Testing Competition in Common Value Auctions: The Case of U.S. Offshore Oil and Gas Lease Auctions*

Andres Aradillas-Lopez
Philip A. Haile
Kenneth Hendricks
Robert H. Porter
Diwakar Raisingh

Preliminary and Incomplete

July, 2022

Abstract

Most of the oil and gas tracts in the Gulf of Mexico that were auctioned by the federal government after 1980 were adjacent to tracts that are already under lease. Owners of neighboring leases enjoy incumbency advantages that act as a barrier to entry for other firms. We ask whether there is evidence that these owners took advantage of their situation and did not compete with each other. To answer this question, we develop two tests of competitive bidding in common value, first-price auctions where rejection suggests collusion. One test examines the affiliation of neighbor firms' participation decisions and their bid levels. The second test examines how the distribution of their pivotal-expected values varies with the level of competition. We implement these tests using data on bids and tract characteristics. The results from both tests indicate that bidding was competitive before 1983, but not thereafter, especially on deep water tracts.

*Aradilas-Lopez: Department of Economics, Pennsylvania State University; Haile: Department of Economics and Cowles Foundation, Yale University, and NBER; Hendricks: Department of Economics, University of Wisconsin-Madison and NBER; Porter: Department of Economics, Northwestern University and NBER; Raisingh: Economic Analysis Group, Antitrust Division, US Department of Justice. The views expressed herein are solely the authors' and are not purported to reflect those of the US Department of Justice.

1 Introduction

The U.S. Bureau of Ocean Energy Management (BOEM), formerly known as the Mineral Management Service, operates under a Congressional mandate to recover "fair market value" from oil and gas leases on federal lands on the Outer Continental Shelf (OCS). The BOEM uses a combination of first-price sealed bid auctions and royalties to capture this value, and there have been few alterations in lease terms and the auction mechanism over the years. The leasing program began in 1954 and has generated considerable revenue for the government. Most of the offshore oil and gas is produced in the Gulf of Mexico. Figure 1 depicts the revenues the government earned from the sale of leases in the Gulf of Mexico and royalty payments from these leases for the period 1954 to 2002. The revenues are reported in two year increments and measured in 1982 dollars. Total sales revenues over the entire period were $64 billion and total royalty payments were $91.6 billion. Sales revenues rose dramatically in the 1970's and early 1980's, and the discoveries on these leases generated a substantial flow of royalty payments for many years thereafter. But sales revenues plummeted after 1983. Prior to 1983, the government earned $47.7 billion from sales of 3,256 leases, or $15 million per lease; after 1983, it earned only $16.3 billion from sales of 15,848 leases, or $1 million per lease. In a previous paper **?**, we document that, not only did the value of bids fall after 1983, but rent capture also fell, especially in deep water areas. While previous empirical work has supported the hypothesis of competitive bidding prior to 1983, [1] the drop in revenues raises the question of whether the same was true in later years. However, because a number of other significant changes occurred in 1983, answering this question requires a more detailed examination of the data. This is the goal of our paper.

We focus primarily on the sale of tracts that have adjacent tracts under lease at time of the sale.[2] We will refer to these tracts as "neighbor tracts" and their owners as "neighbor firms*"*. Neighbor firms have previously invested in seismic surveys of the area and, in some cases, drilled wells. These investments imply that neighbors are potential bidders for adjacent tracts. They may also deter non-neighbors from entering, which in turn would give neighbors an incentive not to compete against each other. The benefits from collusion are clear: lower prices and possibly better information if neighbors are willing to share information. The costs are less obvious, since bidding joint ventures (except for the Big 8) are not illegal.[3] Figure 2 illustrates how the sale of tracts with neighbors have grown in

[1] See, e.g., ***

[2] A tract is considered an adjacent of tract t if it shares a boundary point or an edge with tract t.

[3] All firms were allowed to bid jointly prior to 1975. After 1975, the (then) eight largest oil and gas companies were banned from bidding jointly with each other, although they were free to partner with other firms.

importance over time as more of the Gulf is explored and developed. It reports the number of tracts with neighbors as a fraction of all tracts sold in two year increments for the period 1954 to 2002. Tracts with neighbors have dominated sales since the mid-1970s, typically accounting for between 70 and 90 percent of leases sold in any given sale.

Our approach is to develop tests of competitive bidding in first-price, common value auctions where rejection suggests collusion. We apply those tests to bids by neighbors to determine whether they bid competitively. The tests exploit the fact that first, the number of neighbors is known and arguably exogenous and second, given their investments in the neighborhood, neighbors are potential bidders for adjacent tracts. The answer to this question is important in its own right. If neighbors enjoy incumbency advantages, then rent capture depends critically upon the willingness of neighbors to bid competitively. It is also important as a specification test. Identifying model primitives and evaluating changes in technology and policies through the lens of a competitive model of bidding makes sense only if collusion is not a factor.

Our restriction to neighborhood cartels is motivated by two considerations. Tracts can be viewed as locations in a spatially differentiated market. In the U.S. Gulf of Mexico, the typical tract covers nine square miles and has eight adjacent tracts. Tract values are spatially correlated because of the way in which nature creates oil and gas deposits. As a result, competition for unleased tracts may be limited primarily to the firms that lease tracts in the neighborhood. It is natural to ask if these firms can take advantage of the situation and agree not to compete against each other. Second, in previous work by two of the authors **??**, we documented that firms would frequently bid jointly on tracts in one area but against each other in other areas, both within and across sales. Thus, cooperation appears to be more location-based than sale or firm-based. Of course, cooperation at one location may spill over into other locations, including ones where firms are not neighbors. Our analysis does not preclude such spillovers but neither does it detect them.

Our first test of competitive bidding is a non-parametric test of affiliation of participation and bids by bidders. The standard model of common value auctions assumes that bidders obtain informative signals about the value of the lease and that these signals are affiliated. This information structure is often sufficient for existence of an equilibrium in which bid functions are monotone increasing. Therefore, if bidders bid competitively, their participation decisions and bids are affiliated. Of course, failure to reject does not imply competitive behavior since bidders could be engaged in phantom bidding, such as bids that are proportional to the serious cartel bid. Unobserved heterogeneity also biases the outcome towards not rejecting. However, rejecting affiliation does suggest noncompetitive behavior, specifically models of collusion in which the neighborhood cartel submits only one bid. Our test is non-parametric because bidding is a tail event - i.e., most tracts in a sale do not receive any bids, especially in the more recent portion of our sample. The test is based on inequalities implied by affiliation, and these inequalities continue to hold conditional on the event of at least one bid - i.e., excluding tracts with no bids. This is important because we do not observe the set of tracts considered by the bidders, especially in sales prior to 1983.

Our second test of competitive bidding is a stochastic dominance test of the distributions of pivotal-expected values in the number of neighbors. This test requires stronger assumptions. Specifically, it assumes that bids are best replies to the empirical distribution of the highest rival bid and that the variation in number of neighbors is exogenous. **?** ("HHS") developed this type of test to discriminate between private values and common values models under the maintained hypothesis that bidders are bidding competitively.[4] We use a variation of their approach to discriminate between competitive and collusive behavior when the information structure is common values. Under this hypothesis, the distribution of pivotal-expected values of neighbors should be stochastically increasing in the number of neighbors if they are bidding competitively, and rejection suggests models of collusion in which the neighborhood cartel submits only one serious bid.

We apply the tests separately to shallow water tracts sold before and after 1983, and to deep water tracts which became available after the introduction of area-wide leasing (AWL) in 1983. The tests provide strong support for competitive bidding on shallow water tracts sold before 1983. We fail to reject affiliation of neighbor participation and bids and find that the distribution of pivotal-expected values of neighbors is stochastically increasing in the number of neighbors. These results are consistent with the evidence and results presented in prior work discussed below. The test results for deep water tracts tell a different story. We reject affiliation of neighbor participation and bids and find that the distributions of pivotal-expected values of neighbors are not stochastically increasing in the number of neighbors. Since the information structure on shallow and deep water tracts are similar, these results provide strong evidence that neighbors are not competing on deep water tracts. The results for shallow water tracts with neighbors sold after 1983 are more mixed.

We also examine isolated tracts, those without neighbor tracts under lease, in the same periods. Here we focus on the bidding behavior of the seven most active firms. The tests tell a similar story, with bidding behavior consistent with competition according to both tests before 1983, but not after 1983.

Our paper makes several contributions. The tests provide a way of detecting collusive behavior in common value environments. The previous literature on collusion in auctions has focused on private value environments, where the participation decisions of bidders

[4] See also **?**.

are independent of the number of bidders. On the methodological front, we develop a new method for homogenizing the auctions. The standard approach is to regress bids on a rich set of neighbor covariates and use the residuals to non-parametrically estimate the empirical distribution of the highest rival bid. We could not use this approach because we do not observe non-neighbor bidders who chose not to bid. Instead, we use machine-learning techniques to construct a one-dimensional index of neighborhood covariates for predicting the distribution of the high rival bid, conditional on a bid by a neighbor firm. We use this index to estimate the distribution of pivotal-expected values non-parametrically. On the policy front, our results inform lease and auction design, which is currently a topic of discussion at various government agencies such as the Congressional Budget Office and U.S. Government Accountability Office.

Earlier work by Hendricks, Porter *et al* focused on the pre-1980 period. Hendricks and Porter **?** studied competition in auctions of drainage tracts. These are neighbor tracts where oil and gas have been discovered on adjacent tracts. We showed that bidding for these tracts is consistent with a Bayesian Nash equilibrium of a pure common value auction in which neighbor firms are informed and collude, and non-neighbor firms are uninformed. This paper provides formal tests of the collusion hypothesis. **?** ("HPP") study competition in auctions of wildcat tracts, which are mostly isolated tracts. They find that expected rents on these tracts were mostly competed away and could not reject the null hypothesis that bidding for these tracts are the Bayesian Nash equilibrium of a pure common value auctions with symmetric information. **?** generalize the HHS test for common values and apply it to bidding data from sales of neighbor leases in the Gulf of Mexico prior to 1983. They were able to reject private values in favor of common values when accounting for unobserved heterogeneity and endogenous bidder entry.

We begin in Section 2 with an overview of the offshore leasing program. In Section 3 we describe our model of competitive bidding in first-price, common value auctions, and the that underlie the analysis. Section 4 describes the implications of collusion. Our test procedures are described in Sections 5 and 6, together with the implementation of the tests. Section 7 describes the data. Section 8 presents the results of the tests. Section 9 concludes.

2 Background

To motivate our model, we provide information on the federal offshore leasing program in the Gulf of Mexico and the information structure of the auctions. More detailed descriptions of the program can be found in **?** and **?**. We also discuss joint bidding ventures and how they affect our measure of competition among neighbor firms.

2.1 The Offshore Leasing Program

The program is administered by the Bureau of Ocean Energy Management (BOEM), a division of the US Department of Interior. Federal offshore lands in the Gulf of Mexico are partitioned into blocks or tracts. The typical tract covers a rectangular area of 5,760 acres or 9 square miles and has eight adjacent tracts. The tracts are leased to oil and gas firms. Most leases cover a single tract, although when a hydrocarbon deposit is very likely, BOEM sometimes splits a tract in half and covers it with two leases. BOEM holds yearly sales in which it sells hundreds of leases to oil and gas firms. Prior to each sale, BOEM announces that tracts within a designated area are available for leasing. Firms interested in acquiring leases engage in seismic testing and analysis of the area to determine their willingness-to-pay for individual leases.[5] The leases are sold individually and simultaneously in first-price, sealed bid auctions. There is an announced minimum bid or reserve price, which is typically \$15 or \$25 per acre, with the same reserve price for all leases in a given sale.[6] In addition, the government can reject bids above the reserve price if it believes that there is not enough competition for the lease.[7]

A lease is an option contract: the holder has the right but not the obligation to drill an exploratory well. There is a fixed term during which time drilling must begin; otherwise the lease expires and ownership reverts back to the government for resale. The term is 5 years for *shallow* water tracts and 8-10 years for *deep* water tracts. Tracts are classified as *shallow* if they are located in water depths less than 200 meters and *deep* otherwise. The royalty rate of a lease also depends upon tract depth: it is 1/6 on shallow water leases and 1/8 on deep water leases. The payments that the winning bidder makes to BOEM consist of its bid, payable on the sale date, a small annual rental fee until the lease expires or production begins, and a royalty rate on any revenues it earns from post-sale production.

The only major change to the offshore leasing program in the Gulf of Mexico occurred in 1983 with the introduction of area-wide leasing (AWL). Prior to 1983, BOEM restricted the supply of new tracts in a sale to several hundred shallow water tracts in selected areas. Firms were invited to nominate eligible tracts to ensure that they were included in the final list of tracts offered at the sale. This procedure forced firms to target their investments in information and compete for relatively few tracts. In the context of tracts with neighbor leases, the nomination procedure may have played an important informational role. By not nominating an adjacent tract, neighbors revealed that they were not interested in bidding

[5] They are not permitted to drill wells on tracts prior to their sale.

[6] In a few sales, the reserve price depends on water depth.

[7] It conducts a bid adequacy decision, based in part on its own independent assessment of tract value and the number and values of bids submitted. The rejection rate was approximately 13% during the period 1954-82 but only 4% during the period 1983-2002.

for that tract. After 1983, BOEM did away with the selection and nomination process. Each year, firms could survey and bid for any of the several thousand unleased tracts in the Western and Central Gulf, including deep water tracts.

2.2 Seismic Information

Potential bidders in a sale face considerable uncertainty about the location and size of hydrocarbon deposits. To reduce this uncertainty, they acquire seismic data prior to the sale, which they use to learn about the properties of the subsurface layers of sediment and rock, and to identify structural anomalies that have the potential to trap oil and gas. Most of the seismic surveys conducted before 1990 involved vertical cross sections that the firms would use to develop 2D models of the substrata. The costs of acquiring these data are hundred thousands of dollars per tract, with costs often shared among several companies. In the 1990s, advances in computing power made 3D seismic analysis possible, especially for deep water tracts. These surveys are both more informative and more expensive. But the only way to know whether a tract contains a hydrocarbon deposit and the size of that deposit is to drill exploratory wells.

After acquiring the seismic data, geophysicists at the oil and gas firms spend a lot of time interpreting the data and estimating the value of the tracts offered for sale. This process is a closely guarded secret, and it provides firms with noisy, but qualitatively similar, private signals about deposit locations and sizes. The observed dispersion in bidder participation and bids suggests that the heterogeneities in bidder beliefs are substantial and important. Firms frequently bid on different tracts, and the amounts that they bid vary widely. For example, the second-highest bid is on average only two-thirds of the highest bid on leases that attract at least two bids (Haile et al. 2010).

An important feature of oil and gas exploration is that hydrocarbon deposits are spatially correlated. A field in the Gulf of Mexico typically covers several tracts, and larger fields may cover an even wider area. As a result, prior investments in seismic data by firms who own leases adjacent to an unleased tract may give them a cost advantage over other firms. Neighbor firms who have also drilled wells may also be better informed. However, this information advantage is short-lived. Firms have to submit a drill core report on each well to BOEM and these reports are made public after the lease expires or within 24 months of the drilling date, whichever comes first. Furthermore, production from wells is more or less public information.

To a large extent, the uncertainty faced by bidders is common value, as opposed to private value. Bidders value oil and gas deposits similarly, on the basis of wellhead prices, and have similar drilling costs, which are primarily determined by tract depth and rig rental rates.

2.3 Joint Ventures and Neighbor Bidders

After conducting their analysis of the seismic data and prior to the sale, firms often negotiate joint bidding agreements with each other. These agreements specify the firms' share of costs (including the bid) and revenues, and are legally binding contracts. Most of these agreements are area and sale-specific; that is, firms who bid jointly in one area of the sale may bid competitively in other areas or in other sales. All firms were allowed to bid jointly prior to 1975. After 1975, joint bids involving two or more of the (then) eight largest oil and gas firms (where large is determined by world oil and gas production) were banned, although these firms were free to participate in joint bids with other firms.

After the sale, firms sometimes sign shared work or acquired interest agreements, typically with owners of adjacent leases. These joint drilling and production agreements specify the firms' shares in costs and revenues, and designate an operator. They need to be approved by BOEM. They are especially common on deep-water tracts. Interestingly, all firms, including the major firms, are free to sign these agreements with each other. In other words, even though major firms cannot bid jointly on leases, they can become co-owners of leases, which in turn can affect bidding for adjacent leases in subsequent sales.

Given the prevalence of joint bidding, it is important to identify the set of neighbor firms on tract t in sale σ who are potential competitors. To do this, we partition the set into groups. If firm A is the sole owner of at least one neighbor lease, then it is assigned to its own group. If firms A and B jointly own a lease on a neighbor tract, then they are assigned to the same ownership group. The assumption here is that such firms will agree not bid against each other on tract t and will act as a single firm, either submitting a joint bid or a solo bid by one of the firms. Of course, there can be multiple joint-bidding relationships within a single neighborhood. For example, if firms A and B jointly own a lease of a neighbor tract and firms B and C jointly own a lease on another neighbor tract, we specify firms A, B, and C as belonging to a single ownership group. The assumption here is that if A and B agree not to compete and B and C agree not to compete, then A and C also agree not to compete.[8]

More formally, the set of joint ownership relations in a neighborhood defines graph in which each node (vertex) represents a neighbor firm and each edge indicates joint ownership of at least one tract in the neighborhood. We define ownership groups corresponding to the

[8] We relax this assumption in a robustness exercise.

components of this graph: owners of neighbor tracts who are not connected (i.e., isolated nodes in the graph) are in singleton components, while owners connected to each other through one or more joint venture relations belong to the same component. The number of ownership groups for tract t is equal to the number of components. We will refer to these ownership groups as "neighbor bidders" We observe changes in ownership of a given tract over time. Neighbor firms, and therefore also ownership groups, for a given tract t in sale σ are defined as of the date of the sale. The number of contemporaneous ownership groups for tract t is our measure of the number of neighbor bidders, N_t.

3 Model

In this section, we present our model of competitive bidding and characterize competitive equilibrium behavior. We then discuss the effect of the winner's curse on participation and bids.

3.1 Setup

Here we present a model that underlies our tests of competitive bidder behavior. A lease t is offered in a first-price sealed bid auction with a reserve price r_t. Commonly observed characteristics of the lease and its neighborhood are denoted by $X_t \in \mathbb{X}$, with $r_t \in X_t$. Potential bidders for the lease comprise two types of firms: neighbors and non-neighbors. We denote the set of neighbors by $\mathcal{N}_t$, letting $N_t = |\mathcal{N}_t|$. The value of the lease (net of drilling and production costs) is unknown but common to all firms and denoted by $V_t \in [\underline{v}, \overline{v}]$.[9] A firm can obtain a private signal of the tract value by acquiring and analyzing seismic surveys and other data for the relevant area. We assume that neighbor firms are endowed with signals. This reflects the fact that owners of active neighbor leases will already have undertaken the relevant data analysis for the area of the tract. Non-neighbors, on the other hand, must choose whether to acquire a (costly) signal. We let NN_t denote the number of non-neighbors acquiring a signal. We assume that no firm without a signal will submit a bid above the reserve price.

We refer to firms with signals as "bidders." Bidders are the players in the auction game. We let $K_t = N_t + NN_t$ denote the total number of bidders. We focus on auctions in which there are at least two bidders, i.e., $K_t \geq 2$. Let S_{it} denote the signal of bidder i. Without loss of generality we let each signal S_{it} have a standard uniform marginal distribution.

[9] Although we assume a pure common values model due to the nature of these auctions, the model and testable predictions are essentially unchanged if we generalize the to the symmetric affiliated values model of **?** as long as the limiting case of private values is ruled out. Evidence supporting the presence of a common value component in the pre-AWL period has been shown, e.g., by **?**, **?**, and **?**.

We assume that bidders are symmetric at the auction; thus, while neighbors have a cost advantage in signal acquisition, the signals of all firms i are equally informative ex ante. This reflects the fact that all firms ultimately have access to the same signal technology.[10] We denote the joint distribution of signals and tract value conditional on (X_t, K_t) by $F(S_{1t}, ..S_{K_t t}, V_t|X_t, K_t)$. Let $S_t = (S_{1t}, ..., S_{K_t t})$ denote the signals of all bidders for tract t and let $S_{-it} = S_t/S_{it}$ denote the signals of bidder i's competitors.

Assumption 1. For any $x \in \mathbb{X}$ and $k \in \mathbb{Z}_+$, $F(S_t, V_t|x, k)$ has a continuously differentiable density that is strictly affiliated, exchangeable in the bidder indices, and positive on $[0, 1]^k \times [\underline{v}, \overline{v}]$.

Affiliation is a standard assumption, formalizing the notion that higher values of the signal S_{it} imply stochastically higher values of both V_t and S_{-it}. The strict form of affiliation assumed here requires that this positive dependence be strict, as is natural in a common values setting. The exchangeability component of Assumption **??** implies that bidders are *ex ante* symmetric. Thus, although we allow for asymmetries in signal acquisition costs, we assume that all firms have access to the same signal technology.

We will not model the signal acquisition ("entry") decisions of non-neighbors. Instead we rely on two assumptions on entry outcomes. The first is that the number of bidders K_t is uniquely determined by X_t and N_t.

Assumption 2. Conditional on (X_t, N_t), the distribution of K_t is degenerate with probability one.

This is an important assumption, made necessary by the fact that we do not observe entry decisions of non-neighbor firms. This is a mild assumption when (X_t, N_t) are sufficient statistics for the information available at the entry stage.[11] For example, Assumption 2 holds when one combines the the standard auction model with a standard simultaneous-move, complete information entry game in which non-neighbors have identical signal acquisition costs conditional on X_t, N_t. Following **?** and **?**, it is easily confirmed that under mild regularity conditions, in any Nash equilibrium of this entry game, the number of non-neighbor entrants is unique with probability one (see **?** for details). Intuitively, in this case the number of bidders is determined by a standard free-entry condition. Thus, although the identities of entering non-neighbors is indeterminate, the number of non-neighbor entrants

[10] Appendix G provides a robustness analysis, focusing on the subset of tracts with no prior drilling on neighboring tracts, where the most likely source of any informational asymmetry is absent.

[11] Violations of Assumption 2 could arise if entry decisions reflect unobserved auction-level heterogeneity, selective entry, mixed strategies, or stochastic selection among equilibria that yield different numbers of entrants.

is unique except in the (zero probability) event of indifference. With this assumption, we can write

$$K_t = k(X_t, N_t)$$

for some function k.

Although some testable implications of our model are obtained without additional maintained assumptions, we state two additional conditions that underlie one of our testing strategies. The first condition requires the function k to be weakly increasing in N_t, and strictly so for some values of (X_t, N_t).

Assumption 3. $k(x, n)$ is weakly increasing in n for all (x, n), and strictly increasing in n for some (x, n).

Assumption **??** is a natural assumption, reflecting the fact that, all else equal, adding a neighbor bidder increases the number of bidders that obtain a free signal. The quantifier here reflects the fact that k need not always respond to a change in n If the marginal entrant at (x, n) is a non-neighbor bidder, then an increase in the number of neighbor bidders from n to $n+1$ makes the winner's curse more severe and, as a result, may "crowd out" a non-neighbor bidder, leaving k unchanged. Assumption **??** holds in the example described above as long as non-neighbor entry costs do not increase with N_t. In that case, adding a neighbor bidder will never reduce the total number of bidders; and whenever $k(x, n) = n$, it will be the case that $k(x, n+1) = n+1$ [12] With this property, we may represent $F(S_t, V_t|X_t, K_t)$ instead as $F(S_t, V_t|X_t, N_t)$, which we will do henceforth.

The second assumption requires that the variation in N_t is exogenous conditional on X. In order to make explicit what we mean by this assumption, let $F\left(S_1, \ldots, S_{k(X,N)}, V|X, N, \ell\right)$ denote the joint distribution of the tract valuation and $\ell \leq k(X, N)$ randomly chosen bidder signals, conditional on (X, N).

Assumption 4. For all $x \in \mathbb{X}$, $n \in \mathbb{Z}_+$, and $\ell \leq k(x, n)$, $F(S_1, ..., S_{k(x,n)}, V|x, n, \ell) = F(S_1, \ldots, S_{k(x,n+1)}, V|x, n+1, \ell)$.

Assumption 4 states that, if we add a neighbor bidder but hold all else the same, then the joint distribution of the tract value and the signals of ℓ randomly chosen bidders conditional on X is unaffected. This is an important assumption, but one that may be particularly plausible in our application. A primary reason this type of exogeneity condition could fail is the presence of unobserved tract-level heterogeneity that affects bidders' entry decisions. Such concerns may be mitigated when it is possible to control for the primary sources of

[12] If non-neighbors do not enter with $K = n$, then it cannot be profitable for them to enter when $K = n+1$.

tract-level heterogeneity. For example, because the number of active neighbor leases L_t will be included in X_t, variation in the number of neighbor firms N_t conditional on X_t reflects only variation in the number of distinct firms owning the L_t leases, not the overall attractiveness of the neighborhood. In fact, our data set offers an exceptionally rich set of tract- and neighborhood-level observables, and our machine learning approach is designed to allow us to control flexibly for their impact on the joint distribution of signals and tract value.

3.2 Equilibrium Behavior and the Winner's Curse

By standard results (see, e.g., **?**), Assumptions **??** and 2 ensure existence of a unique symmetric Bayesian Nash equilibrium in increasing bid functions for any realization of (X_t, N_t). Let $\beta\left(\cdot; X_t, N_t\right)$ denote this equilibrium bidding strategy, which is in fact strictly increasing.[13] Let the random variable

$$B_{it} = \beta\left(S_{it}; X_t, N_t\right) \tag{1}$$

denote the equilibrium bid of bidder i in auction t. Let $G\left(B_{1t}, \ldots, B_{k(x,n)t}; x, n\right)$ denote the joint distribution of equilibrium bids conditional on $X_t = x$ and $N_t = n$.

Of course, a bidder may conclude after observing its signal that the lease is not sufficiently valuable to justify a bid above the reserve price. As discussed by **?**, this decision is determined by a threshold signal $s^*(x, n) \in [0, 1]$, known as the "screening level" (**?**). The screening level is defined implicitly by the condition

$$E[V_t | S_{it} = s^*(x, n), \max_{j \neq i} S_{jt} \leq s^*(x, n), X_t = x, N_t = n] = r_t \tag{2}$$

characterizing indifference to winning the auction at the reserve price. By convention, we set $B_{it} = 0$ when bidder i chooses not to bid. Because we have normalized signals to be uniform on $[0, 1]$, the probability that a given bidder submits a (strictly) positive bid in auction t is equal to $1 - s^*(x_t, n_t)$.

The threshold condition (2) incorporates bidders' recognition of the winner's curse. A bidder with signal $S_{it} = s^*(x_t, n_t)$ knows in equilibrium that it it will win the auction only when no other bidders have signals above $s^*(x_t, n_t)$. The winner's curse also appears in the

[13] **?** provide a characterization of this strategy. We exploit a characterization of the inverse bidding strategy below.

calculus determining the level of the nonzero bids. Let

$$M_{it} = \max_{j \neq i} B_{jt}$$

denote the highest bid among bidder i's rivals at auction t. Let

$$G_{M|B}\left(m|b,x,n\right) = \Pr\{M_{it} \leq m|B_{it} = b, X_t = x, N_t = n\},$$

denote the conditional distribution function of the high rival bid, with $g_{M|B}\left(m|b;x,n\right)$ denoting the associated conditional density. Following the key insight in **?**,[14] each equilibrium bid B_{it} must satisfy the first-order condition

$$w(S_{it}; X_t, N_t) = b_{it} + \frac{G_{M|B}(B_{it}|B_{it}, X_t, N_t)}{g_{M|B}(B_{it}|B_{it}, X_t, N_t)}, \tag{3}$$

where we define

$$w\left(s_{it}; x, n\right) = E\left[V_t|S_{it} = s_{it}, \max_{j \neq i} S_{jt} = s_{it}, X_t = x, N_t = n\right]. \tag{4}$$

The term $w\left(s_{it}; n, x\right)$ will play an important role below. It represents the expectation of the tract value conditional on all information available to bidder i and the hypothesis that i is tied with another bidder for having the highest signal at the auction. The monotonicity of equilibrium bidding strategies implies that in equilibrium a tie for the highest signal implies a tie for winning the auction, i.e., that i is "pivotal." As usual, optimal behavior involves conditions defined conditional on pivotality, even though a bidder will turn out to be pivotal with probability zero. Following **?**, we refer to $w\left(s_{it}; x_t, n_t\right)$ as bidder i's "pivotal expected value" at auction t. An important observation for what follows is that (3) immediately implies identification of the pivotal expected values associated with each nonzero bid.[15]

4 Testable Implications of Competition

Our broad strategy is to test falsifiable restrictions of the model of competitive bidding developed above. We consider testable implications of two key aspects of the competitive model.

[14] See also **?**, **??**, **?**, **?**, and **?**.

[15] See, e.g., **?**, **?**,**??**, **?**, **?**, and **?**.

4.1 Affiliation of Bids

The first testable implication of the model follows from the following well-known result[16].

Lemma 1. *If $Z_1, \ldots, Z_k$ are affiliated random variables and $g_1, \ldots, g_k$ are weakly increasing functions, then the random variables $g_1(Z_1), \ldots, g_k(Z_k)$ are affiliated.*

The equilibrium bidding strategies $\beta(\cdot; X_t, N_t)$ derived above are weakly increasing in signals; in fact, they are strictly increasing in signals above the screening level $s^*(X_t, N_t)$. Thus, by Lemma 1, competitive behavior requires affiliation of bids at each auction t.

Proposition 1. *Under Assumptions **??**–2, for any auction t with $K_t > 1$, (B_{it}, B_{jt}) are affiliated conditional on X_t, N_t for any bidders i and j.*

Note that this implication of competitive bidding does not require Assumptions **??** or 4. This prediction applies to the bids of both neighbors and non-neighbors. However, we will focus below on neighbor bids. This reflects both the fact that we do not observe the number of non-neighbor bidders, and our focus on neighbor bidders as the most natural candidate firms for any collusion. Note also that Proposition 1 requires affiliation in both the discrete (positive or not) and continuous portions of the bidding decisions. We will examine each of these below.

4.2 Response to the Winner's Curse

The second aspect of the competitive model we examine is bidders' equilibrium response to the winner's curse. In equilibrium, bidders who are acting competitively adjust their participation and bidding strategies to account for the "bad news" (à la **?**) implied upon learning that one has outbid all competitors. Although we cannot observe such adjustments directly, it is possible to isolate the way those adjustments respond to changes in the severity of the winner's curse that arise from exogenous changes in the level of competition. Building on the insights of **?** and **?**, we consider two manifestations of these comparative statics predictions.

4.2.1 Monotone Screening Levels

In the competitive model, a bidder's willingness to submit a bid of at least the reserve price depends on the number of competitors faced. Intuitively, the bad news implied by discovering that no other bidders had signals above the screening level depends on how many other bidders there are. The larger the number of bidders, the worse is the news

[16] See Theorem 3 of **?**.

about the tract value. Formally, given Assumptions **??**–4, it is straightforward to show that the participation threshold $s^*(x,n)$ is increasing in n (see Lemma 1 of **?**), and strictly increasing in n for some (x,n). Thus, fixing X_t, our model of competition implies that bidders are less likely to submit a bid in auctions with more neighbors. Letting

$$\rho\left(X_t, N_t\right) = \Pr\{B_{it} > 0 | X_t, N_t\}$$

this immediately implies the following testable restriction.

Proposition 2. *Under Assumptions* **??**–*4,* $\rho\left(x,n\right)$ *is decreasing in* n *for all* x*, and strictly decreasing in* n *for some* (x,n)*.*

4.2.2 Stochastic Dominance of Pivotal Expected Values

We also look for evidence of neighbors' equilibrium response to the winner's curse in the level of their positive bids. Our approach here follows from two observations. First, as noted section 3.2, for every auction t and bidder i submitting a positive bid, the pivotal expected value $w\left(s_{it}; x_t, n_t\right)$ is identified *via* the equilibrium first-order condition (**??**). The second observation is closely related to the ideas in section 4.2.1. In particular, the conditioning event $\{\max_{j\neq i} S_{jt} = s_{it}\}$ reflecting the winner's curse in (4) conveys different information depending on the number of bidders. Specifically, while it always implies a tie at s_{it} with one bidder, it also implies that $K_t - 2$ bidders had signals below s_{it}. The latter is always bad news about these signals and, therefore, about the tract value V_t. But the larger is K_t, the worse is this bad news. This was shown formally (given Assumption 4) in **?**. And under Assumptions 2 and **??**, this extends immediately to imply that $w\left(s_{it}; x, n\right)$ is decreasing in n, and strictly so for some x, n.

To state the testable implication formally, define the random variable

$$W_{it} = \begin{cases} w(S_{it}; X_t, N_t) & \text{if } S_{it} \geq s^*(X_t, N_t) \\ 0 & \text{otherwise,} \end{cases},$$

and let $F_W(\cdot|X_t, N_t)$ denote the conditional distribution of W_{it}. For any $x \in \mathbb{X}$ let

$$w^*\left(x\right) = \max_n w\left(s^*\left(x,n\right); x, n\right).$$

Then we have the following testable implication of the competitive model.

Proposition 3. *For all* x*, under Assumptions* **??**–*4,* $F_W\left(w|x,n\right)$ *is weakly increasing in* n *for all* $w \geq w^*\left(x\right)$*. For some* (x,n)*, the monotonicity is strict.*

The need to limit attention to pivotal expected values above $w^*(x)$ arises from the fact the first-order condition (**??**) only allows identification of pivotal expected values associated with signals above the screening level for each auction.[17] Thus, given $X_t = x$ and $N_t = n$, we can identify $F_W(w|x,n)$ only for $w \geq w(s^*(x,n);x,n)$. The value of $w(s^*(x,n);x,n)$ can move up or down with an increase in n: a large n reduces $w(s;x,n)$ at any given signal s, but the threshold signal $s^*(x,n)$ rises with n. This creates ambiguity about the ordering of the distributions $F_W(w|x,n)$ with respect to n for w near $w(s^*(x,n);x,n)$. However, by restricting attention to $w \geq w^*(x)$, we are guaranteed to have $F_W(w|x,n{+}1) \geq F_W(w|x,n)$.

The stochastic ordering is strict when increases in N result in increases in K. But, and unlike the models in **?** or **?**, adding a neighbor bidder can have two effects. As described above, the direct effect is an increase in the number of neighbor bidders, but the indirect effect may be a decrease in the number of non-neighbors due to reduced profitability of entry. The net effect may be no change in K, in which case the distributions do not vary with n. We test the stochastic ordering that reflects the net effect.

4.3 Collusive Alternatives

Because there are many ways neighbor firms might collude, we will not specify a particular model of collusion as an alternative to the hypothesis of competitive behavior. However, it is useful to briefly discuss collusion in broad terms and to suggest forms of collusion that illuminate the potential for the testable implications discussed above to fail in the absence of competition.

The main goal of any collusive agreement is to avoid competing away rents. Neighbor firms enjoy a positive of cost advantage by already having signals of the tract value. And they can earn rents if they agree not to bid against each other. They might also benefit from pooling their data and assessments to obtain a more precise estimate of lease value. Joint bidding ventures among most bidders are not illegal in federal offshore auctions. After the auctions are held and winners determined, all firms can and, as we document later, often do enter joint drilling and production agreements. Finally, owners of neighboring tracts typically cooperate in producing from a common pool if one is discovered. Therefore, neighbors have an incentive to cooperate in bidding for adjacent leases, and the opportunity to do so.

One simple form of collusion involves neighbors coordinating to designate a single neighbor to bid in the auction. In what we refer to as the "designated bidder model," the cartel includes all neighbors and randomly selects one of them to submit a serious bid This bidder

[17] See the related discussion in HHS.

evaluates its signal and bids as if there were no other neighbor bidders. The cartel may or may not submit additional "phantom" bids below the serious bid in order to create the appearance of competition.[18] Non-neighbors are aware of the cartel and rationally act as if there is only one competing neighbor bidder. In this model, the severity of the winner's curse does not vary with the number of neighbor firms—the auction is always as if there were a single neighbor bidder. Likewise, the number of non-neighbors entering will vary only with X_t, not N_t.

A possible variation of the designated bidder model is one in which the neighbor firms pool their information prior to the submission of the single serious bid.[19] In this case, the cartel acts as a single bidder that is better informed than non-neighbors when $N_t > 1$, and this informational advantage increases with N_t. The informational asymmetry between the cartel bidder and non-neighbors makes it difficult to predict how the non-neighbors bid (including the decision to submit a positive bid) will vary with N_t. Not least among the challenges is the potential for multiple equilibria. But the implied violations of our maintained assumptions and the assumption of competitive bidding will lead to multiple forms of misspecification in the empirical model.

We do not commit to either of these models. Both suggest that collusion could lead to failures of the required stochastic dominance ordering of pivotal expected values inferred from the observed bids through the first-order conditions of the competitive model. More broadly, however, observe that the combination of Propositions 1 and 2 offers a difficult hurdle for a cartel to clear. Any form of cartel that submits no phantom bids will mechanically produce neighbor participation probabilities that decline with n, mimicking the pattern required by competitive entry. But in this case the requirement of affiliation in neighbor decisions to submit nonzero bids will fail—indeed, these decisions will exhibit negative dependence. On the other hand, if colluding neighbors always submit one bid each, the observed frequencies of nonzero neighbor bidding will be invariant to n, violating the requirement of Proposition 2.

5 Data and Descriptive Statistics

In this section, we introduce the data and report summary statistics on bidding behavior and on lease outcomes. The statistics provide descriptive evidence that bidding behav-

[18] Colluding neighbors would have an incentive to engage in phantom bidding because BOEM can reject the high bid on a tract if it perceives a lack of competition (i.e., too few bids). Recall that the rejection rate during the pre-AWL period was approximately 13% but only 4% during the AWL period.

[19] One of the bidders who participated in joint bidding agreements told us that the exchange of information is typically limited to summaries of seismic reports. Bidders are careful not to reveal processing secrets to each other for fear that this information will be used against them in auctions where they are rivals.

ior changed after the introduction of the Area Wide Leasing program, and became less competitive (see also Haile et al (2010)).

Our data cover all tracts in the Gulf of Mexico for the period 1954 to 2002, inclusive. For each tract, we observe the tract's water depth, acreage, location, that is, the longitude and latitude coordinates of its boundary points, and the date that it is sold, or all dates if the tract is sold more than once. Each time the tract is offered for sale and received bids, we observe the royalty rate, the lease term, the identity of all bidders and the amounts they bid, and whether the high bid is rejected. If a bid is joint, the participants and their shares are recorded. For any ownership transfer after the original sale, we see the transfer date, the identity of the new owners, and their shares. We know the date, number, and depth of any wells drilled on the tract, as well as monthly production data through 2010 of oil, condensate, natural gas and other hydrocarbons. For the set of leases that are productive beyond 2010, we forecast future output based on an econometric model of historical decay rates in lease production.

The tract-level data allow us to construct a large number of covariates associated with each tract and its neighborhood. We consider covariates falling into three categories. First are characteristics that vary at the level of the sale, including sale date and contemporaneous real oil and gas prices (as measured by the offshore Gulf first purchase price) and real rig rental rates.[20] Second are characteristics of the tract itself, such as water depth, acreage, dummies for different regions (Eastern or Central Gulf), and whether the tract was offered previously (attracting no bids or being relinquished by a prior leaseholder). Finally, we have constructed measures of many neighborhood characteristics. Among these is the number of active neighbor leases, which allows us to isolate variation in the number of neighbor firms conditional on the number of neighbor tracts under lease. We also construct variables capturing the bidding history for neighbor tracts, as well as the exploration and production history for neighbor leases. The neighborhood bidding and drilling histories include tracts within three rings of the tract in question. The Appendix includes a complete list of the variables that we use in estimating the model.

Tables 1 and 2 provide sample statistics for isolated leases and for those with active neighbor leases, respectively, conditional on receiving at least one bid. To capture changes in bidder behavior, we divide the sample into three subsets. The first subsample consists of all leases receiving bids in the period before Area Wide Leasing, from 1954 to 1982, which we denote pre-AWL. The AWL period is from 1983 to 2002, and we divide the AWL leases into two subsets, AWL Shallow and AWL Deep, where the former consists of shallow water

[20] In practice, we use year-of-sale dummies to capture common price and cost shocks that are likely to influence the drilling decisions of leases sold in the same year.

tracts and the latter consists of deep water tracts. Essentially all tracts receiving bids in the pre-AWL period were in shallow water.

Table 1: Summary Statistics for Isolated Leases

Variable	Pre-AWL	AWL-Deep	AWL-Shallow
No. of Leases Bid	2,016	2,574	951
Avg. No. of Bids	3.566	1.310	1.386
Fraction of Leases Sold	0.900	0.985	0.978
Drill Rate	0.780	0.112	0.309
Hit Rate	0.461	0.275	0.331
Avg. Win Bid (if sold)	11.140	0.860	1.881
Avg. Rev $\vert$Hit	221.277	557.258	53.642
Avg. Cost $\vert$Drilled	10.941	21.133	11.642

Table 2: Summary Statistics for Neighbor Leases

Variable	Pre-AWL	AWL-Deep	AWL-Shallow
No. of Leases Bid	1,619	4,815	8,111
Avg. No. of Bids	2.611	1.404	1.508
Fraction of Leases Sold	0.863	0.967	0.946
Drill Rate	0.820	0.156	0.388
Hit Rate	0.547	0.322	0.497
Avg. Win Bid (if sold)	13.684	0.922	1.020
Avg. Rev $\vert$Hit	146.763	354.870	40.210
Avg. Cost $\vert$Drilled	15.360	21.260	11.467

In these tables, the drill rate refers to the fraction of sold leases on which at least one well was drilled. The hit rate refers to the fraction of drilled leases that had any production. We convert production information into revenue estimates using wellhead prices. The reported revenue of a productive tract is computed by converting production flows of oil and gas into revenues using the real wellhead prices at the date of production, and discounting them to the auction date at a 5 percent *per annum* rate.[21] Costs in these tables are based upon

[21]There is an important question of how to treat firms' expectations of future oil and gas prices, since leases can be productive for as long as 50 years. For the purposes of this paper, we are using an *ex post* measure of revenues, but we have experimented with others measures, including revenues at wellhead prices at the time of the sale date (static expectations) and revenues based on a forecasting model of wellhead prices.

the American Petroleum Institute annual survey of drilling costs of wildcat and production wells.[22] All dollar figures are in millions of 1982 dollars.

The differences between the pre-AWL period and the AWL period are striking. The majority of Pre-AWL tracts are isolated, but the pattern is reversed in the AWL period. There are very few isolated tracts in the AWL Shallow subsample. Many more tracts are sold during the AWL period but the number of bids per lease and bid levels fell dramatically. On leases with active neighbor tracts, the number of bids per tract fell from 2.6 to 1.5 and average winning bid fell from $13.6 million to $1 million. There is an even greater drop in the number of bids on isolated tracts, from 3.6 to 1.3, and a comparable fall in average winning bids. The average number of bids and winning bids are similar in the AWL subsamples. Drill rates are near 80 percent on both Pre-AWL samples, 38 percent in the AWL Shallow sample, and 14 percent for AWL Deep. Hit rates are comparable for the Pre-AWL and AWL Shallow samples, around 50 percent, but about 33 percent for AWL Deep. The latter number probably in part reflects the much higher costs for developing deep water tracts, so that there is production only for larger discoveries. Drilling costs on AWL deep water tracts are almost twice as high as drilling costs on AWL Shallow or Pre-AWL tracts. But average revenue per productive tract is much higher on AWL Deep tracts.

Haile et al (2010) provide evidence that the Pre-AWL period also differs from the AWL period in other notable respects, and especially in comparison with AWL Deep sales. In the Pre-AWL period, the government captured most of the value of the tract. Tract value is measured by discounted revenues less discounted drilling costs, and the government share equals bonus bids plus discounted royalty payments. In contrast, firm profits were a substantial share of value on AWL Deep tracts. Moreover, a large share of the profits in the AWL Deep period were earned on tracts with active neighbor leases where a neighbor firm submitted a bid. The vast majority of those tracts were acquired by neighbor firms, which therefore earned the majority of the AWL Deep profits.

In our formal analysis, we will focus on tracts with $1 \leq N_t \leq 5$. Tables 3, 4 and 5 present summary statistics by number of neighbor bidders for each of the three subsamples of leases which have active neighbor tracts, again conditioning on submission of at least one bid. We report the fraction of tracts receiving at least one neighbor bid ("NB" bids), as well as the neighbor win rate and the fraction of tracts won by neighbor firms conditional on leases being sold. We also report the number of non-neighbor bids ("NN" bids). Finally, for the sample of submitted bids, we report the fraction which faced at least one rival bid

[22] We use the API estimates for offshore Louisiana and Texas, and information on the number and depth of wells, to compute drilling costs for each tract drilled, classifying wells as productive if the tract produced hydrocarbons and exploratory if it did not. These costs are also discounted to the auction date at a 5 percent per annum rate.

submission. Within each sample, neighbor bid and win rates increase with N_t. The number of non-neighbor bids (and the value of their bids) per tract falls slightly with N_t, more so in the AWL Deep sample. We do not report variation in outcomes within each sample, which is largely uncorrelated with N_t. For example, drill and hit rates are roughly constant across N_t for each subsample.

Table 3: Summary Statistics for Pre-AWL Leases by Number of Neighbors

Variable	$N_t = 1$	$N_t = 2$	$N_t = 3$	$N_t = 4,5$
Total No. of Leases Bid	552	579	335	147
Total No. of Bids	1,443	1,421	878	462
Fraction of Bids with M>0	0.820	0.823	0.841	0.894
Total No. of NN Bids	1,117	937	483	236
Fraction of Leases with NB bids	0.591	0.686	0.779	0.837
Fraction of Leases Sold	0.891	0.846	0.836	0.884
NB Win Rate on Leases Sold	0.360	0.429	0.568	0.592

The AWL program had a large impact on participation rates and outcomes. Prior to 1983, the average neighbor bid rate was approximately 60% for $N_t = 1$, and drill rates were roughly 85%. After 1983, the neighbor bid rate dropped to roughly 30% for $N_t = 1$, and the drill rate dropped to 38% for shallow water tracts and 15% for deep water tracts. The number of non-neighbor bids averaged a little more than 2 per tract during the pre-AWL period but dropped to a little more than 1 per tract in the AWL period.

Table 6 lists the nine most active firms over the full sample, and designates the seven who acquire the most leases for each subsample. The nine most active firms acquire a large fraction of the Pre-AWL and AWL Deep leases, 88% and 77% respectively, and a relatively small share of AWL Shallow leases, 38%. Bidding for AWL Deep leases is dominated by solo major firm bids, but *ex post* there is more joint lease ownership. For example, the seven most active firms acquire 64 percent of the leases sold. The majority of those bids, 55 percent of the 64%, are solo major firm bids, that is they are not a joint bid with another large firm. (The "solo" bids by major firms may be joint with smaller firms.) The remaining 9 percent of the winning bids by large firms involve two or more majors. However, there are ownership changes after the sale date, resulting in more concentration of ownership and

Table 4: Summary Statistics for AWL-Deep Leases by Number of Neighbors

Variable	$N_t = 1$	$N_t = 2$	$N_t = 3$	$N_t = 4, 5$
Total No. of Leases Bid	$2,091$	$1,573$	815	326
Total No. of Bids	$2,812$	$2,268$	$1,178$	488
Fraction of Bids with M>0	0.415	0.489	0.487	0.512
Total No. of NN Bids	$2,120$	$1,416$	594	250
Fraction of Leases with NB bids	0.331	0.491	0.617	0.620
Fraction of Leases Sold	0.979	0.964	0.946	0.954
NB Win Rate on Leases Sold	0.275	0.396	0.529	0.514

more joint ownership, especially for tracts that are explored. For example, 72 percent of the AWL Deep tracts that are drilled have a major firm owner before the time of the first well, 48 percent with a solo major firm owner, and the remaining 24 percent with two or more major firm owners.

Table 5: Summary Statistics for AWL-Shallow Leases by Number of Neighbors

Variable	$N_t = 1$	$N_t = 2$	$N_t = 3$	$N_t = 4, 5$
Total No. of Leases Bid	$1,212$	$1,893$	$2,057$	$2,537$
Total No. of Bids	$1,693$	$2,707$	$3,081$	$4,049$
Fraction of Bids with M>0	0.465	0.487	0.533	0.582
Total No. of NN Bids	$1,423$	$1,990$	$2,126$	$2,741$
Fraction of Leases with NB bids	0.223	0.347	0.415	0.444
Fraction of Leases Sold	0.949	0.943	0.942	0.946
NB Win Rate on Leases Sold	0.174	0.276	0.322	0.343

Table 6: Leases Won by Top 9 Firms

Name	Pre-AWL	AWL-Deep	AWL-Shallow
Shell	346*	1401*	613*
Getty/Arco/Conoco/Cities	633*	604*	607*
Chevron	394*	560*	415*
BP	265*	710*	259
Texeco	257	602*	320*
Exxon	307*	622*	249
Gulf/Hess	460*	382	332*
Amoco	250	392*	383*
Mobil	270*	379	300*
Fraction Won by Top 9	0.875	0.765	0.384

*Firm is in top 7 for the period.

6 Implementation of the Tests

Here we discuss our formal tests of hypothesis discussed in section 3.2. Throughout we assume that the set of tracts offered for lease constitute an i.i.d. sample of (X_t, N_t, B_t).

revise when resolveIn principle, we may restrict the sample of auctions considered arbitrarily on the basis of the covariates X_t. However, we must always restrict attention to values of (N_t, X_t) for which $K(N_t, X_t) \geq 2$—i.e., those implying that each neighbor faces competition. When $N_t \geq 2$, this is automatic. However, when $N_t = 1$, we must identify values of X_t ensuring $K(N_t, 1) \geq 2$. Appendix B describes how we do this. Our tests are then implemented

6.1 Tract Heterogeneity: A Nonparametric Index

Controlling for tract-level heterogeneity is critical to each of our testing strategies. This is not unique to our approach, but essential to all empirical work relying on inversion of equilibrium first-order conditions a là **?**, or on comparative statics predictions that are known to hold only when conditioning on the information X_t commonly known to all bidders at auction t. Our data set offers an exceptionally rich set of tract-level covariates. This is valuable because it alleviates concerns about the potential for significant unobservable tract heterogeneity. Of course, in finite samples one must control for covariates in a practical way. A common approach in the literature is to assume that tract-level observables alter valuations through a multiplicatively (or additively) separable scalar index, and that the index itself is linear in covariates. For example, this permits "homogenization" of auctions by residualizing bids with a first-stage regression (see **?**).

Here we take a related approach, but relax the separability and linearity assumptions. In particular, although we assume that the joint distribution of (V_t, S_t) is affected by X_t only through a scalar index (conditional on N_t), we require no further functional form assumption beyond continuity.

Assumption 5. For each $n \in \{1, .., \overline{n}\}$, there exists a function $\lambda_n : \mathbb{X} \to \mathbb{R}$ such that $F(V_t, S_t | N_t = n, X_t = x) = F(V_t, S_t | N_t = n, \lambda_n(x))$, with $F(V_t, S_t | N_t = n, \lambda_n(x))$ continuous in $\lambda_n(x)$.

Note that we permit the index function $\lambda_n(\cdot)$ to vary freely with the number of neighbors. Because the bidding strategies in (1) are strictly increasing, an immediate implication is that the joint distribution of equilibrium bids conditional on $X_t = x$ and $N_t = n$ satisfies

$$G\left(B_{1t}, \ldots, B_{nt}; x, n\right) = G\left(B_{1t}, \ldots, B_{nt}; \lambda_n(x), n\right). \tag{5}$$

Thus, if we know the index function $\lambda_n(\cdot)$, we need only condition on the scalar $\lambda_n(X_t)$ rather than the entire vector X_t when examining the properties of equilibrium bids in an n-neighbor auction, or when constructing bidders' equilibrium first-order conditions, allowing us to do so in a flexible way.

To estimated the index functions, we again exploit (5). In particular, we obtain an estimate of the index function $\lambda_n(\cdot)$ for each $n = 1, 2$, and 3 by fitting a random forest to predict the value (including zero) of M_{it} conditional on X_t, N_t, and B_{it} for each positive bid B_{it}. Under Assumption 5, any functional of the joint distribution of bids $(B_{1t}, \ldots, B_{nt})$ could be used as the target outcome (see (5)) for estimating the relevant index functions $\lambda_n(\cdot)$. We focus on the maximum opposing bid M_{it} associated with each positive bid due to the central role that the conditional distribution of this variable plays in the equilibrium conditions used to interpret the equilibrium bids. We provide details in Appendix A.

6.2 Affiliation Test

*****old text from testable predictions dropped here for now***Because most available tracts receive no bids, we focus on the subset of tracts receiving at least one bid. This requires correcting for the implied sample selection. Fix $X_t = x$ and $N_t = n$. If $g_B(\cdot|x, n)$ is the joint density of equilibrium neighbor bids in the population (including tracts receiving no bids), the density in the restricted sample is

$$g_B^*(b|x,n) = \frac{\mathcal{L}(B_t = b,\ \max_i B_{it} > 0|x,n)}{\Pr(\max_i B_{it} > 0|x,n)}$$

where in the numerator is the likelihood of the event $\{B_t = b,\ \max_i B_{it} > 0|x,n\}$. Let $\mathcal{B}(x,n)$ denote the set of all neighbor bid vectors b such that $g_B^*(b|x,n) > 0$. For any $b \in \mathcal{B}(x,n)$, the event $\{B_t = b\}$ implies the event $\{\max_i B_{it} > 0\}$. So

$$\mathcal{L}\left(B_t = b,\ \max_i B_{it} > 0|x,n\right) = g_B(b|x,n).$$

Thus, for $b \in \mathcal{B}(x,n)$, $g_B^*(b|x,n)$ is equal to $g_B(b|x,n)$ divided by a constant that depends only on (x,n). Affiliation is preserved when the density is divided by a constant. Thus, on any subset of $\mathcal{B}(x,n)$ that forms a lattice, we have affiliation of the joint density $g_B^*(\cdot|x,n)$. Although $\mathcal{B}(x,n)$ is not itself a lattice,[23] ****what is our solution?***

[23]In particular, it is not true that whenever b and b' are in $\mathcal{B}(x,n)$, so are $b \vee b'$ and $b \wedge b'$. For example, suppose (x,n) are such that $k(x,n) = n = 2$ (no entry by non-neighbors). Then for some $\tilde{b} > r$ both $b = \left(\tilde{b}, 0\right)$ and $b' = \left(0, \tilde{b}\right)$ will lie in $\mathcal{B}(x,n)$. However, $b \wedge b'$ is equal to $(0,0)$, which cannot be in $\mathcal{B}(x,n)$ (auctions with no bids aren't in the sample). So $g_B^*(b \wedge b'|x,n) = 0$, guaranteeing that the affiliation inequality $g_B^*(b \wedge b'|x,n)\, g_B^*(b \vee b'|x,n) \geq g_B^*(b|x,n)\, g_B^*(b'|x,n)$ fails at these points.

****

We first consider the affiliation of bids required by Proposition 1.Consider the set of tract t for which $N_t > 1$. At each of these tracts, partition the neighbor bids into bins defined by thresholds $b_0 < b_1 < b_2 < \cdots < b_D$, where $b_0 = 0$, $b_1 = r_t$ (recall that $r_t \in X_t$), and b_D is the maximum equilibrium bid conditional on (X_t, N_t). Then for each tract t and neighbor i, define the discrete random variable

$$C_{it} = d \text{ if } B_{it} \in [b_d, b_{d+1}),\ d = 0, .., D-1,$$

indicating the cell of the partition in which B_{it} falls. Let $C_t = (C_{1t}, .., C_{nt})$, with $c_t = (c_{1t}, c_{2t}, .., c_{nt})$ denoting the realization of C_t Because each C_{it} is a weakly increasing transformation of B_{it}, Lemma 1 and Proposition 1 imply that $(C_{1t}, .., C_{nt})$ are affiliated. By definition, this means that for any two possible realizations, c' and c'', of C_t, we have

$$\Pr\{c' \wedge c''\} \Pr\{c' \vee c''\} \geq \Pr\{c'\} \Pr\{c''\}, \tag{6}$$

where $\wedge$ denotes the meet (element-by-element minimum) and $\vee$ denotes the join (element-by-element maximum).

We will test these "affiliation inequalities" (6) using two types of partitions. The first focuses on the affiliation of neighbor decisions to submit a positive bid, setting $D = 2$. The second incorporates affiliation in the continuous bid levels as well. In this case we will set $D = 3$, where the three cells of the partition correspond to $B_{it} = 0$, $B_{it} \in [r_t, m]$, and $B_{it} > m$, where m denotes the median bid in the relevant sample.

To test these affiliation inequalities (6) We developed an algorithm that computes the minimal set of inequalities implied by (**??**) for arbitrary values of D and N_t. The algorithm generates all possible affiliation inequalities and then drops trivial inequalities and redundant inequalities due to symmetry and dependence. We identified the latter by applying linear programming techniques to the log-transformation of the set of inequalities that result from eliminating the trivial inequalities and ones due to symmetry.

Example 1. *Suppose $n = 3$ and $D = 2$. Then $C_{it} = 0$ if bidder i submits no bid, and $C_{it} = 1$ otherwise. By symmetry,*

$$\begin{aligned} \Pr\{1,0,0\} &= \Pr\{0,1,0\} = \Pr\{0,0,1\} \equiv p(1) \\ \Pr\{1,1,0\} &= \Pr\{0,1,1\} = \Pr\{1,0,1\} \equiv p(2). \end{aligned}$$

Define $p(0) = \Pr\{0,0,0\}$ and $p(3) = \Pr\{1,1,1\}$. The affiliation inequality (6) holds trivially *(and with equality) for any* $c \in \{0,1\}^3$ *if* $c' = (0,0,0)$ *or* $c' = (1,1,1)$. *For any distinct* $c, c' \in \{(1,0,0),(0,1,0),(0,0,1)\}$ *or distinct* $c, c' \in \{(1,1,0),(0,1,1),(1,0,1)\}$, *(6) yields*

$$p(0)p(2) \geq p(1)^2 \tag{7}$$

and

$$p(1)p(3) \geq p(2)^2. \tag{8}$$

Finally, for any $c \in \{(1,0,0),(0,1,0),(0,0,1)\}$ *and* $c' \in \{(1,1,0),(0,1,1),(1,0,1)\}$, *(6) either holds trivially or yields*

$$p(3)p(0) \geq p(1)p(2). \tag{9}$$

However, the inequality (9) is already implied by (7) and (8).[24] *Therefore, the minimal set of inequalities is given by (7) and (8).*

We consider first a test of affiliation in neighbor participation, setting $D = 2$. In this case, letting ℓ index the number of nonzero neighbor bids, the minimal set of inequalities for any $N_t \geq 2$ is given by

$$p(\ell - 1|X_t, N_t)p(\ell + 1|X_t, N_t) \geq p(\ell|X_t, N_t)^2 \quad \ell = 1, \ldots, N_t - 1. \tag{10}$$

Given symmetry, it will be convenient to ignore bidder identities and focus on the total number of nonzero bids,

$$A_t = \sum_{i=1}^{N_t} A_{it}$$

where $A_{it} = 1\{B_{it} > 0\}$. Define

$$\pi_{\ell,n}(x) \equiv \Pr(A_{it} = \ell|X_t = x, N_t = n) = \binom{n}{\ell} p\left(\ell|x,n\right).$$

Then, conditioning on $X_t = x, N_t = n$, we can rewrite (10) as

$$\binom{n}{\ell-1}^{-1}\binom{n}{\ell+1}^{-1}\pi_{\ell-1,n}(x)\pi_{\ell+1}(x) \geq \binom{n}{\ell}^{-2}\pi_{\ell,n}^2(x) \quad \ell = 1, \ldots, n-1.$$

In specifying the test statistic, we will use a density-weighted version of the above

[24] (7) implies $p(0) \geq p(1)^2/p(2)$ and equation (8) implies $p(3) \geq p(2)^2/p(1)$. Multiplying these two inequalities yields (9).

inequalities. Define

$$\tau_{\ell,n}(x) = \left[\binom{n}{\ell}^{-2} \pi_{\ell,n}(x) - \binom{n}{\ell-1}^{-1} \binom{n}{\ell+1}^{-1} \pi^2_{\ell-1,n}(x)\pi_{\ell+1}(x) \right] f^2_{X|N}(x,n)$$

where $f_{X,N}$ is the joint density of (X_t, N_t). Let

$$Q_{\ell,n} = E_X[\max\{\tau_{\ell,n}(X), 0\} \cdot \omega(X, \ell, n)]$$

where $\omega(X,\ \ell, n)$ is a pre-specified, positive weighting function. Our test statistic is given by

$$Q = \sum_{n=2}^{\overline{n}} \sum_{\ell=1}^{n-1} Q_{\ell,n}.$$

where $2 \leq n \leq \overline{n}$ is the range of values of N_t over which we will test affiliation. Under the hypotheses of the competitive model, $Q = 0$, When the affiliation inequality (10) is violated for some (x, n), we will have $Q > 0$. The Q statistic is the weighted mean of the extent to which (10) is violated, if at all.

We employ kernel-based nonparametric estimators to construct the test statistic. These estimators use the sample analogue of each $\tau_{\ell,n}(x)$, using kernel smoothing over the value of X_t and aggregating over n. We partition $X_t = (X_t^c, X_t^d)$ into its continuous and discrete components, and let $\phi : \mathbb{R}^q \to \mathbb{R}$ denote the kernel function where q is the dimension of X_t^c. For a given $x = (x^c, x^d)$ and bandwidth $h > 0$, we define[25]

$$\mathcal{H}(x_t - x; h) = \frac{1}{h^q} \phi\left(\frac{x_t^c - x^c}{h} \right) \cdot 1\{x_t^d - x^d\}.$$

Our choice kernel ϕ is a bias-reducing, 12^{th} order polynomial. Our nonparametric estimators of the density weighted event probabilities are of the form

$$\widehat{\pi}_{\ell,n}(x)\widehat{f}_{X,N}(x,n) = \frac{1}{T} \sum_{t=1}^{T} 1\{A_t = l\} \cdot 1\{N_t = n\} \cdot \mathcal{H}(x_t - x^c; h).$$

We use these estimators to compute $\widehat{\tau}_{\ell,N}(x)$, the sample analog of $\tau_{\ell,N}(x)$, and

$$\widehat{Q}_{\ell,N} = \frac{1}{T} \sum_{t=1}^{T} \widehat{\tau}_{\ell,N}(X_t) \cdot 1\{\widehat{\tau}_{\ell,N}(X_t) \geq -b_T\} \omega(X_t, \ell, N)]$$

where $b_T \longrightarrow 0$ is a nonnegative sequence converging to zero at an appropriate rate. Sum-

[25]We can also smooth over the discrete covariates X^d instead of using indicator functions. We focus on estimators of this type for simplicity.

ming $\widehat{Q}_{\ell,N}$ over ℓ and N yields an estimate of the test statistic, $\widehat{Q}$. Under smoothness and regularity conditions, Aradillas-Lopez (2016)***add bibtex reference*** shows that $\widehat{Q}$ satisfies the following key asymptotic property,

$$\widehat{Q} = Q + \frac{1}{T}\sum_{t=1}^{T}\gamma_T(X_t) + \mathcal{O}_p\left(T^{-1/2-\epsilon}\right),$$

where (i) $E[\gamma_T(X_t)] = 0$ and (ii) $\gamma_T(X_t) = 0$ with probability 1 if the inequalities hold almost strictly (see Aradillas-Lopez (2016) for details). Letting $\Sigma_T^2 = Var(\gamma_T(X_t))$, our test-statistic is of the form

$$\widehat{s}_T = \frac{\sqrt{T}\widehat{Q}}{\max\{\kappa_T, \widehat{\Sigma}_T^2\}},$$

where κ_T is a bandwidth sequence that converges to zero slowly (at a logarithmic rate). For a target level α, we reject affiliation if $\widehat{s}_T > z_{1-\alpha}$, where $z_{1-\alpha}$ is the $1-\alpha$ quantile of the standard normal distribution. The critical value is 1.96 for 2.5% significance level and 1.645 for 5% level.

We construct an analogous test statistic for bid levels, with $D = 3$ as discussed above. As in the binary case, the test statistic compares the empirical frequency of the actions for each neighbor against the frequency predicted under the null of affiliation. However, the number of event inequalities to be tested with three bid levels is much larger than with two bid levels. In the participation test, there is only one event inequality when $N_t = 2$, and the number of inequalities increases by one with each additional neighbor. By contrast, with three cells in the partition, there are three inequalities for $N_t = 2$, nine for $N_t = 3$, 18 with $N_t = 4$, and 30 with $N_t = 5$.

6.2.1 Discussion

PH; I think I would cut this whole section*

One reason for using a nonparametric test for affiliation is that it allows for flexible patterns of correlation among bids. This is especially important when submitting a nonzero bid is a tail event, as is the case in our sample. ***PH; I don't understand this claim. Why is nonparametric better in this case?**** Most tracts in a sale do not receive any bids, which suggests that the likelihood of a bidder drawing a signal above the screening threshold is often quite low.

A second, more important reason for choosing a nonparametric test is to deal with the potential selection problem of conditioning on the sample of leases that receive bids. Parametric approaches (e.g., Li & Zhang (2010)) have to account for the probability of auctions getting no bids. ***I think this is misleading. Just as we are doing, in a parametric context one can interpret the parametric specification as apply to the condition (on receiving a bid) distribution rather than the unconditional distribution. I would drop this paragraph and perhaps keep from this section only the discussion of power against collusion*** This requires knowing the set of available leases. We do not observe this set for sales during the pre-AWL period because the locations of the tracts that were nominated but not bid (roughly half of the set) are not reported. Similarly, for sales in the AWL period, we do not observe the bidders' consideration sets of leases. By contrast, our test does not require knowing the set of available leases.

PH: I don't think the claims in this paragraph are true. Without the assumption on non-neighbor participation we might have multiple equilibria conditional on x,n, for example. I'm not sure we need exogenous variation in N; where is that required? Without symmetry, equilibrium may not exist and may not be unique. I propose cutting all of this.**The affiliation test does not require any assumptions on non-neighbor entry nor does it require symmetry between neighbors and non-neighbors. The key assumptions are (i) neighbor bidding strategies are monotone increasing in the private signal and (ii) exogenous variation in the number of neighbors. Our specific formulation of the affiliation test assumes symmetry among neighbors, but this can be relaxed. In estimating the event probabilities, we have restricted the effects of the covariates to be the same across N. This assumption can also be relaxed.

6.3 Tests for Responses to the Winner's Curse

this section should include both the test of monotone screening levels and the FOSD test; below has only the latter

the following bid was moved from above and dropped here for now:* In particular, for any (w, n, x), define

$$\Delta(w; n, x) = F_W(w; n+1, x) - F_W(w; n, x).$$

Then we can formulate the null hypothesis of competition implied by Proposition 3 a

$$H_0 : \Delta(w; n.x) \geq 0 \quad \forall n, x, w \geq w^*(x) \tag{11}$$

with alternative

$$H_1 : \Delta(w, n) < 0 \quad \exists n, x, w \geq w^*(x) \tag{12}$$

Any violation of the null implies violation of some maintained assumption of the model, including the assumption of competitive bidding. Such hypotheses have been used to form specification tests of competitive bidding models by HHS and **?**.

We can also formulate a test for evidence that the stochastic dominance required by our model is present, essentially reversing the direction of the test. In this case we specify the null as

$$H_0 : \Delta(w; n.x) \leq 0 \quad \forall n, x, w \geq w^*(x) \tag{13}$$

with alternative

$$H_1 : \Delta(w, n) > 0 \quad \exists n, x, w \geq w^*(x). \tag{14}$$

Evidence against H_0 in favor of H_1 will indicate satisfaction (rejection of failure) of the requirements of Proposition 1. Note that because the least-favorably null in this case arises when $\Delta(w; n.x) = 0 \forall n, x, w$, testing in this case is equivalent to testing the null of equal distributions $F_w(\cdot|x, n)$ across n against the alternative of stochastic dominance.[26]************

The sample of auctions consists of all auctions of leases with neighbors that received at least one bid (neighbor or non-neighbor). In each auction, we observe the set of neighbors, which neighbors bid and the value of their bids, and which neighbors did not bid. We set $B_{it}^{NB} = 0$ if neighbor i does not bid in auction t. We do not observe the set of potential non-neighbor bidders nor do we observe which non-neighbors entered and did not bid; we only observe which non-neighbors bid and the values of their bids. Let B_{it}^{NN} denote a non-neighbor bid. Our sample consists of an i.i.d. sample $\{B_t, X_t, N_t\}_{t=1}^T$ where

$$B_t = \{B_{it}^{NB}\}_{i=1}^{N_t} \cup \{B_{it}^{NN} : B_{it}^{NN} > 0\}.$$

The main objects to be estimated with these data are the probability that a bidder does not participate, $s^*(x, n)$, the conditional bid distributions $G_{M|B}(M_{it}|B_{it}; n, x)$, and the conditional densities, $g_{M|B}(M_{it}|B_{it}; n, x)$ for $n = 1, 2$, and 3. In estimating the distributions and densities, it will be convenient to use natural log bids with one added to allow for zeros.[27] Let $\widetilde{B}_{it} = \ln(1 + B_{it})$ and $\widetilde{M}_{it} = \ln(1 + M_{it})$.

6.3.1 Bid Distributions

We estimate $s^*(x, n)$ for each n using a kernel regression of the binary outcome $1\{\widetilde{B}_{it}^{NB} > 0\}$ on $\widehat{\lambda}_n(X_t)$. We include only neighbor bids in this regression because we do not observe the number of non-neighbors who entered but did not bid. However, under our symmetry

[26] In HHS, the null of equal distributions was derived as the implications of a private values model. Thus, the hypothesis we test are equivalent to those of the test for common values in HHS.

[27] This allows the bandwidths for the kernel estimators to scale with the value of the bid, which we measure in dollars. The reserve price is approximately $86,000.

assumption, our estimate of the probability of a neighbor not bidding conditional on n and $\widehat{\lambda}_n(x)$ also applies to non-neighbors. We denote our estimate of the probability of not submitting a bid as $\widehat{s}(n, \widehat{\lambda}_n(x))$.

We will consider tests of our stochastic dominance hypotheses for $n = 1, 2$ (comparing auctions with 1, 2, and 3 neighbors). One practical concern is that $w^*(x)$ may be quite high - so high that it would mean looking at a very small part of the distributions of pivotal expected values. But there is no reason why we have to include all observed values of X. We can (and do) consider a subset of the support of X where the participation probabilities are relatively high.

The first-order condition in equation (5) can be rewritten in natural log bids as

$$w(s, s, n, \widehat{\lambda}_n(x)) = \exp(\widetilde{b}) \left(1 + \frac{G_{\widetilde{M}|\widetilde{B}}(\widetilde{b}|\widetilde{b}, n, \widehat{\lambda}_n(x))}{g_{\widetilde{M}|\widetilde{B}}(\widetilde{b}|\widetilde{b}, n, \widehat{\lambda}_n(x))} \right). \tag{15}$$

To obtain an estimate of the ratio $G_{\widetilde{M}|\widetilde{B}}/g_{\widetilde{M}|\widetilde{B}}$, we follow the approach taken by Li, Perrigne, and Vuong (2000, 2002) and use nonparametric estimates of the form

$$\widehat{G}_{\widetilde{M},\widetilde{B}}(\widetilde{m}, \widetilde{b}; n_t, \widehat{\lambda}_n(x_t)) = \frac{1}{T_t \times h_B \times h_\lambda} \sum_{t=1}^{T_n} \sum_{i=1}^{\ell_t} \Phi \left(\frac{\widetilde{m} - \widetilde{M}_{it}}{h_M} \right) K \left(\frac{\widetilde{b} - \widetilde{B}_{it}}{h_B} \right) K \left(\frac{\lambda - \widehat{\lambda}_{n_t}}{h_\lambda} \right)$$

$$\widehat{g}_{\widetilde{M},\widetilde{B}}(\widetilde{m}, \widetilde{b}; n_t, \widehat{\lambda}_n(x_t)) = \frac{1}{T_t \times h_M \times h_B \times h_\lambda} \sum_{t=1}^{T_n} \sum_{i=1}^{\ell_t} K \left(\frac{\widetilde{m} - \widetilde{M}_{it}}{h_M} \right) K \left(\frac{\widetilde{b} - \widetilde{B}_{it}}{h_B} \right) K \left(\frac{\lambda - \widehat{\lambda}_{n_t}}{h_\lambda} \right).$$

where ℓ_t is the number of bids in auction t, T_n is the number of auctions with n_t neighbors (and the sum is over the set of auctions with n_t neighbors, with the total number of observations $T_t = \Sigma_{t=1}^{T_n} \ell_t$), h_B and h_λ are bandwidths, K is a (Gaussian) kernel, and Φ is the standard normal distribution function with precision h_M. We use Φ to smooth the indicator function $1\{\widetilde{M}_{it} < m\}$. The bandwidths are selected using Silverman's Rule of Thumb. The ratio of $\widehat{G}_{\widetilde{M},\widetilde{B}}$ to $\widehat{g}_{\widetilde{M},\widetilde{B}}$ is a consistent estimator of $G_{\widetilde{M}|\widetilde{B}}/g_{\widetilde{M}|\widetilde{B}}$.[28]

Note that, because we set $\widetilde{M}_{it} = 0$ whenever bidder i faces no rival bids, the mass point at 0 will not affect unduly our estimator of $G_{\widetilde{M}|\widetilde{B}}/g_{\widetilde{M}|\widetilde{B}}$ near the reserve price r. Observations at zero will receive some weight in the estimation, but their weight near r will

[28] See HHS (2003) for details. The argument follows from the fact that

$$G_{\widetilde{M},\widetilde{B}}(\widetilde{m}, \widetilde{b}; n, \widehat{\lambda}_n(x)) = G_{\widetilde{M}|\widetilde{B}}(\widetilde{m}|\widetilde{b}; n, \widehat{\lambda}_n(x)) g_{\widetilde{B}}(\widetilde{b}, n, \widehat{\lambda}_n(x))$$

$$g_{\widetilde{M},\widetilde{B}}(\widetilde{m}, \widetilde{b}; n, \widehat{\lambda}_n(x)) = g_{\widetilde{M}|\widetilde{B}}(\widetilde{m}|\widetilde{b}; n, \widehat{\lambda}_n(x)) g_{\widetilde{B}}(\widetilde{b}, n, \widehat{\lambda}_n(x))$$

where $g_{\widetilde{M},\widetilde{B}}$ is the joint density of bid and maximum rival bid and $g_{\widetilde{B}}$ is the marginal density of bids. Note that $G_{\widetilde{M},\widetilde{B}}$ is NOT a distribution function.

be small.

To obtain estimates of the distribution of pivotal expected values, we also need an estimate of the marginal distribution of an equilibrium bid conditional on n, λ and the bid being positive. We estimate $\widehat{G}_B(\cdot|n, \widehat{\lambda}_n(x))$ for each $n = 1, 2, 3$ using kernel regression and all observed positive bids.

6.3.2 Pivotal Expected Value Distributions

****discuss here the choice to drop tracts that receive no bid and the selection correction; given our determination of which tracts were actually offered, we could have just used that entire sample and assigned the zero bids of neighbors to $w^*(x)$; I think the only reason not to do that is that we would need a lot of simulation draws to get a good sample of positive bids, and this is a bit costly inside a bootstrap loop; but maybe we should run that once and just confirm that we get CDFs that look the same***

Given estimates of the various bid distributions, we simulate the model to obtain empirical distributions of the pivotal expected value distributions. The simulation proceeds as follows:

1. Randomly draw a value of x from its empirical distribution and compute $\widehat{\lambda}_n(x)$ for $n = 1, 2$ and 3.
2. For each $n = 1, 2$ and 3, draw S bids b^s from the marginal distribution of bids, $\widehat{G}_B(\cdot|n, \widehat{\lambda}_n(x))$, setting $b^s = 0$ with probability $\widehat{s}(n, \widehat{\lambda}_n(x))$.
3. Set pivotal expected values of zero bids to zero and plug each positive bid into the GPV formula (5) to get a simulated pseudo-(pivotal expected) value $w^s(n, \widehat{\lambda}_n(x))$.
4. Draw another value of x and repeat the process.

As we iterate, we collect the pseudo-(pivotal expected) value draws associated with each n, at different draws of x. In this way, our sample for each n reflects the same mixture of values of x. Note that, although the index differs across n at each x, we are conditioning on the same draw of x for all values of n. This is all we need because the first-order stochastic dominance holds across n conditional on x. The empirical distribution of pseudo-(pivotal expected) values associated with n is a consistent estimator of the distribution function of the pivotal expected values conditional on n (i.e., integrating over x).

In practice, we will need to limit the values of x. We are already restricting attention to the subset of X that imply competition for $N_t = 1$. However, some trimming for $N > 1$ will be needed as well. When a rival bid is highly unlikely, then $\widehat{g}_{M|B}(b|b, n, \widehat{\lambda}_n(x)) \approx 0$

for some simulated bid in some simulated auction. In these cases, estimation error can generate implausibly high estimates of the associated pseudo-(pivotal expected) values. We trim these values of x by requiring that the probability of facing a rival bid (i.e., $\Pr\{M_{it} > 0 | n, \widehat{\lambda}_n(x)\}$) exceeds some cutoff η for $n = 1, 2$ and 3.

Let $\widehat{F}_W(\cdot; n)$ denote the empirical distribution of pseudo-(pivotal expected) values for each $n = 1, 2, 3$. We test the null hypotheses on these distributions. Let

$$\widehat{\Delta}(w, n) = \widehat{F}_W(w, n+1) - \widehat{F}_W(w, n).$$

Define the operator $[u]_- = u \times 1\{u < 0\}$. Following CHS, we compare pairs of distributions using the one-sided Cramer-von Mises type statistic[29]

$$CVM = \int_{w^*}^{\infty} \left(\left[\widehat{\Delta}(w, n) \right]_- \right)^2 dw.$$

Here $w^*(x)$ is the maximum (over n) of the pseudo- (pivotal expected) value $w^s(n, \widehat{\lambda}_n(x))$ associated with the threshold signal $\widehat{s}(n, \widehat{\lambda}_n(x))$, as defined above. Under the null (given consistency of our $\widehat{F}_W$), this statistic converges to zero. Under the alternative, it converges to a strictly positive number[30]. In addition to the pairwise tests, we will also construct a test statistic for the full range of n based on the maximum statistic over the pairwise statistics:

$$CVM^{\max} = \max_n \int_{w^*}^{\infty} \left(\left[\widehat{\Delta}(w, n) \right]_- \right)^2 dw.$$

For inference, we use the bootstrap. As usual, this requires re-centering of the bootstrap test statistics. Let $CVM_{n,0}$ denote the statistic obtained in our full sample. Let $CVM_{n,r}$ denote the statistic obtained by applying the same estimation procedure (i.e., starting from the very beginning, fitting of the random forest, etc.) for bootstrap replication r. We then use the empirical distribution of

$$CVM_{n,r} - CVM_{n,0}$$

to approximate the distribution of our test statistic under the null. Thus, letting H_n denote the empirical distribution of $CVM_{n,r} - CVM_{n,0}$, our bootstrap p-value for the test

[29] In practice we approximate the integral with a sum over simulated draws, e.g., the draws we use to construct the distribution functions that we have been plotting. We could consider other statistics. The KS statistic would simply replace the summation over w with the max over w, while dropping the exponent 2.

[30] Due to the boundary bias of kernel estimators, however, we restrict the integration to an interval bounded away from the limits of the support of the pivotal expected values.

is $1 - H_n(CVM_{n,0})$.

For the bootstrap resampling, we follow the standard auction block bootstrap procedure: draw an auction t from the full sample, include all bids from the selected auction in the bootstrap sample, repeat (with replacement).[31] We perform 1,000 replications.

We also test the hypotheses that the participation thresholds $s^*(x, n)$ are increasing (or constant) in n under the first test (or under the second test), versus the alternative that they are constant (decreasing) under the alternative hypothesis. These tests are relatively straightforward, involving a comparison across n of $\widehat{s}(n, \widehat{\lambda}_n(x))$, the estimated fraction of entrants who do not submit a bid. We also use the boostrap for inference. The Appendix contains more detail.

6.4 Discussion

Our stochastic dominance test relies on a number of maintained assumptions, such as optimal behavior as described by Bayes Nash equilibrium. An important maintained assumption is symmetry conditional on entry. It is possible that there may be asymmetries between neighbor and non-neighbor firms, or within those subsets, especially within the set of neighbor firms if they have differential experience. However, the GPV first order condition, as described in equation (5), only applies in private value environments with symmetric or asymmetric bidders, or in common value environments with symmetric bidders. The GPV first order condition does not characterize equilibrium behavior in asymmetric common value auctions.

As we discussed above, our stochastic dominance tests will not detect some forms of collusion. For example, an all inclusive cartel with a designated bidder and no information sharing will be indistinguishable from competitive behavior under private values, and so the null hypothesis will not be rejected in the first test, if the model is otherwise correctly specified. Such a cartel would also result in the null not being rejected for the second test, however, which we would view as suspicious given our presumption that this bidding environment entails common values.

[31] Unlike HPP and CHS, we will not employ a spatial boostrap because we do not permit any unobservables in our model that could be spatially correlated. In essence, we are assuming that by using the random forest on a large set of covariates, we control for all auction heterogeneity except for random realizations of bidder signals.

7 Test Results

7.1 Affiliation

We begin with the affiliation tests. We apply these tests separately for neighbor bids on leases with neighbors and for bids of the Top 7 firms on leases without neighbors (i.e., isolated tracts). In each case, we apply the tests to three subsamples: Pre-AWL, AWL Shallow, and AWL Deep. To ensure comparability of our results across subsamples, we use the same range of N for each test. Our criterion is to identify values of N that included at least 1% of observations in every subsample. The testing range for the neighbor bids is $N \in [2, 5]$ and, for Top 7 firm bids, it is $N \in [4, 7]$

Table 7 reports the results for affiliation of neighbor bids for the three subsamples. The first row gives the values of the test statistic for the binary case of participation, and the second row gives the values of the test statistic for the trichotomous case of participation with low and high bid. The numbers in parentheses are the p-values of the statistics. The last row gives the number of observations for the three subsamples.

Table 7: Affiliation of Neighbor Bids

	Pre-AWL	AWL Shallow	AWL Deep
{No Bid, Bid}	0.90 (0.18)	1.96 (0.025)	4.05 (0.00)
{No Bid, Low Bid, High Bid}	1.44 (0.075)	3.63 (0.00)	4.31 (0.00)
Number of Bids	1,191	6,485	2,715

The critical values for rejection of the null are 1.645 and 2.33 for 5% and 1% significance levels, respectively. Thus, we fail to reject affiliation of neighbor participation and bids for the Pre-AWL subsample at the 5% level, but we reject affiliation of participation and bids for the AWL-Deep subsample at this significance level. The results for the AWL-Shallow are less definitive. We reject affiliation of neighbor participation at the 5% level, but not at the 1% level. However, we do reject affiliation of bids in the trichotomous case at the 1% level. These results suggest that neighbors bid competitively during the Pre-AWL period but did not do so during the AWL period, and especially not on deep water tracts.

Table 8 reports the values of the test statistic and p-values for bids by the Top 7 firms on isolated tracts.

Table 8: Affiliation of Top 7 Firm Bids on Isolated Tracts

	Pre-AWL	AWL Shallow	AWL Deep
{No Bid, Bid}	1.05 (0.15)	1.70 (0.045)	4.62 (0.00)
{No Bid, Low Bid, High Bid}	0.36 (0.36)	2.12 (0.017)	3.45 (0.00)
Number of Bids	1,723	924	2,574

The test results are similar to those reported for neighbor bids. We fail to reject affiliation of Top 7 firm bids for the Pre-AWL subsample at the 1% level, but we reject the null at this significance level for the AWL Deep subsample. For the AWL Shallow subsample, we reject affiliation of participation by Top 7 firms at the 5% level, but we reject the null at the 1% level in the trichotomous case. We infer from these results that top firms bid competitively for isolated tracts during the Pre-AWL period but did not do so during the AWL period, and especially not on deep water tracts.

The AWL Shallow subsample has relatively few isolated tracts, and many of them became isolated due to neighbor leases expiring. This may explain why the test results for isolated tracts are similar to those of neighbor tracts. However, it cannot explain the results for the AWL Deep subsample. Most of these tracts were isolated because they are located in areas that had not been previously explored. Despite the absence of any asymmetries in entry costs for these leases, the top firms do not appear to have bid competitively for them. It suggests that the lack of competition for leases with neighbors may have spilled over onto isolated leases. This would not be too surprising, since the neighbors on deep water tracts are mostly Top 7 firms. They won roughly 70% of the leases in the AWL Deep subsample and, as we documented in the previous section, they often owned those leases jointly at the time of exploratory drilling. Indeed, the lack of competition among top firms for isolated tracts suggests that they may not have bid competitively for tracts with neighbors, regardless of whether they are neighbors or non-neighbors.

7.2 Stochastic Dominance

We turn next to the stochastic dominance test, which we apply separately to each of the three subsamples. The test consists of comparing the empirical distributions of pseudo (pivotal expected) values (aggregated over x) for $n = 1, 2$, and 3. The distributions depend on the trimming parameters, ζ and η. Recall that, in estimating the bid distributions for $n = 1$, we want to omit auctions in which the neighbor bids and, in equilibrium, is the only entrant. This event is not observable (to us). Instead, we estimate the probability of observing at least one non-neighbor bid conditional on X and drop auctions for which this prediction is less than ζ. Since the neighbor is more likely to participate when it has no rivals, trimming the sample in this way should affect our estimate of the probability of bidder participation. We find that this is indeed the case. As ζ increases and more non-competitive auctions are dropped from the sample, the probability of non-participation increases, with the increases diminishing as ζ gets larger. The increases are essentially zero when ζ lies between 0.3 to 0.4 so, in what follows, we set $\zeta = 0.3$.

Similarly, in simulating the empirical distributions, we focus on auctions that the data indicate are competitive ***need a different term for this*** by only drawing values of X at which the predicted probability of a rival bid is above η. This situation occurs when the probability of drawing a signal above the screening threshold is sufficiently high. It also occurs when neighbors collude and, in equilibrium, non-neighbors enter. As we mentioned previously, our model tends to generate estimates of pseudo-values that are sometimes too high when a bidder believes that it is likely to be the only bidder (i.e..$g_{M|B} \approx 0$). Thus, the main effect of trimming the simulation sample is likely to be on our estimates of the upper tails of the distributions of W. In what follows, we set η equal to 0.4.

Figures 3 to 5 show the estimated distribution functions for the three sub-samples. Table 9 reports the p-values from the formal tests for the ordering of the participation probabilities, and Tables 10 and 11 report the p-values from the tests for the ordering of the distribution functions of the pseudo (pivotal expected) values.

Figure 3 shows the estimated distribution functions for the Pre-AWL subsample with $\zeta = 0.3$ and $\eta = 0.4$. The distributions depicted in Figure 3 exhibit the stochastic ordering implied by competitive bidding in a common values model. The probability of non-participation increases with n, from 0.45 for $n = 1$ to 0.59 for $n = 2$, and then to 0.65 for $n = 3$. Both the ordering of the distributions and the probability ordering are significantly different from zero. We cannot reject the hypothesis that the participation threshold and the distribution of pseudo values are increasing in n, whereas we can reject the hypothesis that they are constant in n. In both cases, the test results are consistent with competitive bidding under common values. As expected, the gap between $\widehat{F}_W(w; n+1)$ and $\widehat{F}_W(w; n)$ is larger for $n = 1$ than for $n = 2$. It is also larger at lower values of w and goes to zero as w approaches the upper bound of its support. These results provide strong support for the hypothesis that auctions in the Pre-AWL subsample are common values and that firms bid competitively in them.[32]

Figure 4 presents the estimated distribution functions for the AWL Deep subsample

[32] Hendricks, Pinkse, and Porter (2003) and CHS (2019) also present evidence of common values and competitive behavior for this subsample.

and $\zeta = 0.3$ and $\eta = 0.4$. The ordering of the distributions for $n = 1$ and $n = 3$ is opposite that predicted under competitive bidding. The probabilities of non-participation are similar (0.73), but larger than the probability of non-participation for $n = 2$ (0.68). The distribution for positive w are also similar for $n = 2$ and $n = 3$, and below that for $n = 1$. These patterns are not consistent with competitive bidding under common or private values.

Figure 5 shows the estimated distribution functions for the AWL Shallow subsample and $\zeta = 0.3$ and $\eta = 0.4$. The distributions for $n = 2$ and $n = 3$ are ordered as implied by competitive bidding with common values. The probabilities of non-participation increase with n (0.78 and 0.83), as predicted. The issue is that the distribution function for $n = 1$ lies between the other two, with an intermediate probability of non-participation (0.82). The patterns of the distributions for positive w are also not consistent with competitive bidding, under common or private values.

The test results are not consistent with competitive bidding under common values in both the AWL samples. The discrepancy is starker in the AWL Deep sample, where we would reject the null of an increasing participation threshold at any conventional significance level, for example.

8 Additional Evidence

this is just an idea, but a section here showing relevant suggestive descriptive evidence that supports (or not) collusion as the reason for rejecting competition would strengthen the paper. Some possibilities:

1. reproduce/re-do what was in the JIE paper re: participation and profits
2. info pooling: neighbor vs. non-neighbor profits, neighbor profits increasing in N [a problem is that both would be explained by other things, including spatially correlated UH]
3. correlation with ex post values: (conditional on X)
    (a) high neighbor bid vs. losing neighbor bids (phantoms)
    (b) high neighbor vs. high non-neighbor (info pooling)

9 Conclusion

In this paper we have developed two tests of competitive bidding in common value, first-price auctions using data on bids. We have applied our approach to offshore oil and gas auctions of tracts where adjacent tracts were already under lease, and for isolated tracts. Specifically, we test whether the owners of the adjacent tracts bid competitively in the former case, and whether the most active firms bid competitively in the latter case. Our two tests lead to similar conclusions. The behavior of neighbor firms is consistent with competitive bidding in the Pre-AWL period on tracts with active neighbor leases, and for major firms bidding on isolated tracts in the same period. The bidding of neighbor firms for deep water tracts in the AWL period is not consistent with competition, nor is the bidding of major firms for isolated tracts in that period. The evidence is mixed for bidding on shallow water tracts in the AWL period.

Appendices

A Estimating the Index Function

For each value of n, we estimate the index function $\lambda_n(\cdot)$ by estimating a random forest to predict the value the maximum rival bid M_{it}, which may be zero, associated with each positive bid B_{it}. We treat this as a prediction (regression) problem, with features (covariates) X_t and outcome $\ln(M_{it}+1)$. The random forest is estimated on the trimmed sample (see Appendix ***), restricted tracts receiving positive bids. We denote the resulting predicted value of the index at $X_t = x$ as $\hat{\lambda}_n(x)$.

The estimation procedure at each value of n is as follows, using parameter values $L = 50$ and $Q = 50$:

1. For $\ell = 1, \ldots, L$
   (a) Randomly split the sample into 2 equal sized folds.
   (b) Estimate 2 random forests which have Q trees each.[33] Each random forest is estimated on 1 of the 2 folds.
   (c) Combine the trees from the 2 random forests from (b) into one random forest. This random forest has $Q \times K$ trees. Call this random forest RF_ℓ.
2. Combine the L random forests $\{RF_1, \ldots, RF_L\}$ into a random forest denoted λ.

This approach has the feature that it will not overfit compared to the standard implementation of the random forest when training and predicting on the same data. When fitting a random forest, each data point in the training sample will be included in approximately 2/3 of the trees due to the bootstrapping. This means that when obtaining the predicted value for a given x, two thirds of the trees in the forest will be using its associated target outcome to generate a prediction. The index algorithm reduces the number of times a data point is used to train a tree. Under the parameters selected, each data point is used to train approximately 1/3 of the trees.[**need more clarity here. What supports the claim of not overfitting with this approach and overfitting with the standard approach? How do we get to the 2/3 and 1/3? What is meant by " reduces the number of times a data point is used to train a tree" (compared to what? and why is this desirable?)***]

The full set of features X_t used in the training of the random forest is as follows: ****insert the list; don't use abbreviations***

[33] See Appendix H for additional detail.

B Estimation Sample for Winner's Curse Tests

In constructing the sample used to estimate the auction model (for our tests for responses to the winner's curse), we must address two issues. One is that for the pre-AWL period we do not observe the set of tracts that is up for bid. We require an estimates of these "offer sets" in order to make valid inferences about the frequency with which bidders obtain signals below the screening value for the auction. Our sample for the pre-AWL period is then limited to the tracts in the estimated offer sets.

The second issue is that for tracts t for which $N_t = 1$, we do not directly observe whether $K(X_t, N_t) \geq 2$, as required for bids to be characterized by the conditions we derived in section 3. Thus, we provide a conservative approach for identifying the set $\mathcal{X} = \{x \in \mathbb{X} : K(x, 1) \geq 2\}$. In particular, we estimate a strict subset of $\mathcal{X}$ and limit our entire estimation sample (i.e., that use for all values of N_t) to tracts for which X_t falls in this subset.[34]

B.1 Offer Sets

For a given sale, define a tract as "available" if it is not under lease at the time of the sale. Define an available tract as "offered" if it is included among those up for bid in the sale. For each sale in the AWL period, the set of available tracts and the set of offered tracts are the same. However, this is not true of sales in the Pre-AWL period. Prior to a sale, MMS restricted leasing to specific areas and invited firms to nominate tracts in those areas. It then selected the set of tracts eligible for bidding in the sale, which we will call the *offer set*. We observe the total number of tracts in the offered set and the location of tracts that received bids, but not the location of tracts that did not receive bids, which is often half of the offered set.

To obtain a sample of Pre-AWL sales for which we have good estimates of the offer set, we first restrict the sample to Pre-AWL sales in which there is a significant number of available tracts with adjacent tracts under lease. There are 17 such sales.[35] If an Area Code in a sale had at least one tract that received a bid, we assume that all tracts not under lease (including isolated tracts) in that Area Code could be nominated by the firms. Define the set of tracts that meet this criterion as the *nomination set*. If no tract in an Area Code received a bid for a given sale, we exclude all tracts in that Area Code from the nomination

[34] Selecting the sample based on the observables X_t does not alter any of the theory or validity of our tests. Of course, the conclusions we reach apply only to the sample analyzed.

[35] Their sale umbers (with the year in parentheses) are 16 (1967), 33 (1974), 36 (1974), 38 (1975), 38A (1975), 41 (1976), 44 (1976), 47 (1977), 45 (1978), 51 (1978), 58 (1979), 58A (1979), A62 (1980), 62 (1980), A66 (1981), 66 (1981), 67 (1982).

set.

Next we estimate a random forest that predicts the event that a given tract in the nomination set receives a bid. We then take the τ tracts with the highest probabilities, where τ is the known number of tracts offered in the sale. These τ tracts are treated as those available in the sale.

The main***why "main" and not the actual list?*** covariates (features) used fore this random forest are as follows: ***we should spell out full versions of the abbreviations***

`year`, `oil` price, gas price, acreage, eastern gulf dummy, central gulf dummy, water depth, drainage dummy
max neighbor tract revenue within 1, 2 or 3 rings, max neighbor tract bid within 1, 2 or 3 rings,
neighbor tract drill dummy, mean number of bids on neighbor tracts, i_produced9, drilled_ever1, drilled_
NBactive_fract_by2, number of neighbor bid groups, num_bid_on_nbr, num_nbr_bid_firms, number o

The first three variables are characteristics of the sale, the next six are characteristics of the tract, and the remaining seventeen variables are characteristics of the tract's broad neighborhood, defined by up to three rings of tracts around the tract.

B.2 Trimming to Drop Single-Bidder Auctions

Under Assumption 2, a neighbor knows the number of other bidders it faces in the auction. Our goal is to drop one-neighbor tracts for which this number is zero, i.e., those for which $K(X_t, 1) = 1$. We do not observe $K(X_t, 1)$ directly for any tract, because at any given auction some bidders may receive signals below the relevant screening level. However, if $K(X_t, 1) = 1$, an implication is that

$$\Pr(M_{it} > 0 | X_t, N_t = 1, B_{it} = b) = 0$$

for all values of b of the neighbor firm's bid.

To exploit this observation, we estimate a random forest that predicts the probability that $M_{it} > 0$ on the subsample of tracts for which $N_t = 1$ and the neighbor firm does bid. Using the predicted probabilities and a parameter $\zeta \in (0, 1)$, we then define the set

$$\mathcal{X}(\zeta) = \{x \in \mathbb{X} | \Pr\{M_{it} > 0 | X_t = x, N_t = 1, B_{it} > 0\} > \zeta\}.$$

Here, larger values of the parameter ζ imply a more conservative restriction of the sample, providing greater certainty that tracts for which $K(X_t, 1) = 1$ are excluded. We set $\zeta =$******* and then restrict attention to the "trimmed sample" in which, for all values of

N_t, we include only tracts for which $X_t \in \mathcal{X}(\zeta)$.

C Estimation

****Revise this to reflect estimation of 2 of the 3, and then calculation of the third******

In this section, we provide the details on how we estimate the screening value functions.

1. Using the trimmed sample (see Appendix B), we estimate the probability that a tract receives a bid using a kernel regression (with a Gaussian kernel)

$$\mathbf{1}(\text{tract } t \text{ received a bid}) = \pi_N(\hat{\lambda}_N(X_t)) + \epsilon,$$

separately for each $N_t = 1, 2, 3$. We use the estimate $\hat{\pi}_N(\cdot)$.to control for sample selection when recovering the pivotal expected value distributions.

2. Using the trimmed sample, we estimate the screening value function $s^*(X, N)$ using a kernel regression (with a Gaussian kernel) of the form

$$1(B_{i,t}^{NB} = 0) = q_N(\widehat{\lambda}_N(X)) + \epsilon$$

separately for $N_t = 1, 2, 3$. We focus only on neighbors because we do not observe when non-neighbors enter and do not bid. Since signals are distributed uniform on the unit interval, $\widehat{q}_N(\widehat{\lambda}_N(X))$ is an estimate of $s^*(X, N)$. Given our symmetry assumption, this function also applies to non-neighbors who enter.

3. The probability that a neighbor firm's signal is below the screening value *conditional* on a tract receiving a bid is estimated on the trimmed sample using a kernel regression (with a Gaussian kernel)

$$\mathbf{1}(B_{i,t}^{NB} = 0) = p_N(\hat{\lambda}_N(X)) + \epsilon$$

separately for each $N_t = 1, 2, 3$.. We use the estimate $\widehat{p}_N(\cdot)$ to generate the mass point of the distribution function of pivotal expected values when we simulate the pivotal expected value distributions.

4. In each of the above regressions, we use the rule-of-thumb for conditional density estimation to compute the bandwidths.[36]

[36] We could use a cross-validation procedure to pick the bandwidth but this is computationally intensive (especially because we have to do multiple starts). If we use this approach in the bootstrap, we will have to use the same bandwidth across all bootstrap iterations.

These screening value functions are related through the law of total probability:

$$\begin{aligned}
1 - q_N(\widehat{\lambda}_N(X)) &= \Pr\{S_{i,t}^{NB} \geq s^*(X,N)\} \\
&= \Pr\{S_{i,t}^{NB} \geq s^*(X,N) | \max_{j=1,..,K_t}\{S_{j,t} \geq s^*(X,N)\} \Pr\{\max_{j=1,..,K_t}\{S_{j,t} \geq s^*(X,N)\} \\
&= (1 - p_N(\widehat{\lambda}_N(X))\pi_N(\hat{\lambda}_N(X_t)
\end{aligned}$$

However, we estimate the three functions separately for efficiency reasons.

D Simulation

We begin by constructing an empirical distribution for X. We first concatenate the X's across $N_t = 1, 2, 3$, and then trim this set by excluding values of X with index values $\widehat{\lambda}_N(X)$ below the 15th percentile and above the 85th percentiles of the distribution of λ_N for *all* $N_t = 1, 2, 3$. The purpose of the trimming is to exclude values of X that general index values in the tails of the distributions of λ_N. If we are close to boundaries of these distributions in the training data, the kernel estimates of the distribution of pivotal expected values are noisy.

1. The simulation procedure is as follows:
2. Draw x (with replacement) from the empirical distribution of X_t in the trimmed sample
3. Loop over $N_t = 1, 2, 3$ and do the following:
   (a) Draw D bids from $\hat{G}_B(\cdot|\widehat{\lambda}_N(x), N)$. We chose $D = 500$.
   (b) Compute the pivotal expected value w associated with each of the bids using GPV (with the LPV trick).
   (c) Let $\mathbb{W}(x, N)$ denote the set of pivotal expected values given $X = x$ and N and set $w^*(x, N) = \min\{W(x, N)\}$.[37]
4. Compute $w^*(x) = \max_N\{w^*(x, N)\} \times 1.05$.
5. For each N, replace each $w < w^*(x)$ in $\mathbb{W}(x, N)$ with $w^*(x)$.
6. Use the modified set of pivotal expected values to estimate the distribution of pivotal expected values for each N with a Gaussian kernel. Denote this distribution by $\widehat{F}_W^{**}(\cdot|x, N, s > s^*(x, N))$.

[37] If we draw enough bids, then $w^*(X, N)$ corresponds to the pivotal expected value of the reserve price.

(a) Define $\overline{w}(x,N) = E[V|S_{it}=1, \max_{j\neq i} S_{jt}=1, x, N]$ and $\overline{w}(x) = \min_N\{\overline{w}(x,N)\}$. Note that this latter value does not depend on N.

7. For each $w \in [w^*(x), \overline{w}(x)]$, estimate $\widehat{F}^*_W(\cdot|x,N)$, the marginal distribution of the pivotal expected values *in the selected sample,* using the formula

$$\widehat{F}^*_W(w|x,N) = \widehat{p}_N(\widehat{\lambda}_N(x)) + 1 - \widehat{p}_N(\widehat{\lambda}_N(x))\widehat{F}^{**}_W(w|x,N,s>s^*(x,N)).$$

8. Let $\widehat{F}_W(\cdot|x,N)$ denote the estimated marginal distribution of the pivotal expected values *in the population.* Then we correct for sample selection in $\widehat{F}^*_W(\cdot|x,N)$ using the formula

$$\widehat{F}^*_W(w|x,N) = \frac{\widehat{F}_W(w|x,N) - (1-\widehat{\pi}_N(\widehat{\lambda}_N(x)))}{\widehat{\pi}_N(\widehat{\lambda}_N(x))}. \tag{D.1}$$

We set $\widehat{F}^*_W(w|x,N)$ equal to 0 for $w < w^*(x)$ and equal to 1 for $w > \overline{w}(x)$.

9. Do the above steps S times.

Given estimates of $\hat{F}_W(\cdot|X,N)$, we compute $\hat{F}_W(\cdot|N)$ using the formula:

$$\widehat{F}_W(w|N) = \frac{1}{S}\sum_{s=1}^{S}\widehat{F}_W(w|x_s,N),$$

where s indexes the simulation draw. When evaluating a CDF at a value outside of the grid points, we use linear interpolation on a dollar scale (the units of the pivotal expected values).

E Stochastic Dominance Tests

We conduct two tests: the *specification test* (the null is that the data was generated via a Bayes-Nash equilibrium) and the *equality test* (the null of private values against the alternative of common values). Let $N_2 > N_1$ be integers representing the number of neighbors.

1. For each neighbor pair, compute the test statistic

$$CVM_{N_1,N_2} = \int_{\underline{w}}^{\overline{w}}\left(\left[\widehat{F}_W(w|N_2) - \widehat{F}_W(w|N_1)\right]_-\right)^2 dw,$$

where $[u]_- = \min\{u,0\}$ for the *specification test* and

$$CVM_{N_1,N_2} = \int_{\underline{w}}^{\overline{w}}\left(\left[\widehat{F}(w|N_2) - \widehat{F}(w|N_1)\right]_+\right)^2 dw,$$

where $[u]_+ = \max\{u, 0\}$ for the *equality test*. The **bounds of integration** are the largest support interval of the recovered pivotal expected value distributions. **Integration** is done using the Riemann sum midpoint rule with 1 million subdivisions.

2. Let $CVM_{N_1,N_2,0}$ be the test statistic computed from the data ("point estimate") and $CVM_{N_1,N_2,r}$ denote the test statistic computed under the bootstrap sample r.

   (a) Testing a pair of neighbor categories.

      i. Use the empirical distribution of $CVM_{N_1,N_2,r} - CVM_{N_1,N_2,0}$ to approximate the distribution of the test statistic under the null.
      The P-value is

$$\Pr(CVM_{N_1,N_2,r} - CVM_{N_1,N_2,0} \geq CVM_{N_1,N_2,0}) = \Pr(CVM_{N_1,N_2,r} \geq 2 \times CVM_{N_1,N_2,0}).$$

      Notice that we use a greater than *or equal to* sign. The "equal to" is included because if all the test statistics are zero and the "equal to" is dropped, the P-value will be zero, which does not make sense in this context.

   (b) Testing all pairs (the three pairs possible when testing pairs for $N_t = 1, 2, 3$) for a sample period.

      i. Since we are testing multiple hypotheses, say H_1 through H_K, we use the Bonferroni correction. A desired overall type 1 error rate of at most α can be guaranteed by testing each null hypothesis at level α/K. In our case, where we test using pivotal expected value distributions for $N_t = 1, 2, 3$, $K = 3$, as we test equivalence or ordering of the distributions for each pair of N's.

      ii. *This approach does not account for the Bonferroni correction.* Let

$$CVM_r = \max_{\text{across pairs } N_1,N_2} \{CVM_{N_1,N_2,r}\}.$$

      Then the **P-value** is

$$\Pr(CVM_r - CVM_0 \geq CVM_0) = \Pr(CVM_r \geq 2 \times CVM_0).$$

F Monotone Screening Tests

1. In the simulation described above, compute and store $\widehat{q}_N(\widehat{\lambda}_N(x_s))$ for each $s = 1, .., S$. Given these estimates, we compute

$$\widehat{q}_N = \frac{1}{S}\sum_{s=1}^{S} \widehat{q}_N(\widehat{\lambda}_N(x_s)).$$

2. Compute $\widehat{q}_{\mathbb{N}_1,r}$ and $\widehat{q}_{\mathbb{N}_2,r}$ for the point estimate and each replication (this integrates out the x's). Let

$$\widehat{\mu}_{N_1,N_2,0} = \widehat{q}_{N_2,0} - \widehat{q}_{N_1,0}$$

be the test statistic computed from the data point estimate and let $\widehat{\mu}_{N_1,N_2,r}$ denote the test statistic computed from bootstrap sample r.

(a) Testing the null of equality of the mass point against the alternative that the screening level is decreasing in the number of neighbors:

$$H_0 : \widehat{\mu}_{N_1,N_2,,r} \geq 0;$$
$$H_1 : \widehat{\mu}_{N_1,N_2,r} < 0.$$

i. The pairwise test P-value is

$$\Pr(\widehat{\mu}_{N_1,N_2,,r} - \widehat{\mu}_{N_1,N_2,,0} \leq \widehat{\mu}_{N_1,N_2,,0}) = \Pr(\widehat{\mu}_{N_1,N_2,,r} \leq 2 \times \widehat{\mu}_{N_1,N_2,,0}).$$

ii. The pooled test involves defining

$$\widehat{\mu}_r = \min_{\text{across pairs } N_1,N_2} \{\widehat{\mu}_{N_1,N_2,r}\}$$

and calculating the P-value as

$$\Pr(\widehat{\mu}_r - \widehat{\mu}_0 \leq \widehat{\mu}_0) = \Pr(\widehat{\mu}_r \leq 2 \times \widehat{\mu}_0).$$

(b) Testing the null of equality of the mass point against the alternative that the screening level is increasing in the number of neighbors:

$$H_0 : \widehat{\mu}_{N_1,N_2,r} \leq 0;$$
$$H_1 : \widehat{\mu}_{N_1,N_2,r} > 0.$$

i. The pairwise test P-value is

$$\Pr(\widehat{\mu}_{N_1,N_2,,r} - \widehat{\mu}_{N_1,N_2,,0} \geq \widehat{\mu}_{N_1,N_2,,0}) = \Pr(\widehat{\mu}_{N_1,N_2,,r} \geq 2 \times \widehat{\mu}_{N_1,N_2,,0})$$

ii. The pooled test involves defining

$$\widehat{\mu}_r = \max_{\text{across pairs } N_1,N_2} \{\widehat{\mu}_{N_1,N_2,r}\}$$

and calculating the P-value as

$$\Pr(\widehat{\mu}_r - \widehat{\mu}_0 \geq \widehat{\mu}_0) = \Pr(\widehat{\mu}_r \geq 2 \times \widehat{\mu}_0).$$

F.1 Bootstrap Implementation (to be done)

The procedure is described for a single sample period (e.g. "Pre-AWL") where each neighbor category N has T_N auctions.

Do the following steps $1,000$ times:

1. Bootstrap sampling, index estimation, and nonparametric estimation.

   (a) For each N, construct a bootstrap sample by randomly drawing T_N auctions (leases) from the trimmed sample with replacement.

   (b) For each N, fit a random forest to obtain $\hat{\lambda}_N(X)$.

   (c) For each N, estimate a kernel regression of $\mathbf{1}(B_t^{NB} = 0)$ on $\hat{\lambda}_N(X)$. Denote this estimate kernel regression that predicts the probability of non-participation of neighbors $\hat{p}_N(\cdot)$. The bandwidth is chosen via the rule of thumb on the bootstrap sample.

   (d) For each N, nonparametrically estimate $G_B(\cdot|\hat{\lambda}_N(X), N)$, $G_{M|B}(\cdot|B, \hat{\lambda}_N(X), N)$, and $g_{M|B}(\cdot|B, \hat{\lambda}_N(X), N)$. The bandwidths h_M, h_B, and h_λ are chosen via the rule of thumb on the bootstrap sample.

G Asymmetry

Our model assumed that bidders have access to the same signal technology. This permits neighbors and non-neighbors to have different costs of signal acquisition—e.g., extending the seismic analysis performed on one tract to an adjacent tract may be less costsly than analysis in a new neighborhood. However, we rule out asymmetries in the quality of information obtained when performing the analysis. Although this is natural with respect to seismic information, posible concern with this symmetry assumption arises from the fact that some of the leases are adjacent to tracts that have have already been drilled. Although drilling outcomes are publicly observable and accounted for in our covariates, a neighbor that has drilled a well itself might have better information than other bidders.[38] As a robustness check, here we we will repeat our analysis restricting the sample to undrilled neighborhoods, where such concerns do not arise.

****do this; could reference similar robustness check in CHS****.

[38] Drill cores are a possible source of such information. Firms submit a drill core report to BOEM but drill cores themselves are not submitted. The drill core information is especially important in evaluating dry holes. However, firms who want to examine cores can sign an agreement to do so with the lease owner.

Supplemental Appendices

H Random Forest Estimation Algorithm

do we need this? isn't this standard?*

The random forest algorithm is as follows (see, for example, Efron and Hastie (2016)). The training data are given by $\boldsymbol{d} = (\boldsymbol{X}, \boldsymbol{y})$, where $\boldsymbol{X}$ is an $\rho \times p$ matrix of covaraties (features) and $\boldsymbol{y}$ is a $\rho \times 1$ vector of outcomes (targets). Let Q denote the desired number of trees in the forest.

1. For $q = 1, \ldots, Q$:

   (a) Create a bootstrap version of the training data $\boldsymbol{d}_q^*$.

   (b) Grow a maximal-depth tree using $\boldsymbol{d}_q^*$, sampling ϕ of the p features at random prior to making each split. For a classification tree, we use $\phi = \lfloor\sqrt{p}\rfloor$; for a regression tree, we use $\phi = \lfloor p/3 \rfloor$. Growing a tree involves recursively splitting variables at a certain point to maximize the reduction in either the mean squared error (if it is a regression tree) or node impurity (if it is a classification tree). The algorithm is "greedy," splitting the variable which will improve the fit the best. The stopping rules stipulate stopping either (i) if splitting a node results in one of the two resulting nodes having too few data points (1 for classification or 5 for regression) and (ii) a branch from the tree stump has hit a predefined number—we use 32—of levels

   (c) let $\hat{y}_q(x)$ denote the resulting prediction at the point x

2. The random forest prediction at any point x is the average

$$\hat{y}_{rf}(x) = \frac{1}{Q} \sum_{q=1}^{Q} \hat{y}_q(x)).$$

I eliminated the references to the bootstrap frequency vector $\boldsymbol{w}_b^*$, which seem to play no role; also eliminated duplicative notation;*

Test Tables

Table 7: Non-participation Test

Sample Period Pair	Point Est	P-value	5%	50%	95%
Pre-AWL_1_2	0.134	0	0.090	0.134	0.180
Pre-AWL_1_3	0.196	0	0.146	0.192	0.237
Pre-AWL_2_3	0.062	0.002	0.019	0.057	0.096
AWL-deep_1_2	−0.054	0.995	−0.090	−0.049	−0.011
AWL-deep_1_3	−0.006	0.637	−0.054	−0.002	0.052
AWL-deep_2_3	0.048	0.067	−0.013	0.048	0.102
AWL-shallow_1_2	−0.036	0.941	−0.075	−0.039	−0.007
AWL-shallow_1_3	0.013	0.207	−0.025	0.010	0.041
AWL-shallow_2_3	0.049	0.004	0.020	0.049	0.079
Pre-AWL	0.196	0	0.146	0.192	0.237
AWL-deep	0.048	0.067	−0.009	0.049	0.102
AWL-shallow	0.049	0.004	0.021	0.049	0.079

Tests the null of equality against the alternative of non-participation increasing in the number of neighbors.

Table 8: Specification Test

Sample Period Pair	Point Est	P-value	5%	50%	95%
Pre-AWL_1_2	0	1	0	0	0
Pre-AWL_1_3	0	1	0	0	0
Pre-AWL_2_3	0	1	0	0	0
AWL-deep_1_2	73,371	0.074	37,879	87,078	156,228
AWL-deep_1_3	32,788	0.197	7,806	37,912	102,088
AWL-deep_2_3	0	1	0	0	1,398
AWL-shallow_1_2	1,399	0.859	669	10,350	36,553
AWL-shallow_1_3	0	1	0	0	160
AWL-shallow_2_3	0	1	0	0	0
Pre-AWL	0	1	0	0	0
AWL-deep	73,371	0.075	41,054	87,978	156,266
AWL-shallow	1,399	0.859	669	10,350	36,553

Pre-AWL integration bounds: [1,578,318; 194,755,842].
AWL-deep integration bounds: [1,042,123; 21,515,052].
AWL-shallow integration bounds: [1,047,976; 22,302,226].

Table 9: Equality vs Positive Inequality Test

Sample Period Pair	Point Est	P-value	5%	50%	95%
Pre-AWL_1_2	785,528	0.245	291,658	1,063,045	2,520,324
Pre-AWL_1_3	2,261,651	0.045	1,089,579	2,301,571	4,400,170
Pre-AWL_2_3	386,017	0.012	48,264	261,373	650,180
AWL-deep_1_2	0	1	0	0	0
AWL-deep_1_3	0	1	0	0	1,572
AWL-deep_2_3	8,152	0.395	46	11,983	53,679
AWL-shallow_1_2	0	1	0	0	3
AWL-shallow_1_3	18,987	0.005	106	6,380	23,994
AWL-shallow_2_3	29,034	0.070	12,770	32,416	61,895
Pre-AWL	2,261,651	0.045	1,089,579	2,301,571	4,400,170
AWL-deep	8,152	0.395	46	11,983	53,679
AWL-shallow	29,034	0.070	12,875	32,462	61,895

Pre-AWL integration bounds: [1,578,318; 194,755,842].
AWL-deep integration bounds: [1,042,123; 21,515,052].
AWL-shallow integration bounds: [1,047,976; 22,302,226].

Figure 1: Sales Revenue and Royalty Payments From Gulf of Mexico in Two Year Increments

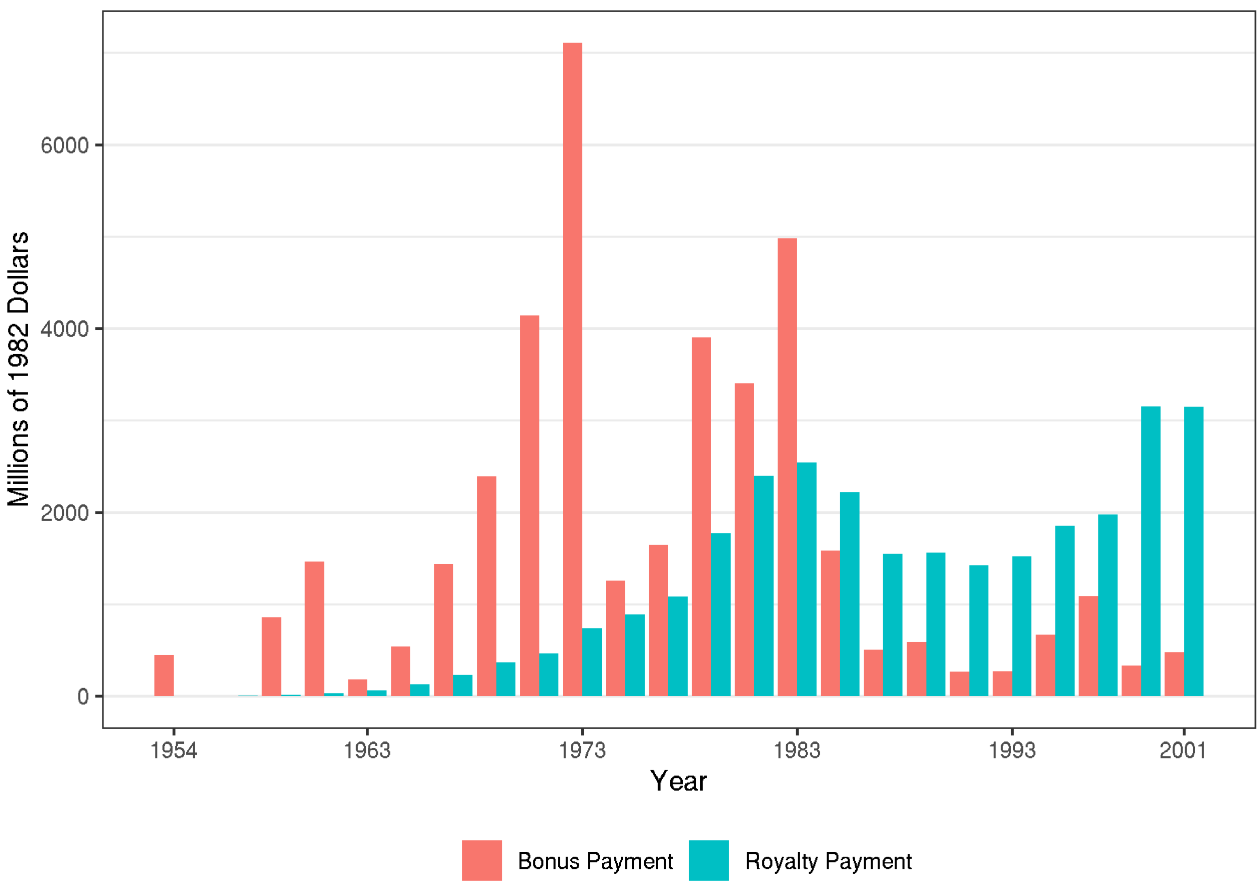

Figure 2: Fraction of Tracts with Neighbors

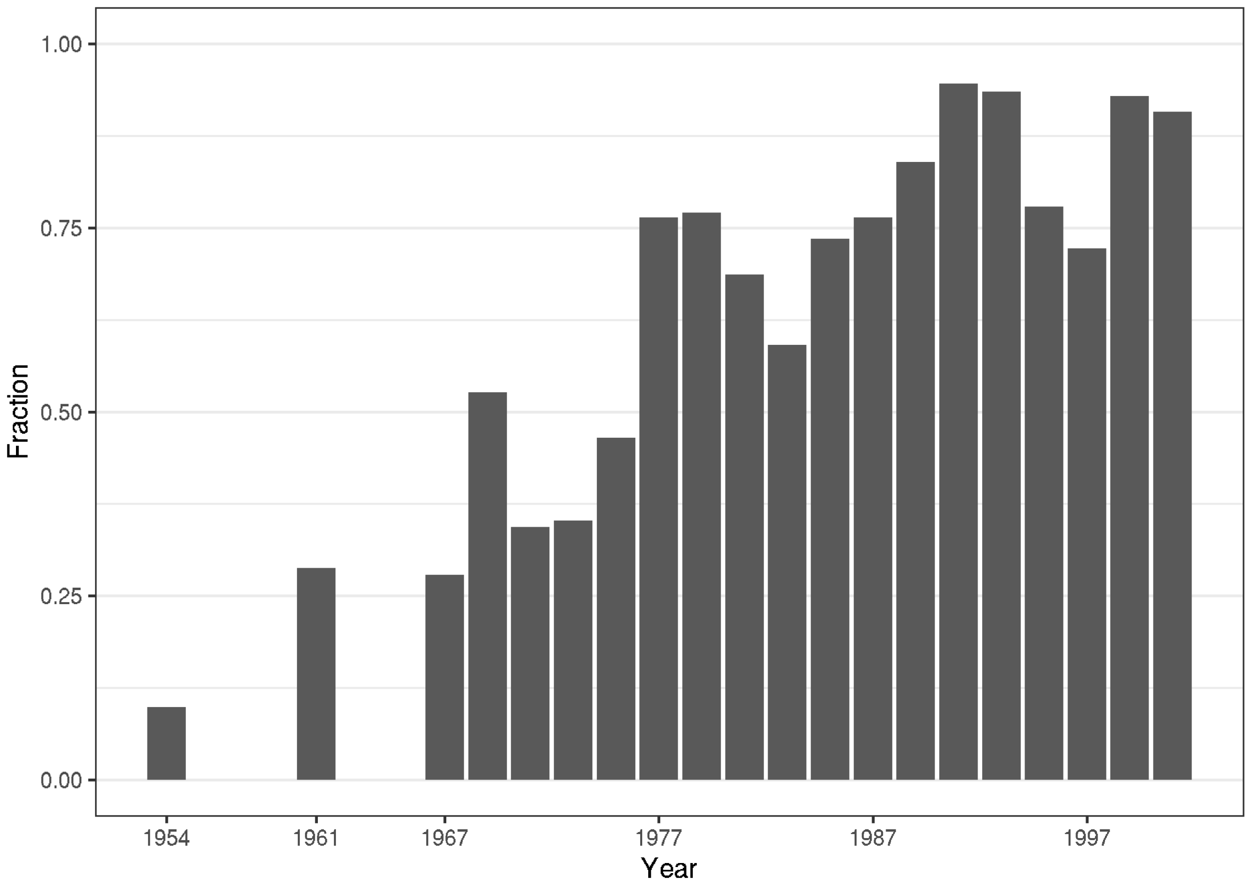

Figure 3: Estimates of the Distribution of Pivotal Expected Values for Pre-AWL

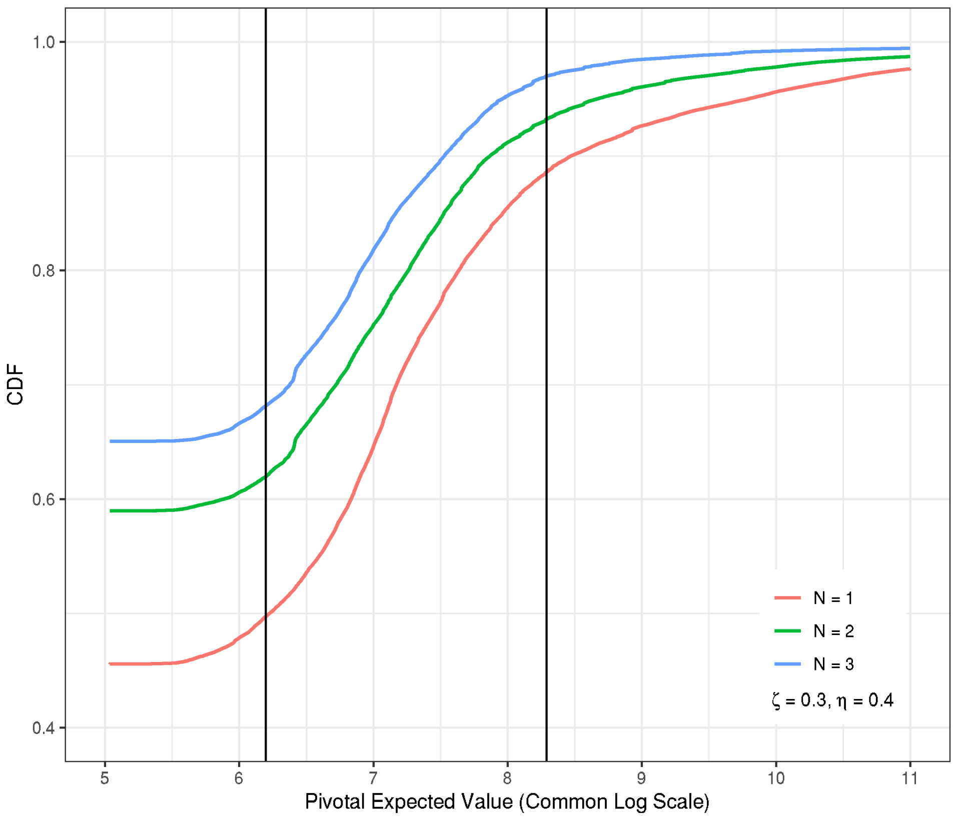

Figure 4: Estimates of the Distribution of Pivotal Expected Values for AWL-Deep

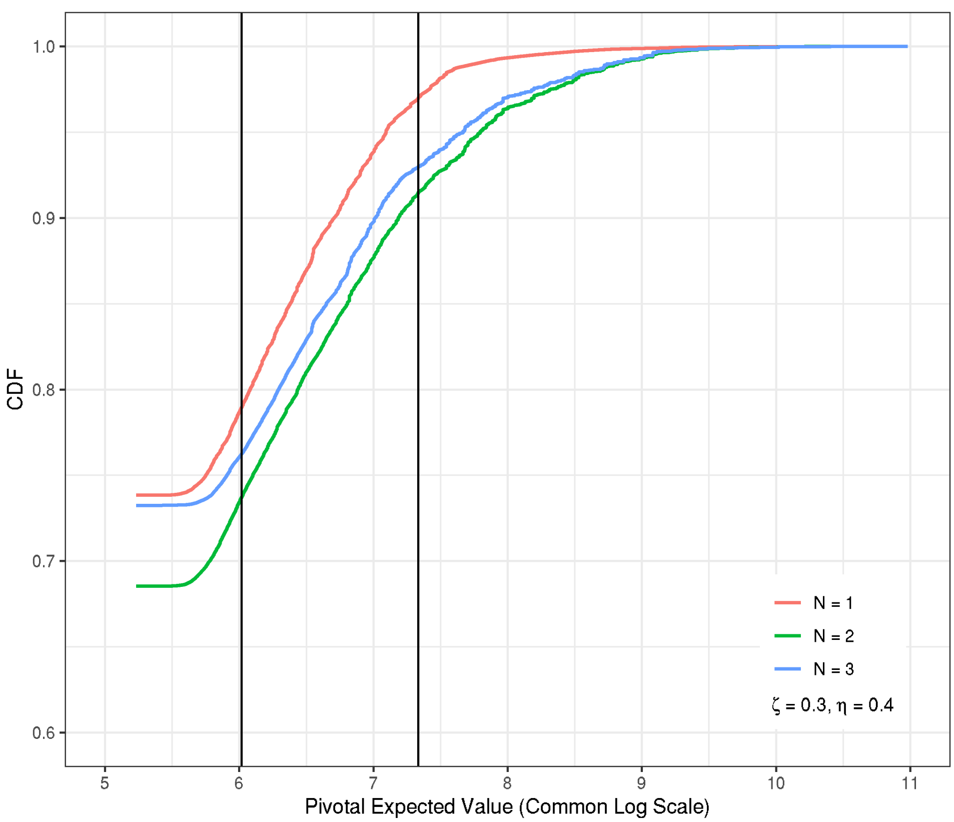

Figure 5: Estimates of the Distribution of Pivotal Expected Values for AWL-Shallow