of risk aversion parameters that can rationalize the observed bid data, even when attention is restricted to a particular functional form.

Following the intuition at the beginning of this section, however, this nonidentification might be overcome with observable exogenous variation in the sets of gambles available to bidders. One possibility is a covariate that shifts bidders' initial wealth or, equivalently, bidders' valuations for the good. Suppose, for example, that each bidder $i$'s utility from winning the auction is

$$\omega\big(h(w_i) + u_i - b_i\big) \tag{6.28}$$

for some increasing function $h(\cdot)$, where the covariate $w_i$ is independent of $u_i$ and is observable to all bidders prior to the auction as well as to the econometrician. Let $\mathbf{w} = (w_1, \ldots, w_n)$. The model then becomes asymmetric, even though bidders' preferences are given by the same function. Let $G_{M_i}(b_i|\mathbf{w}, \mathcal{N})$ be the distribution of the maximum bid of bidder $i$'s opponents, conditional on $\mathbf{w}$ and $\mathcal{N}$. Let $b_{\alpha,\mathbf{w},\mathcal{N}}$ denote the $\alpha$th quantile of the distribution $G_{M_i}(b_i|\mathbf{w}, \mathcal{N})$, while $u_\alpha$ is the $\alpha$th quantile of $F_U(\cdot)$. Then equilibrium requires that

$$u_\alpha = b_{\alpha,\mathbf{w},\mathcal{N}} - h(w_i) + \lambda^{-1}\left(\frac{G_{M_i}(b_{\alpha,\mathbf{w},\mathcal{N}}|x, \mathcal{N})}{g_{M_i}(b_{\alpha,\mathbf{w},\mathcal{N}}|x, \mathcal{N})}\right)$$
$$\forall \mathbf{w} \in \operatorname{supp} \mathbf{W},\ \forall \alpha \in [0, 1]. \tag{6.29}$$

The data can be rationalized by the model only if we can find a $\lambda(\cdot)$ such that (6.29) holds and such that

$$\xi_i(b_i; \mathcal{N}, \lambda, \mathbf{w}) \equiv b_i - h(w_i) + \lambda^{-1}\left(\frac{G_{M_i}(b_i|\mathbf{w}, \mathcal{N})}{g_{M_i}(b_i|\mathbf{w}, \mathcal{N})}\right)$$

is increasing in $b_i$.

This may not be possible, especially within a restricted class of utility functions. To see this, again consider the one-parameter CRRA example and suppose that the vector $\mathbf{W}$ takes on only two values, $\mathbf{w}'$ and $\mathbf{w}''$. Then equilibrium requires that for all $\alpha \in [0, 1]$,

$$\begin{aligned} u_\alpha &= b_{\alpha,\mathbf{w}'',\mathcal{N}} - h\big(w_i''\big) + (1-c)\frac{G_{M_i}(b_{\alpha,\mathbf{w}'',\mathcal{N}}|\mathbf{w}'', \mathcal{N})}{g_{M_i}(b_{\alpha,\mathbf{w}'',\mathcal{N}}|\mathbf{w}'', \mathcal{N})} \\ &= b_{\alpha,\mathbf{w}',\mathcal{N}} - h\big(w_i'\big) + (1-c)\frac{G_{M_i}(b_{\alpha,\mathbf{w}',\mathcal{N}}|\mathbf{w}', \mathcal{N})}{g_{M_i}(b_{\alpha,\mathbf{w}',\mathcal{N}}|\mathbf{w}', \mathcal{N})}. \end{aligned}$$

Thus

$$c = 1 - \frac{b_{\alpha,\mathbf{w}'',\mathcal{N}} - b_{\alpha,\mathbf{w}',\mathcal{N}} - (h(w_i'') - h(w_i'))}{\frac{G_{M_i}(b_{\alpha,\mathbf{w}',\mathcal{N}}|\mathbf{w}',\mathcal{N})}{g_{M_i}(b_{\alpha,\mathbf{w}',\mathcal{N}}|\mathbf{w}',\mathcal{N})} - \frac{G_{M_i}(b_{\alpha,\mathbf{w}'',\mathcal{N}}|\mathbf{w}'',\mathcal{N})}{g_{M_i}(b_{\alpha,\mathbf{w}'',\mathcal{N}}|\mathbf{w}'',\mathcal{N})}}. \tag{6.30}$$

For a given quantile $\alpha$, rationalizing the data with the CRRA model requires that there exist a function $h(\cdot)$ such that this $c$ lies in the interval $[0, 1)$. If no such $h(\cdot)$ exists, then we can immediately reject the CRRA model. Of course, (6.30) must hold for all quantiles $\alpha$. Unless the ratio on the right side of (6.30) is invariant to the quantile $\alpha$, the model will be rejected. Thus, a more flexible specification of risk preferences will typically be required to rationalize the observed bidding data when there are bidder-specific covariates shifting wealth or valuations.

Of course, there is more than one way to relax the structure imposed by (6.28) and CRRA. Campo et al. (2002) maintain the CRRA specification above but assume a functional form for the effect of covariates on valuations only at a single quantile of the distribution of valuations. By leaving the effects at other quantiles unspecified, the problem that the data may reject the model is avoided. We refer readers to their paper for details, as well as an estimation approach.[79]

Another possible approach to identification is to exploit exogenous variation in the number of bidders [e.g., Bajari and Hortaçsu (2005)]. Such exogenous variation changes the equilibrium probability that each given bid wins and, therefore, changes the lotteries available to bidders. Note that unlike the effect of a covariate on the utility gain from winning, this variation in the probability of winning can be determined directly from the equilibrium bid distribution for each $\mathcal{N}$. Using the CRRA model as an example, suppose that there are two groups of bidders, $\mathcal{N}$ and $\mathcal{N}'$, with $|\mathcal{N}| = n$ and $|\mathcal{N}'| = n+1$. Letting $b_{\alpha,\mathcal{N}}$ be the $\alpha$th quantile of $G_{B_i}(b_i|\mathcal{N})$, equilibrium requires

$$u_\alpha = b_{\alpha,\mathcal{N}} + \frac{1-c}{n-1}\frac{G_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})}{g_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})} = b_{\alpha,\mathcal{N}'} + \frac{1-c}{n}\frac{G_{B_i}(b_{\alpha,\mathcal{N}'}|\mathcal{N}')}{g_{B_i}(b_{\alpha,\mathcal{N}'}|\mathcal{N}')}$$

so that

$$c = 1 - \frac{b_{\alpha,\mathcal{N}'} - b_{\alpha,\mathcal{N}}}{\frac{1}{n-1}\frac{G_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})}{g_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})} - \frac{1}{n}\frac{G_{B_i}(b_{\alpha,\mathcal{N}'}|\mathcal{N}')}{g_{B_i}(b_{\alpha,\mathcal{N}'}|\mathcal{N}')}}.$$

Again, for a given $\alpha$, this $c$ may not lie in $[0, 1)$ and, further, the right-hand side may not be constant in $\alpha$ as required. Thus, exogenous variation in participation may allow us to reject the CRRA model (or other parameterized utility functions). This suggests some hope of identifying preferences. In the completely general case, one would need to find a utility function such that

$$b_{\alpha,\mathcal{N}} + \lambda^{-1}\left(\frac{1}{n-1}\frac{G_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})}{g_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})}\right)$$

is invariant to $\mathcal{N}$ for each $\alpha$. Depending on how much variation in $b_{\alpha,\mathcal{N}}$ and $\frac{G_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})}{g_{B_i}(b_{\alpha,\mathcal{N}}|\mathcal{N})}$ is induced by variation in $\mathcal{N}$ and $\alpha$, it may be possible to identify the entire utility function.

[79] Bajari and Hortaçsu (2005) propose an alternative estimation approach.

### *6.4.2. Asymmetric preferences*

As we discussed at the start of Section 6, in many cases the econometrician is faced with several modeling alternatives when attempting to rationalize a given distribution of observables. So far, we have assumed that all bidders had the same preferences (either risk averse or risk neutral), but we have allowed distributions of valuations to vary across bidders. This allows us to reconcile bid distributions that vary across bidders. However, a natural alternative is that the distribution of valuations is the same for all bidders, but preferences differ. As mentioned above, in an ascending auction, behavior depends only on a bidder's valuation for the object, so it is impossible to distinguish these two cases. In a first-price auction with a fixed number of bidders, it is also difficult to distinguish these cases: in particular, as long as $\frac{G_{M_i|B_i}(b_i|b_i)}{g_{M_i|B_i}(b_i|b_i)}$ is increasing for each $i$, it is possible to rationalize bidding data using a model with homogeneous preferences. More generally, following the logic outlined above, there will generally exist homogeneous preferences with sufficient risk aversion such that

$$b_i + \lambda^{-1}\left(\frac{1}{n-1}\frac{G_{M_i|B_i}(b_i|b_i)}{g_{M_i|B_i}(b_i|b_i)}\right)$$

is increasing for all $i$.

However, there may be settings in which institutional information leads the econometrician to believe that a model with heterogeneous preferences is more natural than a model with heterogeneous value distributions. Campo (2002) has recently explored a model in which different bidders are permitted to have different preferences even though they draw their valuations from the same distribution. For $\alpha \in [0, 1]$ let $u_{i,\alpha}$ and $b_{i,\alpha}$ denote the $\alpha$th quantile of bidder $i$'s valuation and bid distributions, respectively. Generalizing our notation from above to allow bidders to have heterogeneous preferences represented by $\omega_i(\cdot)$, let $\lambda_i(s) \equiv \omega_i(s)/\omega_i'(s)$. Then, for all $i$, $j$, $\alpha$, we have

$$u_{i,\alpha} = b_{i,\alpha} + \lambda_i^{-1}\left(\frac{G_{M_i}(b_{i,\alpha})}{g_{M_i}(b_{i,\alpha})}\right),$$
$$u_{j,\alpha} = b_{j,\alpha} + \lambda_j^{-1}\left(\frac{G_{M_j}(b_{j,\alpha})}{g_{M_j}(b_{j,\alpha})}\right).$$

Since the distributions of valuations are assumed to be the same across bidders, it follows that

$$b_{i,\alpha} + \lambda_i^{-1}\left(\frac{G_{M_i}(b_{i,\alpha})}{g_{M_i}(b_{i,\alpha})}\right) = b_{j,\alpha} + \lambda_j^{-1}\left(\frac{G_{M_j}(b_{j,\alpha})}{g_{M_j}(b_{j,\alpha})}\right). \tag{6.31}$$

Campo (2002) shows that a set of observed bid distributions that are independent and satisfy standard regularity conditions can be rationalized using this model if and only if (i) there exist functions $\lambda_1(\cdot), \ldots, \lambda_n(\cdot)$ such that (6.31) holds for every quantile $\alpha \in [0, 1]$ where, for each $i$, $\lambda_i(0) = 0$, $\lambda_i'(\cdot) \geqslant 1$, and (ii) $\xi_i(b) = b + \lambda_i^{-1}(\frac{G_{M_i}(b)}{g_{M_i}(b)})$ is strictly increasing. There is no guarantee that these conditions can be satisfied, because (6.31) must hold for all $\alpha \in [0, 1]$.[80] Indeed, Campo (2002) provides an example where the conditions cannot be satisfied, and establishes that the set of bid distributions that can be rationalized using her model is a strict subset of those that can be rationalized using the model with homogeneous preferences and heterogeneous distributions of valuations.

Rather than analyze conditions under which risk preferences are nonparametrically identified, Campo (2002) takes a semi-parametric approach, with preferences given by $\omega(\cdot;\theta_i)$ (implying an associated $\lambda(\cdot;\theta_i)$), where $\theta_i$ is a finite dimensional parameter. To analyze identification, observe that for all $i, j, \alpha, \alpha'$, we have

$$\begin{aligned}
u_{i,\alpha} &= b_{i,\alpha} + \lambda^{-1}\left(\frac{G_{M_i}(b_{i,\alpha})}{g_{M_i}(b_{i,\alpha})};\theta_i\right),\\
u_{j,\alpha} &= b_{j,\alpha} + \lambda^{-1}\left(\frac{G_{M_j}(b_{j,\alpha})}{g_{M_j}(b_{j,\alpha})};\theta_j\right),\\
u_{i,\alpha'} &= b_{i,\alpha'} + \lambda^{-1}\left(\frac{G_{M_i}(b_{i,\alpha'})}{g_{M_i}(b_{i,\alpha'})};\theta_i\right),\\
u_{j,\alpha'} &= b_{j,\alpha'} + \lambda^{-1}\left(\frac{G_{M_j}(b_{j,\alpha'})}{g_{M_j}(b_{j,\alpha'})};\theta_j\right).
\end{aligned}$$

Suppose, for example, that each $\theta_i$ is a scalar. Since by assumption $u_{i,\alpha} = u_{j,\alpha}$ and $u_{i,\alpha'} = u_{j,\alpha'}$, for a given pair of quantiles $\alpha$ and $\alpha'$, this is a system of four equations in four unknowns $(u_{i,\alpha}, u_{i,\alpha'}, \theta_i, \theta_j)$, so that $\theta_i$ and $\theta_j$ are identified using data from just two quantiles. Once $\theta_i$ and $\theta_j$ are known, $F_U(\cdot)$ is uniquely determined by the first-order conditions and the observed distribution $G_{M_i}(\cdot)$. Similarly, once $F_U(\cdot)$ is identified, the first-order conditions and $G_{M_k}(\cdot)$ determine $\theta_k$ for $k \neq i, j$. Campo (2002) considers the case of CRRA preferences discussed above and gives the nonsingularity conditions for the system of equations above (restrictions on the pair $(G_{M_i}(\cdot), G_{M_j}(\cdot))$) that ensure identification in that case. It is crucial that there are some asymmetries in the bid distributions. She proposes a parametric estimation approach. We refer readers to her paper for details.

Since it is possible to identify $\theta_i$ and $\theta_j$ using data from just two quantiles of the bidding distribution when $\theta_i$ is a scalar, there is no guarantee that, given observed bid distributions, a particular functional form can rationalize the data at every quantile. Indeed, the example considered by Campo (2002) of CRRA preferences $\omega_i(u) = u^{1-c_i}$ requires the existence of constants $c_i, c_j$ on $[0, 1)$ such that

$$b_{i,\alpha} + (1-c_i)\left(\frac{G_{M_i}(b_{i,\alpha})}{g_{M_i}(b_{i,\alpha})}\right) = b_{j,\alpha} + (1-c_j)\left(\frac{G_{M_j}(b_{j,\alpha})}{g_{M_j}(b_{j,\alpha})}\right). \tag{6.32}$$

[80] Campo (2002) requires the condition $\lambda_i'(\cdot) \geqslant 1$ in order to guarantee that the induced preferences satisfy risk aversion; in fact, existence of an equilibrium in increasing strategies requires a slightly weaker condition, namely that $\ln(\omega_i(\cdot))$ is concave in the relevant region, which is guaranteed if $\lambda_i'(\cdot) \geqslant 0$. To see why log-concavity of $\omega_i(\cdot)$ is important, note that in an IPV auction, a bidder's objective function is $\omega_i(b_i - u_i)G_{M_i}(b_i)$. Maximizing this is equivalent to maximizing its logarithm; but, if $\ln(\omega_i(\cdot))$ is strictly convex, then $\frac{\partial^2}{\partial b_i \partial u_i}\omega_i(b_i - u_i) < 0$, so that bidders with higher valuations choose lower bids.

Suppose, for example, that there are just two bidders and that $B_1$ is uniformly distributed on [0, 1]. Then, (6.32) becomes, for all $\alpha \in [0,1]$,

$$\alpha + (1 - c_1)\left(\frac{G_{B_2}(\alpha)}{g_{B_2}(\alpha)}\right) = (2 - c_2) G_{B_2}^{-1}(\alpha). \tag{6.33}$$

Clearly, this places strong restrictions on $G_{B_2}(\cdot)$. For example, this would rule out a distribution of the form $G_{B_2}(b_2) = b_2^{\gamma}$ with $\text{supp}[B_2] = [0,1]$ (unless $\gamma = 1$, which would violate the assumption of asymmetric bid distributions).

Finally, we note that even when the data can be rationalized by both the homogeneous preference-heterogeneous valuations and the heterogeneous preference-homogeneous valuations models, it may be possible to extend the testing approaches described above that exploit exogenous variation in participation or other exclusion restrictions. When each model is identified for fixed $\mathcal{N}$, exogenous variation in $\mathcal{N}$ leads to over-identifying restrictions. In general, even if two different models rationalize the same data for fixed $\mathcal{N}$, the out-of-sample predictions of the models for $\mathcal{N}' \neq \mathcal{N}$ will differ between the two models. When data from auctions with both $\mathcal{N}$ and $\mathcal{N}'$ are observed, the out-of-sample predictions might be tested.

## 7. Common values auctions

While we have discussed a wide range of private values models in the preceding sections, in many applications a common values model may seem more natural. Recall that we use the term "common values" to refer to a broad class of models in which information about each bidder's valuation is dispersed among bidders (see Section 2).We emphasize, however, that the presence of factors affecting all bidders' valuations need not imply common values. For example, if

$$X_i = U_i = V_0 + \varepsilon_i$$

this is a private values specification despite the "common" factor $V_0$. Indeed, in this example each bidder knows his own valuation with certainty.[81] The presence of $V_0$ does introduce correlation of bidders' valuations and of bidders' information, and even causes one bidder's signal to be correlated with another's valuation; however, it does not introduce common values because no opponent has information that is relevant to a bidder's assessment of his own valuation, given that he has observed his own signal. The critical distinction concerns the nature of bidders' private information. When each bidder's private information concerns only idiosyncratic determinants of his own valuation, this is a private values setting.

[81] This is not essential for a private values environment. For example, if $U_i = X_i + \epsilon_i$ with $\epsilon_i$ independent of $X_j$ for all $j \neq i$, this remains a private values setting.

Nonetheless, many auction environments seem likely to fall in the common values category. Often the good for sale will not be consumed immediately (or the procurement contract being bid for will not be fulfilled immediately), and bidders may have different information about future states of the world – e.g., market conditions or the supply and demand of substitute objects. In some applications bidders will naturally have access to different information. A bidder might conduct her own seismic survey of an oil tract or might learn about market conditions from her own customers and suppliers. Furthermore, even if bidders have access to the same market data, they may have different algorithms or rules-of-thumb for using this information to form beliefs about the object's value. The output of one bidder's algorithm (i.e., its signal) might then be useful to another bidder in assessing her own valuation even after seeing the output of her own algorithm. In such cases it may be appropriate to model bidders as having different private information of a common values nature.

Aside from the potential prevalence of common values in practice, common values models are also of particular interest because they provide an example of a market environment in which adverse selection may play an important role. In a private values auction, bidders need only to follow a simple dominant ("bid your value") strategy in an ascending auction or to respond optimally to a distribution of opposing bids in a first-price auction. In a common values auction, bidders must understand the strategies that underlie the competing bids in order to make correct inferences about their informational content; in particular, bidders must account for the information that would be implied by their winning the auction in order to avoid the winner's curse. An important contribution of the empirical industrial organization literature has been to confirm some of the fairly subtle equilibrium predictions of common values auction models.[82] However, a number of positive and normative questions depend not just on whether bidder behavior is broadly consistent with theory, but on the exact structure of demand and information.

For example, typically the seller or auctioneer has some discretion over the auction rules. As first demonstrated by Milgrom and Weber (1982) for symmetric common values environments, the information revealed publicly by losing bidders' exits in an ascending auction reduces both the severity of the winner's curse and the informational rents obtained by the winner, leading to higher expected revenues than with a first-price sealed-bid auction. With asymmetries, first-price auctions may allocate the good inefficiently; however, they tend to raise more revenue in private values settings, may be less susceptible to collusion (detection and response to defections are more difficult than in an ascending auction), and may be less costly to administer. The choice of auction format also affects bidder entry when bidders are asymmetric [Klemperer (2002), Athey, Levin and Seira (2004)]. In trading off these factors, a seller must understand the underlying structure of bidder demand and information that determines the significance

[82] Examples include Hendricks, Porter and Boudreau (1987), Hendricks and Porter (1988), Hendricks, Porter and Wilson (1994), Athey and Levin (2001), and Haile (2001).

of each factor. Even within an auction format, the joint distribution of signals and valuations is important for positive questions (e.g., the division of surplus) and for design issues (e.g., the optimal reserve price, the optimal entry fee, and whether restrictions on participation would be profitable).[83]

### *7.1. Limits of identification with a fixed number of bidders*

In a common values environment, identifying the joint distribution $F_{\mathbf{X},\mathbf{U}}(\cdot)$ requires substantial restrictions on the underlying structure, and/or data beyond bids from a fixed environment. To suggest why, observe that in common values auctions the primitives of the model involve two different random variables for each bidder $i$: $X_i$ and $U_i$. Hence, the joint distribution $F_{\mathbf{X},\mathbf{U}}(\cdot)$ governs $2n$ random variables, yet an auction will reveal at most $n$ bids.[84] Even in the special case of pure common values, where $U_i = U_0$ for all $i$, the primitive of interest, $F_{\mathbf{X},U_0}(\cdot)$ has dimension $n+1$. So some additional structure will be necessary to obtain identification.

We begin by considering first-price auctions. One convenient normalization of signals (recall that this is without loss of generality) is[85]

$$E\Big[U_i \;\Big|\; X_i = \max_{j\neq i} X_j = x, \mathcal{N}\Big] = x. \tag{7.1}$$

With this normalization, (2.3) and the first-order condition (2.4) imply

$$v_i(x_i, x_i; \mathcal{N}) = x_i = b_i + \frac{G_{M_i|B_i}(b_i|b_i; \mathcal{N})}{g_{M_i|B_i}(b_i|b_i; \mathcal{N})}. \tag{7.2}$$

In Section 3.2 we discussed the identification and estimation of the distribution of the random variable

$$B_i + \frac{G_{M_i|B_i}(B_i|B_i; \mathcal{N})}{g_{M_i|B_i}(B_i|B_i; \mathcal{N})}.$$

All that changes when we consider common values settings is the interpretation of this distribution: using (7.2) we now interpret it as the distribution of the random variable $v_i(X_i, X_i; \mathcal{N})$.

This distribution alone will be sufficient for some questions of interest (see, e.g., Section 8), but certainly not all. In particular, it does not provide identification of the joint distribution $F_{\mathbf{U},\mathbf{X}}(\cdot)$. Consider the case in which one observes all bids from auctions with $n$ symmetric bidders. Under the private values assumption, $v(X_i, X_i; n) = X_i = U_i$,

[83] In a common values auction, restricting participation reduces the severity of the winner's curse, leading to more aggressive bidding. This can result in higher expected revenues despite the presence of fewer bidders, depending on the underlying distributions [e.g., Smiley (1979), Matthews (1984), Hong and Shum (2002)].

[84] Note that a normalization of signals does not change this argument, since the normalization cannot address the correlation between signals and valuations.

[85] Note, however, that this normalization cannot be maintained if $\mathcal{N}$ varies.

and the joint distribution $F_{\mathbf{U}}(\cdot)$ is just identified (see Section 3). Under the common values assumption the joint distribution of $(v(X_1, X_1; n), \ldots, v(X_n, X_n; n))$ is $F_{\mathbf{X}}(\cdot)$ under the normalization (7.1). Since this distribution is just identified when $n$ is fixed, it follows that it is impossible to distinguish common values from private values based on bidding data from first-price auctions with no reserve price and a fixed number of bidders [Laffont and Vuong (1996), Guerre, Perrigne and Vuong (2000)]. Thus, it is important to emphasize that any conclusions from data from a fixed set of bidders (and with no reserve price) rely on a maintained assumption of common values or private values. For example, it might be possible to justify a wide range of reserve prices as optimal for the seller under different assumptions about $F_{\mathbf{U},\mathbf{X}}(\cdot)$ that are consistent with the identified marginal distribution $F_{\mathbf{X}}(\cdot)$.

Ascending auctions are even more difficult in the common values setting. First, just as in a first-price auction, it would be impossible to distinguish common values from private values using a data set with a fixed number of bidders, even if all bids (including the planned exit price of the winner) were observed. Any observed distribution of bids could simply be equal to the distribution of private values for the bidders [Laffont and Vuong (1996)]. Second, exactly as in the case of a private values ascending auction, the unobservability of the winner's planned exit price can challenge even the identification of $F_{\mathbf{X}}(\cdot)$. Further, while a normalization like (7.1) can be applied to signals in the *initial* phase of an ascending auction (the period before any bidders drop out), no single normalization can induce the simple strategy $\beta_i(x_i, n) = x_i$ throughout the auction, since bidders modify their strategies each time an opponent exits. The exact forms of these modifications depend on the joint distribution of signals and valuations. While we might hope that this dependence would enable observed bids to provide information about this joint distribution, it also creates serious challenges. Finally, further complications arise from the fact that, when $n > 2$, there is a multiplicity of symmetric equilibria in weakly undominated strategies in common values auctions [Bikhchandani, Haile and Riley (2002)], implying that there is no unique interpretation of bids below the transaction price.

The following result from Athey and Haile (2002) establishes that the common values model is generally not identified in ascending auctions. Here we ignore the multiplicity of equilibria and assume a special case of a pure common values model in which signals are i.i.d. Even this very restrictive common values model is not identified.

THEOREM 7.1. *In an ascending auction, assume the pure common values model, i.i.d. signals $X_i$, and select the equilibrium characterized by* Milgrom and Weber (1982). *With n fixed, the model is not identified* (*even up to a normalization of signals*) *from the observable bids.*

PROOF. Take $n = 3$ and consider two models. In both, signals are uniform on [0, 1]. In the first, the value of the good is

$$u_0 = u(x_1, x_2, x_3) = \frac{\sum_i x_i}{3},$$

while in the second model

$$u_0 = \hat{u}(x_1, x_2, x_3) = \frac{x^{(1:3)}}{3} + \frac{x^{(2:3)}}{6} + \frac{x^{(3:3)}}{2}.$$

Because in both models $E[U_0 \mid X_1 = X_2 = X_3 = x] = x$, equilibrium bidding in the initial phase of the auction is identical in the two models in the Milgrom–Weber equilibrium (see Section 2.2.2); i.e., $G_B^{(1:3)}(b) = F_X^{(1:3)}(b) = 1-(1-b)^3$ in both cases. Similarly, since $b^{(2:3)} = E[U_0 \mid X^{(1:3)} = b^{(1:3)}, X^{(3:3)} = X^{(2:3)} = x^{(2:3)}]$, the fact that $\hat{u}(x, y, y) = u(x, y, y)$ for all $x$ and $y$ implies that $G_B^{(2:3)}(\cdot|B^{(1:3)})$ is identical under the two models. Since $G_B^{(1:3)}(\cdot)$ and $G_B^{(2:3)}(\cdot|B^{(1:3)})$ completely determine the joint distribution of the observable bids, the two models are observationally equivalent. □

This is a strong negative result for common values ascending auctions. Even ignoring the equilibrium selection problem and possible doubts about the interpretation of losing bids in an ascending auction, this most restrictive of common values models is not identified. This nonidentification is important for policy. Continuing the example from the proof of Theorem 7.1, consider the simple problem of setting an optimal reserve price for a second-price sealed-bid auction. Recalling the participation threshold (2.1), the optimal reserve price solves

$$\max_r 3(1 - F_X)\big(x^*(r, 3)\big)F_X\big(x^*(r, 3)\big)^2 r + \int_{x^*(r,3)}^{1} v(y, y; 3)6f_X(y)\big(1 - F_X(y)\big)F_X(y)\,\mathrm{d}y.$$

By construction, $v(x, x; 3)$ is the same for all $x$ in the two models. However, for any $r$, the participation threshold $x^*(r, 3)$ is lower in the second model, due to the reduced dependence of each $U_i$ on $X_{-i}$ when $X_i$ is maximal. Hence, the objective function above differs for the two models and (as can be confirmed directly) implies different optimal reserve prices.

We will see in Section 8 that variation in the number of bidders can be useful for overcoming at least some of these limitations. In particular, this variation can be sufficient to enable discrimination between private and common values models. Whether this kind of variation can go farther to enable nonparametric identification of a common values model is a question not yet explored. Below we will consider identification through additional structure and through additional data.[86]

## 7.2. *Pure common values*

Given the negative identification results obtained thus far, it is natural to consider whether additional assumptions can alleviate the problem. One possible approach is to restrict attention to the *pure* common values model, where $U_i = U_0\ \forall i$.

[86] Parametric models of common values auctions have been estimated by, e.g., Smiley (1979), Paarsch (1992a), Hong and Shum (2002, 2003), and Bajari and Hortaçsu (2003a).

In the pure common values model, the joint distribution $F_{\mathbf{X},\mathbf{U}}(\cdot)$ governs $n+1$ random variables $(U_0, X_1, \ldots, X_n)$; however, at most $n$ bids are revealed in a first-price auction, and only $n-1$ "bids" are revealed in the standard model of an ascending auction. This suggests that the pure common values assumption alone will not be sufficient to obtain identification, and that either additional structure or additional data will be needed. Below, we explore examples of both: we first consider additional restrictions on $F_{\mathbf{X},U_0}(\cdot)$, and then consider cases in which the realization of $U_0$ is observable *ex post*.

### *7.2.1. Identification with additional structure: The mineral rights model*

A special case of the pure common values model given considerable attention in the literature is the symmetric "mineral rights model" defined in Section 2.1. Here, bidders' signals are i.i.d. conditional on the realization of the common value $U_0$. As the name suggests, this model is motivated by auctions in which firms bid for the right to extract oil from an offshore tract. All firms may place the same value on the oil, since it is sold in a common market, but none knows how much oil (if any) there is. Each receives a seismologist's report, providing a (conditionally independent) noisy signal of $U_0$. This structure may be natural in other applications as well.

Even with this structure, identification from bid data is not straightforward since this requires somehow separating the variation in bids due to the randomness of $U_0$ from that due to the randomness of $X_i$ conditional on $u_0$. One possible approach is to assume a separable functional form like $X_i = U_0 + A_i$, where the "errors" $A_i$ mutually independent conditional on $U_0$. This can be useful, although the additive structure need not survive the normalization (7.1) in general. Put differently, while it will be useful for the left-hand side of the first-order condition (7.2) to have a separable form, one must be careful about what underlying structures on $(X_i, U_0)$ can deliver this separability. This is a question that has been explored by Li, Perrigne and Vuong (2000). To discuss their approach, we first define two (nested) special cases of the symmetric mineral rights model.

**Linear Mineral Rights (LMR):** $U_i = U_0$. In addition, for each $n$ there exist two known constants $(C, D) \in \mathbb{R} \times \mathbb{R}_+$ and random variables $(A_1, \ldots, A_n)$ with joint distribution $F_{\mathbf{A}}(\cdot)$ such that, with the normalization $E[U_0 \mid X_i = \max_{j \neq i} X_j = x, n] = x$, either (i) $X_i = \exp(C) \cdot (U_0 \cdot A_i)^D\ \forall i$, with $(A_i, U_0, X_i)$ nonnegative; or (ii) $X_i = C + D(U_0 + A_i)\ \forall i$. Further, conditional on $U_0$, the components of $\mathbf{A}$ are mutually independent and identically distributed.

**LMR with Independent Components (LMR-I):** In the LMR model, $(U_0, \mathbf{A})$ are mutually independent, with all $A_i$ identically distributed.

Li, Perrigne and Vuong (2000) focus on the LMR-I model and provide examples satisfying its assumptions. Under the LMR-I model, taking case (ii), (7.2) simplifies to

$$C + D(u_0 + a_i) = b_i + \frac{G_{M_i|B_i}(b_i|b_i; n)}{g_{M_i|B_i}(b_i|b_i; n)}. \tag{7.3}$$

Since $C$ and $D$ are known and the right-hand side of (7.3) is observable, it follows that the joint distribution of $(U_0 + A_1, \dots, U_0 + A_n)$ is identified from a data set containing all bids in first-price auctions. Li, Perrigne and Vuong (2000) note that standard deconvolution results, such as those used in the literature on measurement error (see Section 6.1.2), can then be used to separately identify the distributions $F_{U_0}(\cdot)$ and $F_A(\cdot)$.[87]

THEOREM 7.2. *Assume that for all $i$, the characteristic functions $\psi_{U_0}(\cdot)$ and $\psi_{A_i}(\cdot)$ of the random variables $U_0$ and $A_i$ are nonvanishing. If all bids are observed in a first-price auction, then the LMR-I model is identified.*

Even with these kinds of strong assumptions, identification is problematic when some bids are unobserved. Bids reveal realizations of order statistics of the form $U_0 + A^{(i:n)}$. Since order statistics are correlated even when the underlying random variables are independent, the identification approach based on the measurement error literature followed by Li, Perrigne and Vuong (2000) fails, unless all order statistics are observed (impossible in an ascending auction).

### *7.2.2. Identification and testing when ex post values are observable*

In some applications, an *ex post* measure of the realized common value $u_0$ will be observable to the econometrician. One notable example is an US outer-continental-shelf auction of drilling rights, where the quantities of oil and other minerals extracted from a tract are metered [e.g., Hendricks and Porter (1988), Hendricks, Pinkse and Porter (2003)]. Another example is a "scaled sale" timber auction, common in the US and Canada, where the quantity of each species of timber extracted from a tract is recorded by an independent agent at the time of harvest [e.g., Athey and Levin (2001)]. In other cases, resale prices can provide measures of realized values [e.g., McAfee, Takacs and Vincent (1999)]. Such additional data can be helpful in the mineral rights model.[88] In practice, the measures of $u_0$ available may be only imperfectly correlated with the true value to the bidders; we discuss this possibility below.

*7.2.2.1. First-price auctions* When we impose the structure of the symmetric pure common values model, the first-order condition (7.2) and the normalization (7.1) give

$$E\Big[U_0 \;\Big|\; X_i = \max_{j \neq i} X_j = x_i, n\Big] = x_i = b_i + \frac{G_{M_i|B_i}(b_i|b_i; n)}{g_{M_i|B_i}(b_i|b_i; n)}. \tag{7.4}$$

[87] Février (2004) has recently proposed an interesting alternative restriction of the mineral rights model that enables identification. He considers the case in which, conditional on $u_0$, each $X_i$ has support $[\underline{u}, u_0]$ and density $f_{x|u_0}(\cdot) = \frac{h(\cdot)}{H(u_0)}$ for some function $H(\cdot)$ with derivative $h(\cdot)$ satisfying $H(\underline{u}) = 0$. With this structure, conditional on having the highest signal, there is no information in the signals of one's opponents. He shows that this structure enables identification up to scale.

[88] Smiley (1979) was the first to suggest the value of such information. In his application he did not have access to an *ex post* measure and instead explored use of a noisy *ex ante* measure.

When all bids and the realization of $U_0$ are observed, (7.4) enables identification of the joint distribution $F_{\mathbf{X},U_0}(\cdot)$.

With knowledge of $F_{\mathbf{X},U_0}(\cdot)$, it is possible to perform counterfactual experiments, quantify the extent to which information is dispersed among the bidders, and characterize the magnitude of the "winner's curse." For example, it is interesting to examine the differences

$$E[U_0 \mid X_i = x_i, n] - E\Big[U_0 \Bigm| X_i = \max_{j \neq i} X_j = x_i, n\Big] \tag{7.5}$$

and

$$E[U_0 \mid X_i = x_i, n] - E\Big[U_0 \Bigm| X_i = x_i,\ \max_{j \neq i} X_j \leqslant x_i, n\Big] \tag{7.6}$$

since these provide a measure bidders' equilibrium responses to the winner's curse under the pure common values assumption.[89]

Hendricks, Pinkse and Porter (2003) were the first to suggest this and also proposed a test of equilibrium bidding in this model. Let

$$\zeta(b_i, b_j, n) = E\Big[U_0 \Bigm| B_i = b_i,\ \max_{j \neq i} B_j = b_j, n\Big]. \tag{7.7}$$

When the equilibrium bid function $\beta(\cdot; n)$ is strictly increasing, $\beta(x_i; n) = b_i$ implies

$$\begin{aligned} &E\Big[U_0 \Bigm| X_i = \max_{j \neq i} X_j = x_i, n\Big] \\ &\quad = E\Big[U_0 \Bigm| \beta(X_i; n) = \max_{j \neq i} \beta(X_j; n) = \beta(x_i; n), n\Big] \\ &\quad = E\Big[U_0 \Bigm| B_i = \max_{j \neq i} B_j = b_i, n\Big] \\ &\quad = \zeta(b_i, b_i, n). \end{aligned}$$

Thus, the first-order condition (7.4) can be written

$$\zeta(b_i, b_i, n) = b_i + \frac{G_{M|B}(b_i|b_i; n)}{g_{M|B}(b_i|b_i; n)} \equiv \xi(b_i, n). \tag{7.8}$$

[89] Hendricks, Pinkse and Porter (2003) point out that a positive value for the difference in (7.6) cannot be used as evidence against a private values assumption. The problem is that the interpretation of the empirical measure of $u_0$ as the realized value of the good relies on the pure common values assumption. For example, consider a symmetric independent private values environment and suppose the measured "*ex post* value" $u_0$ is actually just $\max_j u_j$, i.e., the value to the winner. Then the difference (7.6) is

$$E\Big[\max_j U_j \Bigm| U_i = u_i, n\Big] - E\Big[\max_j U_j \Bigm| U_i = u_i,\ \max_{k \neq i} U_k \leqslant u_i, n\Big]$$

which is positive. We will discuss approaches that can be used to discriminate private from common values models in Section 8.

Note that because the joint distribution of $(U_0, \mathbf{B}, N)$ is observable, $\zeta(b_i, b_i, n)$ is identified directly through Equation (7.7). No behavioral assumption is required for this identification: $\zeta(b_i, b_i, n)$ is simply a conditional expectation of the observable $U_0$ given that the observable bids satisfy $B_i = \max_{j \neq i} B_j = b_i$. Since $\xi(b_i, n)$ is also identified from the bidding data under the assumption of equilibrium bidding, the overidentifying restriction $\zeta(b_i, b_i, n) = \xi(b_i, n)$ can be tested.

To examine the differences (7.5) and (7.6) empirically, Hendricks, Pinkse and Porter (2003) first observe that since $b_i = \beta_i(x_i; n)$ and bidding is strictly monotonic, these differences are equal to the differences

$$E[U_0 \mid B_i = b_i, n] - E\Big[U_0 \;\Big|\; B_i = \max_{j \neq i} B_j = b_i, n\Big]$$

and

$$E[U_0 \mid B_i = b_i, n] - E\Big[U_0 \;\Big|\; B_i = b_i,\ \max_{j \neq i} B_j \leqslant b_i, n\Big].$$

They suggest a univariate local linear estimator $\hat{w}(b; n)$ of $E[U_0 \mid B_i = b, n]$, where $\hat{w}(b; n)$ is the solution for $w$ in the problem

$$\min_{w, \gamma} \sum_{t=1}^{T_n} \frac{1}{n} \sum_{i=1}^{n} \big(u_0 - w - \gamma(b - b_{it})\big)^2 \mathbf{1}\{n_t = n\} K\left(\frac{b - b_{it}}{h}\right)$$

with $K(\cdot)$ denoting a kernel and $h$ a bandwidth [see, e.g., Loader (1999)]. A similar estimator for $E[U_0 \mid B_i = \max_{j \neq i} B_j \leqslant b_i, n]$ is obtained by using only the winning bid $b_i$ from each auction $t$, rather than all bids.

To examine the overidentifying restriction (7.8), the right-hand side can be estimated using the kernel methods described in Section 3.2.1. A bivariate local linear estimator $\hat{v}(b)$ of the conditional expectation $\zeta(b_i, b_i, n) = E[U_0 \mid B_i = \max_{j \neq i} B_j = b_i, n]$ is obtained from the solution to

$$\min_{v, \gamma_1, \gamma_2} \sum_{t=1}^{T_n} \frac{1}{n} \sum_{i=1}^{n} \big(u_0 - v - \gamma_1(b - b_{it}) - \gamma_2(b - m_{it})\big)^2 \mathbf{1}\{n_t = n\} \times K\left(\frac{b - b_{it}}{h_1}\right) K\left(\frac{b - m_{it}}{h_2}\right).$$

Here $m_{it}$ is the maximum realized bid among $i$'s opponents at auction $t$. Hendricks, Pinkse and Porter (2003) suggest the use of the bootstrap to construct confidence intervals for testing.

Because of their focus on testing, Hendricks, Pinkse and Porter (2003) did not explore estimation of the joint distribution $F_{\mathbf{X}, U_0}(\cdot)$. However, with the normalization (7.1), an estimate of

$$E\Big[U_0 \;\Big|\; B_i = b_i,\ \max_{j \neq i} B_j = b_i, n\Big]$$

(such as that obtained using the first-order condition and kernel methods described in Section 3.2.1) provides an estimate of each realized $x_i$. Combining this with the observable realizations of $U_0$ presumably would enable consistent estimation of the joint distribution $F_{\mathbf{X},U_0}(\cdot)$ and/or the associated density $f_{\mathbf{X},U_0}(\cdot)$.

*7.2.2.2. Ascending auctions* Since the common values model is over-identified in a first-price auction when *ex post* values are observed, one might hope for identification in an ascending auction with similar data. However, the partial observability of bids in ascending auctions again presents serious challenges. Consider the case of two symmetric bidders. Recall that in an ascending auction, $\beta_i(X_i; n)$ is bidder $i$'s planned exit price when his opponent has not yet exited the auction. With $n = 2$ there is no problem of multiple equilibria [Bikhchandani, Haile and Riley (2002)], and

$$\begin{aligned} b_i &= \beta_i(x_i; 2) = E[U_0 \mid X_1 = X_2 = x_i,\ n = 2] \\ &= E\big[U_0 \bigm| \beta_1(X_1; 2) = \beta_2(X_2; 2) = \beta_i(x_i; 2),\ n = 2\big] \\ &\equiv \zeta(b_i, b_i, 2). \end{aligned}$$

However, it is not possible to estimate

$$\zeta(b_1, b_2, 2) = E[U_0 \mid B_1 = b_1,\ B_2 = b_2]$$

directly from the data (as is possible in a first-price auction) since in any given auction we observe the exit price of only one bidder. We never observe $B_1$ and $B_2$ in the same auction. We can observe the joint distribution of $(U_0, B^{(1:2)})$. Under (7.1), this is also the distribution of $(U_0, X^{(1:2)})$, but without additional structure this information is not sufficient to recover $F_{U_0,\mathbf{X}}(\cdot)$.

If we impose the additional structure of the mineral rights model, however, then conditional on $U_0$, $X^{(1:2)}$ is an order statistic from a sample of independent draws from $F_{X|U_0}(\cdot)$. Exploiting Equation (4.1), identification is then obtained in the symmetric two bidder case when the transaction price and *ex post* value are observed. Extending this approach to the case with $n > 2$ symmetric bidders is possible as well, using the order statistic $B^{(1:n)}$, although the suitability of this extension may be doubted in practice. To see one possible approach, note that with the normalization

$$x_i = E\big[U_0 \mid X_j = x_i\ \forall j \in \{1, \ldots, n\}\big]$$

we have $B^{(1:n)} = X^{(1:n)}$. Exploiting (4.1) again, we could recover the distribution of $X_i|U_0$, which then delivers identification of $F_{U_0,\mathbf{X}}(\cdot)$. This relies on the interpretation of bids implied by the button auction model, which may be especially dubious when applied to the interpretation of the lowest bid. To apply a similar approach using the transaction price $B^{(n-1:n)}$, it is still necessary to make use of the losing bids because the bidders themselves condition on this information. However, it may be easier to defend an approach which incorporates the information from all bids than one based entirely on the lowest bid. Here, we sketch one possibility.

Fix a set of realized values for the $n-2$ lowest bids at $(b^{(1:n)}, \dots, b^{(n-2:n)})$. Then, the (observable) distribution of

$$B^{(n-1:n)} \mid \{U_0 = u_0, B^{(1:n)} = b^{(1:n)}, \dots, B^{(n-2:n)} = b^{(n-2:n)}\}$$

is equal to the distribution of

$$B^{(n-1:n)} \mid \{U_0 = u_0,\ B^{(n-1:n)} \geqslant b^{(n-2:n)}\} \tag{7.9}$$

since bids (which are strictly increasing functions of the signals) are independent conditional on $U_0$ in the mineral rights model. Consider the following normalization of signals:

$$x = E\big[U_0 \mid X_{n-1} = X_n = x,\ B^{(1:n)} = \dots = B^{(n-2:n)} = \inf\big[\operatorname{supp}[B_i]\big]\big].$$

Imposing this normalization, when $b^{(n-2:n)} = \inf[\operatorname{supp}[B_i]]$, the random variable in (7.9) is equal in distribution to

$$X^{(n-1:n)} | U_0 = u_0.$$

Equation (4.1) then uniquely determines the distribution of $X_i | U_0$ and thus $F_{U_0,\mathbf{X}}(\cdot)$, since $U_0$ is directly observable.

This result is not as strong as one might hope for. It relies on an interpretation of the losing bids in an ascending auction (although it is not essential that the bidders use Milgrom and Weber's (1982) equilibrium) and on an assumption that the econometrician's inferences about exit prices match those that the bidders make during the auction. Furthermore, the identification argument relies on the tails of the distribution of bids. In particular, building a nonparametric estimator based on this argument would seem to require an estimate of the distribution of $B^{(n-1:n)}$ conditional on both the *ex post* value $u_0$ and the event that $n-2$ losing bids are increasingly close to the bottom of the support of the bid distribution. Whether this kind of approach can work well in sample sizes typically available has not been investigated.

*7.2.2.3. Biased or noisy observations of ex post values* So far in this section we have assumed that the econometrician observes the true realization of $U_0$. In the case of an oil auction, for example, this requires that the oil extracted is measured without error and that the econometrician has accurate measures of all costs (including opportunity costs) incurred in extracting the oil. These may be strong assumptions in some applications, so it is useful to consider the degree to which they can be relaxed. With the exception of Yin (2004), which does not fully address identification, the literature has not analyzed the issue of imperfect measures of *ex post* values.[90] Here, we present some initial results.

[90] Yin (2004) obtained descriptions of items auctioned on eBay and recruited volunteers to make subjective assessments of the value of the objects. The mean assessment was then treated as a potentially biased proxy for $U_0$. Smiley (1979, Appendix), explored the use of a noisy *ex ante* measure of $U_0$ within a parametric model.

Consider a first-price auction in which all bids are observable and suppose the available measure of $U_0$ is

$$\widetilde{U}_0 = \gamma_0 + \gamma_1 U_0 + \varepsilon, \tag{7.10}$$

where $\gamma_0$ and $\gamma_1$ are fixed parameters, unknown to the econometrician, and $\varepsilon$ is an unobserved random variable satisfying $E[\varepsilon \mid \mathbf{X} = \mathbf{x}] = 0$ for all $\mathbf{x}$. Recall that, given the normalization (7.1), $F_{\mathbf{X}}(\cdot)$ is identified from the bidding data alone in this setting. Since we observe $\widetilde{U}_0$ in every auction, the distribution $F_{\mathbf{X},\widetilde{U}_0}(\cdot)$ is identified as well. With this, we can compute

$$\eta(x) \equiv E\Big[\widetilde{U}_0 \;\Big|\; X_i = \max_{j \neq i} X_j = x\Big].$$

Given (7.10), the normalization (7.1), and $E[\varepsilon \mid \mathbf{X} = \mathbf{x}] = 0$, we also have

$$\eta(x) = \gamma_0 + \gamma_1 E\Big[U_0 \;\Big|\; X_i = \max_{j \neq i} X_j = x\Big] = \gamma_0 + \gamma_1 x. \tag{7.11}$$

Since $\eta(\cdot)$ is identified, the joint distribution of $(\eta(X), X)$ is identified, as is the bias in the measure of the common value, determined by the parameters $\gamma_0$ and $\gamma_1$.

Identification of $\gamma_0$, $\gamma_1$, and $F_{\mathbf{X},\widetilde{U}_0}(\cdot)$ implies identification of quantities such as

$$E[U_0 \mid X_i = x] = E[\widetilde{U}_0 - \gamma_0 \mid X_i = x]/\gamma_1,$$

and

$$E\Big[U_0 \;\Big|\; X_i = x,\ \max_{j \neq i} X_j \leqslant x\Big] = E\Big[\widetilde{U}_0 - \gamma_0 \;\Big|\; X_i = x,\ \max_{j \neq i} X_j \leqslant x\Big]/\gamma_1,$$

so that the differences (7.5) and (7.6) discussed above are identified, for example. Unfortunately, unless $\varepsilon$ is degenerate, the variance of $U_0$ is not identified, nor is the joint distribution $F_{\mathbf{X},U_0}(\cdot)$. In the setting studied by Yin (2004), where $\widetilde{U}_0$ is the mean estimate from a survey, the assumption that $\varepsilon$ is degenerate may be a reasonable approximation when the number of survey respondents per auction is large.

In ascending auctions, the analysis is more complex. Let us focus on the case of a pure CV model with two bidders. Let us maintain the assumption $E[\varepsilon \mid \mathbf{X} = \mathbf{x}] = 0$ for all $\mathbf{x}$, and in addition, assume that $(X_1, \ldots, X_n)$ are independent conditional on $\widetilde{U}_0$. The joint distribution of $(B^{(1:2)}, \widetilde{U}_0)$ is then identified and, under (7.1), this is equal to the joint distribution of $(X^{(1:2)}, \widetilde{U}_0)$. Given that $X_1$ is independent of $X_2$ conditional on $\widetilde{U}_0$, the parent distribution $F_X(\cdot)$ is identified using (4.1), so that the joint distribution of $(\mathbf{X}, \widetilde{U}_0)$ is identified. This completely determines $\eta(\cdot)$, which in turn yields identification of $\gamma_0$ and $\gamma_1$ through (7.11).

Of course, the assumption that $\mathbf{X}$ is independent conditional on $\widetilde{U}_0$ may be strong, especially if $\widetilde{U}_0$ is a noisy measure of $U_0$. In practice this may be most defensible when $\varepsilon$ is degenerate, so that $\widetilde{U}_0$ is a deterministic function of $U_0$.

## 8. Private versus common values: Testing

Negative identification results for common values models provide one important motivation for formal tests that could distinguish between common and private values models. Distinguishing private values from common values was, in fact, the goal behind Paarsch's (1992a) pioneering work on structural empirical approaches to auctions. The distinction between the two paradigms is central to our understanding of behavior in auction markets and has important implications for market design. For example, revenue superiority of an ascending auction relative to a second-price sealed-bid auction in symmetric settings [Milgrom and Weber (1982)] holds only in a common values environment. Furthermore, a common values environment is one with adverse selection. There is relatively little evidence on the empirical significance adverse selection, and an examination of the prevalence of common values in auctions might be suggestive of the nature of private information in other market environments as well.

It might be surprising that questions about the qualitative nature of private information could be answered at all empirically. In fact, early approaches to testing based on reduced-form relationships between bids and the number of bidders were eventually discovered to be invalid. With the structural approach proposed by Paarsch (1992a), it was possible to test particular common values or private values models, but only with maintained parametric distributional assumptions. More recently, Laffont and Vuong (1996) have pointed out that private values and common values models are empirically indistinguishable, suggesting that testing was impossible (see Section 7.1).[91] However, they did not consider the possibilities created by variation in the number of bidders or a binding reserve price. Below we will show how either of these can offer approaches for discriminating between private and common values.

In the case of variation in the number of bidders, the idea is simple. The winner's curse is present only in common values models and becomes more severe as more competitors are added. Having a signal of the object's value that is the highest among twenty implies a more severely biased signal than does having the highest signal among two, for example. This greater severity manifests itself as a reduction in a bidder's expectation of his valuation conditional on winning a large auction. In particular, while the unconditional expectation $E[U_i \mid X_i = x_i]$ is invariant to the set of opponents $i$ faces, his equilibrium bid reflects a downward adjustment

$$E[U_i \mid X_i = x_i] - v_i(x_i, x_i; \mathcal{N})$$

that accounts for the information implied by his bid being pivotal. In a symmetric environment, this downward adjustment is always larger when $i$ faces more competition.

To make this precise, suppose that the number of bidder is exogenous and let $\mathcal{N}_{+j}$ denote the set of bidders comprised of all members of $\mathcal{N}$ plus bidder $j$.[92]

[91] This can be thought of as a nonidentification result for the affiliated values model, which nests the private and common values models.

[92] Note that because the normalization (7.1) depends on the set of bidders, we could not maintain this particular normalization for all $\mathcal{N}$. Other normalizations, e.g., $x_i = F_{X_i}(x_i)$, are of course possible.

LEMMA 8.1. *Suppose the number of bidders varies exogenously. With private values, $v_i(x, x; \mathcal{N}) = v_i(x, x; \mathcal{N}_{+j})$ for all $x$, $\mathcal{N}$, $i$, $j$. With common values and symmetric bidders, $v_i(x, x; \mathcal{N}) > v_i(x, x; \mathcal{N}_{+j})$ for all $x$, $\mathcal{N}$, $i$, $j$.*

PROOF. The first claim is immediate from the fact that $v_i(x, x; \mathcal{N}) = x$ with private values. With common values and symmetric bidders,

$$v_i(x, x; \mathcal{N}) = E\Big[U_i \;\Big|\; X_i = x,\ \max_{k \in \mathcal{N}_{-i}} B_k = \beta(x; \mathcal{N})\Big].$$

Let $m = \arg\max_{k \in \mathcal{N}_{-i}} B_k$ and suppose $j \notin \mathcal{N}$. Then

$$\begin{aligned}
&v_i(x, x; \mathcal{N})\\
&\quad = E\big[U_i \mid X_i = x,\ B_m = \beta_i(x; \mathcal{N}),\ B_k \leqslant \beta_i(x; \mathcal{N})\ \forall k \in \mathcal{N} \text{ s.t. } k \neq i, m\big]\\
&\quad = E_{X_j}\big[E\big[U_i \mid X_i = x,\ X_m = x,\ X_k \leqslant x\ \forall k \in \mathcal{N} \text{ s.t. } k \neq i, m\big]\big]\\
&\quad > E\big[U_i \mid X_i = x,\ \max\{X_m, X_j\} = x,\ X_k \leqslant x\ \forall k \in \mathcal{N} \text{ s.t. } k \neq i, j, m\big]\\
&\quad = v_i(x, x; \mathcal{N}_{+j})
\end{aligned}$$

where the last two lines follow from Definition 2.1 and the strict monotonicity of equilibrium bidding strategies.[93] □

This result provides the basis for testing using variation in the number of bidders. Although to our knowledge the proof was first given by Athey and Haile (2002) and Haile, Hong and Shum (2003), the idea behind this result and its potential value for detecting the winner's curse goes back at least to Gilley and Karels (1981), who suggested regressing bids from first-price auctions on the number of bidders as a test of a common values model. This reflected a belief that bids must increase with $n$ in a private values auction (since adding bidders makes the auction more competitive) but might decline in $n$ in a common values auction if the winner's curse were sufficiently severe to overcome the competitive effect of adding additional bidders [see, e.g., Brannman, Klein and Weiss (1987), Paarsch (1992a, 1992b), Laffont (1997)]. However, Pinkse and Tan (2005) have recently shown that this is incorrect: bids may increase or decrease in $n$ in both private values and common values models. The regression approach might seem more promising in an ascending auction, due to the simplicity of equilibrium bid functions in the button auction model. The multiplicity of equilibria in common values auctions creates one problem. But even ignoring this [e.g., selecting the equilibrium of the button auction given by Milgrom and Weber (1982)] this approach fails due to the fact that the winner's bid is never revealed. For example, in a private values auction the observable bids reveal $(u^{(1:n)}, \ldots, u^{(n-1:n)})$, but $u^{(n:n)}$ is censored. Because the

[93] Note that the second equality need not hold without symmetry. Conditions under which more competition (appropriately defined) implies a more severe winner's curse have not been fully explored.

distribution of the censored valuation $U^{(n:n)}$ varies with $n$, so does the resulting censoring bias. This makes it impossible to discriminate between private values and common values models based on a regression of bids on $n$.[94]

In spite of this, and in spite of the lack of identification of many common values models, testing is often possible. Lemma 8.1 makes use of the assumption of exogenous (to the distribution of signals and valuations) variation in the number of bidders. As discussed by Athey and Haile (2002) and Haile, Hong and Shum (2003), this can be reasonable in some applications. Furthermore, it is implied by some models of participation (see Section 6.3.2). However, the assumption of exogenous participation is not always necessary. Initially we will maintain this assumption to simplify the exposition of the basic testing approaches. In Section 8.2 we discuss an approach to testing with endogenous participation.

### 8.1. Testing in first-price auctions when all bids are observed

In the common values first-price auction, the first-order condition (7.2) requires that $v(x_i, x_i; \mathcal{N}) = \xi(b_i, \mathcal{N})$. Note that both sides of this equation vary with $\mathcal{N}$. However, because $\xi(b_i, \mathcal{N})$ is identified, it is possible to isolate the effect of $\mathcal{N}$ on $v(x_i, x_i; \mathcal{N})$ when $\mathcal{N}$ varies exogenously. Since $v(x_i, x_i; \mathcal{N})$ does not vary with $\mathcal{N}$ in a private values model, it is possible to distinguish the two models, even though $F_{\mathbf{X},\mathbf{U}}(\cdot)$ is not identified. To see how, let $F_{v_i,\mathcal{N}}(\cdot)$ denote the marginal distribution of the random variable $v_i(X_i, X_i; \mathcal{N})$. Lemma 8.1 implies the following result.

COROLLARY 8.1. *Assume exogenous variation in the number of bidders. Then $F_{v_i,\mathcal{N}}(v)$ is invariant to $\mathcal{N}$ in a private values model for all $v$. In a common values model with symmetric bidders*

$$F_{v_i,\mathcal{N}}(v) < F_{v_i,\mathcal{N}_{+j}}(v) \tag{8.1}$$

*for all $i$, $j$ and all $v$ on the interior of the support of $F_{v_i,\mathcal{N}}(\cdot)$ or $F_{v_i,\mathcal{N}_{+j}}(\cdot)$.*

Haile, Hong and Shum (2003) use this result to develop tests of the null hypothesis of private values against the common values alternative.[95] They focus on the case of symmetric bidders, where $F_{v_i,\mathcal{N}}(\cdot)$ can be more simply represented by $F_{v,n}(\cdot)$, and (8.1) can be written

$$F_{v,n}(v) < F_{v,n+1}(v) \quad \forall n, v. \tag{8.2}$$

[94] However, as Bajari and Hortaçsu (2003a) have pointed out, the recurrence relation (8.6) below implies that with this censoring, the average observed bid must increase in $n$ in the dominant strategy equilibrium of a private values button auction. While this is also possible in a common values auction, it provides a testable restriction of the private values hypothesis.

[95] They apply their tests to two types of auctions held by the US Forest Service. Shneyerov (2005) has recently applied one of their tests to data from municipal bond auctions.

Their approach involves two steps. The first is to form estimates $\hat{v}_{it}$ of each $v(x_{it}, x_{it}; n_t)$ using the methods described in Section 3.2. The second step is to compare the empirical distributions

$$\widehat{F}_{\hat{v},n}(v) = \frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n_t}\mathbf{1}\{n_t = n,\ \hat{v}_{it} \leqslant v\}$$

for different values of $n$.

While tests of equality of distributions (or of the alternative of first-order stochastic dominance) are common in statistics and econometrics, a complication here is the fact that only empirical distributions of the "pseudo-values" $\hat{v}_{it}$ can be compared, not those of the "values" $v(x_{it}, x_{it}; n)$. Hence, the first-stage estimation error (which will be correlated in finite sample for nearby $\hat{v}_{it}$ and $\hat{v}_{jt'}$) must be accounted for. A second complication is the fact that trimming, which must be done separately for each value of $n$, must be done carefully to avoid creating the appearance of a winner's curse when there is none, or hiding the winner's curse in a common values model.

Haile, Hong and Shum (2003) explore two types of tests.[96] The first is a comparison of trimmed means of each empirical distribution $\widehat{F}_{\hat{v},n}(\cdot)$.[97] For $\tau \in (0, \frac{1}{2})$ let $x_\tau$ denote the $\tau$th quantile of the marginal distribution $F_X(\cdot)$ and define the quantile-trimmed mean

$$\mu_{n,\tau} = E\big[v(X_i, X_i; n)\ \big|\ X_i \in [x_\tau, x_{1-\tau}]\big].$$

Trimming at the same quantiles for all $n$ fixes the set of signals $x_i$ implicitly included in each mean. This is important since the first-order stochastic dominance relation in (8.2) extends to the distributions of $v(X_i, X_i; n)$ over any fixed interval in $[\underline{x}, \overline{x}]$ but need not hold for intervals that vary with $n$. One can then test the hypotheses

$$\begin{aligned} &H_0\colon\quad \mu_{\underline{n},\tau} = \cdots = \mu_{\bar{n}}, \\ &H_1\colon\quad \mu_{\underline{n},\tau} > \cdots > \mu_{\bar{n}}, \end{aligned} \tag{8.3}$$

which are implied by Lemma 8.1.

Let $b_\tau$ denote the $\tau$th quantile of the observed bids. Since bids are strictly increasing in signals, $\mu_{n,\tau}$ has sample analog

$$\hat{\mu}_{n,\tau} = \frac{1}{nT_n}\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbf{1}\big\{n_t = n,\ b_{it} \in [b_\tau, b_{1-\tau}]\big\}\hat{v}_{it}.$$

Haile, Hong and Shum (2003) show that the vector $(\hat{\mu}_{\underline{n},\tau}, \ldots, \hat{\mu}_{\bar{n},\tau})$ is consistent and has a multivariate normal asymptotic distribution with diagonal covariance matrix $\Sigma$, enabling adaptation of a standard multivariate one-sided likelihood-ratio test

[96] Section 5 discusses several other hypotheses to which their tests may be adaptable.

[97] The test generalizes to other finite vectors of functionals – e.g., a vector of quantiles. See Haile, Hong and Shum (2003) for details.

[Bartholomew (1959)]. Monte Carlo evidence suggests that size distortions may be reduced by using the bootstrap to estimate the elements of the covariance matrix $\Sigma$.[98]

The second testing approach uses a generalized version of a multi-sample one-sided Kolmogorov–Smirnov test of equal distributions. Given a differentiable strictly decreasing function $\Lambda(\cdot)$, let

$$\Lambda_n(v) = \frac{1}{nT_n} \sum_{t=1}^{T} \sum_{i=1}^{n} \mathbf{1}\{n_t = n\} \Lambda(\hat{v}_{it} - v)$$

and

$$\overline{\delta}_T = \sum_{n=\underline{n}}^{\overline{n}-1} \sup_{v \in [\underline{v}, \overline{v}]} \bigl[\Lambda_{n+1}(v) - \Lambda_n(v)\bigr],$$

where the compact interval $[\underline{v}, \overline{v}]$ is bounded away from the endpoints of the support $F_{v,n}(\cdot)$ under the null. If $\Lambda(\cdot)$ is the smoothed step function

$$\Lambda(y) = \frac{\exp(-y/h)}{1 + \exp(-y/h)}$$

with $h$ denoting a bandwidth, $\overline{\delta}_T$ is easily interpreted as an approximation of a more familiar looking one-sided test statistic

$$\delta_T = \sum_{n=\underline{n}}^{\overline{n}-1} \sup_{v \in [\underline{v}, \overline{v}]} \bigl\{\widehat{F}_{\hat{v},n+1}(v) - \widehat{F}_{\hat{v},n}(v)\bigr\},$$

where $\widehat{F}_{\hat{v},n+1}(\cdot)$ and $\widehat{F}_{\hat{v},n}(\cdot)$ denote empirical distribution functions.

Strict monotonicity of $\Lambda(\cdot)$ and the fact that

$$\Lambda_n(v) \to E\bigl[\Lambda(\hat{v}_{it} - v) \bigm| n_{it} = n\bigr]$$

uniformly in $v \in [\underline{v}, \overline{v}]$ imply that $\overline{\delta}_T \to 0$ under the private values null. Under the common values alternative $\overline{\delta}_T \to \delta > 0$. This is the basis for using $\delta_T$ as a test statistic. Haile, Hong and Shum (2003) show that for an appropriate normalizing sequence $\eta_T$, the generalized Kolmogorov–Smirnov statistic $S_T \equiv \eta_T \overline{\delta}_T$ has a nondegenerate limiting distribution under $H_0$, enabling use of subsampling for estimation of critical values [e.g., Politis, Romano and Wolf (1999)].

Both types of test are easily adapted to the case in which bidders observe only a signal $\eta$ of the number of opponents they face before submitting their bids, as long as the econometrician also observes (or can condition on) $\eta$. In that case estimation of pseudo-values follows the discussion in Section 6.3.3. One could then compare the distribution of pseudo-values in auctions with higher signals to those with lower signals.

[98] The block bootstrap procedure is identical to that discussed in Section 3.2.1. Haile, Hong and Shum (2003) point out that using the bootstrap to estimate the distribution of the test statistic itself would be difficult, due to the need to resample bids under the null hypothesis on the functions $v(\cdot,\cdot;n)$.

### 8.2. Testing with endogenous participation

Haile, Hong and Shum (2003) discuss extensions of their testing approaches to cases in which bidder participation is endogenous. If there is a binding reserve price or a cost of preparing a bid, for example, bidders' participation decisions introduce truncation in the set of types submitting bids. They show how their basic approach can still be applied in such cases by comparing estimated distributions of $v(X_i, X_i; n)$, appropriately adjusted for truncation, on regions of common support. We refer readers to their paper for details.

A more difficult case is that in which participation is affected by unobserved factors that also affect valuations. This leads to two quite different threats to the basic testing approach. First, variation in $F_{v,n}(\cdot)$ with $n$ will arise from variation in the unobserved factors, confounding attempts to detect responses to the winner's curse. For example, if auctions of goods that are more valuable in unobserved dimensions also attract greater participation, this could mask the effects of the winner's curse in a common values auction. The second problem is even more fundamental: unobserved heterogeneity threatens the identification of the distributions $F_{v,n}(\cdot)$ that underlies the approach (recall Section 6.1.2).

Haile, Hong and Shum (2003) have proposed an instrumental variables approach for such situations. Consider a simplified version of their approach in which the number of actual bidders in auction $t$ is a function of two scalar factors $Z_t$ and $W_t$:

$$A_t = \alpha(Z_t, W_t).$$

Here $Z_t$ is an index capturing the effects of factors observable to the econometrician as well as the bidders, while $W_t$ is an index capturing the effects of unobservables.[99] Assume that (i) $Z$ is independent of $(X_1, \dots, X_{\bar{n}}, U_1, \dots, U_{\bar{n}})$ and (ii) $\alpha(\cdot,\cdot)$ is weakly increasing in its first argument and strictly increasing in its second.

Assumption (i) is a standard exclusion restriction: $Z_t$ is an instrument affecting participation but not the distribution of valuations and signals. This instrument might be the number of potential bidders or a proxy for it, like the number of firms in the local market. Of course, in principle there need not be any difference between the potential and actual bidders here, based on our definitions. For example, if there is a cost of acquiring a signal but bidders have access to some information about the distribution of valuations before bearing this cost, the number of potential bidders will be correlated with unobservable factors shifting all bidder valuations. Valid instruments in that case might be the number of firms in the market (those who choose whether to invest in a signal), or factors affecting the cost of acquiring a signal.

Assumption (ii) is a monotonicity restriction. Monotonicity in the instrument $Z_t$ implies that changes in $Z_t$ will provide the exogenous variation in the level of competition

[99] For simplicity we assume there are no auction-specific observables other than $Z_t$, although this is easily relaxed.

that will make it possible to isolate the effects (if any) of the winner's curse. Strict monotonicity in $W_t$ is a key restriction that requires that $W_t$ be discrete (since $A_t$ is). As discussed in Section 6.1.2, this restriction enables identification of the expectations

$$v(x, x; a, z) = E[U_i \mid X_i = x,\ A_t = a,\ Z_t = z]$$

through the first-order condition

$$\begin{aligned} &v(x_{it}, x_{it}; a_t, z_t) \\ &\quad = b_{it} + \frac{\Pr(\max_{j\neq i} B_{jt} \leqslant b_{it} \mid B_{it} = b_{it},\ Z_t = z_t,\ A_t = a_t)}{\frac{\partial}{\partial m}\Pr(\max_{j\neq i} B_{jt} \leqslant m \mid B_{it} = b_{it},\ Z_t = z_t,\ A_t = a_t)|_{m=b_{it}}}. \end{aligned} \tag{8.4}$$

Estimation of the pseudo-values on the left-hand side of (8.4) proceeds by holding fixed both the value of $A$ and the value of the instrument $Z$ to construct estimators of the right-hand side of (8.4). To test for common values, the pseudo-values $v(x_{it}, x_{it}; a_t, z_t)$ are then pooled across realizations of $A_t$ to compare the cumulative distributions

$$F_{v,z}(v) = \Pr\bigl(E\bigl[v(X_{it}, X_{it}; A_t, Z_t)\bigr] \leqslant v \bigm| Z_t = z\bigr)$$

across values of $z$. While these distributions must be the same for all $z$ under private values, the assumptions above imply that $F_{v,z}(v)$ is increasing in $z$ under the common values alternative. Haile, Hong and Shum (2003) provide additional details and an alternative control function estimation approach allowing for multiple instruments. Their application to US Forest Service timber auctions uses the numbers of sawmills and logging firms in a geographic neighborhood of a sale as instruments for the number of bidders.

### 8.3. Testing with incomplete bid data

Athey and Haile (2002) show that testing is also possible in ascending auctions (assuming the button auction model) and in first-price auctions in which not all bids are observable.[100] For the symmetric common values model, recall that the challenge arises because the distribution of $v(X^{(n-1:n)}, X^{(n-1:n)}, n)$ varies with $n$ both due to the winner's curse and because the distribution of the order statistic varies with $n$ even without any winner's curse. However, since $X^{(n:n)}$ is unobserved, the distribution of $v(X_i, X_i, n)$ is not identified.

[100] Haile (2001) develops a different testing approach based on detecting bidders' updating of their willingness to pay as an ascending auction proceeds. The insight is that there is no such updating in a private values auction or in a 2-bidder common values auction. Hence one can compare distributions of bidders' willingness to pay (i.e., $\phi(G_B^{(n-1:n)}(\cdot); n-1, n)$) in 2-bidder auctions to that in auctions with larger numbers of bidders, with a difference suggesting common values. A major limitation of this approach is a requirement of independent signals under both the null and alternative hypotheses. While independence is implied by Haile's model of auctions with resale, this will typically be a strong restriction for a common values auction.

Athey and Haile's (2002) approach exploits the fact that for exchangeable random variables $Y_1, \ldots, Y_n$, the marginal distributions $F_Y^{(i:n)}(\cdot)$ of the order statistics must satisfy the recurrence relation [see, e.g., David (1981)]

$$\frac{n-i}{n} F_Y^{(i:n)}(y) + \frac{i}{n} F_Y^{(i+1:n)}(y) = F_Y^{(i:n-1)}(y) \quad \forall y, n, i \leqslant n-1. \tag{8.5}$$

Intuitively, in an *ex ante* sense, moving from a sample of $n$ draws to a sample of $n-1$ draws is equivalent under exchangeability to taking the $n$ draws and then dropping one at random. When one draw, $Y_j$, is dropped at random from the larger sample, the $i$th order statistic in the smaller sample will be either the $i$th order statistic from the larger sample (when $Y_j$ was one of the $n-i$ highest draws), or the $(i+1)$st order statistic (if $Y_j$ was among the $i$ lowest draws). Note that one direct implication of (8.5) is a recurrence relation between means:

$$\frac{n-i}{n} E\big[Y^{(i:n)}\big] + \frac{i}{n} E\big[Y^{(i+1:n)}\big] = E\big[Y^{(i:n-1)}\big] \quad \forall n, i \leqslant n-1. \tag{8.6}$$

Using (8.5) and (8.6), the private values null can be tested against the common values alternative in both first-price and ascending auctions. This is possible even when not all bids are observable (as is always the case in an ascending auction) and despite the fact that the ascending auction has multiple equilibria in the case of common values. The following theorem combines results originally given in Athey and Haile (2002).

THEOREM 8.1. *Assume exogenous variation in the number of bidders. In an ascending auction or first-price sealed-bid auction, the symmetric private values model is testable against the symmetric common values alternative if one observes the bids $B^{(n-2:n)}$ and $B^{(n-1:n)}$ in the ascending auction, or $B^{(n-1:n)}$ and $B^{(n:n)}$ in a first-price auction.*

PROOF. For the first-price auction, we have seen in Theorem 3.3 that the marginal distributions $F_v^{(n-1:n)}(\cdot)$ and $F_v^{(n:n)}(\cdot)$ of $v(X^{(n-1:n)}, X^{(n-1:n)}; n)$ and $v(X^{(n:n)}, X^{(n:n)}; n)$ are identified for all $n$. In a private values auction, these distributions are $F_U^{(n-1:n)}(\cdot)$ and $F_U^{(n:n)}(\cdot)$ so that (8.5) implies the restriction

$$\frac{1}{n} F_v^{(n-1:n)}(v) + \frac{n-1}{n} F_v^{(n:n)}(v) = F_v^{(n-1:n-1)}(v) \quad \forall v.$$

Under the common values alternative, $v(x, x; n)$ is still a strictly increasing function of $x$, so that the random variables $v(X_i, X_i; n)$ are still exchangeable. But since $v(X_i, X_i; n)$ strictly decreases in $n$ (Lemma 8.1),

$$\frac{1}{n} F_v^{(n-1:n)}(v) + \frac{n-1}{n} F_v^{(n:n)}(v) > F_v^{(n-1:n-1)}(v)$$

for all $v$ on the interior of the support of $F_{v_i,\mathcal{N}}(\cdot)$ or $F_{v_i,\mathcal{N}_{+j}}(\cdot)$.

For the ascending auction, under the private values null, Equation (8.6) implies

$$\frac{2}{n} E\big[B^{(n-2:n)}\big] + \frac{n-2}{n} E\big[B^{(n-1:n)}\big] = E\big[B^{(n-2:n-1)}\big] \quad \forall n > \underline{n}.$$

Under the common values alternative, Athey and Haile (2002, Theorem 9) show that, regardless of the equilibria selected in the $n$-bidder and $(n-1)$-bidder auctions, one obtains the relation

$$\frac{2}{n}E\big[B^{(n-2:n)}\big] + \frac{n-2}{n}E\big[B^{(n-1:n)}\big] < E\big[B^{(n-2:n-1)}\big] \quad \forall n > \underline{n}.$$

□

While Theorem 8.1 relies on exchangeability, Athey and Haile (2002) show how this kind of approach can be adapted to asymmetric ascending auctions as well.[101] To see the key idea, observe that if $(Y_1, \ldots, Y_n)$ have an arbitrary joint distribution, one can obtain a sample of exchangeable random variables $(Y_{R_1}, Y_{R_2}, \ldots, Y_{R_s})$ by taking a random subsample of size $R_s < n$ from the original sample $(Y_1, \ldots, Y_n)$. Hence, even without exchangeability of $(Y_1, \ldots, Y_n)$, a recurrence relation must hold for random subsamples [Balasubramanian and Balakrishnan (1994)]. In a private values auction this implies a recurrence relation between distributions $F_U^{(i:n)}(\cdot)$ in auctions with bidders $\mathcal{N}$ and those from smaller auctions in which the set of bidders is a subset of $\mathcal{N}$.

Formal testing approaches based on these results have not yet been explored. Since the null (alternative) hypothesis can be represented as the hypothesis of equal (stochastically ordered) distributions, it may be possible to adapt the testing approaches of Haile, Hong and Shum (2003), which account for the estimation error arising from the nonparametric estimation of pseudo-values.

### 8.4. Testing with a binding reserve price

While Haile, Hong and Shum (2003) show that their testing approach can be extended to cases in which there is a binding reserve price, Hendricks, Pinkse and Porter (2003) and Haile, Hong and Shum (2003) have each shown that the presence of a binding reserve price can make possible a different sort of test for the winner's curse in first-price auctions. Both approaches rely on observing the number of potential bidders.

Focusing on the symmetric case, recall that participation is determined by the threshold signal $x^*(n)$ defined by (recall Equation (2.1))

$$x^*(n) = \inf\Big\{x\colon\ E\Big[U_i \;\Big|\; X_i = x,\ \max_{j\neq i} X_j \leqslant x\Big] \geqslant r\Big\}. \tag{8.7}$$

The equilibrium bid of a bidder with signal $x^*(n)$ is

$$\beta\big(x^*(n); n\big) = v\big(x^*(n), x^*(n); n\big) = E\Big[U_i \;\Big|\; X_i = x^*(n),\ \max_{j\neq i} X_j = x^*(n)\Big].$$

In a private values model, $E[U_i \mid X_i = x,\ \max_{j\neq i} X_j \leqslant x] = E[U_i \mid X_i = x,\ \max_{j\neq i} X_j = x]$, so that

$$\beta\big(x^*(n); n\big) = r. \tag{8.8}$$

[101] They also discuss extension to cases in which only nonadjacent values of $n$ are observed in the data.

As originally noted by Milgrom and Weber (1982), in a common values model the fact that $E[U_i \mid X_i = x,\ \max_{j \neq i} X_j \leqslant x] < E[U_i \mid X_i = x,\ \max_{j \neq i} X_j = x]$ implies

$$\beta\bigl(x^*(n); n\bigr) > r. \tag{8.9}$$

Hence, a test for common values can be based on the distinction between (8.8) and (8.9). In particular, if we let $\underline{b} = \inf \operatorname{supp} B_i$,

$$\lim_{b \to \underline{b}} b + \frac{G_{M|B}(b|b; n)}{g_{M|B}(b|b; n)}$$

should equal $r$ under the private values hypothesis but should be strictly greater than $r$ with common values. While this idea was first mentioned by Hendricks, Pinkse and Porter (2003), a formal test based on this idea has not yet been developed.

A second possibility, suggested by Haile, Hong and Shum (2003), is to examine variation in the probability $F_X(x^*(n))$ that the reserve price excludes a potential bidder. It is easy to verify (following the proof of Lemma 8.1) that $x^*(n)$ is invariant to $n$ in a private values model but strictly increasing in $n$ in a common values model. By exchangeability,

$$F_X\bigl(x^*(n)\bigr) = F_{\mathbf{X}}\bigl(x^*(n), \infty, \ldots, \infty\bigr) = \sum_{k=1}^{n} \frac{k}{n} \Pr(A = n - k \mid N = n).$$

So if both the number of potential bidders, $N$, and the number of actual bidders, $A$, are observed, $F_X(x^*(n))$ is identified for all $n$, and one can test whether this is constant or decreasing in $n$.

## 9. Dynamics

Until very recently, virtually all structural empirical work on auctions has considered static models, treating each auction in the data as an independent game. There are several reasons this may not be the case. First, even in a stationary environment, dynamic considerations arise if firms engage in collusion.[102] We do not consider collusion in this chapter.[103] Second, bidders' valuation distributions may change over time in a way that

[102] Many models of collusion at auctions are static [e.g. McAfee and McMillan (1992)]. Recently, the theory of tacit collusion in repeated auctions has grown rapidly [Aoyagi (2003), Athey and Bagwell (2001, 2004a), Athey, Bagwell and Sanchirico (2004) and Skryzpacz and Hopenhayn (2004)]. Athey, Bagwell and Sanchirico (2004) show that when only the winning bid, but not the bidder's identity, is revealed by the auctioneer, optimal collusion entails bidding at the reserve price with each bidder having an equal chance of winning, while Athey and Bagwell (2001, 2004a) show that when the bidder's identity is revealed as well, bidders engage in sophisticated rotation schemes so that a bidder's probability of winning is less correlated over time than the bidder's valuation.

[103] For empirical studies of collusion (which typically do not explicitly consider dynamics), see Porter and Zona (1993, 1999), Bajari and Ye (2003), Pesendorfer (2000), Baldwin, Marshall and Richard (1997), and Athey, Levin and Seira (2004). See Bajari and Summers (2002) for a survey.

is exogenous, but is private information to each bidder.[104] This can create dynamic links in bidder strategies. In particular, a bidder's behavior in an auction will affect opponents' beliefs about his valuation distribution in future auctions, changing the equilibrium of the auction game in each period.[105] To our knowledge, there has been no empirical investigation focusing on the dynamics of such models.

Finally, the underlying distribution of valuations might change as a function of auction outcomes, potentially in ways that are observable (or can be directly inferred) by the other bidders. For example, there may be learning-by-doing, so that a firm that wins an auction today might have stochastically lower costs (higher valuations) in the future. Alternatively, firms may have capacity constraints (or more general forms of diseconomies of scale). In that case, a firm that wins an auction today might draw a valuation from a less favorable distribution in the future. In either case, the resulting dynamic considerations for bidders will change the equilibrium at each point in time.

To explore this type of environment, consider a model based on that of Jofre-Bonet and Pesendorfer (2000, 2003).[106] Time is discrete, and firms compete over an infinite horizon. In each period $t$, an item is sold by first-price auction to one of $n$ bidders. For simplicity, assume that there is no reserve price and that all objects to be auctioned have the same observable characteristics [see Jofre-Bonet and Pesendorfer (2003) for extensions]. The distribution of bidder valuations depends on the bidders' capacities (more generally, it could depend on other covariates as well). Conditional on capacities, bidder valuations are independent across bidders and over time. Letting $c_{i,t}$ be bidder $i$'s publicly observable capacity in period $t$, the conditional distribution of bidder $i$'s valuation in period $t$ is denoted $F_U(\cdot|c_{i,t})$, where for simplicity we let this function be the same for all bidders.

The econometrician and the bidders both know the (deterministic) transition function for bidder capacities. In particular, if $k$ is the identity of the winning bidder in period $t$ and $\mathbf{c}_t$ is the vector of bidder capacities in period $t$, then[107]

$$c_{i,t+1} = \omega_i(\mathbf{c}_t, k).$$

The solution concept is Markov-perfect equilibrium. Thus, collusion is ruled out, and dynamic considerations in bidder strategies arise only because bidders anticipate that the identity of today's winner will affect future capacities, which in turn will affect

[104] If the distribution of valuations changes over time in a way that is observed by all bidders, then either the econometrician can observe (and condition on) the factors affecting distribution, or the problem of unobserved heterogeneity discussed in Section 6.1.2 arises. The literature has not explored intertemporal correlation in unobserved auction heterogeneity.

[105] See e.g., Bikhchandani (1988), Bikhchandani and Huang (1989), Haile (1999, 2003), Katzman and Rhodes-Kropf (2002), Das Varma (2003), Goeree (2003), and Athey and Bagwell (2004b).

[106] They analyze a procurement auction. We recast the problem as one in which the bidders are buyers. They also consider two types of bidders, "regular bidders" who bid often, and "fringe bidders" who bid rarely and do not consider the future. We focus on regular bidders to simplify the analysis.

[107] Jofre-Bonet and Pesendorfer (2003) allow a slightly richer specification in which transitions reflect information about the size and duration of projects that have been won in the past.

outcomes in future auctions. Since all asymmetries are captured through capacities, Jofre-Bonet and Pesendorfer (2003) focus on exchangeable strategies. In particular, each bidder's bid in a given period depends on the bidder's own valuation and the vector of capacities, so that strategies can be written $\beta_i(u_{i,t}, \mathbf{c}_t)$.[108]

In this environment, Jofre-Bonet and Pesendorfer (2003) combine the insights of Hotz and Miller (1993) (who studied dynamic discrete choice problems for individuals) with the approach of Guerre, Perrigne and Vuong (2000) in an insightful way, providing very general conditions for identification when the discount factor is known.

The first step in the analysis is to use dynamic programming to represent bidder payoffs.[109] Suppressing $\mathcal{N}$ in the notation, let $G_{M_i}(\cdot|\mathbf{c})$ be the equilibrium distribution of the maximum opponent bid for bidder $i$ when the vector of bidder capacities is $\mathbf{c}$. Let $\delta$ denote the discount factor and let $\omega(\mathbf{c}, k) = (\omega_1(\mathbf{c}, k), \ldots, \omega_n(\mathbf{c}, k))$. Holding opponents' strategies fixed, the interim expected discounted sum of future profits for bidder $i$ is given by

$$W_i(u_i, \mathbf{c}) = \max_{b_i}\Biggl\{(u_i - b_i)G_{M_i}(b_i|\mathbf{c}) + \delta\sum_{j=1}^{n}\Pr(j \text{ wins } |b_i, \mathbf{c})\int_{u_i'} W_i\bigl(u_i', \omega(\mathbf{c}, j)\bigr) f_{U_i}\bigl(u_i' \mid \omega_i(\mathbf{c}, j)\bigr)\,\mathrm{d}u_i'\Biggr\},$$

where the second term sums over the possible identities of the winner to form an expectation of the continuation value to player $i$, given current capacities. One can then define the *ex ante* value function

$$V_i(\mathbf{c}) = \int W_i(u_i, \mathbf{c}) f_{U_i}(u_i|\mathbf{c})\,\mathrm{d}u_i,$$

which can be rewritten as

$$V_i(\mathbf{c}) = \int\Biggl\{\max_{b_i}\Biggl\{(u_i - b_i)G_{M_i}(b_i|\mathbf{c}) + \delta V_i\bigl(\omega(\mathbf{c}, i)\bigr) + \delta\sum_{j\neq i}\Pr(j \text{ wins } |b_i, \mathbf{c})\bigl[V_i\bigl(\omega(\mathbf{c}, j)\bigr) - V_i\bigl(\omega(\mathbf{c}, i)\bigr)\bigr]\Biggr\}\Biggr\} f_{U_i}(u_i|\mathbf{c})\,\mathrm{d}u_i.$$

Note that in equilibrium, the probability that bidder $i$ wins with bid $b_i$ is given by

$$\Pr\Bigl(b_i \geqslant \max_{j\neq i}\beta_j(U_j, \mathbf{c}) \Bigm| \mathbf{c}\Bigr) = G_{M_i}(b_i|\mathbf{c}) = \prod_{j\neq i} G_{B_j}(b_i|\mathbf{c}), \tag{9.1}$$

[108] Jofre-Bonet and Pesendorfer (2000, 2003) establish existence of an equilibrium within the parametric framework they use for estimation, and also sketch an approach for showing existence in general.

[109] Ackerberg et al. (Chapter 63 in this volume) discuss estimation of dynamic strategic models more generally, which relies on very similar ideas.

where $G_{B_j}(\cdot|\mathbf{c})$ is the cumulative distribution of $B_j$ conditional on capacity vector $\mathbf{c}$. The probability that bidder $j \neq i$ wins when bidder $i$ bids $b_i$ is

$$\int_{b_i}^{\bar{b}_j(\mathbf{c})} \left( \prod_{k \neq i, j} G_{B_k}(b_j|\mathbf{c}) \right) g_{B_j}(b_j|\mathbf{c})\, \mathrm{d}b_j,$$

where $\bar{b}_j(\mathbf{c}) = \sup \operatorname{supp} G_{B_j}(\cdot|\mathbf{c})$. Finally, using (9.1) note that

$$\frac{G_{M_i}(b_i|\mathbf{c})}{g_{M_i}(b_i|\mathbf{c})} = \frac{1}{\sum_{j \neq i} \frac{g_{B_j}(b_i|\mathbf{c})}{G_{B_j}(b_i|\mathbf{c})}}.$$

The next step is to solve for the *ex ante* value functions in terms of observables. This requires a significant generalization of the two-step indirect approach proposed by Guerre, Perrigne and Vuong (2000). Consider bidder $i$'s optimization problem in a given auction:

$$\max_{b_i} \Bigg\{ (u_i - b_i) G_{M_i}(b_i|\mathbf{c}) + \delta V_i\big(\omega(\mathbf{c}, i)\big) + \delta \sum_{j \neq i} \left( \int_{b_i}^{\bar{b}_j(\mathbf{c})} \prod_{k \neq i, j} G_{B_k}(b_j|\mathbf{c}) g_{B_j}(b_j|\mathbf{c})\, \mathrm{d}b_j \right) \big[ V_i\big(\omega(\mathbf{c}, j)\big) - V_i\big(\omega(\mathbf{c}, i)\big) \big] \Bigg\}.$$

The first-order condition is

$$u_i = b_i + \frac{G_{M_i}(b_i|\mathbf{c})}{g_{M_i}(b_i|\mathbf{c})} + \delta \sum_{j \neq i} \frac{G_{M_i}(b_i|\mathbf{c})}{g_{M_i}(b_i|\mathbf{c})} \frac{g_{B_j}(b_i|\mathbf{c})}{G_{B_j}(b_i|\mathbf{c})} \big( V_i\big(\omega(\mathbf{c}, j)\big) - V_i\big(\omega(\mathbf{c}, i)\big) \big). \tag{9.2}$$

After substituting this into the *ex ante* value function, a change of variables yields

$$\begin{aligned} V_i(\mathbf{c}) = {} & \int_{\underline{b}_i(\mathbf{c})}^{\bar{b}_i(\mathbf{c})} \frac{G_{M_i}(b_i|\mathbf{c})}{g_{M_i}(b_i|\mathbf{c})} G_{M_i}(b_i|\mathbf{c})\, \mathrm{d}G_{B_i}(b_i|\mathbf{c}) \\ & + \delta \sum_{j \neq i} V_i\big(\omega(\mathbf{c}, j)\big) \Bigg\{ \int_{b_i}^{\bar{b}_i(\mathbf{c})} \prod_{k \neq i, j} G_{B_k}(b_j|\mathbf{c}) g_{B_j}(b_j|\mathbf{c})\, \mathrm{d}b_j \\ & + \int_{\underline{b}_i(\mathbf{c})}^{\bar{b}_i(\mathbf{c})} \frac{G_{M_i}(b_i|\mathbf{c})}{g_{M_i}(b_i|\mathbf{c})} \frac{g_{B_j}(b_i|\mathbf{c})}{G_{B_j}(b_i|\mathbf{c})} G_{M_i}(b_i|\mathbf{c})\, \mathrm{d}G_{B_i}(b_i|\mathbf{c}) \Bigg\}. \end{aligned}$$

For any capacity vector $\mathbf{c}$, this expresses each $V_i(\mathbf{c})$ as a linear function of $V_i(\cdot)$ evaluated at other capacity vectors. The coefficients of this linear relation depend only on the observable bid distributions. Thus, it is possible to solve for the *ex ante* value functions in terms of the observable bid distributions.

Once the *ex ante* value functions have been computed, identification of the distributions $F_U(\cdot|\mathbf{c})$ (assuming the discount factor $\delta$ is known) follows from the first-order condition (9.2). In particular, we can use the observed bid distributions and the *ex ante* value functions to compute the right-hand side of (9.2). Then, if the discount factor $\delta$ is known (for example, from other empirical studies), (9.2) implies that $F_{U_i}(\cdot|\mathbf{c})$ is identified from the observed bids and capacities.

In addition to demonstrating the nonparametric identification of their model, Jofre-Bonet and Pesendorfer (2003) propose a parametric estimation approach, motivated in part by a desire to include covariates in a parsimonious manner. To solve for the value functions, they follow Judd (1998) and discretize the set of possible capacities. Then, calculating the value functions entails solving a system of linear equations. They further simplify the estimation by using a quadratic approximation of the value function. They apply their approach to California highway construction contracts. Using their estimates, they are able to assess the importance of private information, capacity constraints, and the inefficiencies that arise due to the asymmetries induced by capacity differences among bidders under the assumption of forward-looking equilibrium behavior. Note that it is impossible to *test* whether bidders are forward looking in this environment, since the discount factor $\delta$ is not identified.

## 10. Multi-unit and multi-object auctions

### *10.1. Auctions of perfect substitutes*

While most of the empirical literature on auctions focuses on the case of single-unit auctions, auctions of multiple units of identical goods (“multi-unit auctions”) have recently begun to gain significant attention. One motivation is their importance in the public sector. For example, multi-unit auctions have recently been implemented in restructured electricity markets to assign electric power generation to different plants [see, e.g., Wolfram (1998), Borenstein, Bushnell and Wolak (2002), or Wolak (2003)]. Optimal design of such markets is complex: the usual goals of efficiency and surplus extraction in single-unit auctions are complicated by (among other issues) nonlinearities in cost functions, incentives to exercise market power by withholding marginal production capacity, and the need for firms to recover substantial fixed costs [Wolak (2003)]. Empirical analysis of these markets can provide valuable information about the underlying cost structure, market power opportunities, and profitability. Another important policy question that has been the subject of discussion among economists at least since Friedman (1960) is how governments should auction treasury securities to maximize revenues. This question is potentially relevant to the design of markets for other types of securities as well.

In treasury auctions, a large number of identical securities is sold in a mechanism in which each bidder submits an entire “demand function,” i.e., each bidder $i$ offers a (downward sloping) schedule of price-quantity combinations $(b_{ij}, q_j)$ specifying the price he is willing to pay for his $q_j$th unit.[110] Two auction mechanisms are commonly used: *discriminatory* and *uniform-price*. A discriminatory auction is the most common in practice (although recently the US adopted uniform-price auctions after conducting an experiment to evaluate alternative formats). In this mechanism, each bidder who offers more than the market clearing bid for a unit receives that unit at the price he offered. As the name suggests, this results in different prices for different units of the same security – even a given bidder will pay different prices for each unit he wins. In contrast, in a uniform-price auction, the market clearing price (lowest accepted bid) is paid on all units sold. In addition to US treasury bill auctions, electricity auctions are often uniform price, and some firms have used uniform price auctions in initial public offerings.[111]

Of course, bidders will bid differently depending on whether a discriminatory or uniform-price auction is used. The revenue ranking of the two mechanisms is theoretically ambiguous [Ausubel and Cramton (2002)] and can only be determined with knowledge of the true distribution of bidder valuations. To our knowledge, the literature has not yet presented a comprehensive analysis of identification and estimation in uniform-price auctions, although Wolak (2003) provides some initial results.

Before proceeding, we pause to highlight the fact that the theory of multi-unit auctions is much less well developed than the theory of first-price auctions and ascending auctions. Although existence of equilibrium in mixed strategies can be guaranteed quite generally [Jackson et al. (2002), Jackson and Swinkels (2005)], existence of pure strategy Nash equilibria in monotone strategies has been established for only a limited class of models, such as private or common value models where bidder signals are independent [McAdams (2004a)]. In addition, examples have shown that there can be multiple equilibria [e.g. Back and Zender (1993)]. Thus, most existing econometric approaches to these auctions require assumptions on endogenous variables to guarantee that the requisite regularity properties are satisfied, although in practice some of the conditions can be verified empirically.

Hortaçsu (2002) has empirically analyzed the question of which auction mechanism raises higher revenue, and has shown that the relevant primitives can be identified nonparametrically in a private values model of the discriminatory auction.[112] His empirical model is based on the theoretical framework of Wilson (1979).[113] Building on the insight of Guerre, Perrigne and Vuong (2000), he points out that equilibrium bidding

[110] Note that this is a bidder's *strategic* expression of quantities he demands at each price. This need not correspond to the usual notion of a price-taking buyer's demand function.

[111] In the finance literature, these are often referred to as "Dutch auctions," conflicting with economists' use of this term for descending price single-unit auctions.

[112] Parametric structural models have been studied recently by Février, Préget and Visser (2002) and by Armantier and Sbaï (2003), both of which consider common values models. Common values models may be appropriate for many securities auctions, although this is ultimately an empirical question – one for which testing approaches have not been developed. Hortaçsu (2002) discusses institutional details that motivate the assumption of private values in the case of the Turkish treasury bill auctions he studies.

[113] The analysis in Wilson's model relies on an assumption that bidders can bid continuous demand functions. In most applications, bids are restricted (by rule or in practice) to step functions – i.e., finite sets of discrete

strategies can be characterized as best responses by each bidder to the distribution of opposing bids he faces. In this multi-unit setting, the distribution of opposing bids cannot be described by the distribution of the maximum opposing bid (as in a single-unit auction); rather, it is the stochastic *residual supply curve* that characterizes the equilibrium probabilities with which various quantities could be obtained at each possible price.

For a discriminatory auction, suppose that the total quantity of securities to be offered is $Q$. Let $q_i(\cdot)$ denote the demand function offered by bidder $i$; i.e., $q_i(p)$ is the largest quantity for which bidder $i$ is offering a price of $p$ or more for his final unit. For a given set of demand functions $q_1(\cdot), \ldots, q_n(\cdot)$, the market clearing price $p^{\mathrm{c}}$ then equates supply and demand:

$$Q = \sum_i q_i\bigl(p^{\mathrm{c}}\bigr).$$

This market clearing price can be reinterpreted as the price at which $i$'s demand function and his residual supply curve intersect:

$$Q_{R_i}(b) = Q - \sum_{j \neq i} q_j(b).$$

Let $v_i(y; x_i)$ denote bidder $i$'s marginal valuation for a $y$th unit of the good, given the realization of his signal $x_i$. Each bidder $i$'s equilibrium strategy specifies, for each possible realization of $x_i$, a demand function $q_i(b) = \varphi_i(b; x_i)$ expressing the quantity demanded at each price $b$. Let

$$G_i(b, y) = \Pr\biggl(y \leqslant Q - \sum_{j \neq i} \varphi_j(b; X_j)\biggr) \tag{10.1}$$

so that $G_i(b, y)$ is the probability that, given equilibrium bidding by $i$'s opponents, the market clearing price falls below $b$ if $i$ himself demands quantity $y$ at price $b$.

For each $X_i = x_i$, bidder $i$'s optimal strategy $\varphi_i(\cdot; x_i)$ then solves the problem

$$\max_{q_i(\cdot)} \int_0^\infty \biggl( \int_0^{q_i(p^{\mathrm{c}})} \bigl(v_i(y, x_i) - q_i^{-1}(y)\bigr)\, \mathrm{d}y \biggr) \frac{\partial G_i(p^{\mathrm{c}}, q_i(p^{\mathrm{c}}))}{\partial p^{\mathrm{c}}}\, \mathrm{d}p^{\mathrm{c}}.$$

One can show that the optimal bidding strategy can be characterized by the necessary condition

$$v_i\bigl(\varphi(b; x_i), x_i\bigr) = b + \frac{G_i(b, \varphi(b; x_i))}{\frac{\partial}{\partial b} G_i(b, \varphi(b; x_i))}.$$

While this is an Euler–Lagrange condition for a functional optimization problem, this equation closely resembles the first-order condition (2.4) used by Guerre, Perrigne and

price-quantity pairs. Recently, Wolak (2004), McAdams (2005), and Kastl (2005) have explored empirical models explicitly accounting for this discreteness.

Vuong (2000) to show identification of the single-unit discriminatory (i.e., first-price sealed-bid) auction with private values. Its role in the identification argument is similar. Because the demand functions $q_j(b) = \varphi_i(b; x_j)$ are directly observed, $G_i(b, y)$ is identified from Equation (10.1). Then, for any quantity $y$ demanded at price $b$ by bidder $i$, we have

$$v_i(y, x_i) = b + \frac{G_i(b, y)}{\frac{\partial}{\partial b} G_i(b, y)},$$

which uniquely determines the realizations of bidder $i$'s marginal valuations at each quantity $y$. This implies identification of the distributions of each $v_i(y, X_i)$, which are the primitives needed for policy simulations.[114] In particular, if for each quantity $y$ we let $B_i^y$ be a random variable equal in distribution to $\varphi_i^{-1}(y; X_i)$, $v_i(y, X_i)$ must be equal in distribution to

$$B_i^y + \frac{G_i(B_i^y, y)}{\frac{\partial}{\partial b} G_i(b, y)|_{b=B_i^y}}.$$

Hortaçsu (2002) explores several estimation approaches, both parametric and nonparametric. He also finds a clever way to place an upper bound on the revenue that would be obtained under the uniform price auction, while avoiding the difficult problem of solving for the equilibrium given the estimated distribution of valuations: since a bidder will never bid more than her marginal valuation for each unit, the revenue that would be obtained if bidders simply bid their marginal valuations for each unit in a uniform auction provides an upper bound on the equilibrium revenue.

### 10.2. Auctions of imperfect substitutes and complements

One prominent area in which economists' understanding of auctions has been used to guide policy over the last decade is in the design of institutions to allocate spectrum rights [see, e.g., McAfee and McMillan (1996)]. Questions regarding the optimal design of spectrum auctions led to much new theoretical work considering the complications to equilibrium strategies arising in multi-object auctions, where the heterogeneous goods auctioned at the same time may be imperfect substitutes, complements, or combinations of these. Similar issues arise in a number of procurement applications, where complementarities may exist between contracts, and some bundles of contracts may be substitutes for others. Cantillon and Pesendorfer (2003) study one such application: auctions for bus services in London, where it may be cheaper to operate one route if a nearby route is also served. Here, we describe their model and identification results. For

[114] Note that signals play a purely informational role here. Hence, their distribution can be normalized (and assumed symmetric) without loss of generality. Put differently, only the distribution of marginal valuations, not that of the underlying signals, is of economic relevance.

consistency with the rest of the chapter, we treat the auction as one in which the bidders are buyers rather than sellers of services.

Let $S$ be the set of goods offered for sale, with $|S| = m$. Let $U_{i,s}$ be bidder $i$'s valuation for the bundle $s \subseteq S$, with $\mathbf{U}_i \in \mathbb{R}^{2^m-1}$ denoting the vector of his valuations for all possible bundles $s \subseteq S$. Bidders' preferences over combinations of goods may exhibit sub- and/or super-additivity. Let $F_{\mathbf{U}_i}(\cdot)$ be the joint distribution of $\mathbf{U}_i$, while $F_{U_{i,s}}(\cdot)$ denotes the marginal distribution of $U_{i,s}$. Let $B_{i,s}$ denote bidder $i$'s bid on bundle $s$, and let $\mathbf{B}_i$ be the vector of bids placed by bidder $i$. We let $\mathbf{B}_{i,-s}$ denote the vector of bids placed by bidder $i$ on all bundles other than $s$.

For simplicity, we focus on a fixed set of $n$ symmetric bidders. Bidders participate in a sealed-bid discriminatory auction with *combination bidding*: each bidder submits bids on all bundles, and the auctioneer chooses the allocation of all objects that maximizes total revenue, charging each bidder the price he offered for each bundle he is allocated.

Combination bidding enables bids to express complementarities and substitutabilities between objects and/or bundles. Further, we might expect combination bidding to aid efficiency and to encourage less cautious bidding by bidders who desire certain combinations of goods. However, combination bidding also introduces a strategic incentive absent in auctions of homogeneous goods. This arises from the fact that a bidder's bid on one bundle competes with his own bids on other bundles. If a bidder raises his bid for bundle $s$, for example, that will make him more likely to win $s$, but it may reduce his chances of winning a different bundle $t$. This is because an increase in $b_{i,s}$ may make it profitable for the seller to allocate bundle $s$ to $i$ instead of bundle $t$, allocating $t$ to some other bidder instead.

A bidder's problem here turns out to be very closely related to the problem of multiproduct pricing, where it is known that a firm may find it profitable to bundle goods for which demands are independent. Analogously, here a bidder may find it profitable to place bids on bundles (i.e., to make "combination bids" or "bundle bids") even if the goods in the bundle are independent in the sense that $U_{i,s\cup t} = U_{i,s} + U_{i,t}$ when $s \cap t = \emptyset$. This is because the combination bid on the bundle $s \cup t$ can enable bidder $i$ to win bundle $s$ even when bidder $i$'s opponents place a high bid for bundle $s$, unless they also place a high bid for bundle $t$. Thus, the combination bid allows bidder $i$ to "leverage" a high valuation for bundle $s$ into a lower price paid for bundle $t$, or vice versa [cf. Whinston (1989)]. Note that this leads bidder $i$ to bid less aggressively on the individual bundles $s$ and $t$, in order to avoid competing with her combination bid.

Following intuition from the literature on bundling [see McAfee, McMillan and Whinston (1989) or Armstrong and Rochet (1999)], as long as the correlation among opponent bids for $s$ and $t$ is not too high, making a combination bid is profitable for the bidder. Cantillon and Pesendorfer (2003) describe a plausible environment in which allowing combination bids will reduce both expected revenue and efficiency if goods are independent. This provides one motivation for determining whether bidders view the goods as independent, substitutes, or complements.

For the purposes of this section, we will make the following nonprimitive assumptions (Cantillon and Pesendorfer use slightly weaker assumptions)[115]: a pure strategy Nash equilibrium exists, the joint distribution of equilibrium bid vectors $(\mathbf{B}_1, \ldots, \mathbf{B}_n)$ is differentiable almost everywhere in the support of equilibrium bids, and there is zero probability that bidder $i$ uses a bid in equilibrium at which the joint distribution of opponent bids fails to be differentiable.

Given the equilibrium distribution of bid vectors for bidder $i$'s opponents $j \neq i$, let $G^s(\mathbf{b}_i)$ denote the probability that bidder $i$ wins the objects in bundle $s$ when bidder $i$ chooses the bid vector $\mathbf{b}_i$. Note that $G^s(\cdot)$ generally is not a cumulative distribution function and need not even be increasing, since increasing $b_{i,t}$ for a bundle $t$ such that $s \cap t \neq \emptyset$ might lead to a lower probability that $i$ wins all objects in bundle $s$. When there are no reserve prices, bidder $i$ solves the problem

$$\max_{\mathbf{b}_i} \sum_{s \subseteq S} (u_{i,s} - b_{i,s}) G^s(\mathbf{b}_i).$$

If $\mathbf{b}_i$ is the equilibrium bid for bidder $i$ when his type is $\mathbf{u}_i$, then as long as the objective function is differentiable at $\mathbf{b}_i$, the following system of first-order conditions must be satisfied:

$$-G^s(\mathbf{b}_i) + \sum_{t \subseteq S} (u_{i,t} - b_{i,t}) \frac{\partial}{\partial b_{i,s}} G^t(\mathbf{b}_i) = 0 \quad \text{for all } s \subseteq S. \tag{10.2}$$

Let $\mathbf{G}(\mathbf{b}_i)$ denote the $(2^m - 1) \times 1$ vector with components $G^s(\mathbf{b}_i)$, and let $\nabla\mathbf{G}(\mathbf{b}_i)$ be the $(2^m - 1) \times (2^m - 1)$ matrix with $(s, t)$ element $\frac{\partial}{\partial b_{i,s}} G^t(\mathbf{b}_i)$. Then we can rewrite the system of first-order conditions in matrix notation as

$$\nabla\mathbf{G}(\mathbf{b}_i)[\mathbf{u}_i - \mathbf{b}_i] = \mathbf{G}(\mathbf{b}_i).$$

This is a system of linear equations in the vector of valuations $\mathbf{u}_i$. If $\nabla\mathbf{G}(\mathbf{b}_i)$ is invertible, we can rewrite the first-order conditions in a form analogous to the single-unit auction case (2.4):

$$\mathbf{u}_i = \mathbf{b}_i + \left[\nabla\mathbf{G}(\mathbf{b}_i)\right]^{-1} \mathbf{G}(\mathbf{b}_i). \tag{10.3}$$

Invertibility of $\nabla\mathbf{G}(\mathbf{b}_i)$ would then imply that the distribution of (multidimensional) valuations were nonparametrically identified, following the logic developed above for the single-object first-price auction.

[115] They argue that a mixed strategy equilibrium exists, but to our knowledge it is not known what additional assumptions would be required to guarantee that a pure strategy equilibrium exists. Although it might seem that a mixed strategy equilibrium should be inconsistent with identification, that is not necessarily true. In a mixed strategy equilibrium, for at least some valuations, a bidder uses more than one bid vector: the mapping from valuations to bids is one-to-many. Identification of the primitive valuation functions will require that for each bid vector, there is a unique valuation that uses that bid vector; i.e, that the mapping from bids to valuations is many-to-one.

One unresolved question is whether there are useful sufficient conditions on the distribution of bids (or on $\mathbf{G}(\cdot)$) that ensure that observed bidding is consistent with equilibrium behavior (see Section 5.1). First-order conditions are, of course, necessary but not sufficient for equilibrium. In the case of a single-unit first-price auction, Theorem 5.2 ensures that the first-order conditions together with monotonicity of the (inverse) bid function are necessary and sufficient for optimality of each bidder's best response. Thus far there is no analogous result for the multi-object auction considered here. Hence, for an observed bid vector $\mathbf{b}_i$ it is possible that there is a unique $\mathbf{u}_i$ satisfying (10.3), yet for that $\mathbf{u}_i$, $\mathbf{b}_i$ is not a best response to the distribution of $i$'s opponents' bids. However, it should be possible to rule this out in a given application: since the bidder's objective function can be calculated from observables for each vector of valuations, for each observed $\mathbf{b}_i$ and corresponding $\mathbf{u}_i$ satisfying (10.3) it is possible to compute the globally optimal bid vector for $\mathbf{u}_i$ and confirm that it is equal to $\mathbf{b}_i$, thereby verifying that the inverse bid functions implied by (10.3) are mutual best responses.

A second difficulty with using (10.3) arises from the fact $\nabla\mathbf{G}(\cdot)$ will not in general be invertible, since bidders need not make bids on all bundles – not even on all those for which they have positive valuations. Making no bid on a given bundle (or, equivalently, making a bid for this bundle that is sure to lose) can be optimal for a bidder since this ensures that she does not compete with her own bids on other bundles. Given a bid vector $\mathbf{b}_{i,-s}$, Cantillon and Pesendorfer (2003) call a bid $b_{i,s}$ *irrelevant* if

$$b_{i,s} < \inf\bigl\{\tilde{b}_{i,s}\colon\ G^s(\tilde{b}_{i,s}, \mathbf{b}_{i,-s}) > 0\bigr\}.$$

Irrelevant bids are bids that could never win. The problem for identification is that if a bidder places an irrelevant bid on bundle $s$, $\frac{\partial}{\partial b_{i,t}} G^s(b_i) = 0$ and $\frac{\partial}{\partial b_{i,s}} G^t(b_i) = 0$ for all $t \subseteq S$, implying that $\nabla\mathbf{G}(\mathbf{b}_i)$ is not invertible. Indeed, Cantillon and Pesendorfer (2003) establish that $\nabla\mathbf{G}(\mathbf{b}_i)$ is invertible if and only if there are no irrelevant bids. In their application, bidders appear to make many irrelevant bids.[116]

Although irrelevant bids preclude point identification, there is still information in such bids. First observe that if $\mathbf{b}_i$ includes an irrelevant bid for bundle $t$, it is still possible to identify the valuations associated with the bids for other bundles. To see this, note that if for valuation vector $\mathbf{u}_i$ it is optimal to place relevant bids for all bundles in $K \subset 2^S$ and irrelevant bids on other bundles, one obtains the same solution if one treats the bidder's optimization problem as a constrained problem, with the bidder *required* to place irrelevant bids on all bundles $2^S \backslash K$. Formally, let $\mathbf{b}_i^K$ be the subvector of bids on the elements of $K$, and let $G_K^s(\mathbf{b}_i^K)$ denote the probability that bidder $i$ wins bundle $s$ when he places irrelevant bids on bundles $t \in 2^S \backslash K$ and bids $\mathbf{b}_i^K$ on bundles in $K$. Finally, let $\mathbf{G}_K(\cdot)$ denote a vector with elements given by $G_K^s(\cdot)$ for $s \in K$. Then the optimal bid for type $\mathbf{u}_i$ of bidder $i$ in the original game is also the solution to

$$\max_{\mathbf{b}_i^K} \sum_{s\subseteq K} (u_{i,s} - b_{i,s}) G_K^s\bigl(\mathbf{b}_i^K\bigr).$$

116 Irrelevant bids are identified by replacing $\mathbf{G}(\cdot)$ with the empirical analog and directly checking whether each bid has a positive probability of winning.

The solution to this problem will involve no irrelevant bids, so $\nabla \mathbf{G}_K(\mathbf{b}_i^K)$ will be invertible. Hence, the valuations $\mathbf{u}_i^K$ that in equilibrium correspond to bids $\mathbf{b}_i^K$ will be identified.

There is also information in bids about valuations for bundles for which irrelevant bids have been placed. Given a bid vector $\mathbf{b}_{i,-s}$, define the "effective bid"

$$b_{i,s}^{\text{eff}} = \inf\{b_s\colon\ G^s(b_s, \mathbf{b}_{i,-s}) > 0\}.$$

Given continuity of payoffs and the opponent bid distribution, bidder $i$ will always be indifferent between bidding $(b_{i,s}, \mathbf{b}_{i,-s})$, where $b_{i,s}$ is irrelevant, and $(b_{i,s}^{\text{eff}}, \mathbf{b}_{i,-s})$. This implies that increasing $b_{i,s}$ is unprofitable at $b_{i,s} = b_{i,s}^{\text{eff}}$ when $b_{i,s}$ is irrelevant, i.e.,

$$\begin{aligned} &\frac{\partial}{\partial b_{i,s}} G^s(b_{i,s}, \mathbf{b}_{i,-s})\bigg|_{b_{i,s}=b_{i,s}^{\text{eff}}} \left(u_{i,s} - b_{i,s}^{\text{eff}}\right) \\ &\quad + \sum_{t \subseteq S, t \neq s} (u_{i,t} - b_{i,t}) \frac{\partial}{\partial b_{i,s}} G^t(b_{i,s}, \mathbf{b}_{i,-s})\bigg|_{b_{i,s}=b_{i,s}^{\text{eff}}} \leqslant 0 \quad \text{for all } s \subseteq S, \end{aligned} \tag{10.4}$$

where all derivatives are taken from the right. Since $\frac{\partial}{\partial b_{i,s}} G^t(b_{i,s}^{\text{eff}}, \mathbf{b}_{i,-s}) = 0$ (again taking the derivative from the right) for all $t \neq s$ such that $b_{i,t}$ is irrelevant, and since we have just argued that $u_{i,t}$ is identified for all $t$ such that $b_{i,t}$ is relevant, the only remaining unknown in (10.4) is $u_{i,s}$. Thus, (10.4) places an upper bound on the bidder's valuation for bundle $s$. In particular, the true $u_{i,s}$ must be less than the value of $u_{i,s}$ that makes (10.4) hold with equality. This can be used to provide a lower bound on the cumulative distribution of $U_{i,s}$. More generally, a lower bound on the distribution of $\mathbf{U}_i$ is identified using (10.4).

In Cantillon and Pesendorfer's application, two additional constraints are imposed on bids. First, there are reserve prices, denoted $r_s$; bids below the reserve price win with probability zero. Second, the auction rules specify that

$$b_{i,s\cup t} \geqslant b_{i,s} + b_{i,t} \quad \text{for all } s, t \subseteq S \text{ such that } s \cap t = \emptyset. \tag{10.5}$$

This rule is motivated by the idea that if this constraint were violated, the auctioneer could choose to ignore the bid $b_{i,s\cup t}$, and instead accept the bids $b_{i,s}$ and $b_{i,t}$. Thus, bidders can express preferences for complements, but their bids cannot be less for a combination than for the component parts. Cantillon and Pesendorfer (2003) extend the analysis to incorporate these constraints, showing that even in their presence, it is possible to place an upper bound on the extent of the synergies that exist between items.

## 11. Concluding remarks

The prominent role of auctions in allocating a wide range of public and private resources provides one strong motivation for empirical work on auctions. Recent methodological advances have made it possible to address old market design questions (e.g., how to auction Treasury bills), while new policy questions (e.g., how to auction multiple complementary goods) have motivated development of new methodological tools. In addition, auctions hold the promise of shedding light on fundamental questions about the nature of information, preferences, and behavior that are of importance to a much broader scope of economic environments. Like earlier descriptive empirical work on auctions that provided influential evidence on the importance of asymmetric information and strategic behavior, recent empirical work using structural econometric models has also begun to deliver on this promise, addressing such questions as the empirical importance of reputations, entry costs, or adverse selection. Because of the close match between the theory and actual institutions, auctions have the potential to provide insights into fundamental questions that are difficult or impossible to address without the benefit of structure from economic theory. We expect much of the most interesting future work in the empirical auction literature to push farther in this direction.

It is worth noting that the analysis of identification in auction models is useful outside of the realm of econometrics. For example, in some models of learning in games, a central component of the analysis concerns whether it is possible to infer primitives of the game from the distribution of equilibrium outcomes that can be observed by players. The equilibrium concept of self-confirming equilibrium [Fudenberg and Levine (1993), Dekel, Fudenberg and Levine (2003)], motivated by learning models, hinges on just this issue.[117] Recently, Esponda (2004) analyzed self-confirming equilibria in auction games, focusing on the extent to which information revealed by an auctioneer allows bidders to infer the distribution over opponent types. This problem is closely related to the identification problem.[118]

Auctions have long been recognized as providing ideal market institutions for exploring the relationships between economic theory and the actual behavior of economic agents. Since the seminal work of Vickrey (1961) and Wilson (1967), rich theoretical and empirical literatures on auctions have developed. In our view, one of the most exciting advances in this literature is the development of methods for combining theoretical and statistical analysis in order to learn about the primitive features of an auction environment from observed bidding behavior. We have focused our discussion on nonparametric identification, in part because this makes transparent how the relationships derived from theory can be used to make valid inferences from data. We hope that this chapter will be a valuable reference and starting point for researchers who will apply and expand upon these methods to explore the wide range of open questions in the future.

[117] This concept relaxes the common knowledge assumption of Nash equilibrium, but requires that bidders best-respond to beliefs that are consistent with the equilibrium distribution of outcomes that is observable to the bidders. For example, the bidders might observe the distribution of transactions prices, or the distribution of all bids.

[118] Furthermore, this alternative to the standard common knowledge assumption may be an interesting possibility to explore in an empirical model.

## References

Ackerberg, D., Benkard, C.L., Berry, S., Pakes, A. (2007). "Econometric tools for analyzing market outcomes". In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6A. Elsevier (Chapter 63).

Andersen, P.K., Borgan, Ø., Gill, R., Keiding, N. (1991). Statistical Models Based on Counting Processes. Springer, New York.

Aoyagi, M. (2003). "Bid rotation and collusion in repeated auctions". Journal of Economic Theory 112, 79–105.

Armantier, O., Sbaï, E. (2003). "Estimation and comparison of treasury auction formats when bidders are asymmetric". Working Paper. SUNY–Stony Brook.

Armstrong, M., Rochet, J.C. (1999). "Multidimensional screening: A user's guide". European Economic Review 43, 959–979.

Arnold, B., Balakrishnan, N., Nagaraja, H. (1992). A First Course in Order Statistics. Wiley & Sons, New York.

Athey, S. (2001). "Single crossing properties and the existence of pure strategy equilibrium in games of incomplete information". Econometrica 69, 861–890.

Athey, S., Bagwell, K. (2001). "Optimal collusion with private information". RAND Journal of Economics 32, 428–465.

Athey, S., Bagwell, K. (2004a). "Collusion with persistent cost shocks". Working Paper. Stanford University.

Athey, S., Bagwell, K. (2004b). "Dynamic auctions with persistent cost shocks". Working Paper. Stanford University.

Athey, S., Haile, P. (2000). "Identification of standard auction models". MIT Working Paper 00-18.

Athey, S., Haile, P. (2002). "Identification of standard auction models". Econometrica 70, 2107–2140.

Athey, S., Levin, J. (2001). "Information and competition in US Forest Service timber auctions". Journal of Political Economy 109, 375–417.

Athey, S., Bagwell, K., Sanchirico, C. (2004). "Collusion and price rigidity". The Review of Economic Studies 71, 317–349.

Athey, S., Levin, J., Seira, E. (2004). "Comparing open and sealed bid auctions: Theory and evidence from timber auctions". Working Paper. Stanford University.

Ausubel, L., Cramton, P. (2002). "Demand reduction and inefficiency in multi-unit auctions". Working Paper 96–07. University of Maryland.

Avery, C. (1998). "Strategic jump bidding in English auctions". Review of Economic Studies 65, 185–210.

Back, K., Zender, J. (1993). "Auctions of divisible goods: On the rationale for the US treasury experiment". Review of Financial Studies 6, 733–764.

Bajari, P. (1997). "The first-price auction with asymmetric bidders: Theory and applications". PhD Dissertation. University of Minnesota.

Bajari, P. (2001). "Comparing competition and collusion: A numerical approach". Economic Theory 18, 187–205.

Bajari, P., Hortaçsu, A. (2003a). "The winner's curse, reserve prices, and endogenous entry: Empirical insights from eBay auctions". RAND Journal of Economics 34, 329–355.

Bajari, P., Hortaçsu, A. (2003b). "Cyberspace auctions and pricing issues: A review of empirical findings". In: Jones, D. (Ed.), New Economy Handbook. Elsevier.

Bajari, P., Hortaçsu, A. (2004). "Economic insights from Internet auctions". Journal of Economic Literature 17, 457–486.

Bajari, P., Hortaçsu, A. (2005). "Are structural estimates of auction models reasonable? Evidence from experimental data". Journal of Political Economy 113, 703–741.

Bajari, P., Summers, G. (2002). "Detecting collusion in procurement auctions". Antitrust Law Journal 70, 143–170.

Bajari, P., Ye, L. (2003). "Deciding between competition and collusion". Review of Economics and Statistics 85, 971–989.

Bajari, P., Houghton, S., Tadelis, S. (2004). "Bidding for incomplete contracts". Working Paper. Duke University.

Balasubramanian, K., Balakrishnan, N. (1994). "Equivalence of relations for order statistics for exchangeable and arbitrary cases". Statistics and Probability Letters 21, 405–407.

Baldwin, L., Marshall, R., Richard, J. (1997). "Bidder collusion in US Forest Service timber sales". Journal of Political Economy 105, 657–699.

Bartholomew, D. (1959). "A test of homogeneity for ordered alternatives". Biometrika 46, 36–48.

Berman, S. (1963). "Note on extreme values, competing risks, and semi-Markov processes". Annals of Mathematical Statistics 34, 1104–1106.

Berry, S., Reiss, P. (in press). "Empirical models of entry and market structure". In: Armstrong, M., Porter, R. (Eds.). Handbook of Industrial Organization, vol. III. Elsevier.

Berry, S., Tamer, E. (2005). "Identification in models of oligopoly entry". Working Paper. Yale University.

Berry, S., Levinsohn, J., Pakes, A. (1995). "Automobile prices in market equilibrium". Econometrica 63, 841–890.

Bikhchandani, S. (1988). "Reputation in repeated second-price auctions". Journal of Economic Theory 46, 97–119.

Bikhchandani, S., Huang, C. (1989). "Auctions with resale markets: An exploratory model of treasury bill markets". Review of Financial Studies 2, 311–339.

Bikhchandani, S., Haile, P., Riley, J. (2002). "Symmetric separating equilibria in English auctions". Games and Economic Behavior 38, 19–27.

Borenstein, S., Bushnell, J., Wolak, F. (2002). "Measuring market inefficiencies in California's restructured wholesale electricity market". American Economic Review 92, 1376–1405.

Bowman, A., Jones, M., Gijbels, I. (1998). "Testing monotonicity of a regression". Journal of Computational and Graphical Statistics 7, 489–500.

Brannman, L., Klein, D., Weiss, L. (1987). "The price effects of increased competition in auction markets". Review of Economics and Statistics 69, 24–32.

Brendstrup, B., Paarsch, H. (2003). "Nonparametric estimation of Dutch and first-price, sealed-bid auction models with asymmetric bidders". Working Paper. University of Iowa.

Brendstrup, B., Paarsch, H. (2004). "Nonparametric identification and estimation of multi-unit, sequential, oral ascending-price auctions with asymmetric bidders". Working Paper. University of Iowa.

Campo, S. (2002). "Asymmetry and risk aversion within the independent private values paradigm: The case of construction procurement contracts". Working Paper. University of Southern California.

Campo, S., Guerre, E., Perrigne, I., Vuong, Q. (2002). "Semiparametric estimation of first-price auctions with risk averse bidders". Working Paper. Pennsylvania State University.

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

Cantillon, E., Pesendorfer, M. (2003). "Combination bidding in multi-unit auctions". Working Paper. London School of Economics and Political Science.

Chen, X. (2007). "Large sample sieve estimation of semi-nonparametric models". In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6B. Elsevier (Chapter 76).

Chernozhukov, V., Hong, H. (2003). "Likelihood inference in a class of nonregular econometric models". Econometrica 72, 1445–1480.

Chow, Y., Teicher, H. (1997). Probability Theory: Independence, Interchangeability and Martingales. Springer, New York.

Crémer, J., McLean, R. (1988). "Full extraction of the surplus in Bayesian and dominant strategy auctions". Econometrica 56, 1247–1257.

David, H. (1981). Order Statistics. Wiley, New York.

David, H., Moeschberger, M. (1978). The Theory of Competing Risks. Macmillan, New York.

Das Varma, G. (2003). "Bidding for a process innovation under alternative modes of competition". International Journal of Industrial Organization.

Dekel, E., Fudenberg, D., Levine, D. (2003). "Learning to play Bayesian games". Games and Economic Behavior 46, 282–303.

Donald, S., Paarsch, H. (1993). "Piecewise pseudo-maximum likelihood estimation in empirical models of auctions". International Economic Review 34, 121–148.

Donald, S., Paarsch, H. (1996). "Identification, estimation, and testing in parametric empirical models of auctions within the independent private values paradigm". Econometric Theory 12, 517–567.

Donald, S., Paarsch, H., Robert, J. (2006). "An empirical model of the multi-unit, sequential clock auction". Journal of Applied Econometrics 21, 1221–1247.

Einav, L. (2004). "A note on the analogies between empirical models of auctions and of differentiated product markets". Working Paper. Stanford University.

Esponda, I. (2004). "Information feedback and self-confirming equilibrium in first price auctions". Working Paper. Stanford University.

Fermanian, J. (2003). "Nonparametric estimation of competing risks models with covariates". Journal of Multivariate Analysis 85, 156–191.

Février, P. (2004). "Semiparametric identification and estimation of common value auctions". Working Paper. CREST.

Février, P., Préget, R., Visser, M. (2002). "Econometrics of share auctions". Working Paper. CREST.

Flambard, V., Perrigne, I. (2006). "Asymmetry in procurement auctions: Some evidence from snow removal contracts". Economic Journal, Royal Economic Society 116 (514), 1014–1036.

Friedman, M. (1960). A Program for Monetary Stability. Fordham University Press, New York.

Fudenberg, D., Levine, D. (1993). "Self-confirming equilibrium". Econometrica 61, 523–546.

Fudenberg, D., Tirole, J. (1991). Game Theory. MIT Press, Cambridge.

Gallant, R., Nychka, D. (1987). "Semi-nonparametric maximum likelihood estimation". Econometrica 55, 363–390.

Gijbels, I., Hall, P., Jones, M.C., Koch, I. (2000). "Tests for monotonicity of a regression mean with guaranteed level". Biometrika 87, 663–673.

Gilley, O., Karels, G. (1981). "The competitive effect in bonus bidding: New evidence". Bell Journal of Economics 12, 637–648.

Goeree, J. (2003). "Bidding for the future: Signaling in auctions with an aftermarket". Journal of Economic Theory 108, 345–364.

Guerre, E., Perrigne, I., Vuong, Q. (1995). "Nonparametric estimation of first-price auctions". Working Paper #9504. University of Southern California.

Guerre, E., Perrigne, I., Vuong, Q. (2000). "Optimal nonparametric estimation of first-price auctions". Econometrica 68, 525–574.

Haile, P. (1999). "Auctions with resale". Working Paper. University of Wisconsin.

Haile, P. (2001). "Auctions with resale markets: An application to US Forest Service timber sales". American Economic Review 91, 399–427.

Haile, P. (2003). "Auctions with private uncertainty and resale opportunities". Journal of Economic Theory 108, 72–110.

Haile, P., Tamer, E. (2001). "Inference from English auctions with asymmetric affiliated private values". Working Paper. University of Wisconsin-Madison.

Haile, P., Tamer, E. (2002). "Inference with an incomplete model of English auctions". Working Paper. Princeton University.

Haile, P., Tamer, E. (2003). "Inference with an incomplete model of English auctions". Journal of Political Economy 111, 1–52.

Haile, P., Hong, H., Shum, M. (2003). "Nonparametric tests for common values in first-price sealed-bid auctions". NBER Working Paper 10105.

Hall, P., Heckman, N. (2000). "Testing for monotonicity of a regression mean by calibrating for linear functions". Annals of Statistics 28, 20–39.

Harstad, R., Rothkopf, M. (2000). "An 'Alternating Recognition' model of English auctions". Management Science 46, 1–12.

Heckman, J., Honoré, B. (1989). "The identifiability of the competing risks model". Biometrika 76, 325–330.

Heckman, J., Honoré, B. (1990). "The empirical content of the Roy model". Econometrica 58, 1121–1149.

Hendricks, K., Paarsch, H. (1995). "A survey of recent empirical work concerning auctions". Canadian Journal of Economics 28, 403–426.

Hendricks, K., Porter, R. (1988). "An empirical study of an auction with asymmetric information". American Economic Review 78, 865–883.

Hendricks, K., Porter, R. (in press). "Lectures on auctions: An empirical perspective". In: Armstrong, M., Porter, R. (Eds.). Handbook of Industrial Organization, vol. III. Elsevier.

Hendricks, K., Porter, R., Boudreau, B. (1987). "Information, returns, and bidding behavior in OCS auctions: 1954–1969". Journal of Industrial Economics 35, 517–542.

Hendricks, K., Porter, R., Wilson, C. (1994). "Auctions for oil and gas leases with an informed bidder and a random reservation price". Econometrica 62, 1415–1444.

Hendricks, K., Pinkse, J., Porter, R. (2003). "Empirical implications of equilibrium bidding in first-price, symmetric, common value auctions". Review of Economic Studies 70, 115–145.

Hirano, K., Porter, J. (2003). "Asymptotic efficiency in parametric structural models with parameter-dependent support". Econometrica 71, 1307–1338.

Hollander, M., Wolfe, D. (1999). Nonparametric Statistical Methods. John Wiley and Sons, New York.

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

Hong, H., Shum, M. (2000). "Structural estimation of auction models". In: Patrone, F., Garcia-Jurado, I., Tijs, S. (Eds.), Game Practice: Contributions from Applied Game Theory. Kluwer, Boston.

Hong, H., Shum, M. (2002). "Increasing competition and the winner's curse: Evidence from procurement". Review of Economic Studies 69, 871–898.

Hong, H., Shum, M. (2003). "Econometric models of ascending auctions". Journal of Econometrics 112, 327–358.

Hortaçsu, A. (2002). "Mechanism choice and strategic bidding in divisible good auctions: An empirical analysis of the Turkish treasury auction market". Working Paper. University of Chicago.

Hotz, J., Miller, R. (1993). "Conditional choice probabilities and the estimation of dynamic models". Review of Economic Studies 60, 497–529.

Izmalkov, S. (2003). "English auctions with reentry". Working Paper. MIT.

Jackson, M., Swinkels, J. (2005). "Existence of equilibria in single and double private value auctions". Econometrica 73 (1), 93–140.

Jackson, M., Simon, L., Swinkels, J., Zame, W. (2002). "Equilibrium, communication, and endogenous sharing rules in discontinuous games of incomplete information". Econometrica 70, 1711–1740.

Jofre-Bonet, M., Pesendorfer, M. (2000). "Bidding behavior in repeated procurement auctions". European Economic Review 44, 1006–1020.

Jofre-Bonet, M., Pesendorfer, M. (2003). "Estimation of a dynamic auction game". Econometrica 71, 1443–1489.

Judd, K. (1998). Numerical Methods in Economics. MIT Press, Cambridge.

Kagel, J. (1995). "Auctions: A survey of experimental research". In: Kagel, J., Roth, A. (Eds.), The Handbook of Experimental Economics. Princeton University Press, Princeton, pp. 501–585.

Kastl, J. (2005). "Discrete bids and empirical inference in divisible good auctions. Working Paper. Northwestern University.

Katzman, B., Rhodes-Kropf, M. (2002). "The consequences of information revealed in auctions". Working Paper. Columbia University.

Klemperer, P. (1999). "Auction theory: A guide to the literature". Journal of Economic Surveys 13, 227–286.

Klemperer, P. (2002). "What really matters in auction design". Journal of Economic Perspectives 16, 169–189.

Koopmans, T. (1945). "Statistical estimation of simultaneous economic relations". Journal of the American Statistical Association 40, 448–466.

Kotlarski, I. (1966). "On some characterization of probability distributions in Hilbert spaces". Annali di Matematica Pura ed Applicata 74, 129–134.

Krasnokutskaya, E. (2004). "Auction models with unobserved heterogeneity: Application to the Michigan highway procurement auctions". Working Paper. University of Pennsylvania.

Krishna, V. (2002). Auction Theory. Academic Press, San Diego.

Laffont, J. (1997). "Game theory and empirical economics: The case of auction data". European Economic Review 41, 1–35.

Laffont, J., Vuong, Q. (1993). "Structural econometric analysis of descending auctions". European Economic Review 37, 329–341.
Laffont, J., Vuong, Q. (1996). "Structural analysis of auction data". American Economic Review, Papers and Proceedings 86, 414–420.
Laffont, J., Ossard, H., Vuong, Q. (1995). "Econometrics of first-price auctions". Econometrica 63, 953–980.
Lancaster, T. (1990). The Econometric Analysis of Transition Data: An Econometric Society Monograph. Cambridge University Press.
Lebrun, B. (1999). "First price auctions in the asymmetric N bidder case". International Economic Review 40, 125–142.
Levin, D., Smith, J. (1994). "Equilibrium in auctions with entry". American Economic Review 84, 585–599.
Li, T. (2002). "Robust and consistent estimation of nonlinear errors-in-variables models". Journal of Econometrics 110, 1–26.
Li, T. (2003). "Econometrics of first-price auctions with entry and binding reservation prices". Working Paper. Indiana University.
Li, T., Vuong, Q. (1998). "Nonparametric estimation of the measurement error model using multiple indicators". Journal of Multivariate Analysis 65, 139–165.
Li, T., Zheng, X. (2005). "Procurement auctions with entry and an uncertain number of actual bidders: Theory, structural inference, and an application". Working Paper. Indiana University.
Li, T., Perrigne, I., Vuong, Q. (2000). "Conditionally independent private information in OCS wildcat auctions". Journal of Econometrics 98, 129–161.
Li, T., Perrigne, I., Vuong, Q. (2002). "Structural estimation of the affiliated private value auction model". RAND Journal of Economics 33, 171–193.
Lizzeri, A., Persico, N. (2000). "Uniqueness and existence of equilibrium in auctions with a reserve price". Games and Economic Behavior 30, 83–114.
Loader, C. (1999). Local Regression and Likelihood. Springer, New York.
Lucking-Reiley, D. (2000). "Auctions on the Internet: What's being auctioned, and how?". Journal of Industrial Economics 48, 227–252.
Manski, C. (1995). Identification Problems in the Social Sciences. Harvard University Press.
Manski, C., Tamer, E. (2002). "Inference on regressions with interval data on a regressor or outcome". Econometrica 70, 519–546.
Maskin, E., Riley, J. (1984). "Optimal auctions with risk averse buyers". Econometrica 52, 1473–1518.
Maskin, E., Riley, J. (2000a). "Asymmetric auctions". Review of Economic Studies 67, 413–438.
Maskin, E., Riley, J. (2000b). "Equilibrium in sealed high bid auctions". Review of Economic Studies 67, 439–454.
Maskin, E., Riley, J. (2003). "Uniqueness of equilibrium in sealed high-bid auctions". Games and Economic Behavior 45, 395–409.
Matthews, S. (1984). "Information acquisition in discriminatory auctions". In: Boyer, M., Kihlstrom, R.E. (Eds.), Bayesian Models in Economic Theory. North-Holland, New York.
McAdams, D. (2004a). "Monotone equilibrium in multi-unit auctions". Working Paper. MIT.
McAdams, D. (2004b). "Uniqueness in first-price auctions with affiliation". Working Paper. MIT.
McAdams, D. (2005). "Identification and testable restrictions in private value multi-unit auctions". Working Paper. MIT.
McAdams, D. (2007). "Monotonicity in asymmetric first-price auctions with affiliation". International Journal of Game Theory 35, 427–453.
McAfee, R.P., McMillan, J. (1987). "Auctions and bidding". Journal of Economic Literature 25, 669–738.
McAfee, R.P., McMillan, J. (1992). "Bidding rings". American Economic Review 82, 579–599.
McAfee, R.P., McMillan, J. (1996). "Analyzing the airwaves auction". Journal of Economic Perspectives 10, 159–175.
McAfee, R.P., Reny, P. (1992). "Correlated information and mechanism design". Econometrica 60, 395–421.
McAfee, R.P., McMillan, J., Whinston, M. (1989). "Multiproduct monopoly, commodity bundling and correlation of values". Quarterly Journal of Economics 102, 371–383.

McAfee, R.P., Quan, D., Vincent, D. (2002). "Minimum acceptable bids, with application to real estate auctions". Journal of Industrial Economics 50, 391–416.
McAfee, R.P., Takacs, W., Vincent, D.R. (1999). "Tariffying auctions". RAND Journal of Economics 30, 158–179.
McFadden, D. (1989). "Testing for stochastic dominance". In: Studies in the Economics of Uncertainty: In Honor of Josef Hadar. Springer, New York.
Meilijson, I. (1981). "Estimation of the lifetime distribution of the parts from the autopsy statistics of the machine". Journal of Applied Probability 18, 829–838.
Milgrom, P. (1981). "Good news and bad news: Representation theorems and applications". Bell Journal of Economics 12, 380–391.
Milgrom, P. (2004). Putting Auction Theory to Work. Cambridge University Press, Cambridge.
Milgrom, P.R., Weber, R.J. (1982). "A theory of auctions and competitive bidding". Econometrica 50, 1089–1122.
Myerson, R. (1981). "Optimal auction design". Mathematics of Operations Research 6, 58–73.
Ockenfels, A., Roth, A.E. (2006). "Late and multiple bidding in second-price Internet auctions: Theory and evidence concerning different rules for ending an auction". Games and Economic Behavior 55 (2), 297–320.
Olley, G.S., Pakes, A. (1996). "The dynamics of productivity in the telecommunications equipment industry". Econometrica 64, 1263–1297.
Paarsch, H. (1992a). "Deciding between common and private values paradigms in empirical models of auctions". Journal of Econometrics 51, 191–215.
Paarsch, H. (1992b). "Empirical models of auctions and an application to British Columbian timber sales". Research Report 9212. University of Western Ontario.
Paarsch, H. (1994). "A comparison of estimators for empirical models of auctions". Annales d'Economie et de Statistique 34, 143–157.
Paarsch, H. (1997). "Deriving an estimate of the optimal reserve price: An application to British Columbian timber sales". Journal of Econometrics 78, 333–357.
Perrigne, I. (2003). "Random reserve prices and risk aversion in timber sale auctions". Working Paper. Pennsylvania State University.
Perrigne, I., Vuong, Q. (1999). "Structural econometrics of first-price auctions: A survey of methods". Canadian Journal of Agricultural Economics 47, 203–223.
Pesendorfer, M. (2000). "A study of collusion in first-price auctions". Review of Economic Studies 67, 381–411.
Peters, M., Severinov, S. (2006). "Internet auctions with many traders". Journal of Economic Theory 130, 220–245.
Pinkse, J., Tan, G. (2005). "The affiliation effect in first-price auctions". Econometrica 73, 263–277.
Politis, D., Romano, J., Wolf, M. (1999). Subsampling. Springer-Verlag, New York.
Porter, R., Zona, J.D. (1993). "Detection of bid rigging in procurement auctions". Journal of Political Economy 101, 518–538.
Porter, R., Zona, J.D. (1999). "Ohio school milk markets: An analysis of bidding". RAND Journal of Economics 30, 263–288.
Prakasa-Rao, B.L.S. (1992). Identifiability in Stochastic Models: Characterization of Probability Distributions. Academic Press, San Diego.
Quint, D. (2004). "Optimal second price auctions with positively correlated private values and limited information". SIEPR Discussion Paper 03-14. Stanford University.
Reiss, P., Wolak, F. (2007). "Structural econometric modeling: Rationales and examples from industrial organization". In: Heckman, J.J., Leamer, E. (Eds.), Handbook of Econometrics, vol. 6A. Elsevier (Chapter 64).
Reny, P., Zamir, S. (2004). "On the existence of pure strategy monotone equilibria in asymmetric first-price auctions". Econometrica 72, 1105–1125.
Romano, J. (1988). "A bootstrap revival of some nonparametric distance tests". Journal of the American Statistical Association 83, 698–708.

Romano, J. (1989). "Bootstrap and randomization tests of some nonparametric hypotheses". Annals of Statistics 17, 141–159.

Samuelson, W. (1985). "Competitive bidding with entry costs". Economics Letters 17, 53–57.

Schennach, S. (2004). "Estimation of nonlinear models with measurement error". Econometrica 72, 33–75.

Shneyerov, A. (2005). "An empirical study of auction revenue rankings: The case of municipal bonds". Working Paper. University of British Columbia.

Silverman, B. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall, London.

Skryzpacz, A., Hopenhayn, H. (2004). "Tacit collusion in repeated auctions". Journal of Economic Theory 114, 153–169.

Smiley, A. (1979). Competitive Bidding Under Uncertainty: The Case of Offshore Oil. Ballinger, Cambridge.

Song, U. (2003). "Nonparametric estimation of an eBay auction model with an unknown number of bidders". Working Paper. University of Wisconsin.

Song, U. (2004). "Structural analysis of auction data with an unknown number of bidders". PhD Dissertation. University of Wisconsin.

Su, L., White, H. (2003). "Testing conditional independence via empirical likelihood". Working Paper. University of California San Diego.

Thiel, S. (1988). "Some evidence on the winner's curse". American Economic Review 78, 884–895.

Wang, M., Jewell, N.P., Tsai, W. (1986). "Asymptotic properties of the product limit estimate under random truncation". Annals of Statistics 14, 1597–1605.

Whinston, M. (1989). "Tying, foreclosure, and exclusion". American Economic Review 90, 837–859.

Wilson, R. (1967). "Competitive bidding with asymmetric information". Management Science 11, 816–820.

Wilson, R. (1979). "Auctions of shares". Quarterly Journal of Economics 93, 675–689.

Wolak, F. (2003). "Identification and estimation of cost functions using observed bid data: An application to electricity markets". In: Dewatripont, M., Hansen, L., Turnovsky, P. (Eds.), Advances in Economics and Econometrics – Theory and Applications. Eighth World Congress. In: Econometric Society Monographs, vol. 2. Cambridge University Press, Cambridge, pp. 115–149.

Wolak, F. (2004). "Quantifying the supply-side benefits from forward contracting in wholesale electricity markets. Working Paper. Stanford University.

Wolfram, C. (1998). "Strategic bidding in a multi-unit auction, an empirical analysis of bids to supply electricity in England and Wales". RAND Journal of Economics 29, 703–725.

Woodroofe, M. (1985). "Estimating a distribution function with truncated data". Annals of Statistics 13, 163–177.

Vickrey, W. (1961). "Counterspeculation, auctions, and competitive sealed tenders". Journal of Finance 16, 8–37.

Yin, P. (2004). "eBay auctions as markets". Working Paper. Harvard Business School.

Zulehner, C. (2003). "Bidding behavior and bidders' valuations in Austrian cattle auctions". Working Paper. University of Vienna.