Appendix (Not for Publication)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, $F_c^k(.|x_j, u_j)$, and the distribution of entry costs, $G_d^k(.|z_j)$, 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 decison, $I_{ij}(d_{ij}, x_j, u_j, z_j, N_{1j}, N_{2j})$, 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):

$$p_k(x_j, u_j, z_j, N_{1j}, N_{2j}) = \int I_{ij}(d_{ij}, x_j, u_j, z_j, N_1, N_2) dG_d^k(d_{ij}|z_j).$$

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, $b_{ij} = \beta_{k(i)}(c_{ij}|n_{1j}, n_{2j}, F_c^1(.|x_j, u_j), F_c^2(.|x_j, u_j))$, to be submitted to the auctioneer. Here $\beta_k(.|n_{1j}, n_{2j}, F_c^1, F_c^2)$ 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 $\beta_k(.|.)$ and $\tilde{\beta}_k(.|.)$ 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 $\mathrm{E}[\ln(u_j)|x_j, z_j, N_{1j}, N_{2j}] = 0$ and $\mathrm{Var}(\ln(u_j)|x_j, z_j, N_{1j}, N_{2j}) = \sigma_u^2$.

(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:

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

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

(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:

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

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 σ_k^G. The conditional expectation and variance of d_{ij} are given by:

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

(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:

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

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,

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

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

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

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

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

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

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

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:

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

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 η_k.

Finally, to estimate the variance of the unobserved heterogeneity component, σ_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 $k(i_1) \neq k(i_2)$:

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

The empirical counterparts of these moment conditions are given by

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

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 ε_{ij} to add higher-order moment conditions. For a normally distributed random variable X with mean μ and standard deviation σ, 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 $p = 3, ..., P$ that

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

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

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

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 $size_j = \{\text{small,medium,large}\}$.

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

The empirical counterparts to these moment conditions are given by

$$\hat{m}^{kl}_{6+P-2} = \frac{1}{\sum_{j=1}^{J} I(size_j = l)} \sum_{j=1}^{J} I(size_j = l)\Big(n_{kj} - \frac{1}{ns}\sum_{s=1}^{ns} p_k(x_j, z_j, u_s, N_{1j}, N_{2j})N_{kj}\Big)$$

$$\hat{m}^{kl}_{7+P-2} = \frac{1}{\sum_{j=1}^{J} I(size_j = l)} \sum_{j=1}^{J} I(size_j = l)\Big(n^2_{kj} - \frac{1}{ns}\sum_{s=1}^{ns} \big(p_k(x_j, z_j, u_s, N_{1j}, N_{2j})(1- p_k(x_j, z_j, u_s, N_{1j}, N_{2j}))N_{kj} + p^2_k(x_j, z_j, u_s, N_{1j}, N_{2j})N^2_{kj}\big)\Big).$$

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

$$m^{kl}_{8+P-2} = E[n^3_{kj}|size_j = l] = \int\int (N_{kj}p_k(1 - 3p_k + 3N_{kj}p_k + 2p^2_k - 3N_{kj}p^2_k + + N^2_{kj}p^2_k))\, h(u)du\ dF(x_j, z_j, N_{1j}, N_{2j}|size_j = l)$$

$$m^{kl}_{9+P-2} = E[n^4_{kj}|size_j = l] = \int\int (N_{kj}p_k(1 - 7p_k + 7N_{kj}p_k + 12p^2_k - 18N_{kj}p^2_k + + 6N^2_{kj}p^2_k - 6p^3_k + 11N_{kj}p^3_k - 6N^2_{kj}p^3_k + N^3_{kj}p^3_k))\, h(u)du\ dF(x_j, z_j, N_{1j}, N_{2j}|size_j = l).$$

The empirical counterparts to these moment conditions are given by

$$\hat{m}^{kl}_{8+P-2} = \frac{1}{\sum_{j=1}^{J} I(size_j = l)} \sum_{j=1}^{J} I(size_j = l)\Big(n^3_{kj} - \frac{1}{ns}\sum_{s=1}^{ns} \big(N_{kj}p_k(1 - 3p_k + + 3N_{kj}p_k + 2p^2_k - 3N_{kj}p^2_k + N^2_{kj}p^2_k)\big)\Big)$$

$$\hat{m}^{kl}_{9+P-2} = \frac{1}{\sum_{j=1}^{J} I(size_j = l)} \sum_{j=1}^{J} I(size_j = l)\Big(n^4_{kj} - \frac{1}{ns}\sum_{s=1}^{ns} \big(N_{kj}p_k(1 - 7p_k + 7N_{kj}p_k + 12p^2_k - 18N_{kj}p^2_k + 6N^2_{kj}p^2_k - 6p^3_k + 11N_{kj}p^3_k - 6N^2_{kj}p^3_k + N^3_{kj}p^3_k)\big)\Big).$$

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 $F_{\tilde{c}|x}(.)$ 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:*

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

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})*:*

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

Proof:

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

End of Proof

Here, the first and last equalities hold due to the monotonicity of the firm-specific bidding function $\tilde{\beta}_k(.|x,n_1,n_2)$, 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 $\tilde{\beta}_k(.|x,n_1,n_2)$ does not depend on u, and (A-3), together with the monotonicity of $\tilde{\beta}_k(.|x,n_1,n_2)$, yield

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

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

Proof:

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

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, $G(.|z)$, in the presence of unobserved project heterogeneity assuming that $H(.)$ and $F(.|x)$ 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, $n_c = 0, ..., N$ if there are $N + 1$ potential bidders, into the vector

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

It is possible to show that under fairly natural assumptions,

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

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:

$$p_N = ((1 - p)^N, C^1_N p(1 - p)^{N-1}, ..., p^N).$$

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, $supp(G(.|z)) = [\underline{d}(z), \bar{d}(z)]$.

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

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 $\underline{d} \leq r \leq \underline{d}$ 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(.) 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(.).

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

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

First, we transform the terms in the sum.

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

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

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

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

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

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

Note that the boundary of the support of G(.) can be identified as follows:

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

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 $G(.)$ satisfies the restrictions imposed by this problem and thus solves it for every x_1. The result below shows that $G(.)$ is the only solution to this problem.

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

Proof:

Suppose that there exist two solutions $G_1(.)$ and $G_2(.)$ 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.[28] 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

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

[28] 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 $\bar{\pi}_0(., G_1(d_1))$ 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,

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

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

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

Similarly,

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

Therefore, by continuity of the ex-ante profit, interior solutions, $\underline{d} < D(u, x_1^*, G_i) < \overline{d}$, exist for every $u \in (u^*, \overline{u}]$ 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 $u \in (u^*, \overline{u}]$.
(e) Finally,

$$p_1(x_1^*, G_1) = \int_{u^*}^{\overline{u}} G_1(D(u, x_1^*, G_1))h(u)du$$
$$p_2(x_1^*, G_2) = \int_{u^*}^{\overline{u}} G_2(D(u, x_1^*, G_2))h(u)du.$$

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 $N_{small} = N_{large} = 2$.

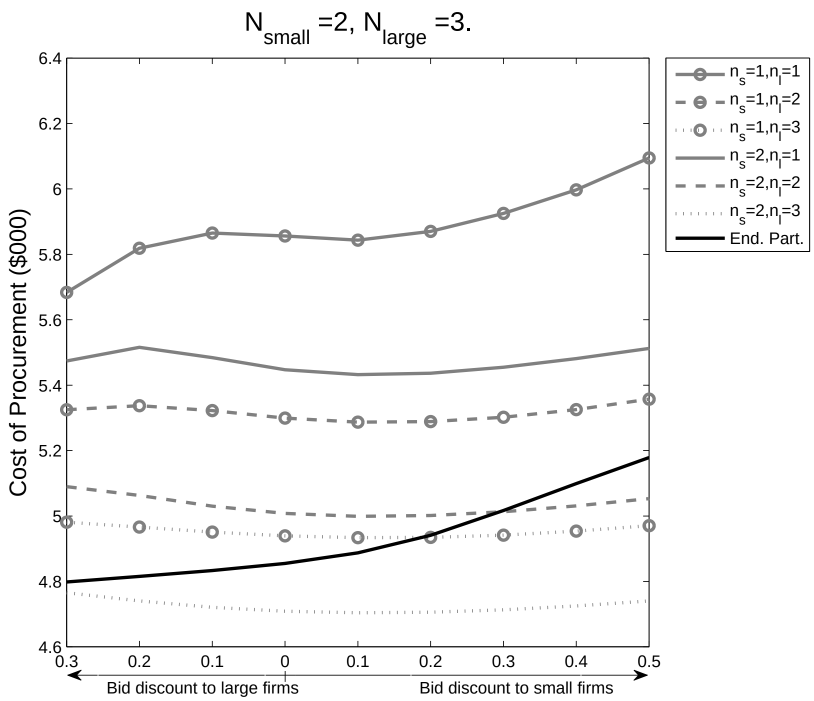
To summarize, the discount allows the government to artificially increase the large-firm probability of winning, thereby increasing large bidders' profitability and inducing higher entry by large bidders. Under the firms' cost structures for project 1, this lowers the price paid by the government.

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

			Entry Prob.		Change, Entry Prob.			
	δ_{small}	δ_{large}	p_{small}	p_{large}	Δp_{small}	Δp_{large}	Gov't cost	Small Prob. of Winning
				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

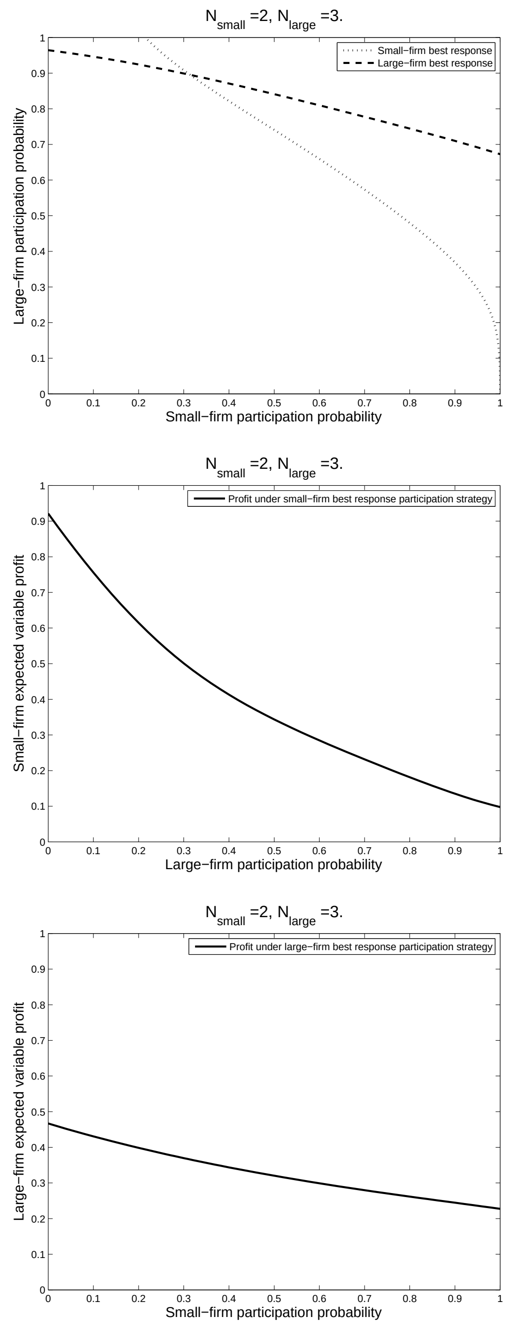
Note: Case (1) depicts equilibrium probabilities of entry, cost to the government, and the small-firm probability of winning for project 1 under the chosen discount level. Cases (2) and (3) consider the impact of increasing the probability of entry for small and large firms by 1 percentage point above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

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



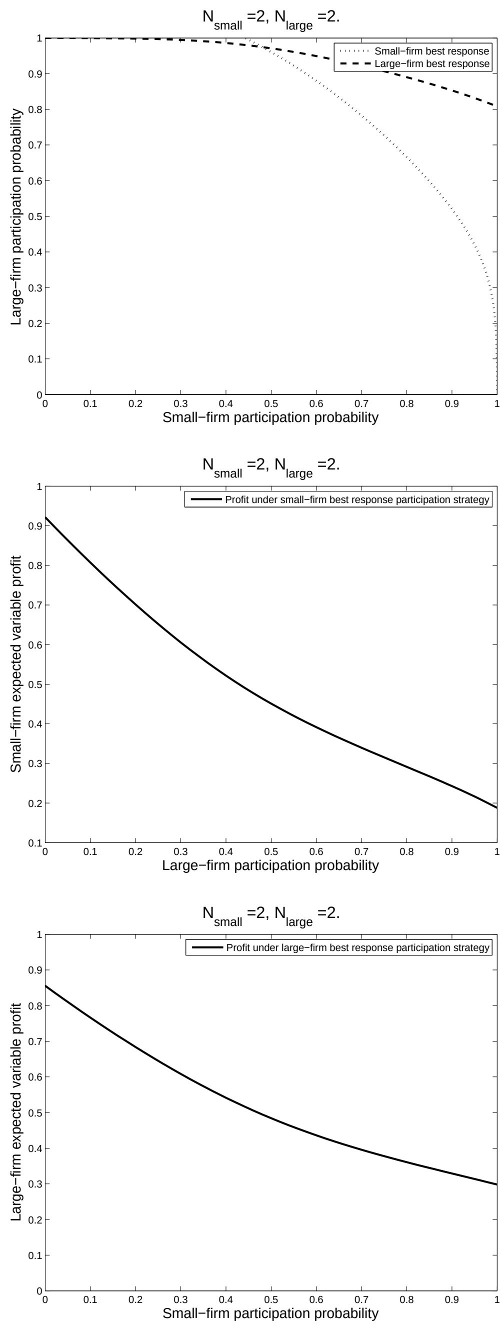
Note: the figure compares the relationship between discount levels and the cost to the government under alternative assumptions on the competitive environment. We depict in gray profiles that arise when regardless of discount, we hold the number of bidders fixed at one of six possible bidder combinations 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

	Entry Probabilities			
	Estimation		Simulation	
Project type	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

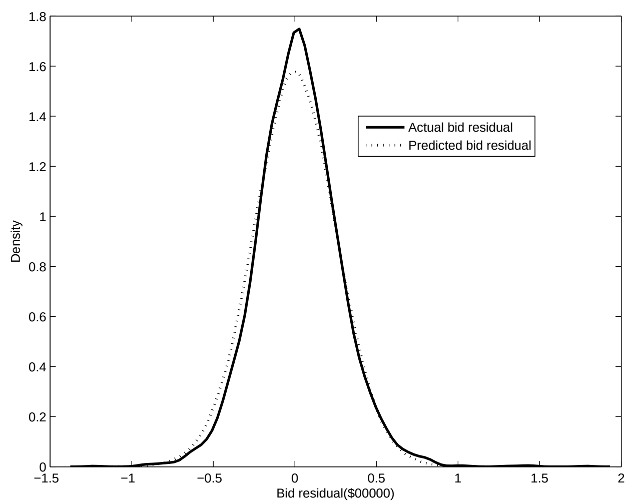
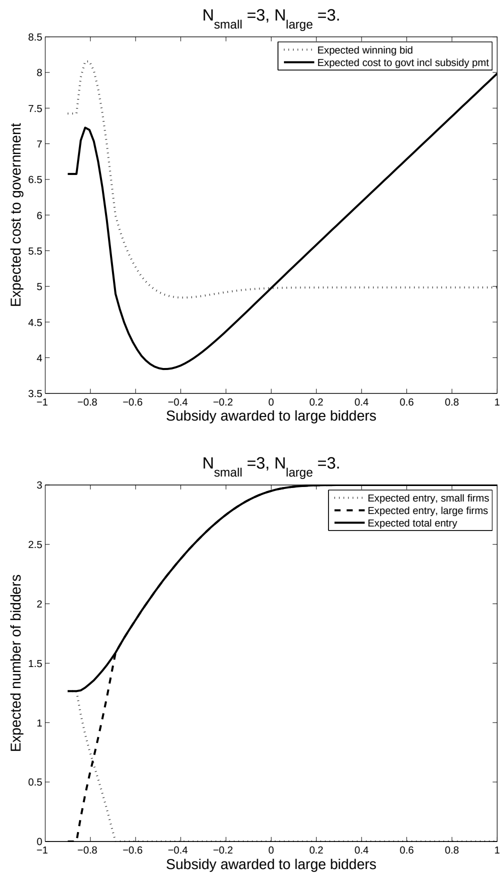


Figure A-5: Expected Cost and Entry under Alternative Subsidy Levels, Sample Project 3



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.