

# Common Values, Unobserved Heterogeneity, and Endogenous Entry in US Offshore Oil Lease Auctions

---

Giovanni Compiani

*University of Chicago*

Philip Haile

*Yale University*

Marcelo Sant'Anna

*EPGE Brazilian School of Economics and Finance*

Although an auction of drilling rights is often cited as an example of common values, formal evidence has been limited by the problem of auction-level unobserved heterogeneity. We develop an empirical approach for first-price sealed-bid auctions with affiliated values, unobserved heterogeneity, and endogenous bidder entry. We show that important features of the model are nonparametrically identified and apply a semiparametric estimation approach to data from US offshore oil and gas lease auctions. We find that common values, affiliated private information, and unobserved heterogeneity are all present. Failing to account for unobserved heterogeneity obscures the evidence

We are grateful to Ken Hendricks and Rob Porter, who generously provided the data, answered our many questions, and gave valuable feedback on early drafts. We also benefited from the comments and suggestions of Ali Hortaçsu, four anonymous referees, Estelle Cantillon, Xiaohong Chen, and numerous seminar and conference participants. Financial support is gratefully acknowledged from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior–Brasil (CAPES), grant 001. Data are provided as supplementary material online.

Electronically published August 28, 2020

[*Journal of Political Economy*, 2020, vol. 128, no. 10]

© 2020 by The University of Chicago. All rights reserved. 0022-3808/2020/12810-0005\$10.00

of common values. We examine implications of our estimates for the interaction between affiliation, the winner's curse, the auction rules, and the number of bidders in determining the aggressiveness of bidding and seller revenue.

## I. Introduction

In many auction settings, it seems likely that important information commonly known among bidders is unobserved by the econometrician. Ignoring such unobserved heterogeneity can lead to a variety of errors. One may infer too much within-auction correlation in bidders' private information, as well as too much cross-auction variation in this information, leading to incorrect conclusions about such issues as bidder market power, the division of surplus, and optimal auction design.<sup>1</sup> In a first-price auction, unobserved heterogeneity presents a particular challenge because standard identification approaches exploit the insight that bidders' equilibrium beliefs about the competition can be inferred from observed distributions of bids;<sup>2</sup> with unobserved auction-level heterogeneity, bidders' beliefs condition on information unavailable to the econometrician. A further problem is that auction-level unobservables are likely to affect not only bids but also bidder participation. Such endogenous bidder entry threatens several identification and testing approaches relying on exogenous variation in the level of competition.<sup>3</sup>

Here we propose an empirical approach for first-price sealed-bid auctions with affiliated values, unobserved heterogeneity, and endogenous bidder entry. We show nonparametric identification of key structural features and propose a semiparametric estimation approach. We apply the approach to auctions of offshore oil and gas leases in the US Outer Continental Shelf (OCS) in order to evaluate important features of these auctions, including the presence of common values.

Although the term "common values" (or "interdependent values") is often associated with auctions, it refers to a classic form of adverse selection that can arise in a broad range of environments where some parties to a potential transaction have private information relevant to others' assessments of their own valuations or costs (e.g., Akerlof 1970; Arrow 1970; Vincent 1989; Maskin and Tirole 1992; Deneckere and Liang 2006). Testing

<sup>1</sup> See, e.g., Athey, Levin, and Seira (2011), Krasnokutskaya (2011), Krasnokutskaya and Seim (2011), and Roberts (2013).

<sup>2</sup> See, e.g., Laffont and Vuong (1993), Guerre, Perrigne, and Vuong (2000), Li, Perrigne, and Vuong (2002), Hendricks, Pinkse, and Porter (2003), and Athey and Haile (2006, 2007).

<sup>3</sup> See, e.g., Gilley and Karels (1981), Athey and Haile (2002), Haile, Hong, and Shum (2003), Guerre, Perrigne, and Vuong (2009), Gillen (2010), and Campo et al. (2011).

hypotheses about the nature of private information necessarily relies on an indirect approach, and auctions often offer settings in which the maintained assumptions relied upon can be tightly linked to actual market institutions.

But while an auction of drilling rights is frequently cited as an example of a common values environment, formal testing for common values has been hindered by the confounding effects of unobserved heterogeneity. Indeed, we reject private values in favor of common values only when accounting for unobserved heterogeneity and endogenous bidder entry. More broadly, we find that affiliated private information, common values, and common-knowledge unobservables—three distinct phenomena with different implications for policy and empirical work<sup>4</sup>—are all present in OCS auctions. We also use our results to quantify Milgrom and Weber's (1982) classic revenue ranking of first- and second-price sealed-bid auctions and to examine the interaction between affiliation, the winner's curse, and the number of bidders in determining the aggressiveness of bidding and seller revenue.

Prior work on testing for common values in auctions includes Paarsch (1992), Athey and Haile (2002), Bajari and Hortaçsu (2003), Haile, Hong, and Shum (2003), Hortaçsu and Kastl (2012), and Hill and Shneyerov (2014). Most of this work, like ours, exploits the fact that in a common values auction the winner's curse becomes more severe as the number of competitors grows (all else equal). Our testing approach is most similar to that of Haile, Hong, and Shum (2003), discussed further below, who studied timber auctions. Our generalizations of their model relax their most restrictive assumptions but make identification substantially more challenging and require a different estimation approach.

Our empirical study of OCS auctions is related to that of Hendricks, Pinkse, and Porter (2003), who focused on testable implications of a pure common values model.<sup>5</sup> Our work is complementary to theirs. We allow the pure common values model but do not assume it, and we neither exploit nor rely on estimates of realized tract values. Hendricks, Pinkse, and Porter (2003) point out that testing for common values would be difficult

<sup>4</sup> For example, affiliation leads to the "linkage principle" (Milgrom and Weber 1982; Milgrom 1987), whereas common values leads to the "winner's curse," each with potentially important implications for auction design. Unobserved heterogeneity, which is held fixed in auction theory, implies neither affiliation nor common values but creates challenges for identification.

<sup>5</sup> Earlier work on OCS auctions includes Gilley and Karels (1981), Hendricks, Porter, and Boudreau (1987), Hendricks and Porter (1988), and Hendricks, Porter, and Spady (1989), as well as that of Li, Perrigne, and Vuong (2000), who called for formal testing for common values.

because of the likely correlation between bidder entry and auction-level unobservables. Our study also complements Haile, Hendricks, and Porter (2010) and the simultaneous work of Aradillas-López et al. (2019), which focus on implications of competitive (vs. collusive) bidding in OCS auctions after the introduction of “area-wide leasing” in 1983. Here we consider only the period 1954–83, where the evidence in Haile, Hendricks, and Porter (2010) supports the assumption of competitive bidding, and we focus on methodological and substantive issues driven by the nature of bidders’ information—both shared and private.

Haile and Kitamura (2019) review existing econometric approaches to first-price auctions with unobserved heterogeneity. All require compromises of some form. Several (e.g., Krasnokutskaya 2011; Hu, McAdams, and Shum 2013; D’Haultfoeuille and Février 2015) require that bidders have independent types, enabling all correlation among bids to be attributed to unobserved heterogeneity.<sup>6</sup> Krasnokutskaya and Seim (2011) and Gentry and Li (2014) have extended these methods to models with endogenous entry. Other approaches to unobserved heterogeneity use a control-function strategy requiring a one-to-one mapping between the unobserved heterogeneity and an observed auxiliary outcome (often, the number of bidders), allowing one to indirectly condition on the unobservable (e.g., Campo, Perrigne, and Vuong 2003; Haile, Hong, and Shum 2003; Guerre, Perrigne, and Vuong 2009; Roberts 2013). Simultaneous work by Kitamura and Laage (2018) proposes a finite mixture approach allowing affiliated types but requiring that the unobservable be discrete (cf. Haile, Hong, and Shum 2003 and Hu, McAdams, and Shum 2013) and enter through a separable structure similar to that in Krasnokutskaya (2011). Finally, while control-function approaches can provide a strategy for isolating exogenous variation in bidder entry, others generally do not.

Our approach requires compromises as well. We rely on an index assumption similar to that in Krasnokutskaya (2011) and Kitamura and Laage (2018). Like Haile, Hong, and Shum (2003), we require an instrument for entry and a reduced form for the entry outcome in which the auction unobservable is the only latent factor. This rules out stochastic bidder participation, including models of selective entry. Our use of a reduced form also implies that additional structure would be needed to evaluate interventions that would alter the map from auction characteristics to entry outcomes.

<sup>6</sup> See also the partial-identification results in Armstrong (2013). An exception among approaches building on the measurement error literature is Balat (2017). His extension of Hu, McAdams, and Shum (2013) exploits observation of potential bidders’ entry decisions at two sequential stages.

But our approach also offers advantages. It avoids the requirement of independent bidder types and provides a strategy for exploiting exogenous sources of variation in bidder entry. This combination of features is particularly important in our application. Common values settings generally demand that we allow correlated types (signals), and our test for common values relies on exogenous variation in entry arising through an instrument. We also avoid requiring a bijection (conditional on covariates) between entry and the unobservable (cf. Campo, Perrigne, and Vuong 2003, Haile, Hong, and Shum 2003, and Guerre, Perrigne, and Vuong 2009)—a requirement that can be difficult to rationalize and that limits the support of the unobservable. In contrast, we show that our empirical model can be derived from a two-stage game motivated by our application—an entry stage à la Berry (1992), in which bidders choose whether to acquire a signal of the good's value, followed by competitive bidding à la Milgrom and Weber (1982). In this example, the underlying unobserved heterogeneity may have arbitrary dimension and unrestricted support, may be correlated with observables, may exhibit spatial dependence, and may affect sample selection.

The next section presents our model. In Section III, we address non-parametric identification. Section IV describes our proposed estimation method. We then narrow our focus to the OCS auctions, with model estimates discussed in section V. We present the tests for common values in section VI, then explore revenue implications of our estimates in section VII. Our conclusion, in section VIII, includes a discussion of several caveats, extensions, and directions for future work.

## II. Model

We consider a standard model of a first-price sealed-bid auctions with symmetric affiliated values, extended to allow for auction-level heterogeneity and endogenous bidder entry. Auction  $t$  is associated with observed characteristics  $X_t \in \mathbb{X}$  and a scalar unobservable  $U_t$ . Without further loss, we let  $U_t$  be uniformly distributed on  $[0, 1]$ . We also assume independence between  $X_t$  and  $U_t$ .

ASSUMPTION 1.  $X_t \perp U_t$ .

The restriction to a scalar unobservable independent of  $X_t$  is less restrictive than it may appear. We show below that this representation can be derived—without loss of generality for most purposes motivating estimation of an auction model—from a model in which auction-level unobservables have arbitrary dimension and arbitrary dependence with  $X_t$ . In that model, the weak monotonicity conditions required below are also obtained as results rather than assumptions.

For each auction  $t$ , we postulate a two-stage process in which entry is followed by bidding. We do not specify a particular model of entry; rather,

we posit a reduced form for the entry outcome and assume Bayes-Nash equilibrium in the auction stage. The number of bidders entering auction  $t$  is denoted by  $N_t$ . Bidders are assumed risk neutral. Bidder  $i$ 's valuation for the good offered is denoted by  $V_{it}$ . Upon entering,  $i$  observes a private signal of  $V_{it}$ , denoted by  $S_{it} \in [\underline{s}, \bar{s}]$ . Let  $V_t = (V_{1t}, \dots, V_{N_t t})$ ,  $S_t = (S_{1t}, \dots, S_{N_t t})$ , and  $S_{-it} = S_t \setminus S_{it}$ .

The bidding stage follows Milgrom and Weber (1982). The realizations of  $(N_t, X_t, U_t)$  are common knowledge among bidders, as is the distribution of  $(S_t, V_t) | (N_t, X_t, U_t)$ .<sup>7</sup> In addition, each bidder  $i$  knows the signal  $S_{it}$ . Let  $F_{SV}(S_t, V_t | N_t, X_t, U_t)$  denote the joint distribution of signals and valuations conditional on  $(N_t, X_t, U_t)$ . We make the following standard assumptions on this conditional distribution.

**ASSUMPTION 2.** Given any  $(n, x, u) \in \text{supp}(N_t, X_t, U_t)$ , (i) the joint distribution  $F_{SV}(S_t, V_t | n, x, u)$  is affiliated and exchangeable in the indices  $i = 1, \dots, n$ , with  $\text{supp } V_{it} = [\underline{v}(x, u), \bar{v}(x, u)]$ ; (ii)  $S_t$  has a continuously differentiable joint density that is positive on  $(\underline{s}, \bar{s})^n$ ; and (iii) the joint distribution function  $E[V_{it} | S_{it}, S_{-it}, N_t = n, X_t = x, U_t = u]$  is strictly increasing in  $S_{it}$ .

Because the bidding stage involves a standard affiliated-values model, it nests a variety of special cases. With *private values*,  $E[V_{it} | S_{it}, S_{-it}, N_t, X_t, U_t]$  does not depend on  $S_{-it}$ . In our setting this is equivalent to bidders' knowledge of their valuations, that is,  $S_{it} = V_{it}$ . When  $E[V_{it} | S_{it}, S_{-it}, N_t, X_t, U_t]$  depends on  $S_{-it}$  we have *common values* (or *interdependent values*). A special case of the common values model is that of *pure common values*, where  $V_{it} = \bar{V}_t$  for all  $i$ .

A conditional expectation of particular relevance for what follows is

$$w(s_{it}; n_t, x_t, u_t) \equiv E \left[ V_{it} | S_{it} = \max_{j \neq i} S_{jt} = s_{it}, N_t = n_t, X_t = x_t, U_t = u_t \right].$$

This is a bidder's expected value of winning the auction conditional on all common-knowledge information, the observed private signal, and the event (typically counterfactual) that this signal ties for highest among

<sup>7</sup> Although standard, the assumption that bidders know the number of competitors is significant and may be inappropriate in some applications. Hendricks, Pinkse, and Porter (2003) point out that in OCS auctions, rivals' joint bidding agreements and participation in follow-up seismic surveys were typically known and that bidders performing a follow-up survey submitted bids on roughly 80% of the tracts analyzed. In our data, we reject the hypothesis that bidders are unaware of the realized  $n_t$  in favor of bidding that is more aggressive (given  $X_t, U_t$ ) for larger  $n_t$ . In addition, two composite tests of our maintained assumptions (discussed below) fail to reject. Nonetheless, these institutional features and empirical findings leave the possibility that bidders in our application have some uncertainty about the realization of  $N_t$ . A general challenge in accommodating such uncertainty in first-price auctions is the need to specify precisely the information bidders do have (see, e.g., Hendricks, Pinkse, and Porter 2003 and Gentry and Li 2014). Simulations indicate that the bias of our test for common values could go either way if our assumption is violated.

those received by all bidders at the auction. This expectation plays an important role in the theory because, when equilibrium bidding strategies turn out to be strictly increasing in signals, tying for the highest signal means that small deviations from one's equilibrium bid will change the identity of the winner. We therefore refer to  $w(s_i; n_t, x_t, u_t)$  as bidder  $i$ 's "pivotal expected value" at auction  $t$ . Pivotal expected values also play a central role in our strategy for discriminating between private values and common values.

We impose the following restriction on how the auction characteristics  $(X_t, U_t)$  affect bidder valuations.

ASSUMPTION 3. (i)  $V_{it} = \Gamma(X_t, U_t) V_{it}^0$ ; (ii) given  $N_t = n$ ,  $(V_{1t}^0, \dots, V_{nt}^0, S_{1t}, \dots, S_{nt})$  are independent of  $(X_t, U_t)$ ; and (iii)  $\Gamma(X_t, U_t)$  is strictly positive for all  $(X_t, U_t)$ , bounded, and weakly increasing in  $U_t$ .

Assumption 3 is an index restriction requiring multiplicative separability in  $(X_t, U_t)$  and weak monotonicity in  $U_t$ .<sup>8</sup> An assumption of multiplicative (or additive) separability has often been relied upon in the auctions literature, including for identification in other settings with unobserved heterogeneity. Our results rely on this assumption as well.<sup>9</sup> Without further loss, we normalize the scale of  $\Gamma$  relative to that of  $V_{it}^0$  by taking an arbitrary point  $x^0 \in \mathbb{X}$  and setting

$$\Gamma(x^0, 0) = 1. \quad (1)$$

We assume initially that the auction is conducted without a binding reserve price, although we also consider an important extension for our application allowing a random reserve price. Under assumption 2, the auction stage of our model admits a unique Bayesian Nash equilibrium in weakly increasing strategies; these strategies, denoted by  $\beta(\cdot; X_t, U_t, N_t) : [\underline{s}, \bar{s}] \rightarrow \mathbb{R}$ , are symmetric and strictly increasing.<sup>10</sup> Let the random variable  $B_{it} = \beta(S_{it}; X_t, U_t, N_t)$  denote the equilibrium bid of bidder  $i$  in auction  $t$ .

A useful fact is that the separability required by assumption 3 is inherited by the equilibrium bidding strategies.<sup>11</sup> Thus, under assumptions 2 and 3 we may write

<sup>8</sup> Without a distributional restriction like that in part ii of the assumption, part i would have no content. And because more "desirable" realizations of the unobservable state can be labeled with larger values, the monotonicity restriction in part iii rules out only variation with  $X_t$  in the partial order on the unobservable implied by desirability (see also example 1 and app. A).

<sup>9</sup> For what follows it is sufficient that the conditional expectations  $E[V_{it}|S_{it}, X_t, U_t, N_t]$  take the multiplicatively separable form. This weaker condition will be more natural when these ex ante conditional expectations are positive even though  $V_{it}$  may take negative values.

<sup>10</sup> See theorem 2.1 in Athey and Haile (2007) and the associated references. Milgrom and Weber (1982) characterize the equilibrium strategies.

<sup>11</sup> See, e.g., Haile, Hong, and Shum (2003), Athey and Haile (2007), or Krasnokutskaya (2011).

$$\beta(S_{it}; X_t, U_t, N_t) = \Gamma(X_t, U_t)\beta^0(S_{it}; N_t), \quad (2)$$

where  $\beta^0$  denotes the symmetric Bayesian Nash equilibrium bidding strategy for a standardized  $N_t$ -bidder auction at which  $\Gamma(X_t, U_t) = \Gamma(x^0, 0) = 1$ . Following Haile, Hong, and Shum (2003), we refer to  $B_{it}^0 = \beta^0(S_{it}; N_t)$  and  $V_{it}^0$  as “homogenized” bids and valuations, respectively.

We link the model of a single auction to the observed sample through assumption 4. Given assumption 3, this is the standard assumption that auctions are i.i.d. (independently and identically distributed) conditional on the auction characteristics,  $(N_t, X_t, U_t)$ . However, we do not require the latent characteristics  $U_t$  to be independent across auctions.

ASSUMPTION 4.  $(V_t^0, S_t) \perp (V_{t'}^0, S_{t'})$  for  $t' \neq t$ .

Finally, we specify the outcome of the entry stage by supposing that the number of bidders at auction  $t$  satisfies

$$N_t = \eta(X_t, Z_t, U_t) \quad (3)$$

for some function  $\eta$  that is weakly increasing in  $U_t$ . Formally, equation (3) is an assumed reduced form for the entry outcome. The weak monotonicity requirement links the interpretation of the unobservable in the entry and bidding stages: unobservables that make the good for sale more valuable also encourage more entry. The new variable  $Z_t$  in equation (3) is an exogenous auction-specific observable that affects bidder entry but is otherwise excludable from the auction model, as formalized in assumption 5.

ASSUMPTION 5. (i)  $Z_t \perp U_t | X_t$ ; and (ii)  $Z_t \perp (S_t, V_t^0) | N_t$ .

The following example, discussed more fully in appendix A, describes one fully specified two-stage game leading to the structure assumed above.

EXAMPLE 1. Consider a model of entry and bidding for an OCS oil and gas lease, where a standard simultaneous-move entry stage à la Berry (1992) precedes a competitive bidding stage à la Milgrom and Weber (1982). Players in the game are firms in the industry. The tract offered for lease is associated with observables  $X_b$ , which include (among other relevant covariates) the number of active leases on adjacent (“neighbor”) tracts and the sets of bidders for those leases. The active neighbor leases are owned by  $Z_t$  distinct *neighbor firms*. Tract-level unobservables are denoted by  $E_b$ , which may have arbitrary dimension, may be correlated with  $X_b$ , and may be spatially correlated. The characteristics  $X_t$  and  $E_t$  scale valuations (multiplicatively) through a bounded index  $\lambda(X_t, E_t)$ . Firms play a two-stage game. They first choose simultaneously whether to enter, with each entering firm  $i$  incurring a signal acquisition cost  $c_i(X_t)$ . These costs are common knowledge and lower for neighbor firms than for nonneighbor firms.<sup>12</sup> Entrants learn their private signals and the number

<sup>12</sup> This structure generalizes that in Hendricks and Porter (1988), where neighbors obtain a private signal for free but nonneighbors face an infinite cost of signal acquisition.



of entrants, then participate in a first-price sealed-bid auction with symmetric affiliated values. Appendix A shows that all pure-strategy perfect Bayesian equilibria with weakly increasing strategies in the bidding stage can be represented by the model and assumptions above. This representation is obtained by defining  $U_i = F_\lambda(\lambda(x_i, E_i)|x_i)$ , where  $F_\lambda(\cdot|x)$  is the cumulative distribution function (CDF) of the random variable  $\lambda(x, E_i)$ . Observe that the distribution of  $U_i$  does not vary with  $X_i$ , although its interpretation does.<sup>13</sup>

### III. Nonparametric Identification

Here we develop sufficient conditions for identification of the entry model, the index function  $\Gamma$ , and key features of the bidding model. Throughout, we assume that the observables include all bids as well as  $X_b, Z_b, N_b$ .<sup>14</sup> Let  $\mathbb{Y}$  denote the support of  $(X_b, Z_b)$ ; let  $\mathbb{Y}(n)$  denote that conditional on  $N_i = n$ . Let  $\underline{n} \geq 0$  denote the minimum value in the support of  $N_b$  with  $\bar{n}$  denoting the maximum. Recalling equation (1), for convenience we take  $x^0$  such that for some  $z$  we have  $(x^0, z) \in \mathbb{Y}(\underline{n})$ .

#### A. Identification of the Entry Model

We show identification of the entry model under the following regularity condition.

**ASSUMPTION 6.** For all  $(x, z) \in \mathbb{Y}$ , there exist  $\underline{n}(x, z)$  and  $\bar{n}(x, z)$  such that  $\eta(x, z, U_i)$  has support  $(\underline{n}(x, z), \underline{n}(x, z) + 1, \dots, \bar{n}(x, z))$ .

Given assumption 6, for any  $(x, z) \in \mathbb{Y}$ , the function  $\eta(x, z, \cdot)$  is characterized by thresholds  $\tau_{\underline{n}(x, z)-1}(x, z) \leq \tau_{\underline{n}(x, z)}(x, z) \leq \dots \leq \tau_{\bar{n}(x, z)}(x, z)$ , where

$$\tau_{\underline{n}(x, z)-1}(x, z) \equiv 0 \quad \text{and} \quad \tau_{\bar{n}(x, z)}(x, z) \equiv 1, \quad (4)$$

and for  $n = \{\underline{n}(x, z), \dots, \bar{n}(x, z)\}$ ,  $\tau_{n-1}(x, z) = \inf\{u \in [0, 1] : \eta(x, z, u) \geq n\}$ . With this observation, identification of  $\eta$  follows easily.

**THEOREM 1.** Under assumptions 1–6,  $\eta$  is identified.

*Proof.* For each  $(x, z)$ ,  $\underline{n}(x, z)$  and  $\bar{n}(x, z)$  equal  $\min\{N_i|X_i = x, Z_i = z\}$  and  $\max\{N_i|X_i = x, Z_i = z\}$ , respectively. For  $n = \underline{n}(x, z), \dots, \bar{n}(x, z)$ ,  $\tau_n(x, z) = \Pr(N_i \leq n|X_i = x, Z_i = z)$ . With (4), this implies the result. QED

<sup>13</sup> In that case, knowledge of the function  $\Gamma$  will not be sufficient to characterize the effects of a ceteris paribus change in  $X_i$  on bidder valuations. See the additional discussion in app. A.

<sup>14</sup> In some applications, including ours, data may be available only for auctions attracting at least one bidder. In app. C (apps. C–K are available online), we show that, within the fully specified model of example 1, our maintained assumptions and analysis remain valid in the presence of such sample selection.

Identification of  $\eta$  determines the effects of  $Z_t$  on bidder entry and provides bounds  $\tau_{n-1}(x_t, z_t)$  and  $\tau_n(x_t, z_t)$  on the realization of each unobservable  $U_t$ . As shown in the following corollary (proved in app. B), it also determines the distribution of  $U_t$  conditional on  $(X_t, N_t)$ .

**COROLLARY 1.** Under assumptions 1–6, the distribution of  $U_t|(X_t, N_t)$  is identified.

### B. Identification of the Index Function

Let  $\gamma(x, u) = \ln \Gamma(x, u)$ . We first provide conditions sufficient to identify  $\gamma(x, u)$  at each  $x \in \mathbb{X}$  and  $u \in \mathcal{U}_x$ , where

$$\mathcal{U}_x = \bigcup_{z: (x, z) \in \mathbb{Y}} \{\tau_{n-1}(x, z), \tau_n(x, z)\}.$$

$$n \in \text{supp } \eta(x, z, U_t)$$

We then give additional conditions guaranteeing that  $\mathcal{U}_x = [0, 1]$  for each  $x$ . We begin with the following result, whose proof illustrates a key strategy.

**LEMMA 1.** Under assumptions 1–6, for all  $n \geq \underline{n}$ , all  $(x, z) \in \mathbb{Y}(n)$ , and all  $(x', z') \in \mathbb{Y}(n)$ , the differences  $\gamma(x', \tau_n(x', z')) - \gamma(x, \tau_n(x, z))$  and  $\gamma(x', \tau_{n-1}(x', z')) - \gamma(x, \tau_{n-1}(x, z))$  are identified.

*Proof.* By equation (2) and monotonicity of the equilibrium bid function,

$$\inf\{\ln B_{it}|N_t = n, X_t = x, Z_t = z\} = \gamma(x, \tau_{n-1}(x, z)) + \ln \beta^0(\underline{s}; n).$$

So under assumptions 1–6, for any  $n$  and all  $(x, z)$  and  $(x', z')$  in  $\mathbb{Y}(n)$ , the differences  $\gamma(x', \tau_{n-1}(x', z')) - \gamma(x, \tau_{n-1}(x, z))$  are identified.<sup>15</sup> Similarly, since

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

we obtain identification of the differences  $\gamma(x', \tau_n(x', z')) - \gamma(x, \tau_n(x, z))$  for all  $n$  and all  $(x, z)$  and  $(x', z')$  in  $\mathbb{Y}(n)$ . QED

Thus far we have not imposed any requirement on the support of  $Z_t$  or its effect on entry outcomes. We do so below in order to obtain point identification of  $\gamma$ . However, even in the case that no instrument is available, theorem 1, corollary 1, and lemma 1 still hold. And once  $\gamma$  is known,  $Z_t$  plays no further role in our identification results. Thus, while we rely on an instrument to obtain point identification, this reliance is formally limited to ensuring that we can move from the partial identification of  $\gamma$  provided by lemma 1 to point identification of  $\gamma$ .

As a step toward point identification, we introduce two additional assumptions. These allow us to show that the first differences obtained above

<sup>15</sup> Both  $\beta^0(\underline{s}; n)$  and  $\beta^0(\bar{s}; n)$  are finite under assumption 2.

can be differenced again, canceling common terms, to obtain a set of first differences sufficient to pin down the value of the index  $\gamma(x, u)$  at all  $x$  and  $u \in \mathcal{U}_x$ .

ASSUMPTION 7. For all  $n \in \{\underline{n} + 1, \underline{n} + 2, \dots, \bar{n}\}$ ,  $\mathbb{Y}(n - 1) \cap \mathbb{Y}(n)$  is nonempty.

ASSUMPTION 8. There exists  $n^*$  such that

- i.  $\forall n \in \{\underline{n}, \dots, n^*\}$ ,  $\mathbb{Y}$  contains points  $(x(n), z(n))$  and  $(x(n), \hat{z}(n))$  such that  $\underline{n}(x(n), z(n)) = n$  and  $\underline{n}(x(n), \hat{z}(n)) = n + 1$ ; and
- ii.  $\forall n \in \{n^*, \dots, \bar{n}\}$ ,  $\mathbb{Y}$  contains points  $(x'(n), z'(n))$  and  $(x'(n), \hat{z}'(n))$  such that  $\bar{n}(x'(n), z'(n)) = n$  and  $\bar{n}(x'(n), \hat{z}'(n)) = n - 1$ .

Assumption 7 requires variation in  $U_i$  that produces local variation in entry. For example, this rules out trivial cases in which  $U_i$  has no effect on  $N_i$ . Assumption 8 requires variation in the instrument  $Z_i$  that can induce local variation in the support of the entry outcomes at some values of  $X_i$ . We prove the following lemma in appendix B.

LEMMA 2. Under assumptions 1–8, for all  $n \geq \underline{n}$  and all  $(x, z) \in \mathbb{Y}(n)$ , the values of  $\gamma(x, \tau_{n-1}(x, z))$  and  $\gamma(x, \tau_n(x, z))$  are identified.

By theorem 1, the values of  $\tau_{n-1}(x, z)$  and  $\tau_n(x, z)$  are known for all  $n$  and  $(x, z) \in \mathbb{Y}(n)$ . Thus, lemma 2 demonstrates identification of  $\gamma(x, u)$  at each  $x \in \mathbb{X}$  and  $u \in \mathcal{U}_x$ . In general, this may still deliver only partial identification of the index function  $\gamma$ , so that in practice one may rely on parametric structure to interpolate between the points  $\{x \in \mathbb{X}, u \in \mathcal{U}_x\}$  at which  $\gamma(x, u)$  is nonparametrically point identified. However, the following conditions are sufficient to ensure that no such interpolation is necessary.

ASSUMPTION 9.

- i. For all  $x \in \mathbb{X}$ ,  $\text{supp } Z_i | X_i = x$  is connected.
- ii. For all  $(x, z, u) \in \mathbb{Y} \times (0, 1)$  and all  $\delta > 0$  such that  $(u - \delta, u + \delta) \subset (0, 1)$ , there exists  $\epsilon > 0$  such that if  $\|z' - z\| < \epsilon$  then  $\eta(x, z', u') = \eta(x, z, u)$  for some  $u' \in (u - \delta, u + \delta)$ .

ASSUMPTION 10. For every  $x \in \mathbb{X}$  there exists a finite partition  $0 = \tau^0(x) < \tau^1(x) < \dots < \tau^{K(x)}(x) = 1$  of the unit interval such that for each  $k = 1, \dots, K(x)$  and some  $z(k), z'(k) \in \text{supp } Z_i | X_i = x$ ,  $\eta(x, z(k), \tau^{k-1}(x)) > \eta(x, z'(k), \tau^k(x))$ .

Assumption 9 requires continuously distributed  $Z_i$  and a type of continuous substitution between  $Z_i$  and  $U_i$  in the “production” of bidder entry: it must be possible to offset the effect (on entry) of a small change in  $Z_i$  with a small change in  $U_i$ . Assumption 10 requires that variation in  $Z_i$  have sufficient effect on participation to offset some discrete variation in the unobservable  $U_i$ . A sufficient condition is that for each  $x$  there exist  $z$  and  $z'$  such that  $\eta(x, z, \tau_{n-1}(x, z)) > \eta(x, z', \tau_n(x, z))$  for all  $n \in \{\underline{n}(x, z), \dots,$

$\bar{n}(x, z)\}$ ; in this case, the set  $\{\tau_{\bar{n}(x,z)-1}, \dots, \tau_{\bar{n}(x,z)}\}$  could define the partition  $\tau^0(x) < \tau^1(x) < \dots < \tau^K(x)$ .

The following lemma, whose proof is provided in appendix B, leads us to the point identification of  $\gamma$  (and therefore  $\Gamma$ ) demonstrated in theorem 2.

LEMMA 3. Under assumptions 1–9,  $\tau_{n-1}(X_t, Z_t)$  is continuous in  $Z_t$  on the preimage of  $(0, 1)$ .

THEOREM 2. Under assumptions 1–10,  $\Gamma$  is identified on  $\mathbb{X} \times [0, 1]$ .

*Proof.* We need only show that  $\mathcal{U}_x = [0, 1]$  for each  $x \in \mathbb{X}$ . For arbitrary  $x \in \mathbb{X}$ , let  $0 = \tau^0(x) < \tau^1(x) < \dots < \tau^{K(x)}(x) = 1$  be as in assumption 10. Take any  $k \in \{1, 2, \dots, K(x)\}$ , and let  $z = z(k)$  and  $z' = z'(k)$  be as in assumption 10. Let  $n = \eta(x, z, \tau^{k-1}(x))$ . Because  $\eta(x, z, \tau^{k-1}(x)) > \eta(x, z', \tau^k(x))$ , we have  $\eta(x, z', \tau^k(x)) < n$  and, therefore,  $\tau_{n-1}(x, z) \leq \tau^{k-1}(x) < \tau^k(x) \leq \tau_{n-1}(x, z')$ . Because the continuous image of a connected set is connected, lemma 3 and assumption 9 (part i) then imply that for every  $\tilde{\tau} \in [\tau^{k-1}(x), \tau^k(x)]$  there exists  $z^{\tilde{\tau}}$  such that  $\tau_{n-1}(x, z^{\tilde{\tau}}) = \tilde{\tau}$ . QED

### C. Identification of the Bidding Model

We now demonstrate identification of the joint distribution of the pivotal expected values  $w(S_{it}; n, x, u), \dots, w(S_{nt}; n, x, u)$  for all  $x \in \mathbb{X}$ ,  $u \in [0, 1]$ , and  $n$  in the support of  $N_i | \{X_i = x, U_i = u\}$ . For a private-values model this is equivalent to identification of the joint distribution of bidder valuations conditional on  $(X_i, N_i, U_i)$ . Thus, theorem 3 below demonstrates identification of the affiliated private-values model. Without the restriction to private values, our result here provides an important form of partial identification.<sup>16</sup> For our empirical application, for example, this is sufficient to allow us to test the hypothesis of equilibrium bidding in the affiliated-values model, to test the hypothesis of private values against the alternative of common values, to examine the potential gains from changes in the auction format, and to assess the effects of competition on bidder market power.<sup>17</sup>

Let  $G_{M|B}(m|b, x, u, n) = \Pr(\max_{j \neq i} B_{jt} \leq m | B_{it} = b, X_t = x, U_t = u, N_t = n)$ , and let  $g_{M|B}(m|b, x, u, n)$  denote the associated conditional density (guaranteed to exist by assumption 2 and strict monotonicity of the equilibrium bid function). Following Laffont and Vuong (1993), Guerre, Perrigne, and Vuong (2000), and Li, Perrigne, and Vuong (2000, 2002), one can characterize the relationship between each realized  $w(s_{it}; n_t, x_t, u_t)$  and the associated equilibrium bid  $b_{it} = \beta(s_{it}; x_t, u_t, n_t)$  in terms of the

<sup>16</sup> Without additional information or structure, common values models are not identified from bidding data. See, e.g., Laffont and Vuong (1993) and Athey and Haile (2002).

<sup>17</sup> Tang (2011) shows how the joint distribution of pivotal expected values can be used to bound counterfactual revenues in standard auctions with binding reserve prices.

joint distribution of equilibrium bids. In particular, each  $b_{it}$  must satisfy the first-order condition (see, e.g., Athey and Haile 2007)

$$w(s_{it}; n_t, x_t, u_t) = b_{it} + \frac{G_{M|B}(b_{it}|b_{it}, x_t, u_t, n_t)}{g_{M|B}(b_{it}|b_{it}, x_t, u_t, n_t)}. \quad (5)$$

Although (5) expresses the pivotal expected value  $w(s_{it}; n_t, x_t, u_t)$  as a functional of a conditional distribution of bids, the presence of  $u_t$  on the right-hand side creates challenges. Because realizations of  $U_t$  are not observable or identified, one cannot directly condition on them to identify the functions  $G_{M|B}$  and  $g_{M|B}$ . This precludes obtaining identification directly from equation (5). With the preceding results, however, we can overcome this problem.

Observe that, like valuations and bids, the pivotal expected values  $w(s_{it}; n_t, x_t, u_t)$  will have the separable structure

$$w(s_{it}; n_t, x_t, u_t) = w^0(s_{it}; n_t)\Gamma(x_t, u_t), \quad (6)$$

where  $w^0(s_{it}; n_t) \equiv E[V_{it}^0|S_{it} = \max_{j \neq i} S_{jt} = s_{it}, N_t = n_t]$ . We refer to  $w^0(s_{it}; n_t)$  as bidder  $i$ 's "homogenized pivotal expected value" at auction  $t$ . The first-order condition (5) can then be written as

$$w^0(s_{it}; n_t) = b_{it}^0 + \frac{G_{M^0|B^0}(b_{it}^0|b_{it}^0, n_t)}{g_{M^0|B^0}(b_{it}^0|b_{it}^0, n_t)}, \quad (7)$$

where  $G_{M^0|B^0}(m|b, n_t) = \Pr(\max_{j \neq i} B_{jt}^0 \leq m|B_{it}^0 = b, N_t = n)$  and  $g_{M^0|B^0}(m|b, n_t)$  is the associated conditional density.

Let  $\tilde{B}_{it} = \ln(B_{it})$  and  $\tilde{B}_{it}^0 = \ln(B_{it}^0)$ . By assumption 3,

$$\tilde{B}_{it} = \tilde{B}_{it}^0 + \gamma(X_t, U_t), \quad (8)$$

with  $\tilde{B}_{it}^0$  and  $\gamma(X_t, U_t)$  independent conditional on  $(N_t, X_t)$ . Lemma 4 (proved in app. B) shows that theorem 2 and a standard deconvolution argument yield identification of the joint distribution of  $(B_{1t}^0, \dots, B_{nt}^0)$  for all  $n$ .

LEMMA 4. Under assumptions 1–10, conditional on any  $N_t = n$ , the joint density of  $(B_{1t}^0, \dots, B_{nt}^0)$  is identified.

This leads directly to our main identification result.

THEOREM 3. Let assumptions 1–10 hold. Then for all  $x \in \mathbb{X}$ ,  $u \in [0, 1]$ , and  $n \geq 2$  in the support of  $N_t|\{X_t = x\}$ , the joint distribution of  $(w(S_{1t}; n, x, u), \dots, w(S_{nt}; n, x, u))$  is identified.

*Proof.* Fix  $n$ . From equation (7), we have  $w^0(S_{it}; n) = \xi(B_{it}^0; n) \equiv B_{it}^0 + G_{M^0|B^0}(B_{it}^0|B_{it}^0, n)/g_{M^0|B^0}(B_{it}^0|B_{it}^0, n)$ . By lemma 4, the joint distribution  $(\xi(B_{1t}^0; n), \dots, \xi(B_{nt}^0; n))$  is known. This implies identification of the joint distribution of  $(w^0(S_{1t}; n), \dots, w^0(S_{nt}; n))$ . The result then follows immediately from equation (6) and theorem 2. QED

#### IV. Estimation

We propose a two-stage semiparametric estimation strategy. The first stage involves semiparametric sieve quasi-maximum likelihood estimation (QMLE) of the entry thresholds  $\tau_t(x, z)$ , the index function  $\gamma$ , and the joint distributions of homogenized equilibrium bids. In the second stage, for each level of bidder entry, we estimate the joint distribution of homogenized pivotal expected values by plugging draws from the estimated distribution of homogenized bids into the auction first-order condition and constructing the empirical distribution of the resulting pseudosample.

##### A. Stage 1: Sieve-QMLE

Let  $\theta_\tau$  denote the parameters of the entry model,  $\theta_\gamma$  the parameters of the index function  $\gamma$ , and  $\theta_B$  the parameters of the joint distributions of log homogenized bids. Let  $L_{1t}(n_t; \theta_\tau) = \Pr(N_t = n_t | X_t = x_t, Z_t = z_t; \theta_\tau)$  denote the (conditional on  $(x_t, z_t)$ ) likelihood for the entry outcome in auction  $t$ . Let  $L_{2t}(b_t | n_t; \theta_\gamma, \theta_B, \theta_\tau)$  denote the likelihood of the observed bids at auction  $t$ , conditional on the entry outcome  $n_t$  (and on  $(x_t, z_t)$ ). Defining  $\theta = (\theta_\tau, \theta_\gamma, \theta_B)$ , the conditional quasi-likelihood<sup>18</sup> function for the observed outcomes  $\{(n_t, b_t)\}_{t=1}^T$  can be written

$$\mathcal{L}(\theta) = \prod_t L_{1t}(n_t; \theta_\tau) L_{2t}(b_t | n_t; \theta_\gamma, \theta_B, \theta_\tau).$$

We give details of our empirical specification and the two components of the quasi likelihood in sections IV.A.1 and IV.A.2 below. Estimates of the parameter vector  $\theta$  can be obtained by maximizing  $\mathcal{L}(\theta)$ . Because  $\theta_\tau$  is identified from the entry outcomes alone, it is also possible to split the QMLE stage, first maximizing  $\prod_t L_{1t}(n_t; \theta_\tau)$  to estimate  $\theta_\tau$ , then maximizing  $\prod_t L_{2t}(b_t | n_t; \theta_\gamma, \theta_B, \theta_\tau)$  conditional on  $\theta_\tau$ . In our data, the two approaches yield very similar estimates. However, because we found the two-step QMLE procedure to be more numerically stable in bootstrap samples, below we report results using the two-step version.

Consistency can be confirmed by adapting the results of White and Wooldridge (1991) for sieve-extremum estimators with weakly dependent time-series data to the case of weak spatial dependence.<sup>19</sup> To conduct inference, we use a nonparametric block bootstrap procedure that captures both dependence among bids within an auction and spatial dependence between the unobservables  $U_i$  across auctions. Specifically, we resample

<sup>18</sup> Recall that we permit spatial dependence.

<sup>19</sup> In particular, we can represent tract locations by points in  $\mathbb{Z}^2$ . Then, under a standard “expanding-domain” asymptotics, White and Wooldridge’s uniform-consistency result for stationary  $\alpha$ -mixing time-series data (corollary 2.6) can be extended using a Bernstein-type inequality for  $\alpha$ -mixing random fields on  $\mathbb{Z}^2$  (e.g., Yao 2003).

auctions with replacement, taking all bids from the selected auction and including in the bootstrap sample all auctions on neighbor tracts as well. Following Hendricks, Pinkse, and Porter (2003), we resample weighting auctions by factors inversely proportional to the number of auctions in the neighborhood of the tract.<sup>20</sup>

### 1. Entry Thresholds

Our entry model above reduces to an ordered-response model where, given  $X_i = x$  and  $Z_i = z$ , we have  $N_i = n$  if and only if  $U_i \in (\tau_{n-1}(x, z), \tau_n(x, z))$ . Equivalently, given any strictly increasing univariate CDF  $H$ , we have

$$\{N_i = n | X_i = x, Z_i = z\} \Leftrightarrow \{A_i \in (\alpha_{n-1}(x, z), \alpha_n(x, z))\},$$

where  $A_i \sim H$  and  $\alpha_n(x, z) = H^{-1}(\tau_n(x, z))$ . We specify a linear threshold function  $\alpha_n(x, z) = \alpha_n - x'\alpha_x - z'\alpha_z$  and specify  $H$  as the standard normal CDF, yielding an ordered-probit model. Letting  $\theta_\tau = (\{\alpha_n\}_{n=\underline{n}}^{\bar{n}-1}, \alpha_x, \alpha_z)$ , we have  $L_{1i}(n; \theta_\tau) = H(\alpha_n(x_i, z_i; \theta_\tau)) - H(\alpha_{n-1}(x_i, z_i; \theta_\tau))$ .

### 2. Index Function and Homogenized-Bid Distribution

Given our focus on testing for common values, we prioritize flexibility in how the joint distribution and density of bids can vary with  $n$  when specifying the second part of the quasi likelihood. We specify the index function  $\gamma$  parametrically as  $\gamma(X_i, U_i; \theta_\gamma)$ ; we use a linear specification below. For each value of  $n$ , the joint density of log homogenized bids is specified semiparametrically, using a parametric copula and a nonparametric (Bernstein polynomial sieve) specification of the common marginal distribution.<sup>21</sup>

We specify the marginal density of a generic bidder's log homogenized bid in an  $n$ -bidder auction as

<sup>20</sup> Similar results are obtained without weighting. Applying the results of van der Vaart and Wellner (1996) and Lahiri (2003), one can verify validity of the bootstrap when we interpret our finite-sample estimator as that for a parametric model. General conditions for consistency of bootstrap inference procedures for sieve M-estimators in the i.i.d. setting can be found, e.g., in Ma and Kosorok (2005) and Chen and Pouzo (2009). See also Chen and Liao (2014), and Chen, Liao, and Sun (2014) in the case of time-series data.

<sup>21</sup> See, e.g., Chen, Fan, and Tsyrennikov (2006). Prior work using Bernstein polynomials in estimation of auction models includes Komarova (2017) and Kong (2020, 2021). Hubbard, Li, and Paarsch (2012) previously combined a parametric copula and nonparametric (kernel) specification of marginal densities to estimate auction models. The Gaussian copula is computationally advantageous for our estimation approach. We discuss computation in app. D.

$$\tilde{g}_{B^0}(\tilde{b}^0; \theta_b, n) = \sum_{j=0}^m \theta_{b,n}^{(j)} q_{j,m}(\Phi(\tilde{b}^0)) \phi(\tilde{b}^0), \quad (9)$$

where  $\Phi(\cdot)$  and  $\phi(\cdot)$  denote the standard normal distribution and density functions, respectively, and  $q_{j,m}(\nu) = \binom{m}{j} \nu^j (1-\nu)^{m-j}$ . Here  $m$  is a parameter, growing with the sample size, that determines the order of the Bernstein polynomial approximation. Let  $\theta_{b,n} = \{\theta_{b,n}^{(j)}\}_{j=0}^m$ . Thus, the parameter vector  $\theta_b$  in equation (9) represents  $\{\theta_{b,n}\}_{n=\underline{n}}^{\bar{n}}$ .

Because Bernstein polynomials approximate functions with domain  $[0, 1]$ , in equation (9) we use Bernstein polynomials to approximate the marginal density of the transformed variable  $\Phi(\tilde{b}^0)$ .<sup>22</sup> This transformation also helps ensure that the nonparametric estimator will offer sensible approximations even in modest sample sizes. With  $m = 0$ , for example, the distribution of log bids would be normal. Thus, the nonparametric component of our specification is based on a sequence of approximating models that starts with a natural (lognormal) parametric specification and adds flexibility as permitted by the sample size.

Let  $\tilde{G}_{B^0}(\tilde{b}^0; \theta_b, n)$  denote the CDF associated with  $\tilde{g}_{B^0}(\tilde{b}^0; \theta_b, n)$ , and let  $\chi(\cdot; \rho_n)$  denote the symmetric Gaussian copula density with covariance parameter  $\rho_n$ . The joint density  $\tilde{g}_{B^0}(\tilde{b}_1^0, \dots, \tilde{b}_n^0; \theta_{b,n}, \rho_n, n)$  of log homogenized bids in  $n$ -bidder auctions is then given by

$$\chi(\tilde{G}_{B^0}(\tilde{b}_1^0; \theta_{b,n}, n), \dots, \tilde{G}_{B^0}(\tilde{b}_n^0; \theta_{b,n}, n); \rho_n) \tilde{g}_{B^0}(\tilde{b}_1^0; \theta_{b,n}, n) \dots \tilde{g}_{B^0}(\tilde{b}_n^0; \theta_{b,n}, n). \quad (10)$$

Letting  $\rho = \{\rho_n\}_{n=\underline{n}}^{\bar{n}}$  and  $\theta_B = (\theta_b, \rho)$ , we have

$$L_{2t}(b_t | n_t; \theta_B) = \int_{\tau_{n-1}(x_t, z_t; \theta_\tau)}^{\tau_n(x_t, z_t; \theta_\tau)} \frac{\tilde{g}_{B^0}(\tilde{b}_{1t} - \gamma(x_t, u; \theta_\gamma), \dots, \tilde{b}_{nt} - \gamma(x_t, u; \theta_\gamma); \theta_{b,n}, \rho_n, n_t)}{\tau_n(x_t, z_t; \theta_\tau) - \tau_{n-1}(x_t, z_t; \theta_\tau)} du, \quad (11)$$

where  $\tau_{n-1}(x_t, z_t; \theta_\tau)$  and  $\tau_n(x_t, z_t; \theta_\tau)$  denote the bounds on  $u_t$  implied by the entry-model parameters  $\theta_\tau$ .<sup>23</sup> We approximate the integral by Monte Carlo simulation.

<sup>22</sup> We use the normalization  $\gamma(0, 0) = 0$  (recall eq. [1]). In estimation, however, we add intercepts  $\gamma_n^0$  for each value of  $n_t$  to the index function  $\gamma(x_t, u_t; \theta_\gamma)$ , implying that in this step the joint densities estimated are those of centered log homogenized bids  $\tilde{b}_i^0 - \gamma_n^0$ . We then adjust the location of each estimated density by the associated intercept estimate to obtain our estimated density of (uncentered) log homogenized bids. This procedure offers practical advantages by centering log homogenized bids at zero before transformation by the normal CDF, ensuring that the location of the estimated bid distribution can move freely with  $n_t$  and freeing the Bernstein coefficients to capture features of the marginal density other than its location.

<sup>23</sup> When  $n_t = 1$ , one may set  $L_{2t}(b_t | n_t; \theta_B) = 1$  by convention, since our baseline model of competitive bidding does not imply an interpretation of the quantity  $\tilde{b}_i - \gamma(x_t, u_t)$  in that case. Alternatively, because we specify different parameters  $\theta_{b,n}$  for each value of  $n$ , eq. (11) gives a correct expression for  $L_{2t}(b_t | n_t; \theta_B)$  for  $n_t = 1$  whenever bids in one-bidder auctions are assumed to inherit the separable structure required of valuations. Below, we develop an



### B. Stage 2: Invert Equilibrium First-Order Conditions

Given the first-stage estimates of the index function  $\gamma$  and joint distribution of homogenized bids, estimation of the relevant auction primitives is straightforward and does not involve further use of the data. We use the equilibrium first-order condition (7), which can be written in terms of the distribution of log homogenized bids as

$$w^0(s_{it}; n_t) = \exp(\tilde{b}_{it}^0) \left( 1 + \frac{\tilde{G}_{M|B}(\tilde{b}_{it}^0 | \tilde{b}_{it}^0, n_t)}{\tilde{g}_{M|B}(\tilde{b}_{it}^0 | \tilde{b}_{it}^0, n_t)} \right), \quad (12)$$

where  $\tilde{G}_{M|B}(\tilde{b}_{it}^0 | \tilde{b}_{it}^0, n_t)$  and  $\tilde{g}_{M|B}(\tilde{b}_{it}^0 | \tilde{b}_{it}^0, n_t)$  are, respectively, the CDF and the probability density function of  $\max_{j \neq i} \tilde{B}_{jt}^0$  conditional on  $\tilde{B}_{it}^0 = \tilde{b}_{it}^0$ .

For each value of  $N_t = n$ , we transform the estimated joint distributions and densities obtained from stage 1 to construct the conditional distributions and densities appearing in equation (12) (see app. D). Then we draw log homogenized bids from their estimated marginal distributions and plug these into equation (12), yielding pseudosamples of the vectors  $(w^0(s_{1t}; n), \dots, w^0(s_{nt}; n))$  for many simulated auctions  $t$ . The empirical distribution of these pseudodraws provides a consistent estimate of the joint distribution of homogenized pivotal expected values for  $n$ -bidder auction. Although these joint distributions will suffice for our application, the pseudodraws can also be scaled by the estimated value of the index  $\Gamma(x, u)$  in order to estimate the joint distributions of (non-homogenized) pivotal expected values, given  $X_t = x$  and  $U_t = u$ . In a private-values setting, for example, this would yield an estimate of the joint distribution of bidder valuations.

## V. OCS Auctions

### A. Background and Data

We examine first-price sealed-bid auctions of oil and gas leases in the OCS held between 1954 and 1983. Extensive discussion of the OCS auctions can be found in, for example, Gilley and Karels (1981), Hendricks and Porter (1988), Hendricks, Porter, and Spady (1989), and Hendricks, Pinkse, and Porter (2003). For a more complete institutional background we refer readers to that work, upon which we rely heavily ourselves. Briefly, however, auctions were held for the right to lease a specified tract for exploration and production of oil, gas, and other minerals. The seller in these auctions was an agency of the US Department of the

---

important extension for our application, incorporating a random reserve price, where this separability is an implication of equilibrium behavior. Therefore, given the large number of one-bidder auctions in our sample, we include these in the QMLE, using eq. (11) for all specifications.

TABLE 1  
SAMPLE SIZES

	1	2	3	4	5	6	7	8	9	10	11–17
$T_n$	814	498	293	229	172	127	100	73	56	56	128

NOTE.— $T_n$  denotes the number of auctions in our sample with  $n$  (columns) bidders.

Interior known as the Mineral Management Service (MMS). Tracts typically covered a rectangular “block” of 5,000–5,760 acres in the Gulf of Mexico.<sup>24</sup> Production on a tract was subject to royalty payments by the leaseholder at a prespecified rate, usually one-sixth. Bids at an auction were offers of an additional up-front “bonus” payment for the right to become the leaseholder.

No exploratory drilling was permitted before the auction, although in some cases exploration and production would already have occurred on neighbor tracts and would be publicly observable. Bids would also reflect information obtained through evaluation of data from magnetic, gravity, and seismic surveys. Although initial collection of survey data was often funded jointly, firms relied on their own experts for modeling and analysis of the data and often performed follow-up surveys of the tracts on which they intended to bid. Differences in expert assessments of the survey data are likely an important source of heterogeneity in bidder beliefs about the value of a given tract (Hendricks, Pinkse, and Porter 2003). These features lead us to treat bidder entry as a decision to acquire a costly signal about the value of the tract (recall example 1).

We have data on all auctions attracting at least one bidder (recall n. 14). Table 1 shows the number of auctions in our sample by number of bidders. We do not separate wildcat, development, and drainage tracts; instead, we account directly for the presence of active neighbor leases and neighbor production, and we allow asymmetry between neighbor and nonneighbor costs of signal acquisition in a way that generalizes the structure considered in Hendricks and Porter (1988; see example 1).<sup>25</sup> Like Aradillas-López et al. (2019), we model bidders as symmetric conditional on acquisition of a signal. Thus, while bidders may decide not to acquire a signal through analysis of the seismic data and may reach different conclusions from such analysis, the technology producing signals is modeled as symmetric across firms. Formal tests, developed in appendix E, fail to reject symmetry.

<sup>24</sup> We limit attention to tracts of at least 4,000 acres. Smaller tracts are typically partial blocks, where several of our measures of neighborhood characteristics would have interpretations different from those on standard tracts. We also drop auctions with missing values for our covariates.

<sup>25</sup> As a robustness check, we also examine the subset of auctions that excludes drainage tracts.

The MMS sometimes announced a small minimum acceptable bid of \$10–\$25 per acre; the MMS also retained the option to reject all bids when it deemed the auction to be noncompetitive. Such rejections were rare in our sample, except at auctions attracting only one bid (see app. H). We initially treat both the announced minimum bid and the MMS bid rejection policy as nonbinding, as in Li, Perrigne, and Vuong (2000). We also estimate a variation of the model in which the MMS bid rejection policy is modeled with a random secret reserve price.

Limited forms of joint bidding were permitted in these auctions. Following the literature, we model each bid as coming from a generic “bidder,” which might be solo firm or a bidding consortium.<sup>26</sup> Typically, a tract will have eight neighbors, only some (or none) of which would be “active” (under lease) when the tract is offered. Our measure of the number of neighbor firms (distinct owners of adjacent leases) accounts for the presence of joint bidding by linking firms that have bid together previously in the same neighborhood, following criteria developed by Aradillas-López et al. (2019).

Our tract characteristics  $X_i$ , all measured as of the time of the auction, include the number of active neighbor leases, whether the tract is isolated (no active neighbors), the number of firms that bid for neighbor leases, whether the tract was offered previously (attracting no bidders or being relinquished by a prior leaseholder), whether a lease has expired on a neighbor tract, the number of neighbor tracts previously drilled, the number of “hits” on neighbor tracts, average water depth (and its square), and the royalty rate associated with the lease. We present a summary of these auction characteristics in table 2. Below we also incorporate year fixed effects.<sup>27</sup>

Following example 1, our instrument for bidder entry is the number of neighbor firms (also in table 2).<sup>28</sup> This variable is likely to affect bidder entry because ownership of a neighbor tract is likely to reduce the cost of assessing the value of the current tract. As discussed in appendix A, when we condition on the number of neighbor tracts and the set of firms that previously bid for those tracts, variation in the number of neighbor firms is determined entirely by the realizations of bidder signals at prior auctions, and therefore independent of  $U_i$  under our maintained assumptions.

<sup>26</sup> Hendricks and Porter (1992) and Hendricks, Porter, and Tan (2008) examine empirical and theoretical aspects of joint bidding in these auctions.

<sup>27</sup> Although we control for time effects and a number of “neighborhood”-level observables, an interesting direction for future work would involve modeling bidders’ dynamic learning about the underlying spatial distribution of oil deposits. See, e.g., Covert (2015) and Hodgson (2019).

<sup>28</sup> Because this instrument has limited discrete support, our theoretical results suggest that we may rely on functional form to fill in the gaps between the points at which the index function  $\Gamma$  is nonparametrically identified.

TABLE 2  
SUMMARY STATISTICS

Variable	Mean	Median	Standard Deviation
No. of active neighbor leases	1.48	.00	1.96
Isolated-lease indicator	.59	1.00	.49
No. of firms that bid for neighbors	2.37	.00	3.56
Reoffered-tract indicator	.20	.00	.40
Neighbor-expired indicator	.31	.00	.46
No. of neighbor tracts drilled	1.69	1.00	2.25
No. of neighbor hits	.65	.00	1.30
Water depth (000 feet)	.20	.14	.28
Royalty rate (%)	16.10	16.00	1.35
No. of neighbor firms	.89	.00	1.27

B. *Baseline Model Estimates*

Here we report estimates of the entry model, index function, and joint distribution of equilibrium bids, using our baseline empirical specification with Bernstein polynomial approximation of order  $m = 4$ .<sup>29</sup> In general, we estimate separate joint distributions for each value of  $n_i$ . However, for large values of  $n_i$  we observe relatively few auctions (recall table 1), leading us to assume that all auctions with  $n_i \geq 11$  share the same marginal distribution of homogenized bids and the same copula correlation parameter. Combined with the ordered-probit specification of the entry model, this leads to a baseline specification with 151 parameters.

1. Entry Model

Table 3 shows our estimated entry-model parameters, with standard errors obtained from the spatial block bootstrap (we use 800 bootstrap replications throughout). As discussed in example 1 (see also app. A), the function  $\eta$  characterizes the effect of  $Z_i$  on entry but will not generally reveal the effects of  $X_i$ . Thus, one must interpret the estimated coefficients on  $X_i$  with caution. However, the coefficient on  $Z_i$  is positive (consistent with the prediction of our motivating example) and statistically significant, supporting its value in providing a source of variation in bidder entry.

2. Index Function and Bid Distribution

Table 3 also reports our estimated index parameters  $\gamma_x$  and  $\gamma_u$ . We again caution that coefficients on  $X_i$  need not have the usual causal

<sup>29</sup> Our choice of  $m$  reflects our experience with the trade-off between flexibility and precision using similar sample sizes in Monte Carlo simulations. We obtain similar results with slightly larger or smaller values of  $m$ , although the precision of the estimates declines as we add parameters.

TABLE 3  
ENTRY-MODEL AND INDEX-FUNCTION ESTIMATES

	ENTRY MODEL		INDEX FUNCTION	
	Estimate	Standard Error	Estimate	Standard Error
X:				
No. of active leases	-.089	.033	.016	.021
1(# neighbor firms = 0)	.518	.100	.042	.087
No. of firms that bid for neighbors	.016	.012	.029	.011
Reoffered-tract indicator	-.201	.075	-.167	.067
Neighbor-expired indicator	-.044	.079	-.280	.067
No. of neighbor tracts drilled (first ring)	-.046	.030	.049	.024
No. of neighbor hits (first ring)	.076	.027	-.022	.025
Depth	-.558	.283	-.224	.269
Depth squared	.120	.132	.092	.118
Royalty rate	-.002	.022	-.006	.017
Time controls	Sale year dummies		Sale year dummies	
Z: no. of neighbor firms	.1743	.0446		
U: unobserved heterogeneity			1.1363	.351

interpretations. However, the estimates indicate a strong effect of the unobserved heterogeneity on bids. Because  $U_i$  is normalized to be uniformly distributed, the coefficient implies that a 1-standard deviation increase in  $U_i$  drives up bids (and valuations) by roughly 33%.

We do not report estimates of the Bernstein polynomial parameters here. However, table 4 shows our estimates of the Gaussian copula correlation parameters  $\rho_n$ . The point estimates are positive, consistent with our assumption of positive dependence between bidders' private information. Wald tests fail to reject the hypothesis of equal copula correlation for all values of  $n$ ,<sup>30</sup> but strongly reject the null that all  $\rho_n$  are zero. Because homogenized bids are strictly increasing functions of signals, this implies rejection of the hypothesis of independent bidder types. This finding is of some importance on its own. Common-knowledge unobservables and correlated private information are two distinct phenomena with different implications for behavior and policy. Often, only one of these has been permitted in applications. Also important, however, is that the estimated correlation is generally small, suggesting modest correlation of bidders' private information.

<sup>30</sup> While not statistically distinguishable from the others, the estimated correlation parameter for the nine-bidder auctions stands out. A close examination of the data revealed no clear explanation for this. We do not use auctions with more than seven bidders in our tests for common values below, and results are similar when we omit auctions with  $n_i > 7$  from the estimation altogether.

TABLE 4  
COPULA CORRELATION ESTIMATES

	<i>n</i> =									
	2	3	4	5	6	7	8	9	10	11–18
Estimate	.044	.031	.089	.136	.112	.111	.133	.295	.161	.109
Standard error	.050	.040	.032	.041	.037	.043	.040	.067	.068	.020

C. *Decomposition of Correlation and Variance*

Bids in the same auction may be correlated because of correlated private information, auction observables, and auction unobservables. We present a decomposition of these factors in table 5, which also shows a decomposition of the overall variance of bids.<sup>31</sup> Figures in column 1, labeled “log  $B_{it}^0$ ,” are for the homogenized log bids. Here the pairwise correlation reflects the correlation among signals, the nonlinearity of the bidding strategy (and log transformation), and the fact that, all else equal, bid levels vary with the number of competitors in the auction. Similarly, the variance in this column reflects the variability in bidders’ assessments of tract values as well as variation in bidding strategies across auctions with different numbers of entrants.

A natural way to characterize the contributions of unobservables is with the correlation/variance arising from variation in  $\gamma(x, U_i)$  at a representative value of  $x$ . With our linear specification of  $\gamma$ , this variation is identical for all  $x$ . In column 2 of table 5, labeled “log  $B_{it}^0 + \gamma_u U_i$ ,” we add the contribution of auction-level unobservables. After accounting for the contributions of the log homogenized bids and unobservable  $U_i$ , all remaining correlation/variance in the log bids reflects auction observables. Again exploiting our linear specification of  $\gamma$ , column 3, labeled “log  $B_{it}^0 + \gamma_u U_i + X_i' \gamma_x$ ,” adds only the variation due to auction-level covariates. The final column, labeled “log  $B_{it}$ ,” adds the contribution of the year fixed effects.<sup>32</sup> Given our wide time span, it is not surprising that the fixed effects account for a substantial portion of the correlation and variance. More interesting is a comparison of the contributions of the auction-level covariates and the auction-level unobservables. The estimated contribution of the unobservables is roughly six times as large as that of the observed covariates. This is particularly noteworthy because we selected covariates  $X_i$  from an unusually rich set of observables in part on the basis

<sup>31</sup> We measure the “within-auction pairwise correlation” with the Pearson correlation coefficient.

<sup>32</sup> Unlike the rest of the decomposition, the order of these last two steps can matter, because of correlation between the covariates and fixed effects. Here, reversing the order has virtually no effect on the implied contributions to the within-auction correlation; however, it increases the contribution of the covariates to the bid variance: the impact of unobservables is then only three times as large as that of the observed covariates.

TABLE 5  
DECOMPOSITION OF LOG BID CORRELATION AND VARIANCE

	$\log B_i^0$ (1)	$\log B_i^0 + \gamma_u U_i$ (2)	$\log B_i^0 + \gamma_u U_i + X_i' \gamma_x$ (3)	$\log B_i$ (4)
Within-auction pairwise correlation	.133	.210	.222	.425
Variance	1.520	1.669	1.693	2.292

of explanatory power in descriptive analysis of bids. Unobserved heterogeneity could be even more important in applications where only a limited set of covariates is available.

VI. Tests for Common Values

A. Testing Approach

We use our estimates to test the null hypothesis of private values against the alternative of common values. We exploit the observation that, all else equal, adding competitors to a common values auction increases the severity of the winner’s curse.<sup>33</sup> The winner’s curse reflects “bad news” about a bidder’s valuation implied when his competitors’ signals are not sufficiently favorable for any of them to outbid him; the larger the number of such unfavorable signals, the worse the news. This effect is intuitive and can distinguish common values from private values, where there is no winner’s curse. Of course, this requires variation in the number of bidders that is not associated with changes in the underlying valuations or information structure that would mimic or reverse the effect on the winner’s curse. In our case, the following is a sufficient condition.

ASSUMPTION 11. For known  $\underline{n} \geq 2$ ,  $\bar{n} > \underline{n}$ , and all  $n = \underline{n}, \dots, \bar{n} - 1$ ,  $F_{SV^0}(S_{1t}, \dots, S_{nt}, V_{1t}^0, \dots, V_{nt}^0 | N_t = n) = F_{SV^0}(S_{1t}, \dots, S_{nt}, V_{1t}^0, \dots, V_{nt}^0 | N_t = n + 1)$ .

This is an assumption that, over a known range of auction sizes,  $U_i$  is the only latent source of dependence between the number of bidders and the valuations/signals.<sup>34</sup> Given this condition, Haile, Hong, and Shum (2003) showed that homogenized pivotal expected values  $w^p(S_{it}; n)$  are unaffected by  $n$  in a private-values auction but decreasing in  $n$  in a common-values auction (see also Athey and Haile 2002).

<sup>33</sup> We use the term “winner’s curse” to describe the adverse selection faced by a bidder competing against others with informative signals, not errors or regret on the part of a bidder.

<sup>34</sup> This condition is stronger than necessary. In our semiparametric specification, it implies (by invariance of copulas to monotonic transformations) that the copula correlation parameter  $\rho_n$  should be the same for  $n = \underline{n}, \dots, \bar{n}$ . Recall that we failed to reject the hypothesis  $\rho_n = \rho$  for all  $n$  and that differences in our point estimates of  $\rho_n$  were small over the range  $n = 2, \dots, 7$  considered below. Our results are very similar if we reestimate the model imposing the restriction.

To construct the tests, we use our estimates of the marginal distributions  $F_w(\cdot; n)$  of homogenized pivotal expected values conditional on  $N_i = n$ . We test the null hypothesis of equality (private values),

$$H_0: F_w(w; n) = F_w(w; n + 1) \quad \forall w, n = 2, \dots, 6,$$

against the one-sided alternative of first-order stochastic dominance (common values),

$$H_1: F_w(w; n) \leq F_w(w; n + 1) \quad \forall w, n = 2, \dots, 6,$$

with the inequality strict for at least some  $n$  and  $w$ . We limit attention to auctions with at most seven bidders in part to ensure that we have a sample of at least 100 auctions for each value of  $n$  considered.<sup>35</sup> An additional reason, however, is that growth in the severity of the winner's curse with the level of competition tends to diminish quickly as  $n$  grows. Intuitively, once a bidder assumes that  $n - 1$  others have low signals, learning that one additional signal is low conveys little "bad news" unless  $n$  is small. Thus, with common values, pivotal expected values are decreasing and, typically, convex in  $n$ .<sup>36</sup> This intuitive feature leads us to expect any evidence for common values to be clearest when comparing distributions at the lowest values  $n$  to those at higher levels of  $n$ .

We compare pairs of distributions using the one-sided Cramér-Von Mises-type statistic

$$\text{CVM} = \int_{-\infty}^{\infty} [\hat{F}_w(w; \mathbb{N}_1) - \hat{F}_w(w; \mathbb{N}_2)]_+^2 dw, \quad (13)$$

where  $[y]_+ = y \times 1\{y > 0\}$  and  $\hat{F}_w(w; \mathbb{N})$  is the estimated distribution of  $w^0(S_{it}; N_i)$  conditional on  $N_i$  lying in a range of values defined by a set  $\mathbb{N}$ . We focus primarily on sets  $\mathbb{N}$  containing two adjacent values of  $n$  ("coarse binning"). This pooling is done to reduce the impact of sampling error. However, by combining  $n = 2$  and  $n = 3$ , where we expect the largest change in the severity of the winner's curse, this pooling may hide the strongest evidence of common values. Thus, we also consider singleton sets ("fine binning"). In addition to pairwise tests, we construct a single test statistic for the full range  $n = 2, \dots, 7$ , based on the maximum (or smoothed maximum) over the pairwise statistics. Below, we report results for the pairwise and "max" tests for coarse binning as well as the max test for fine binning.<sup>37</sup>

<sup>35</sup> Abusing notation slightly, we let  $n$  (without an index) represent the number of bidders in a nonspecific auction rather than referring repeatedly to "the number of bidders."

<sup>36</sup> This convexity holds in all examples of symmetric common values auctions we are aware of. An interesting question is whether additional assumptions are needed to prove this as a general property.

<sup>37</sup> Appendix F provides complete results for fine binning. As expected, the strongest evidence for common values (a  $p$ -value of .001) comes from comparisons of  $n = 2$  to  $n = 3$ .



Haile, Hong, and Shum (2003) also observe that the union of the null and alternative hypotheses requires  $w^0(S_{it}; n)$  to be weakly decreasing in  $n$ . Because violation of this requirement would indicate rejection of at least one of our maintained hypotheses, this allows a specification test.<sup>38</sup> We implement this using the same type of test statistics, just reversing the direction of the one-sided alternative to the hypothesis of equal  $F_w(\cdot; n)$  across  $n$  (the least favorable relation under the null hypothesis of correct specification).

### *B. Results: Baseline Specification*

In figure 1A, we show the estimated CDFs under coarse binning, where we compare “low” ( $n \in \{2, 3\}$ ), “medium” ( $n \in \{4, 5\}$ ), and “high” ( $n \in \{6, 7\}$ ) levels of competition. Under the null, these distributions should differ only as a result of sampling error, whereas the alternative of common values implies that the CDFs will shift “northwest” as  $n$  increases. The estimated distributions shown here exhibit the stochastic ordering implied by the common values model. Further, as expected, the gap between the distributions for low and medium  $n$  is substantially larger than that between the distributions for medium and high  $n$ . Table 6 shows the relative sizes of these shifts (focusing on percentage changes) for the median homogenized pivotal expected values in each of the three bins.<sup>39</sup> As we shift from auctions with  $n \in \{2, 3\}$  to those with  $n \in \{4, 5\}$ , the median pivotal expected value falls by about 14%. When moving from  $n \in \{4, 5\}$  to  $n \in \{6, 7\}$ , the median falls by just under 3%.

Contrast the patterns in figure 1A with those in figure 1B, which shows the estimated distributions obtained when we estimate the model without allowing for unobserved heterogeneity. Here the results suggest stochastic ordering in the direction opposite that predicted by common values. This suggests a misspecified model.

Table 7 shows the  $p$ -values obtained from the formal tests. These results confirm what is suggested by the figures. First consider the comparison between low  $n$  and medium  $n$  when we allow unobserved heterogeneity (“With UH”). The test for common values implies rejection of private values in favor of common values, with a  $p$ -value of .021. The smaller gap between the estimated CDFs for medium and high  $n$  observed in figure 1 cannot be statistically distinguished. However, the max

<sup>38</sup> Another testable implication of the maintained hypotheses is that the inverse homogenized-bid functions—i.e., the right-hand side of the first-order conditions (12)—are strictly increasing as functions of the log homogenized bid (see Guerre, Perrigne, and Vuong 2000). We find no violation of this requirement, even before allowing for sampling error (see fig. 2 below.)

<sup>39</sup> Percentage changes in homogenized and nonhomogenized pivotal expected values are identical.

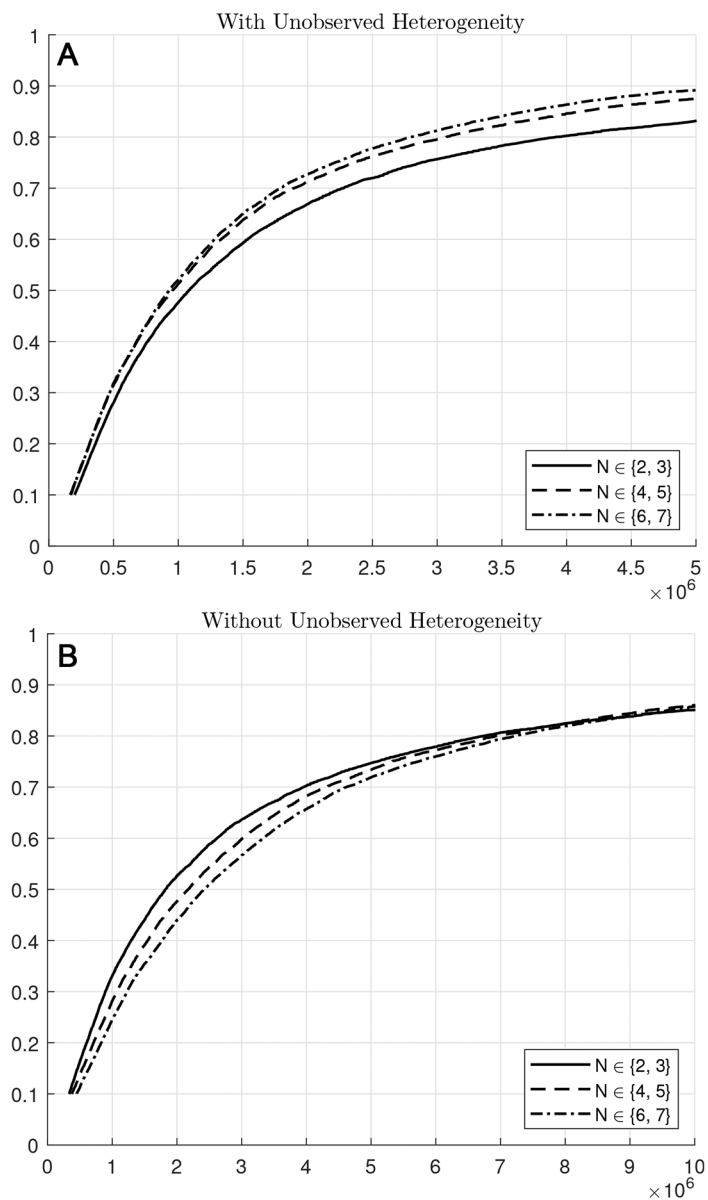


FIG. 1.—Test for common values, baseline specification: estimated cumulative distributions of homogenized pivotal expected values.

TABLE 6  
COMPETITION AND MEDIAN PIVOTAL EXPECTED VALUES

	$n \in \{2, 3\}$	$n \in \{4, 5\}$	$n \in \{6, 7\}$
Median homogenized pivotal expected value	1,067,371	915,437	890,112
Change (\$)		151,933	25,326
Change (%)		14.2	2.8

NOTE.—Median homogenized pivotal expected values are in 1982 dollars.

tests—for both coarse and fine binning—also imply rejection at significance levels around 2%. Consistent with figure 1, the specification test yields no evidence suggesting misspecification in the model with unobserved heterogeneity.

When we consider the model without unobserved heterogeneity (“No UH”), the conclusions are essentially reversed. Not only is there no evidence of common values, but the specification tests suggest misspecification, with all three coarse-binning tests yielding  $p$ -values below .10. This is intuitive but not a necessary implication of ignoring unobserved heterogeneity. When we ignore unobserved heterogeneity and endogenous entry, an “endogenous-treatment” bias works against the winner’s curse effect we are seeking to detect: auctions with more bidders may have a larger winner’s curse, but they also have more favorable unobservables. This suggests that the true effects of  $n$  on pivotal expected values could be masked or even reversed. But this intuition is incomplete. The model is misspecified when unobserved heterogeneity is present but ignored in the first-order conditions used to interpret the data. The cumulative distributions recovered in that case are not those of bidders’ pivotal expected values. Moreover, the direction of the misspecification bias is unclear, and this bias may

TABLE 7  
TEST  $P$ -VALUES: BASELINE SPECIFICATION

	With UH	No UH
	Test for Common Values	
{2,3} versus {4,5}	.021	.299
{4,5} versus {6,7}	.274	.754
Maximum (coarse binning)	.022	.590
Maximum (fine binning)	.019	.541
	Specification Test	
{2,3} versus {4,5}	.906	.066
{4,5} versus {6,7}	.901	.076
Maximum (coarse binning)	.985	.081
Maximum (fine binning)	.922	.239

NOTE.—UH = unobserved heterogeneity.

vary with  $n$ . Nonetheless, the results indicate that ignoring unobserved heterogeneity obscures the presence of common values in our sample.

### *C. Results under Alternative Specifications*

We have explored the same testing strategies under several alternative specifications. In particular, we have examined:

1. an extension of the bidding model allowing for a random reserve price, representing the MMS's option to reject all bids;
2. replacement of the ordered-probit specification of the entry model with a semi-nonparametric estimator, following Gallant and Nychka (1987);
3. dropping the year fixed effects from the index  $\gamma(x_i, u_i)$ , forcing unmeasured time-varying factors affecting the auctions into the unobservable; and
4. dropping all drainage tracts from the sample.

Detailed discussion and results are presented in appendix G. These alternative specifications lead to very similar patterns in the estimated distributions of pivotal expected values and to the same broad conclusions from the formal tests.

## **VII. Affiliation, Common Values, and Seller Revenue**

The presence of affiliation and common values can have important implications for auction design and for the way competition affects outcomes. Our estimates allow us to quantify some of these implications in the context of the OCS auctions.

### *A. Competition, Market Power, and Revenue*

A well-known but counterintuitive feature of common-value auctions is that added bidder competition can lead to less aggressive equilibrium bidding and even to reduced seller revenue.<sup>40</sup> The winner's curse is a key force behind this possibility, although Pinkse and Tan (2005) demonstrated that bids can decline with  $n$  as a result of affiliation of signals alone. Here we use our estimates to examine the equilibrium effects of bidder competition on bid shading, the level of bids, and seller revenue.

In figure 2 we plot, for different values of  $n$ , bidders' homogenized pivotal expected values against the associated homogenized bids implied by the first-order condition (7). Recall that pivotal expected values are strictly

<sup>40</sup> See, e.g., Laffont (1997) and Hong and Shum (2002).

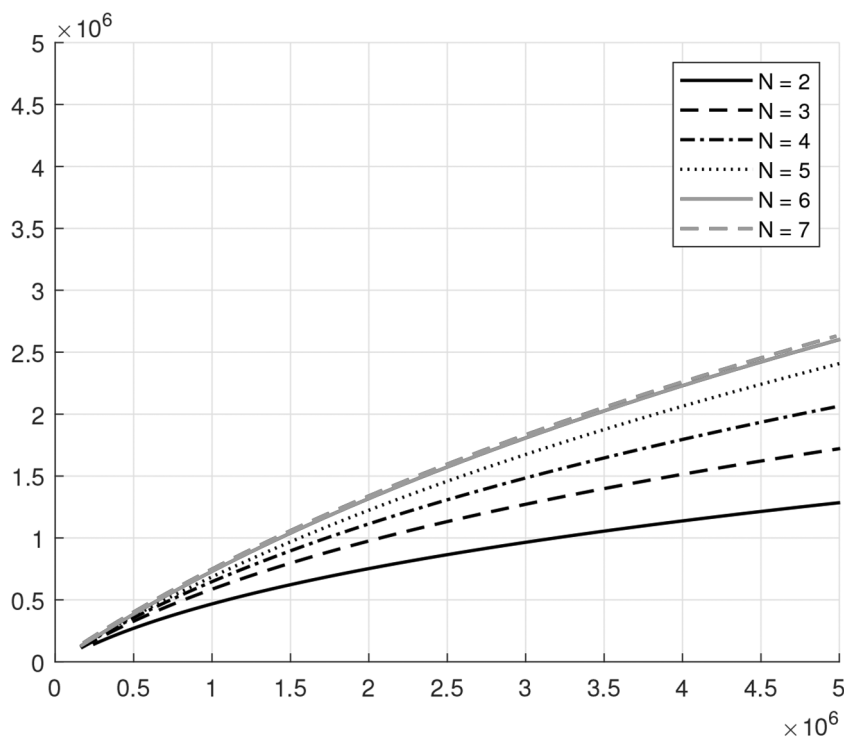


FIG. 2.—Competition and bid shading. Estimated pivotal expected values are on the horizontal axis, with the associated homogenized equilibrium bids on the vertical axis, both in 1982 dollars. A color version of this figure is available online.

increasing in bidder types (signals). We know from theory that all types above the lowest shade their bids below their pivotal expected values and that the degree of bid shading is increasing in type. Here we see that the estimated magnitude of this bid shading (reflecting bidders' market power) is substantial—all curves lie well below the 45° line. However, the gap shrinks as the level of competition rises from  $n = 2$  to  $n = 7$ .

In figure 3, we plot the estimated homogenized equilibrium bidding strategies, normalizing bidder signals to lie on  $[0, 1]$ . The estimated bid functions are strictly increasing, as implied by the model (but not imposed). They also generally increase with  $n$ —that is, the effect of more intense competition generally dominates the accompanying implications for the winner's curse and strategic responses to affiliation. Using our same one-sided testing strategy, we reject (with a 10% significance threshold) the null hypothesis of bid functions that are constant with respect to  $n$  in favor of strategies that increase with  $n$ . Our point estimates suggest that this monotonicity reverses at  $n = 7$ . Such a reversal is consistent with the folk wisdom that equilibrium bids “eventually” decrease with  $n$  in a

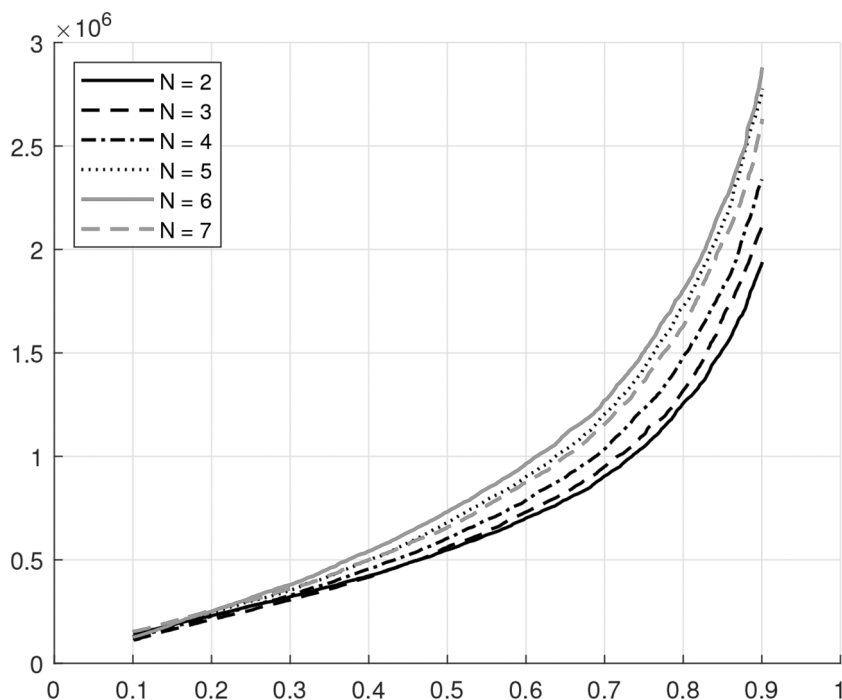


FIG. 3.—Competition and equilibrium bidding strategies. Bidder signals (normalized to  $[0, 1]$ ) are on the horizontal axis, with the associated homogenized equilibrium bids (in 1982 dollars) on the vertical axis. A color version of this figure is available online.

common values auction (see, e.g., Laffont 1997). However, even when comparing  $n = 6$  to  $n = 7$  in isolation, formal tests fail to reject the null that bids are weakly increasing with  $n$  at standard significance levels.

Finally, the second column (labeled “First-Price”) of table 8 shows, at each value of  $n$ , the implications of our estimates for a seller’s expected revenue. To put revenues in a natural scale, we use the median value of the estimated index  $\gamma(X_i, U_i)$  in our sample. Notably, even though our point estimates (fig. 3) suggested a nonmonotonicity of bidding with respect to  $n$ , this effect is overcome by the fact that when  $n$  is larger the winning bid is the maximum among a larger number of bids. Thus, even our point estimates give no indication that a seller would profit from restricting entry to reduce the severity of the winner’s curse faced by bidders.

#### *B. The Linkage Principle and Revenue Rankings*

Milgrom and Weber (1982) and Milgrom (1987) identified the “linkage principle” as a key force determining a seller’s preference among standard auction formats and information revelation policies in an affiliated-values setting. Loosely, the linkage principle states that the information

TABLE 8  
REVENUE GAINS THROUGH THE LINKAGE PRINCIPLE

<i>n</i>	EXPECTED REVENUE		GAIN (%)	STANDARD ERROR (%)
	First-Price	Second-Price		
2	7.02	7.05	.48	1.58
3	9.21	9.31	1.06	2.05
4	11.48	11.79	2.73	1.64
5	14.95	15.64	4.62	1.19
6	16.91	17.51	3.57	1.41
7	16.94	17.43	2.89	1.44
Overall	15.17	15.62	3.04	.49

NOTE.—Simulated expected revenues, in millions of 1982 dollars, based on 1 million simulated auctions of each size  $n$ , scaled by the estimated median value of the index  $\gamma(X_i, U_i)$  in our sample. Bootstrap standard errors are based on 800 replications.

rents obtained by bidders can be limited (and the surplus extracted by the seller enhanced) by linking the price the winning bidder pays to realizations of random variables that are outside the winner's control but affiliated with the winner's private information. Moving from a first-price sealed-bid auction to a second-price sealed-bid auction, for example, enhances expected revenue by letting the second-highest bid determine the price: this bid is a function of the second-highest signal, which is affiliated with the winner's signal.<sup>41</sup> However, despite the central theoretical role of the linkage principle and implied "revenue ranking" results, the existing literature provides few opportunities to quantify this effect empirically.<sup>42</sup> We know of no such evaluation in the case of OCS auctions.

The third column of table 8 presents simulated expected revenues for second-price sealed-bid auctions implied by our estimates.<sup>43</sup> We report results at each value of  $n$  from 2 to 7, as well as an overall average obtained by mixing over the full range of  $N_i$  in our sample according to its empirical distribution. Although the results exhibit the revenue ranking implied by the theory, the estimated revenue gains from switching to a second-price auction are modest, with overall gains of 3% (although still around \$450,000 per auction in 1982 dollars). This is consistent with our finding that correlation among bidders' private signals is relatively weak.

As with the theoretical analysis of Milgrom and Weber (1982) and Milgrom (1987), an important caveat is that we have held bidder entry behavior fixed in our simulation analysis. In reality, changes in auction

<sup>41</sup> Similarly, the MMS's use of royalties, which link the total price paid to the realized value of the tract, can be justified by the linkage principle—at least in the absence of ex post moral hazard (see, e.g., Riley 1988 and DeMarzo, Kremer, and Skrzypacz 2005).

<sup>42</sup> Shneyerov (2006) has done so for a sample of municipal bond auctions.

<sup>43</sup> Equilibrium revenue in a second-price auction equals the second-highest pivotal expected value.

design that reduce bidder surplus at a given level of competition should discourage entry, working against the anticipated gains in revenue. Quantifying the equilibrium effect would require estimation of a structural model of bidder entry (see, e.g., the extension discussed in app. I). However, because we find that even the gains considered by Milgrom and Weber (1982) and Milgrom (1987) would be fairly small, our results are quite informative. At least in the case of these auctions, any gains from moving to a second-price auction to better exploit the linkage principle appear to be limited.

### VIII. Conclusions, Caveats, and Extensions

We proposed an empirical approach to first-price sealed-bid auctions with affiliated values, unobserved auction-level heterogeneity, and endogenous bidder entry. Applying our method to OCS auction data led us to reject the private-values model in favor common values, a conclusion that is robust across a variety of specifications. We found that ignoring unobserved heterogeneity can hide the presence of common values and that models without unobserved heterogeneity can be rejected. Despite the presence of affiliated signals and common values, however, we found that the seller would not gain from limiting competition to soften the winner's curse and that any revenue gain from moving to a second-price sealed-bid auction to better exploit the linkage principle would likely be small.

Although our empirical approach offers several advantages for our study of OCS auctions, it relies on assumptions that will not be suitable for all applications or questions. This is both an important caveat and a call for further work. In appendix I, we discuss some initial extensions, challenges, and directions for further exploration. There we discuss (1) obtaining identification of a structural model of bidder entry, permitting examination of additional counterfactual questions; (2) extending our methods to some types of models permitting bidder asymmetry; (3) allowing a binding public reserve price; (4) testing of additional overidentifying restrictions; and (5) making use of data on realized quantities of oil production. Even where we are able to offer some initial results on these topics, our exploration is necessarily preliminary, leaving important open questions for future work.

Finally, although we obtained a useful form of partial identification for a common values model, the full set of primitives in common values models—even without unobserved heterogeneity—generally is not point identified without assumptions and data beyond those we required. Thus, it will be important to continue pursuit of approaches to identification that exploit the features of particular settings (see, e.g., Hendricks, Pinkse, and Porter 2003 or Somaini 2020) and to explore extensions permitting



unobserved heterogeneity in those frameworks. It may also prove productive to pursue other forms of partial identification that can be used to address positive and normative questions. The recent work of Syrgkanis, Tamer, and Ziani (2018) provides one such approach.

## Appendix A

### Equilibrium Entry in a Model of OCS Auctions

Here we expand on the sketch provided in example 1 to consider a particular extensive-form game of entry and bidding that is motivated by our application and yields an entry outcome satisfying the reduced-form equation (3) and our assumed weak monotonicity conditions. This example also demonstrates how our model can accommodate auction-specific unobservables that are of arbitrary dimension and correlated with auction-specific observables, despite the apparent contradiction to our assumption that  $U_i$  is a scalar and independent of  $X_i$ . Accommodation of such correlation requires that we allow the interpretation of  $U_i$  to vary with the vector  $X_i$ . This precludes identification of (causal) effects of covariates on the auction; but in typical auction applications, auction-level observables are primarily confounding factors to be controlled for rather than factors whose effects are of direct interest. This section also motivates the instrument used in our application.

#### A1. Model

Consider a game of entry and bidding for the lease of a tract  $t$ . Let  $\mathcal{I}$  denote the set of all potential bidders ("firms"), with  $\mathcal{Z}_i \subset \mathcal{I}$  denoting the set of "neighbor firms"—holders of active leases on adjacent ("neighbor") tracts. Let  $I = |\mathcal{I}|$  and  $Z_i = |\mathcal{Z}_i|$ . Let  $V_{it}$  denote the value of the lease to firm  $i$  ( $i$ 's "valuation"). Let  $X_i$  and  $E_i$  denote, respectively, observed and unobserved (to us) characteristics of lease  $t$  that affect bidders' valuations. Let  $X_i$  include (among other relevant characteristics) the number of active leases on neighbor tracts and the set of bidders for each of those leases.<sup>44</sup> We make no restriction on the dimension of  $E_i$  and do not require independence between  $X_i$  and  $E_i$ .

The game consists of two stages. In the first stage, firms simultaneously choose whether to incur an entry cost in order to acquire a signal and participate in the auction. Let  $c_i(x_i)$  denote the entry cost for firm  $i$ . Neighbors have lower entry costs. In particular,  $c_i(x_i) = c(x_i)$  for a neighbor firm, whereas nonneighbor firms have entry costs  $c_i(x_i) = c(x_i) + \delta(x_i)$ , with  $\delta(x_i) > 0$ .<sup>45</sup> Let  $S_{it}$  denote the signal received by firm  $i$ . Firms acquiring signals become "bidders" and learn the number of competitors they face.

<sup>44</sup> In practice, we represent the set of bidders for neighboring tracts more parsimoniously with the number of such bidders.

<sup>45</sup> More generally, one can allow entry costs to depend on the instruments, writing  $c(x_i, z_i)$  and  $\delta(x_i, z_i)$ . For example, in other applications one might have measures of signal acquisition costs that vary across time or location. However, our discussion of sample selection in app. C exploits the exclusion of  $Z_i$  from entry costs.

In the second stage, the lease is offered to bidders by first-price auction with no binding reserve price.<sup>46</sup> Let  $N_i$  denote the number of bidders. Given  $N_i = n$ , let  $S_i = (S_{i1}, \dots, S_{in})$  and  $V_i = (V_{i1}, \dots, V_{in})$ , where without loss we relabel bidders as firms  $i = 1, \dots, n$ . For any conditioning set  $\Omega \subseteq (X_i, Z_i, E_i)$ , let  $F_{SV}(S_i, V_i|N_i, \Omega)$  denote the conditional distribution of bidders' signals and valuations. We assume that  $F_{SV}(S_i, V_i|N_i, X_i, Z_i, E_i)$  satisfies standard smoothness, symmetry, affiliation, and nondegeneracy conditions (see assumption 2 in sec. II).

We assume that  $Z_i$  alters the joint distribution of signals and valuations only through its effect on  $N_i$ —that is, that  $F_{SV}(S_i, V_i|N_i, Z_i, X_i, E_i) = F_{SV}(S_i, V_i|N_i, X_i, E_i)$  and that  $Z_i$  is independent of  $E_i$  conditional on  $X_i$ . We discuss the justification for this conditional independence assumption below. We assume

$$V_{it} = V_{it}^0 \lambda(X_i, E_i), \quad (\text{A1})$$

where the function  $\lambda$  is positive and the random variables  $(V_{i1}^0, \dots, V_{in}^0, S_{i1}, S_{in})$  are independent of  $(X_i, E_i, Z_i)$  conditional on  $N_i = n$ . We assume that, for all  $x \in \mathbb{X}$ ,  $\lambda(x, E_i)$  has a continuous distribution and convex bounded support.

Note that we have not restricted the dimension of  $E_i$ , imposed any monotonicity condition on  $\lambda$ , or required independence between  $X_i$  and  $E_i$ .<sup>47</sup> Nonetheless, we can obtain the model of unobserved heterogeneity in the text by representing the random variable  $\lambda(X_i, E_i)$  in terms of its quantiles conditional on  $X_i$ . In particular, given  $X_i = x$ , let  $F_\lambda(\cdot|x)$  denote the CDF of the random variable  $\lambda(x, E_i)$ , and let

$$U_i = F_\lambda(\lambda(x, E_i)|x). \quad (\text{A2})$$

For  $u \in [0, 1]$ , define  $F_\lambda^{-1}(u|x) = \inf\{\lambda : F_\lambda(\lambda|x) \geq u\}$ , and let

$$\Gamma(x, u) = F_\lambda^{-1}(u|x). \quad (\text{A3})$$

Combining equations (A2) and (A3), for each  $x$  we have  $F_\lambda(\Gamma(x, U_i)|x) = U_i = F_\lambda(\lambda(x, E_i)|x)$ , that is,  $\Gamma(x, U_i) = \lambda(x, E_i)$ . By construction,  $\Gamma$  is strictly increasing (and continuous) in its second argument, and  $U_i$  is uniform on  $[0, 1]$  conditional on  $X_i$ . Because  $U_i$  is a measurable function of  $E_i$  conditional on  $X_i$ ,  $U_i$  is independent of  $Z_i$ , given  $X_i$ .

Note that in this new representation of the model, the distribution of  $U_i$  does not vary with  $X_i$ , but its interpretation generally will. Because  $\Gamma(x_i, u_i) = \lambda(x_i, e_i)$  for all  $i$  by construction,  $\Gamma(X_i, U_i)$  fully characterizes the variation and dependence in valuations and bids that arises from the observables and unobservables. Likewise, controlling for the value of  $\Gamma(X_i, U_i)$  fully controls for the effects of auction observables and unobservables  $(X_i, E_i)$  on valuations, bids, and equilibrium

<sup>46</sup> As is standard in the literature, we assume that only firms incurring the entry cost can submit a bid (see, e.g., Levin and Smith 1994; Li and Zheng 2009; Athey, Levin, and Seira 2011; Krasnokutskaya and Seim 2011; Bhattacharya, Roberts, and Sweeting 2014; or Gentry and Li 2014).

<sup>47</sup> Accommodating dependence may be important, as the nomination process by which tracts were offered for lease in our sample period suggests that a tract with an “undesirable” value of  $X_i$  may have been unlikely to be offered unless the value of  $E_i$  made the tract desirable. See, e.g., Hendricks, Porter, and Boudreau (1987) for a discussion of the nomination process.

first-order conditions. However,  $\Gamma$  does not characterize the effect of a change in  $X_i$  holding unobservables fixed, since our  $U_i$  is redefined at every value of  $X_i$ .

## A2. *Equilibrium*

We henceforth use the representation of the model just derived. The set of firms  $\mathcal{I}$ , the rules of the game, the values of  $(X_i, U_i, Z_i)$ , and the joint distribution  $F_{SV}(S_i, V_i|N_i, X_i, U_i)$  are common knowledge among firms. We consider perfect Bayesian equilibrium in pure strategies, with weakly increasing strategies in the auction stage.

The second stage of the game is identical to the first-price sealed-bid auction with symmetric affiliated values studied by Milgrom and Weber (1982), who characterize the unique Bayesian Nash equilibrium in increasing bidding strategies. Bidder  $i$ 's payoff in the auction stage can be written as a function of the commonly known  $(N_i, X_i, U_i)$  and the realized bidder signals  $S_i$ . As noted in the text, multiplicative separability of valuations is inherited by equilibrium bids. This implies that a bidder's ex post profit, denoted by  $\pi(S_{ii}, S_{-ii}, N_i, X_i, U_i)$ , is strictly increasing in the index  $\Gamma(X_i, U_i)$  and, therefore, strictly increasing in  $U_i$ . Further, we assume the usual case in which the ex ante expected equilibrium payoff  $\bar{\pi}(N_i, X_i, U_i) \equiv E[\pi(S_{ii}, S_{-ii}, N_i, X_i, U_i)|N_i, X_i, U_i]$  is strictly decreasing in  $N_i$ .<sup>48</sup>

In the entry stage, firms make decisions based on the cost of entry and expected profit from participating in the auction. Let  $C_{ii} = c_i(X_i)$ . For firm  $i$ , entering when  $n - 1$  other firms will also enter implies expected profit  $\bar{\pi}(n, X_i, U_i) - C_{ii}$ .

Conditional on  $(X_i, Z_i, U_i)$ , and given equilibrium beliefs about payoffs in the auction stage, the entry stage is then equivalent to the entry game in Berry (1992). Berry showed that a pure-strategy equilibrium exists and that with probability 1 all equilibria exhibit the same number of entrants, given by

$$\eta(X_i, Z_i, U_i) = \max_{0 \leq n \leq I} \{n : \bar{\pi}(n, X_i, U_i) - C_{ii} \geq 0\}.$$

Recall that  $(X_i, Z_i)$  determine the values of  $\{C_{ii}\}_{i \in \mathcal{I}}$ . Thus, in any pure-strategy perfect Bayesian equilibrium (with weakly increasing bidding) we have

$$N_i = \eta(X_i, Z_i, U_i).$$

Because  $\bar{\pi}(N_i, X_i, U_i)$  is strictly increasing in  $U_i$ ,  $\eta$  is weakly increasing in  $U_i$ .

## A3. *The Instrument*

Our instrument for bidder entry  $Z_i$  is the number of neighbor firms. First, consider the exclusion requirement (assumption 5, in sec. II). We have assumed directly that  $Z_i$  is independent of  $(S_i, V_i^0)$  conditional on  $N_i$ , that is, that  $X_i$  are the only observables directly affecting bidder valuations.<sup>49</sup> However, we must verify

<sup>48</sup> We know of no counterexample to strict monotonicity in  $N_i$  under the assumption that  $U_i$  is the only latent source of dependence between the entry and auction stages—i.e., that assumption 11 holds. Nonmonotonicity (within the relevant range of  $N_i$ ) could lead to existence of multiple equilibria with different numbers of bidders.

<sup>49</sup> This could fail here if the number of neighbor firms had a direct effect on tract value (given  $X_i$ ), e.g., by driving up costs of negotiating production from common pools.

that  $Z_i$  is also independent of  $U_i$  conditional on  $X_i$ . A tract with three neighbor leases, for example, may have one, two, or three neighbor firms, depending on which bidders for the neighboring leases won those auctions. Given the number of neighbor leases and the bidders for each neighbor tract (i.e., conditional on  $X_i$ ), the number of distinct winners reflects only random variation in bidders' signals at prior auctions. Recall that signals are assumed independent of tract-specific unobserved heterogeneity and independent across tracts. Thus, even in the case of spatially correlated tract-level unobservables  $E_i$ , the conditional independence requirement  $Z_i \perp\!\!\!\perp U_i | X_i$  will hold.

Regarding the “relevance” requirement for the instrument  $Z_i$ ,<sup>50</sup> observe that changes in the number of neighbor firms affects entry because, for some combinations of  $(X_i, Z_i, U_i)$ , the market will accommodate the  $(n + 1)$ th entrant only if there is a potential bidder with low signal acquisition cost. For example, we will sometimes have two entrants because the market would support entry by a third (low-cost) neighbor, but not by a third firm that is a (high-cost) nonneighbor. Thus, larger values of  $Z_i$  will lead, all else equal, to weakly larger numbers of entrants.

## Appendix B

### Proofs Omitted from the Text

#### B1. Proof of Corollary 1

We can express  $\Pr(U_i \leq u | X_i = x, N_i = n)$  as

$$F_{U|XN}(u|x, n) = \int F_{U|XZN}(u|x, z, n) d\zeta(z|x, n), \quad (B1)$$

where  $F_{U|XZN}$  is the distribution of  $U_i | (X_i, Z_i, N_i)$  and  $\zeta$  is the distribution of  $Z_i | (X_i, N_i)$ . Conditional on  $N_i = n$ ,  $Z_i = z$ , and  $X_i = x$ ,  $U_i$  is uniform on  $[\tau_{n-1}(x, z), \tau_n(x, z)]$ , and by theorem 1 the end points  $\tau_{n-1}(x, z)$  and  $\tau_n(x, z)$  are identified. So  $F_{U|XZN}$  is known. Since  $\zeta$  is directly observed, the result follows from equation (B1). QED

#### B2. Proof of Lemma 2

*Step 1.*—We first show that for all  $n \geq \underline{n}$ , all  $(x, z) \in \mathbb{Y}(n)$ , and all  $(x', z') \in \mathbb{Y}(n)$ ,  $\gamma(x, \tau_n(x, z)) - \gamma(x', \tau_{n-1}(x', z'))$  is identified. For  $n^*$  as defined in assumption 8, take  $n \leq n^*$ , and let  $x(n)$ ,  $z(n)$ , and  $\hat{z}(n)$  be as in part i of assumption 8, so that

$$\begin{aligned} \underline{n}(x(n), z(n)) &= n, \\ \underline{n}(x(n), \hat{z}(n)) &= n + 1. \end{aligned} \quad (B2)$$

Since  $(x(n), z(n)) \in \mathbb{Y}(n)$  and  $(x(n), \hat{z}(n)) \in \mathbb{Y}(n + 1)$ , lemma 1 implies identification of

$$\gamma(x', \tau_{n-1}(x', z')) - \gamma(x(n), \tau_{n-1}(x(n), z(n))) \quad (B3)$$

<sup>50</sup> We state formal “relevance” conditions for nonparametric identification in sec. III.

and

$$\gamma(x'', \tau_n(x'', z'')) - \gamma(x(n), \tau_n(x(n), \hat{z}(n))) \quad (\text{B4})$$

for all  $(x', z') \in \mathbb{Y}(n)$  and  $(x'', z'') \in \mathbb{Y}(n+1)$ . By equations (4) and (B2),

$$\tau_{n-1}(x(n), z(n)) = 0 = \tau_n(x(n), \hat{z}(n)),$$

so subtracting (B3) from (B4) yields identification of

$$\gamma(x'', \tau_n(x'', z'')) - \gamma(x', \tau_{n-1}(x', z')) \quad (\text{B5})$$

for all  $(x'', z'') \in \mathbb{Y}(n+1)$  and  $(x', z') \in \mathbb{Y}(n)$ . By assumption 7, there exists some  $(x'', z'')$  that is in both  $\mathbb{Y}(n+1)$  and  $\mathbb{Y}(n)$ . The claim (for  $n \leq n^*$ ) then follows from lemma 1. A symmetric argument applies for  $n > n^*$ .

*Step 2.*—To complete the proof, we proceed by induction, starting with  $n = \underline{n}$ . By the normalization (1) and our choice of  $x^0$ , we have  $\gamma(x^0, 0) = 0$  and  $(x^0, z) \in \mathbb{Y}(\underline{n})$  for some  $z$ . Lemma 1 then implies identification of  $\gamma(x, \tau_{\underline{n}-1}(x, z))$  for all  $(x, z) \in \mathbb{Y}(\underline{n})$ . Step 1 above then implies identification of  $\gamma(x, \tau_{\underline{n}}(x, z))$  for all  $(x, z) \in \mathbb{Y}(\underline{n})$ . Now take any  $n > \underline{n}$  and suppose that  $\gamma(x, \tau_{n-1}(x, z))$  is known for all  $(x, z) \in \mathbb{Y}(n-1)$ . By assumption 7 there exists a point  $(\tilde{x}, \tilde{z})$  in  $\mathbb{Y}(n-1) \cap \mathbb{Y}(n)$ . Since we have already identified  $\gamma(\tilde{x}, \tau_{n-1}(\tilde{x}, \tilde{z}))$ , by lemma 1 we also know the value of  $\gamma(x, \tau_{n-1}(x, z))$  for all  $(x, z)$  in  $\mathbb{Y}(n)$ . By step 1, this implies identification of  $\gamma(x, \tau_n(x, z))$  for all  $(x, z)$  in  $\mathbb{Y}(n)$ . QED

### B3. Proof of Lemma 3

Take any  $n, x$ , and  $z$  such that  $\tau_{n-1}(x, z) \in (0, 1)$ . Let  $\tau = \tau_{n-1}(x, z)$ , and let  $\nu > 0$  be sufficiently small that  $\tau + \nu < 1$  and  $\tau - \nu > 0$ . We show that for any such  $\nu$  there exists  $\epsilon > 0$  such that for every  $z'$  satisfying  $\|z' - z\| < \epsilon$  we have  $\tau_{n-1}(x, z') \in (\tau - \nu, \tau + \nu)$ . Let  $\delta = \nu/2$ . By the definition of  $\tau_{n-1}(x, z)$  and weak monotonicity of  $\eta$  in  $U$ ,  $\eta(x, z, \tau - \delta) < n$ . So by assumption 9 there exists  $\epsilon_1 > 0$  such that for any  $z'$  satisfying  $\|z' - z\| < \epsilon_1$ ,  $\eta(x, z', \tau') < n$  for some  $\tau' \in (\tau - 2\delta, \tau)$ . Similarly, because  $\eta(x, z, \tau + \delta) \geq n$ , there exists  $\epsilon_2 > 0$  such that for any  $z'$  satisfying  $\|z' - z\| < \epsilon_2$ ,  $\eta(x, z', \tau'') \geq n$  for some  $\tau'' \in (\tau, \tau + 2\delta)$ . Letting  $\epsilon = \min\{\epsilon_1, \epsilon_2\}$ , we have shown that for any  $z'$  satisfying  $\|z' - z\| < \epsilon$ ,  $\eta(x, z', \tau') < n$  for some  $\tau' \in (\tau - \nu, \tau)$ , while  $\eta(x, z', \tau'') \geq n$  for some  $\tau'' \in (\tau, \tau + \nu)$ . Because  $\tau_{n-1}(x, z')$  must then lie in  $[\tau', \tau'']$ , the result follows. QED

### B4. Proof of Lemma 4

Fix  $N_i = n$  and  $X_i = x$ . Let  $\psi_{\bar{B}}$  denote the characteristic function of the log bids  $(\bar{B}_{1i}, \dots, \bar{B}_{ni})$  conditional on  $X_i = x$  and  $N_i = n$ . By assumption 3,  $\gamma(x, U_i)$  is bounded. This implies that the characteristic function of  $\gamma(x, U_i)$ , denoted  $\psi_\gamma$ , is nonzero almost everywhere.<sup>51</sup> By equation (8), for  $(r_1, \dots, r_n) \in \mathbb{R}^n$  we have  $\psi_B(r_1, \dots, r_n) = \psi_{B^0}(r_1, \dots, r_n)\psi_\gamma(r_1 + \dots + r_n)$ , where  $\psi_{B^0}$  is the characteristic function of the log homogenized bids  $(\bar{B}_{1i}^0, \dots, \bar{B}_{ni}^0)$  conditional on  $N_i = n$ . Since the

<sup>51</sup> See, e.g., the proof of theorem 1 in Krasnokutskaya (2011).

distribution of  $U_i|(X_i, N_i)$  is known (corollary 1) and  $\gamma$  is a known function (theorem 2),  $\psi_\gamma$  is known. So the equation

$$\psi_{B'}(r_1, \dots, r_n) = \frac{\psi_B(r_1, \dots, r_n)}{\psi_\gamma(r_1 + \dots + r_n)}$$

uniquely determines  $\psi_{B'}(r_1, \dots, r_n)$  for almost all  $(r_1, \dots, r_n)$ . By continuity of characteristic functions this yields identification of  $\psi_{B'}$ , implying identification of the joint density of  $(\tilde{B}_{1t}, \dots, \tilde{B}_{nt})$ . The result then follows. QED

## References

- Akerlof, George A. 1970. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Q.J.E.* 84:488–500.
- Aradillas-López, Andres, Philip A. Haile, Kenneth Hendricks, and Robert H. Porter. 2019. "Testing Competition in U.S. Offshore Oil and Gas Lease Auctions." Working paper, Pennsylvania State Univ.
- Armstrong, Timothy B. 2013. "Bounds in Auctions with Unobserved Heterogeneity." *Quantitative Econ.* 4:377–415.
- Arrow, Kenneth. 1970. "Political and Economic Evaluation of Social Effects and Externalities." In *The Analysis of Public Output*, edited by Julius Margolis, 1–30. New York: Columbia Univ. Press (for NBER).
- Athey, Susan, and Philip A. Haile. 2002. "Identification of Standard Auction Models." *Econometrica* 70:2107–40.
- . 2006. "Empirical Models of Auctions." In *Advances in Economics and Econometrics, Theory and Applications: Ninth World Congress of the Econometric Society*, vol. 2, edited by Richard Blundell, Whitney K. Newey, and Torsten Persson, 1–45. Econometric Soc. Monographs, no. 42. Cambridge: Cambridge Univ. Press.
- . 2007. "Nonparametric Approaches to Auctions." In *Handbook of Econometrics*, vol. 6A, edited by James J. Heckman and Edward E. Leamer, 3847–965. Amsterdam: North-Holland.
- Athey, Susan, Jonathan Levin, and Enrique Seira. 2011. "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions." *Q.J.E.* 126:207–57.
- Bajari, Patrick, and Ali Hortaçsu. 2003. "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions." *RAND J. Econ.* 34:329–55.
- Balat, Jorge. 2017. "Highway Procurement and the Stimulus Package: Identification and Estimation of Dynamic Auctions with Unobserved Heterogeneity." Discussion paper, Univ. Texas.
- Berry, Steven T. 1992. "Estimation of a Model of Entry in the Airline Industry." *Econometrica* 60 (4): 889–917.
- Bhattacharya, Vivek, James W. Roberts, and Andrew Sweeting. 2014. "Regulating Bidder Participation in Auctions." *RAND J. Econ.* 45:675–704.
- Campo, Sandra, Emmanuel Guerre, Isabelle M. Perrigne, and Quang Vuong. 2011. "Semiparametric Estimation of First-Price Auctions with Risk-Averse Bidders." *Rev. Econ. Studies* 78:112–47.
- Campo, Sandra, Isabelle M. Perrigne, and Quang Vuong. 2003. "Asymmetry in First-Price Auctions with Affiliated Private Values." *J. Appl. Econometrics* 18:179–207.
- Chen, Xiaohong, Yanqin Fan, and Viktor Tsyrennikov. 2006. "Efficient Estimation of Semiparametric Multivariate Copula Models." *J. American Statis. Assoc.* 101:1228–40.

- Chen, Xiaohong, and Zhipeng Liao. 2014. "Sieve  $M$  inference on Irregular Parameters." *J. Econometrics* 182:70–86.
- Chen, Xiaohong, Zhipeng Liao, and Yixiao Sun. 2014. "Sieve Inference on Possibly Misspecified Semi-nonparametric Time Series Models." *J. Econometrics* 178:639–58.
- Chen, Xiaohong, and Demian Pouzo. 2009. "Efficient Estimation of Semiparametric Conditional Moment Models with Possibly Nonsmooth Residuals." *J. Econometrics* 152:46–60.
- Covert, Thomas R. 2015. "Experiential and Social Learning in Firms: The Case of Hydraulic Fracturing in the Bakken Shale." Discussion paper, Univ. Chicago.
- DeMarzo, Peter M., Ilan Kremer, and Andrzej Skrzypacz. 2005. "Bidding with Securities: Auctions and Security Design." *A.E.R.* 95:936–59.
- Deneckere, Raymond, and Meng-Yu Liang. 2006. "Bargaining with Interdependent Values." *Econometrica* 74:1309–64.
- D'Haultfoeuille, Xavier, and Philippe Février. 2015. "Identification of Mixture Models Using Support Variations." *J. Econometrics* 189:70–82.
- Gallant, A. Ronald, and Douglas W. Nychka. 1987. "Semi-nonparametric Maximum Likelihood Estimation." *Econometrica* 55:363–90.
- Gentry, Matthew, and Tong Li. 2014. "Identification in Auctions with Selective Entry." *Econometrica* 82:315–44.
- Gillen, Ben. 2010. "Identification and Estimation of Level- $k$  Auctions." Discussion paper, California Inst. Technology.
- Gilley, Otis W., and Gordon V. Karels. 1981. "The Competitive Effect in Bonus Bidding: New Evidence." *Bell J. Econ.* 12:637–48.
- Guerre, Emmanuel, Isabelle M. Perrigne, and Quang Vuong. 2000. "Optimal Nonparametric Estimation of First-Price Auctions." *Econometrica* 68:525–74.
- . 2009. "Nonparametric Identification of Risk Aversion in First-Price Auctions under Exclusion Restrictions." *Econometrica* 77:1193–227.
- Haile, Philip A., Kenneth Hendricks, and Robert H. Porter. 2010. "Recent U.S. Offshore Oil and Gas Lease Bidding: A Progress Report." *Internat. J. Indus. Org.* 28:390–96.
- Haile, Philip A., Han Hong, and Matthew Shum. 2003. "Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions." Discussion Paper no. 1445, Cowles Found., New Haven, CT.
- Haile, Philip A., and Yuichi Kitamura. 2019. "Unobserved Heterogeneity in Auctions." *Econometrica J.* 22:C1–C19.
- Hendricks, Kenneth, Joris Pinkse, and Robert H. Porter. 2003. "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions." *Rev. Econ. Studies* 70:115–45.
- Hendricks, Kenneth, and Robert H. Porter. 1988. "An Empirical Study of an Auction with Asymmetric Information." *A.E.R.* 78:865–83.
- . 1992. "Joint Bidding in Federal OCS Auctions." *A.E.R.* 82 (2): 506–11.
- Hendricks, Kenneth, Robert H. Porter, and Bryan Boudreau. 1987. "Information, Returns, and Bidding Behavior in OCS Auctions: 1954–1969." *J. Indus. Econ.* 35:517–42.
- Hendricks, Kenneth, Robert H. Porter, and Richard H. Spady. 1989. "Random Reservation Prices and Bidding Behavior in OCS Drainage Auctions," *J. Law and Econ.* 32:S83–S106.
- Hendricks, Kenneth, Robert H. Porter, and Guofo Tan. 2008. "Bidding Rings and the Winner's Curse." *RAND J. Econ.* 39:1018–41.
- Hill, Jonathan B., and Artyom Shneyerov. 2014. "Are There Common Values in First-Price Auctions? A Tail-Index Nonparametric Test." *J. Econometrics* 174:144–64.



- Hodgson, Charles. 2019. "Information Externalities, Free Riding, and Optimal Exploration in the UK Oil Industry." Discussion paper, Yale Univ.
- Hong, Han, and Matthew Shum. 2002. "Increasing Competition and the Winner's Curse: Evidence from Procurement." *Rev. Econ. Studies* 69:871–98.
- Hortaçsu, Ali, and Jakub Kastl. 2012. "Valuing Dealers' Informational Advantage: A Study of Canadian Treasury Auctions." *Econometrica* 80:2511–42.
- Hu, Yingyao, David McAdams, and Matthew Shum. 2013. "Identification of First-Price Auctions with Non-separable Unobserved Heterogeneity." *J. Econometrics* 174:186–93.
- Hubbard, Timothy P., Tong Li, and Harry J. Paarsch. 2012. "Semiparametric Estimation in Models of First-Price, Sealed-Bid Auctions with Affiliation." *J. Econometrics* 168:4–16.
- Kitamura, Yuichi, and Louise Laage. 2018. "Nonparametric Analysis of Finite Mixtures." Discussion paper, Yale Univ.
- Komarova, Tatiana. 2017. "Extremum Sieve Estimation in  $k$ -out-of- $n$  Systems." *Communications Statis.: Theory and Methods* 46:4915–31.
- Kong, Yunmi. 2020. "Not Knowing the Competition: Evidence and Implications for Auction Design." *RAND J. Econ.*, forthcoming.
- . 2021. "Sequential Auctions with Synergy and Affiliation across Auctions." *J.P.E.*, forthcoming.
- Krasnokutskaya, Elena. 2011. "Identification and Estimation in Procurement Auctions under Unobserved Auction Heterogeneity." *Rev. Econ. Studies* 78:293–327.
- Krasnokutskaya, Elena, and Katja Seim. 2011. "Bid Preference Programs and Participation in Highway Procurement Auctions." *A.E.R.* 101:2653–86.
- Laffont, Jean-Jacques. 1997. "Game Theory and Empirical Economics: The Case of Auction Data." *European Econ. Rev.* 41:1–35.
- Laffont, Jean-Jacques, and Quang Vuong. 1993. "Structural Econometric Analysis of Descending Auctions." *European Econ. Rev.* 37:329–41.
- Lahiri, S. N. 2003. *Resampling Methods for Dependent Data*. New York: Springer.
- Levin, Dan, and James L. Smith. 1994. "Equilibrium in Auctions with Entry." *A.E.R.* 84 (3): 585–99.
- Li, Tong, Isabelle Perrigne, and Quang Vuong. 2000. "Conditionally Independent Private Information in OCS Wildcat Auctions." *J. Econometrics* 98:129–61.
- . 2002. "Structural Estimation of the Affiliated Private Value Auction Model." *RAND J. Econ.* 33:171–93.
- Li, Tong, and Xiaoyong Zheng. 2009. "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions." *Rev. Econ. Studies* 76:1397–429.
- Ma, Shuangge, and Michael R. Kosorok. 2005. "Robust Semiparametric M-Estimation and the Weighted Bootstrap." *J. Multivariate Analysis* 96:190–217.
- Maskin, Eric, and Jean Tirole. 1992. "The Principal-Agent Relationship with an Informed Principal, II: Common Values." *Econometrica* 60:1–42.
- Milgrom, Paul R. 1987. "Auction Theory." In *Advances in Economic Theory: Fifth World Congress*, edited by Truman F. Bewley, 1–32. Econometric Soc. Monographs, no. 12. Cambridge: Cambridge Univ. Press.
- Milgrom, Paul R., and Robert J. Weber. 1982. "A Theory of Auctions and Competitive Bidding." *Econometrica* 50:1089–122.
- Paarsch, Harry J. 1992. "Deciding between the Common and Private Value Paradigms in Empirical Models of Auctions." *J. Econometrics* 51:191–215.
- Pinkse, Joris, and Guofu Tan. 2005. "The Affiliation Effect in First-Price Auctions." *Econometrica* 73:263–77.



- Riley, John G. 1988. "Ex Post Information in Auctions." *Rev. Econ. Studies* 55:409–29.
- Roberts, James W. 2013. "Unobserved Heterogeneity and Reserve Prices in Auctions." *RAND J. Econ.* 44:712–32.
- Shneyerov, Artyom. 2006. "An Empirical Study of Auction Revenue Rankings: The Case of Municipal Bonds." *RAND J. Econ.* 37:1005–22.
- Somaini, Paulo. 2020. "Identification in Auction Models with Interdependent Costs." *J.P.E.* 128 (10): 3820–3871.
- Syrsgkanis, Vasilis, Elie Tamer, and Juba Ziani. 2018. "Inference on Auctions with Weak Assumptions on Information." Discussion paper, Harvard Univ.
- Tang, Xun. 2011. "Bounds on Revenue Distributions in Counterfactual Auctions with Reserve Prices." *RAND J. Econ.* 42:175–203.
- van der Vaart, Aad W., and Jon A. Wellner. 1996. *Weak Convergence and Empirical Processes: With Applications to Statistics*. New York: Springer.
- Vincent, Daniel R. 1989. "Bargaining with Common Values." *J. Econ. Theory* 48:47–62.
- White, Halbert, and Jeffrey M. Wooldridge. 1991. "Some Results on Sieve Estimation with Dependent Observations." In *Nonparametric and Semiparametric Methods in Econometrics and Statistics*, edited by William A. Barnett, James Powell, and George E. Tauchen, 459–93. Cambridge: Cambridge Univ. Press.
- Yao, Qiwei. 2003. "Exponential Inequalities for Spatial Processes and Uniform Convergence Rates for Density Estimation." In *Development of Modern Statistics and Related Topics: In Celebration of Prof. Yaoting Zhang's 70th Birthday*, edited by Heping Zhang and Jian Huang, 118–28. River Edge, NJ: World Scientific.