

# Bid Preference Programs and Participation in Highway Procurement Auctions

January 2010

## Abstract

We use data from highway procurement auctions subject to California's Small Business Preference program to study the effect of bid preferences on auction outcomes. Our analysis is based on an estimated model of firms' bidding and participation decisions, which allows us to evaluate the effects of current and alternative policy designs. We show that incorporating participation responses significantly alters the assessment of preferential treatment policies.

**Keywords:** Bid preference programs, auction participation, asymmetric bidders.

**JEL Classification:** D44, L10, H11, H57

# 1 Introduction

Public-sector procurement accounts for over 10% of U.S. GDP. Across levels of government, preferential treatment programs are extensively used in procurement auctions. For example, in 2006, the federal government awarded 20% of its procurement dollars to favored firms.<sup>1</sup> One commonly used preference mechanism, a bid discount or credit, improves the bids of favored firms by a pre-established rate when determining the winner, but uses the actual amount of the winner's bid in the contract.<sup>2</sup> Prominent examples include a 25% bid credit granted to small firms in FCC spectrum auctions and a 50% bid penalty added to foreign bids on defense contracts.<sup>3</sup> The aim of this paper is to improve our understanding of the effects of such preference programs on the government's cost of procurement and the distribution of profits between participants, as well as to provide an assessment of the likely magnitudes of these effects in practice. We do so empirically in the context of the California Small Business Preference program that grants small firms a 5% bid discount.<sup>4</sup>

The stated goal of most preference programs is to facilitate the integration of favored participants into the market place. These are often groups historically discriminated against, or groups considered disadvantaged due to entry barriers, or both. They are also often considered to be less cost efficient. As preference programs result in such high-cost companies performing a larger share of work, one may expect the cost of procurement to increase. At the same time, however, these programs also provide incentives to non-favored firms to bid more aggressively against the strengthened favored group, which mitigates the upward pressure on the cost of procurement. For some discount levels, this last effect is sufficiently strong for the cost of procurement to actually decrease (McAfee and McMillan (1989) and Corns and Schotter (1999) show this theoretically and in experiments, respectively, for assumed numbers of bidders and cost distributions).

The key insight of this paper is that there is a third effect neglected in the literature. Bid preference programs have potentially strong effects on firms' incentives to participate in

---

<sup>1</sup>See the Federal Procurement Report 2007, available at <https://www.fpds.gov/>.

<sup>2</sup>With a 10% bid discount, for example, a bid by a favored firm of \$440,000 is treated as a bid of \$400,000 in comparing it to the remaining, non-favored, firms' bids. If the favored firm wins, its payment is the original amount of the bid, or \$440,000.

<sup>3</sup>See "Implementation of the Commercial Spectrum Enhancement Act and Modernization of the Commission's Competitive Bidding Rules and Procedures," WT Docket No. 05-211, Second Report and Order and Second Further Notice of Proposed Rulemaking, 21 FCC Rcd 4753, 4766 par 36 (2006); and the Department of Defense's "Defense Federal Acquisition Regulation Supplement," Part 225: Foreign Acquisition (2008), available at <http://www.acq.osd.mil/>.

<sup>4</sup>Other empirical studies of preference programs include Marion (2007, 2009) who finds two specific preference programs to be costly to governments; Denes (1997) who provides evidence of cost decreases in some set-aside auctions for dredging work; and Ayres and Cramton (1996) who argue that preference programs yielded significant revenue increases in a small sample of FCC spectrum auctions. These papers use descriptive methods, which allow them to measure the effects of the current programs, but do not permit an evaluation of alternative program designs. Decarolis (2009) analyzes average price auctions that could be interpreted as an extreme form of preference policy where the bid closest to the average wins and the high bid is eliminated.

an auction. We show that accounting for a response in participation behavior significantly alters the assessment of the preference program’s cost to the government and its distributional effects. While it continues to be possible to use bid discounts to lower the cost of procurement as in McAfee and McMillan (1989), both the cost-minimizing level of the discount and the group receiving the discount may change when participation effects are taken into account. The currently accepted practice of evaluating bid preference programs holding participation fixed can yield very misleading results.

The theoretical literature suggests that the magnitudes of the program’s effects crucially depend on the degree of cost asymmetries between favored and other bidders. We thus base our analysis on empirically relevant distributions of firm costs recovered from data on highway procurement auctions that were awarded under a bid preference program. We use a model of firms’ participation and bidding decisions in the presence of a bid discount.<sup>5</sup> The firm’s decision of which bid to submit reflects its private information about its cost of completing the project, which we term “project cost”, and the distributions of its competitors’ project costs. The participation decision instead is based on a comparison of the cost of preparing the bid, or entry cost, to the expected profit from participation. Only firms with entry costs below the expected profit ultimately submit a bid in the auction. We use this model to uncover the underlying distributions of firms’ entry and project costs consistent with observed choices.

The nature and importance of our findings can be seen from Figure 1 that plots changes in the government’s cost of procurement relative to no discrimination at different levels of the bid discount for a typical project in our data.<sup>6</sup> We contrast the cost of procurement implied by a model that does not allow firms to respond to the discount in their participation behavior with one where participation adjusts endogenously. Several patterns emerge:

1. Under fixed participation, the cost of procurement varies only by a limited amount as the discount changes from 50% to large bidders (the leftmost point in the figure) to 50% to small bidders (the rightmost point). The cost of procurement exhibits significantly more variation when we take participation effects into account.
2. The implications for policy design differ significantly in the two cases. To minimize the cost of procurement, the model with fixed participation prescribes a discount of approximately 15% to small bidders. Relaxing the assumption of fixed participation suggests that offering such a discount to small bidders would actually increase the cost of procurement. Instead, a discount of 50% should be offered to *large* bidders to achieve substantial cost savings.
3. California’s Small Business Preference program aims to allocate 25% of procurement dollars

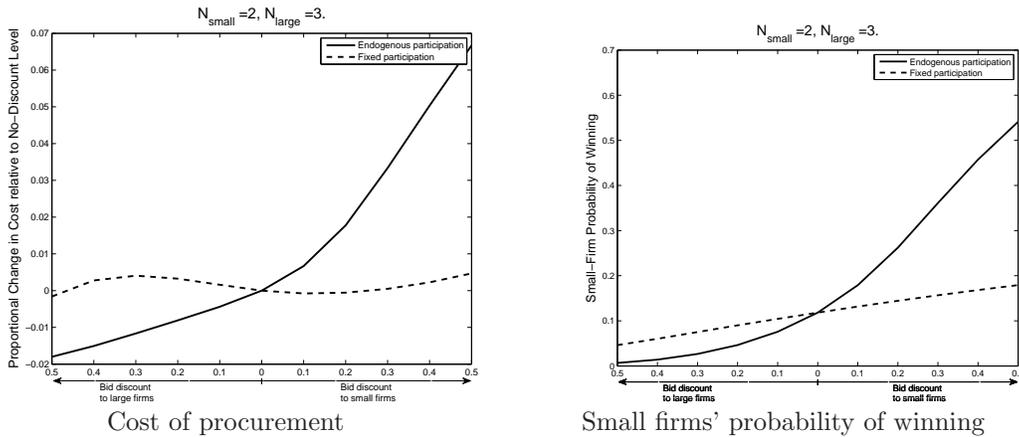
---

<sup>5</sup>Our analysis also contributes to a small, but growing literature that empirically studies the decision to participate in auctions. Athey, Levin and Seira (2008), Bajari and Hortacsu (2003), Li (2005), and Li and Zheng (2009) represent recent contributions to this literature.

<sup>6</sup>The project’s cost distributions are representative of approximately 30% of projects. The remaining projects are discussed in the main body of the paper.

to small firms, which we refer to as the program’s “allocative goal”. The fixed participation model implies that the small-firm discount required to achieve this goal is equal to 50% for this particular project. This model predicts that such a discount yields a 0.6% increase in procurement cost. However, a model that takes participation adjustments into account would recognize that this substantial discount deters large-firm participation and, therefore, that the true cost increase would be 7%. Additionally, preferential treatment increases small-firm participation and in turn the group’s probability of winning, hence, a bid discount of only approximately 20% is sufficient to achieve the allocative goal, raising the government’s cost by 2%.

Figure 1: Cost of Procurement and Probability of Winning under Fixed and Endogenous Participation, Sample Project



This example is based on a particular, albeit common, type of project in our data. An aggregate evaluation of California’s preference policy needs to take into account heterogeneity in project characteristics and the competitive environment, which introduces heterogeneity in the effectiveness of a bid discount across projects. Our empirical results suggest significant differences in the degree of cost asymmetries between large and small firms across projects. For an important subset of projects in our data, we recover cost distributions for large and small firms that are very similar. As a result, small-firm participation and winning rates for these projects are high even in the absence of a bid discount. Because of the particular mix of projects, the aggregate cost of procurement at a discount level that awards 25% of procurement dollars to small firms is only 1.2% higher than the aggregate cost under no preferential treatment. It is important to note, however, that this result is specific to the California market. In other markets where the composition of projects is different, the cost of bid preference programs may be very different.

For California’s current program, which uses a relatively low discount level of 5%, we

find that the cost of procurement is within 1% of the cost of procurement in the absence of discrimination. However, the program induces substantial changes in small and large firms' participation and probabilities of winning. It results in a redistribution of 10 to 18% of profits from large to small firms for typical projects that differ in type of work, location, and size. At the same time the program does not achieve its goal of allocating 25% of procurement dollars to small firms.

Interestingly, we find that an alternative preference mechanism that relies on lump-sum entry subsidies and/or taxes is more cost effective than a discount program. An appropriately chosen entry tax, for example, lowers the cost to the government significantly more than the cost-minimizing bid discount by extracting bidders' full expected surplus. Such a tax does not, however, achieve the State's allocative goal. We show instead that a combination of a subsidy to small firms and a modest tax on large firms can be used to satisfy California's allocative goal at important cost savings relative to a bid discount and equivalent award levels. An entry tax or subsidy, by affecting firms' participation margins regardless of their ultimate cost of completing the project, avoids a distortion associated with bid discounts that grant higher absolute gains to bidders with high cost draws. It is through this channel that lower costs of procurement can be realized.

The paper proceeds as follows. Section 2 provides a brief overview of the highway procurement market in California and the details of the Small Business Preference program. Section 3 outlines the model of firms' joint participation and bidding decisions. Section 4 describes our estimation methodology, the results of which are in Section 5. Section 6 contains an analysis of the current and alternative programs. Section 7 concludes.

## 2 California's Highway Procurement Market

In this section, we describe the California highway procurement market and our data. We focus on highway and street maintenance projects auctioned by the California Department of Transportation ('Caltrans') between January 2002 and December 2005. California's Small Business Preference program is implemented on state-funded projects. During the sample period, Caltrans advertised 869 state-funded projects, of which complete data are available for 697 projects.<sup>7</sup> The data include information on project characteristics, the set of companies that purchased detailed project specifications and their small business status, the set of actual bidders, their bids, and finally, the identity of the winning bidder.

**Letting Process.** Caltrans advertises projects three to ten weeks prior to the bidding date. The project advertisement usually contains only limited information, such as type of work, location, and completion time. Interested contractors must purchase detailed project plans from

---

<sup>7</sup>Caltrans did not preserve lists of companies that purchased bid documents for some projects.

Caltrans' project counter at least one week before the bid opening date. Only those firms that purchased project plans (plan holders) may submit a bid on the project. Our data suggest that purchasing a plan signals interest in bidding; we observe, for example, that in their plan purchases, companies focus on similar projects based on administrative district location and type of work. We therefore assume that the group of potential bidders on a given project coincides with the group of plan holders. The list of companies that purchased plans for a given project is posted on Caltrans' website. Therefore, potential bidders are known to each other at the time when they prepare their bids.

To bid on a project, a company must submit by the bid opening date completed bid documents, which specify the bid amount, the list of subcontractors, their fees, and their tasks. The preparation of bid documents requires time and effort and is, therefore, costly. We treat such bid preparation costs as entry costs in our model below.

During the bid preparation process, companies engage in extensive negotiations with subcontractors. It is likely that participants learn about other companies preparing bids for the same project from subcontractors. Anecdotal evidence confirms that such information leakage occurs. Discussions with industry insiders also suggest that prime contractors are careful not to reveal other information about their bid proposal, such as the cost of other contract items, quotes received from other subcontractors, etc. to potential subcontractors. Price negotiations also typically continue up until the bid submission deadline, limiting the subcontractor's ability to convey any price information to competitors.

As evident from the bid documents, the sets of subcontractors often overlap across companies submitting bids for the same auction. Common subcontractor use can potentially induce affiliation into bidders' costs, i.e. a correlation in their costs in excess of any correlation introduced by factors known to bidders. We investigated empirically how important such subcontractor induced correlation is in explaining bid levels. Using price data and subcontractor information at the level of the individual contract item for a subsample of our bid documents, we find that the identity of the subcontractor explains approximately 6% of the average item price across items and contractors. This combined with the fact that the total value of items for which common subcontractors are used constitutes at most 5% of the overall bid suggests that the extent of affiliation due to common subcontractor use is low.

**Preference Program.** The Small Business Preference program sets a goal of allocating 25% of state procurement dollars to small firms. The program is implemented using a first-price sealed-bid auction mechanism. It grants small firms a bid discount equal to 5% of the low non-favored bid, reducing their bids for comparison purposes only when determining the winner. The winner is then paid the full amount of his bid.

To qualify for the discount, a company has to satisfy three conditions. It has to be independently owned and operated; have fewer than 100 employees; and have average annual gross

receipts limited to \$10 million over the previous three tax years.<sup>8</sup> A common concern with preference programs is the potential for abuse and manipulation. The structure of the procurement market renders such abuse more difficult than in other markets. Strict subcontracting limits are in place and Caltrans monitors projects to ensure that the chosen contractor adheres to these limits. In addition, small contractors' competitors have a vested interest to ensure that the small-business status is used only when applicable. While the instance of abuse is rare, the State also actively prosecutes and penalizes abusers, imposing both monetary penalties and withdrawing the right to participate in future procurement auctions.

We obtained quarterly information on the certification status of companies in our data set from the Department of General Services. In our sample, out of 672 companies that bid on at least one project, 269, or 40%, were certified as small businesses. Caltrans awarded 39.02% of contracts to qualified small businesses. The total value of these contracts accounted for only 15.45% of total procurement dollars, however. Most of the projects allocated to small firms are therefore small. It also means that Caltrans does not meet the program's allocative goal. The bid preference altered the identity of the winning bidder in only 5% of projects.

### 3 Model of Firms' Participation and Bidding Decisions

This section develops a model of firms' participation and bidding decisions that forms the basis for our empirical analysis below. We assume that a total of  $N$  potential contractors express interest in a single standalone project offered for bid. Bidder  $i$ 's decisions reflect two separate costs; entry costs of preparing a bid, denoted by  $d_i$ , and costs of completing the contract (project costs), denoted as  $c_i$ .

We incorporate a preference rule similar to the one used in California into our model. For the purpose of comparison, bids of favored firms are reduced by an amount equal to  $\delta$  percent of the lowest non-favored bid. A favored firm is awarded the project if its reduced bid is below the lowest non-favored bid. For a given lowest non-favored bid of  $b_l$ , a favored firm thus wins the project if its bid is lower than  $(1 + \delta)b_l$ . It receives the full amount of its bid as payment. A preference program thus introduces an asymmetry into the payoffs of favored and other firms. In our analysis we also allow for the possibility that favored (group 1) and other (group 2) firms differ systematically in their costs of preparing bids,  $G_D^k$ , and of completing the project,  $F_C^k$ . Here  $k(i)$  denotes group affiliation of bidder  $i$ . We assume that project and bid preparation costs are private information of each firm and are distributed independently across all firms and identically within group.

---

<sup>8</sup>Such revenue restrictions could affect small firms' entry behavior. For example, a company may decide not to bid on a large project if winning this project brings it over or very close to the revenue threshold. In our data, however, 99% of small firms have yearly revenue below \$5.4m, relative to a large project's typical size of about \$1m. Therefore, in most cases winning one additional large project does not impact the small-firm status of a company and we do not model such dynamic concerns about qualifying for small-firm status.

Similar to other work on auction participation (e.g., Samuelson (1985), Levin and Smith (1994)), we model a potential bidder’s decision as a two-stage process. In the first stage, each potential bidder decides whether to participate in the auction. In the second stage, actual bidders prepare and submit their bids. When deciding over participation, potential bidder  $i$  of group  $k(i)$  knows his own cost of entry,  $d_i$ , the distributions of project and entry costs,  $F_C^k$  and  $G_D^k$ ,  $k = 1, 2$ , and the numbers of potential bidders by group,  $N_{k(i)}, N_{-k(i)}$ . Only firms with an entry cost below the expected profit from participation choose to enter the auction. Firms that decided to enter pay bid preparation costs, become actual bidders, and submit bids. By incurring bid preparation costs, a bidder learns his costs of completing the contract,  $c_i$ , and the numbers of his actual competitors by group,  $(n_{k(i)} - 1, n_{-k(i)})$ .

Our model of entry resembles the setup in Levin and Smith (1994) by relying on two assumptions: (a) a potential bidder does not observe his project cost realization at the time of his participation decision, but learns it through the investment of bid preparation costs; (b) bidders know the numbers of their competitors when they decide on a bid level.<sup>9</sup> An alternative to assumption (a) is presented in Samuelson (1985), where project costs are known at the time of entry. This alternative informational environment finds less support than the assumption we use in empirical tests of entry models.<sup>10</sup> We also carefully considered the applicability of assumption (b) to our setting. We experimented with an alternative informational assumption that firms do not have knowledge of the numbers of bidders throughout the entire bidding process. This model generally produced mark-ups that were significantly higher than typical highway construction mark-ups. Assumptions (a) and (b) greatly facilitate the computation of participation and bidding strategies, in particular given our context of asymmetric auctions where we have to find equilibrium bidding strategies numerically, as we discuss below.<sup>11</sup> This is what allows us to conduct an extensive counterfactual analysis, which would have to be significantly curtailed under either of the two alternative informational environments discussed here.

### 3.1 Characterization of Equilibrium in the Bidding Stage

We begin with an analysis of the bidding stage and then use the results to analyze the participation stage. We focus on group-symmetric equilibria where bidders of group  $k$  follow the same bidding strategy,  $\beta_k(\cdot)$ , mapping project cost,  $c_i$ , into a bid  $b_i$ ,  $\beta_k(\cdot) : [c, \bar{c}] \rightarrow [\underline{b}_k, \bar{b}_k]$ . Due to the bid-preference program, a bidder  $i$  of group  $k$  wins the project if his bid  $b_i$  is below all competing

<sup>9</sup>Athey et al. (2008) also rely on these assumptions.

<sup>10</sup>In the context of symmetric auctions, Marmer, Shneyerov and Xu (2007) and Li and Zheng (2009) perform tests of alternative models of entry using different methodologies. Both sets of authors find more statistical support for a two-stage entry model where firms are initially uninformed or only partially informed about their project costs and pay an entry cost to learn their actual realization than an alternative where project costs are known at the time of entry.

<sup>11</sup>Assumption (a) also simplifies the empirical implementation of the model. The lack of selection on project costs allows us to recover their full (untruncated) distribution in estimation and we are able to more easily incorporate the effect of unobserved project characteristics on firms’ bidding behavior.

bids adjusted by the bid discount  $\delta$  where applicable. Firm  $i$  with project cost  $c_i$  and group membership  $k(i)$  chooses bid  $b_i$  to maximize expected profit conditional on participating:

$$\begin{aligned}\pi_i(c_i) &= (b_i - c_i) \Pr(b_i \leq b_l, \forall l : k(l) = k(i)) \Pr(b_i \leq (1 + \delta)^{1-2I(k=2)} b_l, \forall l : k(l) \neq k(i)) \quad (1) \\ &= (b_i - c_i) \left(1 - F_C^k[\beta_k^{-1}(b_i)]\right)^{n_k-1} \left(1 - F_C^{-k}[\beta_{-k}^{-1}((1 + \delta)^{1-2I(k=2)} b_i)]\right)^{n-k}\end{aligned}$$

where  $I(k = 2)$  is an indicator variable that equals one if firm  $i$  belongs to group 2. The first-order condition of the firm's bidding problem is:

$$\begin{aligned}\frac{1}{b_i - c_i} &= \frac{(n_{k(i)} - 1) f_C^{k(i)}[\beta_{k(i)}^{-1}(b_i)]}{\left(1 - F_C^{k(i)}[\beta_{k(i)}^{-1}(b_i)]\right)} \frac{\partial \beta_{k(i)}^{-1}}{\partial b_i} \quad (2) \\ &+ \frac{n_{-k(i)} (1 + \delta)^{1-2I(k(i)=2)} f_C^{-k(i)}[\beta_{-k(i)}^{-1}((1 + \delta)^{1-2I(k(i)=2)} b_i)]}{\left(1 - F_C^{-k(i)}[\beta_{-k(i)}^{-1}((1 + \delta)^{1-2I(k(i)=2)} b_i)]\right)} \frac{\partial \beta_{-k(i)}^{-1}}{\partial b_i}\end{aligned}$$

The preference program introduces two interesting features into the equilibrium, reflecting the increased competitiveness of favored bidders. First, a single favored bidder with  $c_i = \bar{c}$  finds it optimal to bid above his cost when bidding against several non-favored bidders since the bid discount sufficiently lowers his effective bid to result in a non-zero probability of winning the project.<sup>12</sup> In contrast, with multiple favored bidders, competitive pressure reduces the upper boundary bid to cost. Second, since the highest effective bid submitted by a favored bidder is given by  $\frac{\bar{b}_1}{1+\delta}$ , non-favored bidders with cost  $c_i \in [\frac{\bar{b}_1}{1+\delta}, \bar{c})$  can never win an auction where a small bidder is present and earn positive profit.

The behavior of bidders with boundary cost draws can be summarized as follows.

1. *Right-boundary condition.* Favored bidders with cost level  $\bar{c}$  bid  $\bar{b}_1 = \bar{c}$  if  $n_1 > 1$ . If  $n_1 = 1$ ,  $\bar{b}_1$  is the bid level that maximizes

$$\pi_i = (\bar{b}_1 - \bar{c}) \left(1 - F_2\left(\frac{\bar{b}_1}{(1 + \delta)}\right)\right)^{n_2}. \quad (3)$$

Non-favored bidders with  $c_2 \in [\frac{\bar{b}_1}{1+\delta}, \bar{c})$  have a zero probability of winning and, therefore, bid their cost.

2. *Left-boundary condition.* There exists a bid level  $\underline{b}_1$  such that for all favored firms,  $\beta_1(\underline{c}) =$

---

<sup>12</sup>Note that consistent with Caltrans policy, we do not impose a reserve price. If only a single bidder chose to enter the auction, there are thus no constraints on his bid. We follow Li and Zheng (2009) and assume that in such instances, the government steps in as a second bidder, drawing its project cost from the non-favored cost distribution. This approximates the competitive pressure that Caltrans imposes in such instances through the right to reject a bid and re-scope a project. Since our data do not contain projects with only one bidder, this assumption is only relevant when computing the expected profit from entry by averaging over all possible bidder combinations.

$\underline{b}_1$ . For all non-favored bidders,  $\beta_2(\underline{c}) = \underline{b}_2 = \frac{\underline{b}_1}{(1+\delta)}$ .

The proof of these properties follows the standard reasoning for boundary conditions in first-price auctions. Theorem 2.1 in Reny and Zamir (2004) establishes the existence and uniqueness of the bidding equilibrium in this environment.

### 3.2 Characterization of Equilibrium in the Participation Stage

At the participation stage, firms compare the ex-ante expected profit conditional on entry to their entry cost  $d_i$ . Firms with entry costs below their expected profit decide to incur the entry fee to learn about their cost of completing the project. Ex-ante expected profit from participating is given by

$$\bar{\pi}_k(p_1, p_2) = \sum_{n_k-1, n_{-k} \subset N_k-1, N_{-k}} \left( \int_{\underline{c}}^{\bar{c}} \pi_k(c; n_k-1, n_{-k}) dF_C^k(c) \right) \Pr(n_k-1, n_{-k} | N_k, N_{-k}) \quad (4)$$

where  $\Pr(n_k-1, n_{-k} | N_k, N_{-k})$  is the probability of observing  $(n_k-1)$  competitors of the firm's own group and  $n_{-k}$  competitors of the opposite group, given numbers of potential entrants of  $N_k$  and  $N_{-k}$ .  $\pi_k(c; n_k-1, n_{-k})$  is the expected equilibrium profit of a bidder from group  $k$  with cost realization  $c$ . It reflects that at the participation stage, the firm is uncertain about both its own project cost and the competitive environment it will face upon entry. As a result, the expected profit differs only by group  $k$ , but not by firm  $i$ . The firms assess the probability that there will be  $n_k-1$  and  $n_{-k}$  competitors in the auction as

$$\Pr(n_k-1, n_{-k} | N_k, N_{-k}) = C_{N_k-1}^{n_k-1} C_{N_{-k}}^{n_{-k}} (p_k)^{n_k-1} (1-p_k)^{N_k-n_k-2} (p_{-k})^{n_{-k}} (1-p_{-k})^{N_{-k}-n_{-k}}, \quad (5)$$

where  $C_N^n$  denotes the binomial coefficient of choosing  $n$  firms out of  $N$  potential bidders.

The participation decision is described by group-specific entry cost thresholds,  $D_k$ , such that only firms with entry costs below their group's threshold participate in the auction. They are defined by a zero-profit rule so that  $D_1(p_1, p_2) = \bar{\pi}_1(p_1, p_2)$  and  $D_2(p_1, p_2) = \bar{\pi}_2(p_1, p_2)$ . In equilibrium, bidders' beliefs are correct and the equilibrium entry probabilities solve the system of equations

$$\begin{aligned} p_1 &= G_1 [D_1(p_1, p_2)] \\ p_2 &= G_2 [D_2(p_1, p_2)]. \end{aligned} \quad (6)$$

Brouwer's Fixed Point Theorem guarantees that the group-specific equilibrium of this game exists. In general, the entry equilibrium is not unique. There may be multiple threshold pairs that solve Equation (6). These equilibria are observationally equivalent in terms of submitted bids and differ only in entry probabilities. We verify the uniqueness of the equilibrium

entry probabilities numerically within the estimation routine.<sup>13</sup>

## 4 Estimation

The theoretical model describes group-specific participation and bidding strategies that map firms' project and participation costs and their respective distributions into observed bids and participation behavior. This section outlines the estimation methodology we use to recover parameters of the underlying distributions of entry and project costs from available data. We use a two-step estimation approach. In the first step, parameters of the bid distribution and the distribution of entry costs are estimated without imposing the full set of equilibrium restrictions. In the second step, the distribution of project costs is recovered from the equilibrium bidding first-order conditions following the procedure described in Guerre, Perrigne and Vuong (2000).<sup>14</sup>

### 4.1 Empirical Model

We assume that at announcement, a project is characterized by  $(x_j, z_j, u_j, N_{1j}, N_{2j})$ . Here  $x_j$  and  $z_j$  denote potentially overlapping project characteristics observable to the researcher that affect the distributions of project and entry costs, respectively. There may also exist other project attributes that impact firms' bidding and participation behavior that are not present in the data. These factors are summarized by the variable  $u_j$ . As in Krasnokutskaya (2009a), we assume that bidders' project costs for project  $j$  are given by  $c_{ij} = \tilde{c}_{ij}u_j$ . Here,  $\tilde{c}_{ij}$  is a firm-specific cost component that is private information of firm  $i$ , while  $u_j$  represents a portion of project  $j$ 's cost that is known to all bidders, but is unobserved to the researcher, i.e. unobserved project heterogeneity. The distribution of the firm-specific cost component for group- $k$  firms is given by  $F_c^k(\cdot|x_j)$ , while the distribution of unobserved project heterogeneity is given by  $H(\cdot)$ . We further assume that firms observe the realization of the unobserved project characteristic

---

<sup>13</sup>The equilibrium in the bidding stage results in non-favored bidders with  $c_2 \in [\frac{\bar{b}_1}{1+\delta}, \bar{c})$  having a zero probability of winning. Such firms may decide to drop out of the auction after learning their costs. In this case, Equation (6) should be adjusted to  $p_k = G_k [D_k(p_1, p_2 p^{nb})]$ ,  $k = 1, 2$ , where  $p^{nb}$  denotes the probability of non-favored bidders leaving the auction after learning their project cost realization, which we maintain in estimation. Our estimation procedure accounts for the resulting truncation. We are able to recover the full distribution of large-firm project costs because our data set contains auctions that did not attract any small bidders. The probabilities  $p_2 p^{nb}$  are recovered from ratios of cumulative distribution functions of project costs for projects with  $n_1 = \{1, 2\}$  and those with  $n_1 = 0$  for interior cost levels.

<sup>14</sup>Jofre-Bonet and Pesendorfer (2003) and Athey et al. (2008) use similar estimation methodologies. A standard procedure of estimating the distribution of project costs directly from the data poses severe computational challenges for models with asymmetric bidders. The computational burden is high because these models typically do not yield closed-form solutions for firms' bidding strategies, which are instead found numerically for every parameter guess and every project. A disadvantage of such indirect approaches is that they impose a parametric assumption on the bid distribution, which is not a primitive of the underlying model. To minimize any resulting misspecification bias, we use a flexibly specified bid distribution, controlling for a large number of project characteristics and time trends.

prior to making their entry decisions. It, therefore, affects both firms' participation and bidding behavior.

The firm-specific project cost components,  $\tilde{c}_{ij}$ , are mutually independent conditional on project characteristics,  $x_j$  and  $u_j$ , and are independent of the unobserved auction heterogeneity component,  $u_j$ :

$$F_{\tilde{c}|x}(\tilde{c}_1, \dots, \tilde{c}_{(N_1+N_2)}|x_j, u_j) = F_{\tilde{c}|x}(\tilde{c}_1, \dots, \tilde{c}_{(N_1+N_2)}|x_j) = \prod_{i=1}^{N_1+N_2} F_{\tilde{c}}^{k(i)}(\tilde{c}_i|x_j). \quad (7)$$

The unobserved heterogeneity component,  $u_j$ , is independent of project characteristics  $x_j$  and  $z_j$  and of the numbers of potential entrants,  $N_{1j}$  and  $N_{2j}$ , i.e.  $H(\cdot|x_j, z_j, N_{1j}, N_{2j}) = H(\cdot)$ .

Since we assume that bidders observe the numbers of their actual competitors when preparing their bid, firm  $i$ 's bidding strategy for project  $j$  depends on project characteristics,  $x_j$  and  $u_j$ , and the numbers of actual bidders. Letting  $\beta_k(\cdot)$  and  $\tilde{\beta}_k(\cdot)$  denote the group- $k$  bidding strategies associated with arbitrary draw  $u_j$  and with  $u_j = 1$ , respectively, under our assumed cost structure  $\beta_{k(i)}(c_{ij}|x_j, u_j, n_{1j}, n_{2j}) = u_j \tilde{\beta}_{k(i)}(\tilde{c}_{ij}|x_j, n_{1j}, n_{2j})$ . This implies  $b_{ij} = \tilde{b}_{ij} u_j$ , where  $\tilde{b}_{ij}$  denotes the firm-specific bid component given by  $\tilde{b}_{ij} = \tilde{\beta}_{k(i)}(\tilde{c}_{ij}|x_j, n_{1j}, n_{2j})$ , or  $\ln(b_{ij}) = \ln(\tilde{b}_{ij}) + \ln(u_j)$ . Therefore, the distribution of log-bids for project  $j$  depends on  $x_j, u_j, n_{1j}$ , and  $n_{2j}$ , with the log of the unobserved project heterogeneity acting as an additive mean shifter.

The distribution of firms' bid preparation costs,  $d_{ij}$ , is given by  $G_d^{k(i)}(\cdot|z_j)$ . We assume that firms' bid preparation costs are independent conditional on observed and unobserved project characteristics,  $x_j, z_j$ , and  $u_j$ , and the number of potential bidders,  $N_{1j}$  and  $N_{2j}$ . The theoretical model implies that in the auction for project  $j$  firms' participation behavior is characterized by group-specific thresholds,  $D_{kj}(\cdot)$ , defined by equation (6). The bid preparation cost is private information. Therefore, from the researcher's and the competitors' point of view, the number of actual bidders from group  $k$  is distributed according to a binomial distribution with probability of success of  $p_k(x_j, u_j, z_j, N_{1j}, N_{2j})$  and  $N_{kj}$  trials, where

$$p_k(x_j, u_j, z_j, N_{1j}, N_{2j}) = G_k(D_{kj}(x_j, u_j, N_{1j}, N_{2j})|z_j) \quad (8)$$

**Bid and Entry Cost Distribution Functions.** In estimation, we make parametric assumptions about the distributions of interest because of the relatively small size of our dataset, exploiting instead the availability of a large number of covariates that potentially affect project and entry cost distributions.<sup>15</sup> We assume that the log of the individual bid component  $\ln(\tilde{b}_{ij})$

---

<sup>15</sup>Our parametric assumptions are motivated by prior literature (see for example Hong and Shum (2002), Porter and Zona (1993), etc.), and by results in Krasnokutskaya (2009a) that indicate that the distributions of the firm-specific bid component and of unobserved heterogeneity are close to log-normal.

is normal with mean,  $\mu_{F,kj}$ , and variance,  $\sigma_{F,kj}^2$ , specified as:

$$\begin{aligned} E[\ln(\tilde{b}_{ij})|x_j, n_{1j}, n_{2j}] &= [x_j, n_{1j}, n_{2j}]' \alpha_k \\ Var[\ln(\tilde{b}_{ij})|x_j, n_{1j}, n_{2j}] &= (\exp(y_j' \eta_k))^2 \end{aligned} \quad (9)$$

where  $y_j$  includes some of the project characteristics contained in  $x_j$ . We further assume that  $\ln(u_j)$  is distributed according to a normal distribution with mean zero and standard deviation  $\sigma_u$ .

Last, to ensure that entry costs are positive, we assume that they are distributed according to a normal distribution left-truncated at 0 with mean  $E[d_{ij}|z_j] = z_j' \gamma_k$  and a constant, group-specific standard deviation  $\sigma_k^G$ .

## 4.2 Estimation Approach

Our empirical model yields predictions for equilibrium bids and group-specific participation probabilities. We match these to data using a generalized method of moments estimator. Here, we summarize the theoretical moment conditions that we use to estimate the parameters of the firm-specific bid component, unobserved heterogeneity, and entry cost distributions. The Appendix contains a detailed derivation of the theoretical and empirical moment conditions we use.

**Bid Distribution Parameters.** To estimate the parameters of the mean of  $\ln(\tilde{b}_{ij})$ , we exploit that:

$$\begin{aligned} m_1 &= E[x_j'(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})] = 0 \\ m_2 &= E[n_{kj}(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})] = \\ &\int \int \sum_{n_1, n_2 \subset N_1, N_2} n_{kj} \ln(u_j) Pr(n_{1j}, n_{2j} | x_j, z_j, N_{1j}, N_{2j}, u_j) h(u_j) du_j dF(x_j, z_j, N_{1j}, N_{2j}). \end{aligned} \quad (10)$$

The moment condition for the parameters that correspond to the numbers of bidders reflects the dependence of the joint distribution of  $(n_1, n_2)$  on  $u$  through  $p_k(x_j, u_j, z_j, N_{1j}, N_{2j})$ .

We identify the parameters of the standard deviation of  $\ln(\tilde{b}_{ij})$ ,  $\eta_k$ , from the following second-order moments:

$$\begin{aligned} m_3 &= E[(\ln(b_{i_1j}) - \ln(b_{i_2j}))^2] = \\ &E[(\exp(y_j' \eta_{k(i_1)}))^2 + (\exp(y_j' \eta_{k(i_2)}))^2] + E[(x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)})]^2 \\ m_4 &= E[x_{jl}(\ln(b_{i_1j}) - \ln(b_{i_2j}))^2] = \\ &E[x_{jl}((\exp(y_j' \eta_{k(i_1)}))^2 + (\exp(y_j' \eta_{k(i_2)}))^2)] + E[x_{jl}([x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)})]^2]. \end{aligned} \quad (11)$$

Finally, the standard deviation of the distribution of unobserved project heterogeneity,  $\sigma_u$ , is estimated from a second-order moment condition:

$$m_5 = E[(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})^2] = \sigma_u^2 + E[(\exp(y'_j \eta_k))^2]. \quad (12)$$

In estimation, we also use moments of order three and four for the bid distribution.<sup>16</sup> Their derivation is presented in the Appendix.

**Cost of Entry Distribution Parameters.** A second group of moments is used to recover parameters of the entry cost distributions,  $\gamma_k$  and  $\sigma_k^G$ . We use the first and the second moments of the binomial distribution for the numbers of actual bidders.<sup>17</sup> We specifically consider separate moments for bidder groups,  $k$ , and project size categories,  $size_j$ , where  $size_j = \{small, medium, large\}$ :

$$m_6^{kl} = E[n_{kj} | size_j = l] = \quad (13)$$

$$\int \int p_k(x_j, z_j, u_j, N_{1j}, N_{2j}) N_{kj} h(u) du dF(x_j, z_j, N_{1j}, N_{2j} | size_j = l)$$

$$m_7^{kl} = E[n_{kj}^2 | size_j = l] = \quad (14)$$

$$\int \int (p_k(x_j, z_j, u_j, N_{1j}, N_{2j})(1 - p_k(x_j, z_j, u_j, N_{1j}, N_{2j})) N_{kj} + N_{kj}^2 p_k^2(x_j, z_j, u_j, N_{1j}, N_{2j})) h(u) du dF(x_j, z_j, N_{1j}, N_{2j} | size_j = l).$$

We further include third and fourth order moments of the binomial distribution of the numbers of bidders in estimation. Their derivation is presented in the Appendix.

**Implementation.** In computing the empirical counterparts to the moment conditions in equations (10) through (14), we use Monte Carlo simulation techniques to integrate over the distribution of unobserved heterogeneity. Our reliance on simulation techniques motivates our choice of a simulated GMM estimator over a simulated maximum likelihood estimator, which is highly nonlinear in participation probabilities and therefore more sensitive to simulation error in the at times small participation probabilities (see Ackerberg, Lanier Benkard, Berry and Pakes (2007) for a discussion of the advantages of simulated GMM in similar discrete-choice settings).

To compute the value of objective function for a given guess of parameter values we follow a number of steps. First, for every draw from the distribution of unobserved heterogeneity  $h(u_j)$ , we use the first-order conditions for optimal bidding to recover the project cost distributions implied by the bid distribution,  $F_b^k$ , consistent with the current parameter guess (see Guerre et al. (2000)).

<sup>16</sup>We experimented with moments of higher order as well. However, the estimates were not substantially affected by inclusion of these moments.

<sup>17</sup>We estimate both a specification that relies only on first moments and a specification that uses first and second moments. The results are very similar across the two specifications. We report the estimation results for the first specification, together with predictions for the second moments based on the estimated coefficients.

Next, we numerically solve the equilibrium conditions on the participation side, Equation (6), using a nonlinear equation solver to find the equilibrium entry probabilities. To compute the expected profit from bidding in Equation (6), we use the recovered distribution of project costs to compute the expected profit for every possible combination of competitors ( $\hat{n}_{k(i)j} - 1, \hat{n}_{-k(i)j}$ ),  $\hat{n}_{k(i)j} = 0, 1, \dots, N_{k(i)j} - 1$  and  $\hat{n}_{-k(i)j} = 0, 1, \dots, N_{-k(i)j}$ . Then we combine these values into an expected profit from bidding using bidder  $i$ 's beliefs about the distribution of the numbers of his competitors. We obtain moment conditions by averaging over simulation draws as described in the Appendix. We arrive at the value of objective function by collecting moment conditions into the GMM objective function.

Our routine closely resembles the nested GMM estimation procedure used in the literature on discrete-choice demand estimation (Berry, Levinsohn and Pakes (1995)) since it includes an inner loop that finds the solution to a system of non-linear equations. Berry, Linton and Pakes (2004) show that the GMM estimator used in Berry et al. (1995) is consistent and asymptotically normal.

### 4.3 Model Identification

We conclude this section with a brief discussion of the econometric identification of our parameters. While we rely on parametric assumptions for the bids, entry costs, and unobserved heterogeneity, these distributions can be identified from our data non-parametrically. Krasnokutskaya (2009a) contains a detailed discussion of the non-parametric identification of the firm-specific cost component's distribution in the presence of unobserved project heterogeneity. The identification argument relies on the fact that conditional on the number of actual bidders, the firm-specific cost components are independent across bidders and from the unobserved heterogeneity component. This property holds for our participation model, as we show in the Appendix. Using the procedure from Krasnokutskaya (2009a), we can non-parametrically recover the marginal distributions of the firm-specific cost components and the distribution of unobserved heterogeneity conditional on the numbers of actual bidders. The marginal distribution of unobserved heterogeneity is then obtained by integrating the numbers of bidders, using the empirical distributions of the numbers of bidders in the data.

The distribution of entry costs is also identified non-parametrically. Details of the proof are in Krasnokutskaya (2009b) and are summarized in the Appendix. It can be shown that there is a unique cumulative distribution function  $G$  that could have generated the observed participation behavior under our model of entry. The proof relies on the existence of a full-support variable that affects the distribution of project costs, but not that of entry costs.

Parametric identification of  $G$  hinges on moment condition  $m_6$ , which represents the average numbers of bidders by project size category and group. For each group, the moments trace the average number of bidders as a function of project size. The intercept of this profile

identifies the constant of the distribution of entry costs; the slope identifies the coefficient for project size; and the curvature identifies the variance of the entry cost distribution.

## 5 Empirical Analysis

This section presents results of our empirical analysis. We first summarize descriptive patterns in the data that speak to the presence of cost asymmetries across groups of bidders, the heterogeneity of projects in our data set, and the strategic response of bidders to the bid preference program. We next implement our estimation strategy. We demonstrate that the predicted bid and entry choices based on our estimated parameters fit the data well, including for groups of projects not used in estimation. The estimated parameters of the entry cost distribution imply reasonable entry costs. The results confirm the presence of substantial asymmetries across bidder groups and important variation in the degree of asymmetries that correlates with project characteristics. Small bidders have higher project and entry costs for the majority of projects. However, we also identify a sizable set of projects where small bidders have lower project or entry costs or both.

### 5.1 Descriptive Analysis

Table 1 summarizes the characteristics of the set of state-funded projects that we use in estimation. Important project characteristics include the engineer’s estimate of the project’s total cost, the type of work involved, the project’s location at the level of the administrative district, and the time allocated to complete the project. The engineer’s estimate reflects Caltrans’ assessment of the project’s price based on similar projects auctioned in the past. We follow other procurement auction studies (e.g., Hong and Shum (2002), Jofre-Bonet and Pesendorfer (2003), Porter and Zona (1993)) in using it as a proxy for the size of the project.

We split projects into five work categories: bridge work; landscaping; road repair; signs, signals and lighting; and small building work. Road-repair projects account for 60.26% of contracts; small building work accounts for another 15.93% of contracts, while 10.04% of contracts are for bridge work. The remaining contracts are split roughly equally between landscaping and signs/lighting work. Across projects, the median project has an engineer’s estimate of \$464,000 (standard deviation of \$740,000) and duration of 45 working days (standard deviation of 165 days). Table 1 further highlights significant heterogeneity in the competitive environment. On average a project attracts 4 small potential bidders and 6.5 large potential bidders with 1.7 small and 2.6 large firms submitting bids.

The bottom panel of Table 1 summarizes potential and actual entry separately for small, medium, and large projects, representing the terciles of the distribution of the engineer’s estimate. The small-firm participation rate declines sharply with project size. It drops from 51% of small

potential bidders submitting bids in small projects to only 35% in large projects. In contrast, the participation rate of large firms is roughly constant across project sizes, ranging from 38% to 40%.

To investigate how participation rates vary with project characteristics, we conduct a probit analysis of a potential bidder’s decision to submit a bid (see Table 2). We include proxies for the competitive environment and project characteristics (size, time to completion, type of work, location) and allow the coefficients to differ for small and large plan holders. We control for unobserved project characteristics by including the number of actual bidders. We divide project locations into rural and urban based on the project’s administrative district, defining a project to be rural if it is located in the North Coast, North Central, South Central, or Southern Sierra districts. We also combine bridge and road work into one group, relative to the remaining contracts.

The probit analysis reveals a negative, statistically significant effect of the number of potential competitors on a firm’s participation decision. This is true for potential competitors of the same group as well as of the opposite group. The presence of an additional small potential bidder decreases both a small and a large firm’s propensity to submit a bid by about twice the reduction brought forth by the presence of an additional large potential bidder, a statistically significant difference. This evidence is consistent with companies’ strategic response to the bid preference program.

Table 2 also suggests heterogeneity in participation across locations and type of work. We include interaction variables of the project’s location (urban or rural) and the project’s type of work (road repair/bridge or other) and estimate differences between small and large firms’ participation rates. Across project types, small firms have statistically significantly lower participation rates than large firms. The difference is more pronounced for urban projects, which are larger on average than rural projects, in line with the results in Table 1. Small firms are also less likely to participate in road-repair than in other projects, regardless of project location; however, the difference is statistically significant for urban projects only. Large firms, in contrast, exhibit less heterogeneity in participation choices, and we cannot reject the equality of participation rates across locations and types of work.

These regularities indicate that project size, location, and type of work affect entry in a group-specific way, potentially reflecting differences in the cost of completing a particular project or the cost of preparing bid documents. To investigate the former, we conduct a regression analysis that relates log-bid levels to project characteristics.<sup>18</sup> Table 3 summarizes the results. The estimated coefficients have the expected signs. We find that log-bids increase in the engineer’s estimate and the project’s duration.<sup>19</sup> In addition, we find significant variation in bid levels

---

<sup>18</sup>We include the numbers of potential bidders to control for unobserved project heterogeneity.

<sup>19</sup>In separate regression models (available upon request), we investigate the role of capacity constraints in explaining firm behavior, which would introduce a dynamic element to the firms’ decision making. We follow

across work types and locations, even after controlling for project size. Conditional on project characteristics, the average bid of a small bidder is 8.1% higher than that of a large bidder.

In summary, the descriptive evidence suggests that bidding and entry behavior differ by firm group. We find that the number of small potential bidders affects participation decisions of both groups of bidders more strongly than the number of large potential bidders. This suggests that the Small Business Preference program affects the operation of this market. At the same time, small firms submit bids significantly less frequently, and if they do, bid higher than large firms. Such participation and bidding behavior could arise due to large differences in project costs between small and large firms even if the costs of preparing bids are similar across groups. On the other hand, even without pronounced differences in project costs, small firms' bids may be higher due to the competitive advantage awarded by the preference program, while their less frequent entry is due to larger bid preparation costs.

We now turn to the results of the estimated empirical model that allows us to disentangle the role of the preference program from inherent cost differences between firms, both of which are reflected in the observed firm choices.

## 5.2 Estimation Results

We specify the mean of log bids as a linear function of the log of the engineer's estimate, duration, the numbers of actual and potential bidders and dummies for type of work and location. We also include year dummies to control for cost inflation and monthly dummies to control for seasonal fluctuations in input prices. We allow the effects of most of these covariates to differ by bidder group. The variance of log-bids depends on the log of the engineer's estimate and the bidder's group. We assume that mean entry costs are a linear, group-specific function of the log-engineer's estimate and allow for a group-specific standard deviation.<sup>20</sup>

The results of estimation indicate that there are important differences in project and entry costs across groups of bidders. Table 4 reports the estimated coefficients of the bid distribution. The estimated coefficients are of the expected sign and magnitude. They reflect substantial variation in the means and variances of log-bids across types of work and locations. They also imply substantial differences in log-bids across bidder groups. We estimate that a small firm submits a bid that is, on average, 7.7% higher than a large firm's bid for the same project. We

---

Jofre-Bonet and Pesendorfer (2003) in computing a measure of backlog at the time of each bidding decision, but do not find a statistically significant relationship between the capacity measure and firms' participation decisions or bids. The short time dimension of our data, which is likely to render our measure of capacity utilization imprecise, makes it difficult to interpret these findings.

<sup>20</sup>We also estimated several alternative specifications. First, we estimated a specification where the unobserved project heterogeneity depends on the number of potential bidders. The coefficients for the numbers of potential bidders in the standard deviation of unobserved heterogeneity are not statistically significant; the remaining coefficients are qualitatively similar to our base specification. Second, we estimated specifications that include as additional entry cost shifters a project's number of individual tasks and nonlinear size effects. These variables do not have statistically significant effects on mean entry costs.

find that the variance of the underlying log-normal distribution of bids (which equals  $(\exp \sigma_F^2 - 1) \exp(2\mu_F + \sigma_F^2)$ ) increases in the engineer’s estimate and is lower for small bidders. Unobserved project heterogeneity is important in our data. Increasing the unobserved project characteristic from a value of zero to a value equal to the estimated standard deviation of 0.14 has an effect on mean bids that is equivalent to a 15% increase in the engineer’s estimate. The estimated bid distributions fit the data well; see Figure A-4 in the Appendix for more detail.

We use Guerre et al.’s (2000) methodology to recover the distribution of project costs from the distribution of bids. First, we use the first-order conditions from the firms’ bidding problem to estimate inverse bid functions. Strict monotonicity of bid and inverse bid functions allows us to combine the estimated distribution of bids and inverse bid functions to obtain an estimate of the distribution of project costs. We summarize the estimated distributions of project costs in Table 5, where we report means and variances of project cost distributions as a fraction of the engineer’s estimate for categories of projects defined by size, type of work, and location. With the exception of large rural other work (of which our data contain only nine projects), mean project costs are close to the engineer’s estimate. Table 5 also shows important differences in means and variances of cost distributions across groups of bidders. We test for the statistical significance of these differences next.

We analyze differences in project costs across groups of bidders using a parametric bootstrap technique to test the hypothesis of the equality of the two groups’ means (standard deviations) of their project cost distributions against two-sided and one-sided alternatives. Test results differ across projects. For some projects we cannot reject equality of means or standard deviations, whereas for other projects we reject equality in favor of either group having a lower mean (or standard deviation). We aggregate the test results to the level of the project category (defined in Table 5) to document how cost differences between small and large bidders vary with project characteristics. The aggregation is performed as follows. For every project  $j$  we define a random variable  $I_j$ , which is equal to one if the null hypothesis of equal means is rejected for project  $j$  at the 5% level of significance, and is equal to zero otherwise. The probability that  $I_j = 1$  is thus equal to 0.05. We then calculate the number of rejections  $r_g$  for every category of projects,  $g$ . We report  $\text{Prob}(\text{number of rejections} \geq r_g)$ , the p-value of our test at the project category level, in Table 6.

We find that with a two-sided alternative we reject the equality of means (and standard deviations) across bidder groups for every category of projects. We, therefore, do not report the results of this test in the table. The tests with one-sided hypotheses are more interesting. We can reject an equality of means of the project cost distributions in favor of small bidders having a higher mean than large bidders for most categories. For rural road work and small rural other work, however, we reject the null of equal means in favor of small bidders having lower mean project costs than large bidders. We also reject the null of equal standard deviations in favor of small bidders having lower standard deviations than large bidders for all project categories.

There are thus important project cost differences between small and large firms. However, small firms are not always weaker players in the market. With several exceptions<sup>21</sup>, empirical auction studies rely on the assumption of symmetric bidders. Here, we document significant, and at times unexpected, differences between bidder groups. Such cost differences are important in our application since the use of discrimination is most effective in environments with asymmetric bidders.

Next, we turn to the estimated coefficients for the cost of entry distribution reported in Table 7. All coefficients have the expected signs and are statistically significantly different from zero. We have also estimated specifications that include conditional moments based on the type of work and location in addition to size classes. We use these additional moments to perform a test of overidentifying restrictions. The overidentifying restrictions could not be rejected on the basis of our estimates and the estimated parameters vary little across specifications.

Table 8 reports the fit for our base specification. The top panel shows the fit for the moments that we use in the estimation. The lower panel reports average and predicted numbers of actual bidders for other project groupings that were not used to form moment conditions in estimation. While the literature has not established a benchmark for assessing the fit of the entry part of our model, our fit appears to be good.

Table 9 reports the implied mean cost of entry and mean cost as a fraction of the engineer’s estimate across bidder groups and project size categories. We estimate that mean entry costs amount to 2.7% to 3.7% of the engineer’s estimate. This ratio increases with project size for small bidders but decreases in size for large bidders. Our estimates are comparable to estimates obtained in the academic literature (Bajari, Hong and Ryan 2009) and suggested magnitudes from general construction manuals.<sup>22</sup> We also test the equality of the two groups’ cost of entry distributions. We reject equality for all but one project size category. The results of the test are reported in the last column of Table 9.

## 6 Counterfactual Analysis

We use the estimation results to assess the effect of the preferential treatment of small firms on participation, the cost to the government, and the probability that a project is awarded to a small firm. After a brief overview of the counterfactual approach, we first contrast the outcomes of a preference auction under endogenous and fixed participation under a range of discount values. This allows us to investigate whether a bid discount could serve as an effective tool to lower the government’s cost of procurement or to achieve California’s allocative goal. Due

---

<sup>21</sup>For example, Asker (2009), Athey et al. (2008), Bajari (2001), and Jofre-Bonet and Pesendorfer (2003).

<sup>22</sup>Halpin (2005) and others suggest that estimating costs (cost of time and effort expended to develop a total bid price and submit a proposal) typically range between 0.25% and 2% of the total project cost, but vary widely depending on the complexity, type of job, and type of work being estimated.

to the computational cost of numerically deriving equilibrium bidding strategies, we do so for select representative projects only. We then study the current program as a detailed example of policy effects at a relatively low discount level, before considering an entry tax or subsidy as a preferential treatment mechanism that targets the participation margin directly.

**Methodology.** To compare behavior in alternative environments, we need to derive the appropriate bidding strategies that solve the system of differential equations defined by the first-order conditions in Equation (2). Except for special cases, this system of differential equations does not have a closed-form solution and has to be solved numerically. We apply and extend the method proposed by Marshall, Meurer, Richard and Stromquist (1994) to our setting.<sup>23</sup>

As a performance check, we initially compare the simulated entry probabilities for the 5% discount level to the entry probabilities implied by our estimation routine. The estimation routine computes expected profits conditional on participation using the observed bid distributions directly (see Guerre et al. (2000)), thus avoiding the simulation step. The simulation routine produces entry probabilities that match closely the ones used in estimation. Table A-2 in the Appendix contains a detailed comparison by project category.

**The Role of Participation.** We use the numerical routine to simulate auction outcomes under a large set of discount values for five typical projects that vary in small and large firms' relative project and entry costs. We consider two scenarios; in (1) we hold participation fixed at the zero discount level, and in (2) we allow participation to adjust with the discount level. Figure 2 illustrates for two most typical (yet very dissimilar) types of projects the changes that participation adjustments introduce into the relationship between the discount and auction outcomes, such as the cost of procurement (or the expected winning bid), the small-firm probability of winning, and the expected numbers of bidders. Both the probability of small-firm award and the cost of procurement have flatter profiles under fixed than under endogenous participation. The fixed participation case isolates the response of bidding strategies to alternative discount levels. With endogenous entry, the bid response is enhanced through a decline in large-firm and an increase in small-firm participation associated with increasing discounts to small firms. Hence, the probability of small-firm award rises not only because a given small bidder's probability of winning increases, but also because the proportion of small participants increases.<sup>24</sup> In turn, the

---

<sup>23</sup>Bajari (2001) and Marshall et al. (1994) provide details on numerical solution algorithms for asymmetric auctions. Marshall et al. (1994) use polynomial approximations to the cost distributions and employ a forward recursive algorithm to solve the resulting set of difference equations with an upper boundary condition. We extend their approach as follows. We embed the recursive algorithm in a search routine for a starting point that satisfies the upper boundary conditions. We approximate the estimated cost distributions by polynomial splines, which we found to produce more stable results than the original Taylor approximations. Finally, we extend their setup in which a single asymmetric bidder competes against a second group of bidders to settings with arbitrary numbers of bidders within the two groups, which entails solving a larger-dimensional system of differential equations. As in estimation, we use Monte Carlo simulation to integrate over the distribution of unobserved heterogeneity.

<sup>24</sup>Note that in our counterfactuals, we hold the pool of potential entrants fixed. These results thus do not reflect that the discount may change the incentive of a firm of either group to become a potential bidder, which

cost of procurement increases as a higher proportion of contracts is awarded to small bidders who charge higher prices due to their high costs and the bid discount. The fixed and endogenous participation scenarios also differ in their implications for the discount levels needed to achieve procurement cost minimization or an allocative goal such as California's.

As shown in Figure 2, the implications of accounting for endogenous participation are quite different for the two projects because they are characterized by different degrees of asymmetry between small and large bidders. Project 2 belongs to a group of projects where small firms' project and entry costs are very similar to the costs of large bidders. In addition, in these projects the variances of the project costs distributions tend to be lower than for the average project. As a result, a non-negligible share of large bidders is priced out of the auction and chooses not to bid once they observe their project cost. This effect, which is significant for this group of projects, largely mimics the participation adjustment effect. As a result, the relationships under fixed and endogenous participation are similar. Most small projects, and medium rural road work projects, share these properties.

In contrast, project 1 is exemplary of medium and large urban projects where small firms are the less efficient group in both entry and project costs. Because of substantial asymmetry there is unlikely to be a significant mass of large bidders whose bid is beaten by all, including the highest-cost, small bidders. That is why the share of large bidders dropping out after observing their project costs is much smaller than in the case of project 2. This implies that under fixed participation the group of viable bidders does not change very much as the discount level changes.

**Cost-Minimizing Discount.** We use the five projects to explore the potential of a bid discount to lower the cost of procurement. Table 10 shows that discounts of 10%, 15%, and 22% to large bidders for projects 2, 3, and 4, respectively, minimize government cost under fixed participation since large bidders are less efficient for these projects. The model with fixed participation prescribes 10% and 15% discounts to small bidders for projects 1 and 5 where small firms have higher project costs. In contrast, the cost-minimizing policy that takes participation adjustments into account does not generally favor the group with the highest project costs, but reflects entry cost differences as well. It implies that very high discounts should be given to large firms (which would induce small firms not to participate) on projects 1 and 5; that a 10% discount should be given to large firms on project 2 and small firms on project 3; and that no discount should be awarded to either firm group on project 4. It is worth noting that if the government followed the prescriptions of the fixed-participation model, the cost of procurement would actually increase rather than decrease in four out of the five cases after participation adjusts.

Under the model with participation adjustment the program affects not only bidding behavior but also the composition of bidders. The bid preference generates the highest cost

---

is likely to reinforce the latter effect.

savings to the government for projects 1 and 5 where small firms' project and entry costs are much higher than the corresponding costs of large firms. In these projects large firms impose substantial competitive pressure on the market. In response to a bid discount, the equilibrium participation of large bidders grows and this competitive pressure intensifies, even though the equilibrium participation of small bidders declines at the same time. This holds for all large-firm participation levels that arise in equilibrium under a large-firm bid discount, regardless of the magnitude of the discount. The reverse is true if a discount is granted to small firms: the associated decline in large-firm equilibrium participation adversely affects the competitive intensity despite the increased presence of small firms. It is therefore optimal for the government to use the discount to encourage entry by large bidders (see the Appendix for details on the optimal policy for these projects). Similarly, for project 3, small firms have lower average project costs, but face higher entry costs. The government's cost-minimizing strategy promotes their participation, which lowers the winning bid through the change in the composition of the set of actual bidders.

**Other Discount Levels.** Most preference programs pursue goals other than pure cost minimization and therefore are likely to produce procurement cost increases. As Figure 2 indicates, such increases can be quite large, ranging for both projects from 1% for discounts as low as 10% to as much as 5% with higher discounts of 40%.

We assess the likely magnitudes of cost increases associated with the preference program's objectives using the allocative goal of the California Small Business Preference program as an example. The second panel of Table 10 reports the discount rates necessary to achieve this goal and the associated cost increases for the five projects discussed earlier. The resulting cost increase range from 0% to 2.3%, with the largest increases being associated with projects where small bidders are substantially less efficient than large bidders (in either project or entry costs). Significantly, the fixed participation model suggests that a much higher (by a factor of two) discount level should be used. Choosing this discount level without recognizing the changed participation incentives is costly: procurement costs rise by approximately 7% across projects.

Given the heterogeneity of cost asymmetries across projects in the data, we also compute an aggregate measure of the cost of allocating 25% of the State's procurement load to small firms. We find that a 13% discount approximately satisfies the aggregate award goal (see Table 11). This results in an approximate increase of 1.2% in the cost of procurement relative to no government intervention.<sup>25</sup> In contrast, if we held participation fixed, we would conclude that a much higher discount of 35% is needed to achieve the allocative goal. The model with fixed participation substantially underestimates the cost increase associated with this discount level

---

<sup>25</sup>We find  $\delta$  that sets, across projects  $J$ ,  $\sum_j p_{sm,j}^{win}(\delta) GovCost_j = 0.25 \sum_j GovCost_j$ , where  $p_{sm,j}^{win}$  denotes the small-firm probability of winning auction  $j$ . We approximate the probability of winning for each individual project by the equivalent for a representative project in its project category (for project categories, see Table 12). We similarly approximate the cost of procurement for a given project by the cost for the representative project.

predicting that the cost would go up by only 1.5%. This assessment ignores participation effects, which would bring the cost increase to 4.5%.

The modest aggregate cost increase for discount levels prescribed by the endogenous participation model reflects the composition of projects in the data, which contain a significant share of projects where small and large firms have similar costs, as in project 2. In these projects small firms are efficient competitors and have a high award rate even in the absence of a discount. It seems that the government could reduce its cost of procurement even further by granting discounts only for projects where small bidders are typically inefficient and thus have low participation and award rates.

**Evaluation of Current Policy.** Next we turn to an analysis of the bid preference program currently in place in California, which uses a relatively low discount of 5%. We compare auction outcomes in the current environment to the counterfactual setting where the state does not use a preference program and instead treats bidders equally. Tables 12 and 13 contain the results of this analysis. We simulate auction outcomes for a larger subset of 119 projects to capture project heterogeneity more finely based on project size, location, and type of work.<sup>26</sup>

Table 12 reports changes in the cost to the government measured as an expected winning bid. The cost to the government does not change very much as a result of the preferential treatment of small bidders. While the cost to the government goes up in some cases and goes down in others, these effects barely amount to a 0.5% change for all project categories as well as for most individual projects.

Table 13 compares probabilities of entry by project category. The preferential treatment produces the expected increased small-firm and reduced large-firm participation. The magnitudes of these effects are economically significant, however, and differ substantially across project categories. Small-firm entry probabilities increase by between 2 and 8 percentage points, or 3.1% and 20%, while large-firm entry probabilities decline by between 3 and 6 percentage points, or 4.3% and 10.6%. The changes in the groups' participation are close to off-setting, however: total entry is virtually unchanged across project categories, with increases or decreases in overall participation of only approximately 1%.

The participation effects contribute significantly to the increase in small firms' probabilities of winning. As Table 14 shows, the change in probability of winning under endogenous participation is twice the change generated by the discount only (under fixed participation). Finally, the program also increases small potential bidder's expected profit prior to participating by 9% to 25%, with an average of 16%, while decreasing large firms' profits by 10% to 18%, with an average of 13% (see Table 12). The preferential treatment thus results in a non-trivial redistribution of profits from large to small firms at almost no cost to the government.

The changes in entry and profits differ substantially in magnitude across project types.

---

<sup>26</sup>Ten projects in each category and all nine projects for large rural other work projects.

Two potential sources of such differences are (1) variation in cost asymmetries and (2) differences in market thickness, or the number of potential bidders. We investigate how these factors affect the magnitude of the program’s impact on small bidders’ participation in Table 15. The table reports the results of an OLS projection of the absolute change in small bidders’ probability of entry on project characteristics, potential entry, and the moments of the two groups’ entry and project cost distributions. The results suggest that small-firm participation responds more strongly for larger projects, for projects where small firms have lower average project costs than large firms, and for projects where the within-group variation in entry costs is lower, in particular for large firms. Last, the program has stronger effects for projects with fewer small potential bidders, but a higher number of large potential bidders. These effects are intuitive. Larger projects produce a larger absolute gain from the program that offsets entry costs. Low variance of the entry cost distribution implies that a given change in expected profit from participating affects the entry behavior of a larger mass of firms. The number of potential bidders reflects the competitive intensity and the size of the set over which the program’s profit gains or losses are divided.

The second specification in Table 15 shows that after controlling for small firms’ base probability of entry at  $\delta = 0$ , only project costs and potential competition play a statistically significant, now larger, role in promoting participation. The results also indicate that gains in small-firm participation are larger in projects where their participation would have been low in the absence of preferential treatment. The program thus appears to be more effective for projects where participation of small firms is impeded without preferential treatment.

This analysis is related to Marion (2007) who provides an alternative estimate of the effect of the California Small Business Preference program on the cost of procurement. He measures this effect by comparing a set of state-funded projects where the program is implemented to a set of federally-funded projects. He finds that the average winning bid on state-funded projects exceeds that on federally funded projects by 3.8%. Attributing this difference to the program is complicated by the fact that federally funded projects have another preferential treatment program in place that restricts bidders’ subcontracting choices. In addition, federally and state-funded projects differ along observable dimensions, suggesting that they may also differ on unobservable characteristics that would affect firms’ cost distributions and, thus, the magnitudes of the effects of a preference program. If this were the case, the observed difference in participation patterns between the two sets of auctions would similarly not represent the changes in participation brought forth by the program.

**Subsidy.** Our analysis so far shows that increases in small-firm participation translate into increases in the group’s probability of project award. This additional effect is often stronger than the direct effect of the discount, which works through the change in bidding strategies. Our analysis further suggests that differences in bid preparation costs contribute significantly to the difference in participation rates across bidder groups. Hence, a direct entry subsidy of

small firms (or tax of large firms) could alternatively help to achieve small-firm award goals. A subsidy would increase the cost of procurement to the government, whereas a tax may reduce government outlays.

Table 16 summarizes lump-sum subsidy or tax policies that achieve alternative government goals for the five sample projects discussed above. We find that the unconstrained cost-minimizing policy (panel 1) involves taxing both groups.<sup>27</sup> It reflects the trade-off between tax rate and tax base: higher tax rates increase per-firm tax receipts, but lower participation, putting upward pressure on the expected winning bid. The government minimizes its cost of procurement by choosing tax rates so only the more efficient group participates in bidding, where efficiency reflects both project and entry costs. For a detailed, graphical example of the relationship between the tax and the total cost of procurement, see Figure A-5 in the Appendix.

We next turn to a policy that achieves California's goal of allocating 25% of procurement dollars to small firms (see panel 2 of Table 16). A subsidy to small firms that achieves this objective is less costly to the government than the equivalent bid discount. It realizes cost savings of 8% on average relative to a bid discount that achieves the same small-firm award rate. These cost savings are realized in part by taxing large firms and come at the cost of substantially lowering that group's profits, which decline by between 25% to 100% relative to the bid discount case.

Since such high penalties on large firms may be undesirable, we consider an alternative scheme that achieves the small-firm award goals, but also limits the large-firm profit reductions to those under the bid discount with an identical award level. For all five projects, such a policy is feasible. In all cases, it results in procurement cost savings over the bid discount policy (see panel 3 in Table 16). This implies that the government can achieve the same small-firm award rate at lower cost, even without additional harm to large firms.

These savings arise because subsidies and taxes do not depend on firms' project cost realizations. In contrast, the bid discount disproportionately rewards firms with high project costs that raise their bids by the highest absolute amount and increase their probability of winning most. A subsidy or tax, by targeting the extensive participation margin only, eliminates these distortions, resulting in a lower cost of procurement. It provides incentives for firms to submit uncompetitive bids for the sole purpose of collecting the subsidy payments, however. To avoid this, the subsidy should be awarded only to the winning bidder, while a tax can be applied to all entrants from the taxed group. The subsidy level has to be adjusted to account for this modification. However, the magnitudes of all effects remain unchanged.

---

<sup>27</sup>With high tax levels, participation may drop to zero, resulting in a non-award of the project. The true economic cost of a non-award and later re-advertising of a possibly re-scoped project is difficult to estimate. We assume, however, that it exceeds the cost of awarding the project immediately. In our simulations, we simply set the government's cost in this case equal to the particular project's engineer's estimate, an amount that across projects exceeds the minimum cost to the government.

## 7 Conclusion

This paper provides evidence based on the California Small Business Preference program on the channels through which bid discounts affect procurement outcomes, separating adjustments in firms' participation behavior from those in their bidding decisions. Within our empirical context, we find that the response in firms' bidding behavior (conditional on participation) to alternative discount levels changes aggregate procurement costs only by a limited amount relative to more substantial changes resulting from participation adjustments. This is of critical importance to policy design; we show, for example, that taking firms' participation incentives into account alters the bid discounts that achieve the government's procurement goals and the assessment of the costs increases associated with different discount levels.

California's current program generates only small increases in procurement costs. While promoting small-firm participation - at the expense of large-firm participation and profit, it does not achieve the State's allocative goal. Our results imply that for the set of projects in our data, a higher discount of 13% is needed to reach the allocative target. This discount level does not come at substantial cost increases, however, raising the aggregate cost of procurement by 1.2% relative to no intervention. It is important to note that these results depend crucially on the mix of projects in California's highway procurement market. In other markets allocative goals may lead to larger or smaller cost increases.

We consider the cost implications of broader policy re-design. In line with theoretical predictions for environments with fixed participation, we find that a bid discount can be used to lower the government's cost of procurement. If the degree of cost asymmetries between small and large firms is high, the cost-minimizing auction design prevents the inefficient group - typically small firms - from participating by granting a large discount to the other group. Since projects where small firms are very inefficient are easily identifiable by observable attributes, the government may prefer to use the set-aside auctions common in the timber industry in such instances. In our dataset, however, even the cost-minimizing discount generates only modest cost reductions, while discounts that depart from this level - but remain within the range of typically used bid discounts - can result in significant cost increases for at least a subset of projects.

We find large heterogeneity in the effect of the bid discount across types of projects. This suggests that the government should optimally employ a more nuanced preferential treatment, tailoring the discount rate to the project type, similar to the approach taken by the FCC. This can result in substantial cost savings while facilitating the implementation of the State's allocative goals.

Our findings suggest further that a lump-sum entry fee is more effective than a bid discount at reducing the cost of procurement. We also show that a moderate fee is sufficient to achieve the program's award goal at the same large-firm profit levels as a bid discount with the same award outcomes, while still bringing about an important reduction in the cost of

procurement.

Our results demonstrate that a preference program evaluation depends critically on capturing firms' participation responses to the policy. While our findings are based on the highway procurement market, we believe that this insight, as well as our technique for predicting participation responses, is pertinent to other auction markets where discriminatory policies are used. A number of open questions remain. We focus on the short-run effects of the program, but do not assess its dynamic, long-run implications. This includes adjustments to the set of potential bidders, which we hold fixed throughout our analysis. Due to the complexities of analyzing asymmetric auctions in a dynamic game, we also do not formally consider the importance of capacity constraints that could affect project costs and thus both bids and participation incentives. Following Caltrans' current practice, we do not introduce a reserve price into our analysis. Similar to the bid discount, a reserve price limits the participation of favored and non-favored bidders with high project costs. At the same time, it induces favored firms to bid more aggressively than under the bid discount alone, thus limiting the re-distribution of profits to favored firms. This raises a number of interesting issues, including the optimal policy design in the presence of possibly group-specific reserve prices and the importance of participation responses in the optimal policy. We leave these to future research.

## References

- Akerberg, Daniel, C. Lanier Benkard, Steven Berry, and Ariel Pakes**, "Econometric Tools for Analyzing Market Outcomes," in J.J. Heckman and E.E. Leamer, eds., *Handbook of Econometrics*, Vol. 6 (63), Elsevier, 2007.
- Asker, John**, "A Study of the Internal Organisation of a Bidding Cartel," *American Economic Review*, 2009. Forthcoming.
- Athey, Susan, Jonathan Levin, and Enrique Seira**, "Comparing Open and Sealed Bid Auctions: Theory and Evidence from Timber Auctions," Working Paper, Harvard University 2008.
- Ayres, Ian and Peter Cramton**, "Deficit Reduction Through Diversity: How Affirmative Action at the FCC Increased Auction Competition," *Stanford Law Review*, 1996, 48 (4), 761–815.
- Bajari, Patrick**, "Comparing competition and collusion: a numerical approach," *Economic Theory*, 2001, 18 (1), 187–205.

- and **Ali Hortacsu**, “The Winner’s Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions,” *RAND Journal of Economics*, 2003, *34* (2), 329–355.
- , **Han Hong**, and **Stephen Ryan**, “Identification and Estimation of a Discrete Game of Complete Information,” *Econometrica*, 2009. Forthcoming.
- Berry, Steven, James Levinsohn, and Ariel Pakes**, “Automobile prices in market equilibrium,” *Econometrica*, 1995, *63*, 841–890.
- , **Oliver Linton**, and **Ariel Pakes**, “Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems,” *Review of Economic Studies*, 2004, *71*, 613–654.
- Corns, Allan and Andrew Schotter**, “Can Affirmative Action Be Cost Effective? An Experimental Examination of Price-Preference Auctions,” *American Economic Review*, 1999, *89*, 291–305.
- Decarolis, Francesco**, “When the Highest Bidder Loses the Auction: Theory and Evidence from Public Procurement,” Working Paper, University of Wisconsin 2009.
- Denes, Thomas A.**, “Do Small Business Set-Asides Increase the Cost of Government Contracting?,” *Public Administration Review*, 1997, *57* (5), 441–444.
- Guerre, Emmanuel, Isabelle Perrigne, and Quang Vuong**, “Optimal Nonparametric Estimation of First-Price Auctions,” *Econometrica*, 2000, *68* (3), 525–574.
- Halpin, Daniel**, *Construction Management, 3rd Edition*, Wiley, 2005.
- Hong, Han and Matthew Shum**, “Increasing Competition and the Winner’s Curse: Evidence from Procurement,” *Review of Economic Studies*, 2002, *69* (4), 871–898.
- Jofre-Bonet, Mireia and Martin Pesendorfer**, “Estimation of a Dynamic Auction Game,” *Econometrica*, 2003, *71* (5), 1443–1489.
- Krasnokutskaya, Elena**, “Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity,” *Review of Economic Studies*, 2009. Forthcoming.
- , “Note on the Identification of the Entry Cost Distribution in the Presence of Unobserved Heterogeneity,” Working Paper, University of Pennsylvania 2009.
- Levin, Dan and James L. Smith**, “Equilibrium in Auctions with Entry,” *American Economic Review*, 1994, *84* (3), 585–599.
- Li, Tong**, “Econometrics of first-price auctions with entry and binding reservation prices,” *Journal of Econometrics*, 2005, *126* (1), 173–200.

- and **Xiaoyong Zheng**, “Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions,” *Review of Economic Studies*, 2009, *76*, 1397–1429.
- Marion, Justin**, “Are bid preferences benign? The effect of small business subsidies in highway procurement auctions,” *Journal of Public Economics*, 2007, *91* (7-8), 1591–1624.
- , “How Costly is Affirmative Action? Government Contracting and California’s Proposition 209,” *Review of Economics and Statistics*, 2009, *91* (3), 503–522.
- Marmer, Vaden, Art Shneyerov, and Pai Xu**, “What Model for Entry in First-Price Auctions? A Nonparametric Approach,” Working Paper, University of British Columbia 2007.
- Marshall, Robert C., Michael J. Meurer, Jean-Francois Richard, and Walter Stromquist**, “Numerical Analysis of Asymmetric First Price Auctions,” *Games and Economic Behavior*, 1994, *7* (2), 193–220.
- McAfee, R. Preston and John McMillan**, “Government procurement and international trade,” *Journal of International Economics*, 1989, *26* (3/4), 291–308.
- Porter, Robert H. and J. Douglas Zona**, “Detection of Bid Rigging in Procurement Auctions,” *Journal of Political Economy*, 1993, *101* (3), 518–538.
- Reny, Philip J. and Shmuel Zamir**, “On the Existence of Pure Strategy Monotone Equilibria in Asymmetric First-Price Auctions,” *Econometrica*, 2004, *72* (4), 1105–1125.
- Samuelson, William F.**, “Competitive bidding with entry costs,” *Economics Letters*, 1985, *17* (1-2), 53–57.

# Tables and Figures

Table 1: Summary Statistics, Caltrans Projects and Bidders

	Mean	Std. Dev.	10 <sup>th</sup> Pctile	Median	90 <sup>th</sup> Pctile
Engineer's estimate	615.416	738.560	165.250	464.130	1086
Working days	96.598	165.071	20	45	180
N <sup>o</sup> of small plan holders	3.947	3.485	0	3	9
N <sup>o</sup> of large plan holders	6.574	4.324	3	5	11
N <sup>o</sup> of small bidders	1.745	1.890	0	1	4
N <sup>o</sup> of large bidders	2.623	1.597	1	3	5
Small projects (n=229; median engineer's estimate=\$207,000)					
N <sup>o</sup> of small plan holders	4.886	3.656			
N <sup>o</sup> of large plan holders	5.904	3.501			
N <sup>o</sup> of small bidders	2.502	2.137			
N <sup>o</sup> of large bidders	2.349	1.652			
Medium projects (n=235; median engineer's estimate=\$464,000)					
N <sup>o</sup> of small plan holders	4.260	3.536			
N <sup>o</sup> of large plan holders	6.723	3.922			
N <sup>o</sup> of small bidders	1.762	1.882			
N <sup>o</sup> of large bidders	2.668	1.561			
Large projects (n=233; median engineer's estimate=\$787,186)					
N <sup>o</sup> of small plan holders	2.714	2.829			
N <sup>o</sup> of large plan holders	7.651	6.649			
N <sup>o</sup> of small bidders	0.954	1.219			
N <sup>o</sup> of large bidders	2.929	1.645			

Note: 697 projects. Engineer's estimate reported in \$1000s and duration in days.

Small projects denote the bottom one-third, medium projects the middle one-third, and large projects the top one-third of engineer's estimates. Plan holders are our measure of potential entrants.

Table 2: Discrete Choice Model of the Decision to Bid

	Coefficient	Std Error	Marg. Effect
ln(Eng. Estimate) $\times$ LB	0.0563	0.0291	0.0223
Working Days $\times$ LB	-0.0300	0.0170	-0.0119
N <sup>o</sup> of large plan holders $\times$ LB	-0.0537***	0.0136	-0.0213
N <sup>o</sup> of small plan holders $\times$ LB	-0.0944***	0.0089	-0.0374
ln(Eng. Estimate) $\times$ SB	-0.1458***	0.0338	-0.0578
Working Days $\times$ SB	0.0538***	0.0144	0.0213
N <sup>o</sup> of large plan holders $\times$ SB	-0.0412***	0.0074	-0.0163
N <sup>o</sup> of small plan holders $\times$ SB	-0.0783***	0.0081	-0.0310
Rural district $\times$ Road repair $\times$ SB	-0.7594***	0.1488	-0.2724
Rural district $\times$ Other work $\times$ SB	-0.6199***	0.1609	-0.2273
Urban district $\times$ Road repair $\times$ SB	-0.8823***	0.1466	-0.3135
Urban district $\times$ Other work $\times$ SB	-0.7030***	0.1619	-0.2554
Rural district $\times$ Road repair $\times$ LB	0.2898*	0.1366	0.1152
Rural district $\times$ Other work $\times$ LB	0.2282	0.1804	0.0909
Urban district $\times$ Road repair $\times$ LB	0.2516	0.1490	0.1000
Urban district $\times$ Other work $\times$ LB	0.2062	0.1605	0.0821
Observations	6538		

Dependent Variable: indicator of participation decision. Year and month effects included. Number of competing bidders included to control for unobserved project characteristics. Standard errors account for clustering at the project level. LB denotes a large firm; SB a small firm. Road repair includes bridge projects. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 3: Ordinary Least Squares Model of Submitted Bid

	Coefficient	Std Error
Small-firm indicator	0.0813	0.0357**
ln(Eng. Estimate)	0.9571	0.0127***
Working Days	0.0002	0.0001***
N <sup>o</sup> of small bidders	-0.0320	0.0074***
N <sup>o</sup> of large bidders	-0.0329	0.0069***
N <sup>o</sup> of small plan holders	0.0097	0.0041**
N <sup>o</sup> of large plan holders	0.0190	0.0038***

Observations: 3034. Adjusted  $R^2$ : 0.8996. Dependent Variable: log of submitted bid. Controls for year, month, districts, and type of work by bidder group included. Standard errors account for clustering at the project level. \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

Table 4: Estimated Parameters of Log-Normal Distribution of Bids

	Coefficient	Std Error
Constant	0.0367	0.0410
Small-firm indicator	0.0767	0.0253***
ln(Engineer's Estimate)	0.9528	0.0087***
Working Days	0.0001	4.25E-5***
N <sup>o</sup> of small bidders	-0.0328	0.0045***
N <sup>o</sup> of large bidders	-0.0344	0.0038***
N <sup>o</sup> of small plan holders	0.0009	0.0026
N <sup>o</sup> of large plan holders	0.0018	0.0022
Type of work		
Bridge	-0.1369	0.0182***
Landscaping	-0.1299	0.0212***
Road Repair	-0.0618	0.0147***
Signs, Signals, Lighting	0.0343	0.0245
Location of work		
Central Coast	-0.0265	0.0306
East Central	0.0907	0.0467*
Los Angeles	-0.0072	0.0261
North Central	0.0567	0.0338*
North Coast	0.0322	0.0333
Northern Sierras	0.0484	0.0291*
San Bernardino	0.0097	0.0284
San Diego	-0.0690	0.0303**
San Francisco	-0.0157	0.0266
South Central	0.0956	0.0302***
Southern Sierras	-0.0123	0.0380
North Central × Small	-0.1027	0.0325***
North Coast × Small	-0.0744	0.0337**
South Central × Small	-0.1072	0.0311***
Southern Sierras × Small	-0.0922	0.0419**
Std. dev. of log-bids <sup>1</sup>		
Constant	-1.1525	0.0356***
Small-firm indicator	-0.1188	0.0296***
Engineer's estimate	-0.2606	0.0244***
Std. dev. of unobs. proj. char., $\sigma_u$	0.1433	0.0109***

3,034 observations. Specification includes year and month effects by bidder type. Log-bids and the log of the unobserved project heterogeneity are assumed to be normally distributed.

<sup>1</sup> Standard deviation of log-bids estimated as  $\sigma = \exp(b_0 + b_1 \text{Small} + b_2 \text{Engineer's Estimate})$ .

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01.

Table 5: Estimated Project Costs by Project Type

Project Type	# proj	Mean		Std. Dev.	
		Small Bidder	Large Bidder	Small Bidder	Large Bidder
Small, rural, rd repair / bridge	50	0.8910	0.9540	0.3034	0.3075
Medium, rural, rd repair / bridge	59	0.8917	0.9288	0.2355	0.2435
Large, rural, rd repair / bridge	55	0.8570	0.9062	0.2141	0.2144
Small, urban, rd repair / bridge	73	0.9601	0.9375	0.3035	0.3202
Medium, urban, rd repair / bridge	89	0.9696	0.9435	0.2290	0.2428
Large, urban, rd repair / bridge	88	0.9481	0.9333	0.1998	0.2099
Small, rural, other work	39	0.9518	1.0101	0.3025	0.3121
Medium, rural, other work	13	0.9342	0.9581	0.2059	0.2234
Large, rural, other work	9	0.8192	0.8292	0.1320	0.1594
Small, urban, other work	55	1.0401	1.0084	0.3114	0.3300
Medium, urban, other work	51	0.9683	0.9542	0.2462	0.2579
Large, urban, other work	35	0.9916	0.9850	0.2112	0.2237

Means and standard deviations of project costs are averaged across projects of within project type and scaled by the engineer's estimate before averaging.

Table 6: Summary of Tests of Equality of Means of Cost Distributions by Project Type

Project Type	# proj	$H_1:sm>lg$		$H_1:sm<lg$		conclusion
		# rej	Pr(# rej)	# rej	Pr(# rej)	
Small, rural, rd repair / bridge	50	1	0.92	28	0.00	sm<lg
Medium, rural, rd repair / bridge	59	7	0.13	22	0.00	sm<lg
Large, rural, rd repair / bridge	55	7	0.12	16	0.00	sm<lg
Small, urban, rd repair / bridge	73	35	0.00	4	0.50	sm>lg
Medium, urban, rd repair / bridge	89	44	0.00	2	0.94	sm>lg
Large, urban, rd repair / bridge	88	31	0.00	1	0.99	sm>lg
Small, rural, other work	39	4	0.13	26	0.00	sm<lg
Medium, rural, other work	12	3	0.02	2	0.12	sm>lg
Large, rural, other work	9	3	0.01	1	0.37	sm>lg
Small, urban, other work	55	33	0.00	6	0.16	sm>lg
Medium, urban, other work	51	25	0.00	3	0.47	sm>lg
Large, urban, other work	35	10	0.00	0	0.17	sm>lg

Columns 2 and 4 contain the count of projects for which we reject with 95% confidence the null hypothesis of equality of mean costs against the stated alternatives. Columns 3 and 5 contain the binomial probability of observing that number of rejections or more. Similar tests for the difference in the estimated standard deviations yield rejections of the null hypothesis in favor of the standard deviation of small-firm costs being smaller than that of large-firm costs at above 99% levels of confidence.

Table 7: Estimated Parameters of Truncated Normal Distribution of Entry Costs

	Coefficient	Std Error
Constant - Small	-0.7037	0.0122***
(ln(Eng. Estimate)) $\times$ Small	0.4950	0.0028***
Constant - Large	-0.4215	0.0166***
(ln(Eng. Estimate)) $\times$ Large	0.3490	0.0018***
Std. dev. - Small	0.1707	0.0024***
Std. dev. - Large	0.1751	0.0031***

1,313 observations. \* $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ .

The estimated specification allows for unobserved heterogeneity in the bid distribution, the estimated parameters are recorded in Table 4.

Table 8: Model Fit: Entry Predictions by Project Type

	Number of projects	Small firms		Large firms	
		Predicted	Actual	Predicted	Actual
<i>Moment conditions: number of bidders</i>					
Small projects	229	2.5688	2.5616	2.3919	2.3877
Medium projects	235	1.8991	1.8962	2.6996	2.6766
Large projects	233	1.2067	1.2222	2.8167	2.8528
<i>Out of sample fit: number of bidders</i>					
Bridge projects	70	1.9592	2.0714	2.3929	2.9286
Road-repair projects	420	1.6379	1.5578	2.8178	2.8565
Small & road-repair projects	107	2.5287	2.2844	1.9103	2.3527
Medium & road-repair projects	143	1.5805	1.7215	3.0694	2.8115
Large & road-repair projects	170	2.1374	2.0612	2.5652	2.6321
Rural projects	258	1.7513	1.6167	2.8171	2.9091
Urban projects	439	1.1491	1.1172	2.9766	2.9527
<i>Out of sample fit: second moments</i>					
Small projects	229	1.9496	2.2449	1.7123	1.7164
Medium projects	233	1.2542	1.5427	1.7264	1.6997
Large projects	235	0.9072	0.9237	2.1296	1.7500

Note: The second moments compare the expectation of the number of small and large bidders squared as predicted by the model to the sample equivalent.

Table 9: Estimated Entry Costs by Project Size

Project Size	Small Firms			Large Firms			K-S test (p-val)
	Avg. Cost	SD Cost	Cost/Est	Avg. Cost	SD Cost	Cost/Est	
Small	6.635	6.020	0.032	7.325	6.420	0.035	0.081 (0.002)
Medium	14.954	11.136	0.032	13.300	9.949	0.029	0.078 (0.004)
Large	27.830	15.466	0.037	20.412	12.720	0.027	0.235 (0.000)

Note: Costs reported in \$1,000s. The K-S test reports the test statistic and corresponding p-value of a Kolmogorov-Smirnov test of the equality of the estimated cost distributions within each size category.

Table 10: Bid Preference Program Designs under Alternative Objectives, Sample Projects

	Project				
	1	2	3	4	5
$(N_{sm}, N_{lg})$	(2,3)	(2,4)	(3,3)	(2,4)	(2,4)
Project cost diff.: $\bar{c}_{sm} - \bar{c}_{lg}$	0.50	-0.10	-0.20	-0.40	0.90
Entry cost diff.: $\bar{d}_{sm} - \bar{d}_{lg}$	0.05	-1E-03	0.06	0.18	0.14
Cost to the Government, $\delta_{lg} = \delta_{sm} = 0$	4.85	1.18	4.68	10.50	8.50
<i>Cost-minimizing Policy</i>					
(1) Endogenous Entry					
$(\delta_{lg}^E, \delta_{sm}^E)$ (%)	(90,0)	(10,0)	(0,10)	(0,0)	(90,0)
Change in Gov. Cost (%)	-1.85	-0.28	-0.32	0	-1.69
(2) Entry fixed at $\delta = 0$ -levels					
$(\delta_{lg}^F, \delta_{sm}^F)$ (%)	(0,10)	(10,0)	(15,0)	(22,0)	(0,15)
Change in Gov. Cost under fixed entry (%)	-0.08	-0.22	-1.50	-2.20	-0.50
Change in Gov. Cost under endogenous entry at $\delta^F$ (%)	0.65	-0.28	1.80	1.30	0.82
<i>Policy targeting 25% Small-Firm Award Rate</i>					
(1) Endogenous Entry					
$(\delta_{lg}^E, \delta_{sm}^E)$ (%)	(0,20)	(0,0)	(0,0)	(0,5)	(0,25)
Change in Gov. Cost (%)	2.10	0	0	0.50	2.30
(2) Entry fixed at $\delta = 0$ -levels					
$(\delta_{lg}^F, \delta_{sm}^F)$ (%)	(0,50)	(0,0)	(0,0)	(0,45)	(0,40)
Change in Gov. Cost under fixed entry (%)	0.52	0	0	1.90	0.94
Change in Gov. Cost under endogenous entry at $\delta^F$ (%)	7.00	0	0	6.80	6.70

Note: All costs reported in \$100,000s. Change in government cost computed relative to cost under  $\delta_{lg} = \delta_{sm} = 0$ . The change in government cost with endogenous entry in the fixed-participation panels measured the cost change that results from using  $\delta$  chosen under fixed entry, but allowing participation to respond to the discount. Project 1 is in the category of medium urban road-repair work, project 2 in small rural road-repair work, 3 in medium rural road-repair work, 4 in large urban road-repair work, and 5 in large urban other work.

Table 11: Effect of Discount Level on Aggregate Procurement Cost and Allocation of Work to Small Firms

$(\delta_{lg}, \delta_{sm})$ (%)	Agg. Cost to Gov't	Small Firms' % of Work
(25,0)	0.991	4.5
(15,0)	0.994	6.1
(10,0)	0.996	8.2
(5,0)	0.998	11.5
(0,0)	1.000	13.4
(0,5)	1.005	15.6
(0,10)	1.008	20.3
(0,15)	1.017	25.8
(0,25)	1.036	35.6

Note: The cost to the government is reported as a percent of the cost under no government intervention.

Table 12: Counterfactual Analysis of Preference Program: Comparison of Profit and Government Cost by Project Type

Project type	$\delta_{sm} = 0$			Avg Change (%) $\delta_{sm} = 0 \rightarrow 0.05$		
	E[ $\Pi_{sm}$ ]	E[ $\Pi_{lg}$ ]	E[win bid]	E[ $\Pi_{sm}$ ]	E[ $\Pi_{lg}$ ]	E[win bid]
Small, rural, rd repair / bridge	0.069	0.059	1.739	16.040	-14.188	0.351
Medium, rural, rd repair / bridge	0.148	0.111	4.225	11.089	-16.712	0.381
Large, rural, rd repair / bridge	0.155	0.201	6.497	10.680	-11.172	0.417
Small, urban, rd repair / bridge	0.043	0.063	1.745	13.753	-13.694	0.103
Medium, urban, rd repair / bridge	0.056	0.137	4.005	22.500	-10.542	0.242
Large, urban, rd repair / bridge	0.141	0.234	6.471	19.190	-14.970	0.470
Small, rural, other work	0.046	0.033	1.799	8.696	-18.485	0.234
Medium, rural, other work	0.120	0.139	4.836	25.000	-11.359	0.010
Small, urban, other work	0.042	0.064	1.998	13.735	-13.064	0.005
Medium, urban, other work	0.064	0.167	4.416	25.197	-10.365	0.168
Large, urban, other work	0.206	0.259	7.830	11.779	-11.292	-0.074

Note: E[ $\Pi_{sm}$ ] (E[ $\Pi_{lg}$ ]) denote small (large) firms' expected profits. E[win bid] denotes the expected winning bid, which measures the expected cost of procurement to the government. Expected profits and winning bid in \$100,000s.

Table 13: Counterfactual Analysis of Preference Program: Comparison of Predicted Entry by Project Type  
 Potential  $\delta_{sm} = 0$  Avg Change  $\delta_{sm} = 0 \rightarrow 0.05$

Project type	Bidders		Entry Prob				Actual Bidders				Avg Change $\delta_{sm} = 0 \rightarrow 0.05$			
	small	large	$p_{sm}$	$p_{lg}$	$n_{sm}$	$n_{lg}$	$n_{total}$	$\Delta p_{sm}$	$\Delta p_{lg}$	$\Delta p_{sm}$	$\Delta p_{lg}$	$\% \Delta p_{sm}$	$\% \Delta p_{lg}$	$\% \Delta n_{total}$
Small, rural, rd repair / bridge	3.500	3.500	0.703	0.661	2.420	2.310	4.730	0.034	-0.039	0.034	-0.039	4.895	-6.019	-0.532
Medium, rural, rd repair / bridge	3.714	4.571	0.620	0.586	2.024	2.564	4.588	0.051	-0.042	0.051	-0.042	8.105	-7.338	-0.791
Large, rural, rd repair / bridge	3.143	4.714	0.472	0.636	1.290	2.961	4.251	0.079	-0.050	0.079	-0.050	20.043	-8.552	-0.406
Small, urban, rd repair / bridge	4.250	5.000	0.543	0.648	2.186	2.964	5.150	0.032	-0.029	0.032	-0.029	6.587	-4.306	0.292
Medium, urban, rd repair / bridge	4.667	5.000	0.397	0.655	1.821	3.075	4.896	0.042	-0.033	0.042	-0.033	11.019	-5.111	0.765
Large, urban, rd repair / bridge	4.000	4.000	0.404	0.721	1.516	2.805	4.321	0.056	-0.049	0.056	-0.049	15.418	-6.873	0.383
Small, rural, other work	8.111	4.333	0.553	0.499	4.080	1.940	6.020	0.017	-0.041	0.017	-0.041	3.116	-8.433	-0.576
Medium, rural, other work	6.857	3.143	0.537	0.496	3.309	1.592	4.900	0.020	-0.050	0.020	-0.050	4.311	-10.272	-0.169
Small, urban, other work	6.200	4.300	0.529	0.655	2.968	2.698	5.666	0.027	-0.033	0.027	-0.033	5.279	-5.171	0.245
Medium, urban, other work	4.429	4.571	0.412	0.693	1.713	3.034	4.747	0.047	-0.031	0.047	-0.031	11.553	-4.707	1.092
Large, urban, other work	4.667	3.167	0.518	0.705	2.397	1.813	4.210	0.038	-0.056	0.038	-0.056	8.640	-10.602	0.818

Note:  $p_{sm}$  ( $p_{lg}$ ) denote small (large) firms' entry probabilities.  $n_{sm}$ ,  $n_{lg}$ , and  $n_{total}$  denote the expected number of small, large, and total bidders, respectively.  $\Delta$  denotes absolute changes in going from  $\delta_{sm} = 0$  to  $\delta_{sm} = 0.05$ , while  $\% \Delta$  denotes percent changes expressed in percentages.

Table 14: Counterfactual Analysis of Preference Program: Small-Firm Award Rate by Project Type

	$\delta_{sm} = 0$	$\delta_{sm} = 0.05$	
		Fixed Entry	Endog. Entry
Small, rural, rd repair / bridge	0.461	0.500	0.535
Medium, rural, rd repair / bridge	0.314	0.331	0.349
Large, rural, rd repair / bridge	0.219	0.229	0.243
Small, urban, rd repair / bridge	0.290	0.317	0.330
Medium, urban, rd repair / bridge	0.148	0.160	0.182
Large, urban, rd repair / bridge	0.091	0.101	0.117
Small, rural, other work	0.413	0.435	0.449
Medium, rural, other work	0.174	0.197	0.226
Small, urban, other work	0.312	0.339	0.355
Medium, urban, other work	0.128	0.141	0.161
Large, urban, other work	0.088	0.091	0.099

Note: Column 2 shows average small-firm probabilities of winning when entry is held fixed at the levels under  $\delta_{sm} = 0$ . Column 3 shows probabilities of winning when entry is allowed to adjust to the bid discount.

Table 15: Analysis of the Magnitude of Counterfactual Effects: Small Firms' Entry Response

	Coefficient	Std Error	Coefficient	Std Error
Small-Firm Project Cost	0.0450***	0.0091	0.0610***	0.0110
$\Delta$ Project Costs	0.0590***	0.0110	0.0800***	0.0140
Small-Firm Entry Cost	1.3700***	0.4200	0.7440	0.5800
Small-Firm Std. Dev. Entry Cost	-8.8200***	3.6300	-4.6600	3.9200
$\Delta$ Std. Dev. Entry Costs	-19.1600***	8.3100	-12.1900	8.5400
N <sup>o</sup> of small plan holders	-0.0020***	0.0006	-0.0030***	0.0006
N <sup>o</sup> of large plan holders	0.0051***	0.0008	0.0050***	0.0007
Prob. entry at $\delta_{sm}=0$ , SB			-0.0340***	0.1400

Note: Dependent variable is the change in the probability of entry for small firms.  $\Delta$  Costs denotes the difference between large and small costs.

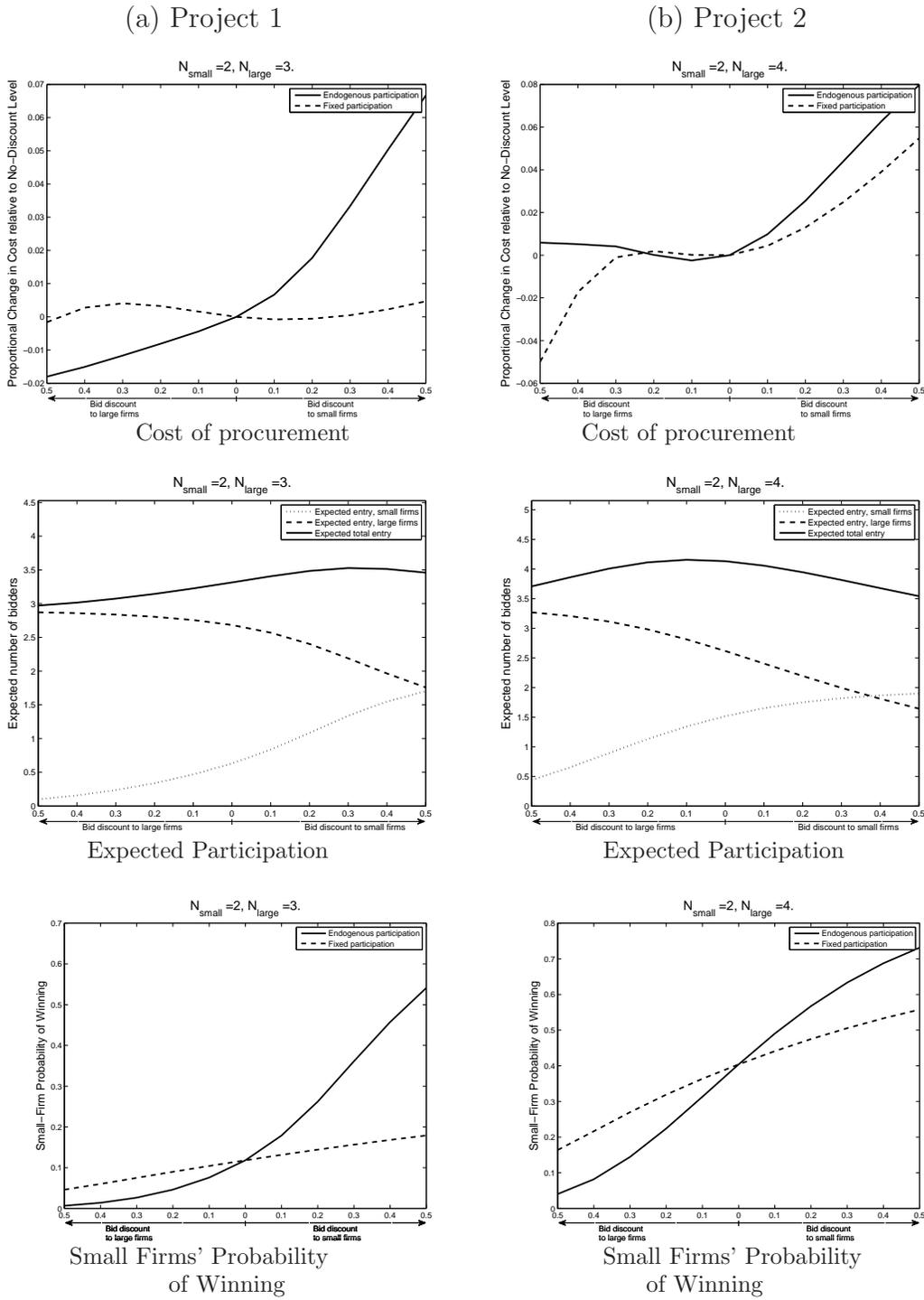
Table 16: Comparison of Alternative Subsidy Programs, Sample Projects

		Project					
		1	2	3	4	5	
Optimal subsidy							
	Gov't Cost	4.21	1.03	4.20	9.42	7.67	
	Subsidy						
		Small	-0.39	-0.09	-0.43	-0.62	-0.51
		Large	-0.35	-0.17	-0.31	-0.73	-0.33
	Exp. N <sup>o</sup> bidders						
		Small	0	1.64	0	1.82	0
		Large	1.84	0	1.89	0	2.51
Subsidy targeting small-firm probability of winning							
Case 1	$\Delta$ Gov't Cost (%)	-7.60	-8.40	-8.01	-8.50	-7.10	
	$\Delta$ Large-Firm Profit (%)	-25.60	-100	-33.20	-100	-37.20	
	Subsidy						
		Small	-0.31	-0.09	-0.39	-0.62	-0.28
		Large	-0.41	-0.17	-0.35	-0.73	-0.35
Case 2	$\Delta$ Gov't Cost (%)	-1.02	-1.27	-1.90	-2.50	-0.75	
	$\Delta$ Large-Firm Profit (%)	0	0	0	0	0	
	Subsidy						
		Small	0.01	-0.04	-0.11	0	0.02
		Large	0.10	0.10	0	-0.08	0.05
Benchmarks							
	Gov't Cost ( $\delta_{sm} = 0$ )	4.85	1.18	4.79	10.54	8.50	
	Gov't Cost (allocation target)	4.95	1.18	4.82	10.6	8.67	

Note: The subsidy amounts denote a subsidy payment to an individual firm was it to enter, with negative amounts denoting a tax. Government cost and subsidy payments in \$100,000s.

Case 1 shows subsidy levels that produce a 25% probability of winning for small firms. Case 2 displays subsidy levels which achieve that same small firm probability of winning as above but also constrain large-firm profits to be at least as large as those under a bid discount with a 25% small-firm probability of winning. The changes in costs and expected profits are computed relative to the respective magnitudes under the above bid discount.

Figure 2: Expected Cost and Entry under Alternative Bid Discounts, Sample Projects



Information on the projects in Table 10.

# Appendix (Not for Publication)

## A.1 Empirical Model: Specification, Estimation, and Identification Details

### A.1.1 Summary of the Empirical Model

At the time of announcement, a procurement project is characterized by a set of observable characteristics  $(x_j, z_j)$  and unobserved characteristic  $u_j$  where  $(x_j, u_j)$  and  $z_j$  denote characteristics that affect the distributions of project cost,  $F_c^k(\cdot|x_j, u_j)$ , and the distribution of entry costs,  $G_d^k(\cdot|z_j)$ , respectively. After the project is announced, firms identify themselves as potential bidders. Denote the numbers of potential bidders for project  $j$  by  $(N_{1j}, N_{2j})$ .

Each potential bidder  $i$  observes  $(x_j, u_j, z_j, N_{1j}, N_{2j})$  and his private entry cost realization,  $d_{ij}$ . On the basis of this information, a potential bidder makes the participation decision,  $I_{ij}(d_{ij}, x_j, u_j, z_j, N_{1j}, N_{2j})$ , where  $I_{ij} = 1$  if bidder  $i$  participates in the auction for project  $j$  and  $I_{ij} = 0$  otherwise. This participation strategy is characterized by a group-specific cut-off point on the support of the entry cost distribution,  $D_k(x_j, u_j, z_j, N_{1j}, N_{2j})$ . The equilibrium participation strategy is consistent with bidders' beliefs about the likelihood of their competitors' participation in the auction (and the observed participation probabilities):

$$p_k(x_j, u_j, z_j, N_{1j}, N_{2j}) = \int I_{ij}(d_{ij}, x_j, u_j, z_j, N_{1j}, N_{2j}) dG_d^k(d_{ij}|z_j).$$

After participation decisions are made, the numbers of actual bidders,  $(n_{1j}, n_{2j})$ , are realized. Conditional on  $(x_j, u_j, z_j, N_{1j}, N_{2j})$  the number of actual bidders,  $n_{kj}$ , is distributed according to a binomial distribution with a probability of success of  $p_k(x_j, u_j, z_j, N_{1j}, N_{2j})$  and  $N_{kj}$  trials.

Participating firms invest into discovering their project costs,  $c_{ij}$ , and prepare their bids,  $b_{ij} = \beta_{k(i)}(c_{ij}|n_{1j}, n_{2j}, F_c^1(\cdot|x_j, u_j), F_c^2(\cdot|x_j, u_j))$ , to be submitted to the auctioneer. Here  $\beta_k(\cdot|n_{1j}, n_{2j}, F_c^1, F_c^2)$  denotes the bidding strategy used by firms of group  $k$  in the auction for project  $j$ . The distribution of bids submitted for a project characterized by  $(x_j, u_j, n_{1j}, n_{2j})$  is given by

$$F_b^k(b|x_j, u_j, n_{1j}, n_{2j}) = F_c^k(\beta_k^{-1}(b|x_j, u_j, n_{1j}, n_{2j})|x_j, u_j).$$

### A.1.2 Assumptions

In this section, we list the assumptions that we impose on bidders' project and entry cost distributions that give rise to the empirical model in the paper. We assume that bidders' project

costs satisfy the following assumptions:

- (A-1)  $c_{ij} = \tilde{c}_{ij}u_j$ , where  $\tilde{c}_{ij}$  denotes the firm-specific component of bidders' costs and  $u_j$  the unobserved project heterogeneity component that is observed by all bidders, but unobserved by the econometrician.

Assumption (A-1) implies that  $\beta_{k(i)}(c_{ij}|x_j, u_j, n_{1j}, n_{2j}) = u_j\tilde{\beta}_{k(i)}(\tilde{c}_{ij}|x_j, n_{1j}, n_{2j})$  where  $\beta_k(\cdot|\cdot)$  and  $\tilde{\beta}_k(\cdot|\cdot)$  denote the group- $k$  bidding strategies associated with an arbitrary  $u_j$  and with  $u_j = 1$ , respectively. Thus,  $b_{ij} = \tilde{b}_{ij}u_j$  and  $\ln(b_{ij}) = \ln(\tilde{b}_{ij}) + \ln(u_j)$ .

- (A-2) The log of the unobserved heterogeneity component is distributed according to a normal distribution. The conditional expectation and variance of  $\ln(u_j)$  are  $E[\ln(u_j)|x_j, z_j, N_{1j}, N_{2j}] = 0$  and  $\text{Var}(\ln(u_j)|x_j, z_j, N_{1j}, N_{2j}) = \sigma_u^2$ .

- (A-3)  $\tilde{c}_{ij}$  are mutually independent conditionally on  $(x_j, N_{1j}, N_{2j})$  and independent of the unobserved project heterogeneity component,  $u_j$ :

$$\begin{aligned} F_{\tilde{c}|x,u}(\tilde{c}_{1j}, \dots, \tilde{c}_{N_{1j}+N_{2j},j}|x_j, u_j) &= \\ &= F_{\tilde{c}|x}(\tilde{c}_{1j}, \dots, \tilde{c}_{N_{1j}+N_{2j},j}|x_j) = \prod_{i=1}^{N_{1j}} F_{\tilde{c}}^1(\tilde{c}_{ij}|x_j) \prod_{i=1}^{N_{2j}} F_{\tilde{c}}^2(\tilde{c}_{ij}|x_j) \end{aligned}$$

for every  $(\tilde{c}_{1j}, \dots, \tilde{c}_{N_{1j}+N_{2j},j})$  that are points of continuity for  $F_{\tilde{c}}^1(\cdot|x_j)$  and  $F_{\tilde{c}}^2(\cdot|x_j)$ .

- (A-4) The log of the firm-specific bid component is distributed according to a normal distribution. The conditional expectation and variance of  $\ln(\tilde{b}_{ij})$  are given by:

$$\begin{aligned} E[\ln(\tilde{b}_{ij})|x_j, n_{1j}, n_{2j}] &= [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)} \\ \text{Var}[\ln(\tilde{b}_{ij})|x_j, n_{1j}, n_{2j}] &= (\exp(y_j' \eta_{k(i)}))^2 \end{aligned}$$

Here,  $y_j$  includes some of  $[x_j, n_{1j}, n_{2j}]$  and, possibly, their squares.

Further, we assume that bidders' entry costs satisfy the following assumptions:

- (A-5) Entry costs  $d_{ij}$  are distributed according to a normal distribution left-truncated at 0 with mean  $E[d_{ij}|z_j] = z_j' \gamma_k$  and a constant group-specific standard deviation  $\sigma_k^G$ . The conditional expectation and variance of  $d_{ij}$  are given by:

$$\begin{aligned} E[d_{ij}|x_j, z_j, N_{1j}, N_{2j}] &= z_j' \gamma_{k(i)} \\ \text{Var}[d_{ij}|x_j, z_j, N_{1j}, N_{2j}] &= \sigma_{k(i)}^2. \end{aligned}$$

(A-6) Entry costs  $d_{ij}$  are private information to firm  $i$  and are mutually independent conditionally on  $(x_j, z_j, N_{1j}, N_{2j})$  and independent of the unobserved project heterogeneity component,  $u_j$ :

$$G_{d|x,z,N_1,N_2}(d_{1j}, \dots, d_{N_1+N_2,j}|x_j, z_j, N_{1j}, N_{2j}, u_j) = \prod_{i=1}^{N_{1j}} G_1(d_{ij}|x_j, z_j, N_{1j}, N_{2j}) \prod_{i=1}^{N_{2j}} G_2(d_{ij}|x_j, z_j, N_{1j}, N_{2j}).$$

### A.1.3 Entry equilibrium and conditional distribution of $u_j$

Recall that a potential bidder  $i$ 's participation strategy is characterized by a group-specific cut-off point on the support of the entry cost distribution,  $D_k(x_j, u_j, z_j, N_{1j}, N_{2j})$ , resulting in equilibrium participation beliefs of  $p_k(x_j, u_j, z_j, N_{1j}, N_{2j})$ . Assumption (A-6) implies that conditional on  $(x_j, u_j, z_j, N_{1j}, N_{2j})$ , the number of actual bidders is distributed according to the product of two binomial distributions with probabilities of success given by  $p_k(x_j, u_j, z_j, N_{1j}, N_{2j})$  and  $N_{kj}$  trials,  $k = 1, 2$ :

$$\Pr(n_{1j} = k_1, n_{2j} = k_2|x_j, u_j, z_j, N_{1j}, N_{2j}) = C_{N_{1j}}^{k_1} C_{N_{2j}}^{k_2} p_1(\cdot)^{k_1} (1 - p_1(\cdot))^{N_{1j}-k_1} p_2(\cdot)^{k_2} (1 - p_2(\cdot))^{N_{2j}-k_2},$$

where  $C_N^k$  denotes the binomial coefficient of choosing  $k$  bidders out of  $N$  potential competitors,  $N!/(k!(N-k)!)$ .

An important and immediate consequence of the endogenously determined numbers of bidders,  $(n_{1j}, n_{2j})$ , is that

$$h(u_j|n_{1j}, n_{2j}) \neq h(u_j)$$

since the joint distribution of  $(n_{1j}, n_{2j})$  depends on  $u$ . Specifically,

$$\begin{aligned} h_u(u_j|n_{1j}, n_{2j}) &= \frac{\tilde{P}(u_j, n_{1j}, n_{2j})}{\tilde{P}(n_{1j}, n_{2j})} = \\ &= \frac{\sum_{N_{1j}, N_{2j}} \tilde{P}(n_{1j}, n_{2j}|N_{1j}, N_{2j}, u_j) h_u(u_j|N_{1j}, N_{2j})}{\int \sum_{N_{1j}, N_{2j}} \tilde{P}(n_{1j}, n_{2j}|N_{1j}, N_{2j}, u_j) h_u(u_j|N_{1j}, N_{2j}) du} = \\ &= \frac{\sum_{N_{1j}, N_{2j}} \tilde{P}(n_{1j}, n_{2j}|N_{1j}, N_{2j}, u_j) h_u(u_j)}{\int \sum_{N_{1j}, N_{2j}} \tilde{P}(n_{1j}, n_{2j}|N_{1j}, N_{2j}, u_j) h_u(u_j) du}. \end{aligned}$$

Here,  $\tilde{P}(u_j, n_{1j}, n_{2j})$  denotes the joint probability of  $(u_j, n_{1j}, n_{2j})$  and  $\tilde{P}(n_{1j}, n_{2j}|N_{1j}, N_{2j}, u_j)$  is the probability of  $(n_{1j}, n_{2j})$  conditional on  $(N_{1j}, N_{2j}, u_j)$ .

### A.1.4 Moment Conditions: Bid Distribution

In this section we use assumptions (A-1) through (A-4) to derive moment conditions to estimate the parameters of the bid distribution.

**First Order Moments.** Assumptions (A-1) and (A-4) imply that

$$\begin{aligned}\ln(\tilde{b}_{ij}) &= [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)} + \varepsilon_{ij} \\ \text{where } E[\varepsilon_{ij} | x_j, n_{1j}, n_{2j}] &= 0, \text{ and} \\ \ln(b_{ij}) &= [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)} + \ln(u_j) + \varepsilon_{ij}.\end{aligned}$$

Then

$$\begin{aligned}m_1 &= E[x_j'(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})] = \\ & E_{x, n_1, n_2}[E[x_j'(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)}) | x_j, n_{1j}, n_{2j}]] = \\ & E_{x, n_1, n_2}[E[x_j'(\ln(u_j) + \varepsilon_{ij}) | x_j, n_{1j}, n_{2j}]] = \\ & E_x[x_j' E[\ln(u_j) | x_j]] + E_{x, n_1, n_2}[x_j' E[\varepsilon_{ij} | x_j, n_1, n_2]] = 0.\end{aligned}$$

An empirical counterpart of this moment condition is

$$\hat{m}_1 = \frac{1}{\sum_{j=1}^J (n_{1j} + n_{2j})} \sum_{j=1}^J \sum_{i=1}^{n_{1j} + n_{2j}} [x_j'(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})].$$

Next,

$$\begin{aligned}m_2 &= E[n_{kj}(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})] = \\ & E_{x, n_1, n_2}[E[n_{kj}(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)}) | x_j, n_{1j}, n_{2j}]] = \\ & E_{x, n_1, n_2}[E[n_{kj}(\ln(u_j) + \varepsilon_{ij}) | x_j, n_{1j}, n_{2j}]] = \\ & E_{x, n_1, n_2}[E[n_{kj} \ln(u_j) | x_j, n_{1j}, n_{2j}]] + E[n_{kj} \varepsilon_{ij} | x_j, n_{1j}, n_{2j}]] = \\ & E_{x, N_1, N_2}[E[n_{kj} \ln(u_j) | x_j, N_{1j}, N_{2j}]] = \\ & \int \int \sum_{n_k=1}^{N_{kj}} \sum_{n_{-k}=1}^{N_{-kj}} n_k \ln(u_j) \Pr(n_k, n_{-k} | x_j, u_j, N_{kj}, N_{-kj}) h(u) du dF_{x, N_k, N_{-k}}(x_j, N_{kj}, N_{-kj}).\end{aligned}$$

Here, we use the notation  $-k$  to denote the opposite group, that is  $-k = 1$  if  $k = 2$  and  $-k = 2$  if  $k = 1$ . The last term arises because of the dependence of the distributions of the number of bidders on the realization of unobserved project heterogeneity.

An empirical counterpart of this moment condition is

$$\begin{aligned}\hat{m}_2 &= \frac{1}{\sum_{j=1}^J (n_{1j} + n_{2j})} \sum_{j=1}^J \sum_{i=1}^{n_{1j}+n_{2j}} (n_{kj}(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})) \\ &\quad - \frac{1}{nS} \sum_{s=1}^{ns} \sum_{n_k=1}^{N_{kj}} \sum_{n_{-k}=1}^{N_{-kj}} n_k \ln(u_s) \Pr(n_k, n_{-k} | x_j, u_s, N_{kj}, N_{-kj}),\end{aligned}$$

where we let  $u_s$  denote a draw from the unconditional distribution of  $u$ ,  $h(u)$ .

**Second Order Moments.** Let  $i_1$  and  $i_2$  indicate two bidders from groups  $k(i_1)$  and  $k(i_2)$ . Then

$$\begin{aligned}m_3 &= E[(\ln(b_{i_1j}) - \ln(b_{i_2j}))^2] = \\ &\quad E_{x,n_1,n_2}[E[(\varepsilon_{i_1j})^2 | x_j, n_1, n_2]] + E_{x,n_1,n_2}[E[(\varepsilon_{i_2j})^2 | x_j, n_1, n_2]] + \\ &\quad E_{x,n_1,n_2}[( [x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}) )^2] = \\ &\quad E_{x,n_1,n_2}[(\exp(y'_j \eta_{k(i_1)}))^2 + (\exp(y'_j \eta_{k(i_2)}))^2] + E_{x,n_1,n_2}[( [x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}) )^2]\end{aligned}$$

This simplifies to  $2E[(\exp(y'_j \eta_{k(i_1)}))^2]$  if  $k(i_1) = k(i_2)$ . Further, letting  $x_{jl}$  denote an element of  $x_j$ , we have that

$$\begin{aligned}m_4 &= E[x_{jl}(\ln(b_{i_1j}) - \ln(b_{i_2j}))^2] = \\ &\quad E_{x,n_1,n_2}[E[x_{jl}(\varepsilon_{i_1j} - \varepsilon_{i_2j})^2 | x_j, n_1, n_2]] + E_{x,n_1,n_2}[x_{jl}([x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}))^2] = \\ &\quad E_{x,n_1,n_2}[x_{jl}E[(\varepsilon_{i_1j})^2 + (\varepsilon_{i_2j})^2 | x_j, n_1, n_2]] + E_{x,n_1,n_2}[x_{jl}([x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}))^2] = \\ &\quad E_{x,n_1,n_2}[x_{jl}((\exp(y'_j \eta_{k(i_1)}))^2 + (\exp(y'_j \eta_{k(i_2)}))^2)] + E_{x,n_1,n_2}[x_{jl}([x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}))^2],\end{aligned}$$

which again simplifies to  $2E[x_{jl}(\exp(y'_j \eta_{k(i_1)}))^2]$  if  $k(i_1) = k(i_2)$ .

The empirical counterparts of these two moment conditions are given by:

$$\begin{aligned}\hat{m}_3 &= \frac{2}{\sum_{j=1}^J n_j(n_j + 1)} \sum_{j=1}^J \sum_{i_1=1}^{n_j} \sum_{i_2=1}^{n_j} ((\ln(b_{i_1j}) - \ln(b_{i_2j}))^2 - (\exp(y'_j \eta_{k(i_1)}))^2 \\ &\quad - (\exp(y'_j \eta_{k(i_2)}))^2 - ([x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}))^2) \\ \hat{m}_4 &= \frac{2}{\sum_{j=1}^J n_j(n_j + 1)} \sum_{j=1}^J \sum_{i_1=1}^{n_j} \sum_{i_2=1}^{n_j} (x_{jl}(\ln(b_{i_1j}) - \ln(b_{i_2j}))^2 \\ &\quad - x_{jl}((\exp(y'_j \eta_{k(i_1)}))^2 + (\exp(y'_j \eta_{k(i_2)}))^2) - x_{jl}([x_j, n_{1j}, n_{2j}]' (\alpha_{k(i_1)} - \alpha_{k(i_2)}))^2),\end{aligned}$$

with  $n_j = n_{1j} + n_{2j}$ .

$\hat{m}_3$  and  $\hat{m}_4$  specify an empirical moment condition for every parameter of the variance of  $\tilde{b}$  and, therefore, allow us to identify and consistently estimate all parameters  $\eta_k$ .

Finally, to estimate the variance of the unobserved heterogeneity component,  $\sigma_u^2$ , two possible moment conditions could be exploited. First, note that

$$\begin{aligned} m_{5a} &= E[(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})^2] = \\ & E_{x, n_1, n_2} [E[(\ln(u_j) + \varepsilon_{ij})^2 | x_j, n_1, n_2]] = \\ & E_{x, n_1, n_2} [E[(\ln(u_j))^2 | x_j, n_1, n_2] + E[(\varepsilon_{ij})^2 | x_j, n_1, n_2]] = \\ & \sigma_u^2 + E_{x, n_1, n_2} [(\exp(y_j' \eta_{k(i)}))^2]. \end{aligned}$$

Additionally, if  $k(i_1) \neq k(i_2)$ :

$$\begin{aligned} m_{5b} &= E[(\ln(b_{i_1j}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i_1)}) (\ln(b_{i_2j}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i_2)})] = \\ & E_{x, n_1, n_2} [E[(\ln(u_j) + \varepsilon_{i_1j}) (\ln(u_j) + \varepsilon_{i_2j}) | x_j, n_1, n_2]] = \\ & E_{x, n_1, n_2} [E[(\ln(u_j))^2 | x_j, n_1, n_2] + E[\varepsilon_{i_1j} \varepsilon_{i_2j} | x_j, n_1, n_2]] = \sigma_u^2. \end{aligned}$$

The empirical counterparts of these moment conditions are given by

$$\begin{aligned} \hat{m}_{5a} &= \frac{1}{\sum_{j=1}^J (n_{1j} + n_{2j})} \sum_{j=1}^J \sum_{i=1}^{n_{1j} + n_{2j}} ((\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})^2 - \sigma_u^2 - (\exp(y_j' \eta_{k(i)}))^2) \\ \hat{m}_{5b} &= \frac{1}{\sum_{j=1}^J \sum_{i_1=1}^{n_j} \sum_{i_2=i_1+1}^{n_j} I(k(i_1) \neq k(i_2))} \sum_{j=1}^J \sum_{i_1=1}^{n_j} \sum_{i_2=i_1+1}^{n_j} I(k(i_1) \neq k(i_2)) \\ & \left( (\ln(b_{i_1j}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i_1)}) (\ln(b_{i_2j}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i_2)}) - \sigma_u^2 \right). \end{aligned}$$

where  $I(\cdot)$  denotes an indicator function. For simplicity, we rely on condition  $\hat{m}_{5a}$  to estimate the variance of  $u$ .

**Higher Order Moments.** We exploit the properties of the normal distributions of  $\ln(u_j)$  and  $\varepsilon_{ij}$  to add higher-order moment conditions. For a normally distributed random variable  $X$  with mean  $\mu$  and standard deviation  $\sigma$ , the centered moment of order  $p$  is given by:

$$E[(X - \mu)^p] = I(p \text{ is even})(p - 1)!! \sigma^p,$$

where

$$(p - 1)!! = \frac{p!}{2^{\frac{p-2}{2}} \frac{p-2}{2}!} \text{ if } p \text{ is even.}$$

Applied to our setting, we have for  $p = 3, \dots, P$  that

$$\begin{aligned}
m_{5+p-2} &= E[(\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})^p] = \\
& E_{x, n_1, n_2} [E[(\ln(u_j) + \varepsilon_{ij})^p | x_j, n_1, n_2]] = \\
& E_{x, n_1, n_2} [E[\sum_{t=0}^p C_p^t \ln(u_j)^t \varepsilon_{ij}^{p-t}]] = \\
& E_{x, n_1, n_2} [\sum_{t=0}^p C_p^t E[\ln(u_j)^t] E[\varepsilon_{ij}^{p-t}]] = \\
& \sum_{t=0}^p C_p^t I(t \text{ is even}) I((p-t) \text{ is even}) (t-1)!! (p-t-1)!! \sigma_u^t E_{x, n_1, n_2} [(\exp(y_j' \eta_{k(i)}))^{p-t}].
\end{aligned}$$

The empirical counterparts of moments  $m_{5+p-2}$  are given by

$$\begin{aligned}
\hat{m}_{5+p-2} &= \frac{1}{\sum_{j=1}^J (n_{1j} + n_{2j})} \sum_{j=1}^J \sum_{i=1}^{n_{1j} + n_{2j}} \left( (\ln(b_{ij}) - [x_j, n_{1j}, n_{2j}]' \alpha_{k(i)})^p - \right. \\
& \left. \sum_{t=0}^p C_p^t I(t \text{ is even}) I((p-t) \text{ is even}) (t-1)!! (p-t-1)!! \sigma_u^t (\exp(y_j' \eta_{k(i)}))^{p-t} \right)
\end{aligned}$$

### A.1.5 Moments: Cost of Entry Distribution

In deriving the second set of moment conditions, we rely on the properties of the binomial distribution of the numbers of small and large bidders, conditional on observed and unobserved project characteristics and the numbers of potential bidders,  $N_{1j}$  and  $N_{2j}$ .

We exploit that

$$\begin{aligned}
E[n_{kj} | x_j, z_j, u_j, N_{1j}, N_{2j}] &= p_k(x_j, z_j, u_j, N_{1j}, N_{2j}) N_{kj} \\
E[n_{kj}^2 | x_j, z_j, u_j, N_{1j}, N_{2j}] &= p_k(x_j, z_j, u_j, N_{1j}, N_{2j}) (1 - p_k(x_j, z_j, u_j, N_{1j}, N_{2j})) N_{kj} \\
& \quad + N_{kj}^2 p_k^2(x_j, z_j, u_j, N_{1j}, N_{2j}),
\end{aligned}$$

where  $p_k(x_j, z_j, u_j, N_{1j}, N_{2j})$  denotes the group-specific equilibrium probabilities of participation. We derive separate moments for bidder groups,  $k$ , and project size categories,  $size_j$ . In our empirical specification, we consider three size categories with  $size_j = \{\text{small, medium, large}\}$ .

$$\begin{aligned}
m_{6+p-2}^{kl} &= E[n_{kj} | size_j = l] = \int \int p_k(x_j, z_j, u_j, N_{1j}, N_{2j}) N_{kj} h(u) du dF(x_j, z_j, N_{1j}, N_{2j} | size_j = l) \\
m_{7+p-2}^{kl} &= E[n_{kj}^2 | size_j = l] = \int \int (p_k(x_j, z_j, u_j, N_{1j}, N_{2j}) (1 - p_k(x_j, z_j, u_j, N_{1j}, N_{2j})) N_{kj} + \\
& \quad N_{kj}^2 p_k^2(x_j, z_j, u_j, N_{1j}, N_{2j})) h(u) du dF(x_j, z_j, N_{1j}, N_{2j} | size_j = l).
\end{aligned}$$

The empirical counterparts to these moment conditions are given by

$$\begin{aligned}\hat{m}_{6+P-2}^{kl} &= \frac{1}{\sum_{j=1}^J I(\text{size}_j = l)} \sum_{j=1}^J I(\text{size}_j = l) \left( n_{kj} - \frac{1}{nS} \sum_{s=1}^{ns} p_k(x_j, z_j, u_s, N_{1j}, N_{2j}) N_{kj} \right) \\ \hat{m}_{7+P-2}^{kl} &= \frac{1}{\sum_{j=1}^J I(\text{size}_j = l)} \sum_{j=1}^J I(\text{size}_j = l) \left( n_{kj}^2 - \frac{1}{nS} \sum_{s=1}^{ns} (p_k(x_j, z_j, u_s, N_{1j}, N_{2j}) (1 - \right. \\ &\quad \left. p_k(x_j, z_j, u_s, N_{1j}, N_{2j})) N_{kj} + p_k^2(x_j, z_j, u_s, N_{1j}, N_{2j}) N_{kj}^2) \right).\end{aligned}$$

**Higher Order Moments.** We further include third and fourth order moments of the binomial distribution of  $n_k$ . These are given by:

$$\begin{aligned}m_{8+P-2}^{kl} &= E[n_{kj}^3 | \text{size}_j = l] = \int \int (N_{kj} p_k (1 - 3p_k + 3N_{kj} p_k + 2p_k^2 - 3N_{kj} p_k^2 + \\ &\quad + N_{kj}^2 p_k^2)) h(u) du dF(x_j, z_j, N_{1j}, N_{2j} | \text{size}_j = l) \\ m_{9+P-2}^{kl} &= E[n_{kj}^4 | \text{size}_j = l] = \int \int (N_{kj} p_k (1 - 7p_k + 7N_{kj} p_k + 12p_k^2 - 18N_{kj} p_k^2 + \\ &\quad + 6N_{kj}^2 p_k^2 - 6p_k^3 + 11N_{kj} p_k^3 - 6N_{kj}^2 p_k^3 + N_{kj}^3 p_k^3)) h(u) du dF(x_j, z_j, N_{1j}, N_{2j} | \text{size}_j = l).\end{aligned}$$

The empirical counterparts to these moment conditions are given by

$$\begin{aligned}\hat{m}_{8+P-2}^{kl} &= \frac{1}{\sum_{j=1}^J I(\text{size}_j = l)} \sum_{j=1}^J I(\text{size}_j = l) \left( n_{kj}^3 - \frac{1}{nS} \sum_{s=1}^{ns} (N_{kj} p_k (1 - 3p_k + \right. \\ &\quad \left. + 3N_{kj} p_k + 2p_k^2 - 3N_{kj} p_k^2 + N_{kj}^2 p_k^2)) \right) \\ \hat{m}_{9+P-2}^{kl} &= \frac{1}{\sum_{j=1}^J I(\text{size}_j = l)} \sum_{j=1}^J I(\text{size}_j = l) \left( n_{kj}^4 - \frac{1}{nS} \sum_{s=1}^{ns} (N_{kj} p_k (1 - 7p_k + \right. \\ &\quad \left. 7N_{kj} p_k + 12p_k^2 - 18N_{kj} p_k^2 + 6N_{kj}^2 p_k^2 - 6p_k^3 + 11N_{kj} p_k^3 - 6N_{kj}^2 p_k^3 + N_{kj}^3 p_k^3)) \right).\end{aligned}$$

### A.1.6 Econometric identification of the project cost distribution

In this section, we derive three properties of the joint distributions of the firm-specific cost and bid components. They form the basis for the nonparametric identification of  $F_{\tilde{c}|x}(\cdot)$  in the presence of unobserved heterogeneity given our model with endogeneous entry. The properties imply that the results in Krasnokutskaya (2009a) can be applied in this environment.

First, recall that in our model potential bidders do not observe the realizations of their firm-specific cost component when deciding whether to participate in the market. Therefore, the following property holds.

**Property 1.** *There is no selection into participation on the firm-specific cost component. That is, firm-specific cost components are independent of the numbers of*

*bidders conditional on project characteristics:*

$$F_{\tilde{c}|x,u,n_1,n_2}(\tilde{c}_{1j}, \dots, \tilde{c}_{N_{1j}+N_{2j},j}|x_j, u_j, n_{1j}, n_{2j}) = F_{\tilde{c}|x,u}(\tilde{c}_{1j}, \dots, \tilde{c}_{N_{1j}+N_{2j},j}|x_j, u_j).$$

At the time when bids are constructed, all participants learn the numbers of actual bidders by group,  $(n_{1j}, n_{2j})$ , and incorporate them into the bids. As a result, firm-specific bid components depend on  $(n_{1j}, n_{2j})$ . Property 1, together with assumption (A-3), implies:

**Property 2.** *Individual bid components are mutually independent conditionally on  $(x_j, n_{1j}, n_{2j})$ :*

$$F_{\tilde{b}|x,n_1,n_2}(\tilde{b}_{1j}, \dots, \tilde{b}_{n_{1j}+n_{2j}}|x_j, n_{1j}, n_{2j}) = \prod_{i=1}^{n_{1j}+n_{2j}} F_{\tilde{b}|x,n_1,n_2}(\tilde{b}_{ij}|x_j, n_{1j}, n_{2j})$$

Proof:

$$\begin{aligned} F_{\tilde{b}|x,n_1,n_2}(\tilde{b}_{1j}, \dots, \tilde{b}_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j}) &= \\ F_{\tilde{c}|x,n_1,n_2}(\tilde{\beta}_{k(1)}^{-1}(b_{1j}|x_j, n_{1j}, n_{2j}), \dots, \tilde{\beta}_{k(n_1+n_2)}^{-1}(b_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j})|x_j, n_{1j}, n_{2j}) &= \\ F_{\tilde{c}|x}(\tilde{\beta}_{k(1)}^{-1}(b_{1j}|x_j, n_{1j}, n_{2j}), \dots, \tilde{\beta}_{k(n_1+n_2)}^{-1}(b_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j})|x_j) &= \\ \prod_{i=1}^{n_1} F_{\tilde{c}|x}^1(\tilde{\beta}_1^{-1}(b_{ij}|x_j, n_{1j}, n_{2j})|x_j) \prod_{i=1}^{n_2} F_{\tilde{c}|x}^2(\tilde{\beta}_2^{-1}(b_{ij}|x_j, n_{1j}, n_{2j})|x_j) &= \\ \prod_{i=1}^{n_1} F_{\tilde{b}|x,n_1,n_2}^1(\tilde{b}_{ij}|x_j, n_{1j}, n_{2j}) \prod_{i=1}^{n_2} F_{\tilde{b}|x,n_1,n_2}^2(\tilde{b}_{ij}|x_j, n_{1j}, n_{2j}). \end{aligned}$$

End of Proof

Here, the first and last equalities hold due to the monotonicity of the firm-specific bidding function  $\tilde{\beta}_k(\cdot|x, n_1, n_2)$ , while Property 1 implies the second equality because of the lack of selection on project cost among entrants. Finally, assumption (A-3) of mutual independence of individual cost components implies the third equality.

Assumptions (A-1), which implies that the firm-specific bidding function  $\tilde{\beta}_k(\cdot|x, n_1, n_2)$  does not depend on  $u$ , and (A-3), together with the monotonicity of  $\tilde{\beta}_k(\cdot|x, n_1, n_2)$ , yield

**Property 3.** *Individual bid components are independent of the unobserved auction heterogeneity component conditionally on  $(x, n_1, n_2)$ :*

$$F_{\tilde{b}|x,n_1,n_2,u}(\tilde{b}_{1j}, \dots, \tilde{b}_{n_{1j}+n_{2j},j}|x_j, n_{1j}, n_{2j}, u_j) = F_{\tilde{b}|x,n_1,n_2}(\tilde{b}_{1j}, \dots, \tilde{b}_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j})$$

Proof:

$$\begin{aligned}
F_{\tilde{b}|x,n_1,n_2,u}(\tilde{b}_{1j}, \dots, \tilde{b}_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j}, u_j) &= \\
F_{\tilde{c}|x,n_1,n_2,u}(\tilde{\beta}_{k(1)}^{-1}(\tilde{b}_{1j}|x_j, n_{1j}, n_{2j}), \dots, \tilde{\beta}_{k(n_1+n_2)}^{-1}(\tilde{b}_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j})|x_j, n_{1j}, n_{2j}, u_j) &= \\
F_{\tilde{c}|x}(\tilde{\beta}_{k(1)}^{-1}(\tilde{b}_{1j}|x_j, n_{1j}, n_{2j}), \dots, \tilde{\beta}_{k(n_1+n_2)}^{-1}(\tilde{b}_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j})|x_j) &= \\
F_{\tilde{b}|x,n_1,n_2}(\tilde{b}_{1j}, \dots, \tilde{b}_{(n_1+n_2)j}|x_j, n_{1j}, n_{2j}). &
\end{aligned}$$

End of proof.

### A.1.7 Econometric identification of the entry cost distribution

This section studies the nonparametric identification of the distribution of entry costs,  $G(\cdot|z)$ , in the presence of unobserved project heterogeneity assuming that  $H(\cdot)$  and  $F(\cdot|x)$  are identified. The full identification proof is developed in Krasnokutskaya (2009b). We summarize the argument here for completeness. We focus on the case of symmetric bidders to simplify exposition.

We assume that  $x_j = [x_{1j}, x_{2j}]$  such that the variables in  $x_{2j}$  are part of  $z_j$  whereas the variables in  $x_{1j}$  are not. In this section we always condition on  $z_j$  and, therefore, suppress  $(z_j, x_{2j})$  going forward.

We employ the following notations. We denote bidder  $i$ 's expected profit conditional on  $x_1$ , the number of bidders,  $n$ , and  $u$  by

$$u\pi_0(x_1, n) = u \int (\tilde{\beta}(\tilde{c}) - \tilde{c})(1 - F(\tilde{c}|x_1))^{n-1} f(\tilde{c}|x_1) d\tilde{c}$$

We assemble profit levels that realize for every possible number of competitors of bidder  $i$ ,  $n_c = 0, \dots, N$  if there are  $N + 1$  potential bidders, into the vector

$$u\pi_0(x_1) = (u\pi_0(x_1, 1), u\pi_0(x_1, 2), \dots, u\pi_0(x_1, N + 1)).$$

It is possible to show that under fairly natural assumptions,

$$\pi_0(x_1, 1) > \pi_0(x_1, 2) > \dots > \pi_0(x_1, N + 1).$$

Here we just assume that.

If  $p$  is an individual bidder's probability of entering the market, then the vector of probabilities for the number of competitors participating in the auction is given by:

$$p_N = ((1 - p)^N, C_N^1 p(1 - p)^{N-1}, \dots, p^N).$$

where  $C_N^k$  again denotes the binomial coefficient of choosing  $k$  bidders out of  $N$  potential competitors,  $N!/(k!(N-k)!)$ .

We denote the ex-ante expected profit of an individual potential bidder from participating by

$$u\bar{\pi}_0(x_1, p) = up'_N\pi_0(x_1).$$

where the firm integrates out the number of competitors using its beliefs over their participation.

The entry threshold that determines the marginal entrant is then given by:

$$D(x_1, u, p) = \begin{cases} u\bar{\pi}_0(x_1, p) & \underline{d} \leq u\bar{\pi}_0(x_1, p) \leq \bar{d} \\ \underline{d} & u\bar{\pi}_0(x_1, p) \leq \underline{d} \\ \bar{d} & \bar{d} \leq u\bar{\pi}_0(x_1, p), \end{cases}$$

and  $p$  is a solution to

$$p = G(D(x_1, u, p)),$$

making it a function of  $x_1$  and  $u$ ,  $p(x_1, u)$ . Finally, the probability of entry at  $x_1$  is given by

$$p(x_1) = \int p(x_1, u)h(u)du.$$

We proceed under the following assumptions:

(B-1) There exists at least one variable  $x_1$  that affects bidders' project costs but not their entry costs.

(B-2) The distribution of entry costs has a bounded support,  $\text{supp}(G(\cdot|z)) = [\underline{d}(z), \bar{d}(z)]$ .

(B-3) The distribution of unobserved heterogeneity has a bounded support,  $\text{supp}(H(\cdot)) = [\underline{u}, \bar{u}]$ .

We make assumptions (B-2) and (B-3) to simplify exposition; they can be relaxed.

(B-4) The expected profit,  $u\pi_0(x_1, n)$ , is continuous in  $x_1$ .

Assumption (B-4) can be obtained easily with minimal assumptions on the primitives. For transparency reasons, we choose to state it here as an assumption.

(B-5) For every  $r$  such that  $\underline{d} \leq r \leq \bar{d}$  there exist  $x_1^*$  and  $x_1^{**}$  that satisfy  $\bar{u}\pi_0(x_1^*, 1) = r$  and  $\bar{u}\pi_0(x_1^{**}, N+1) = r$ .

(B-6)  $G(\cdot)$  is an absolutely continuous distribution.

The condition  $(B - 5)$  is essentially a “full support” type of condition. The proof in the case of a discrete distribution follows very similar steps.

We begin by establishing that the ex-ante expected profit,  $u\bar{\pi}_0(x_1, p)$ , declines in  $p$ , before turning to the proof of identification of  $G(\cdot)$ .

**Proposition 1.** *Ex-ante expected profit is strictly decreasing in the individual probability of participation.*

Proof:

Here we show that  $\bar{\pi}_0(x_1, p)$  is decreasing in  $p$ . From this, Proposition 1 follows immediately.

$$\bar{\pi}_0(p) = (1 - p)^N \pi_0(1) + p^N \pi_0(N + 1) + \sum_{n=1}^{N-1} C_N^k p^n (1 - p)^{N-n} \pi_0(n + 1)$$

Then

$$\begin{aligned} \bar{\pi}'_0(p) &= -N(1 - p)^{N-1} \pi_0(1) + Np^{N-1} \pi_0(N + 1) \\ &+ \sum_{n=1}^{N-1} C_N^m (np^{n-1} (1 - p)^{N-n} - (N - n)p^n (1 - p)^{N-1-n}) \pi_0(n + 1) \end{aligned}$$

First, we transform the terms in the sum.

$$\begin{aligned} \sum_{n=1}^{N-1} C_N^m np^{n-1} (1 - p)^{N-n} \pi_0(n + 1) &= \\ N \sum_{l=0}^{N-2} C_{N-1}^l p^l (1 - p)^{N-1-l} \pi_0(l + 2), \end{aligned}$$

where we perform the change of variables  $l = n - 1$ . Similarly,

$$\begin{aligned} \sum_{n=1}^{N-1} C_N^m (N - n)p^n (1 - p)^{N-1-n} \pi_0(n + 1) &= \\ N \sum_{n=1}^{N-1} C_{N-1}^n p^n (1 - p)^{N-1-n} \pi_0(n + 1). \end{aligned}$$

Substituting the transformed expressions into  $\bar{\pi}'_0(p)$  results in:

$$\begin{aligned} \bar{\pi}'_0(p) &= N \left( (1 - p)^{N-1} \pi_0(2) - (1 - p)^{N-1} \pi_0(1) + \right. \\ & p^{N-1} \pi_0(N + 1) - p^{N-1} \pi_0(N) + \\ & \left. \sum_{l=1}^{N-1} C_{N-1}^l p^l (1 - p)^{N-1-l} (\pi_0(l + 2) - \pi_0(l + 1)) \right). \end{aligned}$$

Since we assume that  $\pi_0(x_1, 1) > \pi_0(x_1, 2) > \dots > \pi_0(x_1, N + 1)$ , it follows that  $\bar{\pi}'_0(p) < 0$ .

End of Proof

Note that the boundary of the support of  $G(\cdot)$  can be identified as follows:

$$\begin{aligned}\underline{d} &= \bar{u}\bar{\pi}_0(x_1^0, 0) \\ \bar{d} &= \underline{u}\bar{\pi}_0(x_1^1, 1),\end{aligned}$$

where  $x_1^0$  is the smallest  $x_1$  such that there is entry into the market and  $x_1^1$  is the smallest  $x_1$  such that all potential entrants enter.

Next, we establish main result of this section. Consider the following problem:

$$p(x_1) = \int G(D(x_1, u))h(u)du \text{ for all } x_1$$

such that

$$D(x_1, u) = u\bar{\pi}_0(x_1, G(D(x_1, u))) \text{ when } \underline{d} \leq u\bar{\pi}_0(x_1, G(D(x_1, u))) \leq \bar{d}.$$

If data are generated by the model described in our paper, then the distribution of entry costs  $G(\cdot)$  satisfies the restrictions imposed by this problem and thus solves it for every  $x_1$ . The result below shows that  $G(\cdot)$  is the only solution to this problem.

**Theorem 1.** *The cumulative distribution function  $G(\cdot)$  is identified.*

Proof:

Suppose that there exist two solutions  $G_1(\cdot)$  and  $G_2(\cdot)$  such that  $G_1(d) \neq G_2(d)$  for some  $d$ . Since the distributions are continuous, there exists for each point  $d'$  with  $G_1(d') \neq G_2(d')$  an open interval around  $d'$  such that for every point in this interval  $G_1 \neq G_2$ . Since the supports of  $G_1$  and  $G_2$  are bounded, there is a finite number of such intervals.<sup>28</sup> Finally, notice that within each of the open intervals either  $G_1 < G_2$  or  $G_1 > G_2$  by the continuity of the distributions.

It is then possible to find such an open subset with unequal distributions closest to the low end of the support. Let us denote it by  $(d_a, d_b)$ . Two distinct cases are possible; case 1:  $d_a = \underline{d}$  and case 2:  $d_a \neq \underline{d}$ . First consider case 1.

**Case 1.** Without loss of generality assume that  $G_1(d) > G_2(d)$  on  $(\underline{d}, d_b)$ . Consider a point  $d_1 \in (\underline{d}, d_b)$ .

(a) There exists a point  $x_1^*$  such that  $\bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1$ .

This follows from Property 1 that  $u\bar{\pi}_0(x_1, p)$  is decreasing in  $p$ , which implies that

$$\bar{u}\bar{\pi}_0(x_1, G_1(d_1)) > \bar{u}\bar{\pi}_0(x_1, 1).$$

---

<sup>28</sup>Indeed, it is possible to choose a closed interval inside each of these open sets. Since the support is bounded, the collection of these closed intervals is compact. The original open intervals create a countable open cover of this set. Therefore, there is a finite subset of this cover that still covers the compact set. From the construction of the compact set, it is clear that the original open cover is finite.

Notice also that  $\bar{\pi}_0(x_1, 1) = \pi(x_1, N + 1)$ . Assumption (B-5) implies that there exist  $x'_1$  such that  $\bar{u}\pi(x'_1, N + 1) \geq d_1$  and, therefore,  $\bar{u}\bar{\pi}_0(x'_1, G_1(d_1)) \geq d_1$ . Similarly,

$$\bar{u}\bar{\pi}_0(x_1, G_1(d_1)) \leq \bar{u}\bar{\pi}_0(x_1, 0) = \bar{u}\pi(x_1, 1)$$

and there exists  $x''_1$  such that  $\bar{u}\pi(x''_1, 1) \leq d_1$  and, therefore,  $\bar{u}\bar{\pi}_0(x''_1, G_1(d_1)) \leq d_1$ . By continuity of  $\bar{\pi}_0(\cdot, G_1(d_1))$  in  $x_1$ , there thus exists  $x_1^*$  such that  $\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1$ .

(b) There exists  $d_2$  such that  $\bar{u}\bar{\pi}_0(x_1^*, G_2(d_2)) = d_2$ .

Indeed, as before,

$$\bar{u}\bar{\pi}_0(x_1^*, G_2(\underline{d})) > \bar{u}\pi_0(x_1^*, 1) > \underline{d}$$

since

$$\underline{d} < d_1 = \bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) < \bar{u}\pi_0(x_1^*, 1).$$

Similarly,

$$\bar{u}\bar{\pi}_0(x_1^*, G_2(\bar{d})) < \bar{u}\pi_0(x_1^*, N + 1) < \bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1 < \bar{d}.$$

Since the ex-ante expected profit,  $\bar{u}\bar{\pi}_0(x_1^*, G_2(d))$ , is continuous in  $d$ , there exists  $d_2 \in [\underline{d}, \bar{d}]$  such that  $\bar{u}\bar{\pi}_0(x_1^*, G_2(d_2)) = d_2$ .

(c) The following holds:  $d_2 > d_1$  and  $G_2(d_2) < G_1(d_1)$ .

This follows again from the ex-ante expected profit,  $\bar{u}\bar{\pi}_0(x_1^*, p)$ , being decreasing in  $p$ , which implies

$$\bar{u}\bar{\pi}_0(x_1^*, G_2(d_1)) > \bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1.$$

Therefore,  $d_1 \neq d_2$ . Moreover, for any  $d < d_1$ :

$$\bar{u}\bar{\pi}_0(x_1^*, G_2(d)) > \bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1 > d.$$

Thus,  $d_2 > d_1$ . Further,

$$\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1/\bar{u}$$

$$\bar{\pi}_0(x_1^*, G_2(d_2)) = d_2/\bar{u}.$$

Therefore,  $\bar{\pi}_0(x_1^*, G_1(d_1)) < \bar{\pi}_0(x_1^*, G_2(d_2))$ . This implies that  $G_1(d_1) > G_2(d_2)$  since the ex-ante expected profit is decreasing in the probability of participation.

(d) Define  $u^* = \underline{d}/\pi(x_1^*, 1)$ . Then for all  $u \in [u^*, \bar{u}]$ ,  $D(u, x_1^*, G_i)$  exists with  $D(u, x_1^*, G_1) < D(u, x_1^*, G_2)$ , while  $G_1(D(u, x_1^*, G_1)) > G_2(D(u, x_1^*, G_2))$ .

Indeed, for an arbitrary  $u \in (u^*, \bar{u}]$ :

$$u\bar{\pi}_0(x_1^*, G_i(\underline{d})) = u\pi_0(x_1^*, 1) > u^*\pi_0(x_1^*, 1) = \underline{d}.$$

Similarly,

$$u\bar{\pi}_0(x_1^*, G_i(\bar{d})) = u\pi_0(x_1^*, N + 1) < \bar{u}\pi_0(x_1^*, N + 1) < \bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1 < \bar{d}.$$

Therefore, by continuity of the ex-ante profit, interior solutions,  $\underline{d} < D(u, x_1^*, G_i) < \bar{d}$ , exist for every  $u \in (u^*, \bar{u}]$  whereas  $D(u^*, x_1^*, G_i) = \underline{d}$  by definition. Finally, point (c) implies that  $G_1(D(u, x_1^*, G_1)) > G_2(D(u, x_1^*, G_2))$  for  $u \in (u^*, \bar{u}]$ .

(e) Finally,

$$p_1(x_1^*, G_1) = \int_{u^*}^{\bar{u}} G_1(D(u, x_1^*, G_1))h(u)du$$

$$p_2(x_1^*, G_2) = \int_{u^*}^{\bar{u}} G_2(D(u, x_1^*, G_2))h(u)du.$$

Therefore,  $p_1(x_1^*, G_1) > p_2(x_1^*, G_2)$ . Thus, both distributions cannot be consistent with the data.

**Case 2.** Now consider  $d_a \neq \underline{d}$ . Since  $(d_a, d_b)$  is an open interval closest to  $\underline{d}$  with  $G_1 > G_2$ ,  $G_1(d_a) = G_2(d_a)$ , but  $G_1(d) > G_2(d)$  for  $d \in (d_a, d_b)$ . Choose  $d_1 \in (d_a, d_b)$ . Find  $x_1^*$  such that the solution of  $\bar{u}\bar{\pi}_0(x_1^*, G_1(d_1)) = d_1$ . After that the steps are the same as in Case 1.

End of proof.

## A.2 Further discussion of the optimal policy results for project 1

The government's cost-minimizing policy for projects such as sample projects 1 and 5 is to choose a sufficiently high large-firm discount rate such that small firms respond by not participating in the auction. Here we provide further details on the intuition behind this result, using sample project 1 as an example.

First note that for this project, the marginal effect of large-firm entry on the cost of procurement is higher than that of small-firm entry. Table A-1 considers the effects on the cost of procurement from a marginal change in the probability of participation. We compare the response in the government's cost to increasing each type's probability of participation by 1 percentage point above its equilibrium participation probability. For discounts of 0%, 10% to small bidders, and 10% to large bidders, the government's cost responds more to increases in large rather than small-firm participation.

Figure A-1 illustrates similar responses conditioning on particular combinations of bidders, suggesting that the marginal effect of an additional large bidder on the cost of procurement is higher than that of a small bidder. For example, moving from the cost profile corresponding to one large and one small bidder to the one with two large bidders and one small bidder entails uniformly a larger decline in cost than a move to the profile with two small bidders and one large bidder. The larger marginal effect of large-firm participation on the cost of procurement suggests then that the government benefits when the presence of large participants increases. Note that these are out-of-equilibrium exercises.

Figures 1 and 2 document similar effects for equilibria associated with different discount levels. Thus, the middle panel of Figure 2 shows that the large-firm probability of participation increases (while the small-firm probability of participation decreases) with the discount level given to large bidders. This effect is accompanied by a decrease in the government's cost of procurement (top panel of Figures 1 and 2). Thus, in equilibrium, the government cost decreases as the large-firm presence increases even though the small-firm presence (and the total number of bidders) decreases at the same time.

The desired high levels of large-firm participation,  $p_{lg}$ , may not be attainable in the unconstrained equilibrium. In Figure A-2 below, we illustrate the participation equilibrium in the absence of intervention ( $\delta = 0$ ) using optimal participation schedules for the two groups of bidders. The optimal participation schedule shows the proportion of bidders by group  $k$  who optimally choose to participate for a given level of participation by the other group of bidders,  $p_{-k}$ .

Recall that equilibrium participation decisions in our model are determined by the relative sizes of the ex-ante expected variable profits and entry costs. To sustain a large-firm probability of participation of  $p_{lg}$  in equilibrium, each large participant needs to earn ex-ante variable profit of at least  $G_{lg}^{-1}(p_{lg})$ . The remaining two panels in Figure A-2 display these ex-ante variable profit

levels earned under each best-response participation strategy by small (middle panel) and large (bottom panel) firms.

The top panel suggests that for high large-firm participation, e. g.  $p_{lg} = 0.95$ , to reflect optimal participation behavior in the unconstrained equilibrium, small-firm participation needs to be very low ( $p_{sm} = 0.10$ ). However, at a level of  $p_{lg} = 0.95$  it is optimal for small firms to participate at a higher level ( $p_{sm} = 0.25$ ) than needed to sustain  $p_{lg} = 0.95$ . Therefore, such high  $p_{lg}$ -levels do not occur in the unconstrained equilibrium. The small-bidder level of participation that is optimal is still quite low, however, reflecting the entry by small firms with very low entry costs only.

In the unconstrained equilibrium, the large-firm participation probability is limited to  $p_{lg} = 0.894$  (see Table A-1 and top panel of Figure A-2), with associated expected profit of 0.368. The middle panel of Figure A-2 suggests that given this equilibrium large-firm participation, the expected ex-ante variable profit levels earned by small firms are only 0.140, consistent with the low amount of entry of only  $p_{sm} = 0.315$  we see from this group in equilibrium. At this level of small-firm entry, large firms do not earn sufficient variable profit to sustain additional entry beyond  $p_{lg} = 0.894$ . Thus, the presence of even a small amount of small-firm entry is sufficient to deter additional large-firm entry.

For increased large-firm participation to be an equilibrium outcome, the group's expected profit needs to rise. Since the expected price (bid) declines as  $p_{lg}$  increases, these profit gains have to be achieved through increases in the probability of winning. A bid discount artificially increases the benefitting group's probability of winning and thus enables the desired increases in large-firm profitability and participation.

Small firms, which have much higher project cost in this example than large firms, have to bid aggressively even without a bid discount, as suggested by the level and flatness of their expected profit profile under optimal participation. In response to a large-firm discount and the associated further reduction in their probability of winning, small firms choose increasingly not to enter. This does not, however, yield price increases in this particular example because of the substantial presence of large firms in the market that counters the incentives generated by the discount to bid less aggressively. Figure 2 illustrates these equilibrium responses to the discount.

Note also that project 1 is characterized by both strong differences in the groups' cost distributions and a tightness difference in the markets for small and large bidders, with  $N_{small} = 2$  and  $N_{large} = 3$ . Figure A-2 reflects the net effect of these cost differences and market tightness differences. The market tightness effect manifests itself in the following properties of the plotted schedules:

1. The large-firm optimal participation schedule is flatter than the small-firm optimal participation schedule.
2. Full participation of large bidders is never achieved. Even with  $p_{sm} = 0$ ,  $p_{lg} = 0.97$ ,

corresponding to 2.91 bidders. At the same time  $p_{sm}$  is close to 1, or the equivalent of two bidders, for  $p_{lg}$  as low as  $p_{lg} = 0.2$ .

3. The profit schedule for small firms (middle panel of Figure A-2) is steeper than that for large firms (bottom panel) since a 1 percentage point increase in the proportion of large participants corresponds to an increase by 0.03 bidders, instead of an increase by 0.02 small bidders as in the case of the large-firm profit schedule.
4. The small-firm variable profit given optimal participation at  $p_{lg} = 0$  is much higher than the large-firm variable profit given optimal participation at  $p_{sm} = 0$ .

These effects disappear when we equalize market tightness across groups of bidders as in Figure A-3 where we replot the optimal participation schedule and associated expected profit levels for the case where  $N_{small} = N_{large} = 2$ .

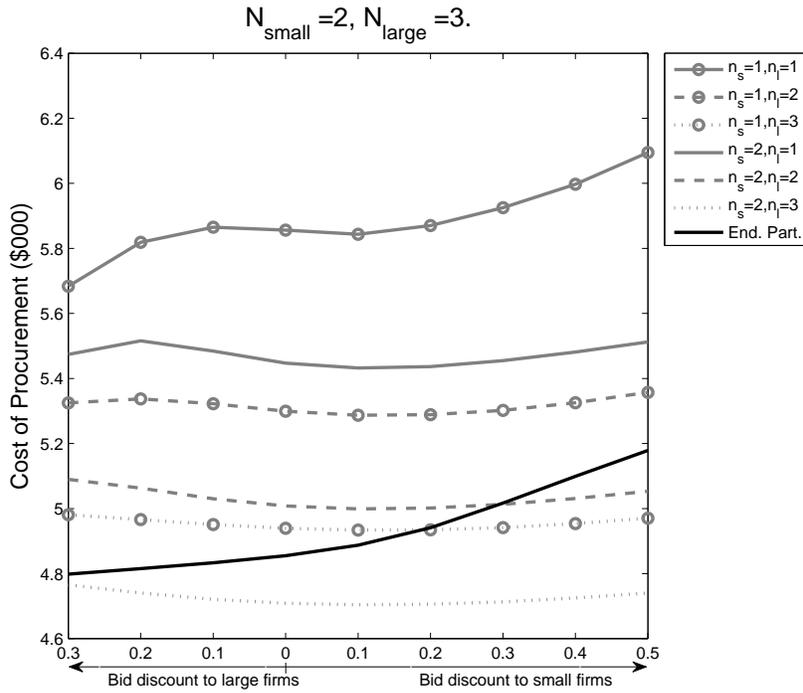
To summarize, the discount allows the government to artificially increase the large-firm probability of winning, thereby increasing large bidders' profitability and inducing higher entry by large bidders. Under the firms' cost structures for project 1, this lowers the price paid by the government.

Table A-1: Marginal Effect of Entry by Group on Expected Cost  
to the Government, Sample Project 1

	$\delta_{\text{small}}$	$\delta_{\text{large}}$	Entry Prob.		Change, Entry Prob.		Gov't cost	Small Prob. of Winning
			$p_{\text{small}}$	$p_{\text{large}}$	$\Delta p_{\text{small}}$	$\Delta p_{\text{large}}$		
No discount								
(1)	0	0	0.315	0.894			4.858	0.118
(2)	0	0	0.325	0.894	0.01	0	4.857	0.118
(3)	0	0	0.315	0.904	0	0.01	4.853	0.115
10% discount to small firms								
(1)	0.1	0	0.418	0.857			4.893	0.179
(2)	0.1	0	0.428	0.857	0.01	0	4.893	0.179
(3)	0.1	0	0.418	0.867	0	0.01	4.889	0.175
10% discount to large firms								
(1)	0	0.1	0.234	0.919			4.835	0.076
(2)	0	0.1	0.244	0.919	0.01	0	4.834	0.076
(3)	0	0.1	0.234	0.929	0	0.01	4.830	0.073

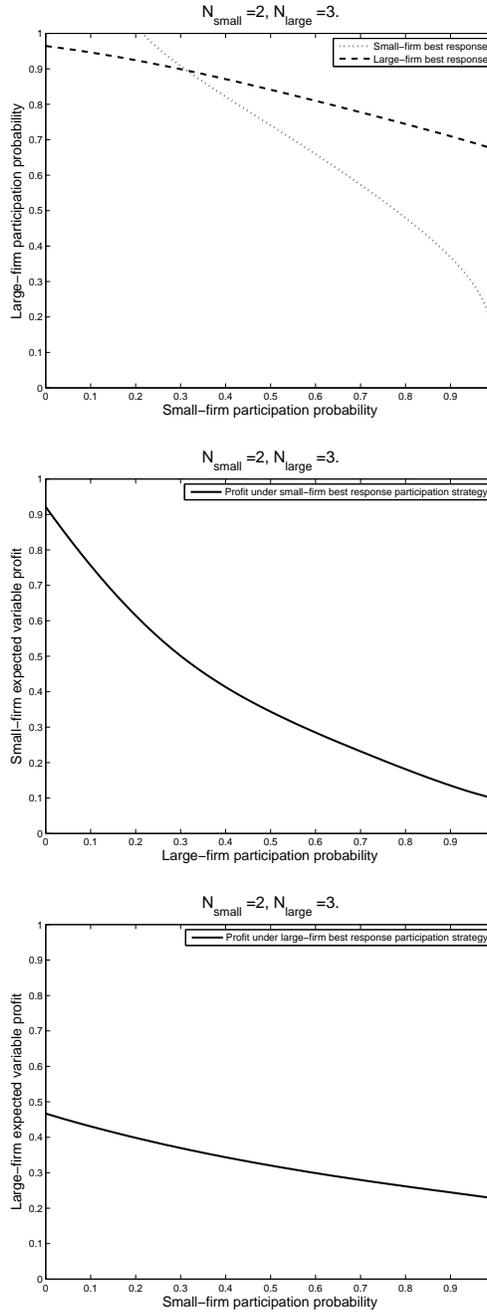
Note: Case (1) depicts equilibrium probabilities of entry, cost to the government, and the small-firm probability of winning for project 1 under the chosen discount level. Cases (2) and (3) consider the impact of increasing the probability of entry for small and large firms by 1 percentage point above the equilibrium, respectively.

Figure A-1: Expected Cost under Fixed and Endogenous Participation, Sample Project 1



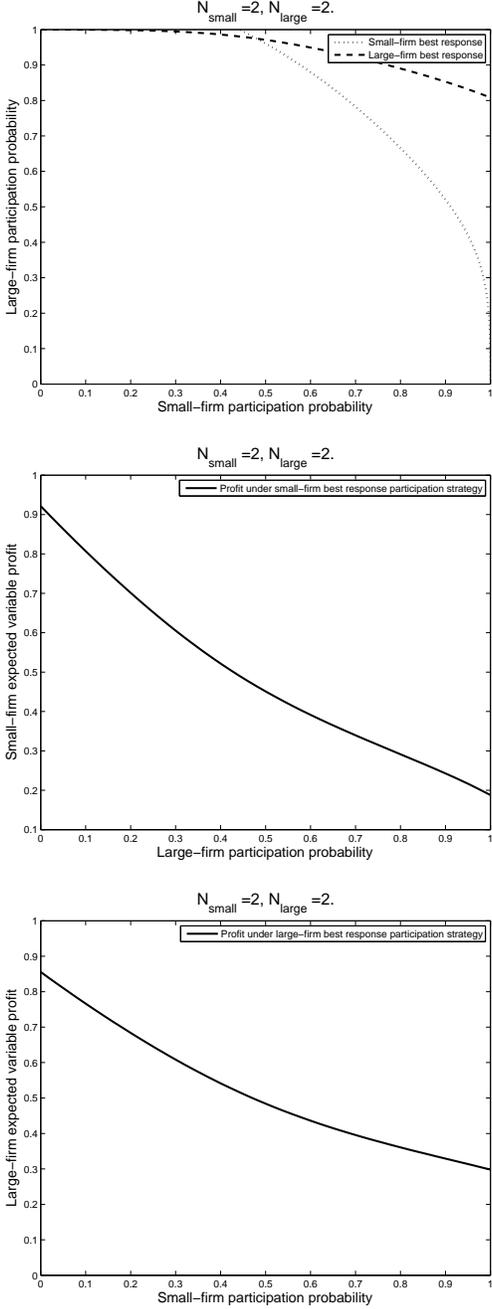
Note: the figure compares the relationship between discount levels and the cost to the government under alternative assumptions on the competitive environment. We depict in gray profiles that arise when regardless of discount, we hold the number of bidders fixed at one of six possible bidder combinations that could arise with 2 small and 3 large potential entrants. We depict in black the profile under endogenous entry. It is steeper than the other profiles, reflecting that as the discount increases, it becomes more likely that the number of bidders is composed of a larger number of small bidders and a lower number of large bidders obtain. These competitive environments correspond to the higher gray profiles.

Figure A-2: Equilibrium under No Bid Discount, Project 1



Note: the top panel depicts the optimal participation schedules for the two groups of bidders when  $\delta = 0$ . An optimal participation schedule reflects the proportion of bidders from group  $k$  who optimally choose to participate for a given level of participation by the other group,  $p_{-k}$ . The bottom two panels show the expected variable profit from participation excluding bid preparation costs associated with optimal participation level for a given level of participation by the other group,  $p_{-k}$ .

Figure A-3: Equilibrium under No Bid Discount, Project 1,  $N_{small} = N_{large} = 2$



Note: this figure replicates the analysis in Figure A-2, but changes the number of potential large entrants to be the same as potential small entry by setting  $N_{small} = N_{large} = 2$ .

### A.3 Additional Figures and Tables

Table A-2: Comparison of Entry Probabilities, Estimation and Simulation Analysis

Project type	Entry Probabilities			
	Estimation		Simulation	
	Small Firms	Large Firms	Small Firms	Large Firms
Small, rural, rd repair / bridge	0.7287	0.5909	0.7377	0.6211
Medium, rural, rd repair / bridge	0.7164	0.5429	0.7018	0.5795
Large, rural, rd repair / bridge	0.6643	0.5492	0.6487	0.5816
Small, urban, rd repair / bridge	0.6196	0.5617	0.6277	0.5924
Medium, urban, rd repair / bridge	0.5590	0.5726	0.5818	0.5996
Large, urban, rd repair / bridge	0.5373	0.5875	0.5624	0.6110
Small, rural, other work	0.5422	0.5546	0.5636	0.5850
Medium, rural, other work	0.5442	0.5409	0.5630	0.5688
Small, urban, other work	0.5362	0.5434	0.5591	0.5730
Medium, urban, other work	0.5223	0.5559	0.5503	0.5810
Large, urban, other work	0.5220	0.5621	0.5507	0.5858

Note: the table compares predicted probabilities of entry generated by our simulation routine with  $\delta = 0.05$  and by the estimation procedure. The small discrepancy in the predicted probabilities of entry arises because in the simulation routine, we have to trim the support of the project cost distribution to ensure that the density is sufficiently far away from zero.

Figure A-4: Predicted and Actual Bid Residuals

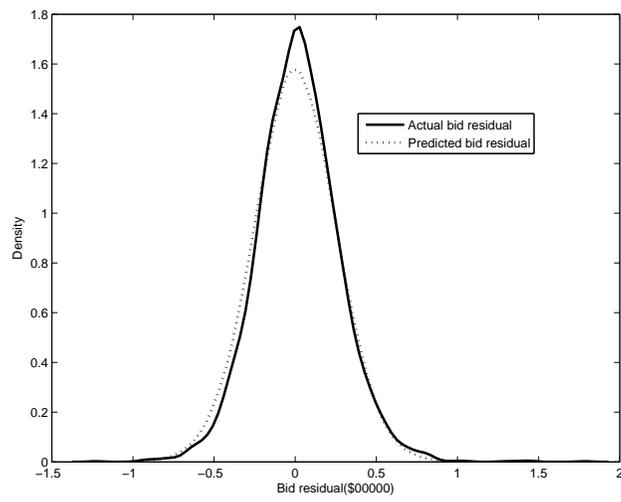
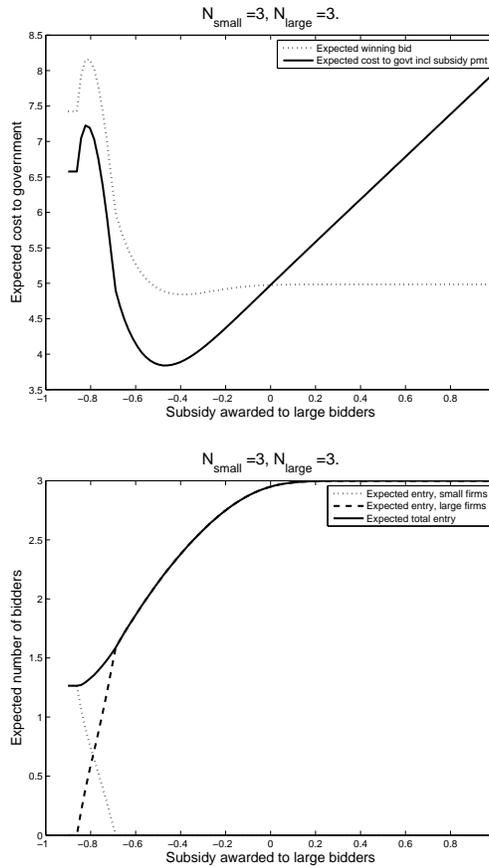


Figure A-5: Expected Cost and Entry under Alternative Subsidy Levels, Sample Project 3



Note: the panels display the cost to the government and entry as a function of the subsidy to large bidders, holding the subsidy for small bidders fixed at the cost-minimizing tax level. Negative subsidy levels correspond to taxes. The expected winning bid reflects the following interplay of participation and bidding decisions. For subsidy levels below -0.85, only small firms are in the market and pay their optimal subsidy, resulting in a constant winning bid. As the tax charged to large bidders starts declining, large bidders begin entering the market and initially replace small bidders. For this particular project, large bidders are less efficient, pushing up the winning bid. Once taxes fall below -0.8, entering large bidders more than displace non-participating small firms, resulting in an overall increase in the number of bidders. This causes the winning bid to begin declining again.