Deciding Between Competition and Collusion

Patrick Bajari and Lixin Ye[1]

July 12, 2002

Abstract

In this research, we develop an approach to identification and testing for bid-rigging in procurement auctions. First, we introduce a general auction model with asymmetric bidders. Second, we study the problem of identification in our model. We state a set of conditions that are both necessary and sufficient for a distribution of bids to be generated by a model with competitive bidding. Third, we discuss how to elicit from industry experts a prior distribution over firm's structural cost parameters. Given this prior distribution, we use Bayes theorem to compare competitive and collusive models of industry equilibrium. Finally, we apply our methodology to a data set of bidding by construction firms in the Midwest. The techniques we propose are not computationally demanding, use flexible functional forms and can be programmed using most standard statistical packages.

[1] We would like to thank Robert S. Pindyck and two anonymous referees for extremely helpful comments and Patrick Muchmore, Nancy Dinh and Susan Yun for their dedicated assistance in collecting the data set. Seminar participants at several universities also provided helpful feedback. We would like to thank the Minnesota, North Dakota and South Dakota Departments of Transportation and numerous City and County governments in the Midwest for their kind assistance in completing the data set. Also, thanks to Construction Market Data and Bituminous Paving Inc. for patiently answering our questions about the construction market. Funding from SIEPR and the National Science Foundation is gratefully acknowledged. Parts of this research were completed while the first author was a National Fellow at the Hoover Institution. This paper is based in part by an earlier paper by the first author, "Econometrics of the First Price Sealed-Bid Auction with Asymmetric Bidders".

1 Introduction.

Bid-rigging is a serious problem in many procurement auctions. According to Engineering News-Record, criminal bid-rigging cases have recently been filed in New York City and Chicago for building public schools, bridge repair, interior remodeling, paving and many other types of construction. A widely publicized instance of bid-rigging occurred in the New York cement industry in the 1980's, where organized crime designed an elaborate bid-rigging scheme that inflated building costs, making the price of poured concrete the highest in the nation.[2] Developing effective and computationally simple tools that can be used by regulators to detect bid-rigging might serve as a deterrent to future collusion, and thereby lower prices and enhance efficiency in some industries.

In this paper, we develop an approach to identification and testing for bid-rigging in procurement auctions. We begin by describing a model of competitive bidding for procurement contracts. A unique aspect of our model is that bidders are asymmetric, that is, ex-ante, the costs of bidders may differ. Asymmetries are commonplace in procurement and may arise due to the location of firms, capacity constraints, or familiarity with local rules and regulations. Next, we study the problem of identification in the asymmetric auction model. We state a set of conditions that is both necessary and sufficient for a distribution of bids to be rationalized by our model and we also discuss how these conditions could be tested. Finally, we use the tools of statistical decision theory to decide between competitive and collusive models of industry equilibrium. We elicit from industry experts a prior distribution over the structural cost parameters that enter into our models. Given this prior distribution over structural cost parameters, we use Bayes Theorem and the laws of conditional probability to decide between competitive and collusive models of industry equilibrium. We apply our methodology to a data set of bidding by construction firms. The techniques needed for the

[2] In his biography, Mafia turncoat Sammy (The Bull) Gravano, who claims to have orchestrated collusion in the construction industry for the Gambino crime family said, "If one of them (contractor) gets a contract for, say, thirteen million, the next thing you know, after he knows he's got it, he jacks up the whole thing before it's over to a sixteen-or seventeen-million-dollar job. Now he's increased the cost thirty-three percent. So our greed (the Mafia) is compounded by the greed of them so-called legitimate guys (contractors)." Maas (1997), p. 271.

computations are not particularly complex and can be programmed using standard statistical packages.

In empirical industrial organization, it is common to consult with industry experts during a research project in order to learn about the nature of demand, costs and strategic interactions in the marketplace. However, information garnered from industry sources is seldom formally incorportated into the estimation. A unique feature of our analysis is that we elicit the beliefs of industry experts about the model parameters and incorporate these beliefs into estimation.

The firms in our application survive in an unforgiving environment: poor decision making is punished with personal financial loss (the firms in our industry are all owner operated). In our data set, over a hundred firms compete, but only 7 firms have market shares that exceed 5 percent. We expect that the a priori beliefs elicited from owner-operators with decades of experience contain important information about industry costs that can help us to estimate parameters and decide between alternative models of industry equilibrium. We do not advocate exclusively estimating models by imposing priors from external sources such as industry experts, consulting engineers or internal cost records of firms. However, we believe that in some circumstances, especially when other identifying information is not available, the data is of poor quality and highly limited, utilizing this type of prior information about the parameters values can be very useful in estimation and decision making.

There are a number of recent empirical papers on the subject of bid-rigging. The first set of empirical papers describe the observed bidding patterns of cartels and compare cartel to non-cartel bidding behavior. Porter and Zona (1993,1999) and Pesendorfer (1996) both analyze data sets where it is known that bid-rigging has taken place. These papers find the following empirical regularities: First, cartel members tend to bid less aggressively than non-cartel members. Second, the bids of cartel members tend to be more correlated with each other than with the bids of non-cartel members. Third, collusion tends to increase prices as compared to a non-collusive control group.

The second set of empirical papers, such as Porter and Zona (1993) and Baldwin, Marshall and Richard (1997) propose econometric tests designed to detect collusive bidding. Baldwin, Marshall and Richard (1997) nest both competition and collusion within a single model to test for collusion. Their model is applicable for oral or second-price auctions with private values. Porter and Zona (1993) propose a procedure where two models of bidding are estimated. The first model is a logistic regression on the identity of the lowest bidder. The second model is an ordered logit regression on the ranking of all the bidders. Under the null hypothesis of no collusion, the parameter values from the two models should be equal. Our companion paper Bajari and Ye (2001) sheds some new light on the analysis of Porter and Zona (1993,1999) and Baldwin, Marshall and Richard (1997) by demonstrating how observable differences across firms such as location and capacity play a key role in the identification of collusion.

Finally our paper is also related to the literature on structural estimation of the first-price auction model. Most closely related to our research is Guerre, Perrigne and Vuong (1999) which estimates the first-price auction model non-parametrically. Also related are Bajari (1997) and Pesendorfer and Jofre-Bonet (2000) which structurally estimate asymmetric models of the first price auction.

We believe that taken together with our companion paper Bajari and Ye (2001), the tests we have proposed are a useful diagnostic for detecting suspicious bidding behavior. While no method for detecting collusion is likely to be fool proof, we believe that our methods can be used as a first step to determine whether suspicious bidding might have occurred and to determine whether further investigation is warranted.

2 The Model

In this section, we develop a model of competitive bidding for a contract to build a single and indivisible public works project. The framework of the model is appropriate for modeling bidding behavior in our seal coating data set as well as bidding for many other types of procurement contracts.

In most auction models, it is usually assumed that firms are symmetric. In procurement settings, this assumption implies that private cost estimates are independently and identically distributed. However, this is not an appropriate assumption for the seal coat industry as well as many other procurement auctions. In the seal coat industry, there are at least four sources of asymmetries. The first is location. Firms that are closer to a construction project are more likely to submit a bid and their bids would be lower, other things being equal. The second source of asymmetries is capacity utilization. Each firm needs to take into account that winning a contract today will limit its available capacity to complete future projects. The third source of asymmetries is different technologies across firms. Firms clearly differ in their sizes, most likely due to technology and managerial efficiency. The fourth source of asymmetry is that firms differ in their success in winning contracts by state. Many firms in our data set complete the majority of their work in a single state. In order to bid successfully, firms must be familiar with local regulations and local procurement officials.

To better capture all the industry features described above, in our model we assume that firms are asymmetric. There are N firms competing for a contract to build a project. Before bidding starts, each firm i forms an estimate of its cost to complete the project. Firm i knows its own cost estimate but does not know the cost estimates of other firms. The cost estimate for firm i is a random variable C_i with a realization denoted as c_i. The random variable C_i has a cumulative distribution function $F_i(\cdot;\theta_i)$ and probability density function $f_i(\cdot;\theta_i)$ where θ_i is a vector of parameters. We will let $\theta = (\theta_1, ..., \theta_N)$ denote the vector of all firm specific parameters. The cost distribution is assumed to have support $[\underline{c}, \overline{c}]$ for all firms.

As a simple example, consider a situation where each firm has a different location and hence different transportation costs. Then one natural specification for firm i's private information is:

$$c_i = constant + \beta_1 * distance_i + \varepsilon_i \tag{1}$$

In equation (1), c_i, firm i's cost estimate is a function of three terms. The first term is a constant, which could reflect attributes of the project that affect all firms identically, such as how many miles of highway must be paved or the tons of concrete that must be poured for the foundation of a new building. The second term, $\beta_1 * distance_i$, is different for different firms due to different locations. The third term, ε_i, serves to model private information about some component of firm i's cost, such as the cost for materials or labor. If ε_i is normally distributed with mean zero and standard deviation σ then $\theta_i = (constant, \beta_1, distance_i, \sigma)$.

Let b_i denote the bid submitted by firm i. Firms are assumed to be risk-neutral. If firm i submits the lowest bid then firm i's payoff is $b_i - c_i$, and if it fails to submit the lowest bid then firm i's payoff is 0. If two or more firms submit the same bid, the contract will be awarded at random among the set of the lowest bidders. But ties only occur with probability zero in equilibrium. Firm i's payoff function can thus be written as:

$$u_i(b_1, ..., b_n, c_i) = \begin{cases} b_i - c_i & \text{if } b_i < b_j \text{ for all } i \neq j \\ 0 \text{ otherwise.} & \end{cases} \tag{2}$$

Firm i has private values since its payoff depends only on c_i and not the private information of other firms. Armantier, Florens and Richard (1997) and Porter and Zona (1993) both use the assumption of private values in their models of procurement auctions. If firm specific factors account for the differences in cost estimates, then the assumption of private values is plausible. In the seal coat industry, it is reasonable to assume that firms have private value components to costs, because labor and material costs are mainly firm specific and to some extent private information.

Firm i's strategy is a function $b_i = \mathbf{b}_i(c_i;\theta)$ which maps firm i's cost draw, c_i, to a bid b_i in the interval $[\underline{c}, \overline{c}]$. LeBrun (1994) and Maskin and Riley (1996a,b) have shown that, in equilibrium, the bid functions $b_i = \mathbf{b}_i(c_i;\theta)$ are strictly increasing and differentiable which implies that the inverse bid functions $\phi_i(b;\theta) = \mathbf{b}_i^{-1}(b;\theta)$ are also strictly increasing and differentiable. To simplify the notation, we will often write firm i's bid function as $\mathbf{b}_i(c_i)$, suppressing its dependence on the vector of parameters θ.

Since $\mathbf{b}_j(c_j)$ is strictly increasing, when firm i has a cost draw of c_i, her expected profit from bidding b_i will be:

$$\pi_i(b_i, c_i; \mathbf{b}_{-i}, \theta) = (b_i - c_i)Q_i(b_i;\theta) \tag{3}$$

where $Q_i(b;\theta) = \prod_{j \neq i} 1 - F_j(\phi_j(b;\theta);\theta_j)$ is the probability that firm i is the lowest bidder. As we can see from equation (3), firm i's expected profit is a markup times the probability that firm i is the lowest bidder.

Static models with competitive bidders may not be sufficiently general in many empirical applications. We now demonstrate that models with non-trivial dynamics and collusion can be viewed as special cases of the model developed above.

2.1 Dynamic Bidding

In the construction industry, firms must usually bid in a sequence of auctions over time. Firms are often capacity constrained due to limited physical capital and limited pool of workers who possess the necessary specialized skills. In a dynamic equilibrium, a firm must account for the fact that winning a project in an early auction will mean that it has less free capacity for future jobs. Capacity constraints have important implications for bidding strategies. A stylized fact in many markets is that firms tend to bid more aggressively early in the construction season when they have more unutilized capacity. Also, according to accounts by industry participants, in order not to lose skilled workers, construction firms may even bid at less than cost to prevent their employees from being idle.

Let s be a vector denoting the state of the industry and assume that the state is publicly observed by all participants in the auction. At any given point in time, this vector might include: the capacities for all firms in the industry, the distances of all the firms to the project locations, the market prices of key materials and so forth.

Assume that $V_{iW}(s)$ is the continuation value attached to winning at state s, and $V_{iL,j}(s)$ is the continuation value attached to not winning (and that firm j is the winning firm) at state s. When i wins the job, its used capacity increases and all the competitors' capacities remain unchanged for the next period. When j $(j \neq i)$ wins the job, firm j's used capacity increases and the capacities of all the other firms (including firm i) remain unchanged for the next period. Note that in general, $V_{iL,j}(s)$ may not be the same for all j $(j \neq i)$.

To simplify the equilibrium analysis, we assume that the option value attached to not winning is the same regardless of the winner's identity, i.e., we assume that $V_{iL,j} = V_{iL}$. With this assumption, the firm's expected payoff will now have the general form:

$$\pi_i(b, c_i; \theta) = (b - c_i + V_{iW}(s))Q_i(b; \theta) + V_{iL}(s)(1 - Q_i(b; \theta)) \quad (4)$$

$$= (b - c_i + V_{iW}(s) - V_{iL}(s))Q_i(b; \theta) + V_{iL}(s) \quad (5)$$

Since constants are irrelevant in a firm's maximization problem, we may write the firm's maximization problem as:

$$\max_b (b - c_i + V_{iW}(s) - V_{iL}(s))Q_i(b; \theta) \quad (6)$$

As in equation (3), a firm's optimization problem is the difference between its bid and a cost estimate times the probability of winning. The only difference now is that a firm's cost includes an estimate of the option value of having free capacity available for future periods. Since both $V_{iW}(s)$ and $V_{iL}(s)$ are common knowledge to all participants in the game, the above equation is equivalent to maximizing utility in a static auction model with cost $c_i + V_{iL}(s) - V_{iW}(s)$.[3]

[3] Once option values are included in the analysis, it might be the case that the distributions of costs no longer have common supports across the firms. As we discuss in the next section, this can cause problems for the existence, uniqueness and other characterizations of the equilibrium. The theoretical problem remains, to the best of our knowledge, unresolved. However, if one makes an arbitrarily small perturbation to the information structure so that both firms have the same support, then the existence, unqiueness and characterization theorems still hold.

The key assumption needed for the validity of this analysis is that $V_{iL,j}(s)$ is independent of j. While this assumption may not be strictly true in a dynamic model of bidding, our empirical work suggest that this assumption may not be a bad approximation. If firm i was concerned about which firm $j \neq i$ wins the procurement in the event that i does not, firm i should bid differently depending on the capacity utilization of her competitors. However, while a firm's own available capacity was significant in reduced form bid functions, the capacity of other firms failed to be significant determinant of firm i's bid. This is consistent with firm i being indifferent between which competing firm $j \neq i$ wins the contract.

2.2 Collusion

Numerous mechanisms for collusion have been found in the construction industry. For example, firms may follow bid rotation schemes to allocate projects, side payments are sometimes made between firms and geographic territories have been established as parts of cartel arrangements.

We will demonstrate that a simple model of collusive bidding where the cartel behaves efficiently is a special case of our model presented at the beginning of this section. Suppose that before the auction begins, all cartel members make a cost draw. The cartel members meet before the auction, compare cost draws and the cartel member with the lowest cost draw submits a real bid, while other cartel members either abstain from bidding or submit phantom bids. Let $C \subseteq \{1, 2, ..., N\}$ denote the cartel. The cost to the cartel (c_c) can be denoted as:

$$c_c = \min_{j \in C} c_j \tag{7}$$

If other bidders are aware of the identity of the cartel, then it is trivial to adapt our previous analysis to the case of a cartel. The cartel is simply modeled as the order statistic of its members' cost draws.

3 Properties of Equilibrium

In this section, we summarize some theoretical properties of the asymmetric auction model. This section can be skipped by those readers primarily interested in the empirical implementation of our procedures.

In the asymmetric auction model, we assume that firms play a Bayes-Nash equilibrium in pure strategies. Firm i first of all makes a cost draw c_i, then taking the cost draw as given, chooses a bid b_i that maximizes (3).

The first order condition for maximizing expected profit in our model is:

$$\frac{\partial}{\partial b}\pi_i(b, c_i; \theta) = (b - c_i)Q_i'(b; \theta) + Q_i(b; \theta) = 0 \quad (8)$$

The equilibrium to the model can be characterized as the solution to a system of differential equations with boundary conditions. This is done by rearranging equation (8):

$$\frac{\partial}{\partial b}\pi_i(b, c_i; \theta) = \prod_{j \neq i}[1 - F_j(\phi_j(b))] - (b - c_i)\prod_{j \neq i} f_j(\phi_j(b))\phi_j'(b)\prod_{k \neq j}[1 - F_k(\phi_k(b))] = 0, \;\; i = 1, ..., N. \quad (9)$$

Equation (9) involves the inverse bid function $\phi_j(b)$ and its derivative $\phi_j'(b)$. Collecting terms and rewriting (9) we can characterize the equilibrium inverse bid functions as the solution to a system of N ordinary differential equations:

$$\phi_i'(b) = \frac{1 - F_i(\phi_i(b))}{(N-1)f_i(\phi_i(b))}\left[\frac{-(N-2)}{b - \phi_i(b)} + \sum_{j \neq i}\frac{1}{b - \phi_j(b)}\right], \;\; i = 1, ..., N. \quad (10)$$

Throughout our analysis we will impose the following regularity conditions on our model primitives:

- Assumption 1. For all i, the distribution of costs $F_i(c_i; \theta)$ has support $[\underline{c}, \overline{c}]$. The probability density function $f_i(c_i; \theta)$ is continuously differentiable (in c_i).
- Assumption 2. For all i, both $f_i(c_i; \theta)$ is bounded away from zero on $[\underline{c}, \overline{c}]$.

Theorem 1 *(LeBrun (1994, 1995a,b) and Maskin and Riley (1996a,b)). If Assumptions 1 and 2 hold, then an equilibrium in pure strategies exists. Furthermore, the equilibrium is strictly monotone and differentiable.*

Another basic result from the theoretical literature is that the inverse bid functions can be characterized as the solution to a system of N differential equations with $2N$ boundary conditions and that the equilibrium is unique.

Theorem 2 *(LeBrun (1995a), Maskin and Riley (1996a)) Suppose that Assumptions 1 and 2 hold. Let $\phi_1(b), ..., \phi_N(b)$ be inverse equilibrium bidding strategies. Then*

(i) For all i, $\phi_i(\overline{c}) = \overline{c}$,

(ii) There exists a constant β such that for all i, $\phi_i(\beta) = \underline{c}$,

(iii) For all i and for all $b \in [\beta, \overline{c}]$, equation (10) holds.

Theorem 3 *(Maskin and Riley (1996a), Bajari (1997), Bajari(2000), LeBrun (2000)) Suppose that assumptions 1 and 2 hold. Then there is a unique equilibrium.*

4 Identification

In this section, we will state a set of conditions that are both necessary and sufficient for a distribution of bids to be generated from the asymmetric auction model of the previous sections. We assume that it is possible to write the cumulative distribution function of firm i's cost in the form $F(c_i|z_i)$ where $z_i \in Z$ is a vector of parameters and covariates that is publicly observable to all other firms. For instance, in equation (1) the vector z_i would be the tuple $z_i = (constant, \beta_1, distance_i, \sigma)$. Let z denote the vector $z = (z_1, ..., z_n)$.

Let $G_i(b; z)$ be the cumulative distribution of firm i's bids and $g_i(b; z)$ be the associated probability density function. Using the Theorems from the previous section, it is straightforward to show that the following conditions must hold in equilibrium.

A1 Conditional on z, firm i's bid and firm j's bid are independently distributed.

A2 The support of each distribution $G_i(b; z)$ is identical for each i.

Conditional on z, each firm's signal c_i is independently distributed and since bids are function of c_i, A1 must hold in equilibrium. Condition A2 holds by the characterization Theorem stated in the previous section.

A third condition that must hold in equilibrium is that the distribution of bids must be *exchangeable* in the z_i. Let π be a permutation, that is, a one-to-one mapping from the set $\{1, ..., N\}$ onto itself. The definition of exchangeability is that for any permutation π and any index i the following equality must hold:

$$G_i(b; z_1, z_2, z_3, ..., z_N) = G_{\pi(i)}(b; z_{\pi(1)}, z_{\pi(2)}, z_{\pi(3)}, ..., z_{\pi(N)}) \tag{11}$$

Equation (11) implies if the cost distributions for the bidders are permuted by π, then the empirical distribution of bids must also be permuted by π. For instance, if we permute the values of z_1 and z_2 holding all else fixed exchangeability implies that $G_1(b)$ and $G_2(b)$ also permute.

A3 The equilibrium distribution of bids is exchangeable. That is, for all permutations π and any index i $G_i(b; z_1, z_2, z_3, ..., z_N) = G_{\pi(i)}(b; z_{\pi(1)}, z_{\pi(2)}, z_{\pi(3)}, ..., z_{\pi(N)})$.

The fourth condition that must hold in equilibrium is that the bid functions must be strictly monotone. First of all note that we can write the first order conditions for equilibrium as:

$$c_i = b - \frac{1}{\sum_{j \neq i} \frac{f(\phi_j(b;z)|z_i)\phi_j'(b;z)}{1-F(\phi_j(b;z)|z_i)}}. \tag{12}$$

Since equilibrium bid functions are strictly monotone it follows by a simple change of variables argument that $G_i(b;z)$ and $g_i(b;z)$ must satisfy:

$$G_i(b;z) = F(\phi_i(b;z)|z_i) \tag{13}$$

$$g_i(b;z) = f(\phi_i(b;z)|z_i)\phi_i'(b;z) \tag{14}$$

Now define $\xi_i(b;z)$ as follows

$$\xi_i(b;z) = b - \frac{1}{\sum_{j \neq i} \frac{g_i(b;z)}{1-G_i(b;z)}}, \text{ then} \tag{15}$$

A4 For all i and b in the support of the $G_i(b;z)$ the function $\xi_i(b,z)$ is strictly monotone.

Finally, from our characterization theorem, the following boundary conditions should also hold:

A5 $\xi_i(\bar{b};z) = \bar{c}$, $\xi_i(\underline{b};z) = \underline{c}$ for $i = 1, 2, \cdots, N$.

We formalize the above observations into theorem 4.

Theorem 4 *Suppose that the distribution of bids $G_i(b;z), i = 1, ..., N$ is generated from a Bayes-Nash equilibrium. Then conditions A1-A5 must hold.*

The next set of result shows that if the conditions A1-A5 hold then it will be possible to construct a distribution of costs $F_i(b|z)$ that uniquely rationalizes the observed bids $G_i(b;z)$ as an equilibrium (see Bajari and Ye (2001) for a proof).

Theorem 5 *Suppose that the distribution of bids $G_i(b;z)$ satisfies conditions A1-A5. Then it is possible to construct a unique set of $F(c|z)$ such that $G_i(b;z)$ is generated from an equilibrium to the game when costs are distributed as $F(c|z)$.*

If there is no variation in z which is observable to the econometrician, it will typically not be possible to determine whether collusion has occurred.

Theorem 6 *If there is no variation in z and A1-A5 hold, then competition is observationally equivalent to a cartel C that is not all inclusive.*

Even if there is variation in z, it may still not be possible to empirically distinguish collusion from competition. A sophisticated cartel that includes all N firms may be able to construct a mechanism for collusion that satisfies conditions A1-A5. For instance, suppose that the cartel operates by first having each firm compute its competitive bid and then submit a bid of 1.1 times its competitive bid. It is straightforward to show that conditions A1-A5 are satisfied if the cartel colludes in this fashion.

5 Competitive Bidding For Seal Coat Contracts

We have compiled a unique data set for the purpose to apply our test for collusive bidding. Our data set was purchased from Construction Market Data (CMD) and contains detailed bidding information for nearly all of the public and private road construction projects in Minnesota, North Dakota and South Dakota. We purchased all archived records of road construction projects for these states awarded during the years 1994-1998. This data set contains nearly 18,000 unique procurement contracts and was a whopping 48 MB in size.[4] For each project, the data set contained a wealth of information including, for example, the project location, the deadline for bid submission, bonding requirements, the identities of all of the bidders, some very detailed project description and many other variables, etc.

We decided to focus on a particular submarket in road construction called seal coating. Seal coating is a maintenance process designed to extend the life of a road. It adds oil and aggregate (sand, crushed rock, gravel or pea rock) to the surface of a road. This gives the road a new surface to wear and also adds oil that will soak into the underlying pavement to slow the development of cracks in the highway. Seal coating is a low cost alternative to resurfacing a highway.[5] There are numerous advantages to focusing on seal coating,

[4] Construction Market Data sells information to general contractors about upcoming construction projects. Many of the general contractors we have spoken with subscribe to *Construction Bulletin*, a weekly periodical published by CMD, to search for work. *Construction Bulletin* also reports bids for contracts that were awarded in previous weeks. From conversations with DOT officials, general contractors, and CMD, we believe that almost all public and private road construction projects exceeding $10,000 are contained in our data set. Some of the data fields for projects owned by city and county governments were not complete. We phoned hundreds of county and city governments throughout the Midwest to fill in missing fields in the CMD database.

[5] In our data set, almost all of the seal coating takes places from late May to mid-September since standard engineering specifications require a temperature of at least 60 degrees before seal coating starts. In fact, most State Departments of Transportations have seasonal limitations on seal-coating that do not allow any sealing to be done before the first week of June or after mid-September.

as spelled out in Bajari and Ye (2001).

5.1 Contract Award Procedures

All public sector seal coat contracts are awarded through an open competitive bidding process.[6] In seal coat projects, contractors do not submit a single bid; rather they submit a vector of bids. This is known as a unit price contract and has the form:

contract item #1	estimated quantity for item #1	unit price for item #1
contract item #2	estimated quantity for item #2	unit price for item #2
contract item #3	estimated quantity for item #3	unit price for item #3
⋮	⋮	⋮

The contract items might include gallons of oil, tons of aggregate and mobilization. Both the contract items and the estimated quantities are established by the owner of the contract (typically a city government or State DOT) and the unit prices are chosen by the contractor.[7,8]

If the contract is awarded, it must by law be given to the lowest bidder. Public officials have the right to reject all bids, but this rarely occurs in practice. Firms also have strong financial incentives to honor their contractual obligations if they are the low bidders. Contractors usually must submit a bid bond of 5 to 10 percent of their total bid guaranteeing that they will not withdraw their bid after the public reading of all bids. After the contract is awarded, the low bidder must submit a performance bond and pay bond to guarantee the completion of the contract and that all subcontractors will be paid. For a more complete discussion of contract procedures see Minnesota Department of Transportation (1999), Bartholomew (1998), Clough and

A typical crew for a seal coat company consists of 2 workers on the chip spreader, one distributor operator, 4 roller operators, 4 flag persons, one person to drive a pilot car, one to drive the broom and one to set temporary pavement markings. On a typical project there can be between 5 to 15 trucks hauling the aggregate to the project site and a loader operator to fill the trucks with aggregate.

According to one company in the industry, who primarily works in the Dakotas, a typical crew (excluding trucks) costs $1,500 per day in labor and $1,000 per day in the implicit rental price for machinery. A crew can typically expect to seal coat 7 to 15 miles of highway per day depending on conditions. The cost of trucking, according to the firm, is $35 per hour (including the driver).

[6] The seal coat contract documents contain 6 major parts: bidding documents, general conditions of the contract, supplementary conditions of the contract, specifications, drawings and report of investigations of physical site conditions. (Please refer to Bajari and Ye (2001) for the detailed descriptions).

[7] In our analysis, we abstract away from the fact that the firms make a vector of bids instead of a single bid. As in Athey and Levin (2000), it would be possible to model bidding as a two-stage procedure. In the first stage bidders choose a total bid and in the second stage they choose unit prices optimally.

[8] The contractor is compensated according to the quantities that are actually used on the job. DOT personel monitor the firm while the work occurs and are responsible for verifying measurements of quantities of material put in place. If actual quantities are 20% less or more than the estimated quantities the price will be renegotiated according to procedures described in the contract.

Sears (1994), and Hinze (1993).

5.2 The Data Set

We observe all public sector seal coat contracts awarded from January 1994 through October 1998 in our data set. There are four owner types: City, County, State and Federal. Most of contracts are owned by City or State. Among all jobs, 230 (46.5%) are owned by Cities, 195 (39.3%) are owned by States, 68 (13.7%) are owned by Counties, and only 2 owned by the Federal Government. The total value of contracts awarded in our data set is $\$92.8$ million.

The size of contracts vary greatly. Of the 495 contracts in our data set, 7 contracts were awarded for more than $1 million, 256 contracts were awarded for less than $1 million but more than $100 thousand, and 232 contracts were awarded for less than $100 thousand. A total of 98 firms bid on at least one of these 495 contracts, with 43 firms never winning any contract for the period reported.

The market concentration can be seen from Table 2, which summarizes the firms' bidding activities, while the firms' identities are listed in Table 1. Among those 55 firms winning at least one job, only 18 firms have a market share exceeding 1%.[9] The largest 7 firms in the market have a market share of 65.6% and are led by Firm 2 (Astech Paving) who alone accounts for 21% of the market shares, attending 66.9% of the auctions conducted.

The owner of the largest firm in the market, Astech (Firm 2) received a one year prison sentence for bid-rigging in the late 1980's. The owners of two other firms, McLaughlin & Schulz Inc. (Firm 5) and Allied Paving (Firm 1) were also fined for bid-rigging with Astech in the seal coat industry. The owners of all three firms were, at one time, banned from bidding for public sector seal coat contracts. Whether these or any other firms in the industry are still engaged in anti-competitive behavior is an interesting question.

In Table 3 we study the concentration of bidders who attend any given auction. The average number of bids in an auction is 3.3 with 29 contracts receiving only one bid. We conjecture that the low participation has to do with bid preparation costs on the part of contractors. The firms we spoke with suggested that significant managerial resources are required to prepare a bid.[10]

[9] A firm's market share is defined as the ratio of the amount of the firm's total winning bids over the amount of total winning bids for all the contracts.

[10] First, firms must gather information about materials prices and find subcontractors for the project which is a time consuming activity. Also, firms must carefully study the project plans and specifications to calculate

Table 4 summarizes values for the 1st lowest bid (BID1), the 2nd lowest bid (BID2), and "money on the table" (BID21=BID2-BID1), when the total number of bids is at least 2. Money left on the table averages \$15,724 and has a maximum of \$352,174. All of the firms in our market are owner operated so whatever profits are made accrue directly to the firm's manager. If firms had complete information about competitors' costs, the amount of money that should be left on the table in equilibrium would be near zero. The amount of money left on the table is consistent with the presence of non-negligible private information about costs.

Based on the winning bids and bidding dates, we construct a new variable "CAP.", which is meant to measure each firm's capacity utilization level. A firm's capacity at a particular bidding time is defined as the ratio of the firm's used capacity (measured by the firm's total winning bids' amount up to that time) over the firm's total winning bids' amount in the entire season.[11]

Another generated variable is distance, which we construct using information about both the location of the firms and the location of the project.[12] For jobs covering several locations, we use the midpoints of the jobs to do the calculation. Table 5 summarizes the distance in miles between the project site and the winning firm (DIST1) up to the distance between the job and the firm submitting the 7th lowest bid (DIST7). Firms with shorter distances from project locations are more likely to win the job.[13] The average distance of the closest firm is 122.3 miles whereas the distances of firms who fail to win projects tends to be considerably higher.[14]

One important control variable for our analysis will be engineer's estimate. This is an estimate formed either by branches of the government or by consulting engineering firms. In speaking with engineers at

expected costs. Many of the projects we have studied are spread out over 10 to 100 miles from endpoint to endpoint. Firms must make a careful (and often tedious) calculation of anticipated transportation costs for the project. Finally, firms will need to get a bond for the project which is also time consuming. Standard references about construction bidding such as Park and Chapin (1992) suggest that bid preparation costs are on average one to two percent of total project costs.

[11] Note that as mentioned earlier, the season during which seal coating can take place lasts from late May to mid-September. So a season mentioned in the above definition starts on September 1 and ends on August 31 of the following calendar year. This measure of capacity was computed using the entire data base of bidding information even though in our econometric analysis we will focus on a subset of these projects.

[12] The calculation is facilitated by using Yahoo's map searching engine http://maps.yahoo.com/py/ddResults.py. Using city and state's names as input for both locations, the map searching engine gives distances automatically. Doing this manually would be too time consuming, so we wrote an "electronic spider" to do the job.

[13] The small mean distance for the firms with the 7th lowest bid is mainly due to the problem of too many missing observations. If those missing observations are recovered, we expect that DIST7 would have much higher mean.

[14] In our complete data set, it was not possible to find the location of all projects or all the firms that bid in these projects. Furthermore, some projects in our original data set had missing information that we could not supplement with information from the relevant branch of City, County, State or the Federal Government. However, we used the available information in calculating firms' capacities even though other parts of the observation might be incomplete.

Minnesota, North Dakota and South Dakota's Departments of Transportation, the engineers claimed that they formed the estimates by gathering information on materials prices, prevailing wage rates and other relevant cost information. The engineer's estimate is supposed to represent a "fair market value" for completion of the project. We found that estimates were available for 139 out of the 441 projects in the data set. Table 6 shows that the engineer's estimate is a useful control for project costs. The normalized winning bid is almost exactly 1 and has standard deviation of 0.1573.

Table 6 suggests that both location and capacity play an important role in bidding. Firms will bid higher the greater the distance to a project and greater capacity utilization implies a higher bid. On the other hand, Probit estimates also suggest that firms with greater capacity utilization and greater distance to a project are less likely to bid all else held constant.

Another important determinant of firm i's success in winning contracts is familiarity with local regulators and local material suppliers. In Table 7, we calculate, for each state and for each firm, the percentage of the firm's total dollar volume done in that state. Our results suggest that the majority of the firms in our data set work primarily in one state. This effect is present even after controlling for distance. For instance, firm 3 is located near the boundaries of Minnesota, North Dakota and South Dakota. Yet it does over 70 percent of its dollar volume of seal coating in South Dakota. Also, firms 6 is located near the Minnesota and South Dakota border. Yet firm 6 has won no contracts in South Dakota.

5.3 Reduced Form Bid Functions

Next, we estimate a set of reduced form bid functions to measure the relationship between a number of variables and the firms' observed bidding behavior. The variables we will use in these regression are as follows:

- $\text{BID}_{i,t}$: The amount bid by firm i on project t.
- EST_t: The estimated value of project t.
- $\text{DIST}_{i,t}$: Distance between the location of the firm and the project.
- $\text{LDIST}_{i,t}$: log($\text{DIST}_{i,t}$+1.0).
- $\text{CAP}_{i,t}$: Used capacity measure of firm i on project t.
- $\text{MAXP}_{i,t}$: Maximum percentage free capacity of all firms on project t, excluding i.
- $\text{MDIST}_{i,t}$: Minimum of distances of all firms on project t, excluding i.
- $\text{LMDIST}_{i,t}$: log($\text{MDIST}_{i,t}$+1.0).

- CON$_{i,t}$: The proportion of work done (by dollar volume) by firm i in the State where project t is located prior to the auction.

We assume that firm i's cost estimate for project t satisfies the following structural relationship:

$$\frac{c_{i,t}}{EST_t} = c(DIST_{i,t}, CAP_{i,t}, CON_{i,t}, \omega_i, \delta_t, \varepsilon_{it}) \tag{16}$$

Equation (16) implies that firm i's cost in auction t can be written as a function of its distance to the project, its backlog, the previous experience that firm i has in this market which we proxy for using $CON_{i,t}$, a firm i productivity shock ω_i, an auction t specific effect δ_t, and ε_{it}, an idiosyncratic shock to firm i that reflects private information she will have about her own costs. The results of Section 3 demonstrate that under certain simplifying assumptions about dynamic competition, a dynamic model with capacity constrained bidders is formally equivalent to a static model where the firm's cost is $c_i + V_{iL}(s) - V_{iW}(s)$, a sum of current project costs, c_i, plus a term $V_{iL}(s) - V_{iW}(s)$ that captures the option value of keeping free capacity. In practice, the measure of backlog $CAP_{i,t}$ will be a good proxy for $V_{iL}(s) - V_{iW}(s)$. Mapping the structural cost function back to the framework of Section 5 implies that $z_i = (DIST_{i,t}, CAP_{i,t}, CON_{i,t}, \omega_i, \delta_t)$.

Firm i's bid function should depend on the entire parameter vector $z = (z_1, ..., z_N)$. However, given the limited number of data points in our sample, it will not be possible to model the bid functions in a completely flexible fashion because z is a vector with $5 * N$ elements. We choose to include a firm's own distance, capacity and concentration. From our conversations with firms who actually bid in these auctions, we believe that the most important characteristics of the other firms to include in the reduced form bid function are the location of the closest competitor and the backlog of the competitor that has the most free capacity. To control for δ_t, we use fixed effects for the auction and to control for ω_i we use firm fixed effects for the largest 11 firms in the market. We are able to identify both our auction fixed effect and firm fixed effects because we do not use fixed effects for all of the firms. This implies that firms that are not the 11 largest have an identical productivity shock ω_i which is probably not a bad assumption in this industry.

Since there are 138 auctions, 11 main firms and one pooled group of non-main firms in our restricted data set, we have 137 auction dummies and 11 firm dummies. The set of regressors thus contains a constant (C), 148 dummy variables, own distance($DIST_{i,t}$), own capacity ($CAP_{i,t}$), maximal free capacity among competitors ($MAXP_{i,t}$), minimal distance among competitors ($MINDIST_{i,t}$), and the job concentration variable ($CON_{i,t}$). To take care of the heteroscedasticity problem, we take the ratio of the bid and the value (the engineer's estimate) as the dependent variable ($\frac{BID_{i,t}}{EST_t}$).

$$\frac{BID_{i,t}}{EST_t} = \beta_0 + \beta_1 LDIST_{i,t} + \beta_2 CAP_{i,t} + \beta_3 MAXP_{i,t} + \beta_4 LMDIST_{i,t} + \beta_5 CON_{i,t} + \varepsilon_{it} \tag{17}$$

The results from the regression, estimated using ordinary least squares, are (with t-statistics in parentheses):

Reduced Form Bid Function

Variable	OLS
C (Constant)	.6809
	(5.95)
$LDIST_{i,t}$ (Own distance)	.0404
	(3.45)
$CAP_{i,t}$ (Own used capacity)	.1677
	(8.51)
$MAXP_{i,t}$ (Maximal free capacity among rivals)	.0255
	(.713)
$LMDIST_{i,t}$ (Minimal distance among rivals)	.0240
	(1.81)
$CON_{i,t}$ (Job concentration)	-.0590
	(-1.866)
Sample Size	450
R^2	.8480

The regression also includes a fixed effect for each project t and fixed effect for each of the 11 largest firms in the market.

The results from our reduced form bid function are consistent with basic economic intuition. Firm i's bid is an increasing function of firm i's distance from the project site and firm i's capacity utilization. As firm i's distance increases so does i's cost. Our theory would lead us to expect a positive coefficient on own distance. The coefficient on $CAP_{i,t}$ is also positive and significant. As firm i's backlog increases, all else held constant, the option value of free capacity will increase because once i becomes completely capacity constrained, firm i will no long have a chance to bid on future projects. The coefficient on $CON_{i,t}$ is negative, indicating that if firm i has more prior experience in the state, firm i will tend to bid more aggressively.

Our reduced form bid function also produces results that are consistent with the strategic interactions implied by the asymmetric auction model. As the distance of firm $j \neq i$ increases or as the capacity utilization of firm $j \neq i$ increases competition will soften and firm i raises her bid. However, the reaction to $MAXPER_{i,t}$ is not significant at conventional levels.

5.4 Test for conditional independence

In this section, we report the result from the test of the conditional independence assumption $A1$ in Section 5. We use a reduced form bid function as in the previous subsection, however, we will allow the model to be more flexible. If firm i is one of the largest 11 firms in the industry we use equation (18) with firm varying coefficients to model i's bid function. If i is not one of the 11 largest firms in the industry we use equation (19) to model i's bid function. We pool equations (18) and (19) in the estimation and include auction fixed effects.

$$\frac{BID_{i,t}}{EST_t} = \beta_{0,i} + \beta_{1,i}LDIST_{i,t} + \beta_{2,i}CAP_{i,t} + \beta_{3,i}MAXP_{i,t} + \beta_{4,i}LMDIST_{i,t} + \beta_{5,i}CON_{i,t} + \varepsilon_{it} \tag{18}$$

$$\frac{BID_{i,t}}{EST_t} = \alpha_0 + \alpha_1 LDIST_{i,t} + \alpha_2 CAP_{i,t} + \alpha_3 MAXP_{i,t} + \alpha_4 LMDIST_{i,t} + \alpha_5 CON_{i,t} + \varepsilon_{it} \tag{19}$$

Suppose the coefficient of correlation between the residual to i's bid function and j's bid function, $\varepsilon_{i,t}$ and $\varepsilon_{j,t}$, is ρ_{ij}. The test of conditional independence is then equivalent to testing the following null hypothesis:

$$H_0 : \rho_{ij} = 0 \tag{20}$$

The Fisher test is applied to test the hypothesis (20) (see Bajari and Ye (2001) for the detail). Among all 23 pairs which have at least 4 simultaneous bids, the null hypothesis cannot be rejected except for four pairs of firms at 5% significance level. These four pairs are: (Firm 1, Firm 2), (Firm 2, Firm 4), (Firm 5, Firm 14), and (Firm 6, Firm 7). However, of these pairs, only the pair (Firm 2, Firm 4) bid against each other more than a handful of times. The Pairs (Firm 1, Firm 2), (Firm 5, Firm 14) and (Firm 6, Firm 7) bid against each other on average no more than two or three times a year in the data set.

5.5 Test for Exchangeability

In this section we use our regression model (18) and (19) to test whether the empirical distribution of bids is exchangeable. Exchangeability implies that capacities and distances should enter the firm's bid-value function in a "symmetric" way. Formally, in the reduced form bid function, let $\beta_{i1,}\beta_{i2}, \beta_{i3}, \beta_{i4}$ be the coefficients of $LDIST1$, $CAP1$, $MAXP$, $MINDIST$ for i, one of the largest 11 firms. Then exchangeability is equivalent to the following hypothesis:

$$H_0{:}\beta_{ik} = \beta_{jk} \text{ for all } i, j, i \neq j, \text{ and for all } k = 1, \cdots, 4. \tag{21}$$

We use the F-test to test for exchangeability. Let SSR_U and SSR_C be the sum of squared errors in the unconstrained and constrained models, respectively. Also let T be the number of observations ($T = 450$ in our data set) , m be the number of regressors, and n be the number of constraints implied by H_0. Then the statistic

$$F = \frac{(SSR_C - SSR_U)/n}{SSR_U/(T - m)} \tag{22}$$

is distributed as an F distribution with parameters ($n, T - m$) under null hypothesis. Note that the F-test is also a variation of the quasi-likelihood ratio test (QLR) on non-linear two- and three-stage least squares.

We conduct two tests of exchangeability in this section. The first set is to test exchangeability for the whole market, i.e., the constrained regression that pools all the 11 main firms together. The second set is to test the exchangeability on pairwise basis, i.e., the constrained regression pools two of the main firms together at each test (hence the number of constraints is 4). We perform this set of tests for each pair of firms with at least 4 simultaneous bids.

The exact result is reported in Bajari and Ye (2001). The result shows that for almost all the tests, we fail to reject the null hypothesis at 5% significance level. In fact, we only reject the null when we pool all the 11 main firms and when we pool Firm 2 and Firm 5.

5.6 Discussion

The results of our tests of exchangeability and conditional independence imply that there are five pairs of firms who exhibit bidding patterns that are not consistent with our characterization of competitive bidding. The four pairs (Firm 1, Firm 2), (Firm 2, Firm 4), (Firm 5, Firm 14), and (Firm 6, Firm 7) fail the conditional independence test and the pair (Firm2,Firm5) fails the exchangeability test. However, of these pairs, only the pairs (Firm 2, Firm 4) and (Firm2,Firm5) bid against each other more than a handful of times. The other three pairs (Firm 1, Firm 2), (Firm 5, Firm 14) and (Firm 6, Firm 7) bid against each other on average no more than two or three times a year in the data set. Also, according to industry participants, these firms function in different sub-markets and would have no reason to view each other as principal competitors. Therefore, the two pairs of firms that might be of concern to regulators are (Firm2,Firm5) and (Firm 2, Firm 4).

Overall, bidding in this industry appears to conform to the axioms A1-A5. This observation is important to policy makers since there is a history of bid-rigging in the seal coat industry. Several of the largest firms in the industry were colluding in the early to mid-1980's and paid damages for bid-rigging. Our analysis suggests that currently most bidding behavior in the industry is consistent with our model of competitive bidding.

We are aware of the following three limitations to our approach:

First, in both our tests for conditional independence and exchangeability, we need to use a correct functional form for the reduced form bid functions. In our analysis, we used a number of different functional forms for the reduced form bid function and the results in the tests for conditional independence and exchangeability were robust across these alternative specifications. The arguments of firm i's bid function is the vector z which has $4 * N$ arguments. Given that we have only 138 auctions in our data set, we can never be certain that our independence and exchangeability results were not influenced by a poor choice of functional form.

Second, our results might be incorrect if there are omitted variables. If there are elements of z that the firms see but which are not present in our data set our regression coefficients will be biased. In our analysis, we use fixed effects for each contract and for the largest 11 firms. Therefore, we should be most worried about omitted variables that are elements of z but which are not co-linear with the firm or contract fixed effects. This could happen, for instance, when there are 3 firms and firm1 and firm 2 always use a quote from a particular subcontractor when computing their cost estimates while firm 3 does not. If this quote is publicly observed, it will then induce positive correlation between the residuals to the bid functions of firms 1 and 2.[15] A similar critique could be made of our test of exchangeability. Omitted cost variables could lead us to falsely conclude that firm 2 and firm 5 fail to have an exchangeable distribution of bids.

Third, if a sophisticated cartel is operating in this market, then, as we mentioned in Section 5, the cartel could satisfy assumptions A1-A5 by generating phony bids in a clever fashion. Therefore, from our tests, we will not be able to identify whether those firms who passed the conditional independence and

[15] However, in our test for conditional independence, we found that the residual between firm 2 and 4 are negatively correlated. If this is due to omitted cost variables, it must be the case that the omitted variable must induce negative correlation between the costs of these two firms. So far, we have not been able to come up with a scenario that would generate this type of cost shock. However, if firm 2 and firm 4 engaged in a scheme of submitting phony bids, this might induce a negative correlation in the residuals since phony bidding implies that when firm 2 bids high firm 4 must bid low.

exchangeability tests are competitive or are smart colluders. In recent empirical papers that document cartel behavior such as Porter and Zona (1993,1999) and Pesendorfer (1995) the authors know from court records and investigations the identity of the cartel. In all these papers, both exchangeability and conditional independence fail. To the best of our knowledge, there is no documented case of cartel bidding where the cartel intentionally submitted phony bids that were both conditionally independent and exchangeable.

Therefore, while collusion is certainly one reason why the tests of conditional independence and exchangeability fail, it is certainly not the only reason. The restrictions of the competitive bidding model are quite stringent and there are a variety of reasons, other than collusion, for our failure to accept the competitive model. Therefore, we shall interpret firms (2,4) and (2,5) as a candidate set of cartels.

In the following section, we shall approach the problem of testing for collusion from an alternative point of view. Our approach shall be to estimate three structural models of industry equilibrium. The first model M1 shall correspond to the case of no collusion, the second model M2 is a model in which firms 2 and 4 collude to maximize joint profits and the third model M3 is a model in which firms 2 and 5 collude to maximize joint profits. We shall elicit the prior beliefs of industry experts about costs and we shall use the tools of statistical decision theory to make a decision between these three alternative models.

6 A Structural Econometric Models.

In this section, we describe how to estimate three alternative structural models of industry equilibrium. We shall refer to the three models as $M1$, $M2$, and $M3$. In the first model all firms determine their bids competitively as in the model outlined in sections 2 and 3. In the second model, firms 2 and 4 collude in an efficient manner. Before the bidding begins, firm 2 and firm 4 compute cost estimates $c_{2,t}$ and $c_{4,t}$ respectively for project t. The cartel operates efficiently so that only the firm with the lowest cost submits a bid. That is, the cost to the cartel is $c_{c,t} = \min\{c_{2,t}, c_{4,t}\}$. The other cartel member either refrains from bidding or submits a "phantom" bid. The cost distribution for the cartel is the minimum of $c_{2,t}$ and $c_{4,t}$. This distribution has a probability density function $f_2(c_{c,t}|z,\theta)(1 - F_4(c_{c,t}|z,\theta)) + f_4(c_{c,t}|z,\theta)(1 - F_2(c_{c,t}|z,\theta))$. In model $M2$, we can treat the cartel as a single firm who has cost $c_{c,t}$. In computing expected utility, it is sufficient for non-cartel firms to only concern themselves with the low bid among all of the cartel members. Therefore, as we mentioned in section 2.2, efficient cartels are a special case of our asymmetric auction model. The third model, $M3$ is analogous to $M2$, with now firm 2 and firm 5 being the cartel members.

We believe that the assumption of an efficient cartel is reasonable for the seal coat industry in the Mid-west. According to industry officials and insiders we have spoken with, before the early 1980's, collusion did occur (and was prosecuted) in the seal coat industry. Firms frequently made side payments to each other. Many side payments were made directly in cash. Another mechanism for making side payments was through the use of falsified invoices. For instance, firm A would rent equipment from firm B on paper. However, the equipment would not actually be rented in practice. The payment that firm A made to firm B for this phantom equipment rental would serve as a sidepayment.

There could be no way to scientifically assess the most appropriate mechanism for collusion or the most common mechanism for collusion when it was present in the industry. However, from our conversations with industry regulators and insiders, we believe that a model of efficient collusion is at least a plausible hypothesis. Furthermore, as we shall see in the next section, it leads to a tractable econometric specification whereas many alternative models of collusion may not be as straightforward to structurally estimate.

Let $g_i(b;\theta,z,M)$ be the p.d.f. of firm i's observed bids and $G_i(b;\theta,z,M)$ be the distribution function in the model $M = M1, M2$ or $M3$. In the competitive model, firm i makes a best response to the distribution of the bids of all other firms. In the collusive model, $M2$, a non-cartel member, firm i needs to make a best response to only the minimum bid of the cartel and the bids of all the other non-cartel firms. In this case, $g_i(b;z,\theta,M2)$ is the minimum of the bid of firm 2 and of firm 4 since any bid that is not the minimum is a "phony" bid that the cartel has submitted in order to deceive regulators. Similarly, in the collusive model $M3$, we treat the cartel as a single firm who submits a bid of the minimum of firm 2's bid and firm 5's bid.

If we assume that the firms in the industry are profit maximizing, then firm i's private information in model M, which we denote as c_i^M must satisfy:

$$c_i^M = b - \frac{1}{\sum_{j\neq i} \frac{g_i(b;\theta,z,M)}{1-G_i(b;\theta,z,M)}}. \tag{23}$$

This relationship must hold for models $m = M1$, $M2$, and $M3$.

Let $\widehat{g}_i(b;\theta,z,M)$ and $\widehat{G}_i(b;\theta,z,M)$ denote the estimated values of $g_i(b;\theta,z,M)$ and $G_i(b;\theta,z,M)$. Let $c_{i,t}^{M,EST}$ denote an estimate of the cost of firm i in project t and let $\mathbf{c}^{M,EST}$ denote the estimated vector of the $c_{i,t}^{M,EST}$. Following Guerre, Perrigne and Vuong (1999) we estimate firm i's private information by evaluating the formula (23) and the empirical estimates of $\widehat{g}_i(b;\theta,z,M)$ and $\widehat{G}_i(b;\theta,z,M)$ of $g_i(b;\theta,z,M)$ and $G_i(b;\theta,z,M)$:

$$c_{i,t}^{M,EST} = b - \frac{1}{\sum_{j\neq i} \frac{\widehat{g}_i(b;\theta,z,M)}{1-\widehat{G}_i(b;\theta,z,M)}}. \tag{24}$$

In what follows, we shall briefly outline the particulars of our approach to the estimation of $c_{i,t}^{M,EST}$. Recovering an estimate of an agent's private information from a first order condition such as (24) is now common practice in the empirical auctions literature. For a complete discussion of issues related to estimation of this equation, please refer to Guerre, Perrigne and Vuong (1999) and Pesendorfer and Jofre-Bonet (2000).

In our problem, the distributions $g_i(b;\theta,z,M)$ depend on a large number of parameters and covariates. The empirical analysis of the previous section indicates that it is important to control for unobserved heterogeneity at the level of the firm and at the level of the project. Obviously, kernel estimation of $g_i(b;\theta,z,M)$ is not practical in our application. Therefore, as a simplifying assumption, we shall assume that it is possible to model the distribution of bids as a linear function of the covariates and an additively seperable error term as in equations (18) and (19). In the competitive model, we include all of the bids. In $M2$ we only include the lower of firm 2's bid and firm 4's bid since if the cartel is profit maximizing, bidders are making a best response to the minimum bid of the cartel. Similarly in model $M3$ we only include the minimum bid of firm 2 and firm 5. In order to estimate the implied costs corresponding to the three models, we must estimate $\widehat{g}_i(b;\theta,z,M)$ and $\widehat{G}_i(b;\theta,z,M)$ for $M = M1, M2, M3$.

As a first stage, we estimate equations (18) and (19). As we mentioned previously, in the case of $M2$ we estimate the cartel's bid as the minimum of the bids of firm 2 and firm4 and in the case of $M3$ the cartel's bid is the minimum of the bid of firm 2 and firm 5. In all three models, we shall assume that the distribution of firm i's bid satisfies:

$$b_{i,t} = (\text{o.l.s. fitted value of } b_{i,t}) + \varepsilon_{i,t}$$

where $\varepsilon_{i,t}$ is modeled as a mixture of four normal distributions. We use the mixture of normals specification because it allows us to flexibly model the distribution of $b_{i,t}$ and be able to compute $\widehat{g}_i(b;\theta,z,M)$ and $\widehat{G}_i(b;\theta,z,M)$ in a simple manner. To estimate the mixture of normals distribution for models $M =$

$M1, M2$, and $M3$ we used maximum likelihood estimation. For models $M1$, $M2$, and $M3$ the estimated distributions of the residuals appear to closely match the empirical distributions of the residuals.

In general, the vector of costs $c^{M,EST}$ will not be estimated exactly, instead the econometrician will have uncertainty about the value of $c^{M,EST}$. It is possible, using standard econometric methods to derive a joint density of the $c^{M,EST}$ and we shall refer to this density as $p(c^{M,EST})$.

7 Testing For Collusion: A Bayesian Framework.

In the previous section, we demonstrated how to estimate the joint distribution of the latent costs $p(c^{M,EST})$ for three alternative models of industry equilibrium. In this section, we will suggest a Bayesian framework for computing a posterior probability for each of our three models.

Each of our three models of industry equilibrium implies a different distribution over structural cost parameters. Let $p(c_{i,t}|z_t, \theta_i, M)$ for $M = M1, M2$, and $M3$ denote a probability density function for the cost distribution of firm i in auction t conditional on covariate z_t, parameters θ_i and model M. For instance, in the collusive model $M2$, the cost distribution for the cartel is $f_2(c_{c,t}|z,\theta)(1 - F_4(c_{c,t}|z,\theta))+$ $f_4(c_{c,t}|z,\theta)(1 - F_2(c_{c,t}|z,\theta))$. In model $M3$, the probability density function of the cost distribution of the cartel is $f_2(c_{c,t}|z,\theta)(1 - F_5(c_{c,t}|z,\theta))+$ $f_5(c_{c,t}|z,\theta)(1 - F_2(c_{c,t}|z,\theta))$. Using the fact that bids are independent across auctions, it is straightforward to compute $p(c^{M,EST}|z_t, \theta, M)$, the probability density function for all the costs in all the auctions:

$$p(c^{M,EST}|z_t, \theta, M) = \prod_{t=1}^{T} \prod_{i} p(c_{i,t}|z_t, \theta_i, M) \tag{25}$$

In Bayesian Statistics, it is necessary to specify a prior distribution over structural parameters. As we shall explain below, we shall elicit the prior distribution of the parameters, $p(\theta)$ from industry experts. Given the likelihood function for each model (25), we define the marginalized likelihood of model M as follows:

$$ML_M = \int \int p(c^{M,EST}|M, \theta)\, p(\theta) p(c^{M,EST}) \tag{26}$$

That is, the marginalized likelihood is the expected value of the likelihood function after margalizing out the distribution of θ and the distribution of $c^{M,EST}$.

In a Bayesian framework, we must specify a prior probability for each model. Let $p(M_i)$ denote the prior belief (probability) that $M = M_i$, and let ML_i denote the value of likelihood function conditional on $M = M_i$. Then according to the Bayesian Theorem, we can compute the posterior probability of $M = M_i$ which we denote as $p^*(M_i)$ as follows:

$$p^*(M_i) = \frac{p(M_i) \cdot ML_i}{p(M_1) \cdot ML_1 + p(M_2) \cdot ML_2 + p(M_3) \cdot ML_3} \tag{27}$$

In the following section, we describe how we elicit the prior distribution over structural cost parameters $p(\theta)$ from industry experts. Given this distribution and a (flexible) functional form for the likelihood $p(c^{M,EST}|z_t, \theta, M)$, to compute the posterior probability of each model, we only need to evaluate the integral (26) and apply Bayes Theorem as in equation (27). In the following sub-section, we describe how we elicit $p(\theta)$ from industry experts and how we flexibly specify the distribution of $p(c_{i,t}|z_t, \theta_i, M)$.

7.1 Eliciting Prior Beliefs.

In this section, we specify a flexible functional form for the distribution of $p(c_{i,t}|z_t, \theta_i, M)$ and we demonstrate how we elicit a prior distribution of beliefs $p(\theta)$ from industry experts. We will model the cost (normalized by the estimate) as the sum of a fitted value plus a flexible error term. That is,

$$\frac{c_{i,t}}{EST_t} = \frac{\overline{c}_{i,t}(\theta, z_t)}{EST_t} + \eta_{it} \tag{28}$$

In equation (28), the normalized cost is a function of a fitted value $\overline{c}_{i,t}(\theta, z_t)$ which depends on the structural parameters θ, a set of auction t co-variates z_t, and an idiosyncratic error term η_{it}. In order to model the error distribution flexibly, we allow it to be a mixture of four normal distributions, that is:

$$\eta_{it} \sim \pi_1 N(\mu_1, \sigma_1) + \pi_2 N(\mu_2, \sigma_2) + \pi_3 N(\mu_3, \sigma_3) + \pi_4 N(\mu_4, \sigma_4) \tag{29}$$

What remains to be specified is how we will use the expert information to compute $\overline{c}_{i,t}(\theta, z_t)$ a fitted value of the cost that is a function of the parameters and how we will determine the vector of parameters $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)$.

The approach we used to elicit the expert's beliefs over the structural parameters was first to elicit a probability distribuiton over markups. We believed that it would be hard for our expert to articulate to us the value of the auction fixed effect in project t. However, we did believe that the experts would have fairly thoughtful views about markups in the industry. Let $m_{i,t}$ be the markup of player i in auction t and let $p(m_{i,t})$ be the prior density function. After extensive discussions with the experts, we elicited their beliefs about the 25, 50, 75 and 99th percentiles of the distribution of markups of auctions in our sample. The views of the experts were quite close and we average their views in the equations below.

$$\begin{aligned} 25^{th} \text{ percentile} &= 3\% \\ 50^{th} \text{ percentile} &= 5\% \\ 75^{th} \text{ percentile} &= 7\% \\ 99^{th} \text{ percentile} &= 15\% \end{aligned} \tag{30}$$

Given the beliefs (30), we assume that markups were independently and identically distributed according to the following distribution:

$$p(m_{i,t}) = \begin{cases} 8.3333 \text{ if } 0 < m_{i,t} < 0.03 \\ 12.5 \text{ if } 0.03 < m_{i,t} < 0.05 \\ 12.5 \text{ if } 0.05 < m_{i,t} < 0.07 \\ 3.125 \text{ if } 0.07 < m_{i,t} < 0.15 \end{cases} \tag{31}$$

It can be easily verified that (31) agrees with the percentiles in equation (30).

If the markup on project t is $m_{i,t}$ and the observed bid in project t is $b_{i,t}$, it follows that the cost $c_{i,t}$ must satisfy:

$$c_{i,t} = (1 - m_{i,t}) * b_{i,t} \tag{32}$$

Next, given the probability density function $p(m_{i,t})$ in equation (30), draw a random vector of markup $\mathbf{m}_T$ for all projects $t = 1, ..., T$ and all bidders i who bid in the t^{th} project. Then using equation (32) we can compute a set of latent costs $c_{i,t}$. We then estimate the following equations using ordinary least squares to find a fitted value $\overline{c}_{i,t}(\theta, z_t)$ for all i and t.

$$\frac{c_{i,t}}{EST_t} = \beta_{0,i} + \beta_{1,i} LDIST_{i,t} + \beta_{2,i} CAP_{i,t} + \beta_{3,i} LMDIST_{i,t} + \beta_{4,i} CON_{i,t} + \varepsilon_{it} \tag{33}$$

$$\frac{c_{i,t}}{EST_t} = \alpha_0 + \alpha_1 LDIST_{i,t} + \alpha_2 CAP_{i,t} + \alpha_3 LMDIST_{i,t} + \alpha_4 CON_{i,t} + \varepsilon_{it} \tag{34}$$

We include firm and auction specific fixed effects in order to capture unobserved heterogeneity across projects and firms. Just as in our estimation of the bid functions in equations (18) and (19), we use equation (33) for the main firms allowing the coefficients to be firm specific and equation (34) for the non-main firms. From equations (33) and (34) we can also find an estimated value $\widehat{\sigma}$ for the standard deviation of ε_{it}. We must next compute the value of $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)$. We shall assume that the parameters of the mixture of normals are drawn uniformly from all of those distributions satisfying:

$$\begin{aligned} \pi_1\mu_1 + \pi_2\mu_2 + \pi_3\mu_3 + \pi_4\mu_4 &= 0 \\ \pi_1\sigma_1^2 + \pi_2\sigma_2^2 + \pi_3\sigma_3^2 + \pi_4\sigma_4^2 &= \widehat{\sigma}^2 \\ \pi_1 + \pi_2 + p_3 + \pi_4 &= 1 \\ -0.2 &\leq \mu_i \leq 0.2 \\ 0.01 &\leq \sigma_i \leq 0.2 \end{aligned} \tag{35}$$

We assume that $-0.2 \leq \mu_i \leq 0.2$and $0.01 \leq \sigma \leq 0.2$ so that the support of the $\mu_i$and $\sigma_i$are compact. We believe that this assumption is appropriate because we have found that the fitted residuals from equations (33) and (34) fall outside the interval $(-.2, .2)$ very infrequently. We also found that $\widehat{\sigma}$ is seldom above 0.2 and below 0.01, therefore the restriction that $0.01 \leq \sigma_i \leq 0.2$ is of small consequence.

To summarize, given a random vector of markups $\mathbf{m}_T$, we compute a vector of fitted costs $\overline{c}_{i,t}$ using equations (33) and (34) and we draw the values of $\boldsymbol{\pi} = (\pi_1, \pi_2, \pi_3, \pi_4)$, $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3, \mu_4)$ and $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3, \sigma_4)$ using equation (35). The likelihood for a vector of costs for the competitive model $M1$ conditional on θ is then:

$$\begin{aligned} p(\frac{c_{i,t}}{EST_t}|\boldsymbol{\theta}, M1) &= \pi_1 N(\frac{c_{i,t} - \overline{c}_{i,t}}{EST_t}, \mu_1, \sigma_1) + \pi_2 N(\frac{c_{i,t} - \overline{c}_{i,t}}{EST_t}, \mu_2, \sigma_2) \\ &\quad + \pi_3 N(\frac{c_{i,t} - \overline{c}_{i,t}}{EST_t}, \mu_3, \sigma_3) + \pi_4 N(\frac{c_{i,t} - \overline{c}_{i,t}}{EST_t}, \mu_4, \sigma_4) \end{aligned} \tag{36}$$

The likelihood for the vector of costs from the model $M2$ and $M3$ are computed similarly except we assume that the cost for the cartel is the order statistic of its member's cost.

7.2 Putting It All Together.

We now summarize all of the steps needed to compute a simulated value of the likelihood function for model M.

1. Draw a random vector $\mathbf{c}_T^{M,EST}$ over the distribution of cost estimates from the distribution $p(\mathbf{c}_T^{M,EST})$ that was estimated in Section 6.
2. Draw a random vector $\mathbf{m}_T$ from the distribution (31).
3. Given $\mathbf{m}_T$, compute fitted values $\overline{c}_{i,t}$ for all auctions t and all firms i.
4. Draw the parameters of the mixture of normal distribution $\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma}$ according to the distribution in the equation (35).
5. Compute the value of the likelihood function, $p(\mathbf{c}_T|\boldsymbol{\pi}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \alpha, \beta, M1)$ as in equation (36).

Using standard methods for computing an integral using simulation, it is possible to evaluate the integral (31). See Judd (2000) for a discussion of standard methods for numerical integration.

Using equation (27), we can compute the posterior beliefs for the 3 competing models. The posterior beliefs can be useful for decision making. Suppose for instance, these beliefs will be used in order to determine whether to investigate and prosecute the candidate cartel. Suppose that the anti-trust authority possesses a loss function $u(a, M)$. Here a is the action that the anti-trust authority must take, for example, to prosecute or not and M takes on values $M1$,$M2$ and $M3$ corresponding to whether or not there is a cartel in the industry. The anti-trust authority's expected loss of an action a, conditional on the data is

$$p^*(M1)u(a, M1) + p^*(M2)u(a, M2) + p^*(M3)u(a, M3) \tag{37}$$

Given the posterior probabilities over models $p^*(M1)$, $p^*(M2)$, $p^*(M3)$, it is then straightforward for the anti-trust official to choose a loss minimizing action.

A difference between using equation (37) and classical methods for model selection is that equation (27) explicitly takes into account the preferences of the decision maker. For instance, an anti-trust authority may not want to choose to investigate firms 2 and 4 just because $M2$ is the largest. We might expect the anti-trust authority to only want to investigate when the probability of collusion is quite large.

As an aside, we note that it is straightforward to simulate the posterior distribution of the structural parameters if we impose the prior beliefs elicited by our industry experts since our specification is just a linear model with a flexible error term. This is an alternative to the methods proposed by Guerre, Perrigne and Vuong (1999). In small samples, such as in our data set, there is often not enough data to implement non-parametric procedures. We would expect a Bayesian approach to have much better small sample properties. For an introduction to Bayesian analysis of many standard models, see Gelman et al. (1995).

8 Results.

In what follows, we shall present the results of the estimation procedures described in Sections 6 and 7. We began by estimating the models $M1$, $M2$ and $M3$. Using equation (24) we can compute an implied distribution of markups under the three alternative models which we summarize in Table 9 below:

Table 9: Distribution of Markups Under Alternative Models.

Percentile	M1	M2	M3
10	0.01229	0.01273	0.0114
20	0.01597	0.01818	0.0182
30	0.02077	0.02422	0.0256
40	0.02536	0.03201	0.0343
50	0.03329	0.04126	0.0447
60	0.04227	0.05434	0.0584
70	0.05692	0.0754	0.0930
80	0.1000	0.1621	0.1756
90	0.2381	0.3354	0.5826

It is clear from Table 9 that markups are consistently higher under the collusive models M2 and M3. The difference between these models becomes particularly striking near the 80th and 90th percentiles. In model M2, 20 percent of the projects have a markup in excess of 15 percent and 10 percent of the projects have markups in excess of 33 percent. In model M3, 20 percent of the projects have markups in excess of 17 percent and 10 percent have markups in excess of 58 percent. Models M2 and M3 generate a much higher frequency of markups over 15 percent than the prior beliefs of our industry experts.

Another difference between models M1, M2 and M3 is the percentages of costs that are found to be negative.

Table 10: Percent of Cost Estimates Found to be Negative.

Model	Percentage Negative
M1	3%
M2	5%
M3	7%

It might come as a surprise to some readers to learn that some bids cannot be rationalized by a positive cost draw. To understand why this might occur, ask what would the bid be by a firm who has a cost of zero. Since costs in our model are assumed to be positive, the firm is certain to win the auction. Would the firm want to bid zero? Certainly not, because with a bid of zero he receives a profit of zero. If the firm bids above zero, she receives positive expected utility. Therefore, the lowest bid consistent with equilibrium is bounded away from zero. This follows formally from Theorem 2.

However, firms in the real world will do sometimes submit bids that are not consistent with our equilibrium model. There are several reasons why this might occur. The first is that the firm simply makes a mistake in calculating its bid. While this occurs infrequently, it is by no means unheard of. The second reason is that the firm does in fact have a negative cost for doing the project. Finding and retaining qualified employees in the construction industry is a challenge for managers. If the firm does not find a sufficient amount of work to occupy its seal coating crew, it risks losing these employees. Therefore, the firm is sometimes willing to bid jobs at a lost in order to retain its employees. A similar phenomenon is documented by Benkard (2001) who argues that Lockheed's production of the L1011 airplane at a loss for 11 years was due to learning curves in the production process. Another third reason that we might infer that costs are negative is because our model is misspecified. In fact, we are almost certain that our model is not an exact representation of the data generating process in the industry. (We do believe, however, that it is a useful representation.)

We believe that it is inappropriate for us to censor the observations in which costs are negative. We interpret the fact that model M3 is unable to assign positive costs to 7 percent of the bids as an evidence in favor of model M1 over M3. However, we do need to modify our likelihood function slightly to allow for negative costs. In computing the marginalized likelihoods, we modified the model of the previous section so that, with 95 percent probability, the likelihood function was constructed as in Section 7. With 5 percent probability, however, we assume that $\frac{c_{i,t}}{EST_t}$ is uniformly distributed between 1.0 and -10. This simple modification of the likelihood function allows costs to be negative. We are confident that our results are robust to alternative specifications that mix the model of Section 7 with a small probability that allows for negative costs in order to account for mis-specification.

Using the procedure described in the previous section, we compute the marginalized likelihoods associated with each of the three competing models. These computations were done using approximately 5 hours of CPU time on a Sun Ultra 60 workstation. In Table 11 we summarize our results.[16]

Table 11: Marginalized Likelihoods and Model Posterior Probabilities.

Model	Expected Value of Log Likelihood	Log of Marginalized Likelihood	Posterior Probability
M1	62.169 (61.916,62.423)	163.4 (161.89,163.9)	1.00
M2	-320.41 (–320,58,320.24)	-185.58 (–186.8,184.75)	0.0
M3	-243.15 (–250.75,235.56)	-186.6 (–187.81,184.55)	0.0

As we can see from Table 11, the competitive model M1 strongly dominates the other two alternatives. This is because, as we saw in Tables 9 and 10, it more closely agrees with the beliefs of our industry insider about markups in the industry. In particular, according to our industry expert, for all practical purposes, none of the markups exceed 15 percent. The collusive models imply a degree of market power that is not consistent with a priori beliefs about the industry. From the results in Table 11, it is clear that the posterior probability (assuming a uniform prior over all models) of the competitive model M1 far exceeds the other models.

9 Conclusion.

In empirical industrial organization, almost no formal use is made of expert opinion in estimation and testing. A contribution of this paper is to make formal use of the a priori beliefs of industry experts in order to decide between non-nested models of industry equilibrium. We believe that using the prior beliefs of our industry experts is a powerful tool that is potentially useful in many applied problems. In the seal coat industry, there are nearly 100 firms who bid at some point in our data set. However, only 6 of the firms have a market share that exceeds 5 percent. In such a competitive environment, we believe that the beliefs of a leading firm will contain important information about markups.

The views of industry experts are particularly useful when the economist only has a small sample. Economists are often forced to make decisions between alternative models or to make predictions with only a handful of data points. We believe that in many cases, using the a priori beliefs of experts will result in improved small sample properties. It is possible to extend the approach we have developed to other problems in industrial organization. If it is possible to elicit a prior distribution over parameters or induce a prior distribution over parameters (in our case by using beliefs about markups) the tools of Bayesian statistics can be used to form a posterior distribution over parameters that take into account the beliefs of

[16] All confidence intervals are based on 500,000 simulated values of the likelihood function.

industry experts.

We believe that the methods proposed in this research, when taken together with our companion paper Bajari and Ye (2001) is a useful methodology to decide between competition and collusion in procurement auctions. We have analyzed a model that allows for asymmetric bidders and non-trivial dynamics that we believe are important strategic considerations in many procurements. We stated a set of conditions that are both necessary and sufficient for a distribution of bids to arise from competitive bidding in the asymmetric auction model. Two of these conditions, conditional independence and exchangeability can be tested in a straightforward fashion. We also argued that deciding between competitive and collusive bidding can be naturally formulated as a statistical decision problem. We developed techniques for structurally estimating asymmetric auction models and for computing the posterior probability of competitive and collusive models. While no empirical techniques for detecting collusion are likely to be flawless, we believe that the tests we propose, taken together can be a useful first step in detecting suspicious bidding patterns.

10 Tables.

Table 1: Identities of main firms

Firm ID	Name of the Company	Firm ID	Name of the Company
1	Allied Blacktop Co.	11	Asphalt Surfacing Co.
2	Astech	12	Bechtold Paving
3	Bituminous Paving Inc.	14	Border States Paving Inc.
4	Lindteigen Constr Co. Inc.	17	Mayo Constr Co. Inc.
5	McLaughlin & Schulz Inc.	20	Northern Improvement
6	Morris Sealcoat & Trucking Inc.	21	Camas Minndak Inc.
7	Pearson Bros Inc.	22	Central Specialty
8	Caldwell Asphalt Co.	23	Flickertail Paving & Supply
9	Hills Materials Co.	25	Topkote Inc.

Table 2: Bidding activities of main firms

Firm ID	No.of wins	Avg. bid	% mkt. share	No. Participation	% of participation
1	92	82790	8.2	145	29.3
2	102	191953	21.1	331	66.9
3	20	363565	7.8	69	14.0
4	35	241872	9.1	114	23.0
5	29	283323	8.9	170	34.3
6	40	77423	3.3	84	17.0
7	45	62085	3.0	121	24.4
8	16	87231	1.5	134	27.1
9	10	237408	2.6	14	2.8
11	4	328224	1.4	28	5.7
12	3	317788	1.0	8	1.6
14	4	754019	3.2	25	5.1
17	5	1018578	5.5	8	1.6
20	13	355455	5.0	38	7.7
21	2	903918	1.9	5	1.0
22	2	903953	2.0	8	1.6
23	2	439619	1.0	4	0.8
25	3	382012	1.2	13	2.6
Total	427		87.7		

Average bid is the average of all bids that a particular firm submitted, No. Participation is the total number of bids that a particular firm submitted and % participation is the fraction of seal coat contracts a particular firm bid for.

Table 3: Bid concentration

Number of bids	1	2	3	4	5	6	7
Number of contracts	29	87	190	118	44	22	5

Table 4: First and second lowest bids

	Observation	Mean	Std Dev.	Min	Max
BID1	466	191355	227427	3893	1772168
BID2	466	207079	244897	4679	1959928
BID21	466	15724	29918	33	352174

Table 5: Distances (in miles)

	Mean	Min	Max		Mean	Min	Max
DIST1	122.3	0	584.2	DIST5	160.3	13	555.2
DIST2	151.9	0	585.2	DIST6	177.9	63	484.4
DIST3	177.9	0	637.6	DIST7	91	44	128.9
DIST4	166.4	11.2	608.6				

Table 6: Summary statistics for restricted data set

Variable	No. Obs	Mean	Std	Min	Max
Winning Bid	441	175,000	210,000	3893	1,732,500
Markup: (Winning Bid-Estimate)/Estimate	139	0.0031	0.1573	-0.3338	0.5421
Normalized Bid:Winning Bid/Estimate	139	1.0031	0.1573	0.6662	1.5421
Money on the Table: 2nd Bid-1st bid	134	15,748	19,241	209	103,481
Normalized Money on the Table: (1st Bid-2nd Bid)/Est	134	0.0776	0.0888	0.0014	0.5099
Number of Bidders	139	3.280	1.0357	1	6
Distance of Winning Firm	134	188.67	141.51	0	584.2
Distance of Second Highest Bidder	134	213.75	152.01	0	555
Capacity of Winning Bidder	131	0.3376	0.3160	0	0.9597
Capacity of Second Bidder	131	0.4326	0.3435	0	1

Table 7: Concetration of Firm Activity by State.

Firm	MN. Concentration	ND. Concentration	SD Concentration
1	1	0	0
2	0.2781	0.7218	0
3	0	0.2377	0.7623
4	0	1	0
5	0.1246	0.5338	0.3414
6	0.8195	0.1804	0
7	0.9572	0.0427	0
8	0.7290	0.2709	0
11	0	0	1
14	0	1	0
20	0	1	0

Table 8: Reduced form bid function

Coefficient	Estimate	Coefficient	Estimate
β_1	.046923 (3.64072)	α_1	.031346 (1.99253)
β_2	.174253 (8.42785)	α_2	.153158 (3.29708)
β_3	.040483 (.555484)	α_3	.036645 (.503462)
β_4	.020369 (1.45323)	α_5	.034269 (2.12646)
R Square	.849138		

Bibliography.

[1] Armantier, Olivier, Florens, Jean-Pierre, Richard, Jean-Francois, "Empirical Game Theoretic Models: Constrained Equilibrium and Simulation", Unpublished Working Paper, University of Pittsburgh 1997.

[2] Bajari, Patrick "The First-Price Auction with Asymmetric Bidders: Theory and Applications.", 1997 University of Minnesota Ph.D. Thesis.

[3] Bajari, Patrick, "Comparing Competition and Collusion: A Numerical Approach", Stanford University Working Paper, forthcoming in *Economic Theory.*

[4] Bajari, Patrick and and Lixin Ye, "Competition Versus Collusion in Procurement Auctions: Identification and Testing", 2001 Stanford University Working Paper.

[5] Baldwin, Laura, Robert Marshall and Jean-Francois Richard, "Bidder Collusion at Forest Service Timber Sales", 0022-3808 v105, n4 (August 1997): 657-99 Journal of Political Economy.

[6] Benkard, C. Lanier, "Learning and Forgetting: The Dynamics of Commercial Aircraft Production,"American Economic Review, (September 2000)

[7] K. E. Brennan, S. L. Campbell and L. R. Petzold,(1989)"The Numerical Solution of Initial Value Problems in Differential-Algebraic Equations", Elsevier Science Publishing Co.

[8] Deltas, George, "Structural Econometric Analysis of Bidding for Government Construction Contracts in Conneticut." 1996 Yale University Working Paper.

[9] Froeb, Luke; Tschantz, Steven; Crooke, Philip, "Mergers in First vs. Second-Price Asymmetric Auctions", 1997 Owen Working Paper.

[10] Gear, William, "Numerical Initial Value Problems in Ordinary Differential Equations", Prentice Hall, 1971.

[11] Gelman, Andrew; Carlin, John B.; Stern, Hal S.; Rubin, Donald B. (1995) "Bayesian Data Analysis", CRC Press.

[12] Guerre, E., Perrigne, I., and Vuong Q. "Optimal Nonparametric Estimation of First-Price Auctions" Econometrica, Volume 68, No. 3 (2000)

[13] Griesmer, J.M, Levitan, R.E., and M. Shubik, "Towards a study of bidding processes, Part Four: Games with unknown costs", Naval Research Quarterly, 14, 415-33 (1967).

[14] Hendricks, K. and Porter, R.H. "Collusion in Auctions", Annales d'Economie et de Statistique, No.

15/16 (1989) pp.217-230.

[15] Hong, H. and Shum, M., "Structural Estimation of First-Price Auction Models", Unpublished Working Paper, Princeton University (1999).

[16] Laffont, Jean-Jaques, Ossard, Herve and Vuong, Quang (1995) "Econometrics of First-Price Auctions", Econometrica 63:953-981.

[17] Lebrun, Bernard, "Existence of an Equilibrium in First Price Auctions", 0938-2259 v7, n3 (April 1996): 421-43 Economic Theory.

[18] Lebrun, Bernard "First-Price Auction in the Asymmetric N Bidder Case: Part I", 1995a Mimeo.

[19] Lebrun, Bernard "First-Price Auction in the Asymmetric N Bidder Case: Part II", 1995b Mimeo.

[20] Lizzeri, Alessandro and Persico, Nicola "Existence and Uniqueness of Equilibrium in First-Price Auction and War of Attrition with Affiliated Values", 1995 Northwestern Working Paper No 1120.

[21] Marshall, Robert C.; Muerer, M.J.; Richard, J.F.; Stromquist, W.; "Numerical Analysis of Asymmetric First-Price Auctions", Games and Economic Behavior, 7, 1994, 193-220.

[22] Maskin, Eric and Riley, John "Asymmetric Auctions". 1993 Mimeo Harvard and UCLA.

[23] Maskin, Eric and Riley, John "Uniqueness in Sealed High Bid Auctions", 1996a Unpublished Working Paper, Harvard and UCLA.

[24] Maskin, Eric and Riley, John "Existence in Sealed High Bid Auctions", 1996b Unpublished Working Paper, Harvard and UCLA.

[25] Pesendorfer, M. "A Study of Collusion in First-Price Auctions", Mimeo, Yale University (1996)

[26] Pesendorfer, M. and Jofre-Bonet "Bidding Behavior in a Repeated Procurement Auction", 2000 Mimeo, Yale University.

[27] Porter, R.H. and Zona, J.D. "Detection of Bid-Rigging in Procurement Auctions", Journal of Political Economy. Vol 101(1993), pp.518-538

[28] Porter, R.H. and Zona, J.D. "Ohio School Milk Markets: An Analysis of Bidding", Rand Journal of Economics, Summer 1999.

[29] Riley, John; Li, Huagang; "Auction Choice", Unpublished Working Paper, UCLA.

[30] Rothkopf, Michael, "A model of rational competitive bidding", Management Science, 1969, 15, pp.362-

73.

[31] Wilson, Robert, “A bidding model of perfect competition”, Review of Economic Studies, 44, pp. 511-18.