Atl Econ J
https://doi.org/10.1007/s11293-020-09690-2

BEST UNDERGRADUATE PAPER

Selective Entry in Highway Procurement Auctions

Sabrina Peng[1]

Abstract When participating in an auction is costly, a potential bidder has to decide whether to enter the auction or not. The extent to which the potential bidders know their private cost before making their entry decisions determines how selective the entry process is. Endogenous selective entry is common in many auctions and it has important implications for designing auctions, in particular, choosing the bid discount policy that is frequently used in public procurements to achieve distributional goals of the government. Prior empirical studies of the bid preferences were based on frameworks that either did not explicitly model endogenous participation or assumed endogenous, but non-selective participation. This study empirically investigated whether the entry process is selective in the highway procurement auctions run by the California Department of Transportation. To this end, the asymmetric affiliated-signal model was adapted to permit endogenous selective entry. Model parameters, including entry costs and distributions of construction costs for regular and fringe companies, were estimated nonparametrically. The results show evidence favoring selective entry of the fringe firms and imply that the level of bid discount required to achieve the procurement buyer's policy objective may be lower than what is previously found in the literature under the assumption of non-selective entry.

Keywords Procurement auctions · Endogenous entry · Selection · Bid discount policy

JEL D44 · H57 · L74

Introduction

In public procurement, bid discount is a policy commonly used by government agencies to promote domestic, local or small firms and companies located in

✉ Sabrina Peng
cp4dr@virginia.edu

1 Department of Economics, University of Virginia, Charlottesville, VA, USA

Published online: 11 December 2020

economically disadvantaged areas, or owned by minority groups. The bid discount policy awards a contract to the favored firm when its bid is within a certain percentage of the lowest bid among the unfavored firms. This rule does not change the price of the contract, which is the winner's bid.

When bidder participation is endogenous and possibly selective, whether there is selection and how selective the entry process is have potentially strong effects on the optimal level of bid discount. For example, Sweeting and Bhattacharya (2015) showed that a seller's revenue-maximizing bid discount level can vary from 2.5 to 12.5% depending on the degree of selection. A weaker player's probability of winning increases with the degree of selection.

In the existing literature, bid-preference programs have been studied under either exogenous entry (Marion 2007) or endogenous but non-selective entry (Krasnokutskaya and Seim 2011) models. Incorrectly assuming non-selection may lead to incorrect estimates of model primitives (Roberts and Sweeting 2010) which in turn bias the policy recommendation, in this case evaluations of bid-preference programs. This study empirically investigated whether the non-selective entry assumption holds in California's highway procurement auction data. Model primitives were estimated using the affiliated-signal (AS) model (Gentry and Li 2014). The AS model is a flexible entry model. It nests a wide range of entry processes depending on the information firms have prior to making their entry decisions. While a non-selective entry model assumes that potential bidders do not have any private information of their costs of project completion before deciding whether to participate in the auction, the AS model allows firms to each receive a private information signal of its project cost prior to the entry decision. A perfectly informative signal implies full selection and a perfectly uninformative signal implies non-selection.

The California Department of Transportation (Caltrans) awards highway construction and repair contracts through first-price sealed-bid auctions. Caltrans aims to promote disadvantaged bidders in their procurement auctions and implements a 5% bid discount for small businesses (SB) in state-funded contracts.[1] California's Small Business Participation Program sets the allocative goal to award 25% of all state-funded contract dollars to SB. The auctions in this sample were not subject to any preference program. Firms were categorized into two types, fringe and non-fringe (Bajari et al. 2014), to account for differences in size and experience. Firms' participation and bidding decisions were modelled as a two-stage game. The equilibrium described by Gentry and Li (2014) was adapted to the setting of low-bid auctions with asymmetric bidders.

If firms have private information about their project costs prior to making the entry decision (i.e. the signal is not perfectly uninformative), only self-selected firms that are more cost-efficient will participate in the auction. Otherwise, the entrants are a random sample of the potential bidders. Thus, selective entry implies that the project cost distribution of the entrants is a truncated distribution of the project cost distribution of all firms, whereas non-selective entry implies that the former is the untruncated distribution. This study investigated whether entry is selective by comparing the two distributions. To recover these distributions, the nonparametric identification strategy of Guerre et al. (2000) was used. The result favors selective entry for fringe firms, which

[1] Firms promoted by Caltrans include SB, disadvantaged business enterprises, and disabled veteran business enterprises. Information can be found on Caltrans website (Caltrans 2020a).

implies that the level of bid discount required for Caltrans to achieve its policy objective may be lower than what is previously found in the literature under the non-selective entry assumption.

This paper relates to the literature of empirical analysis of auctions in three ways. First, the evidence of selection found in the California highway procurement market contributes to the growing literature on empirical testing of different entry models (Li and Zheng 2009, 2012). Second, this study ties the theoretical literature on selective entry and auction design (Sweeting and Bhattacharya 2015) to the empirical literature evaluating bid preference programs (Marion 2007; Krasnokutskaya and Seim 2011). Third, the nonparametric estimation method in Gentry and Li (2014) was appllied to the Caltrans empirical setting, which contributes to the literature on empirical auction studies, where models with partially selective entry are estimated (Roberts and Sweeting 2013, 2016; Bhattacharya et al. 2014). While all three prior studies used fully parametric estimation approaches, assuming no unobserved heterogeneity across auctions permits this study to take a nonparametric approach. Although the conditional distribution of project costs on signals cannot be fully estimated using this approach, important model primitives can still be recovered. This attempt to empirically estimate auction models with endogenous, potentially selective entry nonparametrically is the first in the English language literature to my knowledge.

Data

This study analyzed data from Caltrans on road paving contracts from 1999 to 2005.[2] This is the same sample used by Bajari et al. (2014). The data consist of 819 contracts in 12 districts in California, adding up to a total of $2.21 billion contract value (as measured by the winning bids). 348 unique contractors participated in the auctions and submitted a total of 3666 bids.

The contracts vary in size, and the type of work ranges from small-scale highway resurfacing to four-lane freeway construction. A project includes a number of work items to be completed according to specifications provided by Caltrans. For each work item, engineers at Caltrans provide an estimated quantity needed and an estimated unit price. The engineer's estimate of project cost is thus the sum of unit price times item quantity across all work items. This measure is given to the contractors and is intended to represent the fair and reasonable price the government expects to pay. When deciding how much to bid on a contract, a firm bases the decision on its cost of completing the specified project, which this study refers to as project cost. Project cost depends on factors such as prior experiences in similar projects and a firm's current workload relative to its production capacity. Therefore, the project cost is private information to each firm.

To bid on a contract, interested contractors must submit completed bid documents, which require the contractor to provide the unit price for each item, the list of subcontractors and the work item(s) subcontracted to each subcontractor. Bid preparation takes time and effort and typically involves negotiating with subcontractors.

[2] Excluded were contracts from 2001 and the first half of 2003 whose details are no longer available from Caltrans (Caltrans 2020b).

Therefore, bidding is costly and a contractor may not know the exact project cost before negotiations. In the model below, such bid preparation costs are treated as entry costs. The AS model was used to model the firms' imperfect knowledge about the project cost prior to entry and bidding.

Asymmetry among bidders is a salient feature of the Caltrans procurement market. In this sample, the top 20 firms (as ranked by market share) captured 73.4% of the market share (i.e. share of total contract dollars awarded), whereas the remaining 328 firms each had less than 1% market share. Following Bajari et al. (2014), this study refers to the top 20 firms as non-fringe (or regular) bidders and the remaining firms as fringe bidders to account for asymmetry in size and experience. Of the 819 contracts, 47.4% were awarded to the fringe firms, although half of the fringe firms won only one contract in the sample. These fringe firms likely operate as subcontractors most of the time.

Since the firms' entry behavior plays a central role, it is important to correctly identify potential bidders for each auction in the sample.[3] Taking an approach similar to the one used by Roberts and Sweeting (2013), an auction's potential bidders were defined as the bidders plus the firms satisfying the following conditions: 1) in the prior 90 days, submitted bids on contracts of similar size in the same district, and 2) whose distance to the project site in the contract mentioned in 1) was within 110 miles.[4]

Summary statistics are shown in Table 1. The median number of potential bidders was nine, with three non-fringe potential bidders and six fringe potential bidders, and the median number of fringe and non-fringe bidders were both two. Among potential bidders, on average 45% of fringe firms and 64% non-fringe firms entered (calculated using number of bidders/number of potential bidders for each type of firms). Although fringe firms and non-fringe firms had similar normalized bids (bid/engineer's estimate and winning bid/engineer's estimate), the projects won by fringe firms tended to have smaller sizes (as measured by the engineer's estimate) than those won by non-fringe firms.

Evidence of Selection

This section presents evidence favoring selective entry in the sample. The method used here is that of Roberts and Sweeting (2011).

First, Athey et al. (2011) showed that under non-selective entry, in a type-symmetric entry equilibrium, that the weak type enters with positive probability implies that the strong type enters with probability one. That the strong type enters with probability less than one implies the weak type enters with probability zero. Thus, if the entry process is non-selective, whenever some fringe firms enter, all non-fringe potential bidders should enter. However, for the 728 auctions that some fringe firms entered, in only 34.3% did all non-fringe potential bidders entered. Similarly, in the 520 auctions where not all non-fringe potential bidders enter, only 8.41% had zero fringe potential bidders enter.

[3] Prior studies of Caltrans auctions typically used the number of project plan holders as a proxy for the number of potential bidders, but this measure is not available in the data used for this study.

[4] The third quartile of the same measure among all observed bids in the sample is 110. Eighty-two contracts in the sample did not have any same-district contracts in the past 90 days, in which case, the four most recent of such contracts were used, where four is the median number of the same measure in the sample excluding those 82 contracts.

Table 1 Summary statistics of project characteristics and bids

Variable	Mean	SD	Median	Min.	Max.	Observ.
Across contracts in the sample						
Engineer's estimate ($million)	2.89	7.26	0.95	0.09	105.61	819
when fringe wins	1.64	2.35	0.71	0.09	19.96	388
when non-fringe wins	4.01	9.62	1.33	0.09	105.61	431
Winning bid/engineer's estimate	0.95	0.20	0.93	0.38	2.19	819
Fringe	0.95	0.22	0.92	0.38	2.19	388
Non-fringe	0.94	0.17	0.94	0.83	1.82	431
Number of items	32.78	30.98	21	4	326	819
Number of bidders	4.48	2.16	4	2	19	819
Fringe	2.77	2.20	2	0	17	819
Non-fringe	1.70	1.02	2	0	5	819
Number of potential bidders	9.94	5.50	9	2	34	819
Fringe	7.15	5.23	6	0	29	819
Non-fringe	2.79	1.30	3	0	7	819
Across bids in the sample						
Bid/engineer's estimate	1.05	0.26	1.02	0.38	7.86	3666
Fringe	1.06	0.28	1.03	0.38	7.86	2272
Non-fringe	1.03	0.22	1.00	0.49	2.73	1393

Source: Own calculations using 1999–2005 data from Bajari et al. (2014)

Second, non-selective entry implies that the entrants are a random sample of the potential bidders. This can be tested by estimating a Heckman selection model (Heckman 1976). The regression equation regresses observed bids on project characteristics. The selection equation involves the number of potential bidders. The exclusion restriction is that the number of potential bidders affects entry behavior (since a larger number of potential bidders means more competition for the firms) without affecting bids directly. If the entry process is selective, a higher probability of entry will correspond to a lower project cost and thus a lower bid, implying a negative correlation between the error terms of the two equations. Table 2 compares the ordinary least squares (OLS) results to the Heckman results. Although the estimated inverse Mills ratio λ appeared to be negative (consistent with the negative correlation hypothesis), it was not statistically significant. However, the coefficient for the fringe indicator was much higher in the estimated Heckman model, suggesting that the OLS model not accounting for selection might have underestimated the difference in bids between fringe and non-fringe firms. Note that in both columns of Table 2, the coefficient for the log of the engineer's estimate was significant and close to 1, implying that the bids submitted were centered around the engineer's estimate of a project. This provides some support for using the engineer's estimate to control for project heterogeneity and assuming no unobserved heterogeneity in the estimation approach herein.

These two findings indicated that the entry process is likely to be selective in the sample. To further investigate, a structural model allowing selection was developed and estimated.

Table 2 Regression results with and without accounting for selection

	OLS	Heckman
Constant	0.733***	0.728***
	(0.119)	(0.118)
Fringe	0.038***	0.045***
	(0.006)	(0.010)
ln (engineer's estimate)	0.947***	0.948***
	(0.008)	(0.008)
Working days	0.0001**	0.0001**
	(0.00006)	(0.00006)
Number of fringe bidders	−0.016***	−0.015***
	(0.003)	(0.003)
Number of non-fringe bidders	−0.014*	−0.012
	(0.008)	(0.008)
Number of items	0.001***	0.0009***
	(0.0003)	(0.0003)
λ		−0.024
		(0.028)

[a] Standard errors adjusted for clustering provided in parentheses. Dependent variables of the two columns are both log of bids. Heckman model results were estimated with the full maximum likelihood method. The selection equation included variables from the regression equation plus the number of potential bidders of each type, which were incorporated as a flexible polynomial (up to degree four). Both columns included controls for year, month, and district. λ is the estimated inverse Mills ratio. The R^2 for OLS is 0.972. Statistical significance: 1% (***), 5% (**), 10% (*). Source: Own calculations using 1999–2005 data from Bajari et al. (2014)

Model

Consider a standard first-price sealed-bid auction (with no reserve price or bid subsidy) held by a procurement buyer to allocate a project among N potential suppliers of two types, fringe firms and non-fringe firms (type $\tau \in \{f, n\}$). N_f, N_n are common knowledge. Following Krasnokutskaya and Seim (2011), this study assumed the project costs of the firms are independent private values (IPV) drawn from type-specific distributions F_f, F_n. Although the project costs of each firm are private information, the two distributions are common knowledge. An interested contractor must pay an entry cost K_τ (e.g., opportunity cost and bid preparation cost) before participating in the auction. This entry cost was modelled as an auction-specific cost dependent on the firm's type, similar to the one in Roberts and Sweeting (2013).

The presence of entry cost divides the auction game into two stages: first entry, then bidding. In Stage 1, each potential bidder, i, observes a private signal, s_i, of its (not yet known) project cost, c_i, and all potential bidders choose whether to enter the auction simultaneously. In Stage 2, the n entrants from Stage 1 learn their exact project costs and submit bids. The lowest bidder is awarded the contract at the price of its bid. Following Krasnokutskaya and Seim (2011), this study assumed the bidders observe the number of entrants n_f, n_n when they entered Stage 2.

$F_\tau(c, s)$ denotes the type-specific joint cumulative distribution function of C_i and S_i. Let the support of C_τ be $\left[c_l^\tau, c_u^\tau\right]$. The AS model only assumes that a higher signal implies a potentially higher cost: for each firm i, $s' \geq s$ implies $F_\tau(c|s') \leq F_\tau(c|s)$. Following Gentry and Li (2014), the signals of both types were normalized to have a uniform marginal distribution such that $S_i \sim U[0, 1]$. This normalization was used because the distributions of the signals are not of direct interest and cannot be inferred from the observed entry decisions and bids. Since $F_\tau(c|s)$ is type-dependent and $F_\tau(c, s) = F_\tau(c|s)F(s)$, the variable of interest $F_\tau(c, s)$ is still type-dependent. Because the observed variation in $F_\tau(c, s)$ in the data is fully captured by $F_\tau(c|s)$, this normalization can be done without loss of generality.

In the model, the equilibrium strategy consists of an entry strategy used by firms to decide whether to enter in Stage 1 and a bidding strategy used to decide how much to bid in Stage 2. This study focuses on the following type-symmetric monotone pure strategy Bayesian Nash equilibrium.

Stage 1 Equilibrium Entry Strategy

The entry strategy in Stage 1 was characterized by a type-specific entry threshold $\overline{s}_\tau \in [0, 1]$: potential bidder i of type τ chooses to enter if and only if $s_i \leq \overline{s}_\tau$. The (selected) distribution of project costs among entrants at threshold $\overline{s}_\tau$ is the following truncated distribution:

$$F_\tau^*\left(c; \overline{s}_\tau\right) \equiv F_{C_\tau, \overline{S}_\tau}\left(c | s \leq \overline{s}_\tau\right) = \frac{\Pr\left(C_\tau \leq c, S \leq \overline{s}_\tau\right)}{\Pr\left(S \leq \overline{s}_\tau\right)} = \frac{1}{\overline{s}_\tau} \int_0^{\overline{s}_\tau} F_\tau(c|t) dt \tag{1}$$

where the last equality follows from $S_i \sim U[0, 1]$.

Suppose bidder i, if entering, has Stage 2 expected profit $\Pi_\tau^{II}\left(c_i; n_f, n_n\right)$ (elaborated in the next subsection). In Stage 1, let π_τ^I be the expectation of Π_τ^{II} over all possible (n_f, n_n) (since i gets zero profit if it does not enter, only the cases where i enters are considered):

$$\pi_\tau^I\left(c_i; \overline{s}_f, \overline{s}_n, N_f, N_n\right) = \sum_{\substack{1 \leq n_\tau \leq N_\tau, \\ 0 \leq n_{-\tau} \leq N_{-\tau}}} \Pi_\tau^{II}\left(c_i; n_f, n_n\right) \cdot \Pr\left(n_\tau, n_{-\tau} | \overline{s}_\tau, \overline{s}_{-\tau}, N_\tau, N_{-\tau}, i \; enters\right), \tag{2}$$

where $-\tau$ refers to the type of firm other than bidder i's type. Since $n_\tau \sim Binomial\left(N_\tau, \overline{s}_\tau\right)$, $\Pr\left(n_\tau, n_{-\tau} | \overline{s}_\tau, \overline{s}_{-\tau}, N_\tau, N_{-\tau}, i \; enters\right)$ in the last equation can be rewritten as

$$\Pr\left(n_\tau, n_{-\tau} | \cdot\right) = \binom{N_\tau - 1}{n_\tau - 1} \overline{s}_\tau^{\,n_\tau - 1} \left(1 - \overline{s}_\tau\right)^{N_\tau - n_\tau} \binom{N_{-\tau}}{n_{-\tau}} \overline{s}_{-\tau}^{\,n_{-\tau}} \left(1 - \overline{s}_{-\tau}\right)^{N_{-\tau} - n_{-\tau}}. \tag{3}$$

Now, consider a type-τ bidder i's entry decision in Stage 1. Given its signal s_i, i's expected profit from entry is

$$\begin{aligned} \Pi_\tau^I\left(s_i; \overline{s}_f, \overline{s}_n, N_f, N_n\right) &\equiv E_{C_\tau}\left[\pi_\tau^I\left(c_i; \overline{s}_f, \overline{s}_n, N_f, N_n\right) | s_i\right] \\ &= \int_{c_l^\tau}^{c_u^\tau} \pi_\tau^I\left(c_i; \overline{s}_f, \overline{s}_n, N_f, N_n\right) f_\tau(c_i | s_i) dc_i. \end{aligned} \tag{4}$$

Bidder i enters the auction if and only if its expected profit is no smaller than the entry cost K_τ. The entry threshold is thus determined by the condition of zero net profit of entry. The equilibrium entry threshold s_f^*, s_n^* are uniquely characterized by the system of equations

$$\Pi_f^I\left(s_f^*; s_f^*, s_n^*, N_f, N_n\right) = K_f, \tag{5}$$

and

$$\Pi_n^I\left(s_n^*; s_f^*, s_n^*, N_f, N_n\right) = K_n. \tag{6}$$

Note that if $\Pi_\tau^I\left(1; 1, s_{-\tau}^*, N_f, N_n\right) > K_\tau$, then $s_\tau^* = 1$. If $\Pi_\tau^I\left(0; 0, s_{-\tau}^*, N_f, N_n\right) < K_\tau$, then $s_\tau^* = 0$.

Stage 2 Equilibrium Bidding Strategy

In Stage 2, given c_i, n_f, n_n, type-τ entrant i submits bid b_i to maximize expected profit. The type-symmetric equilibrium strategies for first-price low-bid auctions below are well known in the literature, but under the AS model, the distribution of project costs among entrants is now the truncated distribution in Eq. 1.

Letting $y_\tau(b)$ be the inverse bid function, under the IPV assumption, i's expected profit is

$$\Pi_\tau^{II}\left(c_i; n_f, n_n\right) = (b_i - c_i)\left(1 - F_\tau^*\left(y_\tau(b_i); \bar{s}_\tau\right)\right)^{n_\tau - 1}\left(1 - F_{-\tau}^*\left(y_{-\tau}(b_i); \bar{s}_{-\tau}\right)\right)^{n_{-\tau}}. \tag{7}$$

The first-order condition (FOC) with respect to the bid is thus

$$b = c + \left((n_\tau - 1)\frac{f_\tau^*\left(y_\tau(b); \bar{s}_\tau\right)y_\tau'(b)}{1 - F_\tau^*\left(y_\tau(b); \bar{s}_\tau\right)} + n_{-\tau}\frac{f_{-\tau}^*\left(y_{-\tau}(b); \bar{s}_{-\tau}\right)y_{-\tau}'(b)}{1 - F_{-\tau}^*\left(y_{-\tau}(b); \bar{s}_{-\tau}\right)}\right)^{-1}. \tag{8}$$

Intuitively, a firm's bid is equal to its project cost plus some markup. The boundary conditions of these two differential equations (since $\tau \in \{f, n\}$) can be shown as follows. Let $B_\tau(c)$ be the bid function. If $B_f\left(c_l^f\right) \neq B_n\left(c_l^n\right)$ i.e. $B_\tau\left(c_l^\tau\right) < B_{-\tau}\left(c_l^{-\tau}\right)$, then a type τ bidder drawing project cost c_l^τ and competing with only type $-\tau$ firms can still win by bidding slightly higher than $B_\tau\left(c_l^\tau\right)$ and thus earn a slightly higher profit. Similarly, if $B_\tau\left(c_u^\tau\right) < B_{-\tau}\left(c_u^{-\tau}\right)$, then a type τ bidder with c_u^τ and facing only one type $-\tau$ bidder with $c_u^{-\tau}$ can bid slightly higher than $B_\tau\left(c_u^\tau\right)$ to still win the auction and earn a higher profit. Therefore, the FOC uniquely characterizes the equilibrium bidding strategies $B_\tau(c)$ with the following boundary conditions:

$$B_f\left(c_l^f\right) = B_n\left(c_l^n\right) = \underline{b} > \max\left\{c_l^f, c_l^n\right\}, \tag{9}$$

and

$$B_f\left(c_u^f\right) = B_n\left(c_u^n\right) = \bar{b} > \max\left\{c_u^f, c_u^n\right\}, \tag{10}$$

where $\underline{b}, \bar{b}$ are the minimal and maximal submitted bid, respectively.

In the empirical analysis, this study assumed that the firms play the equilibrium entry and bidding strategies described in this section. A potential bidder enters if and only if its signal is lower than the entry threshold. An intuitive interpretation is that a potential bidder chooses to participate in the auction if according to a vague idea of project cost, its cost-efficient level reaches some threshold so that the firm expects to earn a net profit from entry. Entrants then bid according to the same type-symmetric monotone bid functions to maximize profits.

Note that the type-symmetric entry equilibrium described in this section exists but is not necessarily unique (Athey et al. 2011). That is, there exists $\left(\bar{s}_f, \bar{s}_n\right)$ that characterizes the equilibrium entry behavior, but there may be multiple qualifying $\left(\bar{s}_f, \bar{s}_n\right)$. This study assumed all firms enter according to the same $\left(\bar{s}_f, \bar{s}_n\right)$, which were recovered from the data.

Estimation

This section shows how the project cost, entry threshold and entry cost were recovered from the data. Instead of imposing any assumption on the specific distribution of the project cost, this study combined the nonparametric methods in Guerre et al. (2000) and Gentry and Li (2014) to estimate the marginal distributions of the project costs. The nonparametric approach taken here is the main difference between the empirical analysis in this study and those in the existing literature of endogenous and (possibly) selective entry models.

There are two reasons why this study was able to explore this nonparametric method whereas prior literature did not. First, no unobserved heterogeneity across auctions was assumed in the sample. When there is unobserved heterogeneity, identification becomes difficult and a nonparametric method cannot be used (Roberts and Sweeting 2013). Second, Gentry and Li (2014) showed that in order to identify $F(c|s)$ at every s, the econometrician ideally should observe a continuous auction-level instrument which shifts entry behavior without affecting the underlying distributions. Such continuous entry behavior shifters are limited in practice. The auction-level instrument in the data used in this study was the number of potential bidders of each type, which is a discrete variable. Therefore, $F(c|s)$ cannot be fully identified at every s. However, the marginal distribution $F_C(c)$ can be recovered by combining observations from various levels of s if there is complete variation of s in the data, which was the case for this sample.

Identification of $\mathbf{F}_\tau^*(c; \bar{s}_\tau)$

Assuming all firms bid according to the equilibrium bidding strategy given in the model section, a firm's project cost satisfies Eq. (8). The number of bidders of each type are known and bids are observed. Since the firms' project cost distribution depends on the characteristics of the project and varies by auction, $F_{\tau}^{*}(c;\overline{s}_{\tau})$ is auction-specific. Given an auction k, let $G_{\tau,k}(b)$ be the (cumulative) probability distribution of bids among type τ bidders and let $g_{\tau,k}(b)$ be the corresponding probability density function. It is well known in the literature that a bidder's project cost can be estimated using

$$b = c + \left((n_{\tau} - 1) \frac{g_{\tau,k}(b)}{1 - G_{\tau,k}(b)} + n_{-\tau} \frac{g_{-\tau,k}(b)}{1 - G_{-\tau,k}(b)} \right)^{-1}, \tag{11}$$

where $G_{\tau,k}(b)$ and $g_{\tau,k}(b)$ can be estimated from the bids data. The distribution of the recovered c is the truncated distribution $F_{\tau,k}^{*}(c;\overline{s}_{\tau})$. Thus, $F_{\tau,k}^{*}(c;\overline{s}_{\tau})$ is uniquely identifiable for an auction. However, the sample size of the bids data from a single auction is too small to generate a reasonable estimate, which leads to the next subsection.

Project Heterogeneity

So far, $G_{\tau,k}(b)$ and $g_{\tau,k}(b)$ are auction-specific, but to obtain a large enough sample size for the estimation of the bids distribution, bids data from different auctions need to be pooled together. Following Marion (2007), this study allowed the distribution $F_{\tau,k}^{*}(c;\overline{s}_{\tau})$ to be dependent on project characteristics z_k such that $F_{\tau,k}^{*}(c;\overline{s}_{\tau}) = F_{\tau}^{*}(c;\overline{s}_{\tau}|z_k)$. Specifically, the engineer's estimate was the z_k used to control for project heterogeneity. Since this study assumed that there is no unobserved heterogeneity, the engineer's estimate fully captures the heterogeneity across contracts, and the firms draw project costs from the same distribution $F_{\tau}^{*}(c;\overline{s}_{\tau}|z_k)$ for each auction. Note that $\overline{s}_{\tau}$ is exogenous because it depends only on the number of potential bidders and the entry cost, both of which are exogenous. The variation of $\overline{s}_{\tau}$ in the data results in different truncation levels in $F_{\tau}^{*}(c;\overline{s}_{\tau}|z_k)$. Using the formula of conditional probability, this study first estimated the joint distributions and densities of the bid and engineer's estimate and then estimated $G_{\tau}(b|z) = \frac{G_{\tau}(b,z)}{F_Z(z)}$ and $g_{\tau}(b|z) = \frac{g_{\tau}(b,z)}{F_Z(z)}$. Project costs were then recovered by pooling bids data together:

$$b = c + \left((n_{\tau} - 1) \frac{g_{\tau}(b|z)}{1 - G_{\tau}(b|z)} + n_{-\tau} \frac{g_{-\tau}(b|z)}{1 - G_{-\tau}(b|z)} \right)^{-1}. \tag{12}$$

Estimation Method

Because an entrant's bid depends on the number of bidders of each type, bids data from auctions of different (n_f, n_n) come from different bid functions and $G_{\tau}(b|z)$, $g_{\tau}(b|z)$ need to be estimated separately for each (n_f, n_n) group of contracts. In the estimation, this study used auctions with n_f between one and four and n_n between one and three, which tend to have the largest sample sizes for bids. There were 496 such contracts in total.

Using the nonparametric method in Guerre et al. (2000) and following Marion (2007), this study estimated the distributions of bids and project costs using the standard kernel estimators. The results of this estimation method can be sensitive to the choice of bandwidth (Marion 2007). Following Marion (2007), to estimate the distributions of the recovered project costs, a log transformation was applied to mitigate the effects of skewness due to a few very high bids in the data. Rule of thumb was then used to select bandwidth. When estimating the conditional distributions and densities $G_\tau(b|z)$, $g_\tau(b|z)$, the cross validation maximum likelihood method was used to select bandwidth.

The entry threshold was estimated using the probability of entry among potential bidders of each type, which, in the data, is the proportion of potential bidders who entered the auction. Since the private signal was assumed to be uniformly and independently distributed, the probability a firm enters is equal to the entry threshold. Among the 496 contracts analyzed, in 194 contracts all the non-fringe potential bidders entered, and in 79 contracts, all the fringe potential bidders entered. In other words, the estimated entry threshold in these auctions is one. This means that the truncated distribution in Eq. 1 collapses to the full marginal distribution of project costs, since all potential bidders entered. This study exploited this feature of the sample and used these subsets of the sample to estimate the marginal distribution $F_{C_\tau}(c)$, an important model primitive of interest.

The selected distributions were then estimated using data from auctions with entry thresholds less than one. If the entry process was selective, the selected distribution was a truncated distribution of $F_{C_\tau}(c)$. Otherwise, a non-selective entry process implies that no matter what the entry threshold is, the distribution of the project costs among the entrants is the same as $F_{C_\tau}(c)$. Further, since a lower entry threshold results from lower expected profit and more competition, if the entry process is selective, the truncated distribution would be more skewed to the lower-cost end.

Entry Cost

The entry cost, K_τ, was estimated as a percentage of the engineer's estimate by extending the entry cost estimation approach in Athey et al. (2011). The estimation of entry cost is essential for counterfactual analysis. Given the exogenous variables (N_f, N_n) and the engineer's estimate (which reflects project characteristics of the contract), K_τ can be estimated using Eqs. 5 and 6.

Following Athey et al. (2011), a parametric model was specified for the entry threshold:

$$\overline{s}_\tau(X,N) = \frac{\exp(\alpha X + \beta N)}{1 + \exp(\alpha X + \beta N)}, \tag{13}$$

where X is engineer's estimate, and N refers to (N_f, N_n). Equation 13 can be rewritten as

$$\ln\left(\frac{\overline{s}_\tau}{1-\overline{s}_\tau}\right) = \alpha X + \beta_1 N_n + \beta_2 N_f. \tag{14}$$

Parameters in Eq. 14 were estimated from the data. The entry threshold was then determined using Eq. 13 given the exogenous variables.

K_τ is essentially the expected profit of a potential bidder whose signal is equal to the entry threshold $\overline{s}_\tau$. Using Eq. 4, this expected profit was estimated through simulations. For each simulation, first, c_i was randomly drawn from the distribution $f_\tau(c|\overline{s}_\tau)$. Given each ($n_f$, n_n), c_i was then mapped to the respective $\Pi_\tau^{II}(c_i; n_f, n_n)$. Finally, given ($N_f$,

N_n), the expected profit of a potential bidder with project cost c_i was calculated by taking the expectation of $\Pi_\tau^{II}(c_i; n_f, n_n)$ across (n_f, n_n) using Eqs. 2 and 3. This simulation was repeated 5000 times. The average of the 5000 simulation results was taken, which should yield a consistent estimator of K_τ.

Since $f_\tau(c|\overline{s}_\tau)$ cannot be identified nonparametrically, the approximation below was used:

$$\widehat{f}_\tau\left(c|\overline{s}_\tau\right) = \int_{s_l}^{s_u} f_\tau(c|s)ds, \tag{15}$$

where s_l and s_u are some lower and upper bound for $\overline{s}_\tau$. When s_l and s_u get arbitrarily close to each other, Eq. 15 collapses to the true distribution. In practice, s_l, s_u were chosen to be as close to $\overline{s}_\tau$ as the data allows to ensure a large enough sample size. Equation 15 was estimated by pooling recovered project costs from auctions whose entry thresholds are between s_l and s_u. The appropriateness of this approach depends on how much $f_\tau(c|\, s)$ for different s differs from one another.

To map c_i to $\Pi_\tau^{II}(c_i; n_f, n_n)$ given each (n_f, n_n), Eq. 7 was rewritten:

$$\Pi_\tau^{II}(c_i; n_f, n_n) = (b_i - c_i) \cdot (1 - G_\tau(b_i))^{n_\tau - 1} \cdot (1 - G_{-\tau}(b_i))^{n_{-\tau}}. \tag{16}$$

In other words, a type-τ bidder's Stage 2 expected profit is its markup times its probability of winning, which is the probability that all its rivals bid higher than its bid. $G_\tau(b)$, $G_{-\tau}(b)$ were estimated from the bids data, and b_i was inferred in the following way. Within the range of commonly observed bids (0.2 to 3 times the engineer's estimate), 50,000 "hypothetical bids" were uniformly allocated. Following the method in Athey et al. (2011), Eq. 12 was used to infer the project cost corresponding to each hypothetical bid. This approach created a table mapping bids to project costs. Any given c_i could then be mapped back to b_i by using the closest project cost (and the corresponding bid) in this table as a proxy. Since the 50,000 hypothetical bids were densely allocated, this approach should yield a very close approximation.

A limitation of this entry cost estimation is that K_f, K_n were estimated using separate zero profit conditions as opposed to enforcing the other type's entry threshold $K_{-\tau}$ in the estimation of K_τ, as Eqs. 5 and 6 require. Separate zero profit conditions were used because data that satisfied both $\overline{s}_f$ and $\overline{s}_n$ was too thin. The same approach can be applied if there is a large enough sample that satisfies both $\overline{s}_f, \overline{s}_n$. For this reason, the estimated K_τ should be interpreted as the expected profit of a type-τ threshold firm competing in an average environment as opposed to an environment where firms of the other type enter according to $\overline{s}_{-\tau}$. How different these two are, again, depends on how different the distribution $F_{-\tau}(c|\overline{s}_{-\tau})$ is from $F_{C_{-\tau}}(c)$.

Results

Comparing Density Estimates of the Full and Selected Distributions

Figures 1 and 2 compare the full marginal distribution of project costs (of all firms) with the selected distribution (of the entrants) for each type of firms. The results shown were estimated conditional on the median engineer's estimate. The selected distribution

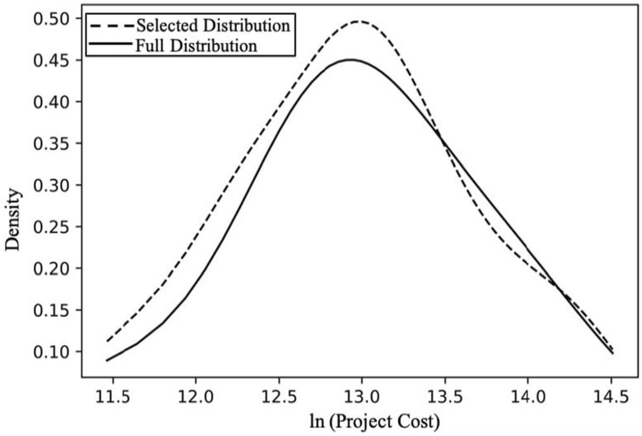

Fig. 1 Fringe Firms' Project Cost Density Comparison. Source: Own calculations using 1999–2005 data from Bajari et al. (2014)

was estimated from auctions with $\overline{s}_f$ from 0.2 to 0.5 for fringe firms and $\overline{s}_n$ less than 0.4 for non-fringe firms. Since auctions of different entry thresholds were combined in the estimation, the result should be interpreted not as one $F^*_\tau(c;\overline{s}_\tau)$ but as various truncated distributions pooled together. In Fig. 1, compared to the full distribution, the selected distribution resembles a truncated distribution skewed to the left, which is consistent with the theoretical predication of selective entry. While Fig. 1 shows evidence favoring selective entry among the fringe firms, such evidence was not found for the non-fringe firms (Fig. 2).

Entry Cost

K_τ was estimated for a representative auction with engineer's estimate $952,000 (the median in the sample), two non-fringe and four fringe potential bidders (the most

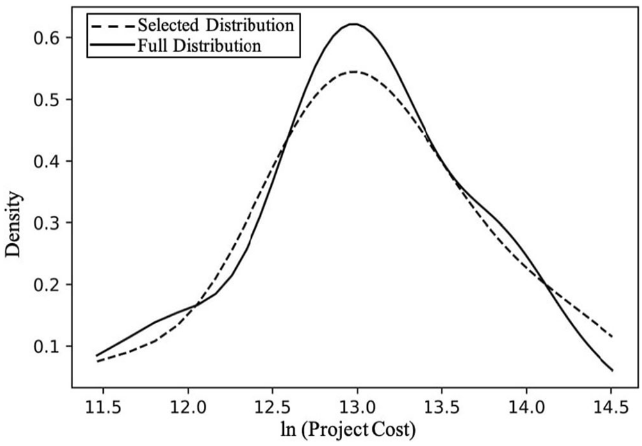

Fig. 2 Non-fringe Firms' Project Cost Density Comparison. Source: Own calculations using 1999–2005 data from Bajari et al. (2014)

common number for each type). Estimated entry thresholds were 0.562 and 0.857 for fringe and non-fringe type, respectively. The results showed that the entry cost is 1.66% of the engineer's estimate for the fringe type and 2.57% for the non-fringe type. Compared with the entry costs estimated by Krasnokutskaya and Seim (2011) under non-selective entry, these results are slightly lower. This difference coincides with the findings in Li and Zheng (2012): entry costs estimated under non-selective entry tend to be higher than those estimated under selective entry. Thus, if unreasonably high entry costs are estimated under a non-selective entry model, the researcher may want to check whether the non-selective entry assumption is appropriate.

Conclusion

This study empirically investigated whether the entry process is selective in the Caltrans procurement auctions data. The results favor selective entry among the fringe firms. Given that prior literature showed a weaker player's probability of winning increases with the degree of selection (Sweeting and Bhattacharya 2015), this finding suggests that SB might have a higher likelihood of winning than what prior empirical studies reported using non-selective entry models. Thus, in order to achieve the policy objective of awarding 25% of total state-funded contract dollars to SB, Caltrans may only need a bid discount level lower than that derived from non-selective entry models in prior literature.

The full and selected project cost distributions were estimated using a nonparametric method under a model allowing endogenous and potentially selective entry. Entry costs were also estimated. The key limitation of the estimation method used in this paper was that the methodology did not account for heterogeneity in project characteristics other than the engineer's estimate. Project heterogeneity that is unaccounted for may lead to different underlying project cost distributions for different (n_f, n_n) bidding pools, which would make combining data from various bidding pools in estimation inappropriate. This methodology also did not provide an estimate of the domain of the bids and project costs.

This paper ties the existing theoretical literature on selective entry and auction design to empirical evaluations of bid preference programs. By comparing the full and selected project cost distributions to test whether the data support selective entry, this paper adds a new method to the existing literature on empirical testing of different entry models. Whereas empirical auction studies in the literature estimate models with partially selective entry, the estimation methodology used in this study provides a first example of empirically estimating such models nonparametrically.

Understanding how selective entry alters the empirical assessment of the effects of bid preference programs on procurement costs and contract allocation is an important direction for future research. With the estimated model parameters, it would be interesting to use numerical analysis (Sweeting and Bhattacharya 2015) to solve for the government's cost minimizing bid discount level that also satisfies its distributional goal.

Acknowledgments I thank my thesis advisor Professor Gaurab Aryal for his invaluable guidance and mentorship, as well as Professor Amalia Miller and Professor John Pepper for their support. I also thank Jiafeng Wu, Charles Moens, and participants of the Distinguished Majors Seminar at the University of Virginia.

References

Athey, S., Levin, J., & Seira, E. (2011). Comparing open and sealed bid auctions: Evidence from timber auctions. *The Quarterly Journal of Economics, 126*(1), 207–257. https://doi.org/10.1093/qje/qjq001.

Bajari, P., Houghton, S., & Tadelis, S. (2014). Bidding for incomplete contracts: An empirical analysis of adaptation costs. *American Economic Review, 104*(4), 1288–1319. https://doi.org/10.1257/aer.104.4.1288.

Bhattacharya, V., Roberts, J. W., & Sweeting, A. (2014). Regulating bidder participation in auctions. *RAND Journal of Econonmics, 45*(4), 675–704. https://doi.org/10.3386/w19352.

Caltrans. (2020a). Civil Rights. Available at: https://dot.ca.gov/programs/civil-rights. Accessed 1 May 2020.

Caltrans. (2020b). Bid Results. Available at: https://dot.ca.gov/programs/procurement-and-contracts/bid-results. Accessed 1 May 2020.

Gentry, M., & Li, T. (2014). Identification in auctions with selective entry. *Econometrica, 82*(1), 315–344. https://doi.org/10.3982/ECTA10293.

Guerre, E., Perrigne, I., & Vuong, Q. (2000). Optimal nonparametric estimation of first-price auctions. *Econometrica, 68*(3), 525–574. https://doi.org/10.1111/1468-0262.00123.

Heckman, J. (1976). The common structure of statistical models of truncation, sample selection, and limited dependent variables and a simple estimator for such models. *Annals of Economic and Social Measurement, 5*(4), 475–492.

Krasnokutskaya, E., & Seim, K. (2011). Bid preference programs and participation in highway procurement auctions. *American Economic Review, 101*(6), 2653–2686. https://doi.org/10.1257/aer.101.6.2653.

Li, T., & Zheng, X. (2009). Entry and competition effects in first-price auctions: Theory and evidence from procurement auctions. *Review of Economic Studies, 76*(4), 1397–1429. https://doi.org/10.1111/j.1467-937X.2009.00558.x.

Li, T., & Zheng, X. (2012). Information acquisition and/or bid preparation: A structural analysis of entry and bidding in timber sale auctions. *Journal of Econometrics, 168*(1), 29–46. https://doi.org/10.1016/j.jeconom.2011.09.004.

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. https://doi.org/10.1016/j.jpubeco.2006.12.005.

Roberts, J. W. & Sweeting, A. (2010). Bailouts and the preservation of competition. *NBER Working Paper No. 16650*. https://doi.org/10.3386/w16650.

Roberts, J. W. & Sweeting, A. (2011). *Competition versus auction design*. Working paper, Duke University.

Roberts, J. W., & Sweeting, A. (2013). When should sellers use auctions? *American Economic Review, 103*(5), 1830–1861. https://doi.org/10.1257/aer.103.5.1830.

Roberts, J. W., & Sweeting, A. (2016). Bailouts and the preservation of competition: The case of the Federal Timber Contract Payment Modification act. *American Economic Journal: Microeconomics, 8*(3), 257–288. https://doi.org/10.1257/mic.20150070.

Sweeting, A., & Bhattacharya, V. (2015). Selective entry and auction design. *International Journal of Industrial Organization, 43*(C), 189–207. https://doi.org/10.1016/j.ijindorg.2015.03.004.