THE UNIVERSITY OF CHICAGO

IDENTIFICATION AND INFERENCE IN FIRST-PRICE AUCTIONS WITH COLLUSION

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF ECONOMICS

BY
KARL EDWARD SCHURTER

CHICAGO, ILLINOIS
JUNE 2017

To my family for their encouragement and

to my wife for her inexhaustible love and support.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I thank Ali Hortaçsu, Azeem Shaikh, and Brent Hickman for their advice and support. I also thank Stéphane Bonhomme, Robert Porter, Michael K. Price, Derek Neal, and members of the IO and econometrics working groups at the University of Chicago for many helpful comments and suggestions. I gratefully acknowledge financial support from the Bradley Foundation's Fellowship Program and the Schultz Endowment.

ABSTRACT

This dissertation develops a method to detect collusion and estimate its effect on the seller's revenue in first-price auctions with independent, private valuations. The challenge is that collusion may be difficult to detect because colluders can use a simple and costless strategy to make their bids appear competitive. If the econometrician observes an exogenous shifter of the level of competition in the auction in addition to the winning bids, a statistical test for collusion by a given bidder can be formulated as a test of independence between the exogenous shifter and the valuations that rationalize its bids under the null hypothesis that it is not colluding. Simulations confirm this test performs well even when colluders attempt to disguise their behavior. I then adopt a multiple hypothesis testing framework to simultaneously test for collusion bidder by bidder. By controlling the probability of making one or more type I errors, the set of rejected hypotheses serves as a lower confidence bound on the set of colluders. To produce a lower confidence bound on the cost of collusion, I use consistent estimates of the bidders' valuation distributions to numerically solve for the seller's expected revenues in auctions with and without collusion. To provide an example of this identification strategy, I use exogenous variation in the reserve prices at British Columbia's timber auctions to estimate the extent of collusion in the years preceding a lumber trade dispute between the United States and Canada.

CHAPTER 1

INTRODUCTION

Consider the possibility that some bidders in a first-price, sealed-bid auction are coordinating their bids to increase the probability that they win at a lower price. Though the seller may suspect bidders have entered into collusive agreements with one another, collusion might be impossible to detect based on a statistical analysis of their bids. In fact, colluders can generally use a simple and costless strategy to ensure their bids appear to be consistent with a model of competitive bidding. Specifically, they could always bid as if they were trying to compete optimally against the non-colluders, even when they know they will lose to one of their co-conspirators. The question is what additional assumptions or data are needed to identify colluders and estimate their effect on the seller's revenue.

Answers to this question are in high demand because collusion poses a significant problem in a wide range of markets and auction formats. For example, Kawai and Nakabayashi (2014) identify 1,000 firms that submitted bids that are inconsistent with a competitive model of Japanese procurement auctions. They estimate collusion increased the government's construction costs by \$720 million (or 8.4%). As another example, Porter and Zona (1999) estimate that a collusive bidding ring inflated the cost of supplying milk to Ohio school districts by 6.5% between 1980 and 1990. And, in a sample of ascending stamp auctions, Asker (2010) estimates that collusion reduced the average revenue for the seller by 4%. These studies are somewhat unusual, however, because Kawai and Nakabayashi were able to exploit idiosyncratic auction rules to catch collusive firms, and Porter and Zona and Asker could verify their findings with legal testimony. Otherwise, when such opportunities do not exist, we have no guarantee, ex ante, that any method would be able to detect collusion in a sufficiently large sample of auctions.

More precisely, collusion might never be detected unless more restrictive assumptions are made. Pesendorfer (2000) and Bajari and Ye (2003) observe that if bidders' willingnesses to pay for the object to be sold are assumed to be identically distributed random variables conditional on observable characteristics competitive models predict bidders' strategies will be symmetric, as well. By contrast, the presence of a collusive bidding ring would create asymmetry among the bidders. Accordingly, they test for collusion by testing for asymmetry in bidding behavior. However, assuming bidders are identical is not plausible in all applications. The case for collusion will be stronger when it can be detected under more general modeling assumptions, especially when supplementary evidence of anticompetitive behavior is unavailable.

Motivated by this problem, I propose a strategy for identifying colluders in standard first-price auctions under the assumption that bidders have independently distributed private valuations (IPV). In particular, the strategy exploits exogenous variation in the level of competition across auctions, which may be induced, for example, by an increase in a binding reserve price or the set of eligible bidders. In response to this increased competition, I show collusive and competitive models predict different comparative statics. Therefore, a collusive bidder can be identified through an analysis of the change in its bids following an exogenous increase in the distribution of its competitors' bids. Based on this identification argument, I suggest a statistic to test the null hypothesis that a given bidder is not colluding, and demonstrate how to use it to diagnose the extent of collusion in first-price auctions in which a ring is suspected but not yet proven to be operating.

I am not the first to use exogenous variation in the level of competition at an auction to test assumptions of the competitive IPV auction model. For example, studies have used variation in the number of bidders to test for interdependence in bidders' valuations (Athey and Haile, 2002; Haile et al., 2003) and to test the assumption that bidders are risk neutral (Bajari and Hortaçsu, 2005; Guerre et al., 2009). I am also not the first to develop tests for collusion in first-price auctions (Porter and Zona, 1993, 1999; Pesendorfer, 2000; Bajari and Ye, 2003; Aryal and Gabrielli, 2013; Kawai and Nakabayashi, 2014). But these tests rely on auxiliary data, symmetry assumptions, and non-standard auction rules to consistently test for collusion. Hence, my contribution to the literature is to suggest commonly available sources of exogenous variation that can be used to identify colluders in standard first-price auctions when bidders have asymmetrically distributed valuations.

The theoretical details of the identification strategy are developed in the next chapter of this dissertation. Section 2.1 works through an example that illustrates the identification problem and the proposed solution, which I then compare with previous approaches to collusion detection in section 2.2. In section 2.3, I present a formal model of first-price auctions with collusion and discuss assumptions about how bidders might collude. The theorems in section 2.4 then use this model to generalize the lessons learned from the example.

In section 2.5, I derive a consistent test for collusion from a test for dependence between the competitively rationalizing valuations and an exogenous instrument. Evidence from simulations indicates this testing procedure has statistical power to detect collusion even when colluders attempt to disguise their behavior. In comparison, when colluders always bid as if they were seriously competing against the non-ring competition, tests for dependence in bidders' strategies reject false null hypotheses with the same probability as true null hypotheses.

Ultimately, the goal is not merely to detect collusion, but also to estimate its effect on prices. In section 2.6, I suggest one such estimate can be obtained by sending the family-wise error rate (FWER)—the probability of falsely rejecting one or more of the null hypotheses—to zero as the data grow.[1] Intuitively, the testing procedure promises to become increasingly conservative as the sample size grows, such that the probability of falsely accusing a non-collusive bidder tends toward zero as the sample of auctions grows. As long as the test does not become too conservative too quickly, the probability of failing to detect collusion will also tend toward zero. Therefore, the set of rejected null hypotheses consistently estimates the members of the collusive ring. Given this estimate of the ring, standard methods from the empirical auction literature will then deliver consistent estimates of each bidder's private

1. Marmer et al. (2016) also adopts a multiple testing procedure that asymptotically controls the FWER to detect collusion in English auctions.

valuation distribution. In turn, the estimated private valuation distributions can be used to predict bids in counterfactual equilibria. Most notably, the econometrician can consistently estimate the cost of collusion by solving for the equilibrium price distribution that would prevail if none of the bidders colluded.

The estimator would be more useful, however, if it were accompanied by an estimate of its precision. Fortunately, a multiple hypothesis testing framework is well equipped to provide lower confidence bounds on the set of colluders. In fact, controlling the FWER is equivalent to controlling the probability that the true set of colluders contains the set of bidders for whom the null hypothesis of no collusion is rejected. Moreover, because the cost of collusion increases with the size of the ring, a lower confidence bound on the ring produces a corresponding lower confidence bound on the cost of collusion.

In chapter 3, I apply the identification strategy to British Columbia's timber auctions, where controversy surrounding the fairness of its timber prices has persisted for decades. Section 3.1 discusses the institutional details of the British Columbian timber industry. In section 3.2, I find evidence that two of the most active firms in the industry are colluding with each other. Together they form a 90%-confidence bound on the collusive bidding ring. Though collusion's overall effect on the region's timber prices is certainly smaller, I estimate it reduced the expected revenue at some auctions by more than 1.2%.

Before arriving at this final answer, I address several issues that commonly arise in empirical applications of structural auction models. Because the methods I use to relate the auction theory to the data are not standard in the literature, I briefly summarize them here.

The first issue is that the set of potential bidders is not directly observed in the data. Instead, the data only contain the set of bidders who actually participated in the auction, which is a lower bound on the set of bidders that considered competing. Because bidders' optimal strategies depend on their beliefs about their competition, the set of potential bidders must be estimated. More precisely, any variation in the potential competition must be reflected by variation in observable covariates. To construct these observables, I use a hierarchical clustering algorithm to produce a nested sequence of markets definitions. These market definitions range from 31 distinct submarkets to a single market and are formed by iteratively joining submarkets in which each bidder's propensity to participate is relatively constant. Assuming bidders have rational expectations and no other means of predicting who might compete in the auction, I argue the bidders' beliefs about the composition of their potential competitors are constant within the submarkets generated by one of the market definitions. I can then condition on these submarkets in order to condition on bidders' expected competition. The alternative market definitions can be used to check for robustness.

Second, I assume that a vector of 18 covariates captures the heterogeneity in timber across auctions. In principle, the distribution of the bidders' valuations could be nonparametrically identified conditional on these covariates, but the curse of dimensionality would lead to very imprecise estimates. To circumvent this issue, I assume that the covariates are independent of the idiosyncratic component of the bidders' valuations and that they enter in a linear, additively separable manner. Under these assumptions, an ordinary least-squares (OLS) regression of bids on covariates would produce consistent estimates of the linear coefficients if the reserve price were not binding (Athey and Haile, 2007). If the reserve price binds, however, OLS will suffer from an omitted-variable bias because, holding the reserve price fixed, more favorable covariates increase the range of valuations with which bidders enter the auction. To accommodate this selection issue, I estimate a partially linear single-index model to obtain consistent estimates of the linear coefficients in the presence of a binding reserve price.

Third, I propose a statistic to test the null hypothesis that valuations are independent of the reserve prices. The statistic is based on a conditional Kendall's τ measure of correlation that Tsai (1990) introduced to test for independence when the data are truncated. I demonstrate that this idea can be extended to incorporate the bidders' participation decisions to add power to the test for dependence between valuations and reserve prices. Though I suggest this approach as a test for collusion, it could also be applied in other contexts. For example, it could provide an additional moment to target in a method of simulated moments or to test the assumption of risk neutrality.

Finally, the purpose of estimating a structural auction model is typically to predict bidding behavior in counterfactual equilibria, which involves solving the system of differential equations that characterizes the Bayes-Nash equilibrium in a first-price auction. This solution must be found numerically because the system of differential equations does not generally have an analytic solution when bidders' valuations are not identically distributed. Moreover, the numerical solution method must allow for the fact that the derivative of the bidders' inverse bidding strategies is unbounded when the reserve price binds. I therefore use a change of variables to transform the problem into a bounded system of differential equations. Then, extending the idea proposed by Fibich and Gavish (2011) to the case in which bidders' valuations have distinct supports, I convert the problem into a boundary value problem in order to solve for the equilibrium bid distributions. Because the transformed system only requires evaluating the quantile function of each bidder's valuation distribution, not its density, and because the quantile function can be nonparametrically estimated at a faster rate, this solution method can improve the estimation of counterfactual equilibria.

CHAPTER 2

A METHOD OF DETECTING COLLUSION IN FIRST-PRICE AUCTIONS

2.1 An Example of the Identification Problem and Its Solution

Before formally introducing the model, I present a simple example that illustrates the identification problem and the proposed solution. To abstract away from issues regarding estimation, I assume the econometrician directly observes the distribution of the bids. From these data, the econometrician's goal is to infer the joint distribution of the bidders' valuations for the object to be sold. I assume these valuations are private and independently distributed.

The first observation to make is that different valuation distributions can rationalize the data depending on which bidders are assumed to be colluding. In fact, this observation proves the collusive model is not identified from the joint distribution of the bids. See Definition 1 and Theorem 1 in section 2.4 for a precise statement of this result.

> Consider an auction with a nonbinding reserve price of zero and three bidders—labeled 1, 2, and 3. Let V_i and B_i denote bidder i's random valuation and bid for $i \in \{1, 2, 3\}$. Suppose the econometrician observes that their bids have a joint distribution function $S(b_1, b_2, b_3) = \sqrt{2b_1} \cdot \sqrt{2b_2} \cdot 2b_3$ for $b_i \in [0, 1/2]$. Because the bids are independently distributed, tests for dependence would not detect collusion.
>
> *Case I:* Suppose none of the bidders are colluding. If bidder i privately observes the realization of its valuation and chooses a bid to maximize its expected profits
>
> $$(v_i - b) \cdot P\{i \text{ wins with a bid of } b\} , \tag{2.1}$$
>
> then there exists a unique distribution of valuations that rationalizes the bid distribution. In this case, bidder 1 (or, symmetrically, bidder 2) wins with probability $(2b)^{3/2}$ when it bids b. The first-order conditions for bidders 1 and 2's profit-maximization problem then implies they will bid $b = 3v/5$. Their marginal

valuation distributions are therefore given by

$$F_1^I(v) = F_2^I(v) = P\{V_i \leq v\} = P\{B_i \leq 3v/5\} = \sqrt{6v/5}$$

for v between 0 and 5/6. Analogously, the probability that bidder 3 wins with a bid of b is $2b$, its optimal bid is $b = v/2$, and its marginal valuation distribution must be uniform between 0 and 1.

Case II: Alternatively, assume bidders 1 and 2 are jointly maximizing their expected profits.[1] Because nothing about bidder 3's profit-maximization problem has changed, no change occurs in the valuation distribution that rationalizes bidder 3's bids. On the other hand, the coalition of bidders 1 and 2 will win with probability $2b$ whenever the greater of their bids is b. The colluder with the higher valuation should therefore optimally bid $b = \max\{v_1, v_2\}/2$, whereas the other may submit any "phantom" bid below this amount.[2] Because these phantom bids may be unrelated to the colluders' valuations, I ignore the lower of b_1 and b_2. Under the assumption that the valuations are independent, however, the higher bid still contains enough information to infer their valuation distributions. In particular, the marginal valuation distributions $F_1^{II}(v) = F_2^{II}(v) = \sqrt{v}$ uniquely solve

$$\begin{aligned} P\{B_1 \leq b, B_2 \leq B_1\} = b = P\{V_1/2 \leq b, V_2 \leq V_1\} \\ P\{B_2 \leq b, B_1 \leq B_2\} = b = P\{V_2/2 \leq b, V_1 \leq V_2\}\ . \end{aligned}$$

Hence, bidders 1 and 2's valuation distribution could be one of two possibilities depending on whether they are colluding.[3]

This example demonstrates the bids generated from competitive and collusive auctions can be observationally equivalent. Though, if a collusive model indeed generated these

1. I discuss this assumption in further detail in section 2.3.

2. If $v_A = v_B$, bidders 1 and 2 can arbitrarily break the tie.

3. Interestingly, collusion affects the support of bidders 1 and 2's inferred valuation distribution. This observation also holds more generally. Therefore, it is possible to identify colluders under the assumption that bidders are symmetric, or even under the weaker assumption that their valuations share the same upper extremity of their supports. A test of the null hypothesis that bidder i is colluding could be formulated as a test of $H_0 : \bar{v}_i = \max_j \bar{v}_j$ versus $H_1 : \bar{v}_i < \max_j \bar{v}_j$, where $\bar{v}_i$ denotes the upper limit of the support of i's valuations.

data, the colluders must have been careful enough to submit phantom bids that could be rationalized without collusion. In particular, despite the fact that any phantom bid below the other's bid would have maximized the colluders' expected profits, bidders 1 and 2 ensured their marginal bid distributions were independent and solved the above system of equations. Because equivalence in distribution is weaker than almost sure equivalence, this fact does not determine their phantom bidding strategies as a function of their private valuations. However, the fairly simple strategy of always bidding one-half of their valuations could have produced these data.

In other words, bidders 1 and 2's strategy may have been to "act natural" and always bid as if they were trying to win against bidder 3. This strategy guarantees that their bids will be independent because each bid is only a function of the bidder's own independent valuation. Moreover, this strategy maximizes the ring's expected profits because bidders 1 and 2 use the same strictly increasing bid function. Therefore, the colluder with the higher valuation will always outbid the other. Disguising collusion is thus costless and does not require any further coordination on the part of the colluders.

Generalizing from the example, independent bid distributions can typically be rationalized by as many valuation distributions as there are possible configurations of the collusive ring. Each configuration of the ring corresponds to different distributions of the bidders' highest competing bid, which in turn imply different optimal bidding strategies. So without knowing who is actually competing against whom, I cannot correctly map bids to valuations. Hence, the distribution of valuations is not identified from the distribution of bids.

Remarkably, however, a competitive bidder's marginal valuation distribution can be correctly inferred regardless of any collusion among the other bidders at the auction. Because competitive bidders do not directly care whether they are competing against a single bidding ring or the maximum bid among many independent bidders, they will employ the same bidding strategy as long as the distribution of their highest competing bid is the same. As a result, the econometrician can infer a given bidder's valuation distribution under the hypothesis that it is not colluding. I formalize this result in Lemma 1 in section 2.4.

But if the bidder is colluding, a competitive model of bidding will underestimate the true valuation distribution because it underestimates how much the colluders "shade" their bids below their valuations. In the example, bidders 1 and 2 were presumed to have bid three-fifths of their valuation under the hypothesis of competition. But when they are in fact colluding, their bids are actually further below their valuations. As a result, bidder 1 and 2's true valuation distributions—F_1^{II} and F_2^{II}—first-order stochastically dominate F_1^I and F_2^I.

Furthermore, the colluders' valuation distributions will be more severely underestimated when the ring is stronger relative to its non-ring competition. Therefore, if a bidder is colluding, the valuations that competitively rationalize its bids will covary with the level of competition at the auction. Loosely speaking, by suppressing a larger share of the potential competition, the ring exaggerates the difference between the ring's competing distribution in the competitive and collusive cases.

Theorem 2 formalizes this discussion to prove the main result. Namely, there exists a unique configuration of the ring that simultaneously rationalizes all of the bidders' responses to an exogenous change in the level of competition. And, in particular, the econometrician can test the null hypothesis that a given bidder is not colluding using a test of independence between the valuations that competitively rationalize its bids and an exogenous shifter of its competition.

The proof of Theorem 2 is provided in the main text for the interested reader. However, an extension of the example succinctly illustrates the argument.

> Suppose bidders 1 and 2 are in fact colluding. Suppose further that they now face stronger competition because bidder 4 enters the auction. Assume that bidder 4's valuation is distributed uniformly between 0 and 1, while the other bidders' valuation distributions are unchanged. In other words, assume that the arrival of bidder 4 is exogenous with respect to each incumbent bidder's valuations.
>
> In the Bayes-Nash equilibrium where bidders 1 and 2 are colluding to maximize their combined expected profits, all serious bidders optimally bid two thirds of

their valuation. If bidders 1 and 2 always bid as if they were trying to win against bidders 3 and 4, the joint distribution of bids will be

$$S(b_1, b_2, b_3, b_4) = \sqrt{3b_1/2} \cdot \sqrt{3b_2/2} \cdot 3b_3/2 \cdot 3b_4/2$$

for $b_i \in [0, 2/3]$.

As in the three-bidder auctions, the standard uniform valuation distribution rationalizes bidder 3's bids under the null hypothesis that Bidder 3 is not colluding. In contrast, under the false null hypothesis that bidder 1 is not colluding, the econometrician observes that the probability it wins with a bid of b is $(3b/2)^{5/2}$. Its optimal bid would therefore be $b = 5v/7$. As depicted in Figure 1, the implied distribution of bidder 1's valuation is then $\sqrt{15v/14}$ for v between 0 and 14/15, which differs from F_1^I, the distribution that was inferred under the null before bidder 4 entered the auction. Because bidder 4's entry is assumed to be exogenous, the competitive IPV model cannot rationalize bidders 1 and 2's responses. Hence, the set of collusive bidders can be revealed by tests of independence between each of the incumbent bidders' competitively rationalizing valuations and the presence of bidder 4.

Alternatively, when bidders 1 and 2 are correctly assumed to be colluding, the ring wins with probability $9b^2/4$. Their optimal strategy is then $b = 2v/3$. Their valuation distribution must therefore be equal to $\sqrt{v}$ on $[0, 1]$, which is the same as the distributions F_1^{II} and F_2^{II} that were obtained before bidder D's arrival.

2.2 Related Collusion-Detection Methods

The identification strategy in this chapter is based on a comparison of a given bidder's behavior across auctions, as opposed to cross-bidder comparisons within auctions.[4] In view of this fact, Aryal and Gabrielli (2013), Price (2008), and Chassang and Ortner (2017) represent the most closely related collusion-detection procedures that have been proposed in the literature. Aryal and Gabrielli suggest testing for stochastic dominance between the valuations that rationalize a firm's bids under the null and alternative hypotheses. They

4. In relation to collusion in non-auction contexts, the identification strategy is in the spirit of Porter (2005), who suggests using firms' responses to changes in the market demand for their product to detect collusion.

argue this within-bidder test asymptotically controls the probability of Type I errors under their modeling assumptions, but such a test would not control size under the assumptions of section 2.3 because the valuations implied by collusion are always greater than the valuations implied by competition, even when the bidder is not actually colluding.[5]

Similarly, Price (2008) looks for evidence of collusion by comparing a firm's bids across auctions. Analyzing the same data that I do in my empirical application, he first uses theoretically motivated criteria to identify pairs of bidders who warrant closer inspection. Next, he regresses bids on a vector of auction covariates, a measure of the firm's capacity utilization, and the number of bidders at the auction. This regression is repeated using (i) the full sample of bids, (ii) the presumed competitive bids, and (iii) bids from suspected colluders when another suspected colluder is located within a certain geographic radius. The results of these regressions indicate suspected colluders bid less aggressively when another suspected colluder is nearby, which is consistent with the hypothesis that the bidders are in fact colluding. On the other hand, asymmetry in the bidders' private valuation distributions might also explain these results.

List et al. (2007) also find suggestive evidence of collusion in the same sample of British Columbia's timber auctions. In their framework, the problem of detecting collusion is a special case of the more general problem of estimating the agents' treatment status when treatment status is not directly observed. In that sense, the bids submitted by collusive firms are the "treated" observations, and the treatment reduces the colluders' bids relative to what they would have bid if they were not colluding. Though some of their findings are inconsistent with the competitive IPV model of bidding, they conclude that the evidence of collusion is mixed. They also suggest further research is needed to quantify the impact the suspected colluders might have on expected revenues.

5. In the example from the previous section, bidders 1 and 2's valuation distributions inferred under the assumption that they were colluding stochastically dominated those inferred under the assumption of competition. Thus, a test for stochastic dominance between F_i^I and F_i^{II} would reject the null hypothesis even if bidders 1 and 2 were not colluding.

Chassang and Ortner (2017) exploit variation in auction environments to detect collusion, but they take a different approach that is more closely related to the literature dating back to Porter (1983) that uses patterns consistent with periodic price wars to identify members of the cartel. Chassang and Ortner observe that, when the auctioneer imposes constraints on the permissible bidding strategies, a collusive bidding ring might not be able to punish defectors as effectively. When these constraints take the form of minimum prices in a procurement auction, an increase in the minimum price could actually cause the equilibrium price paid to decrease if the ring bidders are no longer able to sustain collusion and therefore revert to their competitive strategies. This surprising comparative static can then be interpreted as evidence of collusion.

My proposed tests of cross-auction restrictions are also similar to the cross-mechanism analysis in Athey et al. (2011). They observe that the prices in ascending auctions were lower than predicted given the valuation distributions estimated from a sample of first-price sealed-bid auctions. They then confirm this difference is statistically significant using a test of the null hypothesis that the average observed and predicted prices are equal. Assuming that the choice of auction mechanism is independent of the valuations, this test provides evidence against the null that all bidders are competitive.

The above identification strategies are conceptually distinct from the methods that attempt to detect collusion by testing the within-auction restrictions implied by the competitive model. Because they leverage different comparisons, the within-bidder tests for collusion are complements to, rather than substitutes for, tests of conditional independence among bids as in Porter and Zona (1999) and Bajari and Ye (2003). Relatedly, Porter and Zona (1993) test the restriction that, given a set of covariates, the conditional distribution of bidder i's bids is independent of the event that bidder i wins the auction.

Tests of independence are valid in any IPV model, but when bidders have symmetrically distributed valuations, the competitive model places further restrictions on the distribution of the data. In particular, the joint distribution of the bids must also be symmetric. By contrast, collusion among the bidders would generally create asymmetries. A test for asymmetry in bidding behavior will therefore identify the colluders. For example, Pesendorfer (2000) observes that bidders who collude efficiently operate as though they are a single, stronger bidder whose valuation is distributed as the maximum of the individual bidders' valuations. This induced asymmetry then causes the non-collusive firms to bid more aggressively, and in equilibrium, the non-collusive firms' bid distributions will stochastically dominate each of the colluders' bid distributions.[6] In addition, when the data include bidder-specific covariates, Bajari and Ye suggest a regression-based test of exchangeability in the bidders' strategy functions that may provide further evidence of collusion. Intuitively, this test builds on the insight that all of bidder i's competitors are exchangeable under the null hypothesis. But if bidder i is colluding with bidder j, bidder i's bid distribution will not depend on bidder j's characteristics in the same way that it depends on non-collusive bidders' characteristics.

Collusion detection methods have also been developed for other auction mechanisms, including second-price and ascending auctions (Baldwin et al., 1997; Brannman and Froeb, 2000; Marmer et al., 2016), first-price auctions with secret reserve prices and supplementary rounds of bidding (Kawai and Nakabayashi, 2014), and average-bid auctions (Conley and Decarolis, 2016). Of these methods, the nonparametric identification analysis in Marmer et al. (2016) is closest to the present dissertation. Under assumptions nearly identical to those in section 2.3, they prove the members of a collusive ring and their valuation distributions are nonparametrically identified in an English auction.

6. Colluders are identified if only one ring exists and bidders are ex-ante symmetric. If there is more than one ring, however, the test for stochastic dominance is no longer consistent as a test for collusion. For example, if two rival bidding rings compete in an auction, each consisting of two bidders whose valuations are distributed according to $\sqrt{v}$ for $v \in [0, 1]$, then all serious bidders bid one half of their valuations in equilibrium. If phantom bidders also bid one half of their valuations, a test for asymmetry will not reveal the colluders and their private valuations are not identified. Though, to ease exposition, I maintain the assumption that there is only one collusive bidding ring, I argue in section 2.4 that my identification strategy extends to situations in which multiple bidding rings compete against each other.

2.3 An Asymmetric Model of First-Price Auctions with Collusion

Let $\mathcal{N}$ denote the set of bidders that are eligible to bid for an object in a first-price, sealed-bid auction. Bidder i's private valuation of the object is a random variable V_i distributed according to the distribution function F_i. Assume that the valuations are independent so that the joint distribution of $V = (V_i : i \in \mathcal{N})$ is $F = F_1 \cdot \ldots \cdot F_n$. Further assume that each bidder i's valuation has compact support, $\mathcal{V}_i = [\underline{v}_i, \bar{v}_i]$, and that it has a density, f_i, which is bounded away from zero on $(\underline{v}_i, \bar{v}_i]$. A reserve price, r, may be modeled as an atom of the bidders' valuation distributions for the purposes of characterizing the equilibrium strategies.[7]

If the realization of a bidder's valuation is below the reserve price, then it will not participate in the auction.[8] I model this as a bid of negative infinity, or any other amount less than the reserve price. Otherwise, each bidder i observes its private valuation and, taking the other bidders' strategies as given, chooses b to maximize its expected payoff

$$(v_i - b) \cdot P\{i \text{ wins with a bid of } b\} =: (v_i - b) \cdot G_i(b), \tag{2.2}$$

where $G_i(b)$ is the probability that i wins with a bid of b. I refer to G_i as i's *competing distribution* because it is the distribution function for the highest bid among i's competitors. In equilibrium, G_i depends on the marginal distributions F_j and the strategies of bidders $j \neq i$, as well as the composition and behavior of the collusive ring.

Let $\mathcal{R} \subset \mathcal{N}$ denote the set of bidders in the collusive ring.[9] There are three cases to consider. The first is the competitive case, in which $\mathcal{R}$ is the empty set (or a singleton) and each bidder's competing distribution is the maximum bid among $j \neq i$. The second is when

7. Empirically, however, there will be an atom in bidder i's bid distribution if $\underline{v}_i = r$ and $F_i(r) > 0$, but not if $\underline{v}_i < r$.

8. This model of participation assumes that bidders do not incur any costs when they observe the realizations of their valuations or when they submit their bid.

9. The model easily extends to allow for multiple rings operating in the same auction, but I focus on the case of a single ring for the sake of clarity.

the ring includes all potential bidders. In this case, the model predicts the object will sell at the reserve price (McAfee and McMillan, 1992).

Finally, the third possibility is that $\mathcal{R}$ is a non-empty, non-singleton, proper subset of $\mathcal{N}$. In this case, I assume the ring colludes *efficiently* by nominating the member with the highest valuation to submit a "serious" bid that maximizes the expected profits of the ring. All other ring bidders may submit arbitrary "phantom" bids below the ring's serious bid. These phantom bids would never win but may be intended to create the illusion of competition.

The ring can sustain efficient collusion if it can condition side payments or future cooperation on the actions of its members. Specifically, there exists an incentive compatible mechanism in which the members of the ring report their valuations, bid according to the mechanism's recommendation, and execute the recommended transfers conditional on the other ring members' compliance. On the other hand, if the ring cannot effectively control its members' bidding, Marshall and Marx (2007) show that no such mechanism exists. Thus, efficient collusion is not an innocuous assumption.

Though efficiency is not assumed without loss of generality, it is necessary in order to infer the distribution of colluders' valuations. To demonstrate this necessity, suppose the ring cannot monitor its members' bids because the auction results remain anonymous. Then the serious bidder would sometimes have to bid more than it would have liked in order to deter other ring members from defecting. Consequently, the serious ring bidder would not maximize the expected profits in (2.2). Instead, its optimal bid would be a function of the highest- and second-highest order statistics among the ring's valuations. Because the phantom bidders may submit bids that are entirely unrelated to their valuations, the econometrician's task would be to map the ring's single serious bid to two valuations, making it impossible to nonparametrically identify the bidders' valuation distributions.

Other inefficient forms of collusion similarly impede inferences about the bidders' valuation distributions. For example, suppose the ring members take turns bidding at auctions regardless of their valuations. The non-ring bidders would want to condition their beliefs about their competing distributions on any variables that might help them predict who the serious ring bidder will be. If the econometrician cannot reconstruct their beliefs, their bids will not be correctly mapped to valuations. Therefore, none of the bidders' valuation distributions will be correctly identified.[10]

Though inefficient collusion interferes with identification of the valuation distributions, the identities of the colluders might be revealed under weaker assumptions (see, for example, Marshall and Marx, 2007).[11] This distinction highlights the two different roles that the modeling assumptions play in my identification argument. The first role allows me to detect collusion by providing testable restrictions that will necessarily be violated if a subset of bidders collude. The second role is to enable me to estimate bidding behavior in counterfactual scenarios by specifying the colluders' objective functions so that their inverse bidding strategies and valuation distributions can be recovered. Although colluders may be identified under alternative forms of collusion, I maintain the assumption of efficient collusion as a necessary step toward the ultimate goal of predicting auction prices in the absence of collusion.

To complete the description of the model, I also specify that collusion is common knowledge and that the competitive bidders know the ring maximizes its expected profits by bidding optimally against the non-ring bidders. However, I also note that common knowledge of collusion is perhaps a stronger assumption than is needed. Because bidders only need enough information to maximize their expected profits, the number and strength of

10. A simple bid-rotation scheme may not be as troublesome for identification as the inefficient form of collusion studied in Marshall and Marx (2007). Pesendorfer (2000) analyzed bids from auctions in which each colluding firm only bids in auctions in its designated region. Assuming non-ring firms were aware that the colluders would never seriously compete in another's region, one can recover the distribution of valuations using standard results from the empirical auction literature.

11. If the colluders only have to make side payments to the other ring members in the event that one of them wins, the ring's serious bidder does not have as great an incentive to win. A phantom bidder must therefore bid just below the serious bidder with sufficient probability to discourage it from submitting the lower individually optimal bid rather than the bid that was recommended by the ring. In light of this observation, Marshall and Marx suggest taking very similar bids as evidence of collusion.

its competitors is merely proximate to the distribution of its highest competing bid. A rational-expectations-like argument would justify the fact that bidders know their competing distributions without being specifically aware of any collusion. But because comparative statics play a major role in my identification strategy, I must also justify why bidders are able to respond optimally to changes in the auction environment. Unless it can be argued that all market participants have sufficient experience with auctions spanning the full domain of the covariates, this requires that bidders know who is colluding.

To summarize, the modeling assumptions are enumerated below:

Assumptions (Standard IPV Modeling Assumptions)**.**

MA.1 The set of potential bidders, $\mathcal{N}$, is common knowledge.

MA.2 Private valuations are independently distributed over their compact supports.

MA.3 The valuation densities are bounded away from zero on $(\underline{v}_i, \bar{v}_i]$.

MA.4 A proper (possibly empty) subset of potential bidders, $\mathcal{R} \subset \mathcal{N}$ collude prior to bidding.

MA.5 Each bidder i knows G_i, the distribution of the highest bid among i's competitors.

MA.6 All bidders are risk neutral and bid to maximize their expected profits in (2.2).

In addition, the collusive ring behaves in accordance with $MA.7$

Assumptions (Modeling Assumptions Regarding Collusion)**.**

MA.7 The bidder with the highest valuation among $\mathcal{R}$ bids to maximize the expected profits in (2.2). *All other colluders either abstain from bidding or submit a phantom bid less than the serious colluder's bid.*

Under assumptions MA.1–6, Lebrun (2006) proves that, in a competitive auction, the unique Bayes-Nash equilibrium inverse bidding strategies are differentiable and strictly increasing in the bids whenever they are greater than the minimum bid. Moreover, the support of the equilibrium bid distributions share a common lower bound, though they may have different upper bounds. Nonetheless, the inverse bidding strategies can be continuously extended over the entire compact support of the bids, $[\underline{b}, \bar{b}]$.

Lebrun's result extends easily when some bidders collude efficiently. The competitive bidders' equilibrium strategies will be identical to the ones they would use in an auction in which the colluders were replaced by a single bidder whose valuation is distributed like the maximum of the ring's valuations. Thus, the following proposition characterizes the equilibrium bidding strategies in an auction with collusion.

Proposition 1. *Under assumptions MA.1–MA.6, a non-collusive bidder's equilibrium bidding strategy,* $\sigma_i(v_i)$*,*

(i) is unique,

(ii) is less than or equal to v_i *and is strictly increasing on* $[\underline{v}_i, \bar{v}_i]$ *except possibly for at most one non-collusive bidder,*

(iii) is equal to the minimum bid at the minimum bid, i.e. $\sigma_i(\underline{b}) = \underline{b}$*,*

(iv) is differentiable whenever $\sigma_i(v_i)$ *is strictly greater than the minimum bid, and*

(v) has an inverse that can be continuously extended over the compact support $[\underline{b}, \bar{b}]$*, which is given by the bidder's first-order condition*

$$\sigma_i^{-1}(b_i) = b_i + \frac{G_i(b_i)}{g_i(b_i)}\,, \tag{2.3}$$

where g_i *is the density of i's competing distribution.*

Under the additional assumption MA.7, each collusive bidder's equilibrium strategy, σ_i*,*

(i) satisfies the above properties whenever $v_i > \max_{j \neq i \in \mathcal{R}} v_j$*, hence is the same for all* $i \in \mathcal{R}$*,*

(ii) is otherwise a possibly random function of all of the colluders' valuations, and

(iii) is less than the ring's serious bid whenever $v_i < \max_{j \in \mathcal{R}} v_j$.

In short, the serious equilibrium bids are increasing functions of the valuations and they start at a common minimum bid. The only irregularity is that one of the non-collusive bidders' strategies or all of the colluders' strategies could be constant near the minimum bid. This situation would arise when a binding reserve price causes bidder i's competitors to abstain from bidding with sufficiently high probability and the densities of the competitors' valuations is sufficiently small in a neighborhood of the reserve price. Then, for valuations just above the reserve price, the cost of submitting a higher bid outweighs bidder i's slight benefit of increasing its probability of winning. Thus, bidder i would optimally submit a bid equal to the reserve price even when its valuation is strictly greater than the reserve.

The fact that one of the strategy functions might not be strictly increasing poses a challenge to identification of the valuation distribution at the minimum bid because $F_i(\underline{b}) = S_i(\underline{b})$ if and only if σ_i is strictly increasing. It may be tempting to make further identifying assumptions in order to rule out this case. For example, assuming bidders are type-symmetric, i.e. each bidder of the same type has its valuation drawn from the same distribution, bidding strategies will all be strictly increasing if there are at least two bidders of every type. As an alternative, however, the tail behavior of the bid distributions could be used to test whether $F_i(\underline{b})$ is identified.[12] Specifically, if a bidder's strategy function is constant over some interval, there will be an atom in its bid distribution at $\underline{b}$ and the density of its competing bid will be bounded near the minimum bid. Otherwise, each bidder's bid distribution will have an unbounded density. See Proposition 2 in the appendix for details.

12. The tail-behavior of the bid distribution has also been used to test for common values in symmetric first-price auctions (Hill and Shneyerov, 2013).

2.4 Identification

2.4.1 Identification up to Truncation by the Minimum Bid

I define a private value model to be a collection $\mathcal{M}$ of pairs (F, σ), where F is a joint distribution of valuations and covariates, (V, X), and $\sigma\colon (V, X) \mapsto B$ is a profile of bidding strategies. In the competitive IPV model, $\mathcal{F}$ is restricted to the collection of absolutely continuous valuations distributions that are independent conditional on X, and σ is restricted to the unique strategy profile that satisfies the first-order conditions (2.3) where each bidder i is competing against all other bidders—that is, $G_i(b|x) = \prod_{j \neq i} F_j(\sigma_j^{-1}(b; x)|x)$. When the model is enlarged to include strategy profiles that satisfy the assumptions of the previous section, I refer to $\mathcal{M}$ as an IPV auction model with collusion.

Given the counterfactual questions of interest, the goal is to recover the $F \in \mathcal{F}$ that generated the observed distribution of the data. In nonparametric models, however, the data will never provide information about $F(v)$ at valuations below the minimum bid. Thus, I will say that F is identified (up to the truncation induced by the minimum bid) whenever each $F(v)$ is uniquely determined for v with $v_i > \sigma_i^{-1}(\underline{b})$ for all i. Letting (B, X) denote the random vector of bids and covariates, I formalize this definition as follows.

Definition 1. A model, $\mathcal{M}$ is *identified* from the joint distribution of bids and covariates up to the truncation at $\underline{b}$ if, whenever $(\sigma(V, X), X)$ and $(\sigma'(V', X'), X')$ are equal in distribution for some (V, X) and (V', X') distributed according to F and F' with $(F, \sigma), (F', \sigma') \in \mathcal{M}$, $F(v, x) = F'(v, x)$ for all x and all $v > \sigma_i^{-1}(\underline{b})$.

More generally, a function, $Y(B, X)$, of the bids and covariates may be observed. For instance, the transaction price is sometimes the only observable bid. In this case, I will say the model is identified from Y if the above definition holds when $(\sigma(V, X), X)$ and $(\sigma'(V', X'), X')$ are replaced by $Y(\sigma(V, X), X)$ and $Y(\sigma'(V', X'), X')$.[13]

13. Definition 1 does not require the phantom bidding strategies to be identified since the strategy profile that rationalizes the data need not be unique. The serious bidding strategies are identified, however, because

2.4.2 Identification from Prices through Exogenous Variation in Competition

The IPV model is not identified from bids alone unless the identities of colluders are known *a priori*. But, as the example in section 2.1 suggests, an instrument that induces variation in the level of competition—e.g. an indicator for the presence of bidder 4—can be used to test for collusion. Specifically, under the null hypothesis that the bidder is not colluding, the competitive model correctly infers its valuations regardless of any collusion among the other bidders. The valuations that rationalize the non-colluder's bids are therefore independent of the instrument. In contrast, a colluder's competitively rationalizing valuation distribution is always stochastically smaller than its true valuation distribution; and it is even more so when the colluder bids less aggressively in a sense that will be made precise in the proof of Theorem 2. Because the competitively rationalizing distribution shifts relative to the true distribution, which is constant with respect to the exogenous instrument, its competitively rationalizing valuations will depend on the instrument. Thus, colluders can be detected by tests of independence between the instrument and the competitively rationalizing valuations.

Once all the ring members have been identified in this way, each of the bidders' true competing distributions can be computed. Their valuation distributions can then be inferred from the first-order condition of the bidders' profit maximization problems. Hence, the IPV model with collusion is identified from the distribution of winning bids and exogenous variation in the level of competition.

The proof of Theorem 2 formalizes this argument. Contrary to what the example in section 2.1 might indicate, however, I do not prove that the competitively rationalizing distributions are stochastically ordered by the instrument. Instead, I argue that they are merely unequal in some neighborhood of a particular valuation and, hence, are not independent of the instrument.

they are uniquely determined by F, the identities of the colluders, and the bidders' first-order conditions. This result will be established in the remainder of this section.

To prove this main result, I will use the following lemma.

Lemma 1. *Assume $MA.1$–7.*

(i) Each bidder's marginal serious bid distribution is identified up to the truncation induced by the minimum bid from the prices and the identities of the winners.

(ii) A bidder's competing distribution and valuation distribution are identified under the null hypothesis that it is not colluding with anyone else at the auction.

Proof. I suppress the dependence on the covariates, X, throughout the folllowing proof. Let $M_i(b) = P\big\{B_i \leq b, \max_{j \neq i} B_j \leq B_i\big\}$ denote the probability of the event that bidder i wins the auction and the price is less than or equal to b.[14] If i is a member of the bidding ring, this event is equivalent to the event

$$\left\{\sigma_i(V_i) \leq b, \max_{j \notin \mathcal{R}} \sigma_j(V_j) \leq \sigma_i(V_i), \max_{k \neq i \in \mathcal{R}} v_k \leq v_i\right\},$$

where, by an abuse of notation, I suppress the dependence of σ_i on v_k for $k \neq i \in \mathcal{R}$ and use $\sigma_i(v_i)$ to denote the would-be serious bid that a collusive bidder would have submitted if the other colluders' valuations had been less than v_i. Otherwise, when i is not in the ring, it is equivalent to the event

$$\left\{\sigma_i(V_i) \leq b, \max_{j \neq i \in \mathcal{R}} \sigma_j(V_j) \leq \sigma_i(V_i), \sigma_i(\max_{k \in \mathcal{R}} V_k) \leq \sigma_i(V_i)\right\}.$$

In either case, because each member of the ring uses the same increasing bidding strategy, both events are equivalent to

$$\left\{\sigma_i(V_i) \leq b, \max_{j \neq i} \sigma_j(V_j) \leq \sigma_i(V_i)\right\}.$$

14. The probability that there is a tie in the winning bid is zero, except possibly at a price equal to the minimum bid in the case where the minimum bid is an atom of the bidders' valuations distributions. In that event, ties are broken randomly.

Under the assumption that the valuations are mutually independent, the functions M_i for $i \in \mathcal{N}$ can then be used to construct S_i, the marginal distribution of $\sigma_i(V_i)$. In words, S_i is the distribution of the bid that bidder i would submit if it were trying to win the auction. For each $b > \underline{b}$, $S_i(b)$ is given by

$$S_i(b) = \exp\left\{ -\int_b^{\bar{b}} \frac{1}{M} \, dM_i \right\} , \tag{2.4}$$

where $M = \sum_i M_i$ is the distribution of auction prices. An analogous construction was proven by Berman (1963) and was introduced to the empirical auction literature by Athey and Haile (2002). This proves (i).

Under the null hypothesis that i is not colluding, the distribution of the highest bid among i's competitors is

$$G_i = \prod_{j \neq i} S_j \, , \tag{2.5}$$

Plugging (2.4) into (2.5), the inverse strategy can be written as

$$\sigma_i^{-1}(b) = b + \frac{M(b)}{\frac{\partial M_{-i}(b)}{\partial b}} \, ,$$

where $M_{-i} = \sum_{j \neq i} M_j$. By Proposition 1, σ_i^{-1} is increasing in b for $b \geq \underline{b}$. Therefore, identification of F_i on $[\sigma_i^{-1}(\underline{b}), \bar{v}_i]$ follows from

$$F_i(\sigma^{-1}(b)) = S_i(b) \, . \tag{2.6}$$

This proves (ii). $\square$

In order to correctly infer the bidders' valuations under the null, the key step was to use (2.5) to construct i's competing distribution. But, if bidder i might be colluding with an unknown subset of the other bidders, there are generally $2^{n-1} - 1$ different competing distributions that can be constructed from the other bidders' bid distributions. Each one of these corresponds to a different M_{-i}, and hence, a different valuation distribution. Accordingly, the model is not identified from the winning bids.

In fact, the example in section 2.1 already proved a stronger result. Namely, the model is still not identified when the full vector of bids is observed. For completeness, I formally state this as a theorem.

Theorem 1. *The IPV model with collusion is not identified from the full vector of bids.*

Proof. See the example from section 2.1. □

Interestingly, however, Theorem 2 says that just one binary instrument would be sufficient to identify the true composition of the bidding ring from the $2^n - n - 1$ possible combinations of colluders. Formally, the identifying assumptions are as follows.

Assumptions (Identification Assumptions)**.**

IA.1 The supports of the bidders' valuations have a non-negligible intersection, i.e. $\bigcap_{i \in \mathcal{N}} [\underline{v}_i, \bar{v}_i]$ *has a positive Lebesgue measure.*

IA.2 For each i*, bidder* i*'s competing distribution depends non-trivially on* Z_i*.*

IA.3 $Z = (Z_1, \ldots, Z_n)$ *is observed and, for all* $i \in \mathcal{N}$*, there are covariates,* X_i*, such that* $Z_i \perp\!\!\!\perp V_i \,|\, X_i$ *and* Z_i *is a nondegenerate random variable conditional on* X_i*.*

Assumption IA.1 ensures that each bidder wins with positive probability.[15] Assumption IA.2 guarantees that the instrument is relevant to i's optimal bidding strategy. Assumption IA.3 asserts that the econometrician observes exogenous variation in an instrument for each bidder i. This assertion may be satisfied in one of two ways: either the instruments are independent of the bidders' valuations unconditionally, or there is variation in the instrument for bidder

15. If any two bidders' supports do not overlap, the two could collude and the bidder with the weaker valuations might never appear in the data. In which case, the collusive auction would be observationally equivalent to a competitive auction in which the weaker bidder's valuations are less than the reserve price with probability one. Thank you to Alexander Wolitzky for suggesting this case.

i that is independent of i's valuation conditional on a subset of the observed covariates. For instance, if Z_i is equal to the exogenously varying, binding reserve price for all i, then IA.3 is satisfied. On the other hand, if Z_i is a vector of bidder-specific covariates for i's competitors—e.g. distances of i's competitors to the work site—then Z_i will not be independent of V_i. But it is possible to condition on i's distance to the work site and use the residual variation in Z_i to shift i's competing distribution without affecting i's own valuation distribution. In either case, the following theorem holds.

Theorem 2. *Under the modeling and identification assumptions, the IPV model with collusion is identified from the distribution of prices and covariates.*

Proof. By part (i) of the lemma, the marginal distribution of bidder i's would-be serious bid is identified for each bidder that wins with positive probability, which includes all $i \in \mathcal{N}$ by IA.1. By part (ii) of the lemma, a competitive bidder i's competitively rationalizing distribution will be equal to its true distribution conditional on any value of the covariates. The competitively rationalizing valuations are therefore independent of the instrument. To establish identification of the collusive model, I prove that the converse statement is also true: a colluder's competitively rationalizing valuation distribution must depend on Z_i. Thus, a test for independence between the competitively rationalizing valuations and the instrument can be used to detect collusion. Once all of the colluders have been identified in this way, each of the bidders' true competing distributions can be computed as $\prod_{j \neq i} S_j$ for $i \notin \mathcal{R}$ and $\prod_{j \notin \mathcal{R}} S_j$ for $i \in \mathcal{R}$. The true valuation distributions are then given by (2.6).

The rest of the argument mirrors Figure 1. First, I derive an expression for the horizontal distance between the competitively rationalizing distribution (dashed curves) and the true distribution (dotted curve) in Figure 1. In keeping with part (ii) of the lemma, this horizontal difference is zero for all values of the instrument if the bidder is not colluding. Otherwise, the collusive bidders' true valuation distribution will stochastically dominate the competitively rationalizing distribution.

To see this, I let $\hat{F}_i$, $\hat{G}_i$ and $\hat{g}_i$ denote the valuation distribution and competing distribution and density that are inferred under the null hypothesis that i is not colluding. I also suppress the argument and the conditioning on (Z, X) and let $\mathbf{1}$ denote the identity mapping $b \mapsto b$.

$$
\begin{aligned}
&F_i^{-1} - \hat{F}_i^{-1} \\
&= \left(\mathbf{1} + \frac{G_i}{g_i}\right) \circ S_i^{-1} - \left(\mathbf{1} + \frac{\hat{G}_i}{\hat{g}_i}\right) \circ S_i^{-1} \\
&= S_i^{-1} + \frac{G_i(S_i^{-1})}{g_i(S_i^{-1})} - \left(S_i^{-1} + \frac{\hat{G}_i(S_i^{-1})}{\hat{g}_i(S_i^{-1})}\right) \\
&= \frac{1}{\frac{g_i(S_i^{-1})}{G_i(S_i^{-1})}} - \frac{1}{\frac{g_i(S_i^{-1})}{G_i(S_i^{-1})} + \sum_{j \neq i \in \mathcal{R}} \frac{f_j(b + \frac{G_i}{g_i})}{F_j(b + \frac{G_i}{g_i})}\Big|_{S_i^{-1}} \cdot \frac{\partial}{\partial b}\left(b + \frac{G_i(b)}{g_i(b)}\right)\Big|_{S_i^{-1}}} \\
&= \frac{1}{1 \Big/ \left(F_i^{-1} - \left(\mathbf{1} + \frac{G_i}{g_i}\right)^{-1}\Big|_{F_i^{-1}}\right)} \\
&\quad - \frac{1}{1 \Big/ \left(F_i^{-1} - \left(\mathbf{1} + \frac{G_i}{g_i}\right)^{-1}\Big|_{F_i^{-1}}\right) + \sum_{j \neq i \in \mathcal{R}} \frac{f_j(F_i^{-1})}{F_j(F_i^{-1})} \Big/ \frac{\partial}{\partial v}\left(\left(\mathbf{1} + \frac{G_i}{g_i}\right)^{-1}(v)\right)\Big|_{F_i^{-1}}}.
\end{aligned}
$$

The first equality holds by the GPV equation. The S_i^{-1} cancel in the second line. In the third line, I decompose $\hat{g}_i/\hat{G}_i$ into the reverse hazard rate of i's true competing distribution and the reverse hazard rates of the other colluders' bids. I also use the fact that the ring members would use the same inverse strategy function $(\mathbf{1} + G_i/g_i)$ if they were submitting a serious bid. The last equality again holds by the GPV equation (2.6).

This final expression only depends on the true valuation distribution and equilibrium strategy function. And, because $\sigma_i = (\mathbf{1} + \frac{G_i}{g_i})^{-1} < \mathbf{1}$ and f_j/F_j is greater than zero, it implies that $F^{-1} - \hat{F}^{-1}$ is positive. Moreover, this difference is decreasing in both the slope and the level of σ_i evaluated at the quantile of i's true valuation distribution. That is, the competitive model more severely underestimates the valuation distribution when the colluders bid less aggressively in the sense that the slope and level of their strategy function is lower.

To finish the proof, I will establish that the changes in the slope and level of a colluder's strategy function cannot exactly offset everywhere. Relative to Figure 1, this argument shows that the competitively rationalizing distribution (dashed line) must move relative to the true valuation distribution (dotted line) in response to variation in the instrument. Because IA.3 states that the dotted line is constant with respect to the instrument, the dashed lines cannot coincide for all values of Z.

That is, I argue that there exists a valuation, v, and realizations of the instrument, z and z', such that the horizontal difference between the distribution functions is greater at v, i.e.

$$F_i^{-1}(F_i(v)) - \hat{F}_i^{-1}(F_i(v) \,|\, Z_i = z) \neq F_i^{-1}(F(v)) - \hat{F}_i(F_i(v) \,|\, Z_i = z') \,,$$

which implies that $\hat{F}_i(\cdot \,|\, Z_i = z) \neq \hat{F}_i(\cdot \,|\, Z_i = z')$. The existence of such a v is easy to verify near the minimum bid. To that end, I consider the following cases.

Case i. The instrument does not affect the minimum bid. By IA.2, bidder i's strategy function is a nontrivial function of its instrument. And by Proposition 1, $\sigma_i(\underline{b}; z) = \underline{b}$ for all z. Therefore, there is a "first" point, v_0, where strategies $\sigma_i(\cdot\,; z)$ and $\sigma_i(\cdot\,; z')$ diverge.[16] Then there is a point v_1 in the neighborhood of v_0 such that the slope and level of one of the strategies is greater than the other for, say, $Z = z'$ than $Z = z$. Therefore, $\hat{F}_i^{-1}(F_i(v_1)|Z = z') < \hat{F}_i^{-1}(F_i(v_1)|Z = z)$.

Case ii. The instrument affects the minimum bid. Let $\underline{b}$ and $\underline{b}'$ denote the minimum bids for $Z = z$ and z'. Assume without loss of generality that $\underline{b} < \underline{b}'$. Then $\hat{F}_i(\underline{b}'|Z = z) > F_i(\underline{b}')$ because the competitive model strictly underestimates (in the FOSD-sense) the true valuation distribution. The only exception is at the minimum bid, where $\hat{F}_i(\underline{b}'|Z = z') = F_i(\underline{b}')$ whenever i's strategy is strictly increasing at $\underline{b}'$. This implies that $\hat{F}_i(\underline{b}'|Z = z') < \hat{F}_i(\underline{b}'|Z = z)$.

16. More precisely, let v_0 denote the infimum of the set of v for which $\sigma_i(v\,; z) = \sigma_i(v\,; z')$.

If i's strategy is not strictly increasing at $v = \underline{b}$, then Proposition 1 implies that i will submit a bid exactly equal to $\underline{b}'$ with positive probability. If i is not colluding, then part (i) of Proposition 1 implies that bidder i must be the only bidder with an atom in its bid distribution. In contrast, if i is colluding then all ring members will optimally use the same strategy function, and multiple bidders' bid distributions will have an atom at $\underline{b}'$. In this case, the colluders would be identified even without exploiting variation in the level of competition at the auction. $\square$

Remark 1. Let $V_{i,\mathcal{R}}$ denote the random valuation that rationalizes i's bids under the assumption that the bidders in $\mathcal{R}$ are colluding. Because a competitive bidder i's inferred valuations are uncontaminated by any collusion among the other bidders, $V_{i,\emptyset} = V_{i,\mathcal{R}}$ for each bidder i who is not a member of the alleged ring, $\mathcal{R}$. As a corollary to Theorem 2, one can show that if $V_{i,\mathcal{R}}$ is independent of Z_i (possibly conditional on X_i) for all $i \in \mathcal{N}$, then $\mathcal{R}_0 \subseteq \mathcal{R}$. Consequently, $\mathcal{R}_0$ is the minimal subset of $\mathcal{N}$ that satisfies the independence condition

$$V_{i,\mathcal{R}} \perp\!\!\!\perp Z_i \,|\, X_i \text{ for all } i\,. \tag{2.7}$$

An argument similar to the proof of Theorem 2 can strengthen this statement by establishing the opposite set inclusion relation for a non-singleton $\mathcal{R}$, i.e. $\mathcal{R} \subseteq \mathcal{R}_0$. Namely, if the true ring were a proper subset of the alleged ring, each alleged ring member's implied valuations, $V_{i,\mathcal{R}}$, would be greater than its true valuation $V_{i,\mathcal{R}_0}$. Furthermore, the difference between the true and implied quantiles will be larger the larger is the error in bidder i's implied competing distribution relative to its true competing distribution. If an instrument induces variation in this relative error, $\mathcal{R} = \mathcal{R}_0$ will be the unique non-singleton $\mathcal{R} \subset \mathcal{N}$ satisfying the independence condition (2.7).

The proof of Theorem 2 demonstrates, however, that $V_{i,\emptyset} \perp\!\!\!\perp Z_i \,|\, X_i$ if and only if $i \notin \mathcal{R}$. Thus, even though there is a unique configuration of the ring that rationalizes the data under the modeling and identification assumptions, it is not necessary to search over all possible configurations to find the one subset that satisfies (2.7). Instead, it is sufficient to test for dependence between $V_{i,\emptyset}$ and Z_i for each i.

Remark 2. A version of Theorem 2 extends to auctions in which multiple rings are operating. In this case, the above argument can be used to show that all bidders who are colluding with someone else in $\mathcal{N}$ will be identified. Assuming the members of the bidding rings do not overlap, the next step would be to argue that there is a unique partition of these collusive bidders into rival rings that rationalizes the data. Intuitively, this partition is unique because each partition generally implies a different competing distribution, and hence different predictions about the bidders' responses to the instruments.

An exception is when some of the colluders' valuations are identically distributed. For instance, suppose there are two strong bidders and two weak bidders, and each of the strong bidders is colluding with one of the weak bidders.[17] The data will reject the comparative statics implied by the competitive model but might not reveal which strong bidder is colluding with which weak bidder. Even in this exceptional case, however, the configurations of the rings are identified up to permutations of the identical bidders. Moreover, the competing distributions, and hence the distribution of each bidders' private valuations, will be identified because they are invariant under these permutations.

In principle, given sufficient variation in $\mathcal{N}$, one could identify the exact composition of multiple bidding rings even when bidders' valuations are exchangeable.

2.5 Empirical Framework and Definition of the Estimators

2.5.1 Empirical Framework

Let t index auctions in which an object with characteristics X_t is for sale. In each auction the bidders in $\mathcal{N}_t$ are eligible to bid, but a bid will not be recorded for those whose payoff

17. The same argument holds when the bidders' valuations are all drawn from the same marginal distribution. I assume that there are two types of bidders to illustrate that this problem arises under milder assumptions, as well.

from winning the object, u_{it}, is less than the reserve price, $\tilde{r}_t$. For notational convenience, I suppose that one or both of $\mathcal{N}_t$ and $\tilde{r}_t$ serve as the instrument. Therefore, for each auction $t = 1, \dots, T$, the data consist of the left-censored bids, $\left(\tilde{b}_{it} \vee \tilde{r}_t, \mathbb{1}\left\{\tilde{b}_{it} \geq \tilde{r}_t\right\}\right)$ for $i \in \mathcal{N}_t$, and the auction-level covariates, $(X_t, \mathcal{N}_t, \tilde{r}_t)$. In the asymptotic analysis, I will assume that T tends toward infinity while the support of $\mathcal{N}_t$ is finite so that the set of bidders, $\bigcup_{t=1}^{\infty} \mathcal{N}_t$, is finite and each bidder is observed to participate in many auctions.

If the objects for sale are adequately described by a small number of characteristics, then nonparametrically estimating the distribution of valuations conditional on X_t may be practical. More commonly, however, some assumptions are needed to reduce the dimensionality of the objects' heterogeneity. To this end, I assume that each bidder's utility is additively separable in the object's observable characteristics and an unobservable idiosyncratic private component, V_{it}, which is independently distributed across bidders and independently and identically distributed across auctions:

$$U_{it} = V_{it} + \mu(X_t) \text{ for each } i, \text{ and } V_t = (V_{it})_{i \in \mathcal{N}_t} \sim \prod_{i \in \mathcal{N}_t} F_i \,.$$

One can show that additive separability in the valuations implies additive separability in the equilibrium bidding strategies and the distribution of V_{it} is identified up to location (Athey and Haile, 2002). It is therefore helpful select a benchmark vector of covariates, x_0, and define the homogenized bids, $B_{it} = \tilde{B}_{it} - \mu(X_t) + \mu(x_0)$, and homogenized reserve price, $r_t = \tilde{r}_t - \mu(X_t) + \mu(x_0)$, relative to this benchmark. In words, B_{it} is the bid that would have been observed if $\tilde{r}_t$ had been r_t and the auction-level covariates had been x_0 instead of x_t. If $\mu(x_0)$ is normalized to zero for a particular choice of x_0, then the homogenized equilibrium bidding strategies satisfy

$$B_{it} = V_{it} - \frac{G_i(B_{it}|r_t, \mathcal{N}_t)}{g_i(B_{it}|r_t, \mathcal{N}_t)} \,.$$

In practice, the function μ is typically assumed to be linear in the object's characteristics. Alternatively, data from procurement auctions often include an engineer's estimate which may be considered an estimate of $\mu(X_t)$ or a direct observation of $\mu(X_t)$, itself. Regardless, the relatively fast rate of convergence for estimators of $\mu(X_t)$ asymptotically justifies using the homogenized bids as though they were data for the purposes of the nonparametric estimators defined below. In the next sections, I work exclusively with homogenized bids and reserve prices and refer to them simply as bids and reserve prices.

2.5.2 Definition of the Estimators

The estimators that I define in this section are sample analogues to their population counterparts. In each case, I condition on the set of eligible bidders and use a continuous second-order kernel, $K_h(u) = \frac{1}{h}K(\frac{u}{h})$, with bandwidth h to smooth over the reserve prices. For instance, the sample analogue to $M_i(\cdot, r, \mathcal{N})$ is defined by

$$\mathbb{M}_{iT}(b|r,\mathcal{N}) = \frac{\sum_t \mathbb{1}\{b_{it} \geq p_t\} \cdot \mathbb{1}\{B_{it} \leq b\} \cdot \mathbb{1}\{\mathcal{N}_t = \mathcal{N}\} \cdot K_h(r - r_t)}{\sum_t \mathbb{1}\{\mathcal{N}_t = \mathcal{N}\} \cdot K_h(r - r_t)},$$

where $p_t = \max\{r_t, b_{it} : i \in \mathcal{N}_t\}$ is the price at auction t.[18] To simplify notation, however, I will suppose that $\mathcal{N}_t$ is fixed and suppress its notation as in

$$\mathbb{M}_{iT}(b|r) = \frac{\sum_t \mathbb{1}\{b_{it} \geq p_t\} \cdot \mathbb{1}\{B_{it} \leq b\} \cdot K_h(r - r_t)}{\sum_t K_h(r - r_t)}. \tag{2.8}$$

In addition, let $\mathbb{M}_T(b|r) = \sum_i \mathbb{M}_{iT}(b|r)$ be the estimator for the distribution of the sale price conditional on the reserve. Define $\mathbb{M}_{-iT}(b|r) = \sum_{j \neq i} \mathbb{M}_{jT}(b|r)$.

By analogy to equation (2.4), I then use the $\mathbb{M}_{iT}$ estimators to construct an estimator

18. If none of the bidders' valuations exceed the reserve price, I record this as a price equal to the reserve and the "winner" is the seller.

for the marginal bid distributions,

$$\begin{aligned}\mathbb{S}_{iT}(b|r) &= \exp\left\{-\int_b^{\bar{b}} \frac{1}{\mathbb{M}_T(\cdot|r)}\, d\mathbb{M}_{iT}(\cdot|r)\right\} \\ &= \exp\left\{-\frac{\sum_t \mathbb{M}_T(p_t|r)^{-1}\cdot \mathbb{1}\{b_{it} \geq p_t\}\cdot \mathbb{1}\{p_t > b\}\cdot K_h(r-r_t)}{\sum_t K_h(r-r_t)}\right\}. \qquad (2.9)\end{aligned}$$

Consistent with part (i) of Lemma 1, this estimator does not depend on whether i is assumed to be competitive or collusive.

To estimate $\frac{G_i}{g_i}$ under the null, I first observe that $G_i/g_i = 1/\frac{\partial \log G_i}{\partial b}$. Plugging in the expression for G_i in terms of $\{M_i : i \in \mathcal{N}\}$, this can be written as

$$\begin{aligned}\frac{G_i(b|r)}{g_i(b|r)} = \left[\frac{\partial \log G(b|r)}{\partial b}\right]^{-1} &= \left[-\frac{\partial\left(\int_b^{\bar{b}} \frac{1}{M(p|r)} dM_{-i}(p|r)\right)}{\partial b}\right]^{-1} \\ &= \frac{M(b|r)}{\frac{\partial M_{-i}(b|r)}{\partial b}}.\end{aligned}$$

I estimate this expression by substituting $\mathbb{M}_T(b|r)$ in the numerator and a kernel estimator for derivative of M_{-i} with respect to b. Though, if the reserve price is binding and bidding strategies are all strictly increasing, this density will be unbounded near the reserve price for all i (see Proposition 2 in the appendix). Consequently, the typical kernel density estimator will not be consistent. Instead, I use a boundary corrected kernel density estimator and a change of variables to achieve a consistent estimator for values of b within one bandwidth of the reserve price. Specifically, I first use the change of variables suggested by Guerre et al. (2000) to create the transformed data, $(\max_{j\neq i} \sqrt{b_{jt} - r_t}, r_t)$, whose population density can be shown to be bounded everywhere. I then apply the boundary correction procedure in Karunamuni and Zhang (2008) on these data to obtain an estimate of their density, $\tilde{\mathbb{g}}_{iT}(b, r)$.[19] An estimate of the density of the original data is given

19. Karunamuni and Zhang (2008) uses a combination of transformation and reflection of the data to reduce the order of the bias in kernel density estimation near the boundary. Hickman and Hubbard (2015) first demonstrate its use in application to inference in first-price auctions.

by $\mathrm{g}_{iT}(b,r) = \tilde{g}_i(\sqrt{b-r}, r)/(2\sqrt{b-r})$. Finally, the conditional density is consistently estimated by the ratio of $\mathrm{g}_{iT}(b,r)$ to the marginal density of the reserve prices. Thus, the estimator for $\frac{G_i(b|r)}{g_i(b|r)}$ is given by

$$\frac{\mathbb{G}_{iT,\emptyset}(b|r)}{\mathrm{g}_{iT,\emptyset}(b|r)} = \mathbb{M}_T(b|r)\frac{\mathrm{g}_T(r)}{\mathrm{g}_{iT,\emptyset}(b,r)}\,,$$

where the subscript $\emptyset$ denotes that these quantities are estimated under the hypothesis that the set of colluders is empty.

Bidder i's private valuations can then be consistently estimated under the null by

$$\hat{v}_{it,\emptyset} = b_{it} + \frac{\mathbb{G}_{iT,\emptyset}(b_{it}|r_t)}{\mathrm{g}_{iT,\emptyset}(b_{it}|r_t)}\,, \tag{2.10}$$

and an estimate of its valuation distribution is given by the sample analog to the GPV equation (2.6),

$$\mathbb{F}_{iT,\emptyset}\left(b + \frac{\mathbb{G}_{iT,\emptyset}(b|r)}{\mathrm{g}_{iT,\emptyset}(b|r)}\middle| r\right) = \mathbb{S}_{iT}(b|r)\,.$$

To estimate G_i and F_i under the alternative hypothesis that bidders in $\mathcal{R}$ are colluding, I replace $\mathbb{M}_{-iT}(b|r)$ for each $i \in \mathcal{R}$ with

$$\mathbb{M}_{-\mathcal{R}T}(b|r) = \frac{\sum_t \mathbb{1}\{\max_{j\notin\mathcal{R}} b_{jt} \geq p_t\} \cdot \mathbb{1}\{\max_{j\notin\mathcal{R}} B_{jt} \leq b\} \cdot K_h(r-r_t)}{\sum_t K_h(r-r_t)}$$

and replace the estimator for the conditional density of the highest bid among bidders $j \neq i$ with an estimator of the conditional density of the highest bid among $j \notin \mathcal{R}$. Otherwise, for each competitive bidder $i \notin \mathcal{R}$, the estimators are unaffected. Thus, as in the population equations, the estimators for a competitive bidder's equilibrium bidding strategy and valuation distribution do not depend on whether other bidders are assumed to be colluding.

2.5.3 Uniform Convergence of the Estimators

The random functions $\mathbb{M}_{iT}(\cdot|r, \mathcal{N})$ serve as the building blocks for the estimators defined above. Because these kernel-based estimators' asymptotic behavior is well understood, the estimates of F_i, G_i, and S_i will yield similarly well behaved asymptotics if the transformation ϕ defined by the right-hand side of equation (2.4) is Hadamard differentiable so that the functional delta method applies. Fortunately, when the reserve price is binding, differentiability of ϕ as a map into the space of càdlàg functions follows almost immediately from Lemma 20.10 in van der Vaart (1998). Otherwise, when the reserve price does not bind, some trimming near the minimum observed bid is required in order to bound the integrand in equation (2.4).[20] The standard uniform rates of convergence for kernel-based estimators will therefore carry through. In particular, the above estimators will uniformly converge at the optimal rates derived by Guerre et al. (2000).

To be precise, I assume that the reserve price is continuously distributed, and $M_i(b, r)$ has a nonzero, twice continuously differentiable density on its compact support.[21] Under these regularity conditions, and given a sequence of bandwidths $h \propto (\log T/T)^{1/5}$, the conditional empirical process $\mathbb{M}_{iT}(\cdot|r)$ converges in the space of càdlàg functions to a Guassian process (Stute, 1986):

$$\left(\frac{T}{\log T}\right)^{2/5} (\mathbb{M}_{iT}(\cdot|r) - M_i(\cdot|r)) \rightsquigarrow W_i \,.$$

The functional delta method then implies that $\mathbb{S}_{iT}$ also converges to a Gaussian limit

$$\left(\frac{T}{\log T}\right)^{2/5} (\mathbb{S}_{iT}(\cdot|r) - S_i(\cdot|r)) \rightsquigarrow \phi'_{(M_i,M)}(W_i, W),$$

where $\phi'_{(M_i,M)}$ denotes the Hadamard derivative of ϕ at (M_i, M).

In contrast, the limiting behavior of the estimated equilibrium strategies, and hence $\mathbb{F}_{iT}$,

20. Marmer et al. (2016) use a trimming sequence to resolve a similar issue in the asymptotic behavior of their estimator.

21. Otherwise, if the reserve price is discrete or nonbinding, the estimators $\mathbb{M}_{iT}$ and $\mathbb{S}_{iT}$ converge at the $\sqrt{T}$ rate, and $g_i(\cdot|r)$ can be estimated with a uniform rate of convergence of $(T/\log T)^{2/5}$.

is governed by the slower uniform rate of convergence of $\mathbb{g}_{iT}(b,r)$.[22] To see this, I decompose the distance between $\mathbb{F}_{iT}$ and F_i into two pieces:

$$
\begin{aligned}
\sup_v |\mathbb{F}_{iT}(v|r) - F_i(v|r)| &= \sup_v \left| \mathbb{S}_{iT} \circ \left(\mathbf{1} + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}\right)^{-1}(v) - S_i \circ \left(\mathbf{1} + \frac{G_i}{g_i}\right)^{-1}(v) \right| \\
&= \sup_v \left| \mathbb{S}_{iT} \circ \left(\mathbf{1} + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}\right)^{-1}(v) - S_i \circ \left(\mathbf{1} + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}\right)^{-1}(v) \right. \\
&\qquad \left. + S_i \circ \left(\mathbf{1} + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}\right)^{-1}(v) - S_i \circ \left(\mathbf{1} + \frac{G_i}{g_i}\right)^{-1}(v) \right| \\
&\leq O_p\left((\log T/T)^{2/5}\right) + \\
&\qquad \sup_v \left| S_i \circ \left(\mathbf{1} + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}\right)^{-1}(v) - S_i \circ \left(\mathbf{1} + \frac{G_i}{g_i}\right)^{-1}(v) \right| .
\end{aligned}
$$

The first term converges at the rate derived above for $\mathbb{S}_{iT}$, while the second term converges at the slower optimal rate of convergence $(\log T/T)^{1/3}$ if $h \propto (\log T/T)^{1/6}$ (Stone, 1980).[23] Roughly speaking, the estimation error for the function S_i is eventually small relative to the error in the points at which it is evaluated. As a result, the probability that V_i falls between $b + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}(b|r)$ and $b + \frac{G_i}{g_i}(b|r)$ is the leading term in the limit process of $\mathbb{F}_{iT}$:

$$
\begin{aligned}
&\left(\frac{T}{\log T}\right)^{1/3} \sup_v |\mathbb{F}_{iT}(v|r) - F_i(v|r)| \\
&\quad = \left(\frac{T}{\log T}\right)^{1/3} \sup_v \left| F_i \circ \left(\mathbf{1} + \frac{G_i}{g_i}\right) \circ \left(\mathbf{1} + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}\right)^{-1}(v) - F_i(v) \right| + o_p(1) \\
&\quad = \left(\frac{T}{\log T}\right)^{1/3} \sup_b \left| F_i\left(b + \frac{G_i}{g_i}(b|r)\right) - F_i\left(b + \frac{\mathbb{G}_{iT}}{\mathbb{g}_{iT}}(b|r)\right) \right| + o_p(1) \,. \qquad (2.11)
\end{aligned}
$$

Because the density estimator, $\mathbb{g}_{it}$, uniformly converges to a Gaussian limit, the continuous mapping theorem guarantees that $\mathbb{F}_{iT}$ weakly converges to its absolutely continuous limit

22. I suppress the subscript $\mathcal{R}$ in $\mathbb{g}_{iT,\mathcal{R}}$ and $\mathbb{F}_{iT,\mathcal{R}}$ because this analysis applies to all possible configurations of the ring.

23. Note that the optimal sequence of bandwidths used to estimate $g_i(b|r)$ is generally not the same as the sequence used to estimate M_{iT} and is optimally larger.

in the space of càdlàg functions. Moreover, because g_i is bounded away from zero, the derivative of $g \mapsto F_i \circ (\mathbf{1} + G_i/g_i)$ exists and is bounded on the relevant domain. The limit process of $\mathbb{F}_{iT}$ can then be found by again applying the functional delta method.

2.5.4 Test Statistics

The proof of Theorem 2 demonstrates that bidder i's competitively rationalizing valuation distribution, $F_{i,\emptyset}$, is independent of the instrument, Z_i, if and only if bidder i is not colluding. Depending on whether Z_i is binary, discrete, or continuously distributed, the estimators and asymptotic analysis in subsections 2.5.2 and 2.5.3 can be used to construct consistent tests for collusion in a variety of ways. Looking ahead to the application to British Columbia's timber auctions, the continuously distributed reserve price will serve as the instrument, and I will propose a test for collusion based on a conditional Kendall's τ statistic.

Before discussing the continuous case in detail, however, I briefly consider the case of discrete instruments and formally demonstrate the existence of a consistent test for collusion using only prices, the winning bidders' identities, and exogenous variation in the level of competition. If the instrument is binary, then a test for collusion could be based on a Kolmogorov-Smirnov-type statistic such as

$$D_{iT} = \sup_{\{v\,:\,v>r,r'\}} \left| \mathbb{F}_{iT,\emptyset}(v|Z=z) - \mathbb{F}_{iT,\emptyset}(v|Z=z') \right| ,$$

where $Z = (r, \mathcal{N})$ and $Z' = (r', \mathcal{N}')$ are distinct levels of the instrument and I have reintroduced the dependence of $\mathbb{F}_{iT}$ on $\mathcal{N}$. For instance, D_{iT} would be an appropriate test statistic when the seller randomly decides to "set aside" some auctions for a particular category of bidder. Set-asides are common in government auctions as a means of subsidizing smaller firms (see, for example, Athey et al., 2013; Krasnokutskaya and Seim, 2011) and could be used as a test for collusion in those settings.

Theorem 3 establishes that $(T/\log T)^{2/5} D_{iT}$ has an absolutely continuous limiting distribution. When the optimal bandwidth is chosen, however, this distribution will not be centered about zero. Despite this asymptotic bias, its critical values can be estimated by resampling methods (Politis and Romano, 1994).

Theorem 3. *Assume $MA.1$-$MA.7$ and $IA.1$-$IA.3$. Suppose Z is a discrete random variable.*

(i) D_{iT} converges in probability to zero if and only if $i \notin \mathcal{R}$.

(ii) $\left(\frac{T}{\log}\right)^{2/5} D_{iT}$ converges in distribution to an absolutely continuous random variable if $i \notin \mathcal{R}$ and diverges if $i \in \mathcal{R}$.

Proof. See Appendix A.1. □

When the instrument takes on more than two values, generalizations of the Kolmogorov-Smirnov statistic could be used. Alternatively, Haile et al. (2003) observe that a test based on the mean valuations may converge at a faster rate. They propose an asymptotically chi-square statistic based on the insight that the mean valuation implied by the IPV model should be strictly decreasing in the number of symmetric bidders due to the worsening winner's curse. To adapt this statistic to a test for collusion, the null hypothesis would be the same, but the alternative would be the opposite. That is, the mean valuations should be strictly increasing in the number of symmetric bidders when the bidder is colluding.

Otherwise, when the instrument is continuously distributed, I propose a test based on correlations. Although this test would not detect dependence among higher moments of V_i, covariances may be sufficient. For instance, a collusive bidder's competitively rationalizing valuation may be positively correlated with the reserve price. Intuitively, this follows from the fact that the reserve price is a source of non-ring competition—i.e. competition from the seller. When the reserve price increases, the ring is not able to suppress as much of the competition at the auction, which reduces the difference between their competitively rationalizing valuation distributions and their true valuation distributions. Hence, their competitively rationalizing valuations are positively correlated with the reserve price.

To be more precise, the statistics would be of the form

$$\rho\left(\mathbb{1}\left\{\hat{V}_{i,\emptyset} > v_i^*\right\} \cdot \hat{V}_{i,\emptyset},\ Z_i\right)$$

for each i, where ρ is a measure of correlation and v_i^* is the minimum valuation at which bidder i's valuation distribution is identified. For convenience, I refer to this as the correlation between i's competitively rationalizing valuations and its instrument despite the truncation at v_i^*.

When Z_i is equal to the reserve price, testing for dependence is complicated by the fact that valuations are censored below the reserve, which will introduce spurious correlation between the observations of $V_{i,\emptyset}$ and r. But unlike the typical random censoring problem in survival analysis, the data always include the value at which the valuations would have been censored—i.e. the reserve price. Consequently, the assumption of independent censoring can be tested nonparametrically using a conditional Kendall's τ statistic.

Apart from being able to accommodate a continuously distributed instrument, the conditional Kendall's τ has another appealing feature: it incorporates information contained in the bidder's losing bids and participation decisions. Although, to prevent phantom bids submitted by other firms from biasing the estimate of bidder i's pseudo-valuations, I only use the prices to estimate the inverse bidding strategies because these are the only bids that are known to be serious. Even so, I substantially increase the amount of data available to test the null hypothesis for bidder i by evaluating that inverse bidding strategy at all of bidder i's bids, as opposed to only its winning bids. In fact, in the later application, I observe so few wins by each bidder that using their losing bids is necessary to generate power against the alternative.

Formally, I define the conditional Kendall's τ as

$$\hat{\tau}_i^C = \hat{\tau}^C(\hat{v}_{i,\emptyset}, r) = \frac{\displaystyle\sum_{t \leq s} \text{sign}(\hat{v}_{it,\emptyset} - \hat{v}_{is,\emptyset}) \cdot \text{sign}(r_t - r_s) \cdot \Lambda_{its}}{\displaystyle\sum_{t \leq s} \Lambda_{its}},$$

where Λ_{its} is an indicator for the event that both $(\hat{v}_{it,\emptyset}, \hat{v}_{is,\emptyset})$ and (r_t, r_s) can be ordered.[24] For example, if $\hat{v}_{it} < r_t < r_s < \hat{v}_{is}$, then the summand in the numerator can be evaluated even though $\hat{v}_{it}$ is not observed. Listing each of the other 23 possible permutations, it would be clear that the pairs are "orderable" if and only if $\max\{r_t, r_s\} < \max\{\hat{v}_{it}, \hat{v}_{is}\}$. This implies that one of the pseudo-valuations must be uncensored in order for the pair to enter the summation. Because only one of the observations needs to be uncensored, however, the auctions for which bidder i did not submit a bid will be represented in $\hat{\tau}_i^C$. In this sense, τ_i^C incorporates information contained in bidder i's participation decisions. Thus, unlike the Kolmogorov-Smirnov statistic defined above, this conditional Kendall's τ statistic allows the bidder to incriminate itself through its phantom bidding strategy.

If bidder i's valuations were directly observed, then this statistic would have an expected value of zero under the null hypothesis. This can be easily verified by noting that the summand evaluates to 1 in six of the twelve orderable permutations, and it evaluates to -1 in the other six. The mean of $\hat{\tau}^C(v_{i,\emptyset}, r)$ is zero under the null hypothesis because each of these permutations is equally likely when the bidder's valuations are independent of reserve prices. Moreover, because $\hat{\tau}^C(v_{i,\emptyset}, r)$ is the ratio of two U-statistics of degree two, its asymptotic normality would immediately follow.

Unfortunately inference using $\hat{\tau}_i^C$ is complicated by the fact that the pseudo-observations, $\hat{v}_{it,\emptyset}$, are estimated. As a result, $\hat{\tau}_i^C$ is an asymptotically biased estimator for the parameter

24. Similar statistics have been proposed as a test of independent truncation in medical trials (see, for example, Tsai, 1990; Martin and Betensky, 2005), but I do not believe that this particular statistic has been studied. No doubt, this is due to the unusual nature of the censoring problem induced by binding reserve prices. In the biostatistics literature, it would correspond to the clearly impossible situation in which the statistician observes the times at which subjects would have dropped out of the study if they had not died.

of interest

$$\tau^C(V_{i,\emptyset}, r) = E[\text{sign}(V_{i1,\emptyset} - V_{i2,\emptyset}) \cdot \text{sign}(r_1 - r_2) \,|\, \Lambda_{its} = 1]\,.$$

The analysis of this statistic is considerably more difficult than, for example, the mean pseudo-valuation estimated in Haile et al. (2003) because the estimation error enters through the discontinuous sign and indicator functions. Consequently, I cannot apply standard approaches to asymptotic analysis based on first-order approximations. This problem also falls outside the scope of the standard asymptotic theory for U-statistics because, as equation (2.10) shows, the estimation error in v_t is dependent on r_t by virtue of $\hat{v}_t$ being equal to the estimated inverse strategy function evaluated at (b_t, r_t).

Though I do not prove the bootstrap is consistent for the asymptotic distribution of $\hat{\tau}_i^C$, simulation evidence reported in the next section indicates that bootstrapped critical values perform well in a test for collusion. Specifically, I resample auctions with replacement to produce K samples of T auctions. Letting $\hat{\tau}_i^{C,(k)}$ denote the statistic for bidder i calculated in each of the bootstrapped samples $k = 1, \ldots, K$, the critical value for $\hat{\tau}_i^C$ is estimated as the $(1-\alpha)$-quantile of $\hat{\tau}_i^{C,(k)} - \tau_i^C$. Formally justifying this inference procedure is the subject of ongoing research.

2.5.5 Simulation Results

To demonstrate nonparametric identification from auction prices and exogenous variation in the level of competition, I simulate auctions with collusion and use the Kolmogorov-Smirnov statistic to test the null hypotheses that bidder i is not colluding. Each of the $n = |\mathcal{N}|$ bidders' valuations are drawn uniformly from the unit interval $[0, 1]$. I then numerically solve for the equilibrium bidding strategies when the bidders in $\mathcal{R}$ are colluding.[25] Each data set consists of the auction prices from $T = 1,000$ and 3,000 simulated auctions.

25. The system of ODEs that define the equilibrium strategies is difficult to solve numerically because the strategy function's derivative is indeterminate at the initial condition, i.e. the reserve price. Therefore, I transform the system of ODEs into a boundary value problem and approximate the solution using splines. See Appendix A.2 for details.

Table 1 reports the size and power of the Kolmogorov-Smirnov tests using rule-of-thumb kernel bandwidths proportional to $(\frac{T}{\log T})^{-1/5}$. The first column indicates the number of auctions in the simulated data set. The second and third columns report the total number of potential bidders and the number of bidders who are colluding. The identifying variation comes from either changes in the number of potential bidders or from changes in the number of ring members who are present, e.g. the first row indicates that three bidders were eligible to bid in half of the auctions while four were eligible to bid in the other half. The third column contains the fraction of null hypotheses that were rejected for bidders who were not in the ring, while the fourth reports the fraction for the colluders. In each case, the critical values for each of the null hypotheses were estimated as the $1-\alpha$ quantile of the bootstrap distribution centered around the test statistic. The bootstrap distribution was simulated from 1,500 samples of auctions resampled with replacement. The simulation results confirm that the test of the individual hypothesis that bidder i is not in the ring asymptotically controls size and its power to reject the null hypothesis tends toward one.

In the above simulations, a non-binding reserve price is held constant at zero. In order to simulate samples in which the reserve price is continuously distributed, I must solve for the equilibrium bidding functions at each of many different reserve prices. Because numerically solving for these functions can be computationally expensive, I restrict attention to a special case in which an analytic solution to the bidding strategies has been derived by Kaplan and Zamir (2012). In particular, I simulate auctions in which there is only one non-ring bidder whose valuation is uniformly distributed on $[\underline{v}_1, \bar{v}_1]$. The maximum of the ring's valuations is also uniformly distributed on $[\underline{v}_2, \bar{v}_2]$. For simplicity, I choose to make the ring members symmetric, i.e. their valuations are distributed according to $(\frac{v-\underline{v}_2}{\bar{v}_2-\underline{v}_2})^{1/|\mathcal{R}|}$ on $[\underline{v}_2, \bar{v}_2]$. Then, for each auction, the reserve price is independently drawn uniformly at random from $[0.1, 0.5]$ and all bidders use their optimal serious bidding strategy. Table 2 reports the results from 1,000 simulations.

In the simulation results reported in Table 2, I use 750 bootstrap samples to estimate the critical values for the $\hat{\tau}_i^C$ as described in the previous section. For comparison, the table includes simulations that use three different bandwidth sequences that variously trade off bias for variance in the estimate of $g_i(b, r)$. The optimal bandwidths for uniform convergence of the estimators are proportional to $(T/\log T)^{-1/6}$; $T^{-1/6}$ is the optimal rate for convergence in terms of integrated squared error; and bandwidths proportional to $T^{-1/5}$ undersmooth so that $\hat{v}_t$ is asymptotically unbiased. The results suggest that all three choices lead to tests that control size and have good power in finite samples.

2.6 Confidence Bounds on the Cost of Collusion

The identification argument naturally suggests a statistic to test the null hypothesis that a given bidder is not colluding. Asymptotically, this test will reject the null hypothesis with probability approaching one if the bidder is colluding, but it will also reject true null hypotheses with positive probability. Therefore, to consistently estimate the bidding ring's membership, the probability of rejecting a true null hypothesis must tend toward zero. This can be achieved by sending the family-wise error rate (FWER)—the probability of falsely rejecting one or more true null hypotheses—to zero at an appropriately slow rate as the data grow. If this rate is slow enough, the probability of rejecting false null hypotheses will still tend toward one, while the probability of rejecting true null hypotheses is made arbitrarily small.[26] A consistent estimate of the cost of collusion can then be computed by numerically solving for the price distribution in an auction in which no bidders collude.

To construct confidence bounds on the cost of collusion, let α be the chosen tolerance for the FWER. Typically, this is understood to mean that α asymptotically bounds the

26. The sequence for the FWER can be chosen so that the test makes type I and type II errors with probabilities that approach zero at the same rate. In finite samples, however, the probability of type II errors may be much larger than the conventional tolerance levels for type I errors. For instance, in the simulation for $\hat{\tau}_i^C$ with $T = 500$ auctions and a bandwidth proportional to $T^{-1/5}$, the probability of a type II error is roughly 58% while the probability of a type I error is much smaller at only 2%. Perhaps an intuitively appealing sequence for the FWER would attempt to balance the probability of type I and type II errors in finite samples. This might be approximately achieved by simulating the power of the test using a pilot estimate of the ring and the bidders' valuation distributions, but it will likely require unusually large tolerances for type I errors.

probability of rejecting a null hypothesis for a non-collusive firm. But this is also equivalent to the statement that the set of rejected hypotheses forms a lower confidence bound on the set of colluders, i.e.

$$\liminf_{T\to\infty} P\{\mathcal{R}_T \subseteq \mathcal{R}_0\} \geq 1-\alpha\,, \tag{2.12}$$

where $\mathcal{R}_0$ is the true collusive ring and $\mathcal{R}_T$ is the set of bidders for whom the null hypothesis is rejected given data from T auctions. Then, because the cost of collusion is monotonic in the ring (with respect to set inclusion), this translates into a lower confidence bound on the cost of collusion:

$$\liminf_{T\to\infty} P\{C(\mathcal{R}_T, F) \leq C(\mathcal{R}_0, F)\} \geq 1-\alpha\,, \tag{2.13}$$

where $C(\mathcal{R}, F)$ is the difference between the seller's expected revenues when the bidders in $\mathcal{R}$ are or are not colluding and bidders' valuations are jointly distributed according to F. The function C does not typically have an analytic expression, but its value at $(\mathcal{R}_T, F)$, and hence a lower confidence bound on the true cost of collusion, can be found numerically. Of course, this estimator is infeasible because F is not observed. The unknown joint valuation distribution must be replaced by a consistent estimator, which will generally differ from the joint distribution that rationalizes the data under the assumption that $\mathcal{R}_T$ is the set of colluders. Because equilibrium strategies are continuous in F with respect to the weak topology (Lebrun, 2002), a consistent "plug-in" estimator for $C(\mathcal{R}_T, F)$ is obtained by numerically evaluating $C(\mathcal{R}_T, \mathbb{F}_{T,\hat{\mathcal{R}}})$ for some consistent estimator of the ring, $\hat{\mathcal{R}}$.

CHAPTER 3

AN APPLICATION TO TIMBER AUCTIONS

3.1 The British Columbian Timber Market

3.1.1 Bidding Procedures

The auctions in my data were conducted under the auspices of the Small Business Forest Enterprise Program (SBFEP). The regional SBFEP offices published a list of the timber licenses to be sold at auction. Prior to each auction, the regional office would specify which firms were eligible to bid. To be eligible to bid in one of these auctions, a firm had to be registered with the SBFEP and not hold more than two outstanding timber licenses. In addition, the SBFEP announced whether registrants that owned or leased their own milling facilities would be eligible to participate. These firms, known as Category 2 registrants, were excluded from about 80% of the auctions between 1996 and 2000. The logging firms, known as Category 1 registrants, were eligible to bid in all of the auctions in my sample.

When the regional office solicited bids, it included several documents containing details about the timber license. These included survey maps, plans for extracting the logs, the estimated volume of merchantable timber by tree species, and projected road development costs. Along with these supporting documents, the office calculated a reserve price per cubic meter of harvested timber in accordance with one of the two appraisal methods described in the next subsection. Firms were also invited to inspect the tract themselves.

At any time before the auction closed, interested firms could submit a sealed "bonus" bid equal to the amount that they would pay above the reserve price. The regional office then opened and recorded all of the bids and awarded the license to the highest bidder. All of the bids and the identities of the bidders were announced at this time, as well. Throughout logging operations, the winner paid an amount per cubic meter of harvested timber equal to the reserve price plus its bonus bid.[1]

Following List et al. (2007) and Price (2008), I restrict my analysis to auctions in which the estimated timber volume was greater than 1,000 cubic meters.

3.1.2 Background on the Industry and US Trade Dispute

British Columbia's Ministry of Forests manages 95% of the province's timber supply. Its annual revenues averaged US$1.1 billion between 1996 and 2000, of which $210 million was raised by auctioning timber licenses under the SBFEP. These licenses grant the right to harvest timber from designated areas during a specified period of time, typically lasting about one year and no more than four years. Though these auctions directly accounted for less than 20% of its revenue, the auction prices affected the index that the Ministry used to benchmark its prices for all other timber licenses. Therefore, a natural question is whether the auction prices accurately reflect the fair market value of the harvesting rights.

The region's timber prices most recently came under scrutiny in 2002 when the previous lumber trade agreement between the US and Canada expired. US-led investigations concluded that the provincial Canadian governments had continued to subsidize their timber industries by selling harvesting rights at prices 18.8% below their fair market value. And in May of 2002, the US Department of Commerce (DOC) implemented a countervailing duty in that amount. Additionally, it calculated an 8.4% antidumping duty for a combined tax of 27.2% on lumber imports. In response to Canadian challenges to the US estimation methods, the countervailing duty was repeatedly revised downward and eventually settled at 11.2%.

Throughout negotiations, however, the DOC insisted that provincial governments reform their sales practices, which, in British Columbia, the largest timber harvesting province, involved expanding upon their new hedonic pricing system. Beginning in 1999, the hedonic pricing system was already being used to set the reserve price for SBFEP licenses. It replaced an older method that compared tracts of land based on their estimated profitability and

1. As part of the license, the winner agrees to pay penalties for unharvested timber.

adjusted the reserve prices to meet the Ministry of Forest's target price. Under the new appraisal system, non-SBFEP licenses were still appraised using the old method, but prices of SBFEP licenses were predicted by regressing sale prices on timber characteristics and cost indices.

The hope was that the hedonic pricing system would make the the region's timber prices more responsive to market conditions. As a necessary condition for its success, the SBFEP auctions must therefore provide an informative price signal. This depends in turn on whether collusion among bidders has a significant impact on the auction prices.

Two papers have recently addressed this question in reduced form analyses (List et al., 2007; Price, 2008). Both find evidence that is consistent with the hypothesis that some of the firms in this sample are colluding. In a regression of bids on covariates, List et al. (2007) observe that significantly negative estimated fixed effects for bidders are consistent with collusion by that bidder if all bidders' valuations are identically distributed. According to one of their regression specifications, 30% of the fixed effects were significantly negative at the 0.10 level, while only 12% were significantly positive. In support of their interpretation of these results as evidence of collusion, they also use theoretically motivated criteria to select firms whose behavior may be consistent with collusion. Specifically, they flag pairs of firms for further investigation if they bid in at least six auctions together, both win at least once, and the pairwise ranking of their bids is balanced. Of the 130 *a priori* suspicious firms that they identified, 55% have significantly negative estimated fixed effects. In contrast, 27% of non-suspect firms had significantly negative fixed effects.

List et al. also suggest incorporating these *a priori* guesses in the regression as an indicator for whether the bid was submitted by a suspected colluder. Though this variable is clearly measured with error, a significantly negative coefficient might suggest further evidence of collusion. As an additional proxy for collusion, they also include an indicator for whether firms are likely to interact outside of the auction setting because this would theoretically make anticompetitive behavior easier to sustain in equilibrium. To construct this indicator, they identify logging firms who are located in a district where the four largest firms in the industry have the greatest presence. These large firms do not participate in the SBFEP auctions, but they subcontract with auction participants to conduct much of their harvesting operations. Therefore, the potential for multimarket contact between logging firms is arguably greater in districts where the demand for harvesting labor is highest. When they include both indicators, the coefficient on their interaction is statistically significantly negative in one of the regression models, but they conclude that this approach does not reveal strong evidence for or against the null hypothesis of competitive bidding.

Price (2008) used the same *a priori* criteria to identify suspicious firms. He then showed that, conditional on a vector of covariates including the number of bidders who participate in the auction, the bids submitted by the suspected colluders were lower for pairs of suspects who were in close proximity to each other. Intuitively, this strategy uses the number of bidders at the auction as a proxy for the number of potential entrants. Then, conditional on a higher spatial concentration of suspected colluders, fewer of these potential entrants are actually competitors. The regression results then suggest that the suspected colluders would not bid as aggressively as predicted if another suspect replaced a bidder that is presumed to be competitive. At the same time, an alternative explanation might be that the *a priori* criteria are biased toward labeling firms as suspicious when their local competition is weaker than the average level of competition in the SBFEP auctions.

On the other hand, in technical reports prepared for the Ministry of Forests, Athey et al. (2002) and Athey and Cramton (2005) argue that their proposed auction reforms would reduce the benefit of anti-competitive bidding to the point that collusion would have to be pervasive within a local market and sustained over at least three years in order to significantly influence the market price for timber. Given the large number of small logging firms that could enter the market if firms successfully conspired to keep prices down, they conclude that this extent of collusion is implausible. Supposing that collusions did manage to reduce prices by 10%, however, they estimate that this would affect the market price for timber by

1–6% under their proposed pricing system.

Thus, the question remains whether the identification strategy presented in the preceding chapter can shed more light on the competitiveness of the SBFEP auctions between 1996 and 2000 and what effect any collusion may have had on the auction prices.

3.1.3 Identifying Variation

In the six years leading up to this dispute, substantial variation in the reserve prices allows me to test whether collusion played a role in keeping the Canadian lumber prices below prices in the United States. For timber appraised using the old non-hedonic pricing formula, the reserve price was set at a fixed fraction of the timber's estimated price per cubic meter.[2] This estimated price was computed as the difference between the tract's appraised value and the average value of land sold in the region, plus a base rate that was determined by the Revenue Branch. The base rate was adjusted quarterly depending on how the prices of active licenses compared with the Ministry's target rate, which in turn was a piecewise-linear function of a weighted average of British Columbia's lumber and wood chip price indices. In addition, the Ministry's district offices could raise the reserve price to cover any silvicultural or development costs that they expected to incur while administering the land. As a matter of practice, however, these levies were only added in Category 2 auctions. As a result of all these modifications to the ministry's initial appraisal, I observe auctions in which the appraised value of the licenses are similar while their reserve prices greatly differ.

For timber appraised using the new hedonic pricing formula, the reserve price was almost always set at 70% of the appraised value without the option of adjusting for silvicultural and development expenses. Though the hedonic pricing formula was not modified during the period under investigation, this change in reserve pricing policy relative to the pre-1999 method provides further exogenous variation in the reserve price. Table 3 reports the number

2. The details of this appraisal process are contained in the annually updated Interior Appraisal Manual. The appraisal manual that took effect on October 1, 1999 (Canada. B.C. Ministry of Forests, 1999) describes the reserve pricing policies for timber licenses appraised using either of the two methods.

of auctions that used each appraisal method by fiscal year.

The data also contain variation in the set of eligible bidders, but I choose not to use this as an additional test for collusion. This variation comes from two sources. First, in an effort to foster employment among small logging firms, the Ministry of Forests restricts participation in most auctions to bidders that do not own or lease any milling equipment. Though its decision to set aside auctions for Category 1 firms does not appear to be related to the Ministry's appraisal of the timber licenses, I do find evidence that the decision to restrict participation in the auction is negatively correlated with my own estimates of the licenses' value. If, in addition, the Ministry's decision is somehow correlated with the idiosyncratic component of a firm's valuations, then the test for collusion would not be valid. Regardless, this test would have little power to reject the null due to the small sample sizes for Category 2 auctions.

Second, variation in the set of eligible bidders comes from the limit on the number of outstanding contracts that bidders are allowed to hold at any one time. Unfortunately, I do not observe any non-SBFEP licenses that the bidders might hold. I also do not observe whether firms completed logging operations before the license expired or applied for an extension. Though I can construct a proxy for a firm's eligibility based on its past bids, the observable variation in the set of eligible bidders is unlikely to provide a reliable test for collusion.

3.2 Analysis

To relate the data from the SBFEP's Category 1 auctions to the general empirical framework described in section 2.5, I must defend the assumptions of the asymmetric IPV model as well as make two further assumptions about bidders' beliefs and valuations. First, I assume that variation in bidders' beliefs about the set of potential bidders can be accounted for by conditioning on the geographic districts defined by the Ministry of Forests. Second, I assume that bidders' valuations are linear in the observable license characteristics so that the estimated coefficients can be used to produce a fitted appraisal value for each license, which I then use to homogenize the bids and reserve prices. Though the Ministry of Forests produced its own appraisals for each timber license, I argue that my appraisal method captures more of the variation in observable characteristics that is relevant to the firms' bids.

Under these additional assumptions, I then implement the methods of section 2.5 to measure the dependence between the competitively rationalizing valuations and the reserve prices. But, in an effort to be robust to misspecification in the model of the bidders' participation, I also discuss an alternative version of the conditional Kendall's τ statistic that does not incorporate information contained in the bidders' participation decisions.

3.2.1 The Modeling Assumptions

The valuations are assumed to be private, independent across firms, independent of the district in which the firm competes, and independent across auctions after controlling for the observed auction covariates. The firms' valuations are likely to be private because there is no active spot market for harvested timbers and the winner only pays for the amount of timber that they actually harvest. Loggers typically negotiate bilateral supply agreements with local mills before bidding in an auction. Therefore, the firms know the price per cubic meter at which they would be able to sell the logs if they were to win the auction. Furthermore, because the winners' payments to the Ministry of Forests are based on the actual volume of harvested timber, as opposed to the estimated volume of timber, they do not bear any risk regarding the total merchantable volume covered by the license.

Firms are also largely insured against uncertainty about the composition and quality of the timber because the Ministry fixes the stumpage rate for low-quality timber and timbers that are to be used for fence posts or other specialty products. Consequently, the bids only apply to good-quality coniferous sawlogs.[3] Different tree species within this category

3. If the Ministry sets the price for low-grade and specialty timbers too high (or low) relative to a bidder's

are typically substitutable in their commercial uses, so variation in the species composition should not greatly affect the bidders' valuations for the license. Moreover, as Paarsch (1997) argues, firms are not likely to have significantly different information about the species composition. Therefore, unlike in pure common value auctions, the winner's curse is unlikely to be an important factor in the firms' bidding strategies.

Nonetheless, the firms' willingness to pay for a license are certainly correlated. I must therefore assume that the idiosyncratic components of their valuations are independent conditional on the observable characteristics of the timber license. Given the rich set of covariates in my data, I would argue that this is plausible. Indeed, I am able to condition on the same set of covariates that the Ministry of Forests used to appraise the licenses.[4]

While the IPV assumption is consistent with prior studies of these auctions, I depart from the norm by allowing for arbitrary asymmetries across bidders, which may arise from differences in expertise, harvesting capacity, or relationships with local mills. These asymmetries could be significant considering the fact that the firms in this sample have very different patterns of participation. The majority of firms participated in very few auctions, and won at most once between 1996 and 2000. In contrast, there are only nine firms that won more than 10 auctions. These firms appear to rely on the SBFEP auctions much more than the less active firms, and could have a different business relationship with mills than the fringe competitors do.

Importantly, however, the unobserved heterogeneity in bidders' valuations is independent of the district in which the auction takes place. While this assumption is consistent earlier

valuation, the bidder would have an incentive to decrease (increase) its bid on licenses containing a higher proportion of that timber. In an attempt to determine whether this is an issue, I test whether the proportion low-quality timber that was actually harvested is correlated with the auction price after controlling for the license characteristics using the methods of Subsection 3.2.3. Unfortunately, I only observe the volume of harvested timber for a subset of the auctions in my sample. For these auctions, however, the proportion of low-quality timber is not significantly correlated with the unexplained variation in the winning bids.

4. In Subsection 3.2.3, I estimate my own appraisal of the timber licenses and find that it better predicts the winning bids than the Ministry's appraisal. It is still reasonable, however, that the Ministry selected an appropriate set of variables to include in their hedonic pricing model even though their exact formula could be improved.

work (Paarsch, 1997; List et al., 2007; Price, 2008), it would clearly not be a palatable assumption if the distance from the firm's headquarters to the harvesting site were correlated with the firm's harvesting costs. Because the data include the cycle time—the estimated time necessary to transport the timbers to the nearest point of appraisal—I argue that, conditional on the cycle time, the headquarters-to-site distance is not likely to be correlated with the cost of extracting the timber.

Finally, I acknowledge that inter-auction dynamics may be induced by capacity constraints and the Ministry's cap on the number of outstanding licenses that firms can hold. A firm's bidding strategy would then depend on the current state of the market, which might be summarized by the characteristics of future auctions that have already been scheduled, expected characteristics of auctions that have not been announced, and all of the firms' backlogged workloads. I observe the time series of auction characteristics and could assume that firms have rational expectations for the licenses that are likely to come to auction in the future. But, as discussed in the previous section, I only have noisy measures of the firms' backlogs. If they were observable, perhaps the methods of Jofre-Bonet and Pesendorfer (2003) could be employed to account for auction dynamics in the estimation. This approach, however, would be limited by the infrequency with which bidders participate in auctions, the high dimensionality of the state vector, and the unobservable asymmetries among the bidders. Nonetheless, to the extent that the data will allow, I attempt to address the consequences that auction dynamics would have on the hypothesis test results in Appendix A.3.

3.2.2 Conditioning on the Set of Potential Bidders

As a prerequisite to all the estimators that have been discussed so far, I must be able to condition on the sets of firms that each bidder i deems to be potential competitors. In line with earlier work on SBFEP auctions (Paarsch, 1997), I assume that the 31 geographic districts defined by the Ministry of Forests constitute submarkets in the sense that all bidders hold the same beliefs about the distribution of potential entrants for any auction in a given district. The bidders' equilibrium strategies can then be estimated conditional on each district.

This assumption appears to be supported by the data. Figure 2 depicts the participation patterns of bidders across auctions. Each auction is colored by district and arranged along the horizontal axis, while individual bidders are represented along the vertical axis and grouped according to their propensities to bid in each district. A dot at point (t, i) in the plane represents a bid by bidder i in auction t. To give a sense of the intensity of bidder i's bidding in each district, the opacity of the dot indicates whether bidder i's bid was the highest, second highest, third highest, and so on.

The bands of dots of the same color indicate that the set of firms that are most likely to bid at an auction is fairly constant within each district. At the same time, the large areas of white space in the figure illustrate that the sets of active bidders differ widely across districts.

There are, however, neighboring districts in which the same sets of bidders have similar participation rates. If bidders' beliefs about $\mathcal{N}_t$ are in fact the same for auctions in these districts, then they could be pooled in order to more precisely estimate the competing distributions. But this raises two key questions. First, what is the right notion of similarity between districts? And, second, how similar should districts be to justify pooling them in the estimation?

Regarding the first question, I argue that two districts should be grouped into the same market if the same bidders have similar rates of participation in both districts. For example, suppose there are two bidders. If one bidder participates in 5% of the auctions in district 1 and 4% of the auctions in district 2 while the other bidder participates in 20% of the auctions in district 1 and 18% of the auctions in district 2, then the probability distribution over sets of competitors are approximately the same in both districts. Consequently, the bidders' competing distributions will be similar in the two districts, and it may be reasonable to pool these districts in the estimation.

I therefore define the distance between districts l and k by

$$d(k,l) = \sum_i \frac{|w_{il} - w_{ik}|}{w_{il} + w_{ik}},$$

where w_{il} is the frequency of participation by bidder i in district l, and the sum is taken over all bidders i for whom w_{il} and w_{ik} are not both zero. This measure of distance is known as the Canberra distance and can be viewed as a weighted L_1-distance between the vectors w_l and w_k.[5]

In answer to the second question of how similar is similar enough, I do not take a definitive position. Instead, I use a hierarchical clustering algorithm that produces a sequence of increasingly coarse market definitions. Initially, the algorithm assigns each of the 31 districts to its own market (or cluster). In the first step, the algorithm merges the two districts that are closest to each other and computes a measure of dissimilarity among the 30 resulting markets. It then iteratively merges the two most similar markets and recalculates the dissimilarities until only two markets remain. If the measure of dissimilarity between clusters is chosen appropriately, this algorithm produces a sequence of nested partitions of the auctions.[6] I can then perform the analysis using any one these market definitions. As an example of the partitions that the clustering algorithm produces, Figure 3 shows the participation patterns when the districts are partitioned into nine markets.

These market definitions appear reasonable with the exception of the smallest and last to be merged of the nine clusters. The districts in this cluster are small in the sense that there were very few licenses sold and in the sense that the fewest number of bidders ever

5. I also experimented with other metrics, such as the Euclidean distance, but found that they produced less reasonable results. For example, the Euclidean distance might consider two districts to be very similar even though the identities of the firms that constitute the competitive fringe do not overlap. This under-emphasis on matching based on the zeros in w_l and w_k then led the hierarchical clustering algorithm to group together discontiguous districts. The Canberra distance appears to be better suited to the sparse participation patterns in the data.

6. I use Ward's method of defining dissimilarities between clusters, which is designed to produce clusters that have minimal within-cluster variance.

participated in them. Apart from the fact that most firms never participated in them, however, they do not seem to have anything in common. In fact, none of the firms that participated in one of the districts ever participated in either of the other two districts in the cluster. I will therefore omit them from the subsequent analysis because they are too small to include individually and too dissimilar to justify pooling together.

Other output from the hierarchical clustering algorithm is illustrated by the dendrogram in Figure 4. Each "leaf" of the dendrogram represents a district as defined by the Ministry of Forests. The height at which two leaves join represents the dissimilarity between the markets that were merged to form a new market. The tree can be "cut" at different heights to produce the various market definitions that will be used in the analysis. For example, the dashed horizontal line in Figure 4 shows the cut that partitions the districts into nine markets.

At the bottom of the figure, I sum the number of Category 1 auctions in each district between 1996 and 2000. I also sum the auctions in each of the illustrated markets. This partition into nine submarkets will be the default market definition in the following analysis, so these sums indicate the number of independent observations available to estimate the bidders' market-specific inverse bidding strategy.

3.2.3 Homogenizing the Bids and Reserve Prices

The last assumption that I make about bidders' valuations is that auction level heterogeneity is linear in the observable characteristics:

$$u_{it} = \beta' x_t + v_{it} \, ,$$

where u_{it} is bidder i's payoff from winning auction t and v_{it} is the unobservable idiosyncratic component of i's payoff. If v_{it} is independent of x_t, then the following regression equation will hold:

$$\tilde{b}_{it} = \beta' x_t + \eta(\tilde{r}_t - \beta' x_t, \mathcal{N}_t) + \epsilon_{it} \, ,$$

where

$$\epsilon_{it} = v_{it} - \eta(\tilde{r}_t - \beta' x_t, \mathcal{N}_t) - \frac{G_i}{g_i}(v_{it} | \tilde{r}_t - \beta' x_t, \mathcal{N}_t)$$

is an independent error term with variance $\sigma^2(\tilde{r}_t - \beta' x_t, \mathcal{N}_t)$ and η is an unknown function of the screening level—the threshold below which idiosyncratic valuations will be censored by the reserve price—and the set of eligible participants. If the reserve price were never binding, then none of the bids would be censored and the above equation could be estimated via OLS with fixed effects for $\mathcal{N}_t$. On the other hand, if the reserve price binds, bidder i's competing distribution in auction t will be correlated with x_t to the extent that x_t is correlated with the screening level. The OLS estimates will therefore suffer from an omitted variable bias.

To account for the fact that the reserve price is often binding in SBFEP auctions, I use a slightly modified version of the partially linear single-index model (PLSIM) in Liang et al. (2010). The estimator, $\hat{\beta}$, minimizes the sum of squared residuals in a local linear regression of $\tilde{p}_t - \beta' x_t$ on the scalar index $\tilde{r}_t - \beta' x_t$, where $\tilde{p}_t$ is the observed price at auction t. In their model, however, the covariates that enter linearly are not constrained to be the same as those that enter nonlinearly through η. Additionally, because they do not allow the function η to depend on a categorical variable, I estimate the local linear regression independently for each market, while constraining $\hat{\beta}$ to be constant across markets. Thus, my estimator is the same as theirs, except I impose the obvious equality constraints across markets and on the coefficients in the linear and nonlinear components of the regression. Consequently, my estimator for β inherits the nice asymptotic properties of theirs. In particular, $\hat{\beta}$ is $\sqrt{T}$-consistent and asymptotically normal with a limiting variance of $M^{-1} V M^{-1}$, where

$$M = E[(1 - \eta'(\tilde{r} - \beta' x, \mathcal{N}))\, x x']$$
$$V = E[(1 - \eta'(\tilde{r} - \beta' x, \mathcal{N})\, \sigma^2(\tilde{r} - \beta' x, \mathcal{N})\, x x']$$

and $\eta'(\cdot, \mathcal{N}_t)$ is the derivative of η with respect to its first argument.

For the local linear regression estimator of η, I use a quartic kernel to ensure differentiability of the least-squares objective function. To improve the performance in sparse regions on the data, I also include ridge regression parameters, $\lambda_{\mathcal{N}} > 0$, for each market. Both the kernel bandwidths and the ridge regression parameters were selected to minimize the leave-one-out cross-validated mean integrated squared error at $\beta = \hat{\beta}$,

$$\sum_t \left(\tilde{p}_t - \beta^T x_t - \hat{\eta}_{-t}(\tilde{r}_t - \beta^T x_t, \mathcal{N}_t) \right)^2 \bigg|_{\beta=\hat{\beta}}$$

where $\hat{\eta}_{-t}$ is the local linear ridge regression estimator computed from all auctions excluding auction t.

Descriptions and summary statistics for the variables are in Tables 4 and 5. Estimates for β are reported in Table 6. For comparison, the first column reports the OLS estimates from a regression that includes market fixed effects for eight of the nine clusters defined in the previous section. These estimates will be biased to the extent that the reserve price is binding. The first PLSIM that I estimate includes only the covariates that the Ministry of Forests considered in its original hedonic pricing formula. Though the Ministry used a dummy variable to indicate whether 60% or more of the estimated volume came from a combination of hemlock or balsam trees, I opt to include these as separate continuous variables. The second model reflects changes to the pricing formula that the Ministry of Forests introduced in 2001, after the auctions in my data had been held. The Ministry omitted the quality index variable and added the estimated volume of white pine as a percent of the total volume. It also amended the hemlock/balsam indicator variable to include cedar and adjusted the threshold so that it was equal to one if the fraction of estimated volume from hemlock, balsam, and cedar was greater than 50% of the total estimated volume. I again opted to include the proportion of cedar volume as a separate continuous variable. Finally, the third model combines the variables from the Ministry's two formulas. As is evident in Table 7, the fitted values, $\hat{\beta}' x_t$, from the three models are highly correlated with each other.

Figure 5 plots the winning bonus bid against the estimated appraisal price from Model 3 minus the reserve price, $\hat{\beta}' x_t - \tilde{r}_t$.[7] As $\hat{\beta}' x_t - \tilde{r}_t$ increases, the screening level decreases and bidders are more likely to participate with a wider range of valuations. Consequently, the variance in the winning bonus bid increases. In contrast, at low values of $\hat{\beta}' x_t - \tilde{r}_t$, the bonus bids are predictably smaller and have noticeably smaller variance. Thus, the general upward trend and increasing variance in Figure 5 suggests that the estimated appraisal price accurately summarizes variation in the covariates that is relevant to bidders' valuations. In addition, the wide range in the observed screening levels indicates that there is substantial variation with which it will be possible to test for collusion.[8]

In Figure 6, I replicated Figure 5 using the Ministry's appraisal, which I computed by inverting the formulas for the reserve price in the Ministry's Appraisal Manual. The Ministry's appraisal clearly does not predict the winning bonus bid as well as the PLSIM estimates do. In fact, the appraisals conducted according to the old non-hedonic policy are slightly negatively associated with the winning bonus bid (Spearman's $\rho = -0.066$), while the appraisals obtained through the Ministry's hedonic pricing formula are moderately related (Spearman's $\rho = 0.207$).[9] In contrast, the appraisals obtained from the Model 3 have a rank correlation of $\rho = 0.363$. Although I will also report the results of the tests for collusion using the Ministry's appraisal, I prefer to homogenize the bids and reserve prices

7. Any constant term in the hedonic pricing formula is not identified separately from the level of η. To make Figure 5, I normalize the level of the appraisal so that the overall mean of $\hat{\beta}' x_t - \tilde{r}_t$ is equal to the mean winning bonus bid.

8. To verify that the estimated screening level affects the bidder's participation decisions, I regress the number of participants on the estimated screening level and a constant, fully interacted with indicator variables for each district. On average, a one-standard-deviation increase in the estimated screening level is associated with a decrease of 0.88 in the number of bids submitted. Though measurement error might attenuate the estimated coefficient on the screening level, the estimate was significantly different from zero ($p = 0.05$). In addition, the estimates of $\eta(\cdot, \mathcal{N}_t)$ were monotonically increasing functions of the screening level. If the reserve prices did not bind, then η should be constant within each market.

9. I report the rank correlation because the relationship between the screening level $r - \beta^T x_t$ and the winning bonus bid need not be linear.

with my estimated appraisal.

The time series in Figure 7 provides an alternative view of the identifying variation in the screening level. Before the policy change in 1999, the Ministry of Forests' reserve price varied significantly relative to my estimated value of the timber license. After 1999, the hedonic appraisal method successfully reduced the variance in the reserve price. It also slightly decreased the average. These changes are more drastic when I recreate the plot using the Ministry's appraised values in Figure 8. But, for the reasons discussed in the preceding paragraph, I believe this overstates the amount of exogenous variation in the reserve price.

3.2.4 Using Participation Decisions to Test for Collusion

As mentioned above, the typical bidder participates in a small fraction of the auctions for which it appears to be eligible. Much of this can be explained by geography: bidders typically focus their activities in just one or two neighboring districts. So, as a first approximation to their true participation decision process, I assume that firms never considered participating in auctions outside of the districts in which they were active. Still, some firms only bid in a nearby district a few times, so it may be incorrect to assume that they potentially would have entered every auction in that district if the reserve price had not been too high. I therefore define a firm as being active in a district only if it participated in at least 5% of the auctions for which my data indicates it was eligible.[10,11]

Yet, even in the districts where they are most active, firms typically participate in less than a third of the auctions for which they are eligible. To rationalize this behavior, I therefore rely on the fact that the reserve price is often binding.[12] This creates a censor-

10. To construct a proxy for a firm's eligibility, I track all of the SBFEP licenses it has won but have not yet expired. If I observe a firm bid in an auction when my data suggest it has three outstanding licenses, I presume that it completed logging operations on one of them so that, in fact, it only has two outstanding licenses. If it does not win the present auction, then it will be eligible to bid in all subsequent auctions.

11. This 5%-threshold is arbitrary and can be varied to check for robustness.

12. Alternatively, the lack of participation could be explained if participating in auctions is costly either because firms incur costs in order to evaluate the timber license and learn their valuation or because the process of submitting a bid is costly. In these data, there are no direct costs of participating in an auction,

ing problem in which the valuations are left-censored at the reserve price but the value at which they would have been censored is always observed. Thus, the data consists of $(\max\{\hat{v}_{it}, r_t\}, r_t, \mathbb{1}\{v_{it} > r_t\})$ for auctions t in districts where bidder i was active or in which bidder i participated. Given these data, there are several ways to test for independence between $\hat{v}_i$ and r, but I have found that a conditional Kendall's τ estimator performs well in simulations.

More precisely, this testing procedure performs well in simulations where each bidder is a potential entrant in all of the simulated auctions. In the present application, however, it is possible that a bidder's participation decision was affected by unobservable factors, such as the number of non-SBFEP contracts held by the bidder at the time of bidding. As a more conservative approach to the testing problem, it may be prudent to ignore any information that might or might not be contained in bidder i's participation decision.

In this case, the data would consist only of $(\hat{v}_{it}, r_t)|\hat{v}_{it} > r_t$, which is precisely the typical case of left-truncation discussed in Tsai (1990) and Martin and Betensky (2005). An appropriate test statistic would then be defined analogously to $\hat{\tau}_i^C$ above, except that Λ_{ts} is replaced by an indicator for the event that $\max\{r_t, r_s\} < \min\{\hat{v}_{it}, \hat{v}_{is}\}$. In words, the pair (t, s) is only included in the computation of $\hat{\tau}_i^T$ if the estimated valuations are "comparable" in the sense that i's valuation was uncensored in both auctions and would have remained uncensored if the reserve prices were interchanged. Clearly, this condition is more restrictive, so $\hat{\tau}_i^T$ uses less of the data to test the null hypothesis. As a result, it will generally be estimated with greater variance but will avoid any bias that comes from erroneously assuming that valuations were censored.

I use $\hat{\tau}_i^T$ to denote this statistic, where the superscript T indicates that $\hat{\tau}_i^T$ is a function of the truncated data $(\hat{v}_{it}, r_t)|v_t > r_t$ as opposed to $\hat{\tau}_i^C$, which uses the full sample of censored pseudo-valuations and reserve prices. The population parameter $\tau_i^T(V_{i,\emptyset}, r)$ is defined

and the cost of evaluating the timber license is likely to be significantly reduced because the Ministry of Forests' shares detailed information with all potential bidders.

analogously.

3.2.5 Hypothesis Test Results

I formulate a test of the null hypothesis that bidder i is not colluding either as a one-sided test of $\tau^C(V_{i,\emptyset}, r) \leq 0$ against $\tau^C(V_{i,\emptyset}, r) > 0$ or as a test of $\tau^T(V_{i,\emptyset}, r) \leq 0$ against $\tau^T(V_{i,\emptyset}, r) > 0$. To asymptotically control the probability of making one or more false rejections, I adopt the bootstrap-based multiple testing procedure in Romano and Wolf (2010). This procedure begins by effectively "studentizing" the statistics by estimating marginal p-values, $\hat{p}_i$ and $\hat{p}_i^{(k)}$, for the test statistics and all of the bootstrap replicates indexed by $k = 1, \ldots, K$.[13] The algorithm then proceeds by iteratively comparing the smallest $\hat{p}_i$ for which H_i has not yet been rejected to an estimated critical value. Specifically, if A is the set of unrejected hypotheses, it rejects i if $\hat{p}_i$ is less than the α-quantile from the sample of $\min_{i \in A} \hat{p}_i^{(k)}$. The algorithm stops when no further rejections can be made.

Writing the critical values in this way highlights two features of this testing procedure that increase the probability of detecting collusion. First, it incorporates information about the set of true null hypotheses by restricting the minimum to the provisionally accepted hypotheses. As a result, the critical value increases with each iteration, thereby making it easier to make subsequent rejections. This contrasts with the Bonferroni correction, which compares all the p-values to α/n, where n is the number of hypotheses under consideration. Second, it incorporates information about the joint distribution of p-values by choosing critical values from the distribution of their minimum. Thus, it will make weakly more rejections than the Holm procedure, which would be equivalent to setting the critical value to $\alpha/|A|$ at each step of the algorithm.

As a final consideration before applying the testing procedure, I must decide which null hypotheses to test because there are not enough data to simultaneously test for collusion by

13. Working with p-values as opposed to the raw test statistics asymptotically balances the probability of making false rejections across the individual hypotheses.

all of the firms while controlling the FWER. More than 1,400 firms bid in a SBFEP auction between 1996 and 2000, but most of these firms participated in three or fewer auctions. Moreover, given that the ultimate goal is to estimate the cost of collusion, this comprehensive analysis would not be optimal because the estimated effect on expected revenue is minimally affected by whether one of the small firms is colluding. A better strategy would be to assume that the smallest firms are competitive in order to increase the probability of rejecting the null hypothesis for larger firms. In other words, the point estimate and confidence bound on the cost of collusion can be improved by judiciously allocating statistical power across the individual hypotheses. In fact, a hypothetically optimal testing procedure would weight each of the null hypotheses based on the relative probability of rejecting the null and their effect on the estimated cost.

Developing an adaptive method of determining these optimal weights is outside the scope of this research. Instead, as a rough approximation to the optimal procedure, I simultaneously test the null hypothesis for each of the 17 firms that bid in more than 30 auctions and won more than five of them. The precise thresholds are arbitrary, but as long as they are chosen in a manner that is independent of the test statistics, the procedure described above will asymptotically control the FWER.

The results of the testing procedure are reported in Table 8.[14] The marginal p-values for each firm are estimated from 20,000 bootstrap samples, while the adjusted p-values are equal to the smallest α for which the multiple testing procedure would have rejected the null hypothesis for firm i while controlling the FWER at level α. The panels of Table 8 correspond to the different choices of test statistic and appraisal method. The firms are sorted according to their p-values in the first panel, which reports the results using the hedonic pricing formula estimated in Model 3 and the conditional Kendall's τ that does not incorporate information from bidder i's participation decisions.[15]

14. The diagonal entries in Table 9 indicate the total number of pseudo-valuations that were used to estimate the conditional Kendall's τ statistics.

15. This table was generated using a partition of the region into nine submarkets, but the qualitative

The method of homogenizing the auctions and the choice of test statistic have a large impact on the decisions to be made for each null hypothesis. This is perhaps unsurprising given the differences in Figures 5 and 6 and in the definition of the statistics. As discussed above, I prefer to homogenize the auctions using the hedonic pricing formula estimated from Model 3 primarily because it better predicts the bonus bids and the general trend in Figure 5 is more consistent with theory. In addition, the correlations reported in Table 7 suggest that the Ministry's two different appraisal methods produced substantially different estimates of the timber's value. After 1999, their hedonic pricing formula's produced appraisals that are more closely related to the estimates from Model 3 ($\rho = 0.884$) than the pre-1999, non-hedonic appraisals are ($\rho = 0.707$). Because of this abrupt change in the way the Ministry's appraisals covary with the observed timber characteristics, it is unlikely that both methods produced accurate appraisals. In any event, the estimates from Model 3 seem to better capture the effect of timber characteristics on the observed bids.

Regarding the choice of statistic, I argue that $\tilde{\tau}$ is the more defensible option because it is more robust to misspecification in the bidders' participation decisions. Thus, according to this preferred specification, the null hypothesis for firm 1 can be rejected at the 0.01 level of significance, while the null hypothesis for firm 2 is rejected at the 0.10 level. Together, firms 1 and 2 form a lower 90%-confidence bound on the set of firms that colluded in the SBFEP auctions.

One drawback to my detection method is that the results could be difficult to interpret. If, for example, only one null hypothesis is rejected, then the test does not indicate with whom they might be colluding. Similarly, if the test detects collusion by more than one firm but they only ever bid in very distant markets, then they would have very little incentive to collude. This would cast doubt on the interpretation of the test as a test for collusion. To address this concern, I tabulate the number of times that each of the 17 firms bid in the same auction as one another in Table 9. These patterns demonstrate that firms 1 and 2 are

results are the same for a wide range of market definitions.

indeed active in the same market.[16]

Though Table 9 indicates the markets in which firms' participation patterns overlap, it might not accurately describe the extent to which the firms are in competition with each other because a collusive bidding ring could be manipulating its members' decisions to enter. But, if the ring colludes efficiently, the number of auctions that each firm wins in a market indicates its strength as a competitor and the potential benefits to collusion. To that end, Table 10 counts the number of times that a firm won in a district where the other firms have ever won an auction. These tabulations suggest that firms 7 and 9 could be firms 1 and 2's biggest rivals. On the other hand, they could also be colluding with firms 1 and 2 even though none of the tests would have detected them.

Having rejected the joint null hypothesis that all of the bidders are competitive, a logical question is whether a collusive model can better fit the data. In particular, the question is whether assuming firms 1 and 2 are members of a collusive bidding ring makes the estimated Kendall's τ statistics less significantly positive. To that end, I perform the testing procedure again, but instead use the valuations that rationalize the observed bids when firms 1 and 2 collude efficiently, i.e. $\mathcal{R} = \{1, 2\}$. The test statistic is unaffected for the non-ring bidders $i \notin \mathcal{R}$. In contrast, the collusive bidders' pseudo-valuations are greater to the extent that they win in the same market because their competing distribution is weaker than those estimated under the null hypothesis. Furthermore, the pseudo-valuations are censored at $z_{it} = \max\{r_t, \hat{v}_{jt} : j \in \mathcal{R}, j \neq i\}$ rather than r_t because any phantom bids might be unrelated to the bidders' valuations.[17] Then, under the maintained assumptions that the true valuations are mutually independent and independent of the reserve price, both versions of Kendall's τ—$\tau_i^C(V_{i,\mathcal{R}}, Z_i)$ and $\tau_i^T(V_{i,\mathcal{R}}, Z_i)$—should be equal to zero if $\mathcal{R}$ is the true set of colluders.

16. Additionally, Table 9 shows that the data are insufficient to perform pairwise comparisons of the firms' simultaneous bids.

17. Recall that $\hat{v}_{jt} > \hat{v}_{it}$ if and only if $v_{jt} > v_{it}$ when i and j are members of the ring because their estimated inverse bidding strategies are numerically identical and strictly increasing. That is, $\hat{v}_{jt} > \hat{v}_{it}$ if and only if $b_{jt} > b_{it}$. Hence, there is no error in observing the event that a colluder's valuation is censored.

The results of this testing procedure are reported in the first panel of Table 11. Though their test statistics are slightly less significant, the test still rejects their null hypotheses at the 0.10 level. Perhaps this suggests that the original test rejections were not due to collusion, but it is also consistent with the hypothesis that the collusive bidding ring is larger than just firms 1 and 2. Therefore, I next refer to Table 10 in order to add the firms that win in the same districts as firm 1 and 2 to the set of suspected colluders. Adding firms 5, 6, 8, 11, and 17 does not appreciably affect the test statistics for 1 and 2. If firm 12 is assumed to belong to the ring, however, none of the hypotheses can be rejected at the 0.10 level. In fact, the test fails to reject any null hypotheses if and only if firm 12 is included in the ring with firms 1 and 2.

There are two important caveats in interpreting the results of this exercise. First, the p-values in Table 11 should be treated with caution, because they do not account for the manner in which I selected the various configurations of the ring to test. Rather, they provide qualitative evidence that collusion helps explain the firms' responses to variation in the reserve prices. Second, I cannot eliminate the possibility that alternative modeling assumptions might generate data that are observationally equivalent to collusive equilibria of my model. Thus, the fact that there exists a collusive bidding ring including firms 1 and 2 that fits the data better than the competitive model merely provides suggestive evidence in support of the conclusion that firms 1 and 2 are colluding.

To estimate a 90%-confidence bound on the cost of collusion in the typical auction where suspected colluders are active, I use the estimated valuation distributions that rationalize the bids when firms 1 and 2 are assumed to be colluding to compute the effect of collusion on the expected revenue.[18] In this typical auction, the reserve price is $\$18.18/m^3$ below

18. I cannot avoid modeling the firms' participation decisions for the purposes of estimating the private valuation distributions. Though the private valuations are identified from the winning bids, the data do not include enough wins by each firm to produce reliable estimates. Instead, I use a Kaplan-Meier estimator based on the full vector of bids and participations decisions. I assume that the firm's valuation was censored at the reserve price if it did not bid in a district where it participates at least 5% of the time that it is eligible.

my estimated appraisal. I therefore include bidders 1 and 2 in the set of eligible bidders along with a group of competitors, that, for simplicity, are assumed to be symmetric and to belong to the competitive fringe consisting of firms who did not participate more than 30 times and win more than five licenses. I then solve for the equilibrium bid distributions with and without collusion between bidders 1 and 2 using the numerical methods described in the Appendix A.2. After adding enough competitive fringe bidders to match the empirical price distribution conditional on $r = -18.18$, i.e. $\mathbb{M}_T(\cdot|r = -18.18)$, I find that the expected revenue increases by $\$0.21/m^3$ when firms 1 and 2 bid competitively, which amounts to 1.2% of the average price with collusion.

3.3 Discussion

In first-price auctions, the competitive and collusive models imply different comparative statics because they predict different changes in the distribution of each bidder's highest competing bid. Therefore, a test for collusion may be based upon a test of whether the competitive model rationalizes a bidder's response to exogenous variation in its competing distribution. When a binary instrument creates this identifying variation, as would be the case if the seller randomly "sets aside" a fraction of the auctions for subset of the bidders, I suggest a Kolmogorov-Smirnov type statistic to test the hypothesis that the valuation distributions that competitively rationalize a bidder's bids are equal. When the variation in competition is not binary, I instead suggest a Kendall's τ statistic as a test of the null hypothesis that the competitively rationalizing valuations are independent of the competition. A conditional version of this statistic may be used when the reserve price is binding. Simulation results confirm both of these statistics control the probability of making type I errors and have power to detect collusion even when colluders attempt to disguise their behavior by always bidding according to the ring's optimal strategy.

Of course, any hypothesis test must be interpreted as a joint test of all of the modeling and identification assumptions. Thus, a rejection of the null hypothesis could be a rejection of any of the baseline modeling assumptions or of the exogeneity of the instrument. Nonetheless, my identification strategy can sometimes provide specific evidence against the assumption of perfect competition.

For example, if the reserve price or the set of eligible bidders increases, the competitively rationalizing valuations should be positively correlated with the level of competition if the bidders are colluding, but it will be negatively correlated if the bidders are risk averse. A one-sided test would then control the probability of type I errors even when the bidders' preferences are misspecified.

Furthermore, one can adapt the same procedures used to test the competitive model in order to test whether an IPV model with collusion rationalizes their responses to the observed variation in the level of competition. The other modeling and identification assumptions remain the same, but this second testing procedure reverses the roles of the null and alternative hypotheses. Thus, if the test fails to reject any of the null hypotheses when the bidders in $\mathcal{R}$ are assumed to be colluding but rejects the null hypotheses of those bidders when $\mathcal{R}$ is assumed to be empty, then this would be consistent with the interpretation that the original hypotheses were rejected due to collusion, as opposed to a violation of one of the other assumptions.

Alternatively, one could strengthen the case for collusion by extending the basic identification strategy to more general auction models, such as models in which bidders' private valuations are affiliated. With affiliated valuations, each competitive bidder's inverse strategy function is identified under the null hypothesis that it is not colluding if the two highest bids are observed (Athey and Haile, 2002).[19] Consequently, a test of independence between the competitively rationalizing valuations and an exogenous instrument would control the probability of making a type I error. In practice, estimating these inverse strategies demands more from the data than under the independence assumption because the distri-

19. Under the null hypothesis that bidder i is not colluding, the highest and second-highest bids are both serious bids if one of them belongs to bidder i. Lemma 1 in Athey and Haile (2002) then implies that the inverse strategies are identified.

bution of the highest competing bid must be estimated conditional on the bidder's own bid. A type-symmetry assumption would justify pooling bidders of the same type in order to more precisely estimate their inverse strategy. But this assumption sacrifices the arbitrary asymmetry among bidders. Therefore, it may be necessary to choose between allowing for affiliation in valuations and allowing for unobservable heterogeneity in bidders' valuation distributions. For example, in the above analysis, I argue that valuations are independently distributed conditional on the observable covariates because the data include all the variables that the Ministry of Forests itself uses to estimate the value of the timber licenses. I acknowledge, however, that this assumption could be relaxed if each bidder were observed to win more frequently.

Furthermore, the cost of collusion is generally not identified possible when bidders' private valuations are affiliated. Even if identities of the colluders were known, the possibility of phantom bidding entails that the distribution of a colluder's valuation could only be nonparametrically identified conditional on the event that its valuation is the greatest among all ring members. If the bidders' valuations were symmetrically distributed, this condition would be sufficient because the event in which its valuation is greatest is equivalent to the event in which its bid is greatest in any symmetric, strictly increasing equilibrium.[20] Otherwise, this event is not necessarily equivalent to the event in which its bid is highest in the counterfactual asymmetric equilibrium. Consequently, the joint distribution of bids is insufficient to nonparametrically identify the counterfactual equilibrium in which none of the bidders are colluding.

By contrast, one could account for some degree of correlation among the bids by using losing bids to estimate the distribution of auction-level heterogeneity that bidders observe but the econometrician does not. The difficulty in developing this extension is again that

20. One would then use the differential equation

$$\frac{\partial \sigma(v_i)}{\partial v_i} = (v_i - \sigma(v_i)) \Big/ \frac{\partial \log P\{\max_{j \neq i} V_j \leq v \,|\, V_i = v_i\}}{\partial v}\bigg|_{v=v_i}$$

to solve for the symmetric equilibrium strategy function in the competitive equilibrium.

some of the losing bids could be phantom bids and possibly unrelated to both the bidders' private valuations and the unobserved heterogeneity. As a result, the procedure would have to simultaneously test for collusion and estimate this latent distribution. On the other hand, if some bidders are known to be competitive *a priori*, the winning bid and the other competitive bids can be viewed as independent measures of the unobserved heterogeneity. The deconvolution methods in Krasnokutskaya (2011) can then be used to estimate the latent distribution and the distribution of each bidder's winning bids. Identification of the colluders would then proceed as in the baseline case.

In the application to British Columbia's timber auctions, however, the reserve price is often binding, with the result that only one bid is observed for many of the auctions. And, unlike the automobile auctions studied in Roberts (2013), the reserve price is an explicit function of observable variables and does reflect any heterogeneity the econometrician does not already observe. Thus, to the best of my knowledge, my identification strategy employs a minimal set of assumptions to nonparametrically estimate the cost of collusion in these auctions.

REFERENCES

Aryal, G. and Gabrielli, M. F. (2013). Testing for Collusion in Asymmetric First-Price Auctions. *International Journal of Industrial Organization*, 31:26–35.

Asker, J. (2010). A Study of the Internal Organization of a Bidding Cartel. *American Economic Review*, 100(3):724–762.

Athey, S., Coey, D., and Levin, J. (2013). Set-Asides and Subsidies in Auctions. *American Economic Journal: Microeconomics*, 5(1):1–27.

Athey, S. and Cramton, P. (2005). Competitive Auction Markets in British Columbia. MDI Report for British Columbia Ministry of Forests.

Athey, S., Cramton, P., and Ingraham, A. (2002). Auction-Based Timber Pricing and Complementary Market Reforms in British Columbia. MDI Report for British Columbia Ministry of Forests.

Athey, S. and Haile, P. A. (2002). Identification of Standard Auction Models. *Econometrica*, 70(6):2107–2140.

Athey, S. and Haile, P. A. (2007). Nonparametric Approaches to Auctions. *Handbook of Econometrics*, 6(PART A):3847–3965.

Athey, S., Levin, J., and Seira, E. (2011). Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions. *Quarterly Journal of Economics*, 126(1):207–257.

Bajari, P. and Hortaçsu, A. (2005). Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data. *Journal of Political Economy*, 113(4):703–741.

Bajari, P. and Ye, L. (2003). Deciding Between Competition and Collusion. *The Review of Economics and Statistics*, 85(4):971–989.

Baldwin, L. H., Marshall, R. C., and Richard, J. F. (1997). Bidder Collusion at Forest Service Timber Sales. *Journal of Political Economy*, 105(4):657–699.

Berman, S. M. (1963). Note on Extreme Values, Competing Risks and Semi-Markov Processes. *The Annals of Mathematical Statistics*, 34(3):1104–1106.

Brannman, L. and Froeb, L. M. (2000). Mergers, Cartels, Set-Asides, and Bidding Preferences in Asymmetric Oral Auctions. *Review of Economics and Statistics*, 82(2):283–290.

Canada. B.C. Ministry of Forests (1999). *Interior Appraisal Manual*.

Chassang, S. and Ortner, J. (2017). Collusion in Auctions with Constrained Bids: Theory and Evidence from Public Procurement.

Conley, T. G. and Decarolis, F. (2016). Detecting Bidders Groups in Collusive Auctions. *American Economic Journal: Microeconomics*, 8(2):1–38.

Davydov, Y. A., Lifshits, M. A., and Smorodina, N. V. (1998). *Local Properties of Distributions of Stochastic Functionals*, volume 173 of *Mathematical Monographs*. American Mathematical Society.

Fibich, G. and Gavish, N. (2011). Numerical Simulations of Asymmetric First-Price Auctions. *Games and Economic Behavior*, 73(2):479–495.

Guerre, E., Perrigne, I., and Vuong, Q. (2000). Optimal Nonparametric Estimation of First-Price Auctions. *Econometrica*, 68(3):525–574.

Guerre, E., Perrigne, I., and Vuong, Q. (2009). Nonparametric Identification of Risk Aversion in First-Price Auctions Under Exclusion Restrictions. *Econometrica*, 77(4):1193–1227.

Haile, P. A., Hong, H., and Shum, M. (2003). Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions.

Hickman, B. R. and Hubbard, T. P. (2015). Replacing Sample Trimming with Boundary Correction in Nonparametric Estimation of First-Price Auctions. *Journal of Applied Econometrics*, 30(5):739–762.

Hickman, B. R., Hubbard, T. P., and Paarsch, H. J. (2016). Identification and Estimation of a Bidding Model for Electronic Auctions.

Hill, J. B. and Shneyerov, A. (2013). Are There Common Values in First-Price Auctions? A Tail-Index Nonparametric Test. *Journal of Econometrics*, 174(2):144–164.

Jofre-Bonet, M. and Pesendorfer, M. (2003). Estimation of a Dynamic Auction Game. *Econometrica*, 71(5):1443–1489.

Kaplan, T. R. and Zamir, S. (2012). Asymmetric first-price auctions with uniform distributions: Analytic solutions to the general case. *Economic Theory*, 50(2):269–302.

Karunamuni, R. J. and Zhang, S. (2008). Some Improvements on a Boundary Corrected Kernel Density Estimator. *Statistics and Probability Letters*, 78(5):499–507.

Kawai, K. and Nakabayashi, J. (2014). Detecting Large-Scale Collusion in Procurement Auctions.

Krasnokutskaya, E. (2011). Identification and Estimation of Auction Models with Unobserved Heterogeneity. *The Review of Economic Studies*, 78(1):293–327.

Krasnokutskaya, E. and Seim, K. (2011). Bid Preference Programs and Participation in Highway Procurement Auctions. *American Economic Review*, 101(6):2653–2686.

Lebrun, B. (2002). Continuity of the First Price Auction Nash Equilibrium Correspondence. *Economic Theory*, 20(3):435–453.

Lebrun, B. (2006). Uniqueness of the Equilibrium in First-Price Auctions. *Games and Economic Behavior*, 55(1):131–151.

Liang, H., Liu, X., Li, R., and Tsai, C. L. (2010). Estimation and testing for partially linear single-index models. *Annals of Statistics*, 38(6):3811–3836.

List, J. A., Millimet, D. L., and Price, M. K. (2007). Inferring Treatment Status When Treatment Assignment is Unknown: Detecting Collusion in Timber Auctions.

Marmer, V., Shneyerov, A., and Kaplan, U. (2016). Identifying Collusion in English Auctions.

Marshall, R. C. and Marx, L. M. (2007). Bidder collusion. *Journal of Economic Theory*, 133(1):374–402.

Martin, E. C. and Betensky, R. A. (2005). Testing Quasi-Independence of Failure and Truncation Times via Conditional Kendall's Tau. *Journal of the American Statistical Association*, 100(470):484–492.

McAfee, P. R. and McMillan, J. (1992). Bidding Rings. 82(3):579–599.

Paarsch, H. (1997). Deriving an Estimate of the Optimal Reserve Price: An Application to British Columbian Timber Sales. *Journal of Econometrics*, 78(2):333–357.

Pesendorfer, M. (2000). A Study of Collusion in First-Price Auctions. *Review of Economic Studies*, 67(3):381–411.

Politis, D. N. and Romano, J. P. (1994). Large Sample Confidence Regions Based on Subsamples under Minimal Assumptions. *The Annals of Statistics*, 22(4):2031–2050.

Porter, R. H. (1983). A Study of Cartel Stability: The Joint Executive Committee, 1880–1886. *The Bell Journal of Economics*, 14(2):301–314.

Porter, R. H. (2005). Detecting Collusion. *Review of Industrial Organization*, 26(2):147–167.

Porter, R. H. and Zona, J. D. (1993). Detection of Bid Rigging in Procurement Auctions. *Journal of Political Economy*, 101(3):518.

Porter, R. H. and Zona, J. D. (1999). Ohio School Milk Markets: An Analysis of Bidding. *The RAND Journal of Economics*, 30(2):263.

Price, M. K. (2008). Using the Spatial Distribution of Bidders to Detect Collusion in the Marketplace: Evidence from Timber Auctions. *Journal of Regional Science*, 48(2):399–417.

Roberts, J. W. (2013). Unobserved Heterogeneity and Reserve Prices in Auctions. *RAND Journal of Economics*, 44(4):712–732.

Romano, J. P. and Wolf, M. (2010). Balanced control of generalized error rates. *The Annals of Statistics*, 38(1):598–633.

Stone, C. J. (1980). Optimal Rates of Convergence for Nonparametric Estimators. *The Annals of Statistics*, 8(6):1348–1360.

Stute, W. (1986). Conditional Empirical Processes. *Annals of Statistics*, 14(2):638–647.

Tsai, W.-Y. (1990). Testing the assumption of independence of truncation time and failure time. *Biometrika*, 77(1):169–177.

van der Vaart, A. W. (1998). *Asymptotic Statistics.* Cambridge University Press, New York, NY.

APPENDIX A

PROOFS AND SUPPORTING MATERIALS

A.1 Propositions and Proofs

Proposition 2. *Suppose $r > \underline{v}_i$ and σ_i is strictly increasing for all i. The density of each i's highest competing bid is unbounded near the reserve price. In particular, the rate at which the density diverges is given by $\frac{\partial S_i(b)}{\partial b} = O(1/\sqrt{b-r})$.*

Proof. The system of differential equations that defines the Bayes-Nash equilibrium bid distributions may be written as

$$\frac{1}{F_i^{-1}(S_i(b)) - b} = \sum_{j \neq i} \frac{\partial \log S_j(b)}{\partial b} = \frac{g_i(b)}{G_i(b)} \quad \text{for all } i\,, \tag{A.1}$$

with the initial conditions $F_i(r) = S_i(r)$. As b approaches r from above, the left-hand sides of the differential equations tend toward infinity. Because $G_i(r) > 0$, this implies that g_i must be unbounded near the reserve price. Notice that it is important that each bidder's strategy is strictly increasing. Otherwise, if there is a bidder i whose strategy is constant in a neighborhood of the reserve, then $G_i(b)/g_i(b)$ converges to a nonzero limit as b approaches r. Hence, g_i will be bounded near the reserve, while g_j will be unbounded for all $j \neq i$.

To demonstrate the rate at which the density tends toward infinity as b approaches r, I use the change of variables suggested by Guerre et al. (2000) for the symmetric IPV case. Let $\tilde{B}_i = \sqrt{B_i - r}$ and let $\tilde{S}_i$ denote its distribution. Define $\tilde{G}_i$ and $\tilde{g}_i$ analogously. Then, the differential system may be rewritten as

$$\frac{2\tilde{b}}{F_i^{-1}(\tilde{S}_i(\tilde{b})) - \tilde{b}^2 - r} = \sum_{j \neq i} \frac{\partial \log \tilde{S}_i(\tilde{b})}{\partial \tilde{b}} = \frac{\partial \log \tilde{G}_i(\tilde{b})}{\partial \tilde{b}} \quad \text{for all } i\,, \tag{A.2}$$

with $F_i(r) = \tilde{S}_i(0)$. Note that $\tilde{G}_i(0) = G_i(r) > 0$ and $\tilde{g}_i(\tilde{b})/\tilde{G}_i(\tilde{b}) = 2\sqrt{b-r}\, g_i(b)/G_i(b)$. Therefore, to prove $g_i(b) = O(1/\sqrt{b-r})$, it suffices to show that $\tilde{g}_i$ is bounded near zero.

Applying L'Hospital's rule applied to the ratio on the left-hand side of the transformed system, $\tilde{g}_i$ will be bounded if $\frac{\partial F_i^{-1}(\tilde{S}_i(\tilde{b}))}{\partial \tilde{b}}$ is nonzero and finite at $\tilde{b} = 0$ (or, more precisely, if this derivative has a nonzero right-limit as $\tilde{b}$ approaches zero.) Because $\frac{\partial F_i^{-1}(t)}{\partial t}|_{\tilde{S}(0)} = 1/f_i(r)$ is positive and finite, the density of $\tilde{B}_i$ will determine the behavior of the ratio on the left-hand side of (A.2), and thereby determine the behavior of $\tilde{g}_i$ at $\tilde{b} = 0$.

Suppose by way of contradiction that the density of $\tilde{B}_i$ is unbounded for some i. Then the transformed system implies that the density of $\tilde{B}_j$ tends toward zero for all $j \neq i$ as $\tilde{b}$ approaches zero. Similarly, if the density of $\tilde{B}_j$ tends to zero for any bidder j, then the density of $\tilde{B}_i$ is unbounded for some $i \neq j$. Therefore, if the densities of $\tilde{B}_i$ at $\tilde{b} = 0$ are not all positive and finite, then one of the limits is unbounded and the rest are zero. Thus, we may assume without loss of generality that $\tilde{B}_i$ is the only random bid whose density is unbounded near zero. Also note that the density of $\tilde{B}_j$ tends toward zero at the same rate for all j.

In other words, the density of bidder i's untransformed bids diverges at a faster than quadratic rate, while each bidder j's untransformed bids diverge at the same slower than quadratic rate. This disparity in the rates of divergence leads to a contradiction because it means that bidder i bids much less aggressively than any of its competitors when it has a valuation that is just barely greater than the reserve price. To see this, fix any j with a bounded density and write the ratio between the density of $\tilde{B}_i$ to that of $\tilde{B}_j$ as

$$\frac{\frac{\partial S_i}{\partial b}}{\frac{\partial S_j}{\partial b}} = \frac{\frac{1}{N-1}\sum_{k\in\mathcal{N}} \frac{F_j^{-1}(S_j(b)) - b}{F_k^{-1}(S_k(b)) - b} - \frac{F_j^{-1}(S_j(b)) - b}{F_i^{-1}(S_i(b)) - b}}{\frac{1}{N-1}\sum_{k\in\mathcal{N}} \frac{F_j^{-1}(S_j(b)) - b}{F_k^{-1}(S_k(b)) - b} - 1}.$$

The right-hand side converges to a non-zero limit because $(F_j^{-1}(S_j(b)) - b)/(F_i^{-1}(S_i(b)) - b)$ tends toward zero while, for all $k \neq i$, $(F_j^{-1}(S_j(b)) - b)/(F_k^{-1}(S_k(b)) - b)$ tends to a positive constant as b tends toward r. But the left-hand side of the equation was assumed to diverge.

Thus, the density of B_i cannot diverge at a faster rate than the density for all $j \neq i$. Hence, $\tilde{B}_i$ has a bounded density for all bidders $i \in \mathcal{N}$.

Hill and Shneyerov (2013) provide an alternative argument for the case in which bidders are type-symmetric—i.e. bidders of the same type have valuations drawn from identical marginal distributions—and there are at least two bidders of every type. To relate their result to statement of Proposition 2, notice that type-symmetric valuation distributions imply that bidders will adopt type-symmetric bidding strategies. Because Proposition 1 says that at most one of the serious bidding strategies is not strictly increasing, all of the bidders must use strictly increasing strategies; otherwise, at least two bidders' strategies would be constant near the reserve price. Thus, the type-symmetry assumption is a sufficient condition for the strategies to be strictly increasing for all bidders. $\square$

proof of Theorem 3. Assume $i \notin \mathcal{R}$. Let z and z' denote distinct realizations of Z. By $IA.3$, the lemma, and equation 2.11,

$$\begin{aligned}
D_{iT} &= \|\mathbb{F}_{iT,\emptyset}(v|Z=z) - F_i(v) + F_i(v) - \mathbb{F}_{iT}(v|Z=z')\|_\infty \\
&\leq \|\mathbb{F}_{iT}(v|Z=z) - F_i(v)\|_\infty + \|\mathbb{F}_{iT}(v|Z=z') - F_i(v)\|_\infty \\
&= O_p\left((T/\log T)^{-2/5}\right) + O_p\left((T/\log T)^{-2/5}\right) \\
&= O_p\left((T/\log T)^{-2/5}\right).
\end{aligned}$$

Note that the stochastic order of the statistic is smaller than in equation 2.11 because the reserve price is also assumed to be discrete.

Otherwise, if $i \in \mathcal{R}$, then Theorem 2 implies that $F_{i,\emptyset}(\cdot|Z=z) \neq F_{i,\emptyset}(\cdot|Z=z')$. Hence, D_{iT} converges in probability to some nonzero constant. This proves (i).

To prove (ii), note that $(T/\log T)^{2/5} D_{iT}$ can be viewed as the supnorm of a non-degenerate Gaussian process if $i \notin \mathcal{R}$. Proposition 12.1 in Davydov et al. (1998) then implies that the limit of $(T/\log T)^{2/5} D_{iT}$ is absolutely continuous. Otherwise, part (i) implies that it diverges whenever $i \in \mathcal{R}$. $\square$

A.2 Numerical Solution to Asymmetric First-Price Auctions with Reserve Prices

As noted in Proposition 2, the system of differential equations that characterizes the equilibrium is indeterminate near the minimum bid. Consequently, the "forward" methods of solving initial value problems cannot be applied because there is no way to evaluate the system at the left boundary. As an alternative, one could start with an initial guess of the maximum bid and use backward shooting algorithms to find the inverse strategies that satisfy the initial value conditions. However, Fibich and Gavish (2011) show that these methods become increasingly unstable when there many different types of bidders. Instead, they advocate converting the initial value problem into a boundary value problem by choosing a bidder, say bidder n, as a benchmark and writing the equilibrium bids and the other bidders' strategies as a function of bidder n's valuation. Fixed-point iterations can then be used to find a solution.

While this iterative method appears to work well in several interesting cases, there is no theory to guarantee that the sequence of iterations will converge. And, in practice, I find greater success by making two modification to their solution strategy. First, rather than discretizing the solution and using finite difference approximations, I seek an approximate solution in the space of cubic splines. That is, I represent the solution as a linear combination of basis functions and solve for the coefficients that minimize the residuals in the differential equations.

Second, I solve the system (A.2) instead of the equivalent system of equations that characterize the inverse bidding strategies. This is done partly for convenience because solving (A.2) only requires evaluating the the quantile function for bidder i, which simplifies the expression for the gradient of the objective function. Though, when applied using estimated valuation distributions, a practical advantage of this formulation is that F_i^{-1} can be non-parametrically estimated at a faster rate than the density f_i. For the purposes of computing an analytic gradient, however, it is still useful to evaluate the derivative of F_i^{-1}. To that end, I use a monotonic cubic spline to interpolate the estimated quantile function.

Thus, I approximate the solution to the boundary value problem consisting of the system of differential equations

$$\frac{\partial \tilde{S}_i}{\partial \tilde{S}_n} = \frac{\tilde{S}_i(w)}{\tilde{S}_n} \frac{\sum_j \frac{1}{F_j^{-1}(\tilde{S}_j(\tilde{S}_n)) - w(\tilde{S}_n)^2 - r} - \frac{N-1}{F_i^{-1}(\tilde{S}_i(\tilde{S}_n)) - w(\tilde{S}_n)^2 - r}}{\sum_j \frac{1}{F_j^{-1}(\tilde{S}_j(\tilde{S}_n)) - w(\tilde{S}_n)^2 - r} - \frac{N-1}{F_n^{-1}(\tilde{S}_n) - w(\tilde{S}_n)^2 - r}}$$

$$\frac{\partial w}{\partial \tilde{S}_n} = \frac{N-1}{2 \cdot w(\tilde{S}_n) \cdot \tilde{S}_n} \frac{1}{\sum_j \frac{1}{F_j^{-1}(\tilde{S}_j(\tilde{S}_n)) - w(\tilde{S}_n)^2 - r} - \frac{N-1}{F_n^{-1}(\tilde{S}_n) - w(\tilde{S}_n)^2 - r}}$$

with the boundary conditions

$$\tilde{S}_i(\underline{\tilde{S}}_n) = F_i(r)$$

$$\tilde{S}(1) = 1$$

$$w(\underline{\tilde{S}}_n) = 0.$$

by finding the coefficients a_{il} and a_{0l} such that $\tilde{S}_i(\tilde{S}_n) = \sum_l a_{il}\, \beta_l(\tilde{S}_n)$ and $w(\tilde{S}_n) = \sum_l a_{0l}\, \beta_l(\tilde{S}_n)$ minimize the error in the above system, where each $\beta = \{\beta_l : l = 1, \ldots, L\}$ is a basis of spline functions defined on the interval $[F_n(r), 1]$. In order to impose the boundary conditions, the knot vector used to construct the basis splines can be chosen to include Ord copies of $F_n(r)$ and 1, where Ord is the order of the spline functions. The boundary conditions are then satisfied whenever $a_{01} = 0$, $a_{i1} = F_i(r)$, and $a_{iL} = 1$. In addition, I impose monotonicity in the solution by restricting $a_{i\cdot}$ to be an increasing sequence for each $i = 0, \ldots, n-1$.

To be precise, let $a = (a_{il} : i = 0, \ldots, n-1, j = 1, \ldots, L)$ and $E_i(a, s; \beta)$ for $i = 0, \ldots, n-1$ denote the difference between the left- and right-hand sides of the above equations evaluated at $\tilde{S}_n = s$ in some arbitrarily fine grid, $s = s_1, \ldots, s_d$. The approximation problem is then given by

$$\begin{aligned}
\min_{a} \sum_{i} \|E_i(a,\cdot\,;\beta)\| \qquad & \text{subject to} \\
F_i(r) = a_{i1} < \cdots < a_{iL} = 1 \qquad & \text{for } i = 1,\ldots,n-1 \\
0 = a_{01} < \cdots < a_{0L}\,, \qquad &
\end{aligned}$$

where $\|\cdot\|$ is some norm on $\mathbb{R}^d$. The Euclidean norm is an attractive option because it is differentiable, but the approximate solution might perform badly in a small region of the domain even when the average squared error is small. This does not appear to be an issue in simple cases, but I have found that the supnorm performs at least as well and is not vulnerable to this criticism. To preserve differentiability, I follow Hickman et al. (2016) and formulate the problem as

$$\begin{aligned}
\min_{a,\epsilon} \epsilon \qquad & \text{subject to} \\
E_i(a,s;\beta) < \epsilon \qquad & \text{for } i = 0,\ldots,n-1 \text{ and } s = s_1,\ldots,s_d \\
-E_i(a,s;\beta) < \epsilon \qquad & \text{for } i = 0,\ldots,n-1 \text{ and } s = s_1,\ldots,s_d \\
F_i(r) = a_{i1} < \cdots < a_{iL} = 1 \qquad & \text{for } i = 1,\ldots,n-1 \\
0 = a_{01} < \cdots < a_{0L} \qquad &
\end{aligned}$$

Because the number of inequality constraints grows linearly with d, the grid cannot be arbitrarily fine. In general, d should be at least as large as the number of basis spline functions and could be much greater. In cases where the analytic solution has been derived by Kaplan and Zamir (2012), I find that $d = 50$ already provides a good approximation with 12 cubic basis splines.

Figure 9 shows the approximate and the analytic equilibrium bid functions for an auction with two bidders and a reserve price of 0.3. The first bidder's valuations are uniform on

[0, 1] and the second benchmark bidder's valuations are uniform on [0.2, 0.8]. As depicted in Figure 9, the spline function provides a uniformly good approximation to the equilibrium; the maximum approximation error is 5.6×10^{-3}.

This figure also illustrates the difficulty in approximating the unbounded bid density near the reserve price. If, for example, Chebyshev polynomials in the bids were used to approximate the solution to the untransformed system (A.1), a high degree polynomial would be required in order to simultaneously approximate the steepness near the reserve price and the linearity near the maximum bid. The basis splines that I selected avoid this issue by solving the system in terms of the square-root of the bid minus the reserve price, as opposed to the bid, itself. Basis splines are also more flexible because they are defined piecewise on the partition of the domain created by the knot vector. Thus, in contrast to globally defined polynomials, they can accommodate curvature in the bid distributions in some regions of the domain without affecting the fit in others.

Note that the above system of equations will not be valid if the upper extremity to the support of bidder n's bids is less than the other bidders'. To avoid this situation, the benchmark bidder n can be chosen to be the bidder with the highest upper extremity to the support of its valuations. In this case, the differential equation for bidder i must be multiplied by an indicator for whether S_i is less than one. And, because the left-hand side of bidder i's differential equation might not approach zero as S_i approaches one, the knot vector for the basis spline functions should be chosen so as to allow for a discontinuous derivative at each bidder's maximum valuation.

Lastly, this solution method does not address the possibility that one of the bidders' strategy functions is constant near the reserve price. In this case the lower boundary condition for one of the n bidders must be removed or replaced with $S_i(r) = F_i\Big(r + \frac{G_i}{g_i}(r)\Big)$. This does not sacrifice uniqueness of the solution to the system of differential equations (Lebrun, 2006), but the above change of variables may no longer be appropriate. Further research is needed to determine a reliable solution method in this case.

A.3 Inter-auction Dynamics

To the extent that firms weigh dynamic considerations in their bidding decisions, firms' valuations should be reinterpreted as the contemporaneous payoff from winning minus the opportunity cost of committing their resources to the job. Their valuations should therefore depend on future reserve prices through their opportunity cost, but not on the current reserve price. If reserve prices are not serially correlated, they will therefore be uncorrelated with firms' valuations, and a test of collusion could still be based on a test of independence.

To determine whether serial correlation is likely to bias the test results, Figures 7 and 8 plot the time series of the screening levels—i.e. the reserve prices homogenized by subtracting the appraised value of the license. Figure 7 does not reveal any serial correlation in the homogenized reserve prices conditional on the Ministry's reserve pricing policy when l use the preferred appraisal method in Subsection 3.2.3. Before and after 1999, however, there is a slight decrease in the mean and a significant reduction in the variance of the homogenized reserve price. I next recreate the plot using the Ministry's appraised value in Figure 8. This time series reveals a noticeable year-to-year increase in the homogenized reserve price, and a stark decrease after the policy change in 1999. After controlling for year and appraisal method, however, the time series in Figure 10 does not appear to have any serial correlation. Therefore, if the firms' relevant planning horizon is less than one-year, I argue that the valuations should be independent of the contemporaneous reserve price.

Apart from their interpretation, the valuations in a dynamic auction setting are difficult to estimate because the optimal strategies are functions of the state of the market. Because I cannot observe the all of the state variables, the inverse strategy functions estimated in this dissertation would be better understood as a mixture of state-dependent inverse strategies:

$$\frac{M(b|r,\mathcal{N})}{\frac{\partial M_{-i}(v|r,\mathcal{N})}{\partial b}} = \frac{\int M(b|r,\mathcal{N},y)\,dQ(y)}{\int \frac{\partial M_{-i}(b|r,\mathcal{N},y)}{\partial b}\,dQ(y)} = \int \frac{M(b|r,\mathcal{N},y)}{\frac{\partial M_{-i}(b|r,\mathcal{N},y)}{\partial b}} \frac{\frac{\partial M_{-i}(b|r,\mathcal{N},y)}{\partial b}}{\int \frac{\partial M_{-i}(b|r,\mathcal{N},y)}{\partial b}\,dQ(y)}\,dQ(y)\,,$$

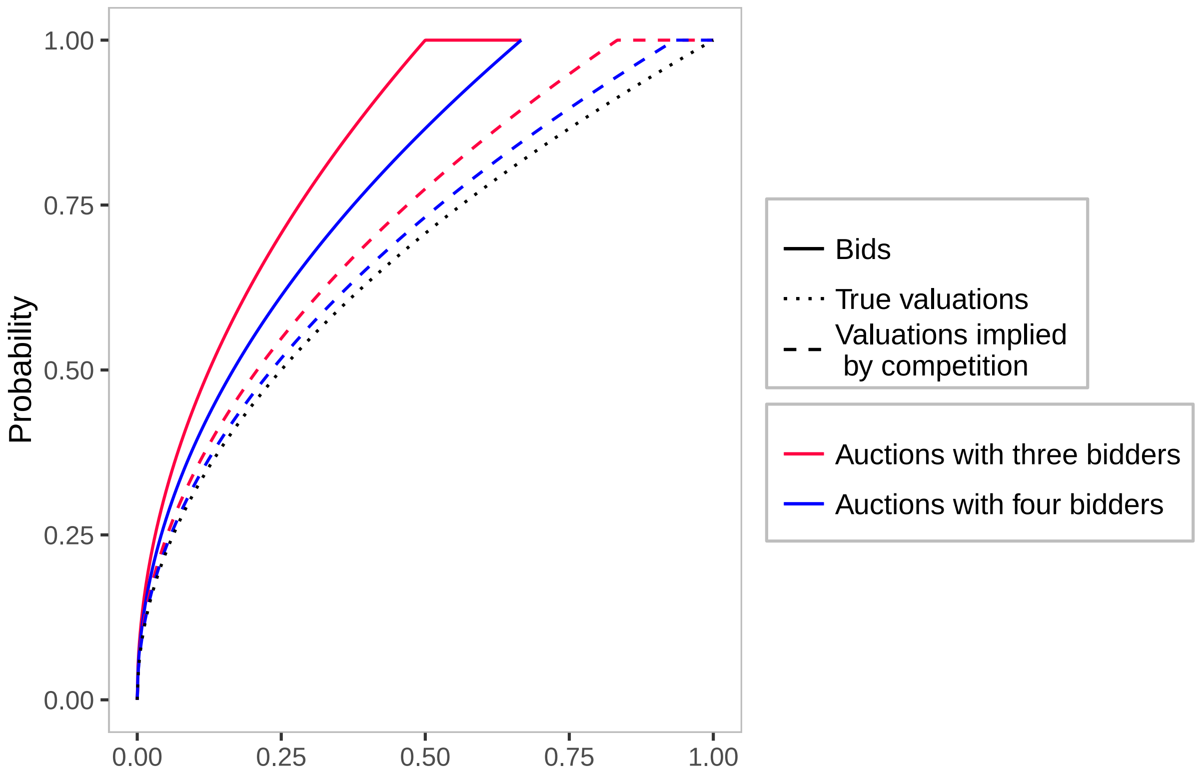
where y is a vector of state variables and Q is its (stationary) distribution.

If the current reserve price is also independent of the difference between the true state-dependent strategies and the estimated strategies, the test for collusion would still be valid. Unfortunately, there is no way to verify this assumption in the data. Indeed, because the expected profit for a firm with a given valuation is greater when the reserve price is lower, firms might plan to have excess capacity available for an upcoming auction with a relatively low reserve price. In which case, the assumption would clearly be false.

If this story is true, however, the already announced reserve prices of future auctions should be correlated with the firms' willingness to pay at the current auction. To test this prediction, I estimate the relationship between bids and covariates using a partially linear-single index model as explained in Subsection 3.2.3. When, in addition to the contemporaneous characteristics, I include the average reserve price for the next five auctions in the same region, the coefficient on this term is estimated to be 0.008 with a standard error of 0.011. I interpret this as suggestive evidence that firms do not take these dynamic considerations into account.

APPENDIX B
FIGURES AND TABLES

Figure 1: *A collusive bidder's bid and valuation distributions before and after the exogenous entry of another bidder.* The level of competition affects the valuation distribution that competitively rationalizes a colluder's bids. In each case, the competitively rationalizing valuation distribution lies above the true valuation distribution because the competitive model overestimates the colluder's competition and therefore underestimates how much the colluder bids below its valuations. Because the arrival of another competitor lessens the difference between the competitive and collusive models' predictions, the valuation distribution implied by competition shifts closer to the truth.



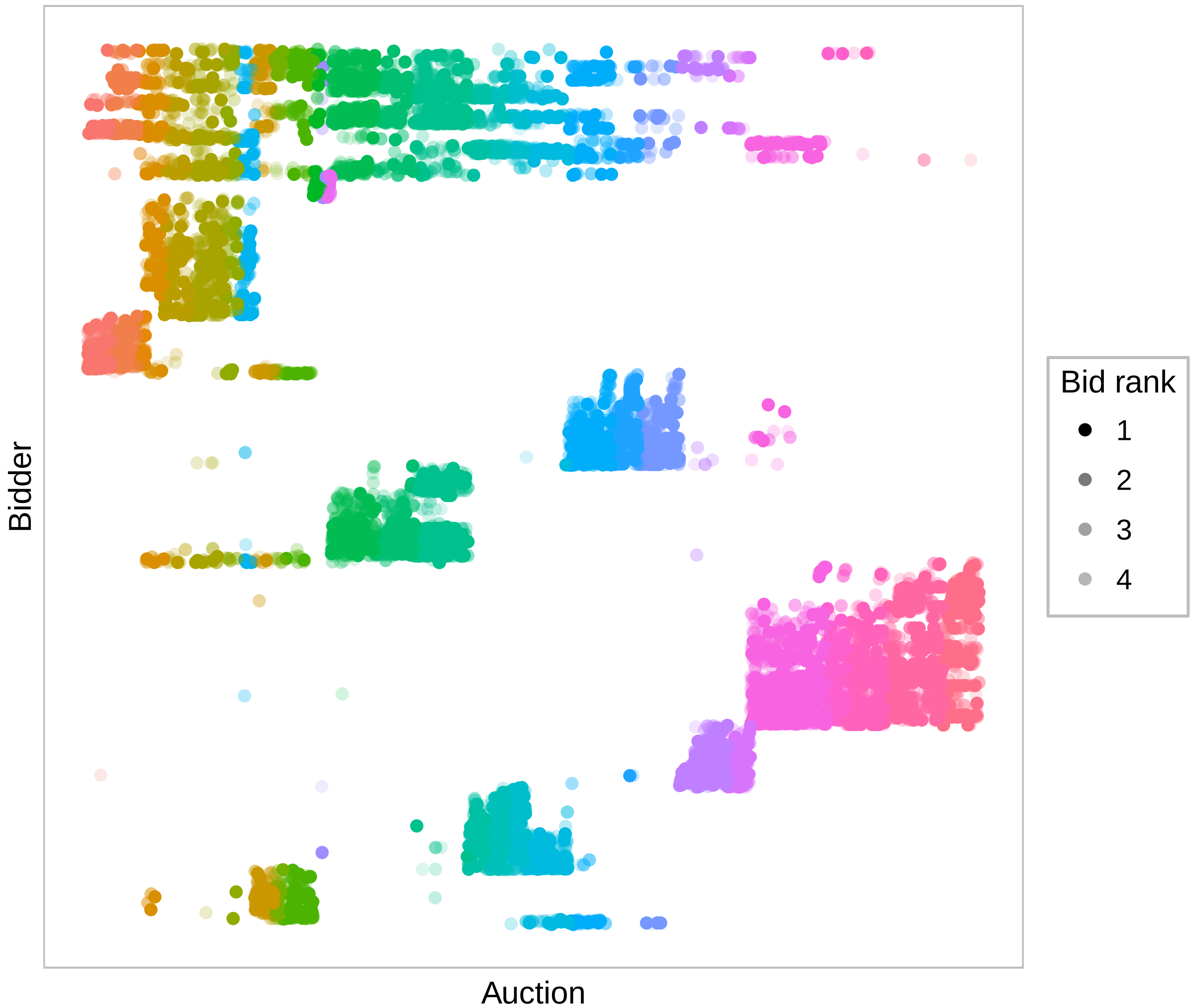


Figure 2: *Bidders' participation patterns across the districts defined by the Ministry of Forests.* Each color corresponds to one of the 31 geographic districts. A dot at auction t and bidder i indicates that bidder i participated in auction t. The opacity of the dot at that point indicates whether i's bid was highest, second-highest, third-highest, etc. at the auction.

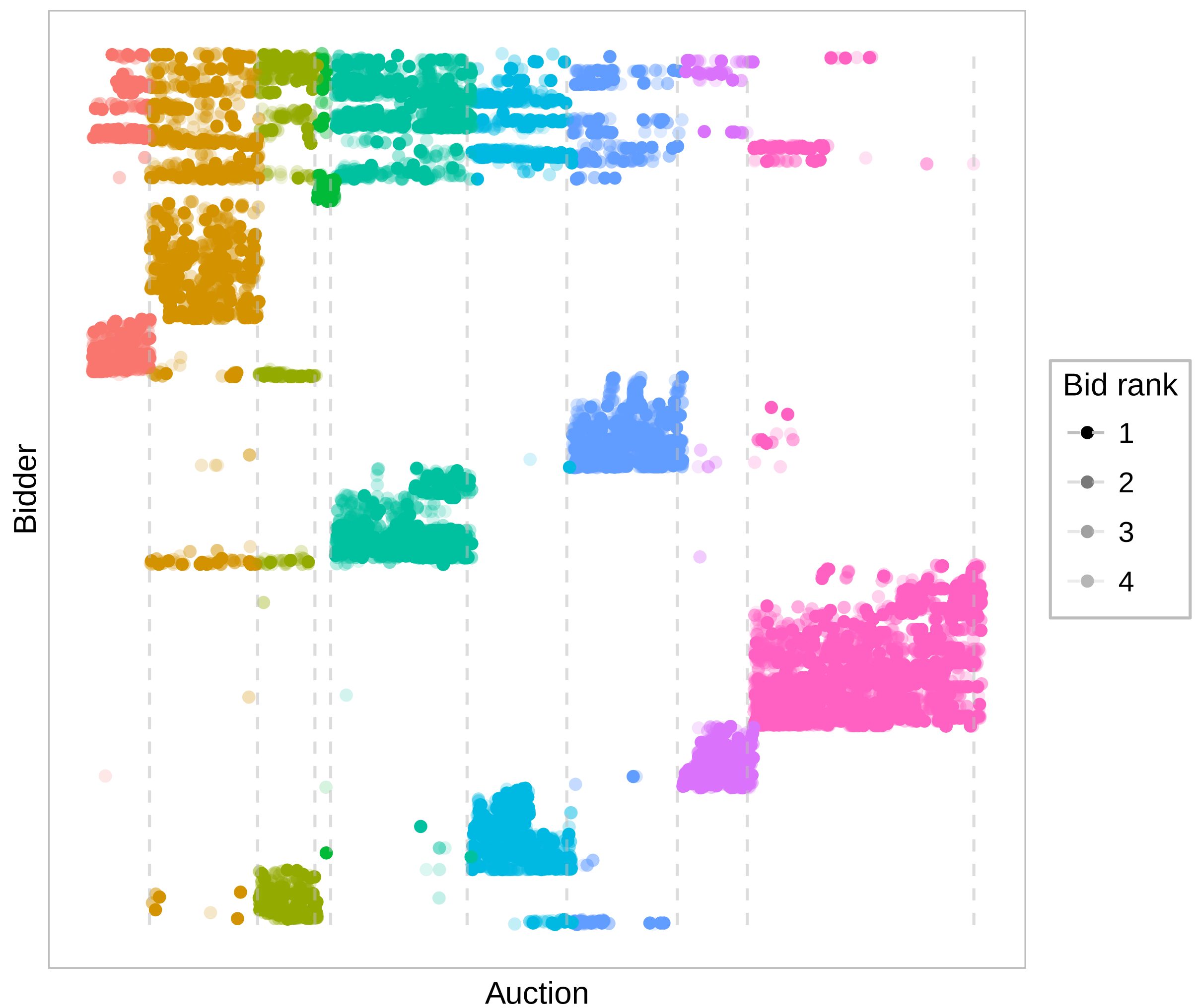


Figure 3: *Participation pattern by market.* Markets were formed by iteratively combining districts for which the same firms have similar propensities to participate.

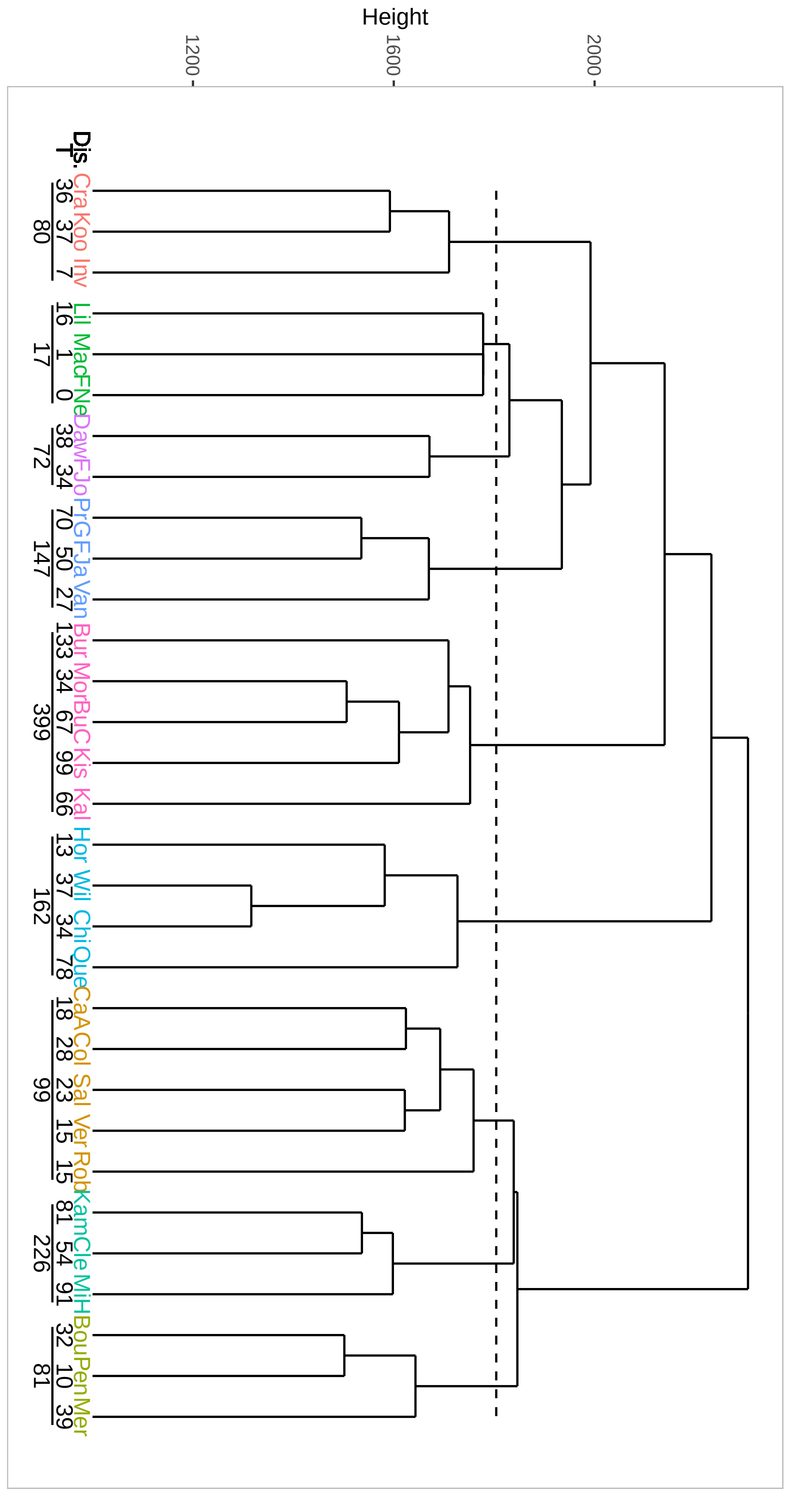


Figure 4: *A dendrogram representing the output of a hierarchical clustering algorithm.* Initially, each district is assigned to its own cluster. The two most similar districts, Williams Lake (Wil) and Chilcotin (Chi), are the first to be merged. At each juncture, the height of the dendrogram represents the dissimilarity between the two clusters that are merged.

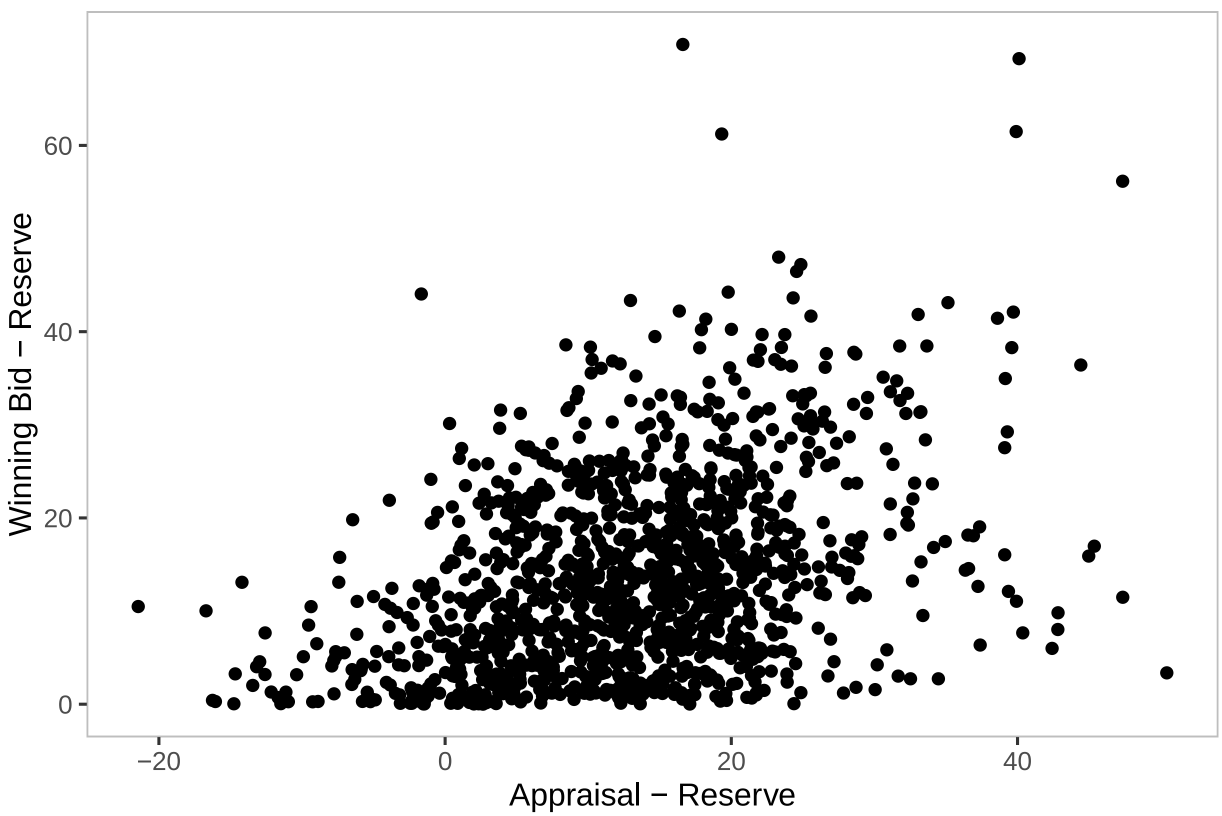


Figure 5: *Homogenized prices as a function of the homogenized reserve prices.* The winning bonus bid is positively correlated with the estimated value of the timber tract. In addition, the variance in the winning bonus bid is greater when the license is more valuable relative to the reserve price because bidders are more likely to enter with a wider range of private valuations.

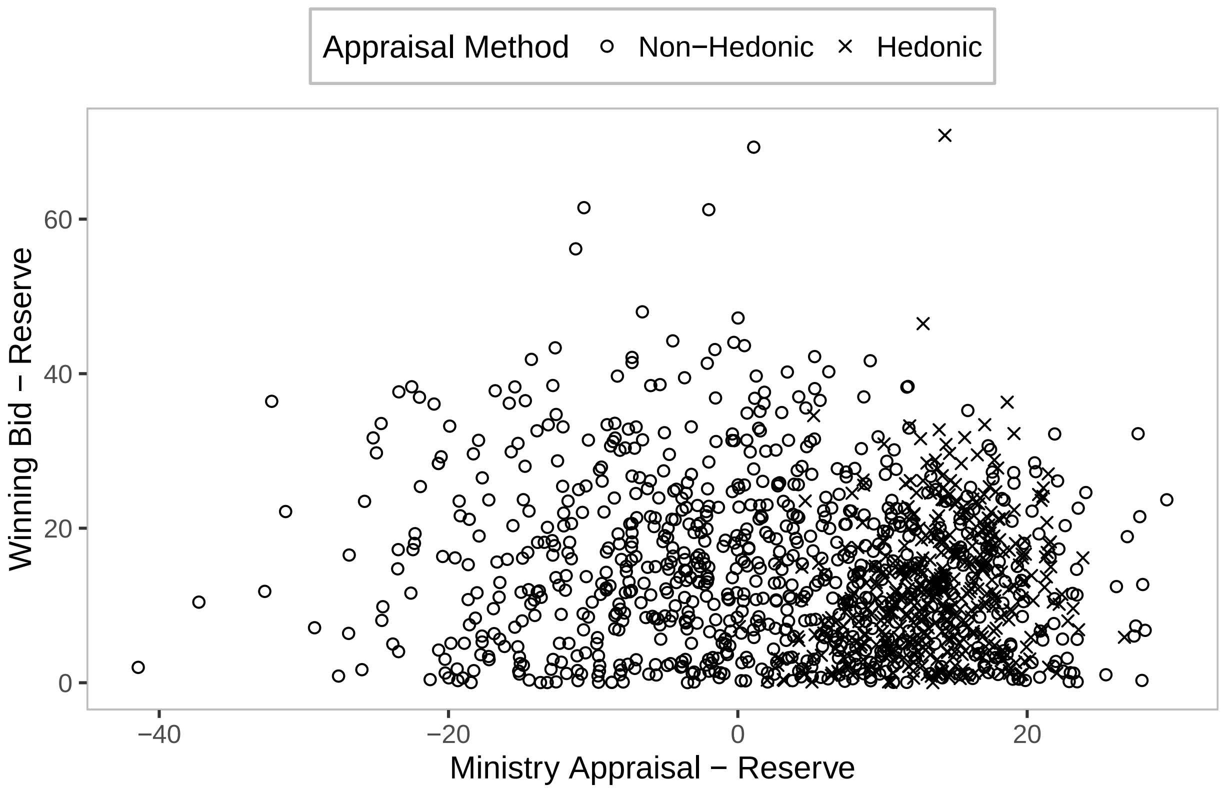


Figure 6: *Prices homogenized using the Ministry's appraisal.* The Ministry of Forests' appraised value of the license does not predict the winning bonus bid as well as the estimates from the partially linear single-index model.

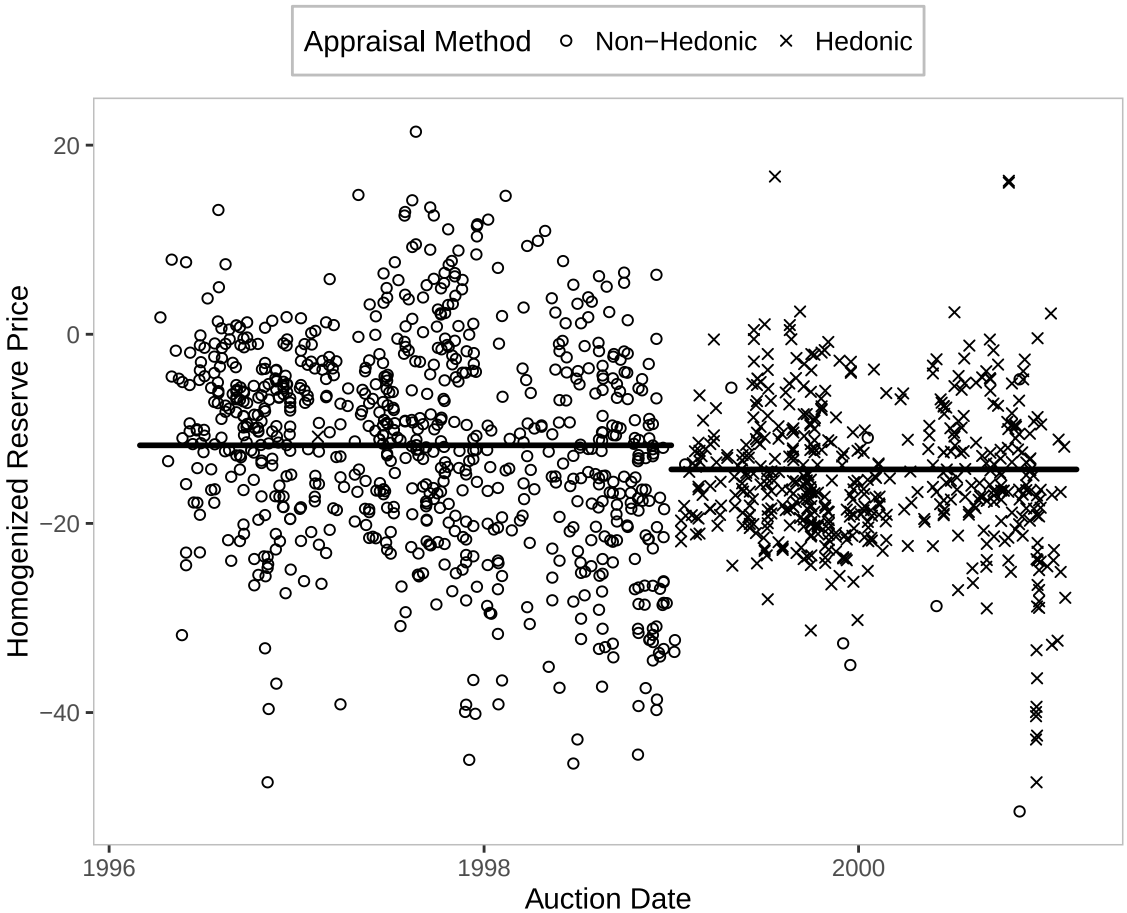


Figure 7: *The time series of homogenized reserve prices.* Conditional on the Ministry's appraisal method, the homogenized reserve prices are not serially correlated. After the policy change in 1999, however, the mean and variance of the homogenized reserve prices are noticeably lower. The outliers at the end of 2000 are explained by the fact that the Ministry experimented with setting the reserve price at 50% of their estimated appraisal as opposed 70%.

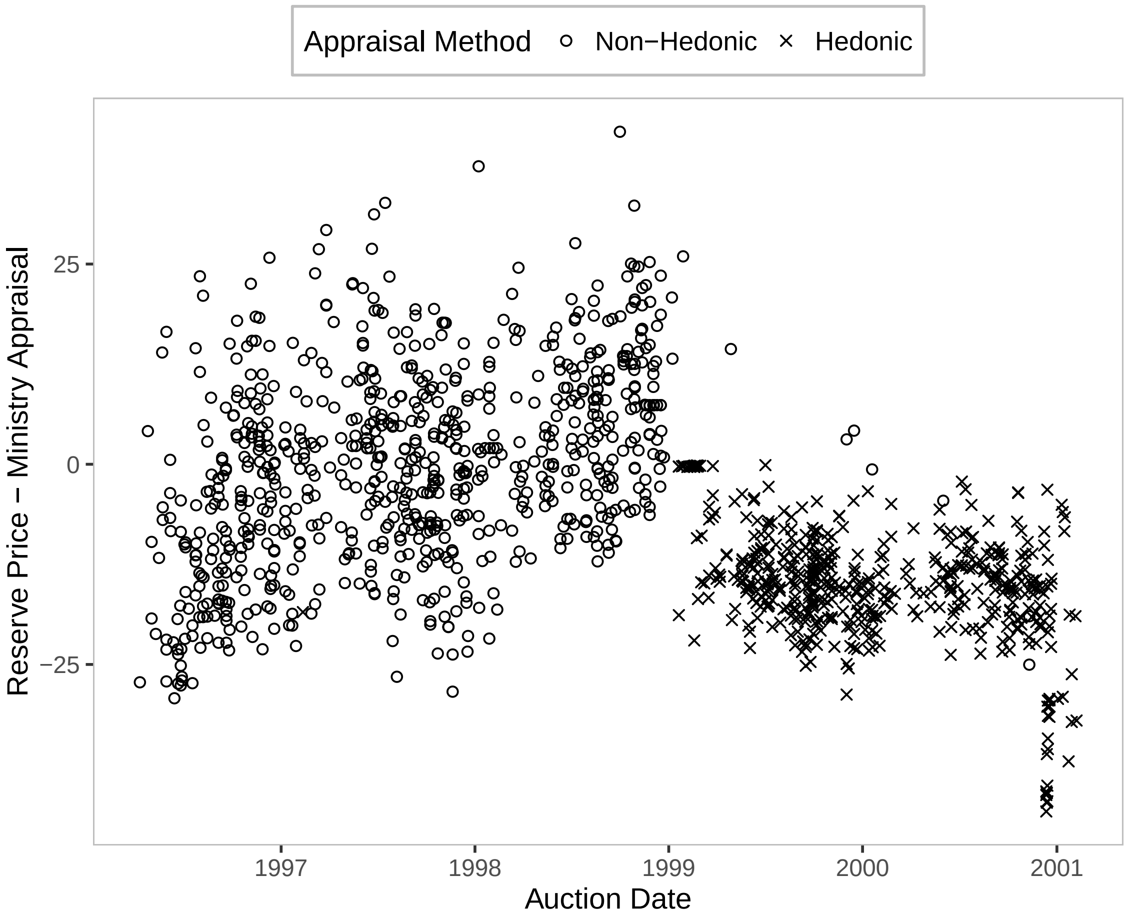


Figure 8: *The time series of homogenized reserve prices using the Ministry of Forests' appraisal method.* Reserve prices increased relative to the Ministry's appraisal each year under the pre-1999 policy. After 1999, there was a sudden decrease in reserve prices. The outliers at the end of 2000 are explained by the fact that the Ministry experimented with setting the reserve price at 50% of their estimated appraisal as opposed to 70%.

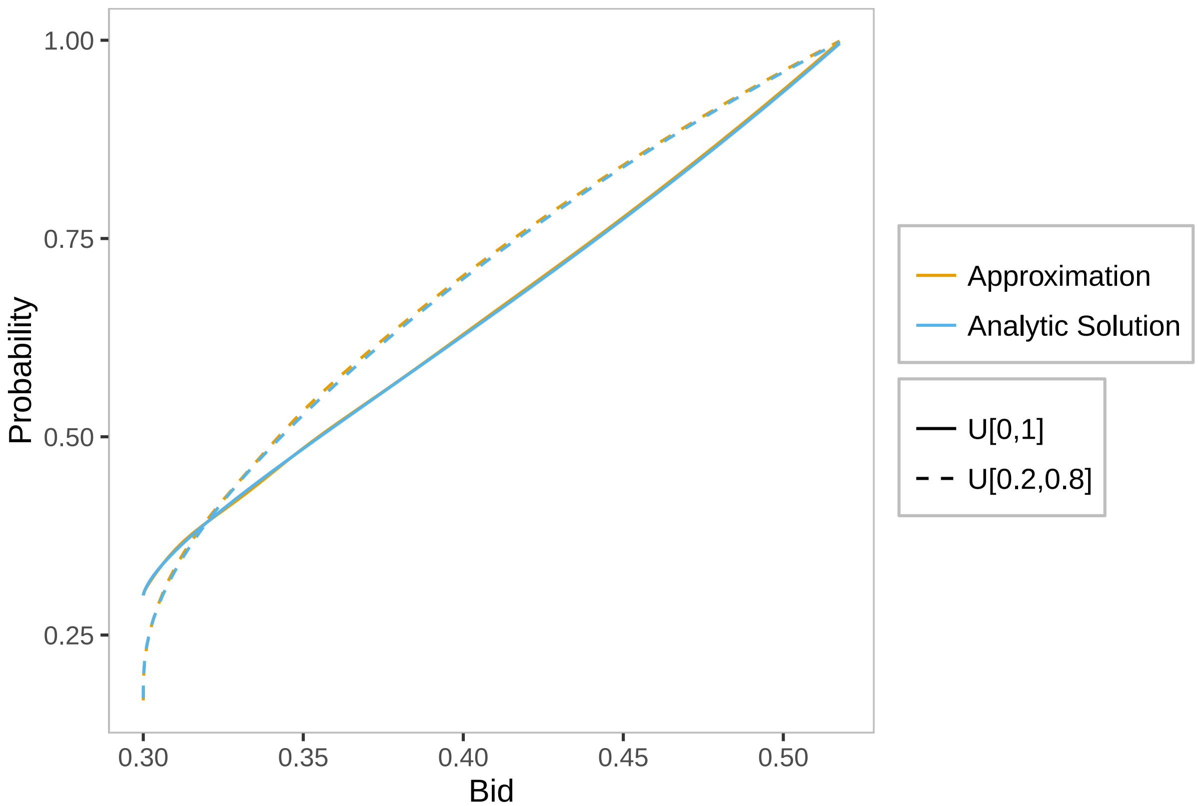


Figure 9: *A spline approximation to equilibrium bid distributions.* Twelve cubic basis spline functions were used to construct an approximate solution to the system of differential equations that defines the Bayes-Nash equilibrium for an auction with two bidders and a reserve price of 0.3. The bidders' valuations are uniformly distributed on $[0, 1]$ and $[0.2, 0.8]$. The splines are defined on a transformed domain in order to reduce the number of basis functions needed to well approximate the equilibrium bid distributions near the reserve price.

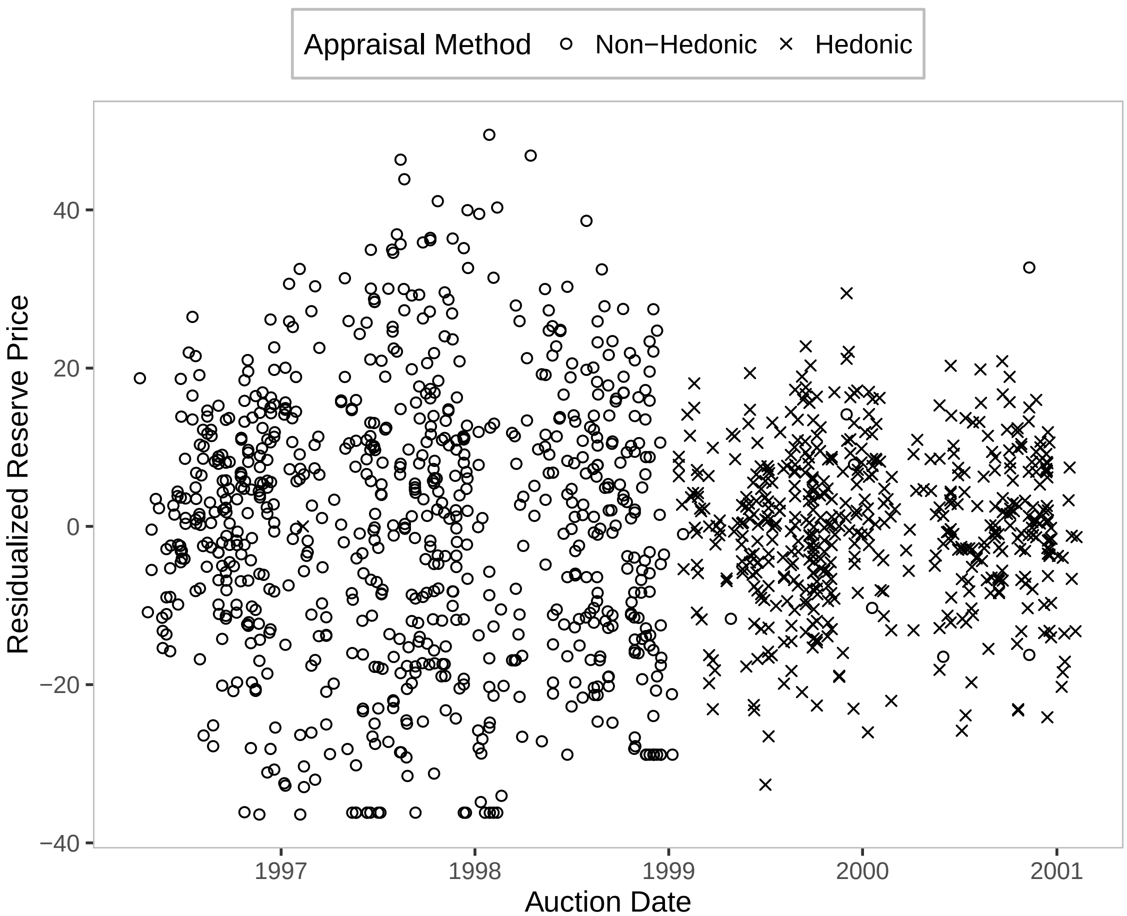


Figure 10: *The times series of residualized reserve prices after controlling for year and appraisal method.* There are no discernible trends in the mean reserve price. The variance is greater in 1997 and 1998.

Table 1: *Simulated Kolmogorov-Smirnov tests for collusion.* Results are based on 1,000 simulations. Critical values for a KS statistic at $\alpha = 0.05$ were estimated from 1,500 bootstrap samples.

T	$\|\mathcal{N}\|$	$\|\mathcal{R}\|$	Size	Power
3,000	3–4	2	0.071	0.891
1,000	3–4	2	0.010	0.158
3,000	3–4	2–3	0.057	0.587
1,000	3–4	2–3	0.072	0.496
3,000	4	2–3	0.091	0.995
1,000	4	2–3	0.072	0.903

Table 2: *Simulated conditional Kendall's τ tests for collusion.* Results are based on 1,000 simulations. Critical values for the conditional Kendall's τ statistic at $\alpha = 0.05$ were estimated from 750 bootstrap samples.

T	Bandwidth	$[\underline{v}_1, \bar{v}_1]$	$[\underline{v}_2, \bar{v}_2]$	$\|\mathcal{R}\|$	Size	Power
200	$T^{-1/5}$	$[0, 1]$	$[0, 1]$	2	0.033	0.273
500	$T^{-1/5}$	$[0, 1]$	$[0, 1]$	2	0.023	0.419
200	$T^{-1/6}$	$[0, 1]$	$[0, 1]$	2	0.037	0.218
500	$T^{-1/6}$	$[0, 1]$	$[0, 1]$	2	0.041	0.324
200	$(T/\log T)^{-1/6}$	$[0, 1]$	$[0, 1]$	2	0.024	0.323
500	$(T/\log T)^{-1/6}$	$[0, 1]$	$[0, 1]$	2	0.032	0.329

Table 3: *Appraisal methods by year.* The hedonic pricing model largely replaced the older appraisal method in 1999.

Time Period	Non-Hedonic	Hedonic
April 1996 to March 1997	571	1
April 1997 to March 1998	544	0
April 1998 to March 1999	453	59
April 1999 to March 2000	131	386
April 2000 to March 2001	97	286

Table 4: *Variable descriptions.*

Statistic	Description
Reserve	Reserve price ($\$/m^3$)
Price	Sale price, i.e. the winning bonus bid plus the reserve price ($\$/m^3$)
Silv. Levy	Amount of silvicultural levy in the reserve price ($\$/m^3$)
Dev. Levy	Amount of developmental levy in the reserve price ($\$/m^3$)
# Bids	Number of bids submitted
Volume	Estimated volume of merchantable timber (1,000 m^3)
Vol. per Week	Average volume per week to be extracted before the license expires
Vol. per Tree	Volume per tree
Vol. per Hect.	Volume per hectare
Price Index	British Columbia's consumer price index
Lumber Price Index	Value of timbers, estimated by multiplying volume by species-specific lumber recovery factors and British Columbia's lumber price indices
Quality Index	Index computed as the ratio of the estimated volume of recoverable lumber to a fixed benchmark defined by the Ministry of Forest's
Cycle Time	Amount of time in hours from the site to the nearest point of appraisal
Avg. Slope	Average slope of land in the tract
Horse	Percent of volume to be extracted by horse
Cable	Percent of volume to be extracted by cable yarding
Helicopter	Percent of volume to be extracted by helicopter
Blowndown	Percent of volume that has been blown down
Burned	Percent of volume that has been burned
Useless	Percent of volume that is dead or useless
Dev. Costs	Anticipated development costs to be paid by licensee ($\$/1,000\ m^3$)
Hemlock	Estimated percent of volume from hemlock trees
Balsam	Estimated percent of volume from balsam trees
Cedar	Estimated percent of volume from cedar trees
White Pine	Estimated percent of volume from white pine trees

Table 5: *Summary statistics.* Summary statistics for Category 1 and Category 2 auctions.

Statistic	Min	Pctl(25)	Median	Pctl(75)	Max	Mean
Reserve	0.25	24.48	34.26	43.48	86.47	33.92
Price	0.68	35.26	46.95	59.33	104.31	47.13
Silv. Levy	0.00	0.00	0.00	6.69	33.24	3.51
Dev. Levy	0.00	0.00	0.00	2.68	28.86	1.80
# Bids	1	2	3	6	19	4.13
Volume	1.00	3.85	6.40	10.79	61.50	8.39
Vol. per Week	0.01	0.06	0.12	0.24	29.28	0.25
Vol. per Tree	0.08	0.34	0.50	0.63	5.01	0.53
Vol. per Hect.	0.004	0.19	0.26	0.34	0.75	0.26
Price Index	0.99	1.00	1.01	1.02	1.05	1.01
Lumber Price Index	50.05	104.02	116.29	128.35	173.49	116.75
Quality Index	0.95	1.17	1.25	1.31	1.54	1.24
Cycle Time	0.00	2.70	3.50	4.60	17.30	3.84
Avg. Slope	0.00	8.00	14.00	23.00	86.00	16.47
Horse	0.00	0.00	0.00	0.00	1.00	0.08
Cable	0.00	0.00	0.00	0.00	1.00	0.07
Helicopter	0.00	0.00	0.00	0.00	1.00	0.02
Blowndown	0.00	0.00	0.00	0.00	1.00	0.02
Burned	0.00	0.00	0.00	0.00	1.00	0.01
Useless	0.00	0.00	0.02	0.07	0.45	0.05
Dev. Costs	0.00	0.00	0.67	1.84	32.31	1.46
Hemlock	0.00	0.00	0.00	0.00	0.99	0.08
Balsam	0.00	0.00	0.01	0.12	1.00	0.11
Cedar	0.00	0.00	0.00	0.00	0.88	0.03
White Pine	0.00	0.00	0.00	0.00	0.32	0.004

$T = 1565$

Table 6: *Ordinary least-squares and partially linear single-index model for bids.* All models are estimated using bids from Category 1 auctions. The OLS regression includes fixed effects for each of nine markets defined using the hierarchical clustering algorithm. The results from all four regressions are robust to the assumed number of markets. The estimates of the partially linear sinbgle-index models are also robust to the choice of bandwidth near the cross-validated optimum.

	OLS	Model 1	Model 2	Model 3
Volume	0.152	0.346	0.380	0.355
	(0.056)	(0.019)	(0.030)	(0.021)
Vol. per Tree	6.963	6.898	6.792	6.895
	(1.408)	(0.417)	(0.657)	(0.417)
Vol. per Hect.	33.966	26.552	26.929	25.366
	(3.619)	(1.119)	(1.542)	(1.181)
Lumber Price Index	0.230	0.168	0.419	0.165
	(0.022)	(0.006)	(0.005)	(0.007)
Quality Index	21.355	33.154		33.408
	(5.665)	(0.663)		(0.725)
Cycle Time	−2.485	−3.004	−2.695	−2.986
	(0.218)	(0.079)	(0.104)	(0.080)
Avg. Slope	−0.177	−0.069	−0.235	−0.091
	(0.043)	(0.016)	(0.017)	(0.017)
Horse	−15.222	−17.413	−19.445	−17.591
	(1.359)	(0.541)	(0.598)	(0.461)
Cable	−10.137	−12.833	−11.255	−12.623
	(1.815)	(0.545)	(0.682)	(0.576)
Helicopter	−48.863	−48.383	−50.716	−48.054
	(2.192)	(1.760)	(0.884)	(1.292)
Blowndown	−9.748	−6.928	−5.928	−6.849
	(2.863)	(0.559)	(0.682)	(0.597)
Burned	−22.544	−14.075	−11.808	−13.995
	(3.036)	(2.969)	(1.938)	(2.853)
Useless	16.362	16.570	17.613	16.817
	(6.030)	(1.437)	(2.697)	(1.674)
Dev. Cost	−0.644	−0.882	−1.005	−0.930
	(0.134)	(0.038)	(0.045)	(0.036)
Hemlock	−13.488	−10.622	−11.442	−10.561
	(2.353)	(0.453)	(1.840)	(0.451)
Balsam	−15.747	−22.092	−22.921	−21.500
	(1.843)	(0.473)	(0.588)	(0.486)
Cedar	−3.987		7.720	4.294
	(4.259)		(1.644)	(2.955)
White Pine	40.559		38.422	34.687
	(14.822)		(9.798)	(9.068)

$T = 1266$

Table 7: *Comparison of appraisal methods.* Linear correlations between the timber license's appraised value using each of the above estimated partially linear single-index models and the Ministry of Forests' appraised value.

	Model 1	Model 2	Model 3	Non-Hedonic	Hedonic
Model 1	1	0.947	0.998	0.712	0.884
Model 2	0.947	1	0.953	0.627	0.819
Model 3	0.998	0.953	1	0.707	0.884
Non-Hedonic	0.712	0.627	0.707	1	
Hedonic	0.884	0.819	0.884		1

Table 8: *Tests for collusion.* Results from simultaneous tests of the null hypotheses that bidder i is not colluding. Each test is formulated as a one-sided test for positive correlation between the reserve price and the valuations that rationalize i's bid. The p-values are estimated from 20,000 bootstrap replicates. The marginal p-values indicate the level of significance in an individual test of the hypothesis for bidder i. To generalize the concept of p-values in the multiple hypothesis testing framework, the adjusted p-values indicate the smallest level of tolerance for one or more type I errors at which the null hypothesis for bidder i can be rejected.

	Model 3, $\hat{\tau}_i^T$		Ministry Appraisal, $\hat{\tau}_i^T$		Model 3, $\hat{\tau}_i^C$		Ministry Appraisal, $\hat{\tau}_i^C$	
	Marginal	Adjusted	Marginal	Adjusted	Marginal	Adjusted	Marginal	Adjusted
1	<.001	**.002**	.004	**.064**	.543	>.999	.356	.999
2	.005	**.075**	.042	.506	.520	>.999	.411	>.999
3	.081	.754	.012	.190	.283	.994	.196	.965
4	.192	.973	<.001	<.001	.664	>.999	.732	>.999
5	.214	.984	.029	.386	.501	>.999	.227	.982
6	.274	.996	.189	.967	.562	>.999	.080	.731
7	.306	.998	.083	.760	.584	>.999	.506	>.999
8	.310	.998	.003	**.049**	.001	**.012**	.291	.995
9	.364	>.999	.166	.949	.142	.914	<.001	**.003**
10	.373	>.999	.015	.219	.574	>.999	.066	.660
11	.491	>.999	<.001	**.005**	.484	>.999	.348	.999
12	.506	>.999	.005	**.081**	.670	>.999	.260	.990
13	.553	>.999	.665	>.999	.008	.120	.246	.987
14	.583	>.999	.384	>.999	.944	>.999	.155	.927
15	.599	>.999	.319	.997	.426	>.999	.583	>.999
16	.606	>.999	.812	>.999	.240	.987	.002	**.033**
17	.807	>.999	.376	>.999	.763	>.999	.540	>.999

Table 9: *Number of bids submitted in the same auctions.* The diagonal contains the number of times each of the most active firms submitted a bid. The off-diagonal entries indicate the number times each pair of firms participated in the same auction.

Firm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	35	4			1	10		5			5	1					3
2	4	31	2			2	1	4		2		1					5
3		2	35	1	1	1	17	10		1	3	5	4				1
4			1	59			4	2		8	1		9				
5	1		1		35	7	7	6		2	5	1	2				
6	10	2	1		7	42	1	9			6	4					
7		1	17	4	7	1	60	10		3	5	5	6				2
8	5	4	10	2	6	9	10	58		5	5	7	3				6
9									42					1	12		
10		2	1	8	2		3	5		37	1		7			1	
11	5		3	1	5	6	5	5		1	38		2				1
12	1	1	5		1	4	5	7				36	1				
13			4	9	2		6	3		7	2	1	40				2
14									1					39	1		
15									12					1	38		
16										1						33	
17	3	5	1				2	6			1		2				34

Table 10: *Number of wins in rivals' territories.* Each row tabulates the number of times a firm won in a district where each of the other firms has ever won an auction. Conversely, each column suggests who the firm's strongest competitors might be. For example, firm 1 won 8 times in districts where firm 2 won an auction, whereas firm 2 won 4 times in districts where firm 1 won an auction. Their strongest competitors are firms 6 and 8, each of which won 10 auctions in the districts where firms 1 and 2 were also competitive.

Firm	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	8	8			4	8		8			8	4					4
2	4	6			1	4		4			4	1					1
3			6		6	6	6	6				6	6				
4				10						9			8				
5	1	1	3		9	4	4	5		1	3	6	4				3
6	10	10	1		8	11	1	11			10	8	1				7
7			8		10	8	10	8		2		8	10				
8	10	10	4		13	14	4	16			11	13	4				8
9									7					2	5		
10				8	1		1			9			6				
11	4	4			8	4		5			10	9					3
12	4	4	3		9	7	3	8			9	13	3				4
13			5	3	7	5	7	5		5		5	10				
14									2					7			
15									7						7		
16																7	
17	5	5			6	5		5			5	5					7

Table 11: *Tests of specific configurations of a bidding ring.* The same procedures are used to test whether a model with collusion better predicts bidders' responses to variation in the reserve prices. When firms 1 and 2 are assumed to be the only colluders, the null hypotheses that their valuations are independent of the reserve prices can still be rejected. When five of the other firms who compete in the same markets as 1 and 2 are included in the ring—firms 5, 6, 8, 11, and 17—the test statistics are still significantly positive. The test only fails to reject any null hypotheses when firm 12 is included in the ring. Firm 12 won 4 times in districts where firms 1 and 2 also won auctions, but only bid against them once.

	Model 3, $\hat{\tau}_i^T$, $\mathcal{R} = \{1,2\}$		Model 3, $\hat{\tau}_i^T$, $\mathcal{R} = \{1,2,5,6,8,11,17\}$		Model 3, $\hat{\tau}_i^T$, $\mathcal{R} = \{1,2,12\}$	
	Marginal	Adjusted	Marginal	Adjusted	Marginal	Adjusted
1	<.001	**.003**	<.001	**.003**	.060	.643
2	.005	**.079**	.006	**.099**	.010	.158
3	.081	.755	.081	.754	.081	.756
4	.192	.973	.192	.971	.192	.974
5	.214	.984	.360	>.999	.214	.984
6	.274	.996	.192	.971	.274	.995
7	.306	.998	.203	.977	.306	.998
8	.310	.998	.310	.998	.310	.998
9	.364	>.999	.364	>.999	.364	>.999
10	.373	>.999	.373	>.999	.373	>.999
11	.491	>.999	.491	>.999	.450	>.999
12	.506	>.999	.392	>.999	.506	>.999
13	.553	>.999	.553	>.999	.553	>.999
14	.583	>.999	.583	>.999	.583	>.999
15	.599	>.999	.599	>.999	.599	>.999
16	.606	>.999	.606	>.999	.606	>.999
17	.807	>.999	.710	>.999	.807	>.999