Endogenous Participation and Local Market Power in Highway Procurement*

Liran Einav and Ignacio Esponda†

March 2008

Please don't circulate or cite without permission.
Preliminary and **very** incomplete draft.
Comments are extremely welcome.

Abstract. We use highway procurement data from Michigan to study the effect of firms' distances to the auction site on participation and bidding decisions. Motivated by "reduced form" evidence, we account for endogenous participation and allow the participation decision to be correlated with subsequent bids. That is, while distant firms face higher costs to complete the project, their participation may indicate otherwise. We develop a structural model of correlated participation and bidding decisions and sketch its non-parametric identification. We then estimate the model to quantify the importance of distance, the extent of local market power in the industry, and the effect of subsidizing participation in remote regions.

JEL classification numbers:
Keywords:

*This is still work progress! We thank Kyle Woodward for excellent research assistance.

†Einav: Department of Economics, Stanford University, and NBER, leinav@stanford.edu; Esponda: Stern School of Business, NYU, iesponda@stern.nyu.edu.

1 Introduction

Location choice and spatial product differentiation is one of the key strategic decisions in many industries, ranging from retail services, such as gas stations, fast food outlets, grocery stores, or hospitals, to manufacturing of goods that are difficult or expensive to transport, such as cement or sugar. In this paper we use highway procurement data from Michigan as an attractive setting that allows us to quantify the importance of local market power, and evaluate alternative mechanisms to reduce it.

A typical empirical paper in industrial organization that analyzes or controls for spatial differentiation does so by splitting the geographical space into discrete local markets (such as counties, MSAs, or states), and assuming that all firms within a market are not spatially differentiated. While this approach is a reasonable approximation in many settings, there are two related reasons why it limits our ability to analyze spatial competition. First, space is continuous, so discretizing the space by "building high walls" that split one market from another could have important implications for the analysis. Markets that are defined too small would tend to under estimate competition, while markets that are defined too big will tend to over estimate competition. Second, the "correct" market definition may be specific to each customer. While a small customer may not be willing to search much and explore many (spatial) options, a large customer may search over a larger area, thus inducing competition from a much broader set of firms. Whether these limitations are important may depend on the context, and it is ultimately an empirical question.

Highway procurement provides an excellent setting to think about spatial competition in a more continuous way.[1] First, highway maintenance projects are quite homogeneous and do not require unique capabilities, arguably making spatial differentiation and distance from the project a first-order consideration for firms. In fact, 48 percent of the projects in our data are awarded to the closest bidder. Second, the locations of firms and projects are easily observed, and auction participation and bidding behavior are observed on a project-by-project basis. This provides good variation in the competitive environment, and allows us to analyze the importance of spatial competition in different situations. Third, dynamics and other "super-game" considerations, which are likely to play an important role in many applications of entry and location choice, are less important for auction participation, so focusing (as we do) on static and simultaneous-move analysis may be less restrictive. Finally, unlike many other cases of entry and location choice applications,[2] auctions provide an attractive setting where post-entry (or, in our case, post-participation) competition is observed, allowing us to "de-link" competitive effects from idiosyncratic shocks to participation costs.[3] Studying highway procurement may also be of a direct interest. In the United States,

[1] We are obviously not the first to explicitly model spatial differentiation. Some recent other examples include Thomadsen (2005), Davis (2006), and Houde (2008), who estimate spatial demand for retail services. One advantage of our setting is that we observe the same firm multiple times, facing different competitive environments, so endogeneity of the firm's location choice is much less of an issue.

[2] E.g., Bresnahan and Reiss (1991), Berry (1992), Seim (2006), and many others.

[3] Motivated by this nice feature, there are few other recent applications that model participation and bidding in auctions. We discuss these in more detail in the next section.

highway construction expenditure has been steadily increasing, surpassing $50 billion annually in recent years.

We focus our attention on two questions. First, we analyze how industry configuration – especially distance – affects bids, auction participation, and ultimately government procurement costs. Besides documenting the importance of distance in the current environment, this analysis helps assess mechanisms by which procurement costs can be reduced. Indeed, our second question evaluates how participation or bidding subsidy to firms that are relatively distant from a particular auction site would affect competition and procurement costs. The answer to this second question requires a model of firm participation and bidding decisions.

We collected data on all auctions that were run by the Michigan Department of Transportation (DOT) from January 2001 to August 2004. In our analysis, we focus on a subset of 865 projects that are primarily highway maintenance projects and fairly homogenous. For each auction, we observe the set of eligible firms that submit a request (but not an obligation) to bid, the set of actual bidders, and all the bids. As usual, the project is awarded to the lowest bidder. We provide some descriptive statistics that motivate the more structural approach we follow later. We show that firm's distance to the project site is statistically and economically important; closer firms submit lower bids and are more likely to submit bids. However, as competitors are further away firms are generally more likely to participate, but, when they do, they bid lower (more aggressively). Taken together, these findings are surprising. We would have expected that if distant firms bid higher and participate less often, then firms would also bid higher in the presence of distant competitors.

When we restrict attention to auctions with nearby competitors, we do obtain results consistent with firms bidding higher when facing more distant competitors. But while the competitor's effect on bidding reverses signs, it is small and insignificant. A possible explanation for these findings, which we explore in detail with our more structural approach, is that there is selection in participation: firms that decide to participate do so because they have relatively low project costs. This effect is likely to attenuate the effect of both own and competitors' distance on participation and bidding decisions.

In principle, with sufficient variation we could identify the effect of distance on bids and participation decisions without imposing much structure. Without a more nuanced model, however, it is not clear what type of parametric restrictions should be imposed when regressing bids on the entire vector of distances, or how selection into becoming a bidder plays out. Hence, even to answer our first question there may be an advantage to developing a more structural model of participation and bidding decisions. Overall, there are three reasons that lead us to develop a structural approach: the need to explicitly account for selection, the goal of quantifying – rather than testing – the effect of spatial differentiation on competition and procurement costs, and our interest in running counterfactuals, such as the effect of a bid preference program.

We consider a model where in the first stage firms that are eligible and that have submitted a request to bid must decide whether to participate in the auction based on a private signal of project cost and on a fixed cost of participation. Those firms that participate then observe their actual project costs and submit a bid in the auction. Potential selection in participation is introduced by allowing the private signal of a firm to be correlated with its actual project cost. Costs and entry signals are drawn independently across firms, but from potentially different distributions. Bidders know the set of firms that submit an intention to bid, but they do not perfectly know the set of firms that end up participating in the auction.

An analysis of the model leads to two relationships that are important for evaluating the effect of a change in distance on equilibrium entry and bidding – a precursor to assessing a bid preference program, as well as other related issues. First, the extent to which firms react to each other's participation decisions affects the composition of firms that participate when a subsidy is given to more distant firms. For example, if participation decisions are strong strategic substitutes, then a subsidy that encourages participation from distant firms may strongly discourage participation of not-so-distant firms, thus increasing procurement costs. On the other hand, if participation decisions are weak substitutes, then a relatively small subsidy may be sufficient to encourage an overall increase in participation and a reduction in procurement costs.

The extent to which participation decisions are substitutes depends on two forces working in opposite directions. When a firm participates more often, it is a stronger competitor and forces others to behave more aggressively, thus decreasing profits. However, if there is selection in participation, then a firm that participates more often is also a relatively weaker firm. This second effect attenuates the extent to which firms react to changes in competitors' participation decisions.

The second relationship that we hope to identify in the data is the extent to which changes in own and competitors' distances affect participation best responses. In turn, this relationship depends on the impact of distance on both project and entry costs. We face two potential problems in trying to identify the above relationships in the data. The first is standard: since we do not observe costs, we need a structural model to disentangle costs from markups given the observed bids. The second problem is specific to our participation model. Even if we were to observe costs of firms that participate in the auction, we would have to worry about selection: those firms who believe will have the lowest project cost will be more likely to participate. To resolve this second problem, we need instruments: something that shifts participation decisions for reasons unrelated to own project costs. In our case, these variables are competitors' distances and, to a lesser extent, the number of eligible firms.

Finally, we proceed to estimate the model. For reasons outlined later, we deviate from most of the literature and propose to estimate a fully parametric specification of the model. We take specific care in allowing enough flexibility to estimate the relationships in the data that we discussed to be crucial above.

[no results from the structural model yet; discussion of results and counterfactuals in the future]

We proceed as follows. We first discuss the related literature, and then in Section 3 we provide a simple numerical example, which analyzes how the direct and strategic effect of distance from the project site may affect participation and bidding decisions. In Section 4 we describe the environment and the data, and in Section 5 we report results from several descriptive regressions. We present the model and discuss its identification properties and estimation in Sections 6 and 7. We present our estimation results *[once we get them!]* in Section 8, and carry out several counterfactual exercises in Section 9. We then conclude. *[unfortunately there are no results yet from the structural model, so none of Sections 7-9 is written yet].*

2 Related auction literature

There exist two standard models of entry or participation in the auction literature, and a recent third that nests both of these and which we take as the basis of our structural model. All three models are for symmetric players with identically distributed private values (or costs). The first model is due to Levin and Smith (1994). They assume that firms decide whether to pay a fixed participation cost to learn their actual value (or cost). Only firms that pay this cost are allowed to bid. Equilibrium participation is in mixed strategies, but equilibrium can be purified as in Athey, Levin, and Seira (2004), where the participation cost is drawn randomly. In both cases, the important aspect of the setup is that participation decisions are not correlated with bidding decisions, hence ruling out the possibility of selection discussed above. The second entry model, as in Samuelson (1985), assumes that firms first privately observe their actual value (or cost) and only then decide whether to pay a participation cost and submit a bid. In this case, there is perfect correlation between the initial signal and the actual value, so selection is assumed to hold.

Recently, Marmer, Shneyerov, and Xu (2007) postulate a model that nests the two standard models above by allowing for a more flexible correlation structure between the participation signal and the actual project cost.[4] We work with this more general model, with the slight difference that we also allow for asymmetries among firms. Asymmetries are important to understand the effect of distances on participation and bidding. In contrast, the main objective of Marmer et al. (2007) is to develop a non-parametric approach that allows them to discriminate among these three models using exogenous variation in the number of firms.

A few empirical papers have estimated a structural model of participation and bidding decisions where players are asymmetric. Athey, Levin, and Seira (2004) allow for endogenous participation to compare revenue and efficiency of sealed bid vs. ascending timber auctions. There are two types of bidders in their auctions, mills (strong) and loggers (weak). Their analyses highlights that it is important to account for participation since auction design may differentially affect the participation decision of each of these types of players. Krasnokutskaya and Seim (2007) also allow for endogenous entry in a setting with two types of bidders, but in order to evaluate a policy that awards small businesses a bidding preference in certain auctions. Both Athey et al. (2004) and Krasnokutskaya and Seim (2007) follow the Levin and Smith (1994) model of participation, extended to allow for asymmetric firms. In particular, selection is ruled out in their models. Moreover, they focus on a comparison of two types of bidders, strong and weak. Hence, while it is plausible that firms that participate tend to be stronger than firms that do not participate, in their context this issue is likely to be of second-order importance compared to the initial differences in strength between weak and strong bidders.

[4] Antecedents to this model also include Hendricks, Pinkse, and Porter (2003) and Ye (2007).

Li and Zheng (2006) also estimate a version of the Levin and Smith (1994) model, but in contrast to the previous two papers they assume bidders are symmetric. One of their main objectives is to understand and measure the effect of an increase in the number of firms on equilibrium bidding. While more competition always increases bidding in an IPV auction, they show that when entry is taken into account there is an additional entry effect of competition that goes in the opposite direction. They then show that in their data the second effect dominates, and that increasing entry costs may actually decrease procurement costs. In a similar vein, we attempt to understand an a priori puzzling result that we see in the data with the help of a model, but our focus is on distance. Hence our need to incorporate both asymmetries and selection.

Some previous papers in the auction literature have also emphasized that distance to the project site may introduce important asymmetries and affect equilibrium bidding in procurement. These papers focus only on a bidding stage, where firms are asymmetric, and abstract from endogenous participation. Among the first to emphasize the relationship between proximity to the site and a higher likelihood of winning the contract are Bajari (1997), Porter (1999), and Bajari and Ye (2003). More recently, Flambard and Perrigne (2006) use data from snow removal auctions in Montreal to show that in a part of the city, where storage costs tend to be high, proximity to the site provides a relative cost advantage.

3 Illustrative Example

We start by providing a simple numerical example that illustrates our framework and guides our empirical analysis. We present the example somewhat loosely and focus on a graphical analysis of the results. The example is a special case of the full model we take to the data. In Section 6, when the full model is presented, we provide full details about the timing of the game, the information structure, and the solution concept.

Consider two firms $i = 1, 2$ that are eligible to submit a bid on a project. Firm i is characterized by its distance to the project, x_i, which is observed by both firms. Closer firms are more likely to have lower costs of project completion c_i. Specifically, let c_i be drawn from $N(x_i, \sigma_c^2 = 0.5)$, truncated to the $[0, 1]$ interval. Throughout, we fix $x_1 = 0.5$ but allow x_2 to vary between 0.1 and 0.9. Thus, from firm 1's perspective the strength of competition changes, while from firm 2's perspective competition is fixed, but its own competitiveness changes.

Initially, firms must simultaneously decide whether to bid in the auction without knowing their realized cost c_i. Rather, firms observe a private signal s_i which is drawn from $N(c_i, \sigma_s^2)$. That is, the signal is imperfectly correlated with the actual cost c_i. We present two cases, one where $\sigma_s^2 = 0.1$ so the signal is relatively informative, while the other where $\sigma_s^2 = 10$ so the signal is not very informative. Each firm incurs a fixed cost of 0.25 for learning its actual costs c_i and submitting a bid. Thus, a sufficiently bad signal or sufficiently bad competitive situation would make a firm unlikely to win the auction and therefore opt out and not submit a bid. Bidders simultaneously submit a bid without observing the set of actual bidders. As long as the lowest bid is below a reserve price of 0.75, the project is awarded to the lowest bidder at the cost submitted by such a bidder. An equilibrium of the game is a participation decision as a function of the private signal and a bidding function conditional on participation for each firm. Section 6 provide details of how we solve for equilibrium, and here we confine ourselves to a graphical analysis of the equilibrium strategies.

Figures 1 and 2 present the results. Figure 1 shows the equilibrium probability of participation for each firm, when $\sigma_s^2 = 0.1$ and when $\sigma_s^2 = 10$, as we move the distance of firm 2, x_2 (recall that $x_1 = 0.5$ throughout). The direct effect of (own) distance is shown by the graph for firm 2. As expected, as firm 2 is more distant, it is less likely to participate. The competitive effect of distant is shown by looking at firm 1. As firm 2 is more distant, entry to the auction is more profitable for firm 1, and it increases its participation probability. Note also that the competitive effect is quantitatively smaller than the direct effect; this could be seen by the smaller slope (in absolute value) for firm 1's graph. The importance of selection into the auction is captured by comparing the case with little selection (dashed lines; $\sigma_s^2 = 10$) and more selection (solid lines; $\sigma_s^2 = 0.1$). As can be seen, selection attenuates both the direct effect and the competitive effect of distance. This can be seen by the flatter graphs when selection is present. Figure 2 is analogous, and it shows the expected bid (conditional on participation) for each firm, when $\sigma_s^2 = 0.1$ and when $\sigma_s^2 = 10$, as we move the distance of firm 2, x_2 (recall that $x_1 = 0.5$ throughout). Both firms increase their expected bids as firm 2 gets further away. Firm 2 does it primarily because its costs are, on average, higher, while firm 1 does it because it faces softer competition. Somewhat surprisingly, the competitive effect of distance seems to be on average larger than the direct effect (but, recall, these are conditional expectations, and the direct effect of distance on participation is higher). The effect of selection here is slightly to attenuate the effect of distance, but primarily to make the auction more competitive. Once selection is stronger (more informative signal; solid lines) the bid curves are slightly flatter, but are primarily shifted down, making firms bid more aggressively. This is because selection makes participating firms more symmetric, so competition harsher. We use these illustrative effects in guiding our discussion of the descriptive empirical findings in the next sections.

4 Data and Environment

Our data comes from highway procurement in Michigan. Each month, the Michigan Department of Transportation (DoT) runs about 50-70 simultaneous auctions on a diverse set of highway procurement projects in the state. Up to 24 hours before the day of the auction, firms submit an intention to bid in a subset of these auctions, allowing the DoT sufficient time to confirm eligibility. Eligibility depends on the types of contract that a firm is pre-qualified to bid in, on the amounts of these contracts, and on the current capacity of a firm. Only a subset of firms (65% in the data we use below) who submit an intention to bid end up submitting a bid. While it has been common in much of the procurement literature to assume that firms have perfect information about the identity of competing bidders, it is not obvious to us – at least in the context of our setting – how well firms can predict who ends up bidding from among those who submit intentions. Therefore, throughout our analysis we will assume that the firms have full information about the identity of competing firms who submit an intention to bid, but may not know at the time of bidding which of these firms actually submits a bid. We also note here that while, in practice, many of these auctions occur simultaneously, so project portfolio considerations may be important, in what follows we abstract from this simultaneity and consider each auction separately. Analyzing the effect of the simultaneity of multiple auctions is beyond the scope of the current paper.

We collected data on all auctions that were run by the Michigan DoT from January 2001 to August 2004. While the original data includes 3,000 auctions, our analysis below restricts attention to the 865 projects that are primarily maintenance projects. This set of projects is fairly homogeneous. It represents fairly simple work, which does not require special capability. As our primary focus is on analyzing the effect of distance to the project, it seems appropriate to focus on simple and homogeneous set of projects, for which distance is more likely to play a first-order role.

Our data set includes details of each project and details about the set of all potential bidders. We also know the set of those firms who submitted an intention to bid, the actual bids by those who decided to participate and submit a bid, and the outcome of the auction. Our main emphasis in the analysis is on the role of distance. We matched each project location and each firm location with a street address, and we measure distance (using www.mapquest.com) between a firm and a project using driving miles and minutes. For firms with multiple locations, we record both the distance of the project to the firm's headquarters as well as the distance to the nearest branch.

Our current data contain information about 865 auctions and the 243 firms that submitted an intention to participate in at least one of those auctions. Table 1 presents summary statistics about the set of auctions. A typical project is estimated to cost from 100,000 dollars to several million dollars, with the typical auction attracting 3-8 participants and 2-5 bids. The winning bid is on average 6 percent lower than the engineers' estimate, although for some auctions it is as low as 25% while in others the winner bids more than 10% above the estimate. The winning bid is on average 7% below the second lowest. Depending on the project location, the nearest bidder could be as close as 1 mile from the project (e.g., when the project is in the Detroit area) or as far as 60 miles (when the project is at the Upper Peninsula of Michigan). As could be expected, distance plays an important role. Firms that are closer to the project are more likely to submit an intention to bid. Among those firms that submitted an intention to bid, closer firms are more likely to submit a bid and closer bidders are more likely to win the auction.

Table 2 presents summary statistics about the set of firms. As is the case in many similar data sets, the distribution of firm's size and activity is very skewed. The median firm participated in only 7 auctions (out of the 865 above), submitted only 3 bids, and won none. In contrast, the most active firm participated in more than 400 auctions (about half), submitted bids in most of them, and won more than 100. Still, distance matters: a firm is more likely to participate and bid on projects that are closer to its locations.

5 Motivating facts

Table 3 presents simple regression analysis of participation and bidding behavior. These results are quite representative of a much larger set of specifications. The top panel of Table 3 reports regressions in which the sample is the set of firms that submit a bid, and the dependent variable is the (logarithm of the) bid amount. The bottom panel reports analogous regressions in which the sample is the set of firms that submitted an intention to bid and the dependent variable is a dummy variable which is equal to one if the firm actually submitted a bid.

Our focus is on the effect of own and competitors' distance. Of course, one could be concerned about firm and auction omitted characteristics, which may affect the interpretation of the distance coefficients. For example, projects near the populated (by firms) area of Detroit are associated with short distances, while projects in the unpopulated Upper Peninsula of Michigan are associated with long distances. To the extent that the region is associated with other omitted characteristics of its projects, this could lead to misleading interpretation. To this end, we report four specifications for each regression. In one specification we use only a single control variable for the auction (we use the engineers' estimate for the project) and a single control variable for the firm (we use the number of projects in our data for which the firm submitted an intention to bid, as a proxy for firm's size). At the other extreme, we use a full set of auction and firm fixed effects. We also report other specifications with only firm fixed effects or with only auction fixed effects, and experimented with many others (not reported). As the tables illustrate, in most cases the effect of distance is quite stable across these very different specifications, so for the most part omitted variables are unlikely to be a concern. All specifications use a logarithmic transformation, so coefficients can be interpreted as elasticities (in the top panel) or semi-elasticities (in the bottom panel). For distance we use in these tables the driving distance in miles from the nearest branch of the firm. Results are qualitatively similar if we measure distance in minutes or if we measure it from the headquarters of the firm.

The top panel of Table 3 reports regressions in which the sample is the set of firms that submit a bid, and the dependent variable is (logarithm of) the bid amount (in thousand of current dollars). Overall, the engineer's estimates are highly correlated with the bids, and bigger (or, more precisely, more active) firms tend to be more competitive and submit lower bids. The effect of own distance on bids is highly significant, and quite stable across specifications. The elasticity of own distance ranges from about 0.02 (without auction fixed effects) to about 0.03 (with auction fixed effects), which is not trivial. This suggests that a bidder who is twice as far will submit a bid which is, on average, 3 percent higher. In more than 25 percent of the auctions (see Table 1), 3 percents increase in the bid would make the winner lose the auction. The fact that closer bidders submit lower bids – presumably reflecting lower costs – is not particularly surprising. As the example in Section 3 illustrates, it seems natural to expect that this should also lead bidders to bid more aggressively when they face a competitor who is near the project location. This is where the top panel of Table 3 reveals a surprising pattern; the distance of the nearest (to the project) competitor affects negatively the bid amount, with elasticities that are about half of the own distance effect. That is, if the nearest competitor (and similar results hold with alternative measures of competitors' distances) is closer to the project – presumably suggesting that he would bid more aggressively – bidders in fact raise their bids and bid less aggressively. This effect is quite stable across specifications and is largely significant (except for the specifications with both firm and auction fixed effects).

The bottom panel of Table 3 reports regressions (linear probability models) in which the sample is the set of firms that submitted an intention to bid and the dependent variable is a dummy variable which is equal to one if the firm actually submitted a bid (about 65% of the cases). The likelihood of submitting a bid (conditional on intention) is lower for larger auctions and higher for bigger firms. Closer firms are more likely to submit a bid, and this effect is, as before, large and stable across specifications. The effect of competitors' distance is less stable, ranging from positive coefficients (which is consistent with the intuitive effect based on the example of section 2) to negative effects, which are less intuitive but arise in what was our ex-ante preferred specifications (columns (10) and (12)).

The somewhat counter-intuitive effects of competitors' distance are puzzling. We believe that these qualitative effects are inconsistent with most competitive static bidding models. It turns out, however, that the negative coefficient on competitors' distance in the bid amount regression seems to be driven by auctions where at least one of the two nearest bidders is relatively far. That is, when we let the effect of distance be discontinuous, we obtain that the extensive margin of competitors' distance (analogous to taking a close competitor and making him far away) generates the negative effect, while the intensive margin of competitors (moving a competitor one mile further) goes in the more intuitive direction. This effect may be consistent with various dynamic incentives (arising from collusion, predation, etc.), which are out of the scope of the current paper. To focus on the patterns in the data that seem more consistent with competitive static bidding models, we therefore restrict attention to projects where at least two bidders are close (within 15 miles). Table 4 replicates the regressions of Table 3 for this set of auctions. Here the distance effects are more consistent with standard intuition, and closer competitors are associated with more aggressive bidding, although largely insignificantly so.

More generally, once firms are allowed to be asymmetric (due to distance), these "reduced form" regressions are hard to interpret. Using the nearest competitor to summarize competition is not satisfactory, and it is easy to think of alternatives (e.g., the average competitors' distance). Similarly, using auction fixed effects may only imperfectly control for auction unobservables, as different (asymmetric) firms may respond to the same effect asymmetrically. Selection may also be a concern: distant firms who intend to bid may do so because they have some unobserved cost advantage. However, these various effects could only be analyzed in the context of a fully specified participation and bidding model, which is our focus in the rest of the paper.

6 Model and identification

Setting and notation There are N firms that are eligible to submit a bid on a particular project. Initially, each of these firms gets a private estimate or participation signal s_i about its cost of completing the project. Given the private signal, each firm decides whether to participate in the auction by submitting a bid. There is a fixed cost of participating in the auction, $e_i > 0$. Only firms that incur this cost observe their actual cost of completing the project, c_i, and become bidders in the auction. Bidders simultaneously submit a bid without observing the set of actual bidders. As long as the lowest bid is below a reserve price r, the project is awarded to the lowest bidder at the cost submitted by such a bidder.

The pairs (s_i, c_i) are realizations of random variables (S_i, C_i) that are independently, but not necessarily identically, distributed across firms $i \in \{1, ..., N\}$. Let $F_i(s_i, c_i)$ denote the corresponding distribution function, with continuous density $f_i(s_i, c_i)$ positive on $[0, 1] \times [\underline{c}, \overline{c}]$, where $0 \leq \underline{c} < \overline{c}$. The reserve price is set below the maximum cost, $r \leq \overline{c}$.[5] In addition, the random variables (S_i, C_i) are affiliated (Milgrom and Weber, 1982), so that higher signals are (weakly) associated with higher costs. Without loss of generality, we can assume that the marginal distribution of the signals is uniform on $[0, 1]$

A strategy for firm i is a set $\mathbb{S}_i$ of participation signals for which a firm participates in the auction and a bidding strategy $\beta_i : [\underline{c}, \overline{c}] \longrightarrow \mathbb{R}_+$ that is followed in case of participation. A profile of strategies is denoted by $(\beta, \mathbb{S})$, where $\beta = (\beta_1, ..., \beta_N)$ and $\mathbb{S} = (\mathbb{S}_1, ..., \mathbb{S}_N)$. We will show that in equilibrium participation decisions are characterized by some threshold $\overline{s}_i$, i.e. $\mathbb{S}_i = \{s_i : s_i \leq \overline{s}_i\}$. Hence, we later use $\overline{s} = (\overline{s}_1, ..., \overline{s}_N)$ to denote participation decisions.

Let $H_i(b, \beta_{-i,} \mathbb{S}_{-i})$ denote firm i's probability of winning the auction when other firms follow strategies $(\beta_{-i}, \mathbb{S}_{-i})$. Expected payoffs of a participating firm with cost c_i that faces other firms' strategies $(\beta_{-i}, \mathbb{S}_{-i})$ and chooses an optimal bid are

$$\pi_i(c_i, \beta_{-i}, \mathbb{S}_{-i}) \equiv \max_b (b - c_i) H_i(b, \beta_{-i,} \mathbb{S}_{-i}). \tag{1}$$

In equilibrium, firms choose participation and bidding strategies that are optimal given the strategies of the other firms.

Definition 1 $(\mathbb{S}, \beta)$ *is an equilibrium of the participation/bidding game if for all* $i \in \{1, ..., N\}$,

- *for each* $c_i \in [\underline{c}, \overline{c}]$, $\beta_i(c_i) \in \arg\max_b (b - c_i) H_i(b, \beta_{-i,} \mathbb{S}_{-i})$,
- *for each* $s_i \in \mathbb{S}_i$, $\int_{\underline{c}}^{\overline{c}} \pi_i(c, \beta_{-i}, \mathbb{S}_{-i}) dF_i(c_i \mid S_i = s_i) - e_i \geq 0$, *and for each* $s_i \notin \mathbb{S}_i$, $\int_{\underline{c}}^{\overline{c}} \pi_i(c, \beta_{-i}, \mathbb{S}_{-i}) dF_i(c_i \mid S_i = s_i) - e_i \leq 0$.

Characterization of equilibrium Finding a best response to other firms' strategies requires a characterization of $\pi_i(c, \beta_{-i}, \mathbb{S}_{-i})$ for each $\left(\beta_{-i}, \mathbb{S}_{-i}\right)$. Instead, we follow a more convenient approach. First, we fix hypothetical equilibrium participation decisions and solve for the equilibrium of the resulting auction game. Second, we require participation decisions to actually be optimal.

For fixed participation decisions $\mathbb{S}$, let $\beta^{\mathbb{S}} = (\beta_1^{\mathbb{S}}, ..., \beta_N^{\mathbb{S}})$ denote an equilibrium profile for the auction game where bidders' primitives are $F_i(c_i \mid S_i \in \mathbb{S}_i)$, i.e. for all i and c_i,

$$\beta_i^{\mathbb{S}}(c_i) \in \arg\max_b (b - c_i) H_i(b, \beta_{-i,}^{\mathbb{S}} \mathbb{S}_{-i}). \tag{2}$$

[5] Since players are uncertain about the number of bidders in the auction, the reserve price will always bind.

As usual, $\beta_i^{\mathbb{S}}$ is increasing and its inverse $\phi_i^{\mathbb{S}}$ is obtained as the solution of a system of differential equations, with boundary conditions $\phi_i^{\mathbb{S}}(r) = r$ and $\phi_i^{\mathbb{S}}(\underline{c}) = \underline{b}$ for all i (uniqueness of equilibrium bidding strategies can be shown using Lebrun (2006) and additional assumptions on primitives). The resulting payoffs of player i depend on the entire profile $\mathbb{S}$ and are given by $\pi_i(c, \beta_{-i}^{\mathbb{S}}, \mathbb{S}_{-i})$.

Next, consider participation decisions. Assuming participation decisions are given by $\mathbb{S}$, if firm i observes signal s_i and decides to participate, its expected profits are

$$\int_{\underline{c}}^{\overline{c}} \pi_i(c, \beta_{-i}^{\mathbb{S}}, \mathbb{S}_{-i}) dF_i(c_i \mid S_i = s_i) - e_i. \tag{3}$$

Since profits from staying out are zero, in order for $\mathbb{S}$ to be an equilibrium it must be that equation (3) is positive for $s_i \in \mathbb{S}_i$ and negative otherwise. Moreover, it is easy to see that π_i is decreasing in c_i. Together with the fact that (S_i, C_i) are affiliated, it follows that the expression in (3) is decreasing in s_i. Hence we can focus, without loss of generality, on equilibria where firms choose participation thresholds $\overline{s}$. Re-writing equation (3) as

$$\Pi_i(s_i, \overline{s}_i, \overline{s}_{-i}) \equiv \int_{\underline{c}}^{\overline{c}} \pi_i(c, \beta_{-i}^{\overline{s}}, \overline{s}_{-i}) dF_i(c_i \mid S_i = s_i) - e_i, \tag{4}$$

equilibrium requires Π_i to be positive for all $s_i \leq \overline{s}_i$ and negative otherwise. Hence, equilibrium thresholds solve, for each i,

$$\widehat{\Pi}_i(\overline{s}_i, \overline{s}_{-i}) \equiv \Pi_i(\overline{s}_i, \overline{s}_i, \overline{s}_{-i}) = 0 \tag{5}$$

for $\overline{s}_i \in (0, 1)$, and $\widehat{\Pi}_i(\overline{s}_i, \overline{s}_{-i}) \geq (\leq) 0$ for $\overline{s}_i = 1$ $(\overline{s}_i = 0)$. Equilibrium does not necessarily exist without further assumptions. By assumption, the function Π_i is continuous in its first argument. If Π_i is also continuous in its second argument (auction comparative statics), then $\widehat{\Pi}_i$ is continuous in its first argument and an equilibrium exists.

To summarize, in order to find the equilibria of the participation/bidding game, we first solve for equilibrium bidding strategies $\beta^{\overline{s}}$ of an auction for fixed thresholds $\overline{s}$. We then use equilibrium profits in the auction to compute $\widehat{\Pi}_i$ for each i and use the system of equations defined by (5) to find equilibrium participation strategies $\overline{s}^*$. The corresponding bidding strategies are then $\beta^{\overline{s}^*}$.

As usual in entry models, equilibrium participation decisions need not be unique. There are two sources of multiplicities. One is the case where a firm never participates because it expects other firms to always participate – to the extent that which firm stays out is arbitrary, there will be multiplicity here. We can get rid of this source of multiplicity by adding restrictions that rule out some non-interior equilibria. For example, this can be achieved by having a reserve price strictly below $\overline{c}$ and by making the lowest signal sufficiently bad news so that a firm would like to stay out of the auction, even if no one else were to participate. Similarly, we could assume that the highest signal is sufficiently good news such that a firm would like to participate even if every other firm were to also participate. The second source of multiplicity occurs when there is more than one interior solution to the system of equations defined by (5). This multiplicity depends on functional form assumptions, and in the parametric specification that we take to the data we make sure to obtain uniqueness.

Modeling asymmetries Motivated by our empirical implementation, we assume that there is a vector of one-dimensional parameters $x = (x_1, ..., x_N)$, and that each player i is characterized by a parameter $x_i \in X_i \subset \mathbb{R}$ that is meant to capture firm asymmetries – in the empirical specification, x_i will depend on a firm's distance to the auction site, among other things. Since the marginal distribution of participation signals is fixed by assumption, we then write the primitives as $F(c_i \mid S_i = s_i, x_i)$ and $e(x_i) = e_i$, where $e : X_i \longrightarrow \mathbb{R}$. In the rest of the paper we write the expression in (5) as $\widehat{\Pi}_i(\overline{s}, x)$.

Sketch of the identification argument Since participation signals are uniformly distributed in the interval $[0, 1]$ without loss of generality, for a fixed x the primitives we would like to identify are the distribution of costs, conditional on each possible entry signal s_i, $F(c_i \mid S_i = s_i, x_i)$,[6] and the entry costs $e_x = (e(x_1), ..., e(x_N))$. For the purpose of this section, we view x_i as measuring the distance of the firm to the auction site. Additional covariates will be included in our estimation.

The argument for non-parametric identification relies on Guerre, Perrigne and Vuong (2000) and has been specifically discussed by Marmer, Shneyerov, and Xu (2007) for a symmetric version of the participation/bidding model presented above. Our participation/bidding model adds two dimensions to the standard auction setting. First, endogenous participation may affect the type of firms that will participate, and we need to identify the relationship between participation signals and project costs. Second, firm asymmetries play a crucial identification role in our model, since we rely on variation in competitors' characteristics, x_{-i}, to identify the primitives for a fixed x_i. In a similar vein, Marmer et al. (2007) rely on exogenous variation in the number of eligible firms, N, to test the extent to which participation signals are correlated with project costs when firms are symmetric.

For fixed x, suppose that firms play an equilibrium $\overline{s}^*(x)$, $\beta^{\overline{s}^*}(x)$ and that we observe x, and everyone's participation and bidding decisions. That is, we make the possibly strong assumption that there are no additional sources of firm heterogeneity which are observed by firms but not by us.[7] Following Guerre et al. (2000), we can estimate the probability $\widehat{H}_i(b, x)$ that each bidder wins with a bid b and use it to back out the project cost that makes the observed bid optimal. Given that we observe x, we can then identify the cost distribution conditional on participation, $F(c_i \mid S_i \leq \overline{s}_i^*(x), x_i)$. Moreover, since participation decisions are observed we can also obtain the probability that each player participates, $\overline{s}^*(x)$.

As pointed out by Marmer et al. (2007), the primitives of interest, $F(c_i \mid S_i = s_i, x_i)$ for all s_i and the vector of participation costs e_x are not identified without additional assumptions on observables. In order to identify the cost distribution conditional on each signal, we need variation in the participation decisions that does not also affect costs. In our specification, this variation can

[6] To the extent that the reservation price r is strictly below $\overline{c}$, we can only identify the primitives for values of c between $[\underline{c}, r]$.

[7] In fact, we don't need to literally observe x, as long as we observe x up to parameters that we estimate. In particular, what we cannot allow for our identification argument is unobserved heterogeneity that affects x. It is well known that the primitives of the auction are not identified in this latter case (Athey and Haile, 2005).

come from sufficient variation in x_{-i}. Suppose that we observe all auctions $x \in \Gamma$, and that for each $\widehat{x}_i$ and $\overline{s}_i \in [0,1]$, there exists $\widehat{x}_{-i}$ such that $\widehat{x} \in \Gamma$ and $\overline{s}_i^*(x) = \overline{s}_i$. Intuitively, this condition is unlikely to be satisfied if there is only variation in N, but seems more plausible when there is variation in both x and N. The source of variation in x also seems clearer, as variation in N may be driven by many unobserved factors. If this condition is satisfied and we do observe a player x_i choosing all possible participation thresholds, then we can identify $F(c_i \mid S_i = s_i, x_i)$ for each s_i and use it to obtain entry costs by setting equation (5) equal to zero. In practice, the extent that variation in x gives us the entire range of equilibrium decisions is determined by our data. To the extent that the data do not provide sufficient variation, identification will rely on some functional-form assumptions, which is what we do in the next section.

To summarize, we are performing the following conceptual exercises, whose analogs we exploit in the data. First, we can think of fixing x_i and moving around x_{-i}. Second, we can think of fixing x_{-i} and moving around x_i. The former variation is the key to identification and allows us to identify the primitive of interest for a fixed x_i. The latter fixes competition and shifts firm i's strength and allows us to non-parametrically identify how the object of interest changes for each value of x_i.

7 Parametric assumptions and estimation

Although the model is, in principle, non-parametrically identified, we continue by making parametric assumption before taking the model to data. We start this section by explaining why we choose to follow a fully parametric approach, where identification, at least partially, may rely on specific functional-form assumptions. To minimize the extent to which important relationships in the data are restricted by parametric assumptions, we continue by intuitively discussing comparative statics in the participation/bidding model. This exercise allows us to isolate four relationships in the data that are crucial for determining the effect of policy changes on equilibrium outcomes. We then present an econometric specification that is flexible enough to capture these relationships. Finally, we describe the estimation strategy.

7.1 Discussion of possible estimation strategies

We discuss three possible estimation strategies, emphasizing both their desirability and feasibility.

Non-parametric estimation Ideally, we would proceed non-parametrically and estimate our primitives following the non-parametric identification argument sketched earlier. As in most application in the literature, we do not have enough data to proceed in this way, and must therefore rely on parametric assumptions. Our setting is likely to be even less attractive for a fully nonparametric approach, because bidder types are continuous, so even with a lot of data, we would be asking a lot from the data to fully identify the primitives non-parametrically.

Semi-parametric estimation. As most of the literature, we could follow the non-parametric identification argument above by making parametric assumptions on the distribution of opponents' bids faced by each player, $\widehat{H}_i(b,x)$. In addition, since we also have a participation stage, we need to estimate equilibrium entry thresholds $\overline{s}^*(x)$ from the data. Again, data restrictions make parametric assumptions more tractable. In particular, we would impose parametric restrictions on the dependence of $\overline{s}^*(\cdot)$ on x. The main advantage of this approach is that it does not require us to solve for the equilibrium of the auction game – once $\widehat{H}_i(b,x)$ is estimated, we can follow the standard procedure of backing out project costs from the first order condition of the bidder's maximization problem. However, we note two drawbacks of following this approach. First, both $\widehat{H}(\cdot)$ and $\overline{s}^*(\cdot)$ are not primitives, but are rather determined endogenously given the primitives. It is not clear what would be a reasonable way in which, say, the entire vector of everyone's distances enters both of these expressions. And it is not clear how exactly such restrictions map to restrictions on the actual primitives of interest. Second, even if we felt comfortable with parameterizing the previous endogenous expressions,[8] we would still need to make sure that the implied primitives are consistent with both parameterizations of $\widehat{H}_i(\cdot)$ and $\overline{s}_i^*(\cdot)$. Of course, if the distribution of costs in the auction, $F(c_i \mid S_i \leq \overline{s}_i^*, x_i)$, were obtained non-parametrically for each $(\overline{s}_i^*, x_i)$, then by definition there would be no need to check for consistency. In practice, however, insufficient data precludes us from non-parametrically estimating this distribution for all values of $\overline{s}_i$ in $[0,1]$ and for all values of x_i in a sufficiently fine grid. Hence, we would need to impose parametric restrictions on $F(c_i \mid S_i \leq \overline{s}_i, x_i)$ as well, but then it is highly unlikely that these restrictions will be consistent with the restrictions we impose on $\widehat{H}_i(\cdot)$ and $\overline{s}_i^*$. In addition, even if we did not impose parametric restrictions on $F(c_i \mid S_i \leq \overline{s}_i, x_i)$, we would still have the problem that the primitives obtained from our original parameterization of $\overline{s}_i^*(\cdot)$ are unlikely to be consistent with the actual equilibrium thresholds.

Parametric estimation For the above reasons, we proceed by specifying a functional form for the primitives $F(c_i \mid S_i = s_i, x_i)$ and $e_i(x_i)$ and estimating the corresponding parameters via maximum likelihood. This approach also involves trade-offs. On one hand, we need to numerically solve for the bidding game, which involves solving for a system of differential equations and can be quite time consuming. Because of this, we need to choose a specification that produces a bidding game that depends on as few parameters as possible. On the other hand, we need to be careful to avoid a priori restrictions on relationships that are important for the questions we seek to answer. We explain how we resolve this trade-off in the remainder of this section.

7.2 Comparative statics

We can use the two-stage characterization of equilibrium in Section 6 to obtain intuition about comparative statics results. It is well-known that equilibrium comparative statics at the auction

[8] For example, we could specify a few primitives, solve numerically for the equilibrium of the game, and get some understanding of how the vector of everyone's distances enters $\widehat{H}(\cdot)$ and $\overline{s}^*$.

stage are hard to characterize and, furthermore, tend to depend on particular parametric assumptions. Hence, the intuition that follows is speculative. Once we obtain results at the auction stage, we can go back to the participation stage and find a fixed point or a solution to the system of equations defined by (5). Here, it is useful to look at how the "best response" of a firm, $br_i(\bar{s}_j, x) = \left\{\bar{s}_i : \widehat{\Pi}_i(\bar{s}_i, \bar{s}_j, x) = 0\right\}$, is affected by changes in the threshold strategies of other firms. It is in this sense that we can discuss the extent to which participation decisions are strategic substitutes. Finally, the effect of changes in parameters, such as distance, can be represented by shifts in firms' participation "best responses." Whether participation decisions are strategic substitutes or complements depends on the effect of $\bar{s}_i$ and $\bar{s}_j$ on $\widehat{\Pi}_i(\bar{s}, x, N)$ – if they affect $\widehat{\Pi}_i$ in the same direction, participation decisions are strategic substitutes; otherwise, they are complements.

First, consider how $\widehat{\Pi}_i(\bar{s}, x, N)$ depends on $\bar{s}_j$, $j \neq i$. Since $\bar{s}_j$ only enters the expression through π_i, it suffices to understand the effect of $\bar{s}_j$ on the outcome of the bidding game. There are two direct effects of an increase in $\bar{s}_j$ on player i's probability of winning**.** The first effect is to increase player j's participation in the auction, while the second effect is to increase the likelihood that player j draws a higher project cost. All players will change their equilibrium strategies as a result. It seems plausible that the first effect would lead players to bid more aggressively, since now they are more likely to find themselves in an auction with an additional bidder. The second effect, however, would probably make bidders bid less aggressively. For future reference, we refer to these effects as the effects of greater competition and the effects of weaker competition, respectively.

As a hypothetical situation, suppose that a change in $\bar{s}_j$ did not affect equilibrium bidding at the auction stage. Then the effect of stronger competition dominates the effect of weaker competition – hence $\widehat{\Pi}_i$ is decreasing in $\bar{s}_j$. The reason for this dominance is that it is always better to face one less potential competitor than to face one more, irrespective of how weak the additional competitor may turn out to be. However, equilibrium bidding strategies will change as a result of the increase in $\bar{s}_j$, thus potentially affecting results in a way that seems hard to predict. The important point here is that, while participation decisions are likely to be strategic substitutes, the effect of weaker competition mentioned above will mitigate the extent to which participation decisions are substitutes. This reduction will be larger whenever a small change in participation threshold produces a sufficiently large increase in costs. Intuitively, this is likely to be the case whenever a firm has initially large uncertainty about its project cost and the signal that it observes is sufficiently informative. Depending on which effect dominates, we would expect $\widehat{\Pi}_i(\bar{s}, x, N)$ to either decrease or increase as $\bar{s}_j$ increases.

Now, consider the effect of $\bar{s}_i$ on $\widehat{\Pi}_i(\bar{s}, x, N)$. As before, we can reason that either more aggressive or less aggressive bidding is plausible as a result of an increase in $\bar{s}_i$. However, there is now an additional effect to take into account. An increase in $\bar{s}_i$ also affects the term $dF_i(c_i \mid S_i = \bar{s}_i)$, thus resulting in higher costs for player i (due to affiliation). Since π_i is decreasing in c_i, it follows that this other effect works to decrease $\widehat{\Pi}_i(\bar{s}, x, N)$.

Finally, consider the equilibrium response to an exogenous increase in the parameter that identifies player i, x_i. Suppose that x_i measures the distance to the auction site, and that roughly, higher distances are associated with higher costs. There are two likely effects from the increase of x_i on

$\widehat{\Pi}_i(\bar{s}, x, N)$. First, facing a weaker opponent, other bidders will bid less aggressively. However, this less aggressive response is unlikely to compensate for the fact that bidder i bids more aggressively and faces higher costs. Second, participation cost $e(x_i)$ is likely to increase with x_i. Hence, an increase in own distance will decrease $\widehat{\Pi}_i(\bar{s}, x, N)$. In contrast, an increase in the distance of an opponent means that one faces weaker competition – only the weaker competition effect is present, since participation thresholds are being held fixed. It then seems likely that best responses are decreasing in own distance and increasing in opponents' distances.

Figure 3 shows the effect of an increase in the distance of player 2 under two different scenarios. In Figure 3a, where the correlation between the participation signal and project cost is weak, the participation "best responses" are relatively steep. In contrast, in Figure 3b the correlation between the participation signal and project cost is strong, and the effect of weaker competition mentioned above makes best responses flatter. In both cases, an increase in firm 2's distance affects firm 2's best responses to a greater extent than firm 1's best responses. As a result, in the first case (Figure 3a) there is a relatively large change, in the opposite direction, in both firm's participation thresholds, while in the second case (Figure 3b) there is a relatively small change in the equilibrium participation decision of firm 1. This difference will be important for some of our counterfactuals. For example, when evaluating a bid preference program for distant firms, the extent to which nearby firms will reduce their participation will affect the efficiency loss created by such a program.

To summarize, the above discussion clarifies that in order to understand the effect of changes in distance (or other parameters) on the equilibrium of the game it is important to obtain the following four relationships from the data: the initial uncertainty about project costs and the precision of the participation signal (to establish the extent to which participation decisions are substitutes), and the effect of distance on both projects and participation costs (to determine the extent to which best responses shift with changes in distance).

7.3 Econometric specification

One of the primitives of interest is $F(c_i \mid S_i = s_i, x_i)$, where x_i will capture both auction and firm characteristics. We assume that $x_i = m\beta + z_i\gamma$, where m is a vector of observed auction characteristics, z_i is a vector of firm characteristics that could vary across auctions (including distance to the project site), and (β, γ) are parameters to be estimated. Note, as we mentioned earlier, that x_i only depends on observables.

We consider the following parameterization. Let the cost of the project be drawn from a normal distribution truncated to the interval $[0, 1]$,

$$C_i \sim N_{[0,1]}(x_i, h_c^{-1})$$

where h_c is the precision of the distribution.

We assume that firms observe a signal k_i that is not drawn from a uniform distribution, and later normalize this signal by defining a participation signal s_i that is a function of k_i and is uniformly distributed on the interval $[0, 1]$ –hence consistent with our setup in Section 6. The signal k_i is drawn from a normal distribution with mean c_i with precision h_s,

$$K_i \mid C_i = c_i \sim N(c_i, h_s^{-1}).$$

Finally, let the other primitive of the game, the participation cost, be given by $e(x_i) = \alpha + x_i\delta$. The parameters to be estimated are then $(\beta, \gamma, \alpha, \delta, h_c, h_s)$.

One convenient property of this specification is that the distribution of C_i conditional on signal k_i is also normally distributed, truncated to the interval $[0, 1]$,

$$C_i \mid K_i = k_i \sim N_{[0,1]}\left(wx_i + (1-w)k_i, \left(h_c^{-1} + h_s^{-1}\right)^{-1}\right),$$

where $w = (1 + h_s/h_c)^{-1} \in (0, 1)$. Part of the convenience lies in the fact that, for fixed (h_c, h_s), the distribution of $C_i \mid k_i$ depends only on one parameter, $wx_i + (1-w)k_i$. Unfortunately, the same is not true for the primitive of the auction game, $C_i \mid K_i \leq k_i$, for which a sufficient statistic for (k_i, x_i) does not exist. Each bidder in the auction stage will then be characterized by two parameters, (k_i, x_i), hence increasing the computational burden when we solve for the equilibrium of each bidding game. Nevertheless, we can show that any specification where the primitive of the auction game depends on a unique sufficient statistic would be too restrictive for our purposes.

To make the setup consistent with Section 6, we can make a normalization and assume that the participation signal observed by firms is $s_i = G(k_i)$, where $G(\cdot)$ is the probability distribution of k_i. Hence, the signal s_i is distributed uniformly on $[0, 1]$, and in equilibrium the participation threshold $\overline{s}_i$ can be interpreted as the probability that player i participates in the auction.

The specification above is rich enough to allow us to determine the relationships that were found in the previous section to be important for comparative statics. In particular, h_c measures the initial uncertainty about project costs, h_s measures the precision of the entry signal (and whether there can be selection at all), and (γ, δ) measure the effect of distance on project and entry costs, respectively.

7.4 Estimation

We estimate the model using maximum likelihood. The key computational burden we need to address is that the equilibrium of the participation/bidding game is solved numerically, which is very time consuming. Following the characterization of equilibrium in Section 6, we need to solve for the equilibrium of an auction game for fixed characteristics of the firms, and we then need to find a fixed point of $\widehat{\Pi}$ to determine equilibrium participation.

There are two ways in which we reduce the computational burden. First, the solution to the auction stage only depends on the precisions (h_c, h_s) and on the vector of firm observables $x = (x_1, ..., x_N)$. In our specification, x depends on parameters (β, γ), but we avoid solving for a particular bidding game more than once when we compute the likelihood. In addition, the participation cost parameters (α, δ) are not inputs to the bidding game. Therefore, we solve "offline" for the equilibrium of auctions as a function of (h_c, h_s, x). We then take (h_c, h_s, x) and the participation cost parameters (β, γ) and solve for the fixed point of $\widehat{\Pi}$. Once we have solved for equilibria as a function of these parameters, we then estimate the parameters of the model using maximum likelihood. The benefit is that instead of solving for the equilibrium of the participation/bidding game at every step of the likelihood maximization routine, we simply look up this solution based on the "offline" results.

The second way in which we reduce the computational burden is to solve "offline," as explained above, but only for a dense grid over (h_c, h_s, x) and over (β, γ). To the extent that solving for the equilibrium of the auction game is more computationally intensive than solving for a fixed point, the first set of parameters will be taken from a coarser grid. We then verify that the solution (participation thresholds and bidding strategies) are smooth in these arguments, and we interpolate them to cover the entire space of parameter values. Since the number of grid points, which is the number of auctions we need to solve, increases exponentially in the number of firms, we have to restrict the auctions we use for estimation to those where the number of firms that submit an intention to bid is not too large. At the moment, we think we can accommodate (in terms of computation time) auctions with up to five firms.

More formally, we can denote the data for a given auction to be given by auction characteristics m and, for each $i = 1, ..., N$, by firm-auction characteristics z_i, participation decisions I_i, and bids b_i that are observed if $I_i = 1$. The parameters that we try to estimate are $\theta = (\beta, \gamma, \alpha, \delta, h_c, h_s)$, so the likelihood function for a given auction is given by

$$L\left(\{I_i\}_{i=1}^N, \{b_i \cdot I_i\}_{i=1}^N \,|\, \{_i z\}_{i=1}^N, m, \theta\right). \tag{6}$$

The likelihood for the entire data is simply the product for the likelihood of each auction. Since the unique role of (β, γ) is to determine the type x_i of firm i in the game, we can write the likelihood function as

$$L\left(\{I_i\}_{i=1}^N, \{b_i \cdot I_i\}_{i=1}^N |h_c, h_s, x, \alpha, \delta\right) \cdot L(x| \{_i z\}_{i=1}^N, m, \beta, \gamma). \tag{7}$$

As explained above, for each $Z = (h_c, h_s, x, \alpha, \delta)$, we can then solve ("offline") for the participation/bidding game to obtain equilibrium participation threshold strategies $\overline{s}^*(Z)$ and equilibrium bidding strategies (conditional on participation) $\beta^{\overline{s}^*}(\cdot|Z)$, with inverse $\phi^{\overline{s}^*}(\cdot|Z)$. We can then use these strategies to write the first part of the likelihood function above as

$$L\left(\{I_i\}_{i=1}^N, \{b_i \cdot I_i\}_{i=1}^N |Z\right) = \prod_{i=1}^N (1 - I_i) L_i\left(I_i = 0|Z\right) + I_i L_i\left(I_i = 1|Z\right) \Pr\left(b_i|Z, I_i = 1\right) \tag{8}$$

where

$$\begin{aligned} L_i\left(I_i = 0|Z\right) &= 1 - \overline{s}_i^*(Z) && (9) \\ L_i\left(I_i = 1|Z\right) &= \overline{s}_i^*(Z) && (10) \\ \Pr\left(b_i|Z, I_i = 1\right) &= f\left(\phi^{\overline{s}^*}(b_i|Z)\right) \frac{d}{db_i} \phi^{\overline{s}^*}(b_i|Z) && (11) \end{aligned}$$

where f is the density that corresponds to $F(c_i \mid S_i \leq \overline{s}_i^*, x_i)$.

[We are now at the point that we are finishing running the first stage, "offline" auctions, and getting ready to obtain estimates using the second stage.]

8 Results

[to be written at some future date]

9 Counterfactuals

[to be written at some future date]

10 Conclusions

[to be written at some future date]

11 References

Athey, Susan, and Philip Haile (2005). "Nonparametric Approaches to Auctions." *Handbook of Econometrics*, Volume 6, Forthcoming.

Athey, Susan, Jonathan Levin, and Enrique Seira (2004). "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions." Mimeo. Stanford University.

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

Bajari, Patrick, and Lixin Ye (2003). "Deciding Between Competition and Collusion." *Review of Economics and Statistics* 85(4), 971-989.

Berry, Steven (1992). "Estimation of a Model of Entry in the Airline Industry." *Econometrica* 60, 889-917.

Bresnahan, Timothy F., and Peter C. Reiss (1991). "Entry and Competition in Concentrated Markets." *Journal of Political Economy* 99(5), 977-1009.

Davis, Peter (2006). "Spatial Competition in Retail Markets: Movie Theaters." *Rand Journal of Economics* 37(4), 964-982.

Flambard, Veronique, and Isabelle Perrigne (2006). "Asymmetry in Procurement Auctions: Evidence from Snow Removal Contracts." *Economic Journal* 116(514), 1014-1036.

Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong (2000). "Optimal Nonparametric Estimation of First-Price Auctions." *Econometrica* 68, 525-574.

Hendricks, Kenneth, Joris Pinkse, and Robert H. Porter (2003). "Empirical Implications of Equilibrium Bidding in First Price, Symmetric, Common Value Auctions." *Review of Economic Studies* 70(1), 115-145.

Houde, Jean Francois (2008). "Spatial Differentiation in Retail Markets for Gasoline." Mimeo. University of Wisconsin.

Krasnokutskaya, Elena, and Katja Seim (2007). "Bid Preference Programs and Participation in Highway Procurement." Mimeo. University of Pennsylvania.

Lebrun, Bernard (2006). "Uniqueness of the Equilibrium in First-price Auctions." *Games and Economic Behavior* 55, 131-151.

Levin, Dan, and James L. Smith (1994). "Equilibrium in Auctions with Entry." *American Economic Review* 84, 585–599.

Li, Tong, and Ziaoyong Zheng (2006). "Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions." Mimeo. Vanderbilt University.

Marmer, Vadim, Artyom Shneyerov, and Pai Xu (2007). "What Model for Entry in First-Price Auctions? A Nonparametric Approach." Mimeo. University of British Columbia.

Porter, Robert H. (1999). "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions." Mimeo. Northwestern University.

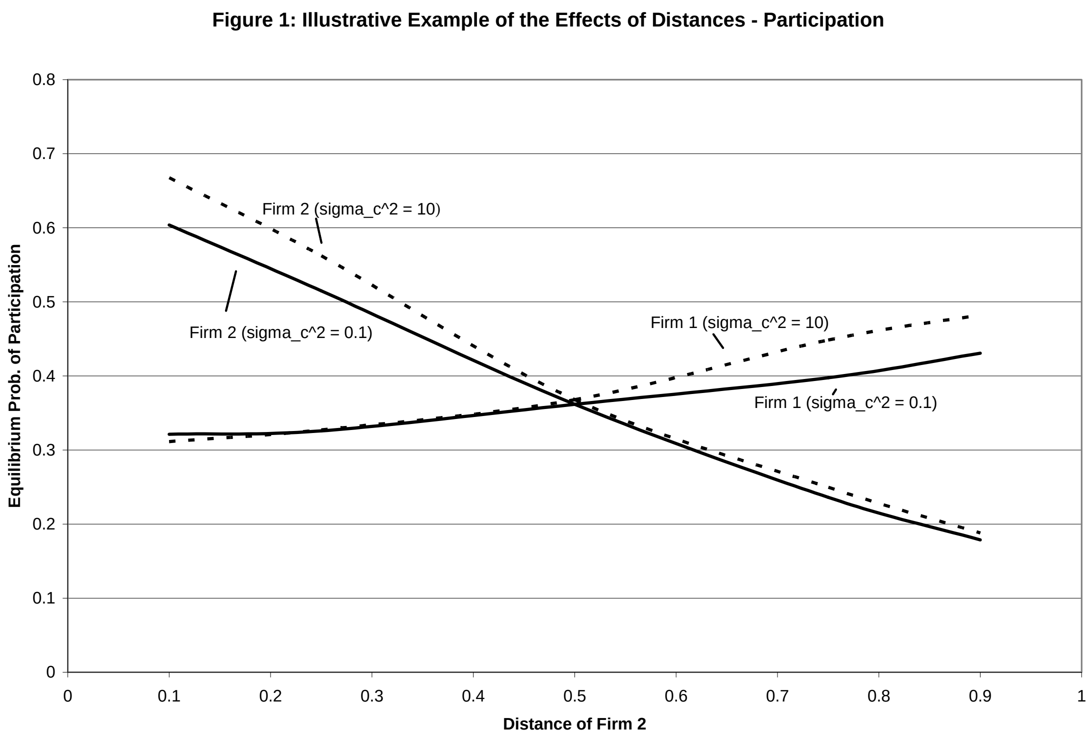
Samuelson, William F. (1985). "Competitive Bidding with Entry Costs." *Economics Letters* 17(1-2), 53-57.

Seim, Katja (2006). "An Empirical Model of Firm Entry with Endogenous Product-Type Choices." *Rand Journal of Economics* 37(3), 619-640.

Thomadsen, Raphael (2005). "The Effect of Ownership Structure on Prices in Geographically Differentiated Industries." *Rand Journal of Economics* 36(4), 908-929.

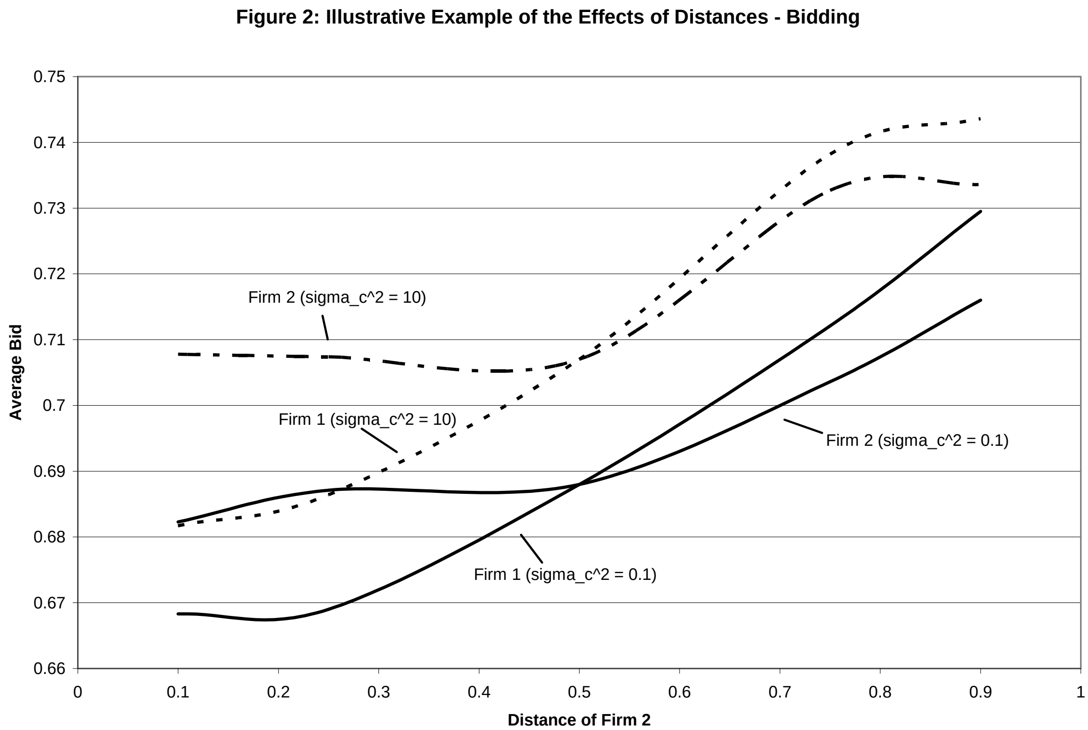
Ye, Lixin (2007). "Indicative Bidding and a Theory of Two-Stage Auctions." *Games and Economic Behavior* 58, 181-207.

Figure 1: Illustrative Example of the Effects of Distances - Participation



The figure presents the equilibrium probabilities of participation from the numerical example presented in Section 3. The horizontal axis is the distance of firm 2 from the site, while the corresponding distance of firm 1 is set to 0.5. The parameter σ_s^2 affects selection: low values imply better information at the time of participation decisions, and therefore more selection.

Figure 2: Illustrative Example of the Effects of Distances - Bidding



The figure presents the equilibrium expected bid (conditional on participation) from the numerical example presented in Section 3. The horizontal axis is the distance of firm 2 from the site, while the corresponding distance of firm 1 is set to 0.5. The parameter σ_s^2 affects selection: low values imply better information at the time of participation decisions, and therefore more selection.

Figure 3: Comparative Statics

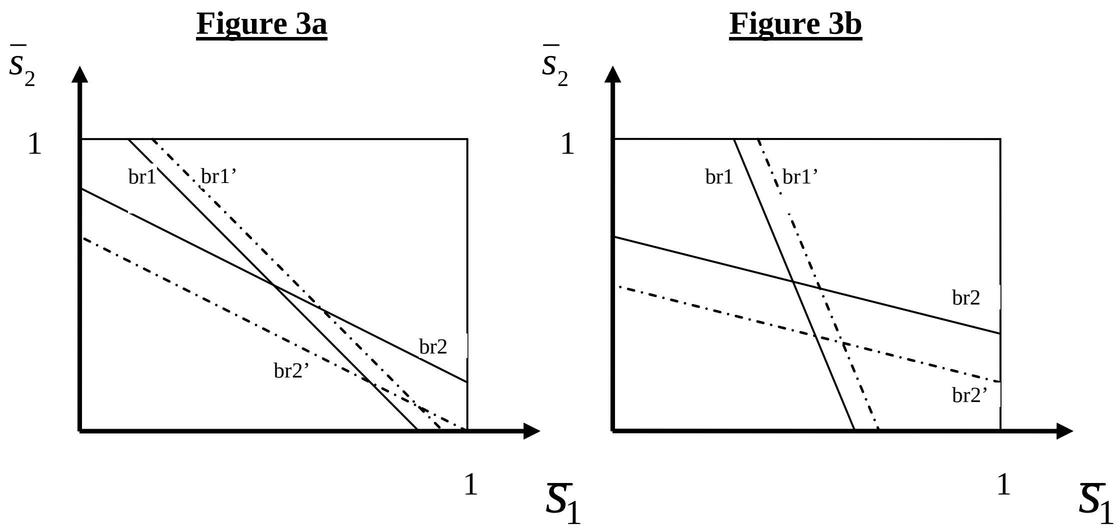


Figure 3a shows the "best response" participation functions with a small selection effect, while Figure 3b shows the case of a large selection effect. The functions $br_i(\bar{s}_j, x)$ are defined as the $\bar{s}_i$ that solves $\widehat{\Pi}_i(\bar{s}_i, \bar{s}_j, x) = 0$. The figure shows how a change in x_j shifts these functions and changes equilibrium participation.

Table 1: Auction charcateristics

	Obs.	Mean	Std. Dev.	5th Pctile	25th Pctile	50th Pctile	75th Pctile	95th Pctile
Number of Firms "Intending to Bid"	865	5.96	4.11	2	3	5	8	15
Number of Actual Bidders	865	3.86	2.29	1	2	3	5	8
Engineers' Estimate (`000$)	865	888.1	1,097.0	108.9	268.0	528.1	1,082.1	2,836.5
Winning Bid / Engineers' Est.	865	0.943	0.140	0.760	0.863	0.948	1.013	1.124
Winning Bid / Second Lowest Bid	809	0.928	0.071	0.801	0.903	0.947	0.976	0.995
Closest Bidder (driving miles)	865	20.6	35.4	1.5	5.9	12.2	25.6	62.5
Second Closest Bidder	809	40.4	47.7	5.5	12.6	29.3	48.6	115.4
Distance Rank of Winner	865	2.14	1.67	1	1	2	3	6
Average Distance (Bidders)	865	44.7	45.0	8.2	20.0	34.5	55.5	107.7
Average Distance ("Intent to Bid")	865	51.6	47.5	11.4	25.4	40.5	61.1	128.2
Average Distance (All)	865	192.1	85.3	128.8	142.3	159.5	191.1	418.5

Table 2: Bidder charcateristics

	Obs.	Mean	Std. Dev.	5th Pctile	25th Pctile	50th Pctile	75th Pctile	95th Pctile	Highest
Number of Firms "Intending to Bid"	243	21.21	44.54	1	2	7	20	91	413
Bids Submitted	243	13.75	34.87	0	1	3	11	67	329
Auctions Won	243	3.56	13.11	0	0	0	2	14	157
Average Distance when bid	197	57.4	67.8	9.3	21.2	34.3	60.8	197.2	
Average Distance when Intend to Bid	243	62.5	66.1	9.3	24.2	39.9	73.2	216.6	
Average Distance (all auctions)	243	192.1	72.1	128.5	153.2	172.2	195.7	371.3	

Table 3: Descriptive regressions for the full data

Dependent Variable:	**Log(Bid Amount)**																							
Dependent Variable:	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)	
Log(Distance to Project)	0.019	(0.003)	0.029	(0.003)	0.019	(0.003)	0.032	(0.003)	0.022	(0.003)	0.024	(0.004)	0.020	(0.003)	0.030	(0.004)	0.023	(0.003)	0.024	(0.003)	0.021	(0.003)	0.029	(0.004)
Log(Distance of Closest Other Bidder)									-0.010	(0.003)	-0.011	(0.006)	-0.006	(0.003)	-0.004	(0.007)								
Log(Distance of Closest Other Participant)																	-0.014	(0.003)	-0.016	(0.007)	-0.011	(0.003)	-0.007	(0.007)
Firm Characteristics:																								
Log(Auctions Participated)	-0.015	(0.002)	-0.006	(0.002)					-0.015	(0.002)	-0.006	(0.002)					-0.014	(0.002)	-0.006	(0.002)				
Firm Fixed Effects					Yes		Yes						Yes		Yes						Yes		Yes	
Auction Characteristics:																								
Average Distance ("Intent to Bid")	0.988	(0.003)			0.988	(0.003)			0.987	(0.003)			0.988	(0.003)			0.986	(0.003)			0.988	(0.003)		
Auction Fixed Effects			Yes				Yes				Yes				Yes				Yes				Yes	
No. of Observations (Bids)	3,341		3,341		3,341		3,341		3,285		3,285		3,285		3,285		3,314		3,314		3,314		3,314	
No. of Auctions	865		865		865		865		809		809		809		809		838		838		838		838	
R-Squared	0.973		0.990		0.977		0.992		0.973		0.990		0.977		0.992		0.973		0.990		0.977		0.992	

Dependent Variable:	**Bid dummy (conditional on participation) - Liear Probability Model**																							
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)	
Log(Distance to Project)	-0.071	(0.006)	-0.103	(0.007)	-0.093	(0.007)	-0.124	(0.009)	-0.083	(0.006)	-0.063	(0.009)	-0.099	(0.007)	-0.086	(0.011)	-0.083	(0.006)	-0.118	(0.009)	-0.100	(0.007)	-0.147	(0.011)
Log(Distance of Closest Other Bidder)									0.058	(0.006)	0.128	(0.019)	0.043	(0.006)	0.100	(0.019)								
Log(Distance of Closest Other Participant)																	0.060	(0.006)	-0.051	(0.021)	0.047	(0.006)	-0.072	(0.021)
Firm Characteristics:																								
Log(Auctions Participated)	0.072	(0.005)	0.057	(0.006)					0.068	(0.005)	0.054	(0.006)					0.066	(0.005)	0.056	(0.006)				
Firm Fixed Effects					Yes		Yes						Yes		Yes						Yes		Yes	
Auction Characteristics:																								
Log(Engineers' Estimate)	-0.045	(0.006)			-0.033	(0.007)			-0.042	(0.006)			-0.042	(0.006)			-0.041	(0.006)			-0.032	(0.007)		
Auction Fixed Effects			Yes				Yes				Yes				Yes				Yes				Yes	
No. of Observations (Intentions to Bid)	5,155		5,155		5,155		5,155		5,099		5,099		5,099		5,099		5,128		5,128		5,128		5,128	
No. of Auctions	865		865		865		865		838		838		838		838		838		838		838		838	
R-Squared	0.078		0.239		0.216		0.373		0.092		0.254		0.222		0.385		0.095		0.238		0.224		0.373	

Table 4: Descriptive regressions for auctions with at least two nearby (< 15 miles) biddres who intend to bid

Dependent Variable:													**Log(Bid Amount)**											
Number of Firms "Intending to Bid"	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)	
Log(Distance to Project)	0.013	(0.004)	0.019	(0.003)	0.011	(0.005)	0.018	(0.004)	0.013	(0.004)	0.019	(0.004)	0.011	(0.005)	0.017	(0.005)	0.013	(0.004)	0.019	(0.003)	0.011	(0.005)	0.019	(0.005)
Log(Distance of Closest Other Bidder)									0.005	(0.005)	0.001	(0.009)	0.002	(0.005)	-0.005	(0.009)								
Log(Distance of Closest Other Participant)																	0.001	(0.005)	0.002	(0.011)	0.002	(0.006)	0.005	(0.013)
Firm Characteristics:																								
Log(Auctions Participated)	-0.006	(0.003)	-0.008	(0.003)					-0.006	(0.003)	-0.008	(0.003)					-0.006	(0.003)	-0.008	(0.003)				
Firm Fixed Effects					Yes		Yes						Yes		Yes						Yes		Yes	
Auction Characteristics:																								
Average Distance ("Intent to Bid")	0.986	(0.004)			0.985	(0.005)			0.986	(0.004)			0.985	(0.005)			0.986	(0.003)			0.985	(0.005)		
Auction Fixed Effects			Yes				Yes				Yes				Yes				Yes				Yes	
No. of Observations (Bids)	1,431		1,431		1,431		1,431		1,427		1,427		1,427		1,427		1,431		1,431		1,431		1,431	
No. of Auctions	298		298		298		298		294		294		294		294		298		298		298		298	
R-Squared	0.973		0.992		0.977		0.993		0.973		0.992		0.977		0.993		0.973		0.992		0.977		0.993	

Dependent Variable:													**Bid dummy (conditional on participation) - Liear Probability Model**											
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)		(11)		(12)	
Log(Distance to Project)	-0.104	(0.009)	-0.096	(0.010)	-0.105	(0.011)	-0.114	(0.012)	-0.103	(0.009)	-0.056	(0.011)	-0.105	(0.011)	-0.071	(0.013)	-0.104	(0.009)	-0.102	(0.012)	-0.106	(0.011)	-0.130	(0.014)
Log(Distance of Closest Other Bidder)									0.020	(0.011)	0.234	(0.031)	0.011	(0.011)	0.204	(0.030)								
Log(Distance of Closest Other Participant)																	0.017	(0.012)	-0.051	(0.045)	0.007	(0.012)	-0.101	(0.046)
Firm Characteristics:																								
Log(Auctions Participated)	0.046	(0.008)	0.030	(0.009)					0.047	(0.008)	0.028	(0.009)					0.046	(0.008)	0.030	(0.009)				
Firm Fixed Effects					Yes		Yes						Yes		Yes						Yes		Yes	
Auction Characteristics:																								
Log(Engineers' Estimate)	-0.031	(0.010)			-0.022	(0.011)			-0.032	(0.010)			-0.022	(0.011)			-0.031	(0.010)			-0.022	(0.011)		
Auction Fixed Effects			Yes				Yes				Yes				Yes				Yes				Yes	
No. of Observations (Intentions to Bid)	2,360		2,360		2,360		2,360		2,356		2,356		2,356		2,356		2,360		2,360		2,360		2,360	
No. of Auctions	298		298		298		298		298		298		298		298		838		838		838		838	
R-Squared	0.076		0.200		0.273		0.421		0.077		0.224		0.273		0.436		0.077		0.200		0.273		0.423	