Spatial Competition and Interdependent Costs in Highway Procurement*

Paulo Somaini†
Stanford University

Job Market Paper

Abstract. I investigate the effect of competition on bidder behavior and procurement cost using highway auction data from Michigan. While a bidder's distance to a project location is important in explaining participation and bid levels, there is no evidence of more aggressive bidding when competitors are located close to the project. This pattern is at odds with theoretical predictions based on first-price auctions with private costs but can be rationalized by a more general model that allows firms to have nonindependent private information and partially common costs. I show that such a model is identified from observable variation in bidder-specific cost shifters, and develop an estimation procedure that exploits variation in project locations. The findings point to significant common costs and to a high degree of correlation in private information. Model estimates are used to show that common costs and information correlation reduce the effect of competition on procurement costs by 28 percent relative to a benchmark that assumes independent private costs. Moreover, subsidies to weak bidders are estimated to cost 27 percent more than in the private costs benchmark.

*This version: November 19, 2011. For the latest version, see `www.stanford.edu/~soma/`. I am grateful to my advisors Frank Wolak, Liran Einav, and Han Hong for their guidance and encouragement. I also thank Tim Armstrong, Arthur van Benthem, Tim Bresnahan, George Bulman, Dominic Coey, Ignacio Esponda, Pedro Gardete, Caroline Hoxby, Jakub Kastl, Theresa Kuchler, Thomas MaCurdy, Roy Mill, Alejandro Molnar, Peter Reiss, Ilya Segal, Johannes Stroebel, Ali Yurukoglu and seminar participants at Stanford for helpful comments and suggestions. Financial support from the Martin Lee Johnson Graduate Fellowship is gratefully acknowledged.

†Department of Economics, Stanford University, 579 Serra Mall, Stanford, CA 94305, soma@stanford.edu

1 Introduction

What are the effects of increased competition in procurement markets? The textbook answer is that competition reduces firms' market power, induces more aggressive bidding, and thus reduces procurement costs. However, this is not necessarily true in markets where bidders have common costs that are uncertain at the time of the auction. The theoretical predictions of a model that allows for both common and private costs are ambiguous—the effect of competition depends on the particular technology and information structure of bidders. I consider a model that allows for asymmetric bidders, nonindependent private information and common costs, and show that its primitives are nonparametrically identified by the distribution of bids conditional on bidder-specific cost shifters. I estimate this model using highway procurement data from Michigan where the distance of a firm to the project location is an observable shifter of its cost, and use the estimated information structure to quantify the effect of increased competition on firm behavior and on expected procurement costs.

Competitive procurement allows governments around the world to obtain goods and services in a cost-efficient and transparent manner. In 2007, federal, state and local agencies in the United States procured goods and services worth over one trillion dollars, and the federal government alone employed competitive procurement for more than 250 billion dollars. Sound policy evaluation requires understanding the extent to which more competitive environments affect bidding behavior and procurement costs. For example, procurement agencies often subsidize minority-owned businesses and local or domestic firms.[1] Some argue that non-subsidized bidders will behave more aggressively because they face increased competition, and as a result, the costs of the subsidies will be mitigated. The costs of subsidies may be higher, however, if the nature of competition is such that bidders respond to increased competition by bidding more cautiously.

Competition may induce cautious bidding due to the adverse selection or winner's curse effect associated with the existence of common costs. When sellers face uncertain costs that are also common to their competitors (e.g., conditions in the input markets or potential difficulties in the execution of the project), they may win the projects that their competitors deem undesirable or too costly. In more competitive environments, sellers will win a more adversely selected set of projects. Thus, rational bidders will realize that more competition implies higher expected costs and will increase their bids.

Analysis of the bidding behavior of firms that participate in Michigan's first-price procurement auctions reveals a pattern that is consistent with the winner's curse. Transportation costs are very important in this market: the median auction participant is located 26 miles from the project, while the median winner is only 17 miles away. Although closer firms bid more aggressively, there is no evidence of more aggressive bidding when competitors are located close to the project: four of the five largest firms do not bid any lower on projects next to their competitors plants. In other

[1] Ayres and Cramton (1996) analyze the effect of subsidies to businesses owned by minorities and women in FCC auctions. Marion (2007) and Krasnokutskaya and Seim (2011) study subsidies to small firms in California highway procurement. Athey et al. (2001) compare the effects of subsidies and set-asides in timber auctions.

words, bidders do not behave more competitively when they face increased spatial competition. Although suggestive, this evidence does not confirm the presence of common costs. Pinkse and Tan (2005) show that this lack of competitive response is also consistent with a class of models in which bidders' private costs are affiliated.[2] To distinguish between the effect of affiliation described by Pinkse and Tan and the effect of the winner's curse, it is necessary to exploit the theoretical predictions of equilibrium bidding in a general model that nests both cases.[3]

Consider a model in which risk-neutral firms participate in a first price auction. Firms are uncertain about their project completion costs but they privately observe a signal, e.g., some private information about their equipment and labor capacity constraints, or about future input or subcontract market conditions. Define a bidder's "full information cost" as the expected cost of the project if the bidder learned the private information of all competitors. For example, bidder A may revise his beliefs about his own completion costs after he learns bidder B's private information about her contracts with firm C that rents equipment. Because the full information cost is allowed to depend on competitors' signals, this is a model of interdependent or common costs.

In an influential paper, Laffont and Vuong (1996) show that this model cannot be identified solely from the joint distribution of bids. However, Athey and Haile (2003) remark that additional data can aid in obtaining testable implications of rational bidding under different information structures.[4] I show that it is possible to identify each bidder's full information cost in the presence of publicly observable cost shifters that introduce asymmetries between bidders. Indeed, asymmetries are not a nuisance that needs to be accounted for to avoid inconsistent estimates, but instead generate the variation needed to identify the model: multi-dimensional variation in the cost shifters identifies multi-dimensional full information costs.

The conditions for identification are that each bidder receives a single-dimensional signal, and that his cost shifter affects only his own cost but not the cost of his competitors. Provided that these conditions are satisfied, the joint distribution of signals and bidders' full information can be recovered from the distribution of bids conditional on cost shifters.

Because plant locations are predetermined and distance to the project is a determinant of costs, firms will be more competitive in projects located close to their plants. Moreover, firms will face more competition and more adverse selection for projects close to competitors' plants. Consider an example with three firms A, B and C. If the project is located close to bidder A and far from bidder B, bidder C should realize that he wins on a very adversely selected set of signals of bidder A, and on a less selected set of signals of bidder B. Similarly, if the project is located far from A and close to B, bidder C wins on a very adversely selected set of signals of bidder B. Notice that

[2] Affiliation is a stronger notion of correlation. See Milgrom and Weber (1982) for additional details

[3] This evidence is also consistent with bid rigging or market allocation agreements among firms. I abstract from this possibility in the absence of charges of collusion in this market.

[4] For example, Paarsch (1992) and Hong and Shum (2003) estimate models that nest both private and common costs models. They identify the model using variation in the number of bidders and functional form assumptions. Haile et al. (2004) and Kastl and Horctacsu (2011) test the private values model. Hendricks et al. (2003) tests the pure common value model with asymmetrically informed bidders.

if bidder C's full information cost depends only on A's and not on B's signals, he should bid more cautiously in the former case, when A is close to the project. This illustrates how bidding behavior for varying project locations is informative about bidders' full information costs.

The estimation procedure follows this identification argument closely. I first estimate the joint distribution of bids conditional on project location and a set of controls. I then use the theoretical predictions of equilibrium bidding to obtain the marginal cost that rationalizes each bid. As first noticed by Milgrom and Weber (1982), the marginal cost is the expected cost conditional on the event that the bid is pivotal—it equals the minimum competitors' bid. Finally, I use the estimates of the marginal costs to test the hypothesis of private costs and to estimate the parameters of each bidder's full-information costs.

I find evidence against the private costs hypothesis for 6 of the 10 largest firms. Moreover, the results suggest that while a bidder's full-information costs are mainly affected by his own signal, the effect of competitors' signals is also positive. I estimate that if bidder A receives a signal that increases his expected costs by \$1000 and bidder B learns that signal from bidder A, B expected costs increase by \$70-\$250, depending on the identities of A and B. These estimates allow me to quantify the effect of increased competition on procurement costs. I simulate the effect of a reduction in the cost of one of the existing firms, and as a result, all other firms facing increased competition. My estimates suggest the effects of competition are weaker than in the independent private costs model. Moreover, my results show an interesting implication of the presence of the winner's curse. Suppose that firm A builds a new plant. Firm A will not only be more efficient in serving projects around its new location, but will also face less of a winner's curse. Therefore, firm A will bid even more aggressively and its competitors will bid less aggressively relative to the independent private costs benchmark.

I quantify the effects of a subsidy policy in favor of weak (high-cost) bidders. Suppose that weak bidders receive a subsidy when they win the auction. Without a subsidy, weak bidders are subject to a large winner's curse because they only win when the low-cost bidders receive very unfavorable signals. With a subsidy they compete on a more even ground with the low-cost bidders and win a less selected set of projects. Relative to the private cost benchmark, subsidized bidders will bid more aggressively, win more often, and because the subsidy is paid only when a subsidized bidder wins, the total cost of the subsidies will be higher. I verify this is the case empirically and estimate that subsidies to weak bidders cost 27 percent more than in the private costs benchmark.

The remainder of the paper is organized as follows. Section 2 describes the data and the institutional setting. Section 3 presents descriptive evidence that show that bidders do not behave more aggressively when they face more competition. Section 4 presents the model and shows its nonparametric identification. Section 5 describes the estimation procedure and Section 6 presents the results. Section 7 shows how procurement costs and bidders' behavior respond to increased competition. The last section concludes.

2 Michigan's highway procurement auctions

2.1 The letting process

Each year the Michigan Department of Transportation (MDOT) awards around 1,050 highway construction and maintenance contracts at a cost of 1.2 billion dollars. On each monthly letting date, 150-200 firms submit a sealed bid for one or more of 50-70 contracts. Firms may participate in as many auctions per letting date as allowed by their work-type and financial prequalification status. The work-type prequalification status of a firm is a list of all the classes of work that the firm can perform. There are 52 work classifications and the typical firm is prequalified for 6-10 of these. The financial rating of a firm is the maximum dollar amount of contracts it can have pending with the MDOT. While the work-type classification is public information, the financial rating of a firm is confidential.

Contracts are advertised for at least 45 days before the letting date, so bidders have detailed information about them. The information about future projects is less precise. The MDOT publishes a 3-month projection of future projects and a 5-year Transportation Improvement Plan, but these are subject to frequent changes and updates.

Prior to submitting a bid, firms download the technical plan and submit a form to become eligible to bid. The MDOT keeps an updated list of both eligible bidders and plan holders that is publicly available on its website. However, eligible bidders and plan holders often choose not to submit a bid. Besides, firms may submit the eligibility form as late as 5:00 p.m. on the day preceding the letting date, and may not appear in the list of eligible bidders prior to the bid submission deadline. As a result, firms are unable to predict with certainty the set of participants in an auction. Thus it is likely that firms base their expectations of competition largely on the location and technical characteristics of the project.

Each contract describes a list of tasks that the contractor has to perform. A task specifies a description and a quantity, e.g., earth excavation, 600 cubic yards. For each task, the MDOT engineer sets a unit price, so that the total estimated cost of the task is the price times the quantity. The engineer's estimate for the contract is the sum of all the tasks' estimated costs. Bidders submit a unit price bid for each task, and the total bid is the inner product of unit prices and quantities. The bidder with the lowest total bid wins. The unit price bid for each task becomes relevant if the quantity has to be modified after the award of the contract, e.g., if the contractor needs to excavate 650 cubic yards. The payment is adjusted by the unit price bid times the difference between the actual and estimated quantities.[5]

Although there is no formal reserve price, the procuring agency has the option to reject all bids if the lowest bid exceeds 110% of the MDOT engineer's estimate. In the case the bids are rejected, the project can be revised and offered in a future letting date. If the agency accepts a bid despite exceeding the estimate, it must justify in writing why the estimate was not correct or why the bids were excessive. From 2001 to 2010, the lowest bid exceeded 110% of the engineer's estimate in 11%

[5]Bidders' incentives to skew their bids are analyzed in Athey and Levin (2001) and Bajari et al. (2011).

of cases and 15% of these were rejected.

2.2 Why distance matters

This paper will focus on contracts where contractors must be prequalified to perform Hot-Mix-Asphalt (HMA) work. HMA is the pavement material used in 96% of all paved roads and streets in the US. It consists of asphalt or bitumen and mineral aggregate that is heated and mixed in the plant. The mix must be trucked to the project site, laid on the road and compacted while the mix is sufficiently hot (above 275°F/135°C). The temperature of the mix at the time of compaction is key to the quality of the pavement mat. Once the mix falls below 175°F/79°C, it cannot be further compacted, and a poorly compacted mat will age faster. HMA pavement projects are rarely performed during winter for this reason.

Trucking time from the plant to the project location is an important determinant of costs not only because of transport costs, but due to the cooling process. During transport, the surface layer of the mix cools faster than the inner mass. Once the mix is dumped into the paver and laid on the road, these temperature differentials may persist and result in cool spots in the pavement mat that cannot be properly compacted. These problems can be mitigated by incurring additional labor and rental costs, for example, by using a Material Transfer Vehicle to remix on site. Thus firms that own plants located close to the project have lower transportation costs and lower costs associated with excess cooling of asphalt.

2.3 Data

Data for all auctions and bids from 2001 to 2010 are available through the MDOT. For each auction, the data includes the project's description, location, prequalification requirements, the engineer's estimate of the total cost of the project, and the list of participating firms and their bids. To obtain the geographical coordinates for each project location, I match the road names in the description to the database of roads available at the Michigan's Geographical Data Library.

The location of each firm's plants and mineral aggregate quarries were obtained from several sources: the MDOT contractor directory; individual firm searches using OneSource North American Business Browser, Duns & Bradstreet's Hoovers and yellowpages.com; firm websites; and the data on firms collected by Einav and Esponda (2008). A firm location was included in the final data set if it appeared in at least two sources or if it was listed explicitly on a firm's website.

Of the 10,522 MDOT auctions that were run from January 2001 to December 2010, 3,851 auctions required the prime contractor or one of the subcontractors to be prequalified to perform work with HMA. Table 1(a) shows the main descriptive statistics of this set of auctions. The median engineer's estimate is around $650,000, while the median winning bid is around $600,000. It is convenient to normalize bids with respect to the engineer's estimate. Let the normalized bid be $b = Bid/Engineer - 1$. The median normalized winning bid is 7% below the engineer's estimate. The median participant is located 38 km (26 miles) from the project, while the median winner is only 27 km (17 miles) away. The average "money-left-on-the-table", or how much higher the second lowest bid is relative to the lowest, is about 7%.

In the empirical analysis I use a set of variables to control for observed project heterogeneity. These variables are dummies for the year of the auction, the size of the project, and other work-type prequalifications that are required to participate in the auction. Table 1(b) shows the distribution of project sizes and Table 1(c) shows the distribution of prequalification requirements.

It is interesting to observe how normalized bids vary with the number of actual participants. Table 2 shows that the average winning bid is decreasing with the number of participants, but the average bid is not. In a standard symmetric independent private cost model with no entry cost, both the expected lowest bid and the expected bid should be decreasing in the number of bidders. Of course, in this setting it is likely that the symmetry and the costless entry assumptions fail. In the next section, I analyze with more detail how individual firms' bidding behavior changes with increased competition.

3 Descriptive evidence

In this section I document two stylized facts. First, firms bid and win more often on projects close to their plants. Second, they do not bid more aggressively on projects close to their main competitors' plants.

I rank firms by the number of bids they submitted in the 3,851 HMA auctions and label them accordingly, i.e., firm 1 is the first-ranked firm and so on. Because firms are differentiated by their locations and technologies, some pair of firms are closer competitors than others. Define the two main competitors of firm i as the two firms that are observed to participate most often in the auctions where i participates.

Figure 1(a) shows the locations of all the auctions in my sample, but with a different marker depending on whether firm 1 won, participated, or did not participate in that auction. It also shows the location of firm 1's plants. The auctions where this firm participated are concentrated around its plants, and the auctions won by the firm are even more concentrated. A similar pattern can be found for other firms. Figure 1(b) shows the same plot for firm 21 that has only one plant.

Table 3(a) shows the results of a least squares regression of observed normalized bids on bidders' own distance to the project, the two main competitors' distances and auction characteristics. I present the results for all firms and for each of the five largest firms individually. While bidders seem to bid lower for projects close to them, their bids are not lower for projects located near their main competitors. The only exception is firm 5, which seems to bid more aggressively when the project is located next to its main competitor, firm 6.

These results appear to be at odds with the predictions of an independent private cost model. Because firms submit lower bids for projects close to their plants, their competitors are less likely to win for any given bid and have less ability to exercise market power. As a result, firms should reduce their markups and submit lower bids in projects close to their competitors.[6]

[6]The effect of a more competitive bidder on competitors' bids can be ambiguous in an independent private cost

Before concluding that the observed bidding pattern of the first four firms is incompatible with a model of rational bidding under independent private costs, it is useful to consider what type of unobserved auction heterogeneity could yield the above results. In principle, if plants are clustered around high-cost locations, firms will submit higher bids around their competitors' locations simply in response to higher costs. In fact, one may expect plants to be clustered in densely populated areas. These areas may be prone to high costs due to logistical reasons and firms may be capacity constrained in areas where they are able to serve a larger volume of projects, i.e., where they face more demand. I construct a measure of road density in a particular location using data on all roads in the state to control for this potential confounding effect. I then assign each project to one of 10 bins according to the road density at its location, and add bin dummy variables to the specification. The results for these specifications are reported in Table 3(b). Even after controlling for this demand shifter, firms do not seem to bid lower on projects close to competitors' plants.

Firms may also respond to increased competition by adjusting their participation decision. Suppose that bidders learn their private information before they decide whether or not to participate in the auction. When competitors are close, they should participate less often and only when their costs are lower. This type of selection implies that even if bidders do not change their bidding strategy, the correlation between bids and competitors' distance should be positive—if competitors are close, the expected bid conditional on participation is lower— and the results above are even harder to reconcile with an independent private cost model. Appendix A presents evidence that the results in this section still hold after controlling for selection.

This section showed that while firms are more likely to bid and win on projects close to their plants, they do not bid more aggressively for projects close to their main competitors' plants. This evidence suggests the presence of common costs but it is also consistent with an affiliated private costs model, as noted by Pinkse and Tan (2005).[7] To distinguish between the two models, I exploit the theoretical predictions of equilibrium bidding in a general model that nests both cases.

4 Model

The model considered in this section is able to rationalize the stylized facts of this market. It allows for asymmetric bidders, nonindependent private information and common costs. Moreover, it is nonparametrically identified from the distribution of bids conditional on cost shifters.

model where bidders need to pay an information acquisition cost in order to learn their completion costs. Consider bidders A, B and C. Bidders A and B are close to the project and C is far away. Suppose that the equilibrium is such that A and B participate with probability 1 and C plays a mixed strategy—he enters with some probability less than 1. If A is closer to the project, he will bid more aggressively, C may enter less often and B may bid higher.

[7]The easiest way to illustrate the affiliation effect described by Pinkse and Tan (2005) is to consider a market that has an unobserved state z. Higher values of z imply that bidders are more likely to draw higher private costs. A bidder anticipates that winning the auction means that the z-value is probably larger than his a priori information suggests that it is. The greater is the number of bidders, the greater is the posterior probability of z being large conditional on winning, and hence the higher is her bid relative to the independent private costs case.

4.1 Set up

The setup is a modified version of Athey and Haile (2006) and Milgrom and Weber (1982) that highlights the role of publicly observable cost shifters. I follow the convention of using upper-case letters to refer to random variables, and lower-case letters to refer to their realized values.

Costs and information structure. The auctioneer, in this case the Department of Transportation, runs a first price auction between n risk-neutral bidders. At the time of the auction, bidders are uncertain about their project completion costs. The cost to bidder i is denoted by the random variable C_i.

Firms possess some private information, e.g., information about their own labor and equipment capacity constraints, and about conditions in the input or subcontract markets. This private information is summarized by a signal. At the time of the auction, bidder i observes his own signal s_i, but he does not observe his competitors' signals. From his perspective the signal of competitor $j \neq i$, S_j, is a random variable. Let $S_{-i} = \{S_j\}_{j \neq i}$ denote the random vector of competitors' signals and $S = (S_{-i}, S_i)$ denote the full random vector of signals.

All bidders have access to the following public information: bidder-specific cost shifters $(d_1, d_2, ..., d_n)$, such as the distance of each bidder to the project, and a set of observable auction characteristics w_0, such as the engineer's estimate. All public information is denoted by $w = (d_1, d_2, ..., d_n, w_0)$.

The expected cost of bidder i given his information at the time of the auction is $E\left(C_i|s_i, w\right)$. Now suppose that bidder i learns the signals of his competitors, then his expected cost becomes $E\left(C_i|s_1, s_2, ..., s_n, w\right)$. This expectation is bidder i's full information expected completion cost, or for short, full information cost.

This general framework nests both the private and the pure common costs paradigms. With private costs, each bidder knows his own expected completion cost, so the full information cost of firm i does not depend on its competitors' signals, i.e., $E\left(C_i|s_1, s_2, ..., s_n, w\right) = E\left(C_i|s_i, w\right)$. With pure common costs, the cost of completing the project is common to all bidders, so they all share exactly the same full information cost. This framework also allows for nonindependent signals. Let $F_{S|w}$ be the joint distribution of signals conditional on the public information w.

The following assumptions are maintained throughout:

1. Firms are risk neutral.

2. Signals are one-dimensional random variables. Without loss of generality, signals are marginally distributed as uniform $[0, 1]$.

3. The full information cost of firm i is continuous and strictly increasing in d_i and s_i but is not affected by d_j for all $j \neq i$, i.e.,

$$E\left[C_i|s_1, ..., s_n, d_1, ..., d_n, w_0\right] = E\left[C_i|s_1, ..., s_n, d_i, w_0\right]. \tag{1}$$

4. Cost shifters and signals are independent: $F_{S|w} = F_{S|w_0}$.

5. The cost of submitting a bid in an auction with observable characteristics w_0 is $k_i(w_0)$. The bidder first observes his signal and then decides whether to enter.

Assumption 1 can be justified on the basis that each project is small relative to the portfolio of contracts that firms hold. For example, firm 1 holds on average \$70 million of pending contracts with the MDOT. This figure does not include contracts that the firm may have with local transport agencies and private parties.

Assumption 2 states that firms summarize all private information in a single dimensional variable. It rules out the possibility that bidder i receives two signals and that each signal differently affects the posterior distribution of competitors' information and costs. Suppose that each bidder receives one signal about own equipment availability and a signal about conditions in the rental equipment market, and only the second changes the expectations about competitors' costs. This violates Assumption 2.

Assumption 3 states that bidders' cost shifters only affect their own full information costs but not competitors'. For example, the distance of bidder j to the project does not affect the full information cost of bidder i. Assumption 4 allows the joint distribution of signals to depend on observable auction characteristics but not on the vector d. These assumptions are less innocuous than they might seem. They imply that j's distance affects neither the technology of i nor the quality of the information of bidder j as perceived by i. They rule out the possibility that bidder i regards the signal of bidder j as more informative when j is close to the project. This assumption is more reasonable in the context of highway procurement, where costs uncertainties depend on the bidder-specific capacity constraints and input market conditions, than in the context of oil drainage leases, where the neighbor firms are better informed than non-neighbors about the value of the tract (Hendricks and Porter 1988,1993).

Assumption 5 states that submitting a bid is costly. Bidders need to spend a few hours reading the contract plan, budgeting each task and preparing the bid. These costs do not seem too high, but they may be high enough to discourage a bidder that has a low probability of winning even when bidding slightly above his costs. While this assumption seems appropriate for this context, the model can also be identified under the assumption that bidders have to incur a cost to learn their private information (Somaini, 2011).

Optimal bidding and equilibrium. First price auctions are modeled as Bayesian games where bidders are privately informed about their signals. A strategy for bidder i is a function $\beta_i(s_i, w)$, that prescribes a bid based on his private and public information. The primitives of the game are the joint distribution of signals and the n payoff functions of bidders. Notice that the payoff function of bidder i is simply his bid minus his own full information cost when he wins the auction minus the entry cost if he decides to submit a bid.

$$\pi_i(b, s, w) = (\ b_i - E[C_i | s_1, ..., s_n, d_i, w_0])\,\mathbf{1}(b_i < b_{-i}) - k_i(w_0)\,\mathbf{1}(b_i < \infty), \tag{2}$$

where $b_i = \infty$ when bidder i does not participate.

The equilibrium concept of the game is Bayesian Nash Equilibrium (BNE). In a BNE, each bidder strategy $\beta_i(s_i, w)$ is a best response given competitors' strategies. Moreover, beliefs about the distribution of competitors' bids have to be consistent with their bidding strategies and the conditional distribution of signals. Let B_j be the bid of bidder j; then i's beliefs about the distribution B_j given its own information is derived from the fact that $B_j = \beta_j(S_j, w)$ and that the distribution $S_j|s_i, w$ is determined by the joint distribution of signals $F_{S|w_0}$. Similarly let $M_i = \min\{B_j\}_{j\neq i}$; in equilibrium, the belief of firm i about the distribution of M_i is derived from competitors' bidding strategies and the conditional distribution of signals.

I make the following monotonicity assumption on the equilibrium bid functions:

6. The data is generated by a unique Pure Strategy Bayes Nash Equilibrium. Each bidder strategy $\beta_i(s_i, w)$ is a differentiable strictly increasing function of s_i for all $s_i \in [0, \pi_i(w))$, and $\beta_i(s_i, w) = \infty$ for $s_i \in (\pi_i(w), 1]$, where $\pi_i(w)$ is the participation cutoff signal.[8]

Observable data. The observable data include bids and all public information. Let $G_{B|w}$ be the joint distribution of bids conditional on w, $G_{B_i|w}$ be the marginal distribution of bidder i's bids conditional on w, and $G_{M_i|B_i,w}$ be the distribution of M_i conditional on B_i and on w. These three distributions can be estimated from the data.

4.2 Identification

The model above is identified if it is possible to learn its primitives from the joint distribution of bids conditional on all public information. The information and technology structure in this model is summarized by the joint distribution of costs and signals conditional on w: $F_{C,S|w}$. It is not possible to recover this joint distribution from bidding behavior. Because bidders are risk neutral, the primitives of the Bayesian game are given by the n full information costs and the joint distribution of signals. If there are two different joint distributions of costs and signals that imply the same primitives of the Bayesian game, the equilibrium strategies and the observed behavior of bidders will be identical. Therefore, the data will be unable to tell them apart. Thus the primitives that can be identified are an array of n full information cost functions, and a joint distribution of signals:

$$\left\{\{E(C_i|s_1, ..., s_n, d_i, w_0)\}_{i=1}^n, F_{S|w_0}\right\}. \tag{3}$$

The information contained in these primitives is enough to analyze the effects of most policy changes— rules of the auction, reserve prices, subsidies— on outcomes such as bidding behavior, project allocation and procurement costs. However, there are some questions that cannot be answered without knowledge of $F_{C,S|w}$. Consider a policy change that allows the winning bidder to subcontract the project after it is awarded. Suppose that at the subcontracting stage there is no private information and that each firm's publicly known cost is its full information cost evaluated at

[8] This assumption can be weakened to allow at most one bidder who may, with positive probability, bid the highest bid that allows him to win with strictly positive probability. See Athey and Haile (2006).

the realized signals plus an idiosyncratic ex-post shock. The winner may make a take-it-or-leave-it offer to the firm that has lowest ex-post costs. His ex-post market opportunities are more favorable when the variance of the ex-post shock is larger because the minimum competitors' cost will be lower. Because bidders should take into account the possibility of subcontracting (or resale as in Haile, 2001), this policy results in a more competitive auction environment if the variance of the ex-post shock is large. This variance, however, is not a primitive that can be recovered from bidding data in an environment that does not allow ex-post subcontracting.

In the remainder of this section I show that the joint distribution of signals and the full information cost functions are identified by the joint distribution of bids conditional on public information. I assume that the entry costs are zero to simplify the exposition. I show that these arguments are also valid with positive entry costs in Appendix B.

4.2.1 Identification of the joint distribution of signals

By monotonicity of the strategy functions (Assumption 6) and the normalization that signals are marginally uniformly distributed, $s_i = \beta_i^{-1}(b_i, w) = G_{B_i|w}(b_i)$.[9] In other words, it is possible to invert the bid function to obtain the signal that prompted each bid. After each signal is recovered, it is straightforward to compute their joint distribution:

$$F_{S|w}(s) = G_{B|w}\left(G_{B_1|w}^{-1}(s_1), ..., G_{B_n|w}^{-1}(s_n)\right). \tag{4}$$

Notice that this is the copula of bids conditional on w.[10]

The model is in fact overidentified. Because $F_{S|w}$ is recovered from the data, it is possible to test whether $F_{S|w} = F_{S|w_0}$, i.e., if Assumption 4 holds or not. The assumption will be rejected if the copula of bids differs across different values of d.

On the other hand, $F_{S|w_0}(s)$ may not be identified for some value of s if the support of d is not large enough. Suppose that there is no variation in d, and that the observed probabilities of participation are given by the vector $\pi(w)$, then $F_{S|w_0}(s)$ is not identified for any s that does not satisfy $s \leq \pi(w)$. Variation in d expands the region where the joint distribution of signals is identified. If there is a d where all firms bid with certainty, the joint distribution of signals is identified for all s.

[9] By definition $G_{B_i|w}(b_i) = P(B_i < b_i|w)$. By the assumption that the observed data is the result of equilibrium behavior: $G_{B_i|w}(b_i) = P(\beta_i(S_i, w) < b_i|w)$. By monotonicity of β: $G_{B_i|w}(b_i) = P\left(S_i < \beta_i^{-1}(b_i, w)|w\right)$, then by uniformity of S_i, $G_{B_i|w}(b_i) = \beta_i^{-1}(b_i, w)$, which equals s_i.

[10] Consider the random vector of bids $(B_1, ..., B_n|w)$, and suppose that the marginal distributions $G_{B_i|w}$ are continuous. The random variable $U_i = G_{B_i|w}(B_i)$ has a uniform distribution. The copula of $(B_1, ..., B_n|w)$ is defined as $C(u_1, ..., u_n|w) = P(U_1 \leq u_1, ..., U_n \leq u_n|w)$. By the definition of U_i, the right hand side becomes $P\left(G_{B_1|w} \leq u_1, ..., G_{B_n|w} \leq u_n|w\right)$, and by monotonicity of $G_{B_i|w}$, $P\left(B_1 \leq G_{B_1|w}^{-1}(u_1), ..., B_n \leq G_{B_1|w}^{-1}(u_n)|w\right)$, which yields the desired result.

4.2.2 Identification of the full information costs

In a Bayesian Nash Equilibrium each firm's bidding strategy is a best response to competitors' strategies. Bidder i's optimal bid given his information (s_i, w) solves the following expected profit maximization problem:

$$\max_b E\left[(b - C_i)\,\mathbf{1}\,(b \le M_i)\,|s_i, w\right], \tag{5}$$

where $M_i = \min_{j \neq i} \beta_j (S_j, w)$, the expectation is taken over M_i and C_i, and $\mathbf{1}\,(.)$ is an indicator function. The bidder's basic tradeoff is that submitting a low bid reduces the profits when he wins, while submitting a high bid reduces the probability of winning. It is illustrative to think of the bidder as choosing his probability of winning q instead of his bid. His problem is to choose q to maximize:

$$\mathcal{D}_i^{-1}(q|s_i, w)\, q - \mathcal{C}_i(q|s_i, w), \tag{6}$$

where $\mathcal{D}_i^{-1}(q|s_i, w)$ is the inverse expected residual demand function: it returns the bid that i has to submit in order to win with probability q, and $\mathcal{C}_i(q|s_i, w)$ is the expected cost conditional on winning times the probability of winning.[11]

Variation in competitors' cost shifters displaces the residual demand curve. For example, firm A bids more aggressively in projects located close to its plants, which implies that bidder i faces a lower residual demand curve— for each bid he wins with lower probability.

In the general interdependent model, variation in d_{-i} also shifts the expected cost. Consider a project located close to bidder A and far from bidder B. Firm i should realize that it wins on an adversely selected set of signals of A, and on a less selected set of signals of B. Now consider another project that is located close to bidder B and far from A. In this case, i wins on a selected set of signals of B. If i's costs depend only on A's signals and not on B's, the expected cost should be higher in the former case, when A is closer to the project. This example also illustrates that if i's cost does not depend on competitors' signals (the private costs hypothesis) the expected cost does not depend on d_{-i}.

Bidder i's expected residual demand curve is identified from the observable data:

$$\begin{aligned} \mathcal{D}_i(b|s_i, w) &= P(b \le M_i|s_i, w) \\ &= 1 - P(M_i < b|b_i, w) \\ &= 1 - G_{M_i|B_i, w}(b|b_i), \end{aligned} \tag{7}$$

where the first equality follows from the definition of the expected residual demand curve, the second from monotonicity of the bidding strategies ($b_i = G^{-1}_{s_i|w}$), and the third from the definition of the cumulative distribution function.

[11] Formally, the expected residual demand function of bidder i is $\mathcal{D}_i(q|s_i, w) = P(b \le M_i|s_i, w)$, and its inverse is: $\mathcal{D}_i^{-1}(q|s_i, w) = \{b : P(b \le M_i|s_i, w) = q\}$. Similarly, define i's expected cost as $\mathcal{C}_i(q|s_i, w) = E\left[C_i|s_i, w, \mathcal{D}_i^{-1}(q|s_i, w) \le M_i\right] q$.

Now consider the first-order condition of problem (6):

$$E\left[C_i|M_i = \mathcal{D}_i^{-1}\left(q^*|s_i, w\right), s_i, w\right] = \mathcal{D}_i^{-1}\left(q^*|s_i, w\right) + \frac{d\mathcal{D}_i^{-1}\left(q^*|s_i, w\right)}{dq} q^*. \tag{8}$$

It states that bidder i should choose q^* such that the expected marginal cost of increasing the probability of winning (left hand side) equals the expected marginal revenue of doing so (right hand side).

Under Assumption 6 each observed bid b_i satisfies the first-order condition, so $b_i = \mathcal{D}_i^{-1}\left(q^*|s_i, w\right)$, $q^* = 1 - G_{M_i|B_i,w}\left(b_i|b_i\right)$, and condition (8) becomes:

$$E\left[C_i|M_i = b_i, s_i, w\right] = b_i - \frac{1 - G_{M_i|B_i,w}\left(b_i|b_i\right)}{g_{M_i|B_i,w}\left(b_i|b_i\right)}. \tag{9}$$

The right hand side of this expression, the expected marginal revenue, is identified from the distribution $G_{M_i|B_i,w}$. Figure 2(a) shows the intuition. The marginal revenue can be derived from the demand curve. For each observed bid it is possible to find the q^* corresponding to that bid from the demand curve, and then find the marginal revenue associated with q^*. Denote the expected marginal cost that rationalizes each observed bid by:

$$mc_i = b_i - \frac{1 - G_{M_i|B_i,w}\left(b_i|b_i\right)}{g_{M_i|B_i,w}\left(b_i|b_i\right)}. \tag{10}$$

This first-order condition is standard in the empirical auction literature. Campo et al. (2003) derive an analogous expression for a private values model and interpret the left hand side as the bidder's valuation. In a model with common costs (or values), the remaining challenge is to link the distribution of marginal costs with the full information cost.

Under the private cost hypothesis, the full information cost does not depend on competitors' signals, and the expected marginal cost becomes just $E\left[C_i|s_i, d_i, w_0\right]$. A testable implication is that the distribution of mc_i should not depend on competitors' cost shifters d_{-i}. For example, suppose that firm i is observed bidding repeatedly in two projects: one 10 miles east from its location, and the other 10 miles west. Firm B is located close to the latter project which means that is 20 miles away from the former. Although firm i's bid distribution could be different across locations because it faces different intensity of competition, the implied distribution of expected marginal costs mc_i should be identical.

In the more general model of interdependent costs, the marginal cost is the expected full information cost conditional on the event that the bid is pivotal:

$$E\left[C_i|M_i = b_i, s_i, w\right] = E\left(E\left(C_i|s_1, S_{-i}, d_i, w_0\right)|M_i = b_i\right) \tag{11}$$

The event "bid b_i is pivotal"— that at least one competitor submits an equally low bid and the rest submit higher bids— is composed by the following set of signals:

$$\left\{S_{-i} : \min_{j \neq i} \beta_j\left(S_j, w\right) = b_i\right\}. \tag{12}$$

Variation in competitors' cost shifters change the set of competitors' signals in this "pivotal set", and therefore, the marginal cost. For example, consider a project located close to bidder A and far from bidder B. Firm i should realize that the marginal projects it wins are ones where he ties with A that received a high signal, or ties with B who received a low signal. Figure 2(b) illustrates the pivotal set of signals in this case. The opposite holds in a project located close to B and far from A.

It is possible to recover the full information cost functions from the distribution of marginal costs conditional on w. The intuition behind the proof can be illustrated using the space of competitors' signals. To simplify the exposition assume that bidders participate with probability one. Define the "win-set" of competitor i when he bids b_i as the set of competitors' signals for which he wins the auction: $\{S_{-i} : \min_{j\neq i} \beta_j (S_j, w) > b_i\}$. Similarly, define the "tie-point" as the vector of competitors' signals for which all of them bid exactly b_i: $\left\{\beta_j^{-1}(b_i, w)\right\}_{j\neq i}$, which is identified from the data by $\left\{G_{B_j|w}(b_i)\right\}_{j\neq i}$. In the case of $n = 3$, the win-set is a rectangle in the space of competitors' signals with the southwest corner at the tie-point and northeast corner at $(1, 1)$ as shown in Figure 2(c). The graphical intuition carries to higher dimensions: the win-set is an hyperrectangle with the southwest corner at the tie-point and northeast corner at $(1, 1, ..., 1)$. The expected marginal cost is equal to the expected cost conditional on the pivotal set that was defined as $\{S_{-i} : \min_{j\neq i} \beta_j (S_j, w) = b_i\}$. In the $n = 3$ case, the pivotal set is composed by the signals in the sides of the win-set that are adjacent to the tie-point as shown in figure 2(c).

Intuitively, the identification proof involves integrating the expected cost conditional on pivotal sets (marginal costs) along a curve to obtain the expected cost conditional on a win-set. Figure 2(d) illustrates this integration. In fact, the proof shows the existence of such a curve. The role of competitors' cost shifters is to shift the pivotal sets along the space of competitors' signals while w_0, d_i and s_i are held constant. w_0 is omitted to ease notation. The signal s_i is held constant by considering the s_i-th quantile of the distribution of bids and marginal costs. Let $\Gamma(d_{-i}|s_i, d_i)$ denote the mapping that returns the tie-point as cost shifters d_{-i} change:

$$\Gamma(d_{-i}|s_i, d_i) \equiv \left\{G_{B_j|w}\left(G^{-1}_{B_i|w}(s_i)\right)\right\}_{j\neq i}, \text{ where } w = \{d_i, d_{-i}\}. \tag{13}$$

The first condition to identify i's expected costs is that the support of d_{-i} is large enough so that it is possible to find a d_{-i} such that $\Gamma(d_{-i}|s_i, d_i) = \sigma$ for all σ in the win-set. The second condition requires that i always faces some competition along the curve.

Proposition 1 *The expected cost* $E[C_i|S_{-i} \geq s_{-i}, s_i, d_i]\, P[S_{-i} \geq s_{-i}]$ *is identified if there exists an* ε *such that for all* $\sigma \geq s_{-i}$*, there are a* d_{-i} *and a* $j \neq i$ *such that* $\Gamma(d_{-i}|s_i, d_i) = \sigma$ *and* $g_{B_j|(d,d_{-i})}\left(G^{-1}_{B_i|(d,d_{-i})}(s_i)\right) > \varepsilon$.
Proof. *See Appendix B.* ■

Because $F_{S|w_0}$ is identified, it is straightforward to identify the full information cost (the expectation conditional on a point) from the expected cost in the neighborhood of the point.

Proposition 2 *The full information cost* $E[C_i|s_{-i}, s_i, d_i]$ *is identified if*

$$E[C_i|S_{-i} \geq s_{-i}, s_i, d_i] \, P[S_{-i} \geq s_{-i}] \tag{14}$$

is identified for all σ *in a neighborhood of* s_{-i}.

Proof. *Let* $\phi(s_{-i}) = E[C_i|S_{-i} \geq s_{-i}, s_i, d_i] \, P[S_{-i} \geq s_{-i}]$. *Differentiating with respect to* s_{-i}:

$$\frac{d^{n-1}\phi(s_{-i})}{ds_{-i}} = E[C_i|s_i, s_{-i}, d_i] \, f(s_i, s_{-i}), \tag{15}$$

where $f(s_i, s_{-i})$ *is the density of the joint distribution of signals which is identified.* ∎

5 Estimation

5.1 Overview

I use bids on the 3,851 HMA projects described above to estimate the full information costs and the joint distribution of signals for the 50 most active firms in the sample. Although there were 411 different bidding firms, only 50 participated in more than 100 auctions. I pool together the other 361 firms and treat them as a single fringe bidder. For all practical purposes, the number of potential participants in every single auction is $n = 51$; however, depending on the auction location and characteristics, some of these 51 bidders may have a very low probability of participation.

The data contains bids, auction covariates w_0, and project and firm locations. Bids are normalized by the engineer's estimate. The vector w_0 contains several sets of dummy variables: year of the auction; project size; and the following prequalification requirements: HMA, aggregate construction, bridges and special structures, concrete items, and sewers and water mains. Tables 1(b) and 1(c) summarize the observed frequencies of these control variables. Own distance to the project is discretized in seven bins: 0-10, 10-25, 25-50,50-100, 100-150, 150-200, and 200-500 km.

Parameters of the model. Although the joint distribution of signals and full information costs are identified without any parametric or functional form assumptions, any fully nonparametric estimation would be plagued by the curse of dimensionality. I address this problem assuming some functional forms. The joint distribution of signals is assumed to be a Gaussian copula with covariance matrix Σ. Moreover, Σ is assumed to have a factor structure: $\Sigma = LL' + \Lambda$, where L is a 51 x l loading matrix and Λ is a diagonal matrix (Λ is restricted so that the elements on the main diagonal of Σ are all one). Restricting $l < 50$ reduces the number of free parameters. Full information costs are assumed to be additively separable in auction covariates and competitors' signals, but own signals are allowed to enter more flexibly:[12]

$$E(C_i|s, d_i, w_0) = \delta_{wi}(s_i) \, w_0 + \delta_{di}(s_i) \, d_i + \sum_{j \neq i} \delta_{ji}(s_i) \, \psi(s_j) + \delta_{ii}(s_i), \tag{16}$$

[12] For example, full information costs are additively separable if the joint distribution of costs and signals is elliptical (Kelker, 1970).

where $\psi(S_j)$ rescales j's signal so that it is marginally distributed as a standard normal, and $\delta_{ji}(s_i)$ captures the effect of one standard deviation in j's (rescaled) signal on i's full information costs. Notice that the parameters δ are indexed by s_i; I will use a quantile regression approach and estimate δ for different values of s_i. In sum, the parameters of the model are those in L and the δs in the full information costs for $i = 1, ..., 51$.

The functional form (16) implies that the expected marginal cost $mc_i = E[C_i | M_i = b_i, s_i, w]$ can be written as:

$$mc_i = \delta_{wi}(s_i) w_0 + \delta_{di}(s_i) d_i + \sum_{j \neq i} \delta_{ji}(s_i) \tilde{\psi}_j(b_i, w) + \delta_{ii}(s_i), \tag{17}$$

where $\tilde{\psi}_j(b_i, w) = E[\psi(S_j) | M_i = b_i, w]$, i.e., the expectation of $\psi(S_j)$ conditional on the pivotal set. While mc_i and $\tilde{\psi}_j(b_i, w)$ are not directly observed, they can be estimated for each observed bid. This equation is used to estimate the δs.

Estimation summary. The estimation procedure has two stages. In the first stage, I estimate the main features of the joint distribution of bids conditional on observables. The location and scale parameters of each firm's distribution of bids and its probability of participation are estimated semiparametrically to allow for a flexible effect of project locations. The covariance matrix of the joint distribution of signals is estimated using an Exploratory Tobit Factor Analysis (Kamakura & Wedel 2001). These first stage estimates are used to compute mc_i and $\tilde{\psi}_j(b_i, w)$ in equation (17). In the second stage, I use quantile regression methods to test the null hypothesis of private costs, and instrumental variable quantile regression to estimate the parameters in (17). Figure 3 illustrates and summarizes the estimation procedure.

The first stage has two steps. The first one is to estimate each firm's marginal distribution of bids conditional on observable characteristics $w = (d_1, ..., d_n, w_0)$, where d_i is the distance of bidder i to the project. For fixed plant locations, the vector of distances can be derived from the geographical coordinates of the project $(east, north)$. I reduce the dimensionality of the marginal bid distribution by $n - 2$ variables by conditioning on $\tilde{w} = (east, north, w_0)$. In other words, I condition on the geographical coordinates of the project instead of the vector of distances.

Each firm is observed to bid in projects that are geographically close to each other. Some important features of the distribution of bids conditional on $\tilde{w}$ can be estimated by semiparametric methods where bids for neighboring projects are weighted according to their distance to $\tilde{w}$. I estimate the following features of the distribution of bids: the probability of participation, $\hat{\pi}_i(\tilde{w})$; the expected bid conditional on participation, $\hat{\mu}_i(\tilde{w})$; and the variance of bids conditional on participation, $\hat{\sigma}_i^2(\tilde{w})$. I standardize each observed bid by the estimated mean and variance, and compute the empirical distribution of the standardized bids conditional on participation, $\hat{H}_i$, assuming that it does not vary across different geographical locations. The estimated marginal distribution of firm i's bids is:

$$\hat{G}_{B_i|\tilde{w}}(b) = \hat{P}(B_i < b|\tilde{w}) = \hat{H}_i\left(\frac{b - \hat{\mu}_i(\tilde{w})}{\hat{\sigma}_i(\tilde{w})}\right)\hat{\pi}_i(\tilde{w}). \tag{18}$$

This estimation procedure allows the geographical coordinates to affect the location and scale of the distribution of bids conditional on participation as well as the probability of participation, but restricts the shape of the distribution of bids to be the same across different project locations. If $\hat{\pi}_i$, $\hat{\mu}_i$, and $\hat{\sigma}_i^2$ are restricted to be constant across different locations, $\hat{G}_{B_i|\tilde{w}}(b)$ becomes the empirical distribution of bids conditional only on w_0, which is the estimator frequently used in the empirical auction literature when there are no observable cost shifters.

The second step is to estimate the joint distribution of signals across auctions. For each observed bid, I obtain $\hat{s}_i = \hat{G}_{B_i|\tilde{w}}(b_i)$, an estimate of the signal that prompted bid b_i. When a bidder is observed not to bid, it is only possible to infer that the signal was above the censoring point which is estimated by $\hat{\pi}_i(\tilde{w})$. The joint distribution of signals is a Gaussian copula with correlation matrix $LL' + \Lambda$, thus the likelihood of a vector of censored and uncensored observations depend on the parameters in the matrix L. I estimate L by Simulated Maximum Likelihood.

The second stage of the estimation procedure involves the estimation of the parameters in the full information cost functions. The estimation of (17) presents two challenges: First, the terms $\tilde{\psi}_j(b_i, w)$ are estimated with error which may result in biased estimates of the coefficients δ in small samples. I address the problem using competitors' distance as an instrument for $\tilde{\psi}_j(b_i, w)$. Notice that competitors' distances are excluded from the full information cost by Assumption 3, and they shift the terms $\tilde{\psi}_j(b_i, w)$ by their effect on the pivotal set. The main advantage of this approach is that that neither of the instruments is an estimated quantity. However, in small samples there is no guarantee that the bias introduced by this instrumental variables approach is smaller than the bias due to measurement error that it tries to avoid. The second challenge is that mc_i and $\tilde{\psi}_j(b_i, w)$ can only be estimated when the firm was observed to bid. Most censored data methods require the knowledge of the censoring point, in this case, the maximum marginal cost under which the bidder participates. Alternatively, Buchinsky and Hahn (1998) propose an censored quantile regression estimator that only requires a consistent estimate of the probability of censoring. This method is better suited for my purposes because the probability of censoring is just the probability of nonparticipation that can be consistently estimated from the data.

Under the private costs hypothesis, the distribution of marginal costs should not depend on competitors' distance (or signals). I test the hypothesis including competitors' distance in equation (17) instead of the terms $\tilde{\psi}_j(b_i, w)$. The coefficients are estimated using Buchinsky and Hahn's (1998) censored quantile regression method.

5.2 First stage: distribution of bids conditional on observables

5.2.1 Marginal distribution of bids

The estimation of i's marginal distribution of bids conditional on covariates $\tilde{w}$ is divided in four steps: the estimation of the probability of participation, $\hat{\pi}_i(\tilde{w})$; the expected bid conditional on participation, $\hat{\mu}_i(\tilde{w})$; the variance of bids conditional on participation, $\hat{\sigma}_i^2(\tilde{w})$; and the distribution of standardized bids conditional on participation, $\hat{H}_i(b^s)$ where $b^s = (b - \hat{\mu}_i(\tilde{w})) / \hat{\sigma}_i(\tilde{w})$.

Probability of participation. Let Y_i be 1 if firm i participates in the auction and 0 if it does not. $P(Y_i = 1|\tilde{w})$ is the probability that firm i participates in an auction with observable characteristics $\tilde{w} = (east, north, w_0)$. Assume that:

$$P(Y_i = 1|\tilde{w}) = \Phi\left(w_0 \gamma_{\pi i} + h_{\pi i}(east, north)\right), \tag{19}$$

where Φ is a standard normal cumulative distribution function and h_i is a flexible function of the location of the project. For each firm i, I estimate $\gamma_{\pi i}$ and $h_{\pi i}$ using an iterative procedure that is explained in Appendix C. The estimated probability is $\hat{\pi}_i(\tilde{w}) = \Phi\left(w_0\hat{\gamma}_{\pi i} + \hat{h}_{\pi i}(east, north)\right)$

Expected bid conditional on participation. Assume that the expected bid conditional on participation is:

$$E(b_i|\tilde{w}, Y_i = 1) = w_0\gamma_{\mu i} + h_{\mu i}(east, north), \tag{20}$$

where $h_{\mu i}$ is a smooth flexible function of the geographical coordinates of the project. This is a standard partial linear model and $\gamma_{\mu i}$ and $h_{\mu i}$ can be estimated using standard semiparametric methods.

First, $E(b_i|east, north, Y_i = 1)$ and $E(w_0|east, north, Y_i = 1)$ are estimated nonparametrically using a kernel smoothing over the location of the project with a 10 km bandwidth. $\hat{\gamma}_{\mu i}$ can be obtained directly by regressing $b_i - \hat{E}(b_i|east, north, Y_i = 1)$ on $w_0 - \hat{E}(w_0|east, north, Y_i = 1)$, where $\hat{E}$ denotes the nonparametric estimates. Finally, the estimated expected bid conditional on entry for bidder i is:

$$\mu_i(\tilde{w}) = \left(w_0 - \hat{E}(w_0|east, north, Y_i = 1)\right)\hat{\gamma}_{\mu i} + \hat{E}(b_i|east, north, Y_i = 1). \tag{21}$$

Variance of bids conditional on participation. The estimation of the bid variance conditional on participation is complicated by the fact that the specification should predict only positive variances. Consider the following specification:

$$E\left((b_i - \mu_i(\tilde{w}))^2 |\tilde{w}, Y_i = 1\right) = \exp\left(w_0\gamma_{\sigma i} + h_{\sigma i}(east, north)\right), \tag{22}$$

where $h_{\mu i}$ is a smooth flexible function of the geographical coordinates of the project. The estimation of $\gamma_{\sigma i}$ and $h_{\sigma i}$ follows the same iterative procedure as the estimation of the probability of entry and is discussed in Appendix C. Plugging the estimates $\hat{\gamma}_{\sigma i}$ and $\hat{h}_{\sigma i}$, I obtain the fitted values:

$$\hat{\sigma}_i^2(\tilde{w}) = \exp\left(\tilde{w}_0\hat{\gamma}_{\sigma i} + \hat{h}_{\sigma i}(east, north)\right) \tag{23}$$

Other features of the distribution. Bids are standardized by their estimated mean and standard deviation: $b_i^s = \frac{(b_i - \hat{\mu}_i(\tilde{w}))}{\hat{\sigma}_i(\tilde{w})}$. Let $\hat{H}_i(b^s)$ be the empirical distribution of b_i^s across auctions for bidder i. The estimated distribution of firm i's normalized bids B_i conditional on $\tilde{w}$ is:

$$\hat{G}_{B_i|\tilde{w}}(b) = \hat{P}(B_i < b|\tilde{w}) = \hat{H}_i\left(\frac{b - \hat{\mu}_i(\tilde{w})}{\hat{\sigma}_i(\tilde{w})}\right)\hat{\pi}_i(\tilde{w}) \tag{24}$$

5.2.2 Joint distribution of signals

The joint distribution of signals is assumed to be a Gaussian copula with correlation matrix $\Sigma = LL' + \Lambda$, where L is 51 x l and Λ is a diagonal matrix. The random vector $Z = \{\psi(S_1), ..., \psi(S_n)\}$ is distributed multivariate normal with covariance matrix Σ. Consider the following factor structure for Z:

$$Z = LX + \varepsilon \tag{25}$$

where X is a vector of l orthonormal Gaussian factors and ε is a vector of n Gaussian random variables with covariance Λ. ε and X are independent. Notice that $E(ZZ') = LL' + \Lambda$ as desired.

I compute the estimated (rescaled) signal that prompted each observed bid as $\psi\left(\hat{G}_{B_i|\tilde{w}}(b_i)\right)$. Moreover, if firm i did not participate in the auction I only estimate a lower bound for the (rescaled) signal $\psi(\hat{\pi}_i(\tilde{w}))$. Let $y_{ia} = 1$ and $\hat{z}_{ia} = \psi\left(\hat{G}_{B_i|\tilde{w}}(b_i)\right)$ if bidder i participated in auction a, and $y_{ia} = 0$ and $\hat{z}_{ia} = \psi(\hat{\pi}_i(\tilde{w}))$ otherwise. The likelihood of $((\hat{z}_{1a}, y_{1a}), ..., (\hat{z}_{na}, y_{na}))$ given L is:

$$\int \prod_{i:y_{ia}=1} \frac{1}{\lambda_i}\phi\left(\frac{\hat{z}_{ia} - Lx}{\lambda_i}\right) \prod_{i:y_{ia}=0} \left[1 - \Phi\left(\frac{\hat{z}_{ia} - Lx}{\lambda_i}\right)\right] \phi(x, I)\, dx, \tag{26}$$

where λ_i is the i-th element in the main diagonal of Λ and $\phi(x, I)$ is the density of an l-variate normal with covariance matrix I evaluated at x. This likelihood is approximated by:

$$T^{-1} \sum_{t=1}^{T} \prod_{i:y_{ia}=1} \frac{1}{\lambda_i}\phi\left(\frac{\hat{z}_{ia} - Lx_t}{\lambda_i}\right) \prod_{i:y_{ia}=0} \left[1 - \Phi\left(\frac{\hat{z}_{ia} - Lx_t}{\lambda_i}\right)\right], \tag{27}$$

where x_t is drawn T times from $N(0, I_l)$. The simulated log-likelihood function for the whole sample becomes:

$$\mathcal{L} = \sum_{a=1}^{A} \log\left(T^{-1} \sum_{t=1}^{T} \prod_{i:y_{ia}=1} \frac{1}{\lambda_i}\phi\left(\frac{\hat{z}_{ia} - Lx_t}{\lambda_i}\right) \prod_{i:y_{ia}=0} \left[1 - \Phi\left(\frac{\hat{z}_{ia} - Lx_t}{\lambda_i}\right)\right]\right), \tag{28}$$

where $A = 3,851$ auctions.

Any orthogonal rotation of L yields the same distribution of Z, so L has to be restricted so that no pair of admissible matrices are an orthogonal rotation of each other. This is achieved by restricting $l(l-1)/2$ coefficients on L. Because Σ is a copula, I restrict Λ so that $\hat{L}\hat{L}' + \hat{\Lambda}$ has all ones in the main diagonal. Kamakura & Wedel (2001) show that L and Λ can be consistently estimated by Simulated Maximum Likelihood.

The choice of l, the number of factors, trades off model fit and model complexity. A large l allows for a flexible covariance matrix, but increases the number of parameters to estimate and may result in overfitting. An appropriate choice should capture the main correlations in the sample with fewer parameters. The Bayesian Information Criterion selects the the model that better fits the data penalizing the number of parameters.

$$BIC = -2\ln \mathcal{L}_l + (nl - l(l-1)/2)\ln(nA) \tag{29}$$

I estimate the model for $l = 0, 1, 2, 3$ and 4. The minimum BIC is achieved when $l = 2$.[13] The estimated joint distribution of signals is denoted by $\hat{F}_S$.

5.2.3 Marginal costs and effect of competitors' signals

The estimates of the joint distribution of signals and each firm's marginal distribution of bids are used to compute the marginal costs $m\hat{c}_i$ and the effect of competitors' signals $\tilde{\psi}$. The expected marginal cost that rationalizes each bid is estimated by

$$m\hat{c}_i = b_i - \frac{\left(1 - \hat{G}_{M_i|B_i,\tilde{w}}\left(b_i|b_i\right)\right)}{\hat{g}_{M_i|B_i,\tilde{w}}\left(b_i|b_i\right)}, \tag{30}$$

where $\hat{G}_{M_i|B_i,\tilde{w}}\left(b_i|b_i\right)$ and $\hat{g}_{M_i|B_i,\tilde{w}}\left(b_i|b_i\right)$ are derived from those estimates as shown in Appendix D.

The assumptions that $\psi\left(S_j\right)$ is distributed as a standard normal and that the joint distribution of signals is a Gaussian copula reduce the problem of calculating $\tilde{\psi}_j\left(s_i, w\right)$ to calculating the expectations of truncated multivariate normal distributions. These expectations can be computed by quasi-Monte Carlo integration methods. Moreover, because the correlation matrix of signals can be decomposed in a factor model of dimension $l < n$, the numerical integration has to be performed on an l-dimensional space. Appendix D.1 presents the algorithm used to perform these integrations in detail.

5.3 Second stage: testing and estimation of the model primitives

Private costs hypothesis test. The private cost hypothesis and the exclusion restrictions imply that the distribution of marginal costs should not depend on competitors' distances. Under the alternative hypothesis, competitors' distances affect the distribution of marginal costs through the effect on the pivotal set. I estimate a quantile regression for each firm's marginal costs. The specification includes the controls w_0 described above, the bidder own distance and the two main competitors' distance to the project. Under the null hypothesis the coefficient associated with competitors' distances should be zero.

Because the marginal costs are only observed when the firm bids, and the firm bids when it has low signals, the observed distribution of marginal costs is right-censored. The censoring point is unobserved but the censoring probability is identified. I therefore use the following general censored quantile regression approach due to Buchinsky and Hahn (1998) that only requires a

[13] The Akaike Information Criterion (AIC) imposes a weaker penalty on the inclusion of additional parameters. This criterion tries to select the model that most adequately describes an unknown, high dimensional reality. $AIC = -2\ln\mathcal{L}_l + 2\left(nl - l\left(l-1\right)/2\right)$. The minimum AIC is achieved for l=3.

consistent estimate of this probability for each non-censored observation:[14]

$$\begin{aligned} \hat{\kappa}_i^{(\tau)} &= \arg\min_{\kappa} \sum_{a=1}^{A} y_{ia} \rho_{\min\left(\frac{\tau}{\hat{\pi}_{ia}},1\right)} \left(m\hat{c}_{ia} - \kappa_{wi} w_{0a} - \kappa_{di} d_i - \kappa_{dj} d_j \right) \qquad (31) \\ \text{where } \rho_{\kappa}(u) &= u\left(\kappa - 1\left(u<0\right)\right). \text{ Null hypothesis: } \kappa_{dj} = 0. \end{aligned}$$

where $A = 3,851$. The probability that firm i participates in auction a, $\hat{\pi}_{ia}$, was estimated semi-parametrically as described in section 5.2.1. Buchinsky and Hahn (1998) show that this estimator is consistent and asymptotically normal. The main intuition is that the τ-th quantile of the uncensored distribution of the dependent variable is the $\frac{\tau}{\pi_{ia}}$-th quantile of the censored distribution.

Estimation of full information costs. The identification results suggest that it should be possible to find a unique set of full information costs that rationalizes the observed $G_{B|w}$ so no further estimation should be required. I made several assumptions to overcome the curse of dimensionality while allowing enough flexibility to capture the main moments and features of the bid distribution. It is nonetheless hard to know what restrictions these assumptions impose on the underlying full information costs. To address this concern, I assume a functional form for the full information cost given by (16), and obtain an estimable equation that only requires estimating the marginal cost $m\hat{c}_i$, and the effect of competitors' signals $\left\{\tilde{\psi}_j(b,w)\right\}_{j\neq i}$ for each observed bid. The estimation of the marginal cost $m\hat{c}_i$ involves estimating the markup each firm adds to its marginal cost that can be inferred from the main features of competitors' bidding behavior. Similarly, the terms $\left\{\tilde{\psi}_j(b_i,w)\right\}_{j\neq i}$ capture the relative strength of each bidder at each auction which also is derived from those main features.

As mentioned above, estimation of (17) presents two challenges: first stage estimation error of $\tilde{\psi}_j(b_i,w)$, and censoring introduced by nonparticipation. To address the first challenge, competitor j's distance is used as an instrument for $\tilde{\psi}_j(b_i,w)$. While j's distance is excluded from the full information cost by Assumption 3, it affects the signals in the pivotal set and $\tilde{\psi}_j(b_i,w)$. The estimation method accounts for censoring following Buchinsky and Hahn's (1998) censored quantile regression approach.

I run the censored quantile regression of the τ-th quantile of $m\hat{c}_i - \sum_{j\neq i}\delta_{ji\tau}\tilde{\psi}_j(b_i,w)$ on covariates w_0, own distance dummies d_i and dummies for two main competitors' distances d_{-i}. Following Chernozhukov and Hansen (2007), I run this regression for a grid of values for $\delta_{ji\tau}$, and choose those that minimize the Wald test for the coefficients on the instruments d_{-i}. For the main set of results, it is assumed that $\delta_{ji\tau} = \delta_{-ii\tau}$ for all $j \neq i$. In other words, each competitor's signal affect bidder i's full information cost symmetrically.

[14] Buchinsky and Hahn (1998) use trimming functions for observations with predicted probabilities $\hat{\pi}_{ia}$ close to τ. However, they find little impact on the trimming functions on the estimators. I performed a Monte Carlo exercise and found that using min $\left(\frac{\tau}{\hat{\pi}_{ia}},1\right)$ in the check function instead of the trimming functions produced more precise estimators. In any case, the results presented here do not change much if I use trimming functions.

6 Results

This section presents the estimation results of the marginal distribution of bids, the joint distribution of signals and the full information costs. I also show the results of the private cost hypothesis test.

6.1 Marginal distribution of bids

I estimate the distribution of bids for each of the 50 largest firms in the sample and for a single entity that groups all the other smaller firms. As shown in Section 3, firms participate and bid lower on projects close to its plants. The semiparametric estimates of the effect of geography confirm these results. Table 4 shows the semiparametric estimation results on the probability of entry for firms 1 to 5. Notice that firms 1-3 participate more often in projects that require the prime contractor to be prequalified to perform HMA work. However, firms 4 and 5 participate with higher probability when construction prequalification is required. The analysis of mean bids conditional on entry, however, does not show strikingly different behavior of firms across different types of projects.

6.2 Correlation of signals

Figure 4 shows the estimated matrix of correlation of signals for the largest 10 firms. Consider two groups of firms: firms 1, 2, 3, 7, 8 and 10 in one group, and firms 4, 5, and 6 in the other. The signals of any two firms within the same groups are very correlated, but the correlation between two firms in different groups is quite low. Firm 9's signals seem to be correlated with firms in both groups.

The evidence from the semiparametric estimates of the probability of entry also point towards this differentiation between firms. Firms from the first group participate more often when the project requires the prime contractor to be prequalified to do HMA-work, and less often when the projects requires construction prequalification. The opposite holds for firms in the second group. Firm 9, however, actively participates when either type of work is required. While all 10 firms are prequalified in both work-types, it seems that firms are specialized in one of the two types of work. In fact, firms in the first group (Rieth-Riley Construction, Ajax Paving, Barret Paving Materials) are larger and own many asphalt plants in the region. Firms in the second group (Kamminga & Roodvoets, C & D Hughes) do not own asphalt plants, but own paving equipment. In a project that requires paving, they can buy asphalt from the first group of firms.

These correlation patterns are consistent with an affiliated private cost model, a common cost model or a model with heterogeneity unobserved to the econometrician but observed by bidders. The difference between these three models lies on the information structure of signals and costs (Hong and Shum, 2003). Although I cannot rule out the presence of unobserved heterogeneity, it seems that the main heterogeneity between auctions is driven by the type of work involved, which I am controlling for. Below, I show that the affiliated private cost model is rejected by the data for 6 of the 10 most active firms, while the common cost model rationalizes the observed data well.

6.3 Private costs hypothesis test

The null hypothesis of private cost is tested by regressing mc_i on own distance, competitors' distances and auction observed characteristics. Tables 5(a) to 5(c) show the censored quantile regression results for firms 1 to 3 for three different quantiles: 0.25, 0.5 and 0.75. The first row shows that the constant term is increasing with the quantile as expected. The first set of coefficients, denoted by $ownd$, show that the marginal cost is increasing in the firm's own distance to the project for each quantile (the omitted category is 0 to 10 km). The second set of coefficients, denoted by $Firm$ j, show how the marginal cost of the firm changes with competitor's j's distance. I include the two main competitors' distance and the omitted category is more than 50 km.

Competitors' distance should not affect the bidders' marginal costs under private costs. However, in the presence of a common cost component, competitors' proximity should increase marginal costs because it worsens the winners' curse. Therefore, the coefficients on $Firm\ j$ should be positive.

Table 5(a) shows that firm 1's marginal costs are higher in projects close to the locations of firms 2 and 3 which is consistent with the presence of a common cost component. Table 5(b) shows that firm 2 has higher marginal costs close to 1, but not next to firm 8. The evidence for firm 3 is less convincing; firm 3's marginal costs seem to be higher in projects close to 7. Of the 10 most active firms, the null of private costs can not be rejected for firms 5, 6, 9 and 10.

6.4 Full information expected cost

The parameters in equation (17) are estimated by censored quantile regression—to account for nonparticipation—and the terms $\tilde{\psi}_j(S_i, X)$ are instrumented with competitors' distance. The specification also includes controls for observed auction characteristics. The coefficient δ_{ji} captures the effect of one standard deviation change in $\psi(S_j)$ on i's full information cost. Because $\psi(S_j)$ is distributed standard normal, δ_{ji} is the effect of an increase in S_j from the 0.31 to the 0.69 quantile. The value of the constant term at the different quantiles indicates the effect of the firm's own signal at the omitted category. To facilitate the comparison between the effect of own and competitors' signals, the regression results are shown for three different quantiles: 0.31, 0.5 and 0.69.

I estimate a single coefficient for competitors' signals $\delta_{-ii} = \delta_{ji}, \forall j \neq i$. In other words, the effect of j's and k's signals on firm i's cost is restricted to be the same. Tables 6(a)and 6(b) show the results for firms 1 and 2. As expected, the full information cost increase with own distance. For example, firm 1's costs increase by less than 3% of the engineer's estimate when projects are located between 25-50km, by 15% for projects located between 50-100km and by 35% between 100-150km.

The effect of competitors' signals seems to be positive but not as large as the effect of one's own signal. For firm 1, a change in its own signal from the 0.31 to the 0.69 quantile increases its full information costs by 10% of the engineer's estimate, while a similar increase in a competitor's signal increases its costs only by 3%. The results for firms 2 are similar. Table 7 summarizes the effect of a similar change own and competitors' signals for the 10 largest firms. The first noticeable pattern is that own signal effects are larger suggesting that there is an important component of costs that is idiosyncratic to the firm, but the effect of competitors' signals is non-negligible suggesting that there is a common cost component as well.

According to the coefficients in Table 7, if firm 2 receives a signal that increases its expected costs by \$1,310 (\$1,000 x the coefficient of firm 2's own signal's effect), firm 1's expected costs increase by \$270 (\$1,000 x the coefficient of firm 1's competitors' signal's effect) upon learning it. Similarly, if firm 1 receives a signal that increases its costs by \$1,000, firm 2's expected costs increase by \$290. The effects for firms 4-6 are smaller: if firm 4 receives a signal that increases its expected costs by \$1,150, firm 5's expected costs would increase by \$80 if it learns it.

Finally, I estimate a specification where the effect of the signal of the main competitor is allowed to differ from that of the other firms. While the coefficients on own distances are similar to those in the first set of results, the coefficients on competitors' signals are not significantly different from each other. It seems that there is not enough statistical power to distinguish the effect of different competitors' signals on the full information cost.

7 The effect of competition

The estimates from Section 6 are used to evaluate the effects of increased competition on bidding behavior and procurement costs. I simulate the effect of a reduction in one of the existing firms' cost. As a result, all other firms face increased competition, while the information structure and the set of potential bidders are kept fixed. Such a reduction in costs can be thought to be the result of the firm owning a plant closer to the project site, a technological change or a subsidy, i.e., an extra payment that the firm receives if it wins the auction. The case of a subsidy is of special interest because it is a tool available to procurement agencies.[15]

The auction I analyze is a 3.51 mile repaving project 10 miles southeast of Grand Rapids that requires the prime contractor to be prequalified to perform hot mix asphalt work. The engineer's estimate was \$502,029 and only two firms participated: firm 1 won with a bid of \$ 491,668 (98% of the engineer's estimate) while firm 2 bid \$528,126 (105%). Given the auction's characteristics and location, out of the 51 firms only firms 1, 2, 5 and 6 had an estimated probability of entry greater than 0.03. Table 8 shows this set of bidders, their initial locations, probabilities of entry, and median completion costs. I also include firm 3, the firm that becomes more competitive in the counterfactual.

I simulate the effect of a \$182,000 (36% of the engineer's estimate) subsidy to firm 3, which is equivalent to a reduction in its distance from 149 km to 25-50 km, under three different information structures. In the first, the coefficients estimated above are used to simulate the interdependent or common costs (CC) model.[16] In the second, all information interdependence is removed by setting $\delta_{ji} = 0$ for all $i \neq j$, but the correlation of signals given by $\hat{\Sigma}$ is kept; this is the affiliated private costs (APC) model. Finally, all information interdependence is removed and the matrix of correlation of signals is set equal to the identity matrix; this is the independent private costs

[15]The MDOT has no subsidy scheme in place but some projects require the prime contractor to subcontract to disadvantaged businesses for a percentage of the total contract value.

[16]Appendix E describes the the numerical algorithm to solve for the equilibrium bidding strategies.

(IPC) model. I compute the effect on expected bids, winning bid, subsidies paid ($182,000 times the probability that firm 3 wins) and net costs of the subsidy (the effect on the winning bid plus the expected subsidies paid) for each of the three models. Table 9 presents the results.

The subsidy reduces firm 1's expected bid by $6,900 in the CC model, $9,200 in the APC model, and $11,400 in the IPC model. The difference between the reduction in the CC model and the APC model can be attributed to the effect of the common cost components, while the difference between the APC and IPC can be attributed to the nonindependence of signals. The effect of increased competition on firm 1's bids is 40% lower in the CC model relative to the IPC benchmark. The results for firm 2 are similar. Firm 3 becomes more competitive not only because of the subsidy but also because it is subject to a lesser winner's curse. Without the subsidy it would only win on a very selected set of signals of bidders 1 and 2. With the subsidy it wins on a less selected set of signals which reduces its cost conditional on winning. Firm 3 expected bid conditional on participation falls by $76,400 in the CC model and by only $58,100 in the APC model, i.e., the winner's curse accounts for a $18,300 reduction in its average bid.

The winning bid decreases by $20,900 under an IPC model, but by only $15,200 under the APC and CC models. The effect of competition on prices is thus 28% lower in the CC model relative to the IPC benchmark. Notice that the effect of the winner's curse on the winning bid is almost zero due to the pro-competitive effect on firm 3 that offsets the anti-competitive effect on firms 1 and 2. The expected subsidies paid are higher under the IPC and CC models, where firm 3 wins with higher probability. Due to the winner's curse, firms 1 and 2 bid more cautiously and firm 3 bids more aggressively. As a result, firm 3 wins and the subsidy is paid more often in the CC model. The total cost of the subsidy is greatest under the CC model, followed by the APC model, and lowest under the IPC model: $26,500, $23,100 and $21,000, respectively. In sum, the effect of nonindependent signals increases the cost of the subsidy by 11% and the effect of common components by an additional 16%.

8 Conclusion

This paper examines the effect of competition on bidder behavior and prices in the context of Michigan highway procurement. A bidder's distance to the project is an important determinant of costs and an important predictor of his participation and bids. However, there is no evidence of more aggressive bidding when competitors are located close to the project. This pattern is hard to reconcile with predictions based on first-price auctions with private costs, but can be rationalized by a model with nonindependent private information and common costs.

The empirical findings point to non-negligible common cost components. I estimate that if a bidder receives a signal that increases his expected costs by $1000, his competitor expected costs increase by $70-$250 upon learning it. I also find that there exist two types of bidders: those that specialize in projects that only require HMA-work and those that specialize in projects that require construction work. The coefficient of correlation of signals for bidders of the same type is between 0.3 and 0.5, but bidders of different types show little pairwise correlation. To understand how common costs and signal correlation affect bidders reaction to increased competition, I simulate the effect of a reduction in costs to one of the existing firms under different competitive settings. The simulation results suggest that relative to the benchmark of independent private costs, these features reduce the effect of more intense competition on competitors' bids by 40% and on procurement costs by 28% .

Procurement agencies often subsidize a subset of high-cost bidders at a cost that depends on the nature of competition in the market. In the independent private cost benchmark, subsidies provide incentives to non-subsidized bidders to bid more aggressively which reduce or even offset the costs of the subsidy. These incentives are weakened in models with common costs or correlated private information. Moreover, in models with an important common component of costs, the subsidized bidder is able to compete on a more even ground with the non-subsidized bidders and suffers a less severe winner's curse, it bids even more aggressively and wins more often. Because the subsidy is paid only when the subsidized bidder wins, the overall cost of the subsidy may be significantly higher than in the benchmark model. Given my estimates on common costs and signal correlation, a subsidy that benefits weak bidders costs 27% more than in the independent private cost benchmark.

Common value components could well be a pervasive feature of procurement markets. Contractors willing to offer their services often face cost uncertainties that may also be shared by their competitors, e.g., conditions in the input markets or potential difficulties in the execution of the project. It is therefore important to devise methods to analyze the procurement process under these market features. In this paper, I consider a model that allows for asymmetric bidders, private and common cost components, and nonindependent signals, and show that the model primitives are nonparametrically identified from the joint distribution of bids conditional on bidder-specific cost shifters (such as distance of each bidder to the project). The variation in observable cost shifters introduces observable asymmetries, and as a result, generates variation in the level and composition of competition. The common components of costs are measured by the response (or lack thereof) of each bidder to varying competitive conditions.

The method proposed in this paper requires that the researcher observes enough variation in firm-specific cost shifters. This type of data is admittedly not available for every market environment. However, there are some markets that have been studied where the geographical location of the bidders is relevant for their competitiveness, e.g., timber (Athey and Levin, 2001; Haile, 2001; Haile, Hong and Shum, 2004; Roberts and Sweeting, 2010; Athey, Coey and Levin, 2011), off-shore drilling (Hendricks and Porter 1988, 1993; Hendricks, Pinkse and Porter, 2003), highway procurement (Porter and Zona, 2003; Marion, 2007; Einav and Esponda, 2008; Krasnokutskaya and Seim, 2008; Bajari and Tadelis, 2011; Groeger, 2011; Krasnokutskaya, 2011), snow removal (Flambard and Perrigne, 2006) and plastic recycling (Kawai, 2011). Moreover, cost shifters need not be related to geographical location, and may result from some exogenous variation in the data or some other firm characteristics that vary across auctions.

References

Athey, S. (2001): "Single Crossing Properties and the Existence of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69(4), 861–889.

Athey, S., D. Coey, and J. Levin (2011): "Set-asides and Subsidies in Auctions," Discussion paper, National Bureau of Economic Research.

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

——— (2006): "Empirical Models of Auctions," *National Bureau of Economic Research Working Paper Series*, No. 12126.

Athey, S., and J. Levin (2001): "Information and Competition in U.S. Forest Service Timber Auctions," *Journal of Political Economy*, 109(2), 375–417.

Ayres, I., and P. Cramton (1996): "Deficit Reduction Through Diversity: How Affirmative Action at the FCC Increased Auction Competition," *Stanford Law Review*, 48(4), 761–815.

Bajari, P., and A. Hortasu (2003): "The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions," *The RAND Journal of Economics*, 34(2), 329–355.

Bajari, P., S. Houghton, and S. Tadelis (2011): "Bidding for incomplete contracts: An empirical analysis," University of Minnesota, Texas A&M University and UC Berkeley.

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

Buchinsky, M., and J. Hahn (1998): "An Alternative Estimator for the Censored Quantile Regression Model," *Econometrica*, 66(3), 653–671.

Bulow, J., and P. Klemperer (2002): "Prices and the Winner's Curse," *The RAND Journal of Economics*, 33(1), 1–21.

Campo, S., I. Perrigne, and Q. Vuong (2003): "Asymmetry in firstprice auctions with affiliated private values," *Journal of Applied Econometrics*, 18(2), 179–207.

Chernozhukov, V., and C. Hansen (2006): "Instrumental quantile regression inference for structural and treatment effect models," *Journal of Econometrics*, 132(2), 491–525.

Chernozhukov, V., C. Hansen, and M. Jansson (2007): "Inference approaches for instrumental variable quantile regression," *Economics Letters*, 95(2), 272–277.

Einav, L., and I. Esponda (2008): "Endogenous Participation and Local Market Power in Highway Procurement," Stanford University and NYU Stern.

Flambard, V., and I. Perrigne (2006): "Asymmetry in Procurement Auctions: Evidence from Snow Removal Contracts," *The Economic Journal*, 116(514), 1014–1036.

Gayle, W., and J. F. Richard (2008): "Numerical Solutions of Asymmetric, First-Price, Independent Private Values Auctions," *Computational Economics*, 32(3), 245–278.

Genz, A. (1992): "Numerical Computation of Multivariate Normal Probabilities," *Journal of Computational and Graphical Statistics*, 1(2), 141–149.

Genz, A., and F. Bretz (2009): *Computation of Multivariate Normal and T Probabilities.* Springer Verlag, Berlin.

Groeger, J. (2011): "A Study of Participation in Dynamic Auctions," Carnegie Mellon Tepper School of Business.

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

Haile, P. A. (2001): "Auctions with Resale Markets: An Application to U.S. Forest Service Timber Sales," *The American Economic Review*, 91(3), 399–427.

Haile, P. A., H. Hong, and M. Shum (2004): "Nonparametric Tests for Common Values in First-Price Sealed-Bid Auctions," Yale U., Duke U. and Johns Hopkins U.

Hendricks, K., J. Pinkse, and R. H. Porter (2003): "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions," *The Review of Economic Studies*, 70(1), 115–145.

Hendricks, K., and R. H. Porter (1988): "An Empirical Study of an Auction with Asymmetric Information," *American Economic Review*, 78(5), 865–83.

Hendricks, K., and R. H. Porter (1993): "Bidding behaviour in OCS drainage auctions: Theory and evidence," *European Economic Review*, 37(2-3), 320–328.

Hill, J., and A. Shneyerov (2010): "Are There Common Values in First-Price Auctions? A Tail-Index Nonparametric Test," Concordia University.

Hong, H., and M. Shum (2002): "Increasing Competition and the Winner's Curse: Evidence from Procurement," *Review of Economic Studies*, 69(4), 871–898.

Hortacsu, A., and J. Kastl (2011): "Testing for Common Values in Canadian Treasury Bill Auctions," Discussion paper, Stanford Instititute for Economic Policy Research.

Hubbard, T. P., T. Li, and H. J. Paarsch (2011): "Semiparametric Estimation in Models of First-Price, Sealed-Bid Auctions with Affiliation," Discussion paper, Center for Economic Institutions, Institute of Economic Research, Hitotsubashi University.

Kamakura, W., and M. Wedel (2001): "Exploratory Tobit factor analysis for multivariate censored data," *Multivariate Behavioral Research*, 36(1), 5382.

Kawai, K. (2011): "Auction Design and the Incentives to Invest: Evidence from Procurement Auctions," NYU Stern.

Kelker, D. (1970): "Distribution Theory of Spherical Distributions and a Location-Scale Parameter Generalization," *Sankhy: The Indian Journal of Statistics, Series A*, 32(4), 419–430.

Krasnokutskaya, E. (2011): "Identification and estimation of auction models with unobserved heterogeneity," *The Review of Economic Studies*, 78(1), 293.

Krasnokutskaya, E., and K. Seim (2008): "Bid preference programs and participation in highway procurement auctions," University of Pennsylvania, Wharton.

Laffont, J., and Q. Vuong (1996): "Structural Analysis of Auction Data," *American Economic Review*, 86(2), 414–20.

Marion, J. (2007): "Are bid preferences benign? The effect of small business subsidies in highway procurement auctions," *Journal of Public Economics*, 91(7-8), 1591–1624.

Matthews, S. (1984): "Information Acquisition in Discriminatory Auctions," in *Bayesian Models in Economic Theory*, ed. by M. Boyer, and R. Kihlstrom, vol. 49, p. 14771500. North Holland.

Milgrom, P. (1989): "Auctions and Bidding: A Primer," *The Journal of Economic Perspectives*, 3(3), 3–22.

Milgrom, P. R., and R. J. Weber (1982): "A Theory of Auctions and Competitive Bidding," *Econometrica*, 50(5), 1089–1122.

Paarsch, H. J. (1992): "Deciding between the common and private value paradigms in empirical models of auctions," *Journal of Econometrics*, 51(1-2), 191–215.

Pinkse, J., and G. Tan (2005): "The Affiliation Effect in First-Price Auctions," *Econometrica*, 73(1), 263–277.

Porter, R. H., and J. D. Zona (1993): "Detection of Bid Rigging in Procurement Auctions," *Journal of Political Economy*, 101(3), 518–38.

——— (1999): "Ohio School Milk Markets: An Analysis of Bidding," *The RAND Journal of Economics*, 30(2), pp. 263–288.

Samuelson, W. F. (1985): "Competitive bidding with entry costs," *Economics Letters*, 17(1-2), 53–57.

Somaini, P. (2011): "Nonparametric Identification of Interdependent Values Auction Models," Stanford University.

APPENDIX

A Stylized Facts. Controlling for selection

This appendix shows that the results in Section 3 still hold after controlling for selection due to nonparticipation. The sample of observed bidders may be selected because only bidders with low realizations of costs decide to participate.

Consider a standard independent private cost auction model where bidders choose to participate in the auction only if the expected profit (at the optimal bid) exceeds some non-negative entry cost. Define the latent optimal bid as the bid that maximizes the bidder's profit after he the entry cost is sunk. The latent bid is a monotone function of his completion costs. Because firms participate only when their project completion cost is below some threshold, the observed distribution of bids in the data is a censored distribution of latent optimal bids.

I rely on a censored quantile regression procedure to estimate the relationship between the latent optimal bid and competitors' distance controlling for own bidder distance and the set of covariates described above. One of the advantages of this procedure is that it does not depend on distributional assumptions about the error term. More importantly, it allows me to estimate the coefficients of a linear specification without observing the censoring point. I use the estimator proposed by Buchinsky and Hahn (1998).

Let b_{ia} denote the bid of firm i in auction a, w_{0a} a vector of auction characteristics, and d_{ja} the distance from project a to the nearest plant of firm j. Moreover, let $y_{ia} = 1$ if firm i submitted a bid in auction a and 0 otherwise. Finally, let $\hat{\pi}_{ia}$ be a consistent estimate of the probability that bidder i participates in auction a (see Appendix C). For each firm i, I estimate:

$$
\begin{aligned}
\hat{\delta}_{\tau}^{(i)} &= \arg\min_{\delta} \sum_{a=1}^{A} y_{ia} \rho_{\min\left(\frac{\tau}{\hat{\pi}_{ia}},1\right)} \left(b_{ia} - w_a \delta\right) \\
\text{where } w_a &= \left[w_{0a}, d_{ia}, d_{ja}\right] \text{ and } \rho_{\kappa}(u) = u\left(\kappa - 1\left(u < 0\right)\right)
\end{aligned}
\tag{A.1}
$$

Table A shows the results for the 5 largest firms in the sample. As before, only the fifth largest firm bids consistently lower close to its competitors' plants.

One of the regressions of bids on competitors' distances in Section 3 and the specification in this appendix include a demand shifter as a control variable. I used geographical information on all roads available at the Michigan Center for Geographical Information to construct it.[17] I compute each road segment's length and assign a weight based on its classification (Interstates, 3.5; Freeways, 3; Principal Arterials, 2.5; Minor Arterials, 2.2; Major Collectors, 2; Minor Collectors, 1.5 and Local Roads, 1). For each project and plant location, I compute a measure of the density of roads using a 10 km bandwidth kernel.

[17] http://www.mcgi.state.mi.us/mgdl/framework/statewide/allroads_mi.zip

B Nonparametric identification

This appendix contains the proof of Proposition 1 of Section 4.2.2 (see Remark 5). The proposition states that the expected cost (the expected cost conditional on winning times the probability of winning) is identified from the distribution of bids conditional on cost shifters assuming that bidders participate with probability one and that entry costs are zero. The results presented below generalize this proposition to the case where entry costs are different from zero and bidders may choose not to participate in equilibrium.

Let the space of cost shifters be $D = \mathbb{R}^m$, $m \geq n$. Consider the function $\pi : D \to [0,1]^n$ such that $\pi(d) = \{P(b_j < \infty | d)\}_{j=1}^n$. This function returns the probabilities of participation given the vector of cost shifters d. To ease notation w_0 is omitted. It is assumed that there is a random ceiling or reserve price B_0 with cumulative distribution function R. R is differentiable for every b in the support $[\underline{b}_0, \bar{b}_0]$. B_0 is distributed independently from the vector of signals and costs, and $b - (R'(b))^{-1}(1 - R(b))$ is increasing in b. These assumptions are used to identify each bidder's entry cost (the cost of preparing a bid) and can be dispensed with if that cost is known. The MDOT, for example, always accepts bids below 110% of the engineer's estimate and rejects with some probability bids above that threshold. The maximum accepted winning bid was 215% of the engineer estimate. Whether the lower limit $\underline{b}_0$ binds depends on the vector of cost shifters. Redefine $M_i = \min(\min_{j \neq i} \beta_j(S_j, d), B_0)$ to simplify the notation.

Lemma 3 shows that entry costs are identified. The proof relies on the assumption that firm i's full information cost does not depend on the realization of the random reserve price. Lemma 4 states that if the entry cost is known, the expected cost is identified under some regularity conditions. This lemma is useful to show identification of the more general case that allows for nonparticipation but also to show Proposition 1. The full information cost for each bidder can be identified using Proposition 2 (in the main text).

B.1 Entry costs

Lemma 3 *The bid preparation cost k_i is identified if there exists a $d \in D$ such that: $\pi_k(d) = 0$ for all $k \neq i$ and $0 < \pi_i(d) < 1$.*

Proof. *Bidder i's expected profit is:*

$$(b - E[C_i|s_i, d_i])(1 - R(b)) - k_i, \tag{B.1}$$

where $E[C_i|s_i, d_i]$ is i's full information cost integrated over all competitors' signals. The first-order condition is:

$$E[C_i|s_i, d_i] = b - \frac{1 - R(b)}{R'(b)}. \tag{B.2}$$

Because $E[C_i|s_i, d_i]$ is continuous and increasing in s_i, i stops participating when his signal is such that his expected profits are zero. The cutoff signal is the probability of participation $\pi_i(d)$, and the maximum bid he submits is $b^* = G^{-1}_{B_i|d}(\pi_i(d))$, then

$$k_i = \frac{[1 - R(b^*)]^2}{R'(b^*)}. \tag{B.3}$$

■

B.2 Full information costs

According to the intuition in Section 4.2.2, identification is achieved integrating the expected cost conditional on pivotal sets (marginal costs) along a curve. The role of competitors' cost shifters is to shift the pivotal sets along the space of competitors' signals. The tie-point associated to each pivotal set as a function of cost shifters d_{-i} is:

$$\Gamma(d_{-i}|s_i, d_i) = \left\{G_{B_j|d}\left(G^{-1}_{B_i|d}(s_i)\right)\right\}_{j\neq i} \tag{B.4}$$

The following lemma formalizes that intuition.

Lemma 4 *The expected cost $E[C_i|S_{-i} \geq s_{-i}, s_i, d_i] P[S_{-i} \geq s_{-i}]$ is identified if:*

(i) s_{-i} is in the range of Γ,

(ii) there exist an $\varepsilon > 0$ such that for all $\sigma \geq s_{-i}$, if $\sigma = \Gamma(d_{-i}|s_i, d_i)$ there is at least one competitor j for which $g_{B_j|(d_i,d_{-i})}\left(G^{-1}_{B_i|(d_i,d_{-i})}(s_i)\right) > \varepsilon$,

(iii) for every $\sigma \geq s_{-i}$ if $\sigma = \Gamma(d_{-i}|s_i, d_i)$ then either $\pi_i(d_i, d_{-i}) = s_i$ or σ is an interior point of $Range(\Gamma) \cup \{\tau : \tau \geq \sigma\}^c$, and

(iv) the entry cost is known.

Proof. *To ease notation, let $\phi(\sigma) = E(C_i|s_i, d_i, S_{-i} \geq \sigma) P(S_{-i} \geq \sigma|s_i)$ and $\gamma(\sigma) = \Gamma^{-1}(\sigma|s_i, d_i)$ (if $\sigma \in Range(\Gamma)$, otherwise $\gamma(\sigma)$ does not exist). Define a parametric curve in the space of competitors' signals $\eta : \mathbb{R} \to [0,1]^{n-1}$ such that: $\eta(0) = s_{-i}$ and for each $j \neq i$*

$$\eta'_j(t) = \frac{\partial \beta_j^{-1}(b_t, d_t)}{\partial b} = g_{B_j|d_i,d_t}(b_t), \tag{B.5}$$

where $d_t = \gamma \circ \eta(t)$ and $b_t = \beta_i(s_i|d_i, d_t)$. By construction, if b_t and d_t exist, $\beta^{-1}(b_t, d_i, d_t) = \eta(t)$. Both d_0 and b_0 exist because of (i). By the non-vanishing competition condition (ii), there exist a j such that $\eta'_j(t) > \varepsilon$. Thus the curve does not converge to any point in the space of competitors' signals. By (iii), the curve $\eta(t)$ stays within the interior of $Range(\Gamma)$ until t reaches some t_1 where either $\eta_j(t_1) = 1$ for some j or $\pi_i(d_i, d_{t_1}) = s_i$. If the curve reaches $\eta_j(t_1) = 1$ for some j, $\phi(t_1) = 0$. If it reaches $\pi_i(d_i, d_{t_1}) = s_i$, because bidder i is indifferent between participating and not participating,

$$\phi(t_1) = \frac{b_{t_1}\left(1 - G_{M_i|b_{t_1},d_i,d_{t_1}}(b_{t_1}|b_{t_1})\right) - k_i}{(1 - R(b_{t_1}))}. \tag{B.6}$$

Firm i's problem is to maximize:

$$b\left(1 - G_{M_i|s_i,d}(b|s_i)\right) - \phi\left(\beta^{-1}(b, d)\right)(1 - R(b)). \tag{B.7}$$

The first-order condition can be written as:

$$b_i g_{M_i|B_i,d}(b|s_i) - \left(1 - G_{M_i|B_i,d}(b|b)\right) = -\frac{d\phi\left(\beta^{-1}(b, d)\right)}{db}(1 - R(b)) + \phi\left(\beta^{-1}(b, d)\right) R'(b). \tag{B.8}$$

Let $\phi(t) = \phi\left(\beta^{-1}(b_t, d_t)\right)$. *Similarly,*

$$\frac{d\phi\left(\beta^{-1}(b_t, d_t)\right)}{db} = \nabla\phi\left(\beta^{-1}(b_t, d_t)\right)\cdot\frac{d\beta^{-1}(b_t, d_t)}{db} = \nabla\phi(\eta(t))\cdot\eta'(t) = \phi'(t)\,. \tag{B.9}$$

For any $t \in [0, t_1]$ *the first-order condition (B.8) becomes:*

$$-\left[b_t g_{M_i|B_i,d_t}(b_t|s_i) - \left(1 - G_{M_i|B_i,d_t}(b_t|b_\tau)\right)\right](1 - R(b_t))^{-1} = \phi'(t) + \left(-R'(b_t)(1 - R(b_t))^{-1}\right)\phi(t) \tag{B.10}$$

Therefore, $\phi(t)$ *is defined by this ordinary differential equation with terminal condition given by (B.6).* ■

Remark 5 *Proposition 1 follows from Lemma 4. While premise (ii) is directly invoked, (i) and (iii) are implied by the fact that every* $\sigma \geq s_{-i}$ *is in the range of* Γ. *The solution to the differential equation is*

$$\phi(s_{-i}) = \int_0^{t_1} b_\tau g_{M_i|B_i,d_\tau}(b_\tau|b_\tau) - \left[1 - G_{M_i|B_i,d_\tau}(b_\tau|b_\tau)\right]d\tau + b_{t_1}\left(1 - G_{M_i|B_i,d}(b_{t_1}|b_{t_1})\right). \tag{B.11}$$

If the assumption that bidders participate with probability one is relaxed, the mapping Γ is not defined when $\pi_i(d) < s_i$, i.e., when bidder i participates with probability below s_i. The domain of Γ is $A = \{d_{-i} \in D_{-i} : \pi_i(d_i, d_{-i}) \geq s_i\}$. Premise (iii) requires that the points that Γ maps to the boundary of its range are mapped from the boundary of its domain. The following two conditions ensure the premise holds:

Condition 6 $\pi(d) : D \to [0,1]^n$ *is a continuous map.*

Condition 7 $\Gamma(d_{-i}|s_i, d_i) : A \to [0,1]^{n-1}$ *is an open map.*

Lemma 8 *Under Conditions 6 and 7, if* $s_{-i} = \Gamma(d_{-i}|s_i, d_i)$ *then either* s_{-i} *is an interior point of Range*(Γ) *or* $\pi_i(d_i, d_{-i}) = s_i$.

Proof. *If* $s_{-i} = \Gamma(d_{-i}|s_i, d_i)$, $s_{-i} \in$ *Range*(Γ) *and* $d_{-i} \in A$. *Suppose that* s_{-i} *is not an interior point of Range*(Γ), *then* d_{-i} *is not an interior point of* A *because* Γ *is an open map. Because* π *is a continuous map,* A *is closed, and boundary points satisfy* $\pi_i(d_i, d_{-i}) = s_i$. ■

Lemma 4 states that

$$E[C_i|S_{-i} \geq s_{-i}, s_i, d_i]\,P[S_{-i} \geq s_{-i}]\,. \tag{B.12}$$

is identified under Conditions 6, 7 and non-vanishing competition. Proposition 2 states that the full information cost is identified if (B.12) is identified in a neighborhood of s_{-i}.

C Semiparametric estimation of the probability of entry

This appendix discusses the semiparametric estimation of the probability of entry and the variance of bids conditional on participation. As mentioned in Section 5, both estimation procedures use a similar iterative routine.

The random variable Y_i denotes whether firm i participates in the auction, and $P(Y_i = 1|\tilde{w})$ is the probability that firm i participates in an auction with observable characteristics $\tilde{w} = (east, north, w_0)$. I assume that:

$$P(Y_i = 1|\tilde{w}) = \Phi\left(w_0 \gamma_{\pi i} + h_{\pi i}(east, north)\right), \tag{C.1}$$

where Φ is a standard normal cumulative distribution function and h is a flexible function of the location of the project. For a fixed $h_{\pi i}(east, north)$, the model (C.1) can be estimated using maximum likelihood:

$$\sum_{a=1}^{A} y_{ia} \log \Phi\left(w_{0a}\gamma + h_a\right) + (1 - y_{ia}) \log\left[1 - \Phi\left(w_{0a}\gamma + h_a\right)\right] \tag{C.2}$$

where y_{ia} is a dummy variable indicating if firm i participated in auction a, and $h_a = h_{\pi i}(east_a, north_a)$.

Notice that by the law of iterated expectations:

$$P(Y_i = 1|east, north) = E\left[\Phi\left(w_0 \gamma_{\pi i} + h_{\pi i}(east, north)\right)|east, north\right] \tag{C.3}$$

I estimate the nonparametric probability of participation conditional only on the geographical location, i.e., $P(Y_i = 1|east, north)$, using a bivariate kernel with a 10km bandwidth:

$$\begin{aligned} \hat{e}_{ia} &= \sum_{r=1} w(a, r)\, y_{ir}, \text{ where} \\ w(a, r) &= \frac{K\left(\frac{east_a - east_r}{h}\right) K\left(\frac{north_a - north_r}{h}\right)}{\sum_{t=1} K\left(\frac{east_a - east_t}{h}\right) K\left(\frac{north_a - north_t}{h}\right)} \end{aligned} \tag{C.4}$$

The estimator I use is defined as the pair $(\gamma, \{h_a\}_{a=1..A})$ that maximizes the likelihood in (C.2) subject to the sample analogue of restriction (C.3):

$$\hat{e}_{ia} = \sum_{r=1}^{A} w(a, r)\, \Phi\left(w_{0r}\gamma + h_a\right) \tag{C.5}$$

Using a Taylor expansion on (C.3) around $\left(\bar{\gamma}, \left\{\bar{h}_a\right\}\right)$ solving for h_a and plugging into (C.2) yields

$$\sum_{a=1}^{A} y_{ia} \log \Phi\left(\left(w_{0a} - \bar{w}_a\right)\gamma + \bar{h}_a + \bar{\gamma}\bar{w}_a\right) + (1 - y_{ia}) \log\left[1 - \Phi\left(\left(w_{0a} - \bar{w}_a\right)\gamma + \bar{h}_a + \bar{\gamma}\bar{w}_a\right)\right], \tag{C.6}$$

where:

$$\bar{w}_a = \frac{\sum_r w_{0r} w(a, r)\, \phi\left(w_{0r}\bar{\gamma} + \bar{h}_a\right)}{\sum_r w(a, r)\, \phi\left(w_{0r}\bar{\gamma} + \bar{h}_a\right)} \tag{C.7}$$

Estimation: Estimate (C.4), fix $h_a^{(0)} = \Phi^{-1}(\hat{e}_{ia})$ for every a and $\gamma^{(0)}$. These are the initial conditions of the iterative algorithm. For each $t = 1, 2, ...$:

1. Estimate $\hat{\gamma}$ in (C.6) by Maximum Likelihood where the Taylor expansion is taken around $\left(\gamma^{(t-1)}, \left\{h_a^{(t-1)}\right\}_{a=1..A}\right)$. Fix $\gamma^{(t)}$ equal to this estimate.

2. Obtain $\left\{h_a^{(t)}\right\}_{a=1..A}$ that satisfies (C.5) given $\gamma^{(t)}$.

3. Stop if $\left\|\gamma^{(t)} - \gamma^{(t-1)}\right\|$ is below some tolerance level, otherwise repeat 1-3.

The variance of bids is estimated semiparametrically using the same procedure. For each observed bid calculate the squared error from the estimated mean: $y_{ia} = (b_{ia} - \mu_i(\tilde{w}_a))^2$. Consider the specification

$$E\left((b_i - \mu_i(\tilde{w}))^2 | \tilde{w}, Y_i = 1\right) = \exp\left(w_0 \gamma_{\sigma i} + h_{\sigma i}(east, north)\right). \tag{C.8}$$

I follow steps 1-5 but I use the exponential function instead of Φ and ϕ, and estimate the coefficients $\hat{\gamma}$ by non-linear least squares instead of MLE.

D Marginal costs and effect of competitors' signals

This appendix shows the details of the numerical computation of the marginal cost and the effects of competitors' signals. These terms are used to estimate the model as described in Sections 5 and 6. Define

$$\begin{aligned} \hat{s}_i &= \hat{G}_{B_i|\tilde{w}}(b_i), \\ s_j^* &= \hat{G}_{B_j|\tilde{w}}(b_i). \end{aligned} \tag{D.1}$$

Then

$$\begin{aligned} \hat{G}_{M_i|B_i,\tilde{w}}(b_i|b_i) &= P\left(S_{-i} > s_{-i}^* | \hat{s}_i\right) \\ \hat{g}_{M_i|B_i,\tilde{w}}(b_i|b_i) &= \sum_{j \neq i} \hat{g}_{B_j|\tilde{w}}(b_i) f\left(\hat{s}_i, s_j^*\right) P\left(S_{-ij} > s_{-ij}^* | \hat{s}_i, s_j^*\right), \end{aligned} \tag{D.2}$$

where $f\left(s_i, s_j^*\right)$ is the density of the bivariate copula of (S_i, S_j) evaluated at $\left(\hat{s}_i, s_j^*\right)$. The estimate of the marginal cost is:

$$m\hat{c}_i = b_i - \frac{1 - P\left(S_{-i} > s_{-i}^* | \hat{s}_i\right)}{\sum_{j \neq i} \hat{g}_{B_j|\tilde{w}}(b_i) f\left(\hat{s}_i, s_j^*\right) P\left(S_{-ij} > s_{-ij}^* | \hat{s}_i, s_j^*\right)} \tag{D.3}$$

Therefore, the problem of calculating $m\hat{c}_i$ can be decomposed into the problem of obtaining:

$$\begin{aligned} & P\left(S_{-i} > s_{-i}^* | \hat{s}_i\right) \\ & P\left(S_{-ij} > s_{-ij}^* | \hat{s}_i, s_j^*\right) \text{ for all } j \\ & \hat{g}_{B_j|\tilde{w}}(b_i) \\ & f\left(\hat{s}_i, s_j^*\right). \end{aligned} \tag{D.4}$$

The last two elements can be obtained from the estimates of the marginal distribution of bids and the joint distribution of signals.

Now consider the computation of the effect of competitors signals based on the estimates. Define:

$$p_j = \frac{\hat{g}_{B_j|\tilde{w}}(b_i) f\left(\hat{s}_i, s_j^*\right) P\left(S_{-ij} > s_{-ij}^* | \hat{s}_i, s_j^*\right)}{\sum_{k \neq i} \hat{g}_{B_k|\tilde{w}}(b_i) f\left(\hat{s}_i, s_k^*\right) P\left(S_{-ik} > s_{-ik}^* | \hat{s}_i, s_k^*\right)}. \tag{D.5}$$

Recall,

$$\begin{aligned} \tilde{\psi}_j(b_i, w) &= E\left[\psi(S_j) | M_i = b_i, \hat{s}_i\right] \\ &= \sum_{k \neq i} E\left[\psi(S_j) | S_{-ik} > s_{-ik}^*, s_k^*, \hat{s}_i\right] p_k \\ &= p_j \psi(s_j) + \sum_{k \neq i,j} E\left[\psi(S_j) | S_{-ik} > s_{-ik}^*, s_k^*, \hat{s}_i\right] p_k \end{aligned} \tag{D.6}$$

The calculation of $\tilde{\psi}_j(s_i, w)$ requires the same components in (D.4) and:

$$E\left[\psi(S_j) | S_{-ik} > s_{-ik}^*, s_k^*, \hat{s}_i\right] \text{ for all } j \neq i \text{ and } k \neq i, j. \tag{D.7}$$

To compute the marginal cost and the effect of competitors' signals it is necessary to integrate over the support of a truncated normal. Bretz and Genz (2009) discuss how to obtain these integrals numerically.

D.1 Numerical integration

Suppose that X is an $n-$dimensional random vector distributed as multivariate normal with mean 0 and variance matrix $\Sigma = \Lambda + LL'$, where Ψ is a diagonal matrix with elements λ_i and L is a $n \times l$ matrix with $l \leq n-1$ and elements denoted l_{ij}. Then, following Bretz and Genz (2009) p.16, $P(q < X < r)$ for $q, r \in \mathbb{R}^s$, can be written as:

$$\int_{\mathbb{R}^l} \phi_l(y, I_l) \prod_{i=1}^{n} \left[\Phi\left(\frac{r_i - \sum_j^l l_{ij} y_j}{\sqrt{\lambda_i}}\right) - \Phi\left(\frac{q_i - \sum_j^l l_{ij} y_j}{\sqrt{\lambda_i}}\right)\right] dy \tag{D.8}$$

where $\phi_l(y, I_l)$ is the pdf of a circular l-dimensional standard normal. Notice that the term in the square brackets can be calculated analytically for each vector $y \in \mathbb{R}^l$. It is therefore only necessary to integrate numerically over l dimensions. In practice, I randomly draw $y_t \in \mathbb{R}^l$ for $t = 1...T$ and for each draw I compute:

$$\eta_t = \prod_{i=1}^{n} \left[\Phi\left(\frac{r_i - \sum_j^l l_{ij} y_{jt}}{\sqrt{\lambda_i}}\right) - \Phi\left(\frac{q_i - \sum_j^l l_{ij} y_{jt}}{\sqrt{\lambda_i}}\right)\right]. \tag{D.9}$$

Because $T^{-1} \sum_t \eta_t \to P(q < x < r)$ as $T \to \infty$, the probability $P(q < x < r)$ can be obtained numerically for the desired precision.

Similarly, $E(x_i | q < x < r) P(q < x < r)$ can be obtaining computing and averaging:

$$\begin{aligned} \varphi_t &= \eta_t \sqrt{\lambda} \left[\frac{\sum l_{ij} y_{jt}}{\sqrt{\lambda_i}} + \frac{\phi\left(\frac{q_i - \sum l_{ij} y_{jt}}{\sqrt{\lambda_i}}\right) - \phi\left(\frac{r_i - \sum l_{ij} y_{jt}}{\sqrt{\lambda_i}}\right)}{\Phi\left(\frac{r_i - \sum l_{ij} y_{jt}}{\sqrt{\lambda_i}}\right) - \Phi\left(\frac{q_i - \sum l_{ij} y_{jt}}{\sqrt{\lambda_i}}\right)}\right], \qquad \text{(D.10)} \\ T^{-1} \sum_t \varphi_t &\to E(x_i | q < x < r) P(q < x < r). \end{aligned}$$

E Algorithm to solve for the equilibrium bid functions

The equilibrium inverse bid functions are calculated using a numerical algorithm that is similar to that in Gayle and Richard (2008). The main difference with their approach is that I allow for interdependent costs and correlated signals. Although the computation time grows fast with the number of bidders, auctions with less than 10 bidders are solved within 5 hours. The algorithm solves the system of differential equations implied by bidders' first-order conditions. Following Gayle and Richard (2008), I guess the initial conditions and, after the system is solved forward, I verify if the terminal conditions are consistent with equilibrium bidding behavior.

E.1 Technology and information

Each bidder full information cost is:

$$E\left(C_i|S\right) = c_i + \delta_{-ii}\sum_{j\neq i}^{n}\psi\left(S_j\right) + \delta_{ii}\psi\left(S_i\right), \tag{E.1}$$

where c_i is a constant that depends on the estimated effect of bidder i's distance and project characteristics, δ_{-ii} is the estimated effect of competitors' signals and δ_{ii} is the estimated own signal effect. The quantile regressions presented in Section 6 result in a different set of parameters for each quantile. For simplicity and because the parameters were not too different across different quantiles, I take the specification with $\tau = 0.5$ for this simulation. The (rescaled) signals $\psi\left(S\right)$ are assumed to be jointly multinomial normal with covariance matrix $\hat{\Sigma}$. The distribution of signals is truncated to avoid numerical problems: instead of using $\psi\left(S_i\right)$, I use $\psi\left(\left(S_i - 0.5\right)0.999 + 0.5\right)$.

E.2 System of differential equations

Define the expected cost,

$$\phi\left(s_i, s_{-i}\right) = E\left(C_i|S_i = s_i, S_{-i} \geq s_{-i}\right)P\left(S_{-i} \geq s_{-i}|s_i\right). \tag{E.2}$$

Bidder i's problem is

$$\max_b bP\left(s_{-i}\left(b\right)\right) - \phi\left(s_i, s_{-i}\left(b\right)\right). \tag{E.3}$$

The first-order condition is

$$P_i + b\nabla P_i s'_{-i}\left(b\right) - \nabla\phi_i s'_{-i}\left(b\right) = 0, \tag{E.4}$$

where ∇P_i and $\nabla\phi_i$ are the gradients of $P\left(S_{-i} \geq s_{-i}\right)$ and $\phi\left(s_i, s_{-i}\right)$ with respect to s_{-i}. Notice that P_i and b are scalars, ∇P_i and $\nabla\phi_i$ are $1 \times (n-1)$ vectors, and $s'_{-i}\left(b\right)$ is an $(n-1) \times 1$ vector.

If all bidders' first-order conditions are considered together:

$$P = Ms', \tag{E.5}$$

where P is an $n \times 1$ vector with typical element: P_i, M is an $n \times n$ matrix with zeros in the main diagonal at typical row: $b\nabla P_i - \nabla\phi_i$, and s' is an $n \times 1$ vector.

E.3 Algorithm

The algorithm makes an initial guess on the lowest bid that firms are willing to submit when they receive signal $s = 0$. So the following variables are initialized: $b^{(0)}$, and $s^{(0)} = [0, 0, ..., 0]'$. The state variables are the scalar $b^{(t)}$ and the vector $s^{(t)}$.

At step t, each bidder's perceived probability of winning $P_i^{(t)} = P\left(S_{-i} \geq s_{-i}^{(t)} | s_i^{(t)}\right)$, expected cost $\phi\left(s_i^{(t)}, s_{-i}^{(t)}\right)$ and the gradients ∇P_i and $\nabla \phi_i$ are calculated by numerical integration as discussed in Appendix D.1. A bidder is "active" if it finds it profitable to win at the current state— if $b^{(t)} P_i^{(t)} - \phi\left(s_i^{(t)}, s_{-i}^{(t)}\right) \geq 0$. Bidders that do not satisfy this condition are "inactive" and their signal state in the next step is: $s_j^{(t+1)} = s_j^{(t)}$. Form vector $P^{(t)}$ and matrix $M^{(t)}$ with all "active" bidders and compute $s' = \left(M^{(t)}\right)^{-1} P^{(t)}$. If all s' are positive, set $s_i^{(t+1)} = s_i^{(t)} + s_i' \Delta b$ and $b^{(t+1)} = b^{(t)} + \Delta b$, where Δb is a predetermined bid step. Some elements of the resulting s' may be negative and there is nothing pathological about it. It means that some firms would prefer to bid higher even if their profits are positive at the current bid. I discuss how to obtain a subset of "active and willing" bidders from the set of "active" bidders below.

The simulation is stopped when a bidder reaches $s = 1$, when all but 1 bidders have nonpositive expected profits, or when the system diverges. The system diverges when all bidders expected profits increase with t. This is the case if inverse bid functions are not steep enough to counteract the effects of an increase in the bid with a reduction in the probability of winning. If bidders reach their zero profit conditions at low signals, the resulting bidding strategies are consistent with a reserve price equal to $b^{(T)}$, where T is the terminal step. If the system is solved for higher initial $b^{(0)}$, the terminal bid $b^{(T)}$ and signals $s^{(T)}$ are larger. For some initial $b^{(0)}$ there is terminal vector of signals such that for some i and j, $s_i^{(T)}$ and $s_j^{(T)}$ are close enough to one. In this case, the resulting bidding strategies are consistent with a nonbinding reserve price. If $b^{(0)}$ is too large, the system diverges and the resulting strategies are not consistent with any equilibrium bidding behavior. To obtain the results in Section 7, I find the initial $b^{(0)}$ consistent with an equilibrium with no reserve price.

E.4 "Active and willing" bidders

In a general asymmetric model the equilibrium bidding strategies may be such that the support of bids is different across bidders. For example, in an independent cost model suppose that there are three bidders A, B and C such that A's and B's costs are distributed on $[\underline{c}, \overline{c}]$ while C's costs are distributed on $[\underline{c}^*, \overline{c}]$, for $\underline{c} < \underline{c}^*$. The equilibrium bids can be such that A and B bid $\underline{b} < \underline{c}^*$ when their costs are $\underline{c}$, but C never bids $\underline{b}$. The optimal bid of C when his costs are $\underline{c}^*$ are above $\underline{c}^*$; as a result, there is a range of bids for which bidder C has positive profits, but finds it unprofitable to submit such a bid—he is "active but unwilling". The first-order condition of "active and unwilling" bidders have to be positive, i.e. they would increase their expected profit by bidding higher. It follows that if k_i is a slack variable,

$$P_i - k_i + b \nabla P s'_{-i}(b) - \nabla \phi s'_{-i}(b) = 0, \tag{E.6}$$

where $k_i = 0$ if i is willing to submit bid b, and $k_i > 0$ otherwise. The system of equations for all bidders can be written as:

$$P - k = Ms'. \tag{E.7}$$

where $k_i = 0$ and $s'_i > 0$ for willing bidders, and $k_i > 0$ and $s'_i = 0$ for unwilling bidders. Therefore,

the problem of finding all "willing" bidders is the problem of choosing a set of indices J such that:

$$\begin{bmatrix} M_s & 0 \\ M_{sn} & I \end{bmatrix}^{-1} P = \begin{bmatrix} s_{j\in J} \\ k_{j\notin J} \end{bmatrix} \geq 0, \tag{E.8}$$

where M_s is a square matrix $\#J \times \#J$ that contains element m_{ij} only if $i, j \in J$, while M_{sn} is a $(n - \#J) \times \#J$ matrix that contains element m_{ij} only if $i \notin J$ but $j \in J$. This is a combinatorial problem that can be solved by a brute force approach if there are only a few bidders. Instead, I consider the following algorithm: denote $D_p = diag(P)$ and let $\tilde{k} = D_p k$ and $\tilde{M} = D_p^{-1}$. Equation (E.7) becomes:

$$1 - \tilde{k} = \tilde{M}s. \tag{E.9}$$

I find the Perron–Frobenius eigenvector of $\tilde{M}$ and try J equal to the indices of its largest elements. I try first with the first two, then the first three largest elements and so on. This algorithm guides the brute force approach and finds the right set of "willing" bidders faster.

Figure 1: Participation and outcomes in the 3,851 auctions in the sample

(a) Firm 1

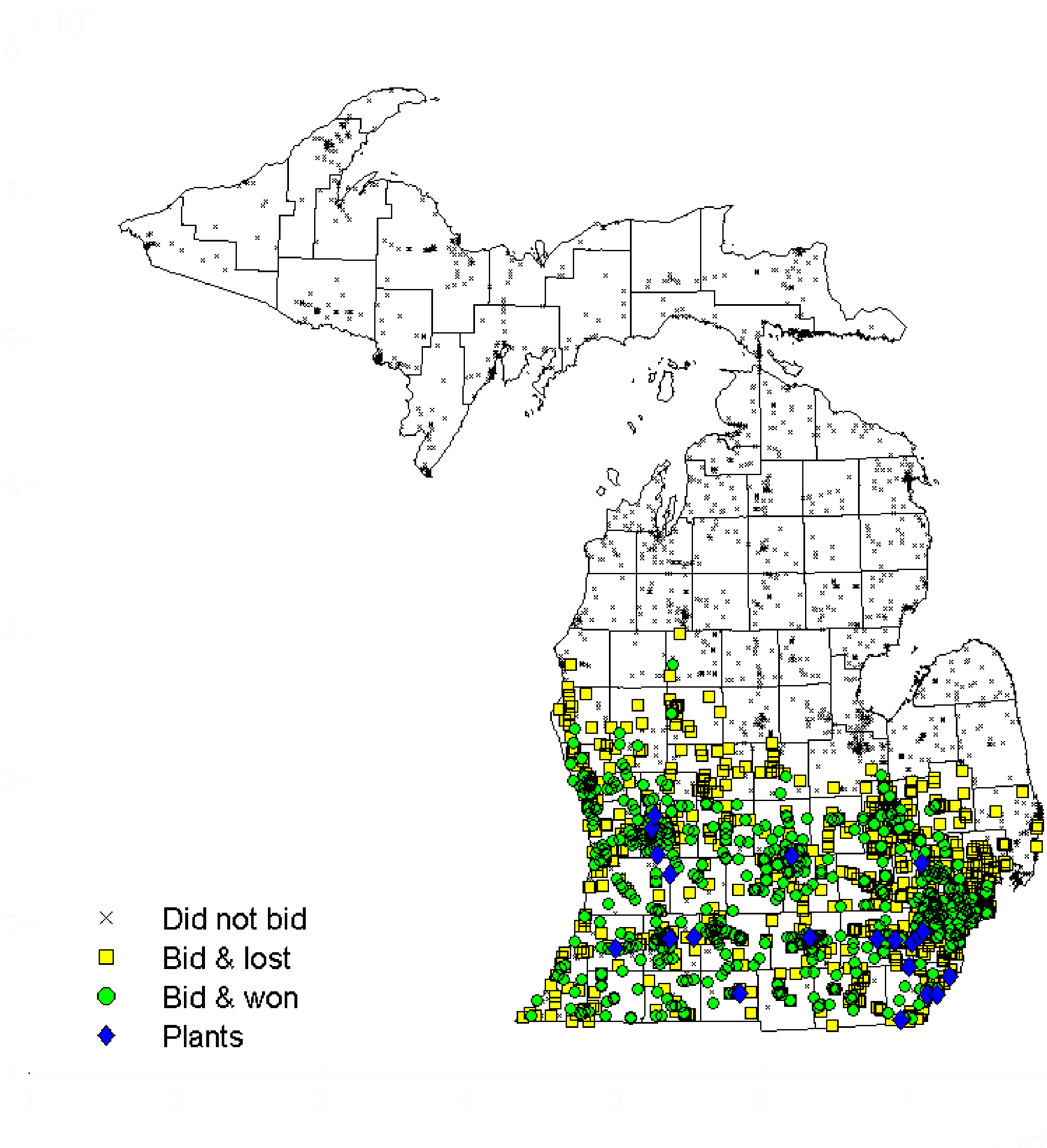


Figure 2: Participation and outcomes in the 3,851 auctions in the sample

(b) Firm 21

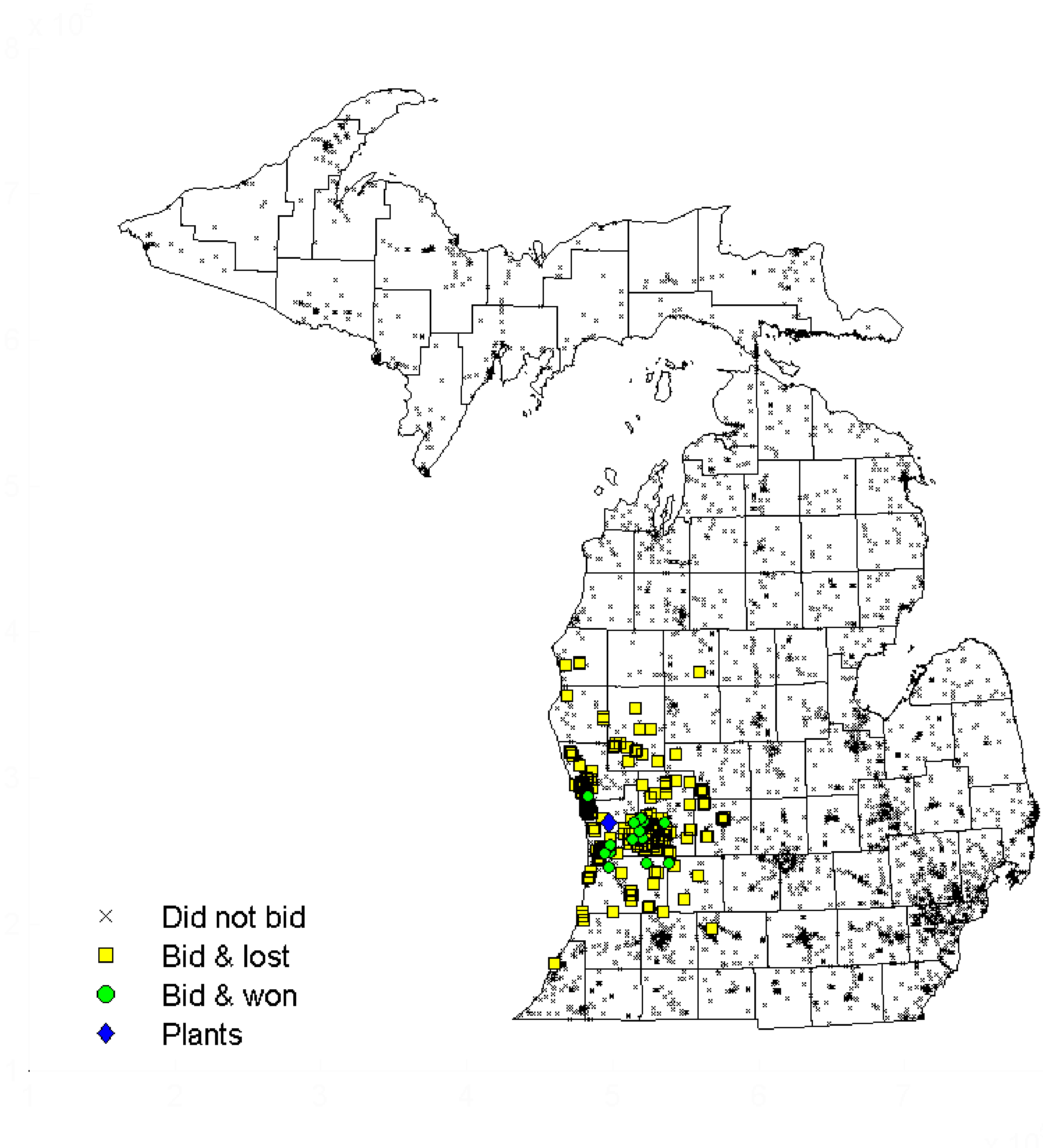
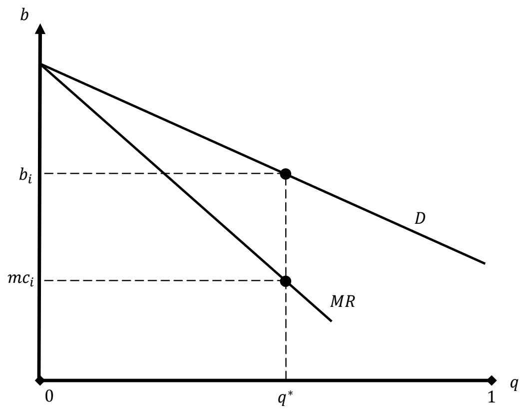
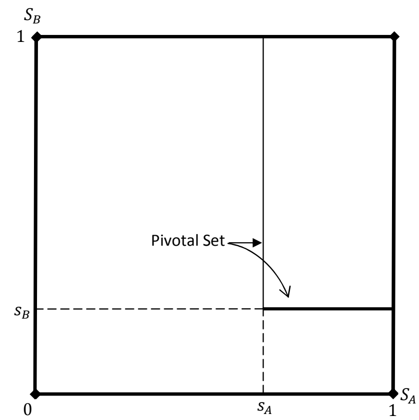


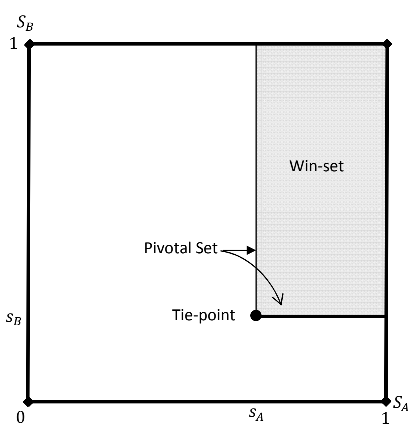
Figure 2: Identification of the full information cost



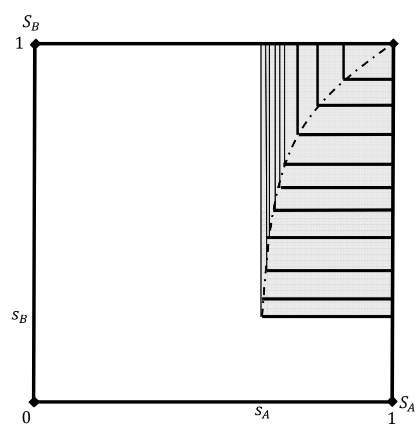
(a) Expected residual demand (D) and marginal revenue (MR) curves. b denotes bid and q denotes probability of winning. D is identified from the data, and MR can be derived from D. For an observed bid b_i find q^* in the D curve, and find the marginal revenue at q^*. mc_i is the marginal cost that rationalizes b_i.



(b) Space of competitors' signals. The pivotal set is composed by the signals that imply that there is at least one bidder that ties with bidder i (i.e., that submits b_i). Suppose that A and B bid b_i when they receive signals s_A and s_B, respectively. The pivotal set, shown by the solid lines, is given by (S_A, S_B) such that $S_A = s_A$ and $S_B \geq s_B$, or $S_A \geq s_A$ and $S_B = s_b$.



(c) Space of competitors' signals. Suppose that A and B bid b_i when they receive signals s_A and s_B, respectively. The tie-point is (s_A, s_B). The pivotal set, shown by the solid lines, is (S_A, S_B) such that $S_A = s_A$ and $S_B \geq s_B$, or $S_A \geq s_A$ and $S_B = s_b$. The win-set is (S_A, S_B) such that $S_A > s_A$ and $S_B > s_B$.



(d) Space of competitors' signals. The expected cost conditional on a pivotal set is identified by the marginal cost. The expected cost conditional on the shaded region is identified integrating the pivotal sets over a curve.

Figure 3: Estimation Procedure.

<u>Estimation: First Stage</u>

- Probability of Participation: $\hat{\pi}_i(\widetilde{w})$ *(semipar. probit).*

- Mean bid conditional on participation: $\hat{\mu}_i(\widetilde{w})$
  (semipar. partial linear model).

- Variance of bid conditional on participation: $\hat{\sigma}^2{}_i(\widetilde{w})$
  (semipar. nonlinear model).

- Distribution of the standardized bid: $\widehat{H}_i(\cdot)$
  (smoothed empirical cdf).

\- Marginal distribution of bids: $\hat{G}_{B_{i|\widetilde{w}}}(b_i) = \widehat{H}_i\left(\frac{b_i - \hat{\mu}_i(\widetilde{w})}{\hat{\sigma}_i(\widetilde{w})}\right)\hat{\pi}_i(\widetilde{w})$.

- Joint distribution of signals: $\widehat{F}_S(\cdot)$ *(exploratory Tobit Factor Analysis).*

\- Joint distribution of bids:

$$\hat{G}_{B_{|\widetilde{w}}}(b) = \widehat{F}_S\left(\hat{G}_{B_{1|\widetilde{w}}}(b_1), \ldots, \hat{G}_{B_{n|\widetilde{w}}}(b_n)\right).$$

\- Effect of competitors' signals

$\hat{\psi}(b_i, w)$ and marginal cost $\widehat{mc}_i = b_i - \frac{1 - \hat{G}_{M_{i|B_i,\widetilde{w}}}(b_i|b_i)}{\hat{g}_{M_{i|B_i,\widetilde{w}}}(b_i|b_i)}$ (See Appendix D).

<u>Estimation: Second Stage</u>

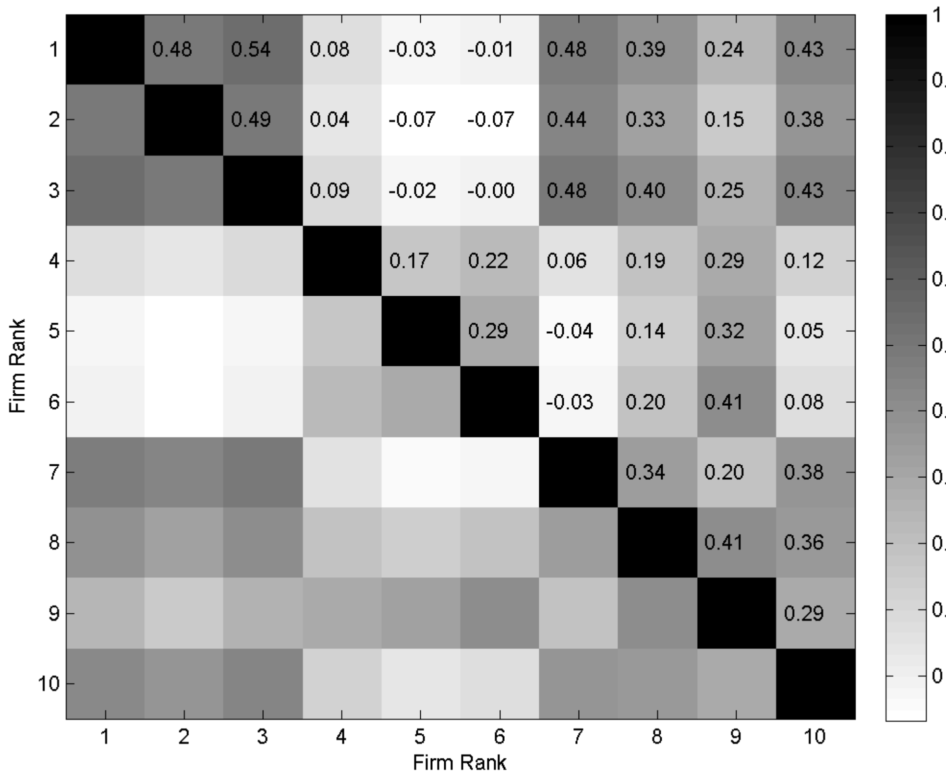
- Test the null hypothesis of private costs: *(Buchinsky and Hahn, 1998, quantile regression).*
- Estimate the coefficients of the full information costs *(Buchinsky and Hahn, 1998, quantile regression; Chernozhukov and Hansen, 2007, IVQR).*

Note:

- Denotes an estimation step

\- Denotes a calculation performed with the estimates

Figure 4: Correlation Matrix of Signals. Top 10 firms.



Each cell represents a correlation coefficient between the firms in the rows and columns. For example, the correlation of signals between firms 2 and 3 is 0.49. Darker cells represent higher correlation. Firms 1, 2, 3, 7, 8 and 10 exhibit high correlation of signals between them. Similarly, firms 4, 5 and 6 also show high correlation. The correlation between firms of different groups is lower.

Table 1(a): Descriptive Statistics. Engineer's estimate, bids and distances.

Variable	N	Mean	Sd	P5	Median	P95
Engineer's estimate ($000)	3,851	1,398	3,117	125	656	4,477
Lowest bid ($000)	3,851	1,320	2,983	118	603	4,236
Participants	3,851	5.08	3.45	2	4	12
(2nd Lowest/Lowest bid-1)×100	3,770	6.9	7.7	.4	4.7	20.9
(Lowest/engineer-1)×100	3,851	-6.4	12.6	-25.6	-7	14.5
Distance of Winner (km)	3,662	40	48	2	27	122
Distance of Bidder (km)	18,778	51	50	4	38	138

Note: Pct stands for percentile. 2nd Lowest: the second lowest bid. engineer: engineer's estimate. In 81 auctions there was only one bid. There were 189 auctions won by a firm for which I did not find any verifiable location.

Table 1(b): Descriptive Statistics. Size of the project.

	No.	%
Very Small (less than 150k)	278	7.22
Small (150k-500k)	1,243	32.28
Normal (500k-1.5m)	1,440	37.39
Large (1.5m-3m)	529	13.74
Very Large (more than 3m)	361	9.37

Note: Distribution of project sizes according to the engineer's estimate.

Table 1(c): Descriptive Statistics. Prequalification Requirements

	No.	%
Hot Mix Asphalt/Bituminous Paving		
Not included	502	13.04
Included	1,432	37.19
Required	1,917	49.78
Grading & Drainage Structures & Agg. Cons		
Not included	2,096	54.43
Included	1,348	35.00
Required	407	10.57
Bridges And Special Structures		
Not included	3,741	97.14
Included	77	2.00
Required	33	0.86
Miscellaneous Concrete Items		
Not included	3,398	88.24
Included	441	11.45
Required	12	0.31
Sewers and Watermains		
Not included	3,371	87.54
Included	459	11.92
Required	21	0.55

A work-type prequalification is required if only contractors prequalified for that work-type can participate in the auction. A work-type is included if the work-type classification can be substituted by other work-type prequalification. For example, if the project requires prequalifications A or (B and C), then A, B and C are included but none is required. However, if the requirement is (A or B) and C, then A and B are included and C is required. Some of the projects do not include or require HMA work, but they require a subcontractor with such prequalification.

Table 2: Number of Participants and Normalized Bids

Participants	Auctions	Winning Bid		Average Bid	
		mean	s.e.m	mean	s.e.m
1	81	1.4	0.9	1.4	0.9
2	747	-4.1	0.5	1.2	0.6
3	913	-5.5	0.4	2.8	0.4
4	496	-4.4	0.6	6.4	0.8
5	356	-6.7	0.6	3.7	0.7
6	271	-7.8	0.7	4.0	0.9
7	217	-8.8	0.8	3.3	0.8
8	189	-8.4	0.8	4.2	0.9
9	132	-8.9	0.9	2.9	1.0
10	118	-10.6	1.0	1.3	1.1

Note: Bids were normalized by the engineer's estimate: (Bid/engineer's estimate-1)$\times$100. The average winning bid and average submitted bid are tabulated by the number of participants in the auction. For example, there were 742 auctions with only two participants. In these auctions the normalized winning bid was on average 4.1% below the engineer's estimate, while the average submitted bid was 1.1% above. While the average winning bid decreases with the number of participants, the average submitted bid does not. s.e.m stands for standard error of the mean.

Table 3(a): OLS of normalized bids on own and competitors' distance.

Variable [a]	All Firms		Firm 1		Firm 2		Firm 3		Firm 4		Firm 5	
	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.
Constant [b]	-9.8***	0.5	-13.0***	1.5	-14.1***	1.7	-11.3**	5.6	-15.0***	5.8	-16.8*	9.2
Own Dist (10- 25km)	1.1***	0.4	-0.1	1.1	1.0	1.4	1.0	1.3	6.6	5.0	4.5	7.3
Own Dist (25- 50km)	2.0***	0.4	2.0**	1.0	1.8	1.3	3.4*	2.0	9.5**	4.8	7.3	7.0
Own Dist (50-100km)	3.7***	0.4	9.7***	1.5	5.0***	1.4	7.2**	3.3	10.9**	5.1	9.7	7.1
Own Dist (100-150km)	4.9***	0.5	15.8*	8.9	11.9***	2.2	26.6***	6.1	9.9*	5.2	12.1*	7.3
Own Dist (150-200km)	5.4***	0.9			24.9***	7.2	5.1	6.9	12.8**	5.4	17.1**	8.0
Own Dist (200-500km)	7.1***	1.4			22.1**	10.3	2.1	10.4	14.9**	6.9	23.1***	8.1
Competitor A (0 -10 km)	-0.2	0.4	3.1**	1.6	5.9***	1.4	0.6	2.5			-5.9***	1.9
Competitor A (10-25 km)	-0.8**	0.4	-0.2	1.4	0.2	1.5	3.3	2.5	-0.0	4.0	-5.9***	2.3
Competitor A (25-50 km)	0.1	0.3	1.2	1.1	3.6***	1.2	2.1	2.4	-0.2	2.3	3.4**	1.6
Competitor B (0 -10 km)	2.2***	0.4	6.6***	1.3	-5.9**	2.5	1.7	4.1	6.0	4.0	-3.6	3.1
Competitor B (10-25 km)	2.0***	0.4	5.1***	1.2	-3.5	2.2	3.2	4.1	3.1	3.1	1.8	2.9
Competitor B (25-50 km)	0.7**	0.3	4.3***	1.5	-0.9	1.7	4.6	4.6	3.9	2.5	1.1	1.8
Competitors' id [c]			2 and 3		1 and 8		1 and 7		5 and 6		6 and 4	
Number of Observations	14,071		1,497		1,248		603		583		577	

Note: P-values: * $p<0.1$, ** $p<0.05$ and *** $p<0.01$. Heteroskedasticity-robust s.e. are reported.
[a] Additional included controls: project size, year, and prequalification requirements. Dependent variable: (Bid/engineer's estimate-1)$\times$100. The regression in the first column includes firm fixed effects.
[b] Omitted category: projects located within 0-10 km of the bidder and more than 50 km away from the two main competitors that require prequalification in HMA-work.
[c] Two main competitors of firm i: the two firms that are observed to participate most often in the auctions where i participates.

Table 3(b): OLS of normalized bids on own and competitors' distance. Includes a set of demand controls.

Variable [a]	All Firms		Firm 1		Firm 2		Firm 3		Firm 4		Firm 5	
	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.
Constant [b]	-9.7***	0.7	-14.4***	3.2	-13.7***	2.4	14.7**	6.6	-31.3***	7.5	-34.2***	9.6
Own Dist (10- 25km)	1.1***	0.4	0.2	1.1	1.8	1.5	0.8	1.3	5.3	5.1	4.3	7.1
Own Dist (25- 50km)	2.2***	0.4	2.1*	1.1	2.9*	1.5	4.5**	2.3	9.2*	4.9	7.6	7.0
Own Dist (50-100km)	3.8***	0.4	10.2***	1.6	5.0***	1.6	7.2**	3.5	11.2**	5.2	10.0	7.1
Own Dist (100-150km)	4.9***	0.5	15.8*	9.5	11.0***	2.6	33.7***	4.4	9.6*	5.3	12.2*	7.3
Own Dist (150-200km)	5.3***	0.9			24.3***	7.4	17.8**	7.6	15.6***	5.8	17.6**	8.1
Own Dist (200-500km)	7.0***	1.4			19.8*	10.6	13.7	10.1	18.0**	7.8	23.1***	7.9
Competitor A (0 -10 km)	-0.2	0.4	2.2	1.7	3.6*	2.2	2.9	2.6			-6.3**	2.9
Competitor A (10-25 km)	-0.8**	0.4	-0.5	1.5	-1.1	2.1	5.2*	2.7	-0.1	4.2	-6.2**	2.7
Competitor A (25-50 km)	0.2	0.3	1.5	1.1	2.4*	1.4	3.9	2.6	-0.1	2.3	3.5**	1.6
Competitor B (0 -10 km)	1.3***	0.5	3.7**	1.7	-4.4	2.8	13.3***	4.5	4.8	4.1	-3.1	3.2
Competitor B (10-25 km)	1.2***	0.4	3.0**	1.5	-3.0	2.3	14.8***	4.4	1.8	3.3	1.5	3.6
Competitor B (25-50 km)	0.6*	0.3	3.8**	1.5	-0.7	1.7	16.9***	4.8	4.3*	2.5	1.0	1.9
Competitors' id [c]			2 and 3		1 and 8		1 and 7		5 and 6		6 and 4	
Number of Observations	14,071		1,497		1,248		603		583		577	

Note: P-values: * $p<0.1$, ** $p<0.05$ and *** $p<0.01$. Heteroskedasticity-robust s.e. are reported.
[a] This specification includes a set of 9 dummies for levels of road density around the project. Additional included controls: project size, year, and prequalification requirements. Dependent variable: (Bid/engineer's estimate-1)$\times$100. The regression in the first column includes firm fixed effects.
[b] Omitted category: projects located within 0-10 km of the bidder and more than 50 km away from the two main competitors that require prequalification in HMA-work.
[c] Two main competitors of firm i: the two firms that are observed to participate most often in the auctions where i participates.

Table 4: Probability of Participation. Semiparametric Estimation.

Variable	Firm 1		Firm 2		Firm 3		Firm 4		Firm 5	
	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.
Year: 2001	-0.237	0.187	-1.780***	0.194	-0.817***	0.285	-0.335*	0.201	-0.952***	0.225
2002	-0.16	0.172	-1.651***	0.189	-0.533**	0.272	-0.093	0.175	-0.574***	0.204
2003	-0.631***	0.172	-1.232***	0.181	0.222	0.249	-0.334*	0.18	-0.647***	0.203
2004	0.117	0.17	-1.000***	0.175	-0.485**	0.24	0.224	0.165	-0.336*	0.202
2005	-0.430**	0.172	-1.333***	0.174	-0.682***	0.253	0.268	0.166	-0.560***	0.196
2006	0.039	0.164	-0.716***	0.162	-0.412*	0.23	0.309*	0.158	-0.500**	0.195
2007	-0.226	0.154	-0.802***	0.16	-0.165	0.212	0.615***	0.149	-0.452**	0.185
2008	-0.215	0.168	-0.901***	0.177	-0.017	0.244	0.045	0.173	-0.292	0.197
2009	-0.06	0.155	-0.371**	0.161	0.174	0.221	0.501***	0.154	0.089	0.184
Size: v. small	-0.1	0.166	-0.339**	0.161	-0.914***	0.254	0.322**	0.155	-0.238	0.199
small	-0.085	0.099	-0.082	0.093	-0.467***	0.143	0.119	0.092	-0.162	0.115
large	0.097	0.121	0.092	0.132	0.202	0.175	0.105	0.115	-0.226*	0.136
v. large	0.302**	0.144	0.456***	0.153	0.917***	0.19	-0.202	0.144	-0.608***	0.188
HMA req	3.321***	0.221	2.551***	0.252	1.545***	0.268	-0.720***	0.2	-1.333***	0.279
HMA inc	1.852***	0.241	0.922***	0.293	1.270***	0.329	0.094	0.206	0.394*	0.235
Constr req	-0.257	0.21	0.189	0.245	-1.283***	0.269	1.124***	0.201	2.591***	0.262
Constr inc	-0.518***	0.18	-0.072	0.249	-0.699***	0.26	1.046***	0.196	1.754***	0.252
Bridge req	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-Inf	-0.082	0.656
Bridge inc	-0.964***	0.228	-1.000***	0.319	-1.328***	0.335	-0.964***	0.275	-0.276	0.249
Concrete req	-0.777	0.641	0.622	0.726	-1.672**	0.808	1.250**	0.581	-0.211	0.89
Concrete inc	-0.197*	0.115	-0.129	0.141	-0.457**	0.217	-0.069	0.108	-0.12	0.115
Sewers req	0.018	0.48	-0.585	0.723	-1.498**	0.697	-0.432	0.494	1.525***	0.509
Sewers inc	-1.394***	0.127	-0.833***	0.154	-1.937***	0.223	-0.069	0.108	0.014	0.118

P-values: * p<0.1, ** p<0.05 and *** p<0.01. Num Obs=3851. Omitted category: projects located within 0-10 km of the bidder and more than 50 km away from the two main competitors that require prequalification in HMA-work. Firms 1 to 4 never participated in a project that requires the prime contractor to be prequalified in Bridges and Special Structures. Firms 1 to 3 participate more often when the project includes or requires HMA prequalification. Firms 4 and 5 participate more often when the project requires construction work.

Table 5(a): Firm 1. Quantile regression of marginal costs on own and competitors' distance.

Variable	$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.75$	
	coef	s.e.	coef	s.e.	coef	s.e.
const	0.537***	0.031	0.688***	0.030	0.840***	0.036
ownd (10- 25km)	-0.007	0.025	0.012	0.020	0.003	0.021
ownd (25- 50km)	0.053**	0.022	0.046**	0.018	0.042**	0.021
ownd (50-100km)	0.248***	0.030	0.208***	0.028	0.184***	0.032
ownd (100-150km)	0.826***	0.198	0.698***	0.215	0.590***	0.192
Firm 2 (0 -10 km)	0.134***	0.032	0.073**	0.030	0.046	0.034
Firm 2 (10-25 km)	0.070**	0.030	0.045*	0.027	0.009	0.030
Firm 2 (25-50 km)	0.055**	0.026	0.060**	0.023	0.005	0.025
Firm 3 (0 -10 km)	0.191***	0.029	0.170***	0.027	0.123***	0.030
Firm 3 (10-25 km)	0.202***	0.025	0.155***	0.023	0.099***	0.027
Firm 3 (25-50 km)	0.152***	0.027	0.138***	0.028	0.094***	0.032

Note: Additional included controls: project size, year, and prequalification requirements. P-values: * $p<0.1$, ** $p<0.05$ and *** $p<0.01$. Num Obs=1,497. Under the private costs hypothesis the distribution of marginal costs should not depend on competitor's distance. Firm 1's marginal costs are higher in projects close to the locations of firms 2 and 3. These results are consistent with an interdependent cost model.

Table 5(b): Firm 2. Quantile regression of marginal costs on own and competitors' distance.

Variable	$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.75$	
	coef	s.e.	coef	s.e.	coef	s.e.
const	0.457***	0.035	0.635***	0.033	0.775***	0.034
ownd (10- 25km)	0.014	0.026	0.015	0.025	0.018	0.027
ownd (25- 50km)	0.083***	0.024	0.061**	0.025	0.042	0.026
ownd (50-100km)	0.208***	0.027	0.162***	0.027	0.122***	0.030
ownd (100-150km)	0.316***	0.036	0.254***	0.041	0.200***	0.052
ownd (150-200km)	0.557***	0.090	0.681***	0.172	0.545***	0.155
ownd (200-500km)	0.882***	0.174	0.689***	0.181	0.540***	0.162
Firm 1 (0 -10 km)	0.250***	0.025	0.194***	0.025	0.151***	0.029
Firm 1 (10-25 km)	0.153***	0.027	0.124***	0.025	0.108***	0.028
Firm 1 (25-50 km)	0.199***	0.024	0.139***	0.020	0.119***	0.024
Firm 8 (0 -10 km)	-0.107***	0.039	-0.091**	0.042	-0.089**	0.042
Firm 8 (10-25 km)	-0.026	0.054	-0.005	0.047	0.012	0.049
Firm 8 (25-50 km)	-0.040	0.037	-0.020	0.030	0.001	0.032

Note: Additional included controls: project size, year, and prequalification requirements. P-values: * $p<0.1$, ** $p<0.05$ and *** $p<0.01$. Num Obs=1,248. Firm 2's marginal costs are higher in projects close to firm 1, but not close to firm 3. These results are consistent with a model where firms 1 and 2 have a common cost component.

Table 5(c): Firm 3. Quantile regression of marginal costs on own and competitors' distance.

Variable	$\tau = 0.25$		$\tau = 0.50$		$\tau = 0.75$	
	coef	s.e.	coef	s.e.	coef	s.e.
const	0.684***	0.063	0.679***	0.077	0.745***	0.245
ownd (10- 25km)	0.008	0.019	0.023	0.020	0.027	0.026
ownd (25- 50km)	0.123***	0.023	0.108***	0.026	0.123**	0.048
ownd (50-100km)	0.222***	0.037	0.298***	0.110	0.571	0.543
ownd (100-150km)	0.537***	0.097	0.477***	0.104	0.416*	0.227
ownd (150-200km)	0.300***	0.082	0.385	0.402	0.420	6.034
ownd (200-500km)	0.262***	0.087	0.269**	0.116	0.164	0.262
Firm 1 (0 -10 km)	-0.006	0.033	0.027	0.040	0.028	0.070
Firm 1 (10-25 km)	0.023	0.033	0.055	0.038	0.031	0.066
Firm 1 (25-50 km)	0.026	0.030	0.039	0.037	0.013	0.065
Firm 7 (0 -10 km)	0.024	0.048	0.111**	0.055	0.163	0.206
Firm 7 (10-25 km)	0.068	0.045	0.119**	0.054	0.146	0.205
Firm 7 (25-50 km)	0.029	0.053	0.148**	0.059	0.163	0.207

Note: Additional included controls: project size, year, and prequalification requirements. P-values: * p<0.1, ** p<0.05 and *** p<0.01. Num Obs=603. Firm 3's marginal costs are higher in projects close to firm 7.

Table 6(a): Firm 1. Full Information Cost.

Variable	$\tau = 0.31$		$\tau = 0.50$		$\tau = 0.69$	
	coef	s.e.	coef	s.e.	coef	s.e.
const	0.665***	0.031	0.712***	0.026	0.765***	0.023
δji	0.030***	0.002	0.025***	0.002	0.027***	0.004
ownd (10- 25km)	-0.004	0.024	0.014	0.018	0.021	0.013
ownd (25- 50km)	-0.014	0.024	0.012	0.017	0.032**	0.013
ownd (50-100km)	0.138***	0.032	0.158***	0.025	0.172***	0.024
ownd (100-150km)	0.407**	0.160	0.377	0.252	0.288***	0.103

Note: Additional included controls: project size, year, and prequalification requirements. P-values: * p<0.1, ** p<0.05 and *** p<0.01. Num Obs=1,497. Firm 1's expected cost of completing a project located within 10km of its plants is 66% of the engineer estimate for a signal in the 31st percentile, 71% for the 50th and 76% for the 69th. Firm 1's costs are ∼15 percentage points higher in projects located 50-100km away and ∼35 points higher 100-150km away. If signals were public, a one standard deviation in the (rescaled) signal received by a competitor would increase firm 1's expected costs by ∼2.7 percentage points. Under the assumption of normally distributed (rescaled) signals, an increase from the 31st to the 69th percentile is a one standard deviation increase. Compare the ∼2.7-point competitor-signal effect with the 10-point own-signal effect.

Table 6(b): Firm 2. Full Information Cost.

Variable	$\tau = 0.31$		$\tau = 0.50$		$\tau = 0.69$	
	coef	s.e.	coef	s.e.	coef	s.e.
const	0.638***	0.031	0.708***	0.016	0.769***	0.023
δji	0.029***	0.003	0.020***	0.003	0.030***	0.006
ownd (10- 25km)	-0.023	0.028	-0.003	0.013	-0.019	0.015
ownd (25- 50km)	0.019	0.023	0.027**	0.012	0.01	0.015
ownd (50-100km)	0.063***	0.021	0.078***	0.014	0.070***	0.023
ownd (100-150km)	0.113***	0.039	0.131***	0.026	0.110**	0.05
ownd (150-200km)	0.427***	0.133	0.651***	0.082	0.509***	0.119
ownd (200-500km)	0.796	2.606	0.721	207.114	0.611	1.209

Note: Additional included controls: project size, year, and prequalification requirements. P-values: * p<0.1, ** p<0.05 and *** p<0.01. Num Obs=1,248. Firm 2's expected cost of completing a project located within 10km of its plants is 64% of the engineer estimate for a signal in the 31st percentile, 71% for the 50th and 77% for the 69th. Firm 2's costs are ∼11 percentage points higher in projects located 50-100km away and ∼50 points higher 100-150km away. If signals were public, a one standard deviation in the (rescaled) signal received by a competitor would increase firm 1's expected costs by ∼2.9 percentage points. Compare the ∼2.9-point competitor-signal effect with the 13-point own-signal effect.

Table 7: Summary: Full information cost estimation results for the 10 most active firms.

Firm	Effect of signals	
	Own	Competitors
1	0.100	0.027
2	0.131	0.029
3	0.075	0.017
4	0.115	0.014
5	0.143	0.008
6	0.025	0.007
7	0.115	0.015
8	0.138	0.014
9	0.183	0.011
10	0.087	0.014

This table summarizes the estimates of the full information cost for the 10 most active firms. The first column shows the estimated own-signal effect—the effect of an increase in bidder i's signal from the 31st to the 69th percentile on its own full information cost. As discussed in tables 6(a) and 6(b), it is the difference between the estimates of the constant term at these two quantiles. The second column shows the estimated competitor-signal effect—the effect of an increase in bidder j's signal from the 31st to the 69th percentile on bidder i's full information cost. In table 6(a) the competitor-signal effect was estimated to be 0.030, 0.025 and 0.027 for percentiles 31st, 50th and 69th. I report the median estimate from these three: 0.027. The p-values on these coefficients are below 0.01 for firms 1, 2, 3, 4 and 10; below 0.05 for firms 5, 6 and 7; and below 0.1 for firm 8. According to these coefficients, if firm 2 receives a signal that increases its expected costs by $1,310 ($1,000 x the coefficient of firm 2's own signal's effect), firm 1's expected costs increase by $270 ($1,000 x the coefficient of firm 1's competitors' signal's effect) upon learning it.

Table 8: Market conditions. Potential Bidders

Firm	Probability [a]	Dist.(Km) [b]	Median Cost ($) [c]
1	0.876	9	356,971
2	0.956	24	355,296
3[d]	0.001	149	554,081
5	0.092	39	569,657
6	0.172	16	579,825

I analyze the effect of competition in a particular auction. Given its characteristics and location only 4 firms had an estimated probability of entry greater than 0.03.

[a] Predicted probability of participation according to the semiparametric estimates.

[b] Distance of the closest firm's location.

[c] Full information cost evaluated at $\psi(S_j) = 0$ for all j.

[d] In the counterfactual simulation firm 3 receives a $182,000 subsidy if it wins the auction. With the subsidy the firm is able to compete with firms 1 and 2 on a more even ground.

Table 9: Counterfactual. The effects of competition.

Effect on expected...	IPC[a]	APC[b]	CC[c]
Firm 1's bid ($)	-11,438	-9,201	-6,942
Firm 2's bid ($)	-12,413	-9,167	-5,840
Firm 3's bid ($)	-64,726	-58,112	-76,448
Winning bid ($)	-20,921	-15,164	-15,161
Subsidy paid ($)	41,877	38,293	41,615
Subsidy net cost ($)	20,957	23,129	26,453

[a] Independent Private Cost model: No correlation of signals or cost interdependence.

[b] Affiliated Private Cost model: Estimated correlation of signals. No cost interdependence.

[c] (Partially) Common Cost model: Estimated correlation of signals and cost interdependence.

The subsidy reduces the expected equilibrium bid of firm 1 by $11,438 under the IPC model, but only by $9,201 under the APC and $6,942 under the CC model. The difference between the reduction in the CC model and the APC model can be attributed to the effect of the common cost components, while the difference between the APC and IPC can be attributed to the nonindependence of signals. Firm 3 bids more aggressively in the CC model not only because of the subsidy but also because it is subject to a lesser winner's curse. The effect of the common components on the winning bid is almost zero due to the pro-competitive effect on firm 3 that offsets the anti-competitive effects on firms 1 and 2. The expected subsidies paid are higher under the IPC and CC models, where firm 3 wins with higher probability. Due to the winner's curse, firms 1 and 2 bid more cautiously and firm 3 bids more aggressively. As a result, firm 3 wins more often in the CC model. The total cost of the subsidy is greatest under the CC model, followed by the APC model, and lowest under the IPC.

Table A: Quantile Regression of the latent normalized bids on own and competitors' distance. Includes a set of demand controls.

Variable [a]	Firm 1		Firm 2		Firm 3		Firm 4		Firm 5	
	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.	coef	s.e.
Constant [b]	-10.8	8.8	-19.3***	3.4	3.9	13.1	-7.3	29.7	-18.4	39.7
Own Dist (10- 25km)	0.8	1.6	1.6	2.1	1.7	1.9	10.1	10.6	6.1	12.1
Own Dist (25- 50km)	2.8*	1.5	3.1	2.1	12.8***	3.9	17.9	12.1	8.3	11.6
Own Dist (50-100km)	13.5***	2.6	9.9***	2.4	29.8***	9.0	25.1*	13.1	11.3	11.6
Own Dist (100-150km)	47.2***	15.8	22.1***	4.2	49.6***	9.4	24.7*	12.8	12.7	11.8
Own Dist (150-200km)			51.9***	14.3	36.1**	15.0	35.7**	15.6	39.9	26.9
Own Dist (200-500km)			55.7***	15.1	51.4**	20.4	43.7**	17.6	29.4*	16.2
Competitor A (0 -10 km)	1.5	2.4	10.8***	3.5	1.7	4.5			-17.3***	5.2
Competitor A (10-25 km)	-1.7	1.9	5.1	3.2	5.1	4.5	1.7	12.4	-14.0***	4.4
Competitor A (25-50 km)	1.0	1.6	9.0***	2.0	2.5	4.6	6.4	6.9	-2.7	2.4
Competitor B (0 -10 km)	2.9	2.5	-10.0**	4.2	20.2**	7.9	35.7	39.7	-7.5	6.1
Competitor B (10-25 km)	1.0	2.2	-7.8*	4.3	21.1***	7.7	11.1	17.2	-3.9	7.5
Competitor B (25-50 km)	1.3	2.2	-2.2	2.6	27.1***	7.7	10.6*	5.6	-0.3	3.5
Competitors' id [c]	2 and 3		1 and 8		1 and 7		5 and 6		6 and 4	
Number of Observations	1,497		1,248		603		583		577	

Note: P-values: * $p<0.1$, ** $p<0.05$ and *** $p<0.01$. Heteroskedasticity-robust s.e. are reported.

[a] This specification includes a set of 9 dummies for levels of road density around the project. Additional included controls: project size, year, and prequalification requirements: HMA, construction, concrete, bridge, water mains. Dependent variable: (Bid/engineer's estimate-1)×100.

[b] Omitted category: projects located within 0-10 km of the bidder and more than 50 km away from the two main competitors that require prequalification in HMA-work.

[c] Two main competitors of firm i: the two firms that are observed to participate most often in the auctions where i participates.