Small Business Set-asides in Procurement Auctions: An Empirical Analysis

by

Jun Nakabayashi
November 2009

Small Business Set-asides in Procurement Auctions:
An Empirical Analysis

Jun NAKABAYASHI[*]

First Draft: March 2009
This Version: December 2010

Abstract

As part of public procurement, many governments adopt small business programs to provide contract opportunities for businesses often with preferences for firms operated by members of groups designated as disadvantaged. The redistribution arising from such programs, however, can introduce significant added costs to government procurement budgets. In this paper, the extent to which small business set-asides increase government procurement costs is examined. The estimates employ data on Japanese public construction projects, where approximately half of the procurement budget is set aside for small and medium enterprises (SMEs). Applying a positive relationship between profitability and firm size obtained by the non-parametric estimation of asymmetric first-price auctions with affiliated private values, a counterfactual analysis is undertaken to demonstrate that approximately 30 percent of SMEs would exit the procurement market if set-asides were to be removed. Surprisingly, the resulting lack of competition would increase government procurement costs more than it would offset the production cost inefficiency.

Key words: procurement auctions, small business set-asides, nonparametric estimation
JEL classification: D44, H23, H57, L74

1   Introduction

As part of public procurement, many governments adopt a program for encouraging small businesses to participate in procurement auctions.[1] In the United States, the Small Business Administration suggests almost all agencies in the federal government spend an overall proportion of 23 percent of their procurement budget with small firms.[2] For some departments, such as the Department of Transportation, the expenditure for small firms in 2005 was approximately $670 million, which accounted for 45 percent of the total annual expenditure. A similar program is seen in public procurement in Japan. For the central government, the spending target to small and medium-sized enterprises

[1]Bannock (1981) identifies the United States, Germany, Switzerland, and Japan as the countries in which governments strongly support small businesses.

[2]The Federal Acquisition Regulation (FAR), Subpart 19.5. states that if the contracts are no more than $100,000, they are automatically reserved exclusively for small business concerns and shall be set aside for small businesses.

(SMEs)[3] was 50.1 percent in 2007.[4] As in the case of the United States federal government, the goal is achieved almost every year.

Reserving contracts to small businesses restricts competition, which can result in the market being inefficient. Nevertheless, some of the theoretical literature on auctions predicts that set-asides may not hurt procurement budgets as much as had been anticipated. For instance, Ayres and Cramton (1996) investigate the affirmative action program in FCC spectrum auctions and observe that setting aside some contracts for disadvantaged bidders enhances competition among advantaged bidders, which can compensate the efficiency loss. Milgrom (2004) points out the analog of set-asides for price discrimination conducted by a multi-market monopoly seller.

Nonetheless, empirical literature in this field is somewhat lacking. In particular, to the best of our knowledge, there is no existing work that estimates the extent that set-asides hurt government budgets.

This paper is the first attempt to investigate the effect of set-asides on government budgets by using structural estimation techniques. In particular, the degree to which government procurement costs are changed and the extent to which SMEs' entry into procurement markets is promoted by small business set-asides are quantified.

The data set used in this research is from Japanese public procurement auctions for civil engineering works conducted by the Ministry of Land, Infrastructure and Transportation (MLIT). From April 2005 to March 2009, the ministry spent nearly $20 billion[5] for approximately 15,000 civil engineering contracts, having accepted nearly 130,000 bids. The ministry set aside approximately 60 percent of the procurement budget of civil engineering projects for SMEs. Another source of data is the government database for certified contractors. It provides contractors' information about their annual sales, amounts of capital and debt, and number of engineers and employees. Based on the information, controls are established for firm size in order to measure the quantitative relationship between firm size and profitability from competitive bidding processes.

To examine the effect of a small business program on procurement costs, knowledge of what the contract prices would be should the government eliminate the program from the procurement market is necessary. However, such data are not available. Therefore, a counterfactual analysis is required to conduct comparative statics analysis of small business set-asides.

Because of set-asides, the number of sample auctions in which large firms and SMEs compete with each other is considerably limited. Hence, the counterfactual analysis begins by creating the competition between large firms and SMEs.[6] However, the size of SMEs participating differs from one to the other even in set-aside auctions. The approach taken in this study is to regress the recovered production costs and profitabilities on firm sizes in each sample auction in order to measure the overall quantitative relationship between profitability and firm size in procurement auctions.

Therefore, our empirical analysis consists of the following three-steps. First, a procedure is used of nonparametric estimation for asymmetric first-price sealed-bid auctions with affiliated private values (APV) to identify the bidders' costs from observed bids. Then, as a second step, a regression analysis is used to find the quantitative relationship between firm size and profitability in procurement auctions, in which profitability (expected payoffs) is defined by the profit margin (bid minus cost) times the probability of winning. Finally, a static entry model is constructed in which the obtained relationship between expected payoffs and firm size is employed. Regarding the estimated *ex ante* expected profits as a payoff from entry, the entry model predicts how many SMEs would drop out because of large firm

---

[3]SMEs are defined as those that hire fewer than three hundred employees and are capitalized at equal to or less than 100 million Yen in Japan. These criteria are applied to the manufacturing, construction, and transportation industries. Service businesses and some others have slightly different criteria.

[4]The law "Ensuring Opportunities for the Procurement of Receiving Orders from Government" encourages each ministry to employ set-asides to achieve the goal.

[5]This is calculated by $1 = ¥ 105.

[6]Although limited, there are auctions in which large firms and SMEs compete with each other since government procurement laws do not allow contract officers to use set-asides in the case in which there are too few SMEs to provide sufficient competition.

entry into a market that was previously reserved exclusively to SMEs under the set-aside program. Furthermore, comparing the winning bid data and the number of participants in each auction, the degree to which the resulting lack of competition affects government procurement costs is estimated.

The model of auctions with entry is based on a two-stage game: potential bidders decide whether to enter the first stage, and the second stage is a first-price auction. The first stage relies on the assumption that entry is sequential and the number of firms is treated as a continuous variable. As in the case of McAfee and McMillan (1987b), an assumption is made that all actual bidders must incur a fixed cost prior to bidding in order to know their own signal. In this setting, relevant estimates from the empirical analysis are used to predict a counterfactual situation in which the set-aside program was to be ineffective.

Surprisingly, the estimation results suggest that the program indeed saves government procurement costs. Applying the quantitative relationship between firm size and productivity to the average difference in firm size between large firms and SMEs, on average, the production cost of SMEs is 1.3 percent higher than that of large firms. Similarly, based on the quantitative relationship between firm size and winning frequency, an SME would win 5.5 percent less frequently than a large firm if an SME and a large firm competed one-on-one. These small differences in costs and winning probability lead to a non-trivial difference in profitability between the two groups of bidders. The expected payoff of an SME would be 48 percent lower than that of a large firm when both compete in the same auction. The simulation result indicates that, due to the disadvantage in profitability, the participation of SMEs would drop by 29 percent on average were set-asides to be removed. Consequently, the large firms' shifting their entry to originally set-aside projects would cause the following two competing effects on procurement costs. The prices of the originally set-aside projects would fall due to the entry by cost-efficient large firms, whereas the prices of the related projects that would have been reserved exclusively to SMEs under the set-aside program would rise because of an approximately 30 percent decline in the number of large firms. The simulation studies suggest that the latter effect dominates the former in our simulation so that the program should decrease the procurement costs by 0.17 percent.

The empirical results conclude that the set-aside program has been successful. It improves equity between advantaged and disadvantaged firms without substantial increase of procurement costs. The results not only correspond to the prediction by the theoretical literature on asymmetric auctions but also are in line with the seminal empirical work of Denes (1997) on set-aside programs, despite the difference in approach and data. In addition, our structural estimation further illustrates that the subsidized SMEs drive non-subsidized bidders to give up more of the gain on the contracts they award. The large firms' expected net gain is 0.27 percent of the estimated project cost while it would be 0.85 percent without the small business program. In other words, set-asides squeeze more rents from large firms, which enables the procurement buyer to lower procurement costs more than offsetting the resulting production cost inefficiency.

The remainder of this paper is organized as follows. Section 2 addresses the related literature. Section 3 provides a brief explanation on public construction procurement markets in Japan. Section 4 includes a description of the data. A theoretical model of asymmetric first-price sealed-bid auctions is provided in Section 5. Section 6 includes a description of the theoretical and empirical models about auctions with endogenous participation. Section 7 provides a illustrates the estimation and simulation results. Section 8 is the discussion. The final section contains further discussion and the conclusion. The proofs are given in the Appendix.

2 Related literature

As a development of theoretical studies on auctions, the issues in procurement auctions have been examined and given many insights in the literature (*e.g.* McAfee and McMillan (1986), (1989), Hansen (1988), Laffont and Tirole (1993), Jofre-Bonet and Pesendorfer (2003), Milgrom (2004), Asker and

Cantillon (2008)).

The development of these theoretical studies has inspired the recent development of empirical studies on procurement auctions based on structural estimation techniques. For instance, Jofre-Bonet and Pesendorfer (2003) established a theoretical model of dynamic auctions in which bidders choose their bids to maximize their discounted payoffs from the sequence of infinitely repeated auctions. Then, an estimation procedure is proposed for the bidder's cost distribution inferred from the first-order condition of the optimal bid. Bajari and Ye (2003) studied collusion in highway procurement auctions. Assuming that bids by cartel members are correlated and that the members bid aggressively against the outsiders but softly within the members, they proposed tests for detecting suspicious bidding behavior of bid rigging. **?** measured the effect of unobserved heterogeneity in asymmetric first-price IPV auctions. Empirical studies of firm decisions to participate in an auction have also been developed (*e.g.*, Athey et al. (2008) and Li and Zheng (2009)). Empirical analyses for a preference program in public procurement are provided by Marion (2007) and Krasnokutskaya and Seim (2009).

In particular, Marion (2007) first investigated the effect of the bid preference program in highway procurement auctions by the California Department of Transportation. Then, Marion argued that by granting a bid credit to higher-cost bidders, the government loses surplus from lower-cost bidders by awarding contracts to likely higher-cost competitors. At the same time, the preferential treatment increases the competitive pressure exerted by favored bidders. In descriptive regressions, Marion found that the auctions with bidding credits increase procurement costs by 3.5 percent, possibly because the likelihood of large firm participation is smaller for preference auctions than for non-preference auctions.

Krasnokutskaya and Seim (2009) also conducted an empirical investigation on bid preference programs in highway procurement auctions. A remarkable extension from Marion (2007) is that their model includes the bidder's endogenous participation. Then, they found that the program substantially raises the small bidders' probabilities of winning and participation but results in a small increase in the government procurement cost.

A laboratory experiment on preference programs in procurement auctions is conducted by Corns and Schotter (1999). They observed that, although preferences generally increased the disadvantaged bidder's winning probability, the procurement costs dropped only with a 5-percent price-preference rule and not with any 10 percent or above price preferences. Corns and Schotter reported that these results are consistent with theoretical predictions introduced by McAfee and McMillan (1989) and concluded that designing a cost-effective preference program is possible if a procurement buyer can access the information about the bidder's cost distribution.

The first economic study on small business set-asides is conducted by Ayres and Cramton (1996). Their case studies focus on the affirmative action program in United States FCC spectrum auctions, in which disadvantaged bidders, such as small businesses and female- or minority-owned firms, are granted set-aside licenses.[7] Their simulation suggests that this effective set-aside program could increase the government's revenues by approximately $45 million or 12 percent of the government's total auction revenue.[8]

Denes (1997) provided the first thorough analysis for the impact of small business set-asides in public procurement. He investigates the federal dredging contracts during 1990 and 1991 and examines the mean values of set-aside (or restricted) bids compared with the mean values of the unrestricted bids on the data in eight categories and performs a series of paired *t*-tests. He found that in all but one instance, there is no significant difference between the bids submitted for set-asides and the bids submitted on the related auctions with unrestricted solicitations. According to his study, 3.6

---

[7]The preference program gives disadvantaged bidders a 40 percent bidding credit on ten of the thirty narrowband licenses as well as a subsidy for their interest payments. Since the combination effect is that favored bidders had to pay the government only 50 percent of a winning bid, they consider that the credit is large enough to discourage entry by advantaged firms. See Ayres and Cramton (1996) for more details.

[8]They also note that set-aside auctions are able to raise the expected auctioneer's welfare if 1) there is insufficient competition among strong bidders; 2) the seller is able to identify who is strong or weak; 3) resale is prohibited.

firms bid on set-asides, whereas only 3.1 firms bid in the remaining auctions. Although the marginal effect of a set-aside program on the procurement cost is ambiguous, he suggested that the increased competition induces either no change or a lower-bid price on set-asides.

The thorough empirical analysis on auctions with endogenous participation by Li and Zheng (2009) is noteworthy. The theoretical analysis of auctions with endogenous participation consists of two groups of literature. One group investigates either an asymmetric equilibrium (e.g., McAfee and McMillan (1987b)) or a symmetric equilibrium (e.g., Levin and Smith (1994)), assuming that the potential bidders decide whether to enter the auction before acquiring their signal. In contrast, the other group analyzes an entry equilibrium in which potential bidders first obtain a signal and then make their entry decision Samuelson (1985). Li and Zheng developed a fully structural framework for entry and bidding admitting these theoretical models and applied the methods to quantify the effect of the bidder's entry decision on government procurement costs using data on procurement auctions of Texas Department of Transportation. Model 2 in their analysis is closest to the entry model employed in this paper.

The literature on structural estimation of auctions began with Paarsch (1992), and the non-parametnric estimation of the first-price auction model is first introduced by Guerre et al. (2000). Most closely related to our research is that of Campo et al. (2003), which proposed an ingenious way to estimate the model of asymmetric first-price auctions with affiliated private values (APV) nonparametrically.

3   Public construction markets in Japan

3.1   Overview

Investment in the construction industry accounts for nearly 20 percent of the country's GDP and employs more than 10 percent of the working population in Japan. The percentage of public investment as a portion of all construction investment was 45.6 in 2001.

Public account law requires that all government and public entities practice competitive bidding when they acquire construction works exceeding 2.5 million Yen. Three types of bidding systems are used in the public sector: 1) open competitive bidding, 2) invited bidders, and 3) contract at discretion. Although not a majority, scoring tenders are also used in the awarding mechanism, in which bidders submit not only the price but also another variable, such as the term of work or quality of work.

An idiosyncratic feature of the Japanese public procurement system is in the screening process for bidders. Contractors must take a preliminary qualification exam in order to bid for projects. The exam measures a firm's technological, financial, and geographical status and gives them scores as a result of the evaluation. For each auction, the procurer selects, or makes an announcement to, a set of legislated contractors as qualified bidders, and the selection is based on the exam results.

In procurement auctions, governments face the risk of awarding the contract to less-qualified or inferior firms that might default. Some projects demand advanced technologies and skills as well as a sufficient amount of capital to complete.[9] To mitigate such an asymmetric information problem, screening processes for selecting qualified bidders are essential to the success of the auction.[10] The preliminary qualification exam works in the same manner as the bonding system in the United States public construction market. A brief discussion of the preliminary qualification examination in Japan is presented in the next section.

Another major difference in the Japanese procurement system is in the contract principle. Unlike in the United States and many other countries, construction contracts are based on total price con-

---

[9]The possibility of default or non-performance can have perverse effects on the bidding in an auction; a bidder with a high likelihood of default tends to be chosen as a winning bidder. See Zheng (2001) for more details.

[10]See also Bajari and Tadelis (2001) and Laffont and Tirole (1993) for further discussion on the importance of the screening processes in procurement auctions.

tracts, in which bidders submit only a total price without necessarily itemizing unit prices. Instead, engineering offices regularly update market price lists and use them in the event that a change order is called for during a certain performance. The yearly updates on these price lists are based on hearings, but the survey is conducted independently from procurement auctions. Unfortunately, there is no formal theory that analyzes the effect of contract formats on bidding behaviors. Therefore, the empirical analysis here ignores the contract format effect.

Finally, the announcement policy of the reservation price and engineer's estimated costs differs from that of many other countries, in which these are typically opened prior to bidding in auctions. On the other hand, in most public procurement auctions in Japan, such information is secret until the auction is over. However, the secrecy of the reservation price is mitigated with the auction design. If no bid is below the reservation price, the next round auction begins immediately with the same member. This process goes on at most three times. The project is reserved unless any contractor bids below the reservation price at the third round. In this sense, reservation prices are not binding in the first round.

3.2   Preliminary qualification examination

Preliminary qualification certifies a set of firms as *bona fide* bidders in procurement auctions to protect the owner of a project against the risk of non-performance. Similar screening processes are widely used at public procurement auctions in Europe and work in the same manner as the bonding system in the United States public construction auctions in terms of reducing the risk of contractor's default.

The preliminary qualification in Japan is based on the firm's disclosure of information with respect to their financial and technological performance. In particular, information includes the annual sales, number of engineers in each area of expertise, experience, and business history. Based on the set of information as well as the evaluation of work performed, governments measure the firm's overall ability to perform. As a result of this evaluation, the qualified firms typically obtain two kinds of scores for each area of their expertise.

The first score is called the "Business Evaluation" (BE) score, which is essentially a weighted average of 1) the annual value of completed construction works by license classification, 2) the number of technical staff, 3) the business conditions (based on financial statement analysis), 4) the number of engineers, and 5) the record of safety performance. For the qualified bidders of MLIT, the maximum and minimum scores are 1,859 and 329, respectively, with an average of 851.1. The detailed summary statistics on the BE score are available in Section 4. The BE score is given through the countrywide criteria of measurement specified in construction industry law; thus, each firm has a unique score value for each expertise.[11]

On the other hand, the second score, which is called the "Technology Evaluation" (TE) score is the past performance evaluation measured by each procurement buyer.[12] Unlike the BE score, the measurement criteria vary across procurement buyers. Hence, it is possible, although not common, that BE and TE scores both are not high. If a government has multiple local divisions, each may have a different evaluation criterion for the TE score.

The assessment on whether a firm is favored in the set-aside program is based on the sum of the BE and TE scores.[13] However, the BE and the total scores are strongly correlated with each other.[14]

---

[11]The number of expertise is 28, which is specified in the construction industry law. Firms must obtain a license for each area of expertise to operate.

[12]The criteria typically reflect the firm's past works such as contribution to the quality of projects and schedule of works.

[13]More precisely, governments assign grade for each firm based on the total score. For instance, MLIT gives either "A", "B", "C", or "D" for each certified contractor with civil engineering expertise, where A is the top grade. Large contractors are likely to have a grade of "B" or higher and are likely to have a grade of "A" if the firm operates countrywide. Based on the grade, governments implement the set-aside program in such a way that firms with a grade of A or B are excluded to bid for low-end projects.

[14]The correlation between the total and BE scores is 0.91. The *t*-statistic of the regression of the total score on the

To avoid the heterogeneity of the TE score across locations, the analysis only uses the BE score as the control variable for the corporation size.

3.3 Set-asides in the public construction market

The selection rule for bidders is primarily based on the "size-matching rule." When a particular project is auctioned, a set of bidders is chosen so that their sizes will match the project size. For instance, only large firms are qualified to participate in the auctions for large and high-end projects and are not allowed to bid on small and low-end projects, which are reserved for SMEs. Set-asides are implemented as part of the size-matching rule. In the case of MLIT, it also grades every civil engineering work from A to D according to the size, where grade A is the highest end. The engineer's estimated costs are typically used as a proxy to determine the project size. Under the size-matching rule, contractors are selected or allowed to participate in the auction so that their grades match the project grade. The size-matching rule has priority in the selection of bidders unless the number of designated bidders is too small to provide adequate competition. Table 1 demonstrates how the government solicits firms for each project size.

Set-asides are the only explicit method to favor SMEs in Japanese public procurement auctions. Every year, the Japanese central and local governments determine the objective set-aside budgets by which the governments should assign contracts to SMEs.[15] In 2005, central governments and public entities spent ¥8.8 trillion to purchase land and items, construction works, and services. ¥4.1 trillion was expended to SMEs, which accounted for 46.9 percent of the total budget (the target amount was ¥4.3 trillion, accounting for 46.7 percent). For the Ministry of Land, Infrastructure and Transportation, 50.8 percent of the entire expenses were allocated to SMEs in the year. To achieve the goal, approximately two thirds of civil engineering contracts were set aside for SMEs.

4 The data

4.1 Overview

The data used in the analysis contain the bid results of the procurement auctions for civil engineering projects from April 2005 through March 2009. The number of contracts awarded was 15,020 during this period.

MLIT posts the bid results on the website, Public works Procurement Information service (PPI).[16] The information available in PPI includes the names of procurement buyers (local branch names), project names, project types, date of auctions, reservation prices, auction formats (open competitive bidding or invited bidders), and submitted bids with the bidder's name.[17] PPI also provides the lists of all qualified firms, which consist of the address of the firm's headquarters, the name of the owner, BE scores as well as grades for each area of expertise. All data in this empirical study are from the website.

MLIT procures 21 types of construction works including civil engineering (or heavy and general construction works), buildings, bridges, paving, dredging, and painting. The amount of civil engineering projects is approximately ¥ 750 billion a year, which accounts for approximately 54 percent of the entire expenditure of the ministry as indicated in Figure 2 and 3 as well as for 7 percent of the public construction investment in the country.

MLIT has 9 regional development divisions in 9 regional districts. The data includes the civil engineering projects in 8 districts indicated in Figure 1. Each of the regional development divisions has a certified firms' list from which it chooses the bidders for each procurement auction. The lists

---

BE score is 786.34 with 0.83 of the R-squared.

[15]This policy is specified by the "Law on ensuring the receipt of orders from small and medium-size enterprises."

[16]The address is "http//www.ppi.go.jp."

[17]The information concerning work location is not generally available.

are updated every two years. The total number of firms on the lists was 43,522 in April 2007. Since large firms typically operate across several regions, it is often the case that a particular firm is listed on two or more of these lists. The number of firms without such duplication is 32,993, which accounts for approximately 20 percent of all the licensed civil engineering construction firms in Japan.[18]

The data have some limitation in the identification of contractors. The bid results provide the bidder's company name only. Therefore, in the case that two or more different firms have an identical company name, the bidder's identity can be guessed but not ensured.[19] The way to narrow down the candidate list is on the basis that whether i) the location (prefecture) of the project matches the location of headquarters, and ii) the bidder's size matches the project size according to the size-matching rule. Through this process, almost all contractors on the bid results are identified.[20] The remaining unidentified firms in the auction are assumed to be the average-sized firm in the auction.

4.2 Summary statistics of bids and scores

4.2.1 Normalization of the bidder's size

In the observations, each auction has a unique set of bidders in general. Hence, a firm with a higher score can be a smaller bidder if the opponents have a much higher score and vice versa. To model the firm's size in comparison to the size of its opponents, the BE score is normalized (hereafter, normalized score) in the following procedure.

Let as assume that there are n bidders in an auction in which each bidder is indexed by $i = 1, \ldots, n$. Let X_i be the value of the BE score of bidder i in the auction and $\bar{X}$ be the mean score of the bidders in the auction, which satisfies $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$. Then, bidder i's normalized score x_i is calculated as

$$x_i = \frac{X_i - \bar{X}}{\bar{X}}. \tag{1}$$

In other words, the value x_i represents the firm i's relative size in the auction. x_i is positive, for instance, if bidder i's firm size is greater than the average firm size.

None of the bidders is informed of who and how many the competitors are. In fact, however, the bidders may speculate about their competition based upon the project location, the project size, and the auction date. Hence, as mentioned later on, the structural estimation model of auctions used in this research assumes that a bidder knows his own score, the mean score of the bidders in the auction, and the distributions of the opponent bidders' scores but not the exact values of the opponent bidders' scores. This assumption makes it easy for not only the structural estimation analysis to be realistic, but also for more samples to be used to estimate the pseudo values of bidders' production costs.

Table 4 provides the summary statistics on X_i and x_i of the actual bidders in the data. Figure 2 depicts the histogram for the normalized score. The effect of the set-aside program is glimpsed from the fact that the coefficient of variation (CV) on X_i, which is defined by the standard deviation divided by the mean of X_i, is approximately 13 percent. Therefore, if bidders are randomly picked in each auction, the standard deviation of x_i would be 13 percent. However, the actual standard

---

[18]The total number of licensed civil engineering firms is 167,896 in 2005 (MLIT, 2005).

[19]For example, there are seven "Showa Kensetsu Co., Ltd" on the contractor list of Kanto District Development Bureau. The bid results do not indicate which "Showa Kensetsu" in fact bid.

[20]Eity-seven percent of bidders are exactly identified from the firm list data by the uniqueness of the firm name on the list, by the address or by the project size (Large firms never enter the auctions for small-size projects). Then, letter grades are used for the further identification. Most of the auctions accept bids from the firms with a certain letter grade. If the majority of the identified firm's letter grade is "C," then the rest of the firms' grades should be "C." In this way, approximately 8 percent of bidders are additionally identified. The rest are guessed by the distance between the firm's headquarters and the construction site; if it is too far, the firm should not be a bidder for the auction. Then, 99.1 percent bidders are identified. Although the final process includes a slight amount of ambiguity, we carefully choose the firm whose letter grade is the same as that of the remaining bidders in the auction, meaning that the BE score of the guessed firm is close to the score of the true firm.

deviation is 7.6 percent, which suggests that the participation restriction by government reduces the asymmetry of bidders.

4.2.2 Percentage bids

Figure 3 contains a description of the histogram on the project size. Since each construction project is unique, there remains a great deal of heterogeneity in project size. The most typical contract is for approximately ¥100 million measured in the engineer's estimated costs. The largest is approximately ¥12 billion, while the smallest is less than ¥1 million. Table 5 is a breakdown of the summary statistics of project size.

To eliminate project heterogeneity, all bids in the empirical analysis are described by the percentage with respect to the engineer's estimated cost. Let Est be the engineer's estimated cost of a project and Bid_i be the value of the bidder i's bid. If the procurement auction is held with the price-only format, then the percentage bid of the bidder i's bid is given by

$$\frac{\text{Bid}_i}{\text{Est}}. \tag{2}$$

In the data, 3,405 out of 15,020 procurement auctions are undertaken with price-only auctions, which account for approximately 23.0 percent. The rest are auctioned off with scoring auctions, in which bidders submit not only the price-bid but also some other factors, such as quality and completion time.

According to the MLIT scoring auction procedure, the bidder with the highest-score wins the project in which the score is calculated by the factor-bid divided by the price-bid. The factor-bid consists of multiple components, such as noise level, completion time, and experience.

The properties of scoring auctions have been investigated in the literature (e.g. Che (1993), Asker and Cantillon (2008)). The assumption that existing studies rely on is that the scoring rule is quasi-linear, which, unfortunately, does not hold in our case. Hence, an alternative model of scoring auctions is established in which the scoring rule is based on division, as will be discussed in Subsection 5.3.

To incorporate the data of scoring auctions into the model, let Score_i be the value of the price-bid divided by the factor-bid submitted by bidder i.[21] Furthermore, let Base be the base score that is the engineer's estimated cost (the highest possible price-bid) divided by the factor-bid evaluating nothing (the lowest possible factor-bid). The base score is set in each scoring auction, and the winning scoring bid must be below the base score.[22] Then, the percentage scoring bid of bidder i is defined as

$$\frac{\text{Score}_i}{\text{Base}}. $$

Since the scoring bid and base score in the model are the inverse of the scoring bid and the base score in the data, the winner of the scoring auction in our model is the lowest-score bidder conditional on the scoring bid being below the base score.

4.2.3 Regression results for bids on corporate size

It is evident that, in each auction, larger firms bid lower prices than smaller ones. Table 6 contains a description of the result of regression for the percentage bids on normalized scores. Auction-specific effects are taken into account by fixed-effect and random-effect models. As in Corns and Schotter (1999), the auctions that contain "throw-away bids," *i.e.*, bids larger than 200 percent of the engineer's estimated cost, are dropped off in order to prevent those minor but extreme samples from dominating

---

[21]As shown above, the actual scoring bid in the data is the factor bid divided by the price-bid.

[22]In fact, the base score in our model is the inverse of the base score in the data. Hence, the winner's scoring bid in real-world scoring auctions must be equal to or above the actual base score.

the estimation.[23] Then, the negative relationship between the normalized bids and size is significant (*t*-value : 6.15 in FE estimation). The number of observed bids after exclusion equals 110416 . Figure 4 shows that the bid density of larger firms (the score is 10 percent greater than the average) is shifted downward when compared to that of smaller firms (the score is 10 percent smaller than the average). Table 6 indicates that the bidder's size yields a small but statistically significant difference in bids.

Finally, the production capacity utilization is explored in procurement auctions. Many small businesses on the qualified-firm lists do not bid even with the set-aside program. Table 7 reports that approximately 80 percent of the large firms on the lists bid during the period whereas only approximately 25 percent of the small businesses on the lists bid. The distribution of the scores of actual bidders and that of qualified firms on the lists are shown in Figure 5 and Figure 6, respectively. The density of the qualified-firms' BE scores shifts toward the left. This tendency is trivial if the distributions between the qualified LBs' and actual bidder LBs' scores are compared as in Figure 10 and Figure 9. However, a significant leftward shift is observed in the distribution of the actual SB bidders' scores in Figure 10 and Figure 9. This indicates that, despite set-asides, a sufficient volume of production capacity remains available in small businesses.

5 Recovery of the bidders' cost distribution

5.1 Overview

Our nonparametric estimation of first-price sealed-bid auctions is based on Campo et al. (2003), which is an extension of Guerre et al. (2000) to cases with asymmetric bidders with the APV model. The bidders' costs could be correlated or even affiliated, since some bidders may use the same subcontractor and employ materials and workers in the same market. Hence, it seems that the affiliated private value is one of the most reasonable settings to analyze auctions for construction projects.[24]

The approach of Campo et al. (2003) relies on the assumption that the bidder's asymmetry is represented by a finite number of segments. Hence, if the number of segments is equal to d, a $(d+1)$-dimensional kernel estimation is required. Therefore, if an empirical model assumes that the bidder's asymmetry is attributed to a continuous variable, then kernel estimation is not implementable.

More recently, Zhang and Guler (2005) proposed a simplified approach in which the only requirement is a two-dimensional kernel estimation regardless of the structure of bidder asymmetries. The essence of their approach is to estimate the bidder's signal separately for each bidder, expressing each bidder's payoff function in terms of the equilibrium distribution of rival bids. They claim that the detrimental effect from the dimensionality of kernels can be avoided as long as the set of bidders in the sample is identical. Unfortunately, their approach results in another problem when the data involves heterogeneity in the set of participants across auctions, as it does in this case.

Hence, a model of asymmetric auctions is reconstructed to utilize more samples in kernel estimation, assuming that bidders know their own firm size but have limited information about their competitors. In particular, we assume that bidder i knows X_i, his own score, F_X, the cumulative distribution function of the set of the bidders' scores, and $\bar{X}$, the mean of the score in the auction but not the exact value of the set of opponent bidder's scores $X_{-i} = (X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n)$. As shown in the next subsection, the bidders are still *ex ante* asymmetric on this assumption. Furthermore, this assumption is more realistic in actual procurement auctions, in which the participants are endogenously determined and nobody knows who the actual opponents are upon bidding.

---

[23]The number of such sample auctions in our data is 306, which accounts for 2.8 percent of 11,114, the total number of sample auctions.

[24]Another non-trivial source that correlates observed bids is auction-specific unobserved heterogeneity, which is observable for bidders but unobservable to an econometrician. A seminal paper by Krasnokutskaya (2009) introduced an innovative methodology to identify a model of auctions with unobserved heterogeneity and quantified the impact on the results of estimation as well as policy-relevant instruments. Furthermore, Krasnokutskaya's approaches have been applied to many empirical analyses on auctions, such as Athey et al. (2008).

5.2   An asymmetric APV model of first-price auctions

A single and indivisible project is auctioned to n risk-neutral bidders, indexed by $i = 1, \ldots, n$. Let X_i be the bidder i's firm size, which is randomly and independently distributed following a cumulative distribution function F_X. Let $\bar{X}$ be the mean of the realized firm sizes in the auction, defined by $\bar{X} \equiv (1/n) \sum_{i=1}^{n} X_i$. Then, the bidder i's normalized firm size in the auction is computed by $x_i = \left(X_i - \bar{X} \right) / \bar{X}$.

Let $H(\cdot)$ be an n-dimensional cumulative distribution function. The vector of each bidder's normalized score $x \equiv (x_1, \ldots, x_n)$ is a realization of a random vector with a joint distribution $H(\cdot)$. Then, for each $i \in \{1, \ldots, n\} \equiv \mathcal{N}$, the conditional distribution of $x_{-i} \equiv (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ and its density can be denoted by $H_{-x_i|x_i}(x_{-i}|x_i)$ and $h_{-x_i|x_i}(x_{-i}|x_i)$, respectively. Let us assume that, for all i, $H_{-x_i|x_i}(x_{-i}|x_i)$ has support $[\underline{x}, \overline{x}]^{n-1}$ and that the probability density function $h_{-x_i|x_i}(x_{-i}|x_i)$ is continuous in x_{-i}.

Let us assume that F_X and $H(\cdot)$ are common knowledge but X_i is known only to bidder i. Let us further assume that the mean of the realized scores in the auction, $\bar{X}$, is known to every bidder. Hence, the conditional distribution of x_{-i}, denoted by $H_{-x_i|x_i}(x_{-i}|x_i)$, and its density, $h_{-x_i|x_i}(x_{-i}|x_i)$, are common knowledge but x_{-i} is not.

The asymmetric APV model with risk-neutral bidders is defined by an n-dimensional distribution with a cumulative distribution function $F(\cdot|x)$. The vector of private information $(c_1, ..., c_n)$ is a realization of a random vector with joint distribution $F(\cdot|x)$. The asymmetry of bidders is captured by x such that x_i affects the marginal distribution of c_i but not the distribution of c_j for any $j \in \mathcal{N} \backslash \{i\}$. In other words, the marginal distribution of c_i is represented by $F_{c_i}(c_i|x_i)$ for all $i \in \mathcal{N}$. The affiliation is captured as follows: let us assume that the ith bidder's signal is c_i, then for some j, the marginal distribution of c_j and its density are given by $F_{c_j|c_i}(c_j|c_i, x_j)$ and $f_{c_j|c_i}(c_j|c_i, x_j)$.

Using $b_i = \beta(c_i|x_i)$ and $\beta^{-1}(b_i|x_i)$, let us denote the equilibrium bidding strategy and its inverse, respectively. In equilibrium, the joint distribution of valuations $F(\cdot|x)$ and the distribution of bids $G(\cdot|x)$ are related with $G(b_1, ..., b_n|x) = F(\beta^{-1}(b_1|x_1), ..., \beta^{-1}(b_n|x_n)|x)$. Let us assume that the marginal distribution of costs $F_{c_j|c_i}(c_j|c_i, x_j)$ has support $[\underline{c}, \overline{c}]$ for any i, j, and x and that the probability density function $f_{c_j|c_i}(c_j|c_i, x_j)$ is continuously differentiable (in c_j). Finally, let us assume that for all $i \neq j$, $f_{c_j|c_i}(\cdot|c_i, x_j)$ is bounded away from zero on $[\underline{c}, \overline{c}]$. Then, firm i's conditional payoff can be written as

$$u(b_i|c_i, x_i) = \max_{b_i}(b_i - c_i) \Pr\{b_i \leq B_i|c_i, x_i\}, \tag{3}$$

where B_i is the bidder i's minimum rival bid, namely $B_i \equiv \min\{b_1, \ldots, b_{i-1}, b_{i+1}, \ldots, b_n\}$.

Then, an increasing Bayesian-Nash equilibrium is considered in pure strategies. Let us assume that there is an increasing equilibrium such that each firm i bids according to a strictly increasing function $\beta(c_i|x_i)$. An equilibrium in pure strategies is an n-dimensional strategy profile $(\beta(\cdot|x_1), \ldots, \beta(\cdot|x_n))$ such that $\beta(\cdot)$ maximizes $u(b_i|c_i, x_i)$ in b_i for all i and c_i in its support.

Then, for any $i \in \mathcal{N}$ and $j \in \mathcal{N} \setminus \{i\}$, $G_{b_j|b_i}(b|b_i, x_i, x_j) \equiv F_{c_j|c_i}(\beta^{-1}(b|x_j)|\beta^{-1}(b_i|x_i), x_j)$ is defined as the probability that an opponent bidder j's bid, b_j, is equal to or greater than b given b_i and x. Note that $G_{b_j|b_i}(\cdot)$ satisfies the property of probability distribution since $\beta(\cdot)$ is strictly increasing.

For bidder i, the minimum rival bid B_i is a random variable conditional on b_i and x_i. Therefore, $G_{B_i|b_i}(B_i|b_i, x_i, x_{-i})$ is used to denote the conditional cumulative distribution of B_i.[25] If x_{-i} were known to bidder i, then bidder i's winning probability would be $1 - G_{B_i|b_i}(\cdot)$ conditional on other bidders following $\beta(\cdot)$.

We have assumed, however, that bidder i knows the conditional distribution $h_{x_{-i}|x_i}(x_{-i}|x_i)$ but does not know the exact values of the opponents' firm sizes, x_{-i}. Therefore, the bidder i's *expected*

---

[25]By affiliation of c, b_i influences $G_{B_i|b_i}$, while, by heterogeneous distribution of c, x_i affects $G_{B_i|b_i}$.

winning probability, $1 - \bar{G}_{B_i|b_i}(\cdot)$, is given by [26]

$$1 - \bar{G}_{B_i|b_i}(b|b_i, x_i) = \int_{x_{-i}} [1 - G_{B_i|b_i}(b|b_i, x_i, x_{-i})] h_{x_{-i}|x_i}(x_{-i}|x_i) dx_{-i}.$$

Hence, if other bidders follow $\beta(\cdot|x_j)$, then (3) is rewritten as

$$u(b_i|c_i, x_i) = \max_{b_i}(b_i - c_i)[1 - \bar{G}_{B_i|b_i}(b_i|b_i, x_i)].$$

The first-order condition gives

$$c_i = b_i - \frac{1 - \bar{G}_{B_i|b_i}(b_i|b_i, x_i)}{\bar{g}_{B_i|b_i}(b_i|b_i, x_i)}, \qquad (4)$$

where $\bar{g}_{B_i|b_i}(\cdot)$ is the density of $\bar{G}_{B_i|b_i}(\cdot)$.

The right-hand side of (4) gives a unique inverse bid function $\beta^{-1}(b_i|x_i)$, implying that i's strategy is also represented by $\beta^{-1}(b_i|x_i)$. Hence, it is a Bayesian-Nash equilibrium in asymmetric first-price auctions with APV. The bidding function can be obtained by solving the system of differential equation represented by $\beta^{-1}(b_i|x_i)$ for all i.

5.3 Application to scoring auctions

The model can be applied to scoring auctions. For simplicity, an independent private value (IPV) is assumed to begin with, and the assumption is relaxed later on.

For all $i \in \{1, \ldots, n\}$, bidder i obtains two-dimensional private information $c_i \in [\underline{c}, \bar{c}]$ and $q_i \in [\underline{q}, \bar{q}]$ distributed independently following the publicly known cumulative distribution F_c and F_q, where q is interpreted as the quality level the bidder will provide to complete the project.

The obtained cost-quality pair is considered as the one optimally selected from a set of feasible cost-quality pairs, where the set can be considered as the firm's cost function. The discussion on how the bidder ends up choosing the cost-quality combination made in, for instance, Che (1993) and Asker and Cantillon (2008), is out of the scope of this research. However, we assume that, given q_i, bidder i estimates the production cost c_i by incurring the information acquisition expense e.

In the scoring auction, bidder i submits a price-bid b_i and a factor-bid q_i. The scoring rule $S(b, q) \in \mathbb{R}_+$ evaluates the paired bid and designates the lowest-scored bidder as the winner. Let S_i be the minimum rival scoring bid for bidder i. Then, bidder i's maximization problem is given by

$$\max_{b_i}(b_i - c_i) \Pr\{S(b_i, q_i) \leq S_i\},$$

where the scoring rule $S(\cdot)$ is given by $S(b_i, q_i) = b_i/q_i$. Then, as in Asker and Cantillon (2008), the i's pseudo-type, θ_i, is defined as $\theta_i = c_i/q_i$. Finally, $s_i = b_i/q_i$ is defined for notational convention.

---

[26]The right-hand side is more formally expressed as

$$\int_{x_1} \cdots \int_{x_{i-1}} \int_{x_{i+1}} \cdots \int_{x_n} \left[1 - G_{B_i|b_i}(b_i|b_i, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)\right] h_{-x_i|x_i}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n|x_i) dx_1 \ldots dx_{i-1} dx_{i+1} \ldots dx_n.$$

Then, the maximization problem is rewritten as[27]

$$\max_{s_i} (s_i - \theta_i) \Pr\{s_i \le S_i\}$$

for some q_i. It implies that the multi-dimensional scoring auction game can be reduced into the single-dimensional first-price sealed-bid auction in which bidder i receives a signal θ_i, submits the bid s_i, and obtains the expected payoff given above for all $i = 1, \ldots, n$.

The way that this specification is taken to the asymmetric APV paradigm is given as follows. Let us assume that the set of pseudo-types $(\theta_1, \ldots, \theta_n)$ satisfies all the properties $(c_1, \ldots, c_n)$ in the price-only auction. Then, let $F_\theta(\cdot|x)$ be the cumulative distribution function and $F_{\theta_j|\theta_i}(\theta_j|\theta_i, x_j)$ with the support $[\underline{\theta}, \bar{\theta}]$ be the marginal distribution of θ_j. The correlation between q and θ_i is admitted in the model so that the cumulative distribution of q is given by $F_q(q_i|\theta_i)$ taking into account the case that the bidder with higher θ is likely to have a higher-quality standard q. The above maximization problem can then be rewritten as

$$\max_{s_i} (s_i - \theta_i) \Pr\{s_i \le S_i | \theta_i, x_i\}. \tag{5}$$

Hence, the maximization problem of the asymmetric APV scoring auctions is equivalent to that of price-only auctions if s_i is replaced with b_i and θ_i is replaced with c_i in Equation (5).

Finally, we verify that the rest of the argument after Equation (3) in the previous subsection holds in the case of scoring auctions. Suppose that there is an increasing function $\sigma(\theta_i|x_i)$ which maximizes Equation (5) for any $\theta \in [\underline{\theta}, \bar{\theta}]$. We can then define $G_s(\cdot|x)$ as $G_s(s_1, \ldots, s_n|x) = F_\theta\left(\sigma^{-1}(s_1|x_1), \ldots, \sigma^{-1}(s_n|x_n)|x\right)$. Therefore, the conditional cumulative distribution function of S_i is given by $G_{S_i|s_i}(S_i, x_i, x_{-i})$ and the i's expected winning probability in the scoring auction is given by $1 - \bar{G}_{S_i|s_i}(\cdot|b_i, x_i, x_{-i})$. Thus, the first-order condition of the maximization problem in the scoring auction, which is a counterpart of Equation (4), is given by

$$\theta_i = s_i - \frac{1 - \bar{G}_{S_i|s_i}(s_i|s_i, x_i)}{\bar{g}_{S_i|s_i}(s_i|s_i, x_i)},$$

where $\bar{g}_{S_i|s_i}(\cdot)$ is the density of $\bar{G}_{S_i|s_i}(s_i|s_i, x_i)$.

As is discussed in Subsection 9, the bid margin $s_i - \theta_i$ in the scoring auction is different from the bidder's conditional payoff $b_i - c_i$. To identify the bidder's expected payoff as well as the *ex ante* expected gain from participating in the auction, the difference has to be dealt with carefully. With this exception, all the arguments in the following apply to scoring auctions by replacing b_i with s_i and c_i with θ_i.

5.4 Nonparametric estimation

Campo et al. (2003) show that the latent value c_i can be estimated by using the inverse bid function $\beta^{-1}(\cdot)$. They show that the estimator for costs can be obtained by computing the bid distribution $G_{B_i|b_i}$ and its density $g_{B_i|b_i}$ without solving the system of differential equations.

As in Zhang and Guler (2005), the first step is to interpret (4). By definition, $1 - G_{B_i|b_i}(b|b_i, x_i)$ is the probability that the minimum rival bid B_i is greater than b conditional on b_i, and $\bar{g}_{B_i|b_i}(b|b_i, x_i)$

---

[27]Since q_i is given and constant for bidder i and the scoring rule is specified in the above expression, the maximization problem can be rearranged as

$$\max_{b_i} q_i \left(\frac{b_i}{q_i} - \frac{c_i}{q_i}\right) \Pr\left\{\frac{b_i}{q_i} \le S_i\right\}.$$

b_i/q_i and c_i/q_i are replaced, respectively, with s_i and θ_i, and the remaining q_i is suppressed due to the fact that it is redundant for the optimization problem. The rearranged optimization problem is then obtained. Note that the optimal bid b_i uniquely determines the optimal score s_i, since q_i is given for bidder i.

is the derivative of $\bar{G}_{B_i|b_i}(b|b_i, x_i)$.

For the estimation, let us assume that there are $k = 1, \ldots, m$ auctions and that n bidders bid in each. Unlike the standard estimation model, the assumption that the set of bidders in each sample is the same is relaxed in the way that bidder i in auction k can be different from bidder i in auction k' in our model. Let $B_{i,k} = \min_{j \neq i} b_{j,k}$ denote the i's minimum rival bid for any sample auction k. Then, the empirical correspondent of (4) is given by

$$c_{i,k} = b_{i,k} - \frac{1 - \bar{G}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k})}{\bar{g}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k})}. \tag{6}$$

Although the number of combinations of $x_k \equiv (x_{1,k}, \ldots, x_{n,k})$ in the observations is infinitely large, $\bar{G}_{B_{i,k}|b_{i,k}}$ and $\bar{g}_{B_{i,k}|b_{i,k}}$ depend only on $x_{i,k}$ but are independent of

$$x_{-i,k} \equiv (x_{1,k} \ldots, x_{i-1,k}, x_{i+1,k}, \ldots, x_{n_k,k}).$$

Therefore, to know the latent value of bidder i in auction k, the values of bid i' in auction k' can be used if the counterpart bidder's score $x_{i',k'}$ is the same or close enough to $x_{i,k}$. The nonparametric estimation equations for the numerator and denominator in (6) are thus given by

$$\begin{cases} 1 - \bar{G}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k}) = \frac{1}{mh_G h_x} \sum_{l=1}^{m} \sum_{\tau=1}^{n} \mathbf{1}(b_{i,k} \leq B_{\tau,l}) K_G \left(\frac{b_{i,k} - b_{\tau,l}}{h_G}, \frac{x_{i,k} - x_{\tau,l}}{h_x} \right), \\ \bar{g}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k}) = \frac{1}{m(h_g)^2 h_x} \sum_{l=1}^{m} \sum_{\tau=1}^{n} K_g \left(\frac{b_{i,k} - B_{\tau,l}}{h_g}, \frac{b_{i,k} - b_{\tau,l}}{h_g}, \frac{x_{i,k} - x_{\tau,l}}{h_x} \right). \end{cases} \tag{7}$$

These hold to the extent that the number of bidders is identical in the sample and there is no heterogeneity in the characteristics of the projects. In fact, the observations in the paper involve significant heterogeneity in the number of bidders. The next subsection is an explanation of how to control for heterogeneity.

5.4.1 Heterogeneity

Here, we essentially follow Guerre et al. (2000) to control the heterogeneity in the number of bidders and the characteristics of each auction. Guerre et al. (2000) report that these are tractable in nonparametric identification by introducing additional dimensions on kernels. The data taken here involve considerable heterogeneity in both the number of bidders[28] and the auction format (scoring auctions or price-only auctions). The procedure is described as follows.

Let z_k denote the vector of associated characteristics in project k. Let us assume that the bidders' cost distribution for the auction, *i.e.*, auction k, is given by the conditional distribution $F(\cdot|z_k)$ for some z_k. Then, the distribution of observed bids in auction k is given by $G(\cdot|n_k, z_k)$. Hence, (4) is rewritten as

$$c_{i,k} = b_{i,k} - \frac{1 - \bar{G}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k}, n_k, z_k)}{\bar{g}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k}, n_k, z_k)}, \tag{8}$$

---

[28]The smallest number is two and the largest 53.

and (7) becomes

$$
\begin{cases}
1 - \bar{G}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k}, n_k, z_k) \\
\quad = \frac{1}{m h_G h_x h_n h_{\mathbf{z}}} \sum_{l=1}^{m} \frac{1}{n_l} \sum_{\tau=1}^{n_l} \mathbf{1}(b_{i,k} \leq B_{\tau,l}) K_G \left(\frac{b_{i,k} - b_{\tau,l}}{h_G}, \frac{x_{i,k} - x_{\tau,l}}{h_x}, \frac{n_k - n_l}{h_n}, \frac{z_k - z_l}{h_{\mathbf{z}}} \right), \\
\\
\bar{g}_{B_{i,k}|b_{i,k}}(b_{i,k}|b_{i,k}, x_{i,k}, n_k, z_k) \\
\quad = \frac{1}{m (h_g)^2 h_x h_n h_{\mathbf{z}}} \sum_{l=1}^{m} \frac{1}{n_l} \sum_{\tau=1}^{n_l} K_g \left(\frac{b_{i,k} - B_{\tau,l}}{h_g}, \frac{b_{i,k} - b_{\tau,l}}{h_g}, \frac{x_{i,k} - x_{\tau,l}}{h_x}, \frac{n_k - n_l}{h_n}, \frac{z_k - z_l}{h_{\mathbf{z}}} \right),
\end{cases}
\tag{9}
$$

where K_G is a four-dimensional kernel and K_g is a five-dimensional kernel. The regularity assumption for F and G is provided in Guerre et al. (2000).

As usual, the bandwidth is given by the so-called rule of thumb; $h_g = c_g (\sum_{k=1}^{m} n_k)^{-1/6}$ and $h_G = c_G (\sum_{k=1}^{m} n_k)^{-1/5}$, where $c_G = c_g = 2.978 \times 1.06 \hat{\sigma}_b$ and $\hat{\sigma}_b$ is the sample variance of the normalized bids. The following triweight kernel is used in the nonparametric identification:

$$
K(u) = \frac{35}{32}(1 - u^2)^3 \mathbf{1}(|u| < 1).
$$

The calculation is executed using a program written in C++; it takes approximately an hour to obtain 100,000 latent variables.

The informational rents obtained by bidders decrease as the number of bidders increases. Figure 11 shows the bidding function in the case of a small number of participants (5 bidders), and Figure 12 describes the case of many participants (between 22 and 28 bidders). In both figures, the dark plots represent the bidding function and the light plots represent the 45-degree line. The bid margins are larger in the case of a smaller number of competitors.

Table 6 shows the regression result for the estimated costs as a function of a firm's size. Again, the fixed and random effects control for the auction-specific heterogeneity, and all the throw-away bids (greater than 200 percent of the reservation price) are dropped in the regression. Table 6 suggests the statistical significance (t-value : 6.57 in the FE regression) that large firms have a cost advantage.

Literature on asymmetric first-price auctions predicts that disadvantaged bidders bid more aggressively than advantaged bidders in an auction. Table 9 shows the regression result of a log bid margin (a submitted bid minus the estimated cost) on bidders' relative sizes. It is statistically significant (t-value : 6.22 in the fixed-effect regression) that a smaller bidder in an auction is likely to bid with a thinner margin than a larger bidder.

6 A model for auctions with entry

6.1 Setting

Our stylized entry model considers that a government procures K^H identical *high-end* projects, denoted by H, and K^L identical *low-end* projects, denoted by L. There are two groups of risk-neutral firms, large ones, denoted by LB, and small businesses, denoted by SB. Let us assume that every firm is normalized to have a unit production capacity regardless of the group it belongs to. As in real-world procurement auctions, a winning bidder who fails to start performing right after the auction will be severely punished *i.e.*, getting a sufficiently large amount of negative profit. Therefore, also because of the unit production capacity, no firm will bid for two or more procurement auctions simultaneously.[29]

---

[29]In reality, a company may bid to multiple projects within a short time. Such a company is represented by multiple firms in our model, each of which submits a bid to a single project.

Based on the fact observed at the end of Subsection 4.2.3, the number of large firms (LBs) is assumed to be limited to a finite number N_{LB}, whereas there is an infinitely large number of small businesses (SBs). Furthermore, let us assume that the size of project H and L is identical.[30] The only difference between the two types of projects lies in the fact that H projects are so technologically demanding that SBs are not allowed to bid. The two types of projects are auctioned through $K^H + K^L$ independent first-price sealed-bid auctions that take place simultaneously.

The procurement proceeds in the following two-stage game: potential bidders decide their entry in the first stage and auctions take place in the second stage. Once a potential bidder decides to participate in an auction, it will incur a participation cost e, obtain its own private information c, know the number of competitors, and submit a bid following a Nash bidding strategy in the second-stage auction game. The participation cost e is interpreted as an information acquisition cost and, hence, is reasonably assumed to be identical and common knowledge for all potential bidders and for any type of projects. If the set-aside program is implemented, the low-end project is exclusively offered to SBs so that LBs cannot bid. Otherwise, an LB can be a recipient of an L project.

Despite the simplification, the entry game has many pure and mixed equilibria depending on the entry process.[31] Therefore, it is further assumed that entry takes place sequentially, as in McAfee and McMillan (1987a), and that the number of bidders is treated as a continuous variable.

Then, the number of players in our entry game can be reduced to the following two representative agents, LB and SB. Each agent $t \in \{LB, SB\}$ decides the number of participants N_t^s in the auctions for type $s \in \{H, L\}$ projects subject to the capacity constraint of LBs in the procurement market, i.e., $N_{LB}^H + N_{LB}^L \leq N_{LB}$ and the participation constraint, i.e. $n_{SB}^H = 0$, and if the set-aside program is implemented, $N_{LB}^L = 0$. It is also assumed that LB decides his entry first, and, successively, SB makes his participation decision. Upon making their entry decision N_t^s, each representative player t incurs participation costs $e \times N_t^s$ for the auctions of type s projects. This setting gives us a unique asymmetric Nash entry equilibrium. The timeline is described in Figure 13.

Let n_t^s be the number of group t bidders in an auction for type s project. Since K^s identical type s projects are auctioned off simultaneously, n_t^s satisfies

$$n_t^s = \frac{N_t^s}{K^s}, \tag{10}$$

for all s and t. Given the equilibrium numbers of participants, (n_{LB}^H, n_{SB}^H) and (n_{LB}^L, n_{SB}^L), they submit a bid following the optimal bidding strategies as will be discussed in the next subsection.

6.2 Analysis for the auction stage

In this subsection, we calculate the expected profit of a potential bidder who has decided to participate in a procurement auction.

Consider a procurement auction in which there are two groups of risk-neutral bidders, LBs and SBs. For each $t \in \{SB, LB\}$, let n_t^s denote the number of group t bidders in auction $s \in \{H, L\}$, i.e., the auction for project s. Bidder i in group t has cost c_t^i, which is drawn from cumulative distribution function $F_c(\cdot|t)$ on $[\underline{c}, \bar{c}]$.

Note that the bidders in the same group are *ex ante* symmetric, drawing their cost from an identical distribution. However, we assume that $F_c(\cdot|SB)$ has conditional first-order stochastic dominance over $F_c(\cdot|LB)$ as in Maskin and Riley (2000). Hence, the cost of an LB who participates in auction L is on average lower than that of an SB in the auction. The cost of an LB for project H is on average the same as that for project L.

---

[30]This assumption does not violate the definition of high- and low-end projects used in our empirical specification. An L project here is assumed to consist of multiple low-end projects in data.

[31]Levin and Smith (1994) show that the number of actual bidders will be stochastic if the entry is simultaneous. Nakabayashi (2010) also analyzes that there are multiple equilibria in the simultaneous entry game if the potential entrants are not *ex ante* identical.

For all $i \in \{1, \ldots, n_t\}$, let b_t^i be the bid value of bidder i in group t. Then, the bidder's problem is to maximize his conditional interim expected payoff u_t given c_t^i, n_{LB} and n_{SB} by choosing b_t^i;

$$u_t(c_t^i | n_{SB}^s, n_{LB}^s) = \max_b (b - c_t^i) \prod_{i' \neq i} \Pr \left(b \leq b_t^{i'} \right) \prod_{i=1}^{n_{t'}^s} \Pr \left(b \leq b_{t'}^i \right), \; t' \in \{SB, LB\} \text{ and } t' \neq t.$$

Suppose that there is an increasing and group symmetric equilibrium in which for each $t \in \{SB, LB\}$, bidders in group t follow an identical bidding function $\beta_t(\cdot)$. If all other bidders follow β_t, then, bidder i in group t's problem (3) is rewritten as

$$u_t(c_t^i | n_{SB}^s, n_{LB}^s) = \max_b (b - c_t^i) \left[1 - F_c \left(\beta_t^{-1}(b) | t \right) \right]^{n_t^s - 1} \left[1 - F_c \left(\beta_{t'}^{-1}(b) | t' \right) \right]^{n_{t'}^s}. \tag{11}$$

As usual, we take the derivative with respect to c_t^i to apply the envelop theorem, impose the symmetry condition such that $b = \beta_t(c_t^i)$, take integral on both sides from c_t^i to $\bar{c}$, and suppress the superscript i. Then, we obtain

$$u_t(c_t | n_{SB}^s, n_{LB}^s) = \int_{c_t}^{\bar{c}} [1 - F_c(c|t)]^{n_t^s - 1} \left[1 - F_c \left(\beta_{t'}^{-1} \left(\beta_t(c) \right) | t' \right) \right]^{n_{t'}^s} dc.$$

Taking the expectation yields $U(\cdot)$, the interim expected payoff of a bidder in group t from the auction with n_{LB}^s and n_{SB}^s bidders is given as

$$U_t(n_{SB}^s, n_{LB}^s) \equiv E[u_t(c_t | n_{SB}^s, n_{LB}^s)]$$
$$= \int_{\underline{c}}^{\bar{c}} F_c(c|t) [1 - F_c(c|t)]^{n_t^s - 1} \left[1 - F_c \left(\beta_{t'}^{-1} \left(\beta_t(c) \right) | t' \right) \right]^{n_{t'}^s} dc. \tag{12}$$

This equals each potential entrant's conditional expected revenue from participating in the auction provided that the numbers of competitors are n_{LB}^s and n_{SB}^s.[32] For any n_t^s and $n_{t'}^s$, U_t could be structurally estimated using Equation (12) and the cumulative distribution function of the recovered pseudo-value c.

Unfortunately, this approach is not feasible here since, as will be discussed, the counterfactual analysis requires a function $U_t(\cdot)$ on continuous variables, n_t^s and $n_{t'}^s$. To acquire U_t numerically (12), the number of necessary computations goes to infinity. Hence, a reduced form approach is adopted to obtain the function $U_t(\cdot)$ as follows.

First, it is assumed that there exists a function $V(\cdot)$ such that

$$V(x_t(n_{SB}^s, n_{LB}^s), n^s) \equiv U_t(n_{SB}^s, n_{LB}^s), \tag{13}$$

where $n^s = n_{SB}^s + n_{LB}^s$. The identity indicates that the *ex ante* interim expected payoff $v(\cdot)$ can be decomposed into two components, i) the number of competitors represented by n^s, and ii) the firm size represented by the normalized score x_t.

The function $x_t(\cdot)$ is defined in the same manner as (1). Let $\bar{X}_t$ be the average size of group t firms in the procurement market. In addition, let $\bar{X}^s$ denote the bidders' average size in auction s, formulated by $\bar{X}^s = (\bar{X}_{LB} \cdot n_{LB}^s + \bar{X}_{SB} \cdot n_{SB}^s)/n^s$. Then, the normalized score of a group t firm in auction s is obtained by

$$x_t(n_{SB}^s, n_{LB}^s) = \frac{\bar{X}_t - \bar{X}^s}{\bar{X}^s}. \tag{14}$$

---

[32]Other entry models, such as Li and Zheng (2009), assume that the number of bidders is unknown to bidders until the auction is over. In this case, the *ex ante* expected gain integrates the interim expected profit over the possible combination of the numbers of participants. See Li and Zheng (2009) for more details.

The explicit form of $x_t(\cdot, \cdot)$ is given in the Appendix.

By linear approximation, the logarithm of $V(\cdot)$ is given by

$$\log V\left(x_t(\cdot), n^s\right) = \log V(0, 0) + V_1 \cdot x_t(\cdot) + V_2 \cdot n^s,$$

where $V_1 = \partial \log V(0,0)/\partial x_t$ and $V_2 = \partial \log V(0,0)/\partial n^s$. Let $\log V(0,0) = \alpha_0$, $V_1 = \alpha_1$, and $V_2 = \alpha_2$. Then,

$$\log U_t(n^s_{SB}, n^s_{LB}) = \alpha_0 + \alpha_1 \cdot x_t(n^s_{SB}, n^s_{LB}) + \alpha_2 \cdot n^s \tag{15}$$

is obtained. The coefficient α_1 is the firm-size elasticity of the *ex ante* interim expected payoff from the auction, measuring the quantitative impact of the bidder's relative size on the profitability from the auction. Note that the *ex ante* interim expected payoff $U_t(\cdot)$ is the payoff from participating in an auction given the number of entrants n^s. Based on these values, each potential entrant decides whether to enter a procurement auction taking into account the expected gain of participation and the participation costs as discussed in the next subsection.

6.3 Analysis for an entry equilibrium

Potential bidders continuously participate in an auction until their expected gain from entry exceeds the costs of entry. A crucial assumption in the entry mode is that each firm has a unit production capacity regardless of the group it belongs to.

First, the case in which set-asides are in use is analyzed. LBs may obtain positive profits from participating in a procurement auction since their production capacity is limited, whereas the marginal SB obtains zero expected payoff because of participation by an unlimited number of SB. Therefore, a unique entry equilibrium consists of the numbers of actual bidders in auction s, which satisfy

$$\begin{cases} U_{SB}(n^{L,r}_{SB}, 0) = e \\ U_{LB}(0, n^{H,r}_{LB}) \geq e, \end{cases} \quad \text{subject to } n^{H,r}_{LB} \leq \frac{N_{LB}}{K^H}, \tag{16}$$

if set-asides are implemented. If LBs obtain positive rents, the constraint is binding such that $n^{H,r}_{LB} = N_{LB}/K^H$.

Assuming that the constraint is binding, a counterfactual situation is considered. Let $(n^{L,u}_{SB}, n^{L,u}_{LB})$ and $(0, n^{H,u}_{LB})$ be the number of large and small bidders in auction L and H, respectively, if set-asides were removed (The second superscript u stands for unrestricted participation.). Without set-asides, low-end projects receive bids from large firms as well. The rent of SBs is, however, still fully extracted because of the unlimited number of SBs. Hence, the SBs' optimal entry decision $n^{L,u}_{SB}$ satisfies

$$U_{SB}(n^{L,u}_{SB}, n^{L,u}_{LB}) = e \tag{17}$$

for some $n^{L,u}_{LB}$. Solving (17) for $n^{L,u}_{SB}$ gives the SBs' best response to the LB's entry decision $n^{L,u}_{LB}$. Now, let $\Gamma(n^{L,u}_{LB})$ be the SB's best response function. Since U_{SB} is decreasing in both $n^{L,u}_{SB}$ and $n^{L,u}_{LB}$, $\Gamma'(n^{L,u}_{LB}) < 0$ is obtained.

In addition, the number of large firms in the market is given and finite, and each bidder with a unit production capacity can bid only once. Therefore, the number of large bidders in high-end projects $n^{H,u}_{LB}$ is a decreasing function of $n^{L,u}_{LB}$.

Let $\Lambda(n_{LB}^{L,u})$ be the decreasing function. It can be simply derived explicitly as [33]

$$n_{LB}^{H,u} = \Lambda(n_{LB}^{L,u}) \equiv n_{LB}^{H,r} - \frac{K^L}{K^H} n_{LB}^{L,u}. \tag{18}$$

In equilibrium, the *ex ante* expected gain of each large bidder from participation must be the same between the two projects. Hence, $n_{LB}^{L,u}$ satisfies

$$U_{LB}(\Gamma(n_{LB}^{L,u}), n_{LB}^{L,u}) = U_{LB}(0, \Lambda(n_{LB}^{L,u})) \tag{19}$$

subject to $\ 0 \leq n_{LB}^{L,u}$ and $0 \leq \Lambda(n_{LB}^{L,u})$,

where the left- and right-hand sides, respectively, represent the *ex ante* payoff of an LB from low- and high-end projects. A unique solution of $n_{LB}^{L,u}$ is thus obtained from the equation. Then, the unique $n_{SB}^{L,u}$ and $n_{LB}^{H,u}$ are obtained using $\Gamma(\cdot)$ and $\Lambda(\cdot)$.

6.4   An empirical model for auctions with entry

We first obtain $\Lambda(\cdot)$ empirically. Let ρ be the proportion of L projects in volume in the total budget for civil engineering contracts during the period. According to MLIT (2007), civil engineering projects with their engineer's estimated costs being less than ¥300 million are set aside for SMEs.[34] Consequently, with this model, we consider that a project is H if the engineer's estimated cost is no less than ¥300 million and low-end if the estimated cost is less than ¥300 million. Then, ρ is approximately 59 percent.

Since we normalize the project size as identical for both H and L projects, it is reasonable to consider that the proportion of the number of (normalized) L projects is equal to ρ so that we have $K^L/(K^L + K^H) = \rho$ and $K^H/(K^L + K^H) = 1 - \rho$. Using these, $\Lambda(\cdot)$ is empirically obtained as

$$\Lambda(n_{LB}^{L,u}) = n_{LB}^{H,r} - \frac{\rho}{1-\rho} n_{LB}^{L,u}. \tag{20}$$

The bidders are then divided into either large firms or SMEs. In fact, the distinction between SMEs and large firms in the data is somewhat ambiguous. The set-aside program allows large firms to participate in relatively small projects unless a sufficient competition among SMEs is expected. Consequently, quite a few large firms submit their bids in low-end projects. In addition, some SMEs that met a quality standard were able to participate in some high-end projects. Hence, one dependable way to distinguish these two groups of firms would be to assume that those that bid on high-end projects are large firms and those that bid on low-end projects are SMEs. Since the average scores in high- and low-end projects are 1419.3 and 990.3, respectively, $\bar{X}_{SB} = 990.3$ and $\bar{X}_{LB} = 1419.3$ were set. Table 10 is a summary of the statistics of the bidders' scores in both high- and low-end projects.

The equilibrium participation under the set-aside program, $n_{SB}^{L,r} = 7.46$ and $n_{LB}^{H,r} = 8.49$, is obtained from the data.[35] Then, a counterfactual analysis is conducted to predict $(n_{SB}^{L,u}, n_{LB}^{L,u})$ and $(0, n_{LB}^{H,u})$.

First, the bidders' *ex ante* expected payoff U_k^s is estimated for all k auction samples.[36] Let $b_{(1),k}^s$

---

[33]From Equation (10), we have $n_{LB}^{H,u} = N_{LB}^H/K^H, n_{LB}^{L,u} = N_{LB}^L/K^L$. Since every firm can bid only once, we have

$$n_{LB}^{H,u} K^H + n_{LB}^{L,u} K^L = N_{LB}.$$

Solving for $n_{LB}^{H,u}$ and replacing N_{LB}/K^H with $n_{LB}^{H,r}$ give the decreasing function.

[34]More precisely, ¥300 million is the threshold value with which the government determines whether a project is auctioned for contractors that are grade B or above or C or below. Although the contractors with grade C or below may not satisfy the exact criteria of "SMEs" in Japan, the empirical analysis used here considers them as SMEs for simplicity.

[35]They are estimated by the sample means of the number of bidders in each type of the project as $n_t^s = \sum_{k=1}^m \mathbf{1}\{s_k = s\} \cdot n_{k,t}/K^s$ for each $s \in \{H, L\}$, where $n_{k,t}$ is the number of group t bidders in sample auction k.

[36]As has been discussed, U_k could be structurally estimated. However, solving n_t^s numerically from the obtained V

be the lowest bid in auction k for a type s project, and, with a slight abuse of notation, let $c^s_{(1),k}$ be the cost of the lowest bidder.[37] In addition, let $Q^s(x^s_{(1),k}, n^s_k | \mathbf{z}^s)$ be the prior probability that a bidder with the relative score $x^s_{(1),k}$ becomes the lowest bidder in the auction given that the number of bidders is equal to n^s_k. Then, the bidder's *ex ante* expected profit U^s_k is characterized as

$$U^s_k = \left(b^s_{(1),k} - c^s_{(1),k} \right) Q^s(x^s_{(1),k}, n^s_k | \mathbf{z}^s). \tag{21}$$

In this specification, the probability $Q^s_{(1),k}$ can be identified from a simple linear probability regression model as follows.

Let $Q_{i,k}$ be the index of bidder i in auction k for a type s project where $Q^s_{i,k} = 1$ if the bidder wins and $Q^s_{i,k} = 0$ otherwise. Then, the probability that i wins in the auction is estimated by

$$Q^s_{i,k} = \delta^s_1 \frac{1}{n^s_k} + \delta^s_2 x^s_{i,k} + \delta^s_3 \mathbf{z}^s_k + \nu^s_{i,k}, \tag{22}$$

where $\mathbf{z}^s_k$ controls for scoring auction and project size as $\mathbf{z}^s_k = (\mathrm{SCORE}^s_k, \mathrm{EST}^s_k)$.[38] Table 11 shows the regression results of Equation (22). Fixed effects control the unobserved heterogeneity in project locations. Due to the fact that the mean difference in scores for SMEs is 43 percent lower than that for large firms,[39] it can be concluded that the mean difference in the frequency of winning for SMEs is approximately 5.5 percent lower than that for large firms (t-value : 8.48 with FE).

Let $\hat{\delta}_1, \hat{\delta}_2$ and $\hat{\delta}_3$ denote the least square estimates of (22), and let $x^s_{(1),k}$ denote the normalized score of the lowest bidder in auction k for a type s project. Then, the estimator for the *ex ante* winning probability, $\hat{Q}(\cdot)$ is obtained as

$$\hat{Q}^s(x^s_{(1),k}, n^s_k | \mathbf{z}^s) = \hat{\delta}^s_1 \frac{1}{n^s_k} + \hat{\delta}^s_2 x^s_{(1),k} + \hat{\delta}^s_3 \mathbf{z}^s_k. \tag{23}$$

Since $b^s_{(1),k}$ is observable and $c^s_{(1),k}$ can be replaced with the estimates $\hat{c}^s_{(1),k}$ obtained in Section 5, U^s_k is estimated for all k and s by using (21) as

$$\hat{U}^s_k = \left(b^s_{(1),k} - \hat{c}^s_{(1),k} \right) \hat{Q}^s(x^s_{(1),k}, n^s_k | \mathbf{z}^s).$$

Then, using the estimated $\hat{U}$ and observed data, x, and n, the following linear regression is constructed:

$$\log \hat{U}^s_k = \alpha^s_0 + \alpha^s_1 x^s_{(1),k} + \alpha^s_2 n^s_k + \alpha_3 \mathbf{z}^s_k + \epsilon^s_k,$$

where ϵ^s_k is assumed to be an *i.i.d.*, mean zero random variable. The regression results are shown in Table 12.

Let $\hat{\alpha}^s \equiv (\hat{\alpha}^s_0, \hat{\alpha}^s_1, \hat{\alpha}^s_2, \hat{\alpha}^s_3)$ be the least square estimates of $\alpha^s \equiv (\alpha^s_0, \alpha^s_1, \alpha^s_2, \alpha^s_3)$. Replacing α with $\hat{\alpha}$ and x^s_t with $x_t(n^s_{SB}, n^s_{LB})$ on (15) yields its empirical counterpart as

$$\log U^s_t(n^s_{SB}, n^s_{LB} | \mathbf{z}^s) = \hat{\alpha}^s_0 + \hat{\alpha}^s_1 x_t(n^s_{SB}, n^s_{LB}) + \hat{\alpha}^s_2 n^s + \hat{\alpha}^s_3 \mathbf{z}^s. \tag{24}$$

Specifically, $\hat{\alpha}_1$ measures the marginal effect of the number of bidders on the profitability while $\hat{\alpha}_2$

---

requires an infinitely large number of calculations because of the assumption that the number of bidders is continuous. A reduced-form specification is thus employed to approximate a bidder's probability of winning and the winner's payoff in the counterfactual analysis.

[37] Since our model assumes asymmetric first-price sealed-bid procurement auctions, it is possible that the lowest bidder does not have the lowest signal.

[38] Although percentage bids are used in order to control the project size heterogeneity, there still remains project size heterogeneity in percentage bids. To further eliminate such heterogeneity, EST is included in $\mathbf{z}$.

[39] $\frac{\bar{X}_{BB} - \bar{X}_{SB}}{\bar{X}_{BB}} = 0.39$.

captures the marginal effect of bidder's relative size on the profitability.

It is also used to identify the participation (information acquisition) cost e.[40] As shown in (16), $U_{SB} = e$ holds in equilibrium. The empirical counterpart of the zero-profit condition gives the estimator of the entry cost for a type s project, e^s as:

$$\hat{e}^s = \hat{U}_{SB}^L \left(n_{SB}^{L,r}, 0 | \mathbf{z}^s\right),$$

assuming that the difference in the participation cost between H and L stems only from $\mathbf{z}$.[41] Finally, the individual rationality condition for large firms in the case of set-asides:

$$\hat{U}_{LB}^H \left(0, n_{LB}^{H,r} | \mathbf{z}^H\right) \geq \log \hat{e}^H$$

is checked. The data show that the exponential of the left-hand side of 1.41 percent is greater than that of the right-hand side of 1.14 percent, which suggests that large firms obtain positive gains on average in the procurement market. The individual rationality condition indeed holds.

The above argument can be applied to the scoring auction. The procedure is described in the Appendix.

7 Counterfactual analysis

7.1 The procedure

With (24) and $\hat{e}$, a counterfactual analysis is simply implemented as follows. Replacing e with $\hat{e}$ in Equation (17) and solving for n_{SB} give the explicit form of the best response function

$$n_{SB}^{L,u} = \Gamma(n_{LB}^{L,u}).$$

The complete derivation is provided in the Appendix. Together with $\Lambda(\cdot)$ defined in (20), Equation (19) which characterizes the condition that LBs earn the same *ex ante* expected payoff from H and L projects becomes

$$\hat{\alpha}_0^L + \hat{\alpha}_1^L \cdot x_{LB}(\Gamma(n_{LB}^{L,u}), n_{LB}^{L,u}) + \hat{\alpha}_2^L \cdot [\Gamma(n_{LB}^{L,u}) + n_{LB}^{L,u}] + \hat{\alpha}_3^s \mathbf{z}^L = \frac{1-\rho}{\rho} [\hat{\alpha}_2^H \cdot \Lambda(n_{LB}^{L,u}) + \hat{\alpha}_3^H \mathbf{z}^H], \quad (25)$$

where $\rho \simeq 0.59$. The left-hand side describes the large businesses' *ex ante* expected gain from participating in low-end projects, whereas the right-hand side equals the *ex ante* expected gain from high-end projects. Since low-end projects are greater in value terms than high-end projects, a weight variable $(1-\rho)/\rho \simeq 0.69$ is introduced so that (25) describes an equilibrium in which the gain of a large firm from entering the low-end market is identical to that from entering the high-end market. For simplicity in the calculation, $\Gamma(\cdot)$ and $\Lambda(\cdot)$ are linearized in Equation (25). The details are described in the Appendix.

Finally, the comparative statics of the winning bid are described with respect to the participation restriction. Let p_k be the lowest bid in auction k. p_k is considered to be a random variable conditional on the numbers of bidders, the normalized score of each bidder, and exogenous variables, such as the auction-specific effect.

Hence, assuming that ϵ_p follows an *i.i.d.* distribution, a linear model was established for the lowest

---

[40]Since bids and costs as well as profits here are represented in the percentage of the engineer's estimated cost of the project, the participation cost e is also expressed as the percentage. The participation (information acquisition) cost in dollar terms is a proportion of the engineer's estimated cost of the project.

[41]Since the firm's profit is expressed as a percentage of the engineer's estimated cost of a project, the participation cost e^s for the project is also expressed as the percentage of the engineer's estimated cost. Hence, the e^s is heterogeneous for every project auction, depending on the size and the auction characteristics *i.e.,* a scoring or price-only auction.

bids as

$$p_k = \theta_0 + \theta_1 \cdot x_{(1),k} + \theta_2 \cdot n_k + \theta_3 \cdot \mathbf{y}_k + \epsilon_k^p. \tag{26}$$

$\mathbf{y}_k = (\mathrm{SCORE}_k, \mathrm{EST}_k, \log \mathrm{EST}_k, \mathrm{AUC\_FORMAT})$ controls for auction-specific effects such as a scoring auction option, determination of project size by an engineer's estimate, and choice of open competition or invited bidders, and $x_{(1),k}$ denotes the lowest bidder's normalized score. Then, $\hat{\theta}_1$ measures the difference of the winning bid between LBs and SBs.

Table 13 shows the result of the regression of the lowest bids on x. Fixed effects control the area-specific effects.

Let $\hat{\theta} \equiv \left(\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3 \right)$ be the least square estimates of $\theta \equiv (\theta_1, \theta_2, \theta_3)$. Using these estimates, the mean winning bids are estimated in the low-end projects under the unrestricted participation situation as follows.

It is assumed that the mean winning bids of LBs and SBs can be described by

$$p_t(n_{SB}^{L,u}, n_{LB}^{L,u}) = \hat{\theta}_0 + \hat{\theta}_1 \cdot x_t(n_{SB}^{L,u}, n_{LB}^{L,u}) + \hat{\theta}_2 \cdot (n_{SB}^{L,u} + n_{LB}^{L,u}).$$

Let Q_t be the winning probability of a group t bidder. By using (23), Q_t is obtained empirically as

$$Q_t(n_{SB}^{L,u}, n_{LB}^{L,u}) = \hat{\delta}_1 \frac{1}{n_{SB}^{L,u} + n_{LB}^{L,u}} + \hat{\delta}_2 x_t(n_{SB}^{L,u}, n_{LB}^{L,u}),$$

for each $t \in \{LB, SB\}$. Given the number of bidders $(n_{SB}^{L,u} + n_{LB}^{L,u})$ in a procurement auction for a low-end project, the probability that some LB wins the project is equal to $Q_{LB}(\cdot) \cdot n_{LB}^{L,u}$, and the probability that some SB wins is equal to $Q_{SB}(\cdot) \cdot n_{SB}^{L,u}$. Then, the mean winning bids in the low-end projects are given as

$$p(n_{SB}^{L,u}, n_{LB}^{L,u}) = p_{SB}(n_{SB}^{L,u}, n_{LB}^{L,u}) \cdot Q_{SB}(n_{SB}^{L,u}, n_{LB}^{L,u}) \cdot n_{SB}^{L,u} + p_{LB}(n_{SB}^{L,u}, n_{LB}^{L,u}) \cdot Q_{LB}(n_{SB}^{L,u}, n_{LB}^{L,u}) \cdot n_{LB}^{L,u},$$

and those of the high-end projects are given as

$$p(0, n_{LB}^{H,u}) = p_{LB}(0, n_{LB}^{H,u}),$$

The percentage change of the procurement cost for low-end projects is estimated by

$$p(n_{SB}^{L,u}, n_{LB}^{L,u}) - p(n_{SB}^{L,r}, 0),$$

and that for high-end projects is given by

$$p(0, n_{LB}^{H,u}) - p(n_{LB}^{H,r}, 0).$$

7.2 Results

The empirical results suggest that the set-aside program likely decreases the procurement costs. The counterfactual analysis predicts what the bidder's entry decision and bidding behavior would be were the program to be eliminated. The program yields the competing effects in terms of government procurement costs, the cost reduction in set-aside projects and the cost increase in the remaining projects.

The analysis suggests that, were the program to be eliminated, 2.51 large firms on average would switch their entry from high-end to low-end projects so that their *ex ante* payoff from these two projects must be identical in equilibrium. Since there is a difference in volume for each category of projects, represented by $\rho = 0.59$, the mean number of large firms in low-end projects would be 1.72,

which is obtained by 2.51 times $(1 - \rho)/\rho$.

The serious problem by removing the participation restriction is that the number of participants would decrease in both high- and low-end projects. In high-end projects, the number of large firms would drop from 8.49 to 5.98, which would raise the procurement costs of those projects by 0.99 percent. At the same time, the large firms' participation in low-end projects would depress SME entry into low-end projects. The mean number of SME participants would decline from 7.46 to 5.29.[42] The number of both large-firm and SME participants in low-end projects would drop from 7.46 to 7.01 on average since, according to the static entry model, the participation of one more large firm in the low-end projects would eliminate 1.26 SME participants on average.[43] The procurement costs of low-end projects would fall by 0.40 percent, despite the presence of fewer participants, because of the entry of cost-efficient large firms. The average score of bidders would be increased from 990.3 to 1095.7.

Surprisingly, the resulting lack of competition would drive up government procurement costs. There are two competing effects that set-asides have on government procurement costs, increasing competition versus the participation of cost-inefficient SMEs. Taking also into account the fact that the government spent approximately 60 percent of the procurement budget on low-end projects, the effect of increasing competition would overcompensate for the effect of production inefficiency cost. The counterfactual analysis suggests that set-asides would decrease government procurement costs by 0.17 percent.

It is interesting to observe how the *ex ante* expected profits of large firms are changed by set-asides. Without set-asides, large firms obtain a positive expected gain (2.12 percent of the engineer's estimated cost for each auction), and the net positive gain from entry is almost 0.83 percent of the project's estimated cost. Set-asides completely squeeze the positive net gain from large firms so that the expected gain of large firms with set-asides is almost zero (-0.15 percent). Obviously, this rent extraction from large firms contributes to lowering government procurement costs more than to offset the resulting production cost inefficiency.

8 Discussion

Most small business programs declare that the importance of providing more contract opportunities for disadvantaged businesses lies in the encouragement of their long-run growth. The long-run benefit on an economy has been assumed to outweigh the short-run cost of supporting small businesses. In fact, even in the short run, the program can benefit the procurement buyer, as our analysis has illustrated. Upon designing a public procurement policy, the non-trivial short-run gain should be more carefully considered.

In addition, set-asides are robust against collusion in procurement auctions. Our simulation results indicate that both high-end and low-end auctions receive more participants when set-asides are in use. Obviously, more participants in auctions suggest fewer chances for bidders to be cooperative. Procurement buyers may, therefore, have further short-run benefit from set-asides.

The assumption that the firm has unit production capacity can be relaxed so that multiple units of production and, hence, participating in more than two auctions at the same time are possible without changing the obtained results in this analysis. However, the model does rely on the production capacity, especially, the capacity constraint of the cost-efficient businesses. It is easy to imagine that procurement costs would always be lower by inviting only the cost-efficient firms if their production always exhibits constant returns to scale, although the situation is unrealistic for many procurement buyers.

---

[42]This outcome implicitly assumes that each group of bidders follows a Nash equilibrium bidding strategy. Should the large firms intentionally make a low-ball bid to deter entry by SMEs, the decrease of SMEs would be much more significant.

[43]The coefficient is given by $\gamma = 1.26$.

Regarding the entry model, Li and Zheng (2009) show a series of models of auctions with endogenous participation. Among these, Model 2 is closest to the entry model employed in this paper. However, their method does not apply to our case because it requires the assumption that potential bidders are identical to estimate the distribution of the potential bidder's participation cost.[44] In fact, the empirical specification of the distribution is extremely complicated or even impossible when the asymmetry of bidders varies across the auctions as exhibited in our data.

Due to the limitation, the reduced-form approach is chosen, where the expected gain from participation is assumed to be a decreasing function of the expected numbers of large and small participants in the auction. Although the obtained value of the expected gain using such mean numbers differs from the expected gain from participation computed by integrating all possible profits over the possible numbers of bidders, these two are strongly correlated. Hence, the former satisfies the sufficient condition to become a proxy for the latter. The reduced-form approach is thus chosen with the consideration that it is the most efficient way to predict the potential bidder's participation decisions without small business set-asides.

The entry model assumes a positive participation cost incurred by every bidder. The interpretation of such sunk cost comes from the fact that information acquisition for articipating in a procurement auction is quite a costly process especially for construction projects. Bidders must carefully read drawings and specifications and thoroughly examine them, including all the notes and references. When subcontracts are needed, bidders must solicit bids from subcontractors. The estimation of subcontract costs is not available until all subcontract bids are opened and pre-award subcontract agreements are made with the chosen subcontractors. These processes are extremely time-consuming. Furthermore, the accuracy of estimation is very important since an error could lead to losing the contract or getting a winner's curse. Hence, those sunk costs crucially affect the bidder's participation decision into a procurement auction.

9   Conclusion

Set-asides are widely used in real-world public procurement. The encouragement of SMEs has evoked a controversy on the extent of the extra cost society is paying. However, there is no previous systematic analysis to measure the impact on procurement costs.

In this paper, we provide the first systematic analysis of the effect of small business set-asides on government procurement costs, bidding behaviors, and bidder participation in competitive bidding processes. The simulation study suggests that the program dramatically increases SME participation but is almost neutral with respect to the procurement costs. The production inefficiency caused by set-asides is overcompensated by the increased entry and resulting enhancement of competition by large firms. The set-aside program was observed to increase SME participation in procurement auctions by approximately 30 percent.

The empirical results show that the set-aside program has been successful. It improves equity between advantaged and disadvantaged firms and reduces government procurement costs. The results also suggest that the government cost of set-aside auctions is exaggerated if only the excess amount on contracts allocated to SMEs is considered. The theoretical literature suggests that, despite the efficiency loss, the encouragement of less advantaged bidders in the auction can reduce procurement costs. For instance, Bulow and Roberts (1989) and McAfee and McMillan (1989) insist that bidding credits (or bid discounts in procurement auctions) for disadvantaged bidders increase the auctioneer's welfare, yielding more competitive pressure on advantaged bidders. Similarly, subsidized SMEs drive non-subsidized bidders to give up more of the gain on the contracts they award.

The conclusion also provides an economic rationale on why several countries, such as the United States and Japan, opt out of SMEs from the Government Procurement Agreement (GPA) of the

---

[44]Furthermore, if potential bidders are not symmetric, random participation results in multiple equilibria as shown in Nakabayashi (2010). Empirical specification is, then, impossible or extremely difficult.

World Trade Organization (WTO). Although Article 4 in the GPA prohibits the member countries from giving unfavorable treatment to any company, the set-aside programs are exempted in the GPA Appendix. EU countries have also been renegotiating with the WTO to obtain the exclusion of their SMEs. An important question, however, is whether those practices are robust for corruption or favoritism. Further theoretical and empirical consideration is needed.

A limitation of this study is that it does not consider the long-term effect of set-asides. In the long run, there are positive and negative effects of set-asides on procurement costs. If SMEs could win more auctions, they would have more chances to develop their production skills through learning by doing. On the other hand, subsidization of SMEs may discourage them to develop their businesses to a stage in which they could not be favored in the preference program. Given the sheer volume of public sector procurement, it is clear that more serious research and evaluation are needed to investigate the long-run effect of the set-aside program.

Appendix

Linearization of $\Gamma(\cdot)$

The linear approximation for $n_{SB}^{L,u} = \Gamma(n_{LB}^{L,u})$ at $n_{LB}^{L,u} = 0$ is given by,

$$n_{SB}^{L,u} = \Gamma(0) - \Gamma'(0)n_{LB}^{L,u}.$$

Since $U_{SB}(n_{SB}^{L,r}, 0) = e$ by (17), $\Gamma(0) = n_{SB}^{L,r}$. Therefore,

$$n_{SB}^{L,u} = n_{SB}^{L,r} - \Gamma'(0)n_{LB}^{L,u}. \tag{A-1}$$

To obtain the explicit form of $\Gamma'(0)$, take total derivative of (17) with respect to n_{SB}^L and n_{LB}^L, and the following is obtained

$$0 = \hat{\alpha}_2^L \cdot \frac{\partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)}{\partial n_{SB}^L} \Delta n_{SB}^L + \hat{\alpha}_2^L \cdot \frac{\partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)}{\partial n_{LB}^L} \Delta n_{LB}^L + \hat{\alpha}_3^L \cdot (\Delta n_{SB}^L + \Delta n_{LB}^L). \tag{A-2}$$

By the chain rule, $\partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)/\partial \bar{n}_t^L = \partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)/\partial \bar{X}_L \cdot \partial \bar{X}_L/\partial \bar{n}_{t,L}$ holds for each $t \in \{SB, LB\}$. Since $\partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)/\partial \bar{X}_L = -\bar{X}_{SB}/(\bar{X}_L)^2$, $\partial \bar{X}_L/\partial n_{SB}^L = 0$ and $\partial \bar{X}_L/\partial n_{LB}^L = (\bar{X}_{LB} - \bar{X}_{SB})n_{SB}^L/(n_L)^2$ with $\bar{X}_L = \bar{X}_{SB}$ and $n_L = n_{SB}^L$,

$$\frac{\partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)}{\partial n_{LB}^L} = -\frac{\bar{X}_{LB} - \bar{X}_{SB}}{\bar{X}_{SB} \cdot n_{SB}^L} \text{ and}$$

$$\frac{\partial \bar{x}_{SB}(n_{SB}^L, n_{LB}^L)}{\partial n_{SB}^L} = 0$$

are obtained. Plug them into Equation (A-2), and the following is obtained

$$-\Delta n_{SB}^L = \Gamma'(0)\Delta n_{LB}^L.$$

Hence,

$$\Gamma'(0) = \frac{\hat{\alpha}_2^L}{\hat{\alpha}_3^L} \frac{\bar{X}_{LB} - \bar{X}_{SB}}{\bar{X}_{SB} \cdot n_{SB}^L} + 1 = 1.26$$

is obtained.

Application to scoring auctions:

In particular, it is shown that there is one-to-one mapping between the percentage markup, $b_i - \hat{c}_i$ and the informational rent in the scoring auction, $s_i - \theta_i$:

Proof. Let $\bar{S}$ be the base score which is defined as $\text{Est}/\underline{q}$. Let π_i be the i's conditional informational rent in the scoring auction that satisfies $\pi_i = s_i - \theta_i$.

First, we show that the informational rent measured with the dollar term bid and cost is identical to that measured with the percentage bid and cost. Let $\$b_i$ and $\$c_i$ be the bid and cost in dollar terms. Let $\$s_i$ and $\$\theta_i$ be the scoring bid and the pseudo-type measured with the dollar term bid and cost, which satisfy

$$\$s_i \equiv \frac{\bar{S}}{q_i/\$b_i}, \quad \$\theta_i \equiv \frac{\bar{S}}{q_i/\$c_i}.$$

Note that $\$s_i$ and $\$\theta_i$ are identical to s_i and θ_i.[45] Given q_i, there is one-to-one mapping between the scoring bid and the price-bid and between the pseudo-type and the cost:

$$(s_i, \theta_i) = \left(\frac{b_i \underline{q}}{q_i}, \frac{c_i \underline{q}}{q_i} \right)$$
$$= \frac{\underline{q}}{q_i}(b_i, c_i)$$

Hence, the percentage conditional payoff, $\hat{\pi}_i$, in the scoring auction is estimated by

$$\hat{\pi}_i \equiv b_i - \hat{c}_i = \frac{q_i}{\underline{q}}(s_i - \hat{\theta}_i).$$

$\square$

Using the transformation, the percentage conditional payoff of the winning bidder in the scoring auction is obtained.

References

Asker, John and Estelle Cantillon, "Properties of scoring auctions," *RAND Journal of Economics*, 2008, *39* (1), 69–85.

Athey, Susan, Jonathan Levin, and Enrique Seira, "Comparing Open and Sealed Bid Auctions: Evidence from Timber Auctions," NBER Working Papers, National Bureau of Economic Research, Inc December 2008.

Ayres, Ian and Peter Cramton, "Deficit Reduction Through Diversity: How Affirmative Action at the FCC Increased Auction Competition," Technical Report 1996.

---

[45]

$$(\$s_i, \$\theta_i) \equiv \left(\frac{\bar{S}}{q_i/\$b_i}, \frac{\bar{S}}{q_i/\$c_i} \right)$$
$$= \left(\frac{\underline{q}/\text{Est}}{q_i/(b_i\text{Est})}, \frac{\underline{q}/\text{Est}}{q_i/(c_i\text{Est})} \right)$$
$$= \left(\frac{\underline{q}}{q_i/b_i}, \frac{\underline{q}}{q_i/c_i} \right)$$
$$\equiv (s_i, \theta_i).$$

Hence, the scoring bid and pseudo-type are, respectively, the same regardless of whether they are measured with the percentage bid, b_i, and cost, c_i or with the dollar bid $\$b_i$ and the dollar cost $\$c_i$.

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

__ **and Steven Tadelis**, "Incentives versus Transaction Costs: A Theory of Procurement Contracts," *RAND Journal of Economics*, Autumn 2001, *32* (3), 387–407.

Bannock, Graham, *The Economics of Small Firms: Return from the Wilderness*, Oxford: Basil Blackwell, 1981.

Bulow, Jeremy and John Roberts, "The Simple Economics of Optimal Auctions," *Journal of Political Economy*, October 1989, *97* (5), 1060–90.

Campo, Sandra, Isabelle Perrigne, and Quang Vuong, "Asymmetry in first-price auctions with affiliated private values," *Journal of Applied Econometrics*, 2003, *18* (2), 179–207.

Che, Yeon-Koo, "Design Competition through Multidimensional Auctions," *RAND Journal of Economics*, Winter 1993, *24* (4), 668–680.

Corns, Allan and Andrew Schotter, "Can Affirmative Action Be Cost Effective? An Experimental Examination of Price-Preference Auctions," *American Economic Review*, March 1999, *89* (1), 291–305.

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.

Hansen, Robert G., "Auctions with Endogenous Quantity," *RAND Journal of Economics*, Spring 1988, *19* (1), 44–58.

Jofre-Bonet, Mireia and Martin Pesendorfer, "Estimation of a Dynamic Auction Game," *Econometrica*, 09 2003, *71* (5), 1443–1489.

Krasnokutskaya, Elena, "Identification and Estimation in Highway Procurement Auctions under Unobserved Auction Heterogeneity," *Review of Economic Studies*, 2009. forthcoming.

__ **and Katja Seim**, "Bid Preference Programs and Participation in Highway Procurement Auctions," *working paper*, 2009.

Laffont, Jean-Jacques and Jean Tirole, *A theory of incentives in procurement and regulation*, Cambridge, MA: The MIT Press, 1993.

Levin, Dan and James L Smith, "Equilibrium in Auctions with Entry," *American Economic Review*, June 1994, *84* (3), 585–99.

Li, Tong 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*, August 2007, *91* (7-8), 1591–1624.

Maskin, Eric and John Riley, "Asymmetric Auctions," *Review of Economic Studies*, July 2000, *67* (3), 413–38.

McAfee, R. Preston and John McMillan, "Bidding for Contracts: A Principal-Agent Analysis," *RAND Journal of Economics*, Autumn 1986, *17* (3), 326–338.

_ **and** _ , "Auctions and Bidding," *Journal of Economic Literature*, June 1987, *25* (2), 699–738.

_ **and** _ , "Auctions with entry," *Economics Letters*, 1987, *23* (4), 343–347.

_ **and** _ , "Government procurement and international trade," *Journal of International Economics*, May 1989, *26* (3-4), 291–308.

Milgrom, Paul, *Putting Auction Theory to Work*, Campridge University Press, 2004.

MLIT, Ministry of Land, Infrastructure, Transportation and Tourism, *Construction Contracts: Annual Report: Tyokkatsu Koji Keiyaku Kankei Shiryou* 2007. available at http://www.mlit.go.jp/chotatsu/contractsystem/keiyaku/.

Nakabayashi, Jun, "Equilibria in Asymmetric Auctions with Entry," January 2010.

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

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

Zhang, Bin and Kemal Guler, "Nonparametric estimation of asymmetric first price auctions: A simplified approach," *Economics Letters*, September 2005, *88* (3), 318–322.

Zheng, Charles Z., "High Bids and Broke Winners," *Journal of Economic Theory*, September 2001, *100* (1), 129–171.
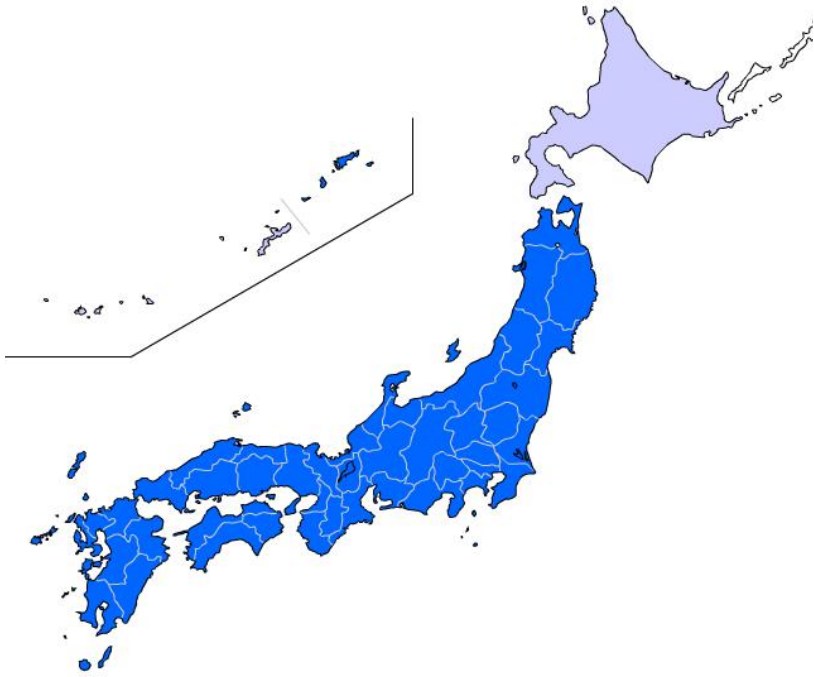
Figure 1: Data area
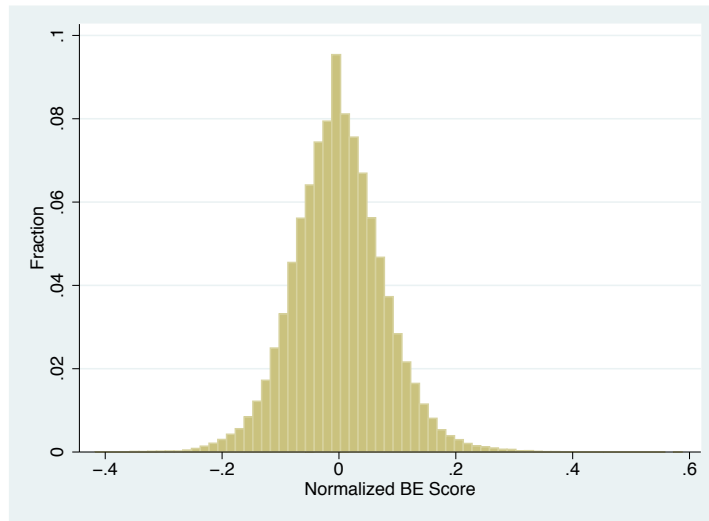
Figure 2: Normalized score of actual bidders
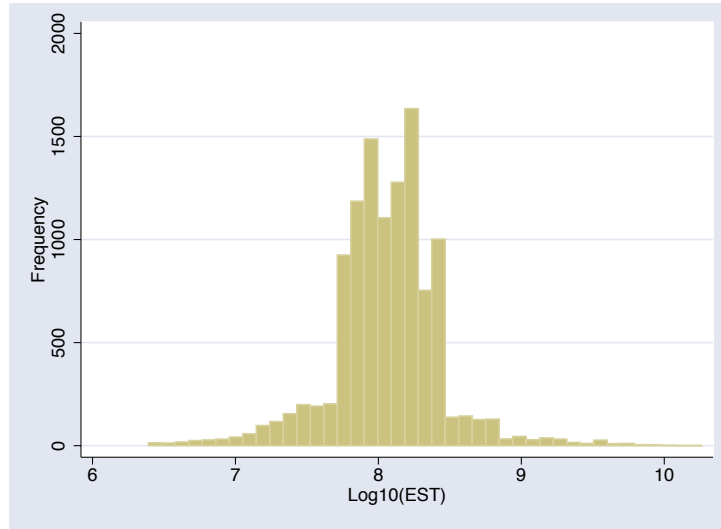
Figure 3: Project size ($\log_{10}$ of the engineer's estimate)
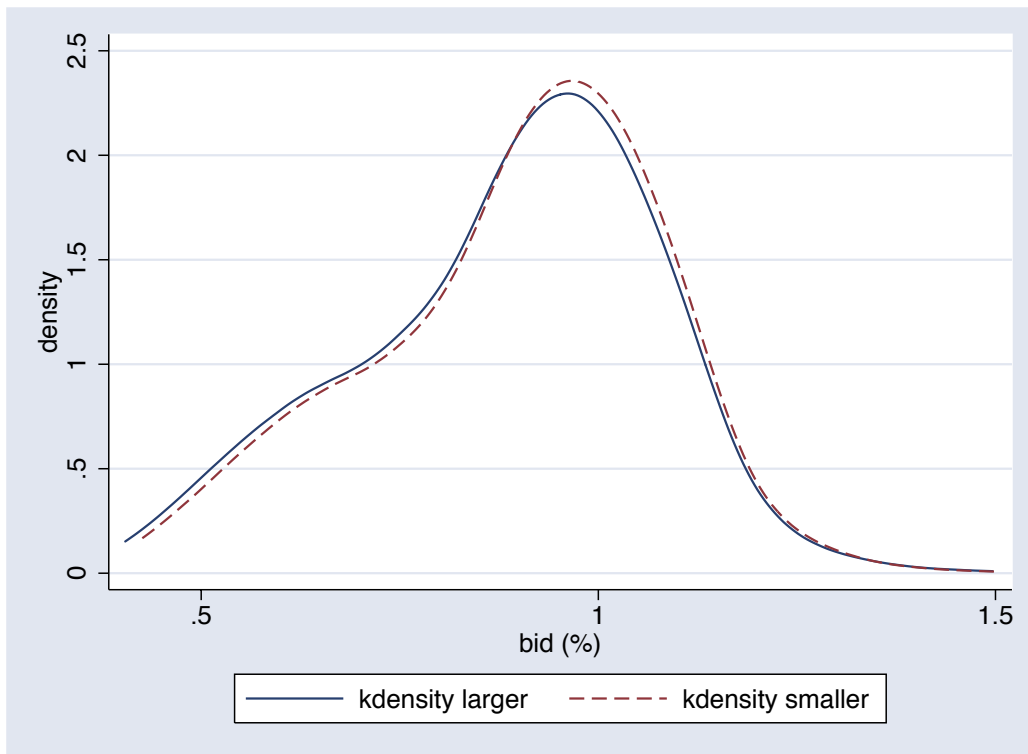
Figure 4: Densities (Percentage bids)

Figure 5: BE score of actual bidders

Figure 6: BE score of all firms on the certified contractor lists

Figure 7: BE score of LB bidders

Figure 8: BE score of all LBs on the lists

31

Figure 9: BE score of SB bidders

Figure 10: BE score of all SBs on the lists

Figure 11: Few bidders: $n = 5$

Figure 12: Many bidders: $22 \leq n \leq 28$

Figure 13: The model of auctions with entry

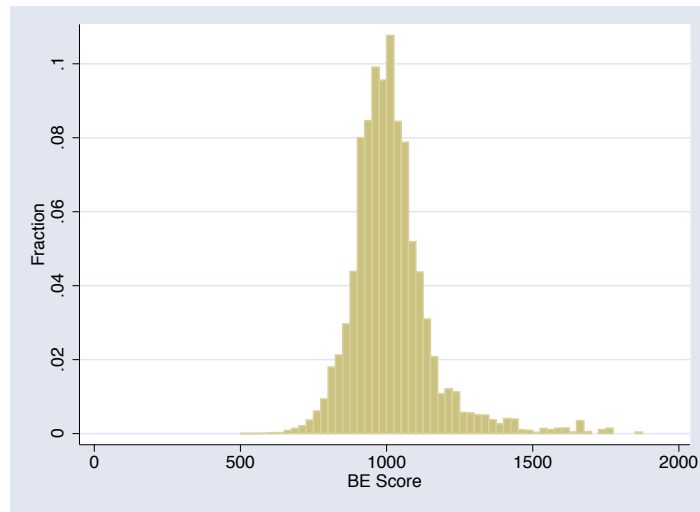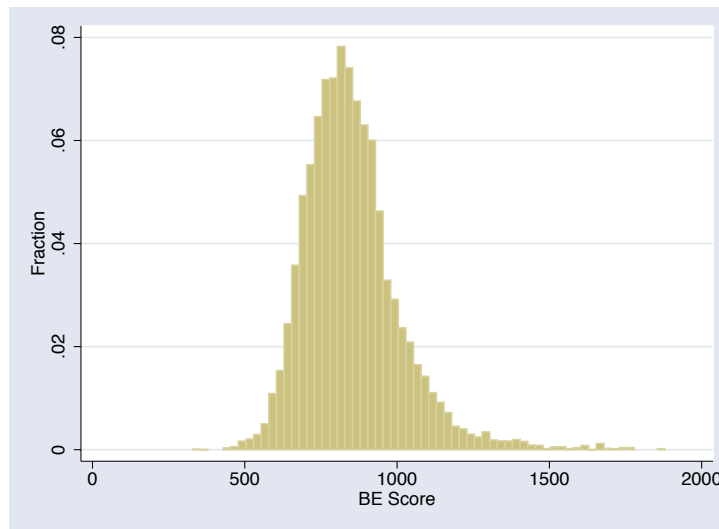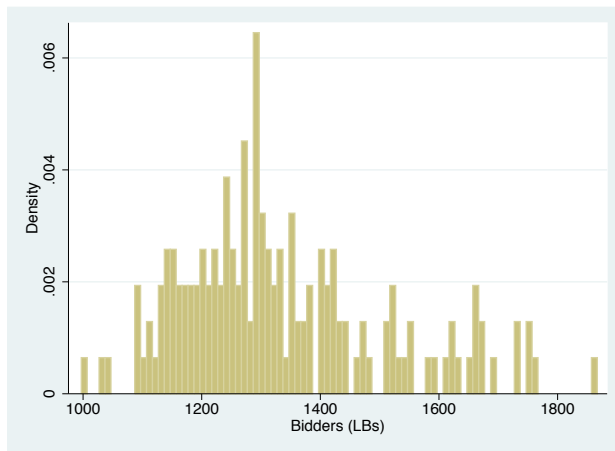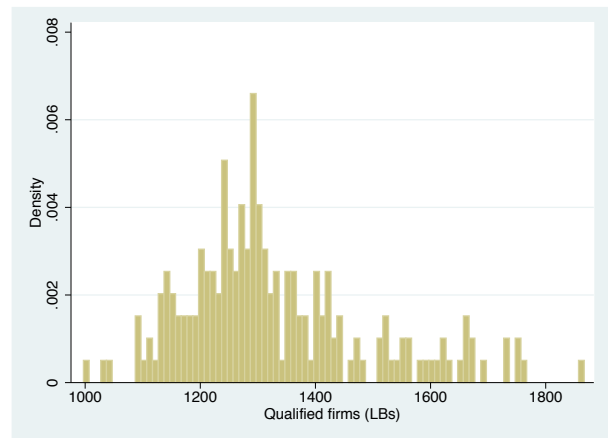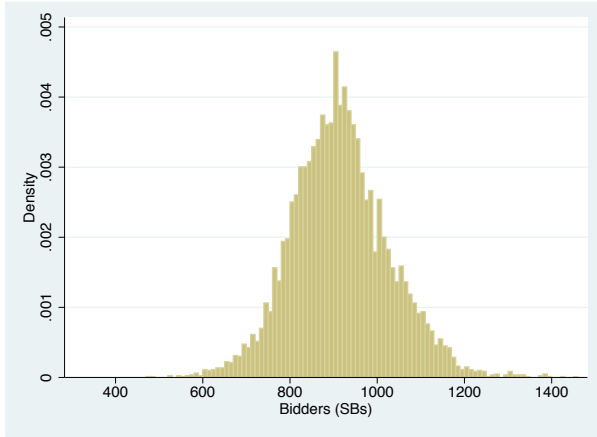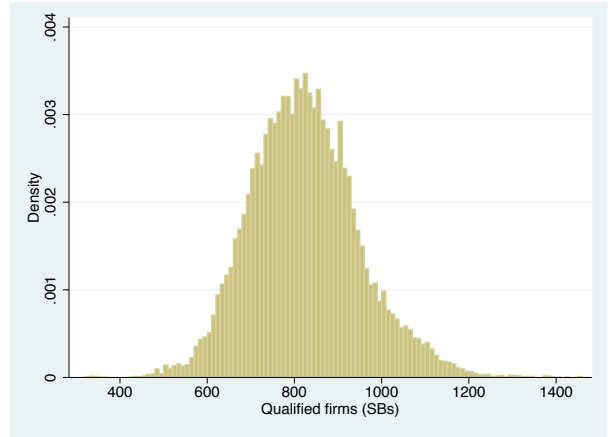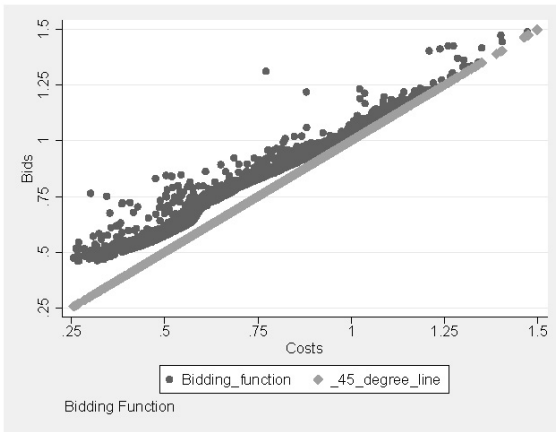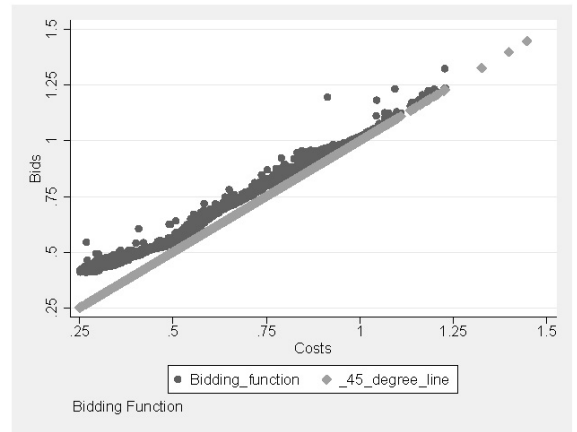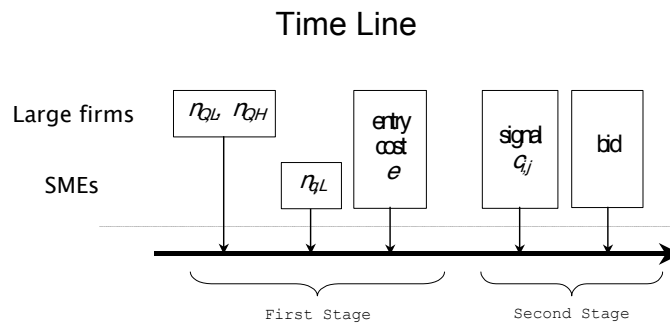|	Project A, Est. ≤ 730		Project B, 300 < Est. ≤ 730		Project C, 60 < Est. ≤ 300		Project D, Est. ≤ 60	
	obs.	percentage	obs.	percentage	obs.	percentage	obs.	percentage
Firm's letter-grade								
A	3,269	86.92 %	781	14.5 %	128	0.12 %	0	0.00 %
B	430	11.43 %	4,218	78.31 %	2,588	2.38 %	16	0.12 %
C	62	1.65 %	387	7.19 %	103,093	94.99 %	4,365	31.51 %
D	0	0.00 %	0	0.00 %	2,716	2.5 %	9,472	68.38 %
Total	3,761	100.00 %	5,386	100.00 %	108,525	100.00 %	13,853	100.00 %

Observations count the total number of participants by the letter grade.

Table 1:

	FY2008				FY2007			
Category	Amount	(%)	Count	(%)	Amount	(%)	Count	(%)
General & Heavy	963,309	(57.1)	4,219	(38.5)	806,667	(54.3)	4,092	(36.4)
Paving	140,078	(8.3)	846	(7.7)	113,252	(7.6)	789	(7.0)
Bridge	95,166	(5.6)	184	(1.7)	118,931	(8.0)	218	(1.9)
Landscaping	10,485	(0.6)	276	(2.5)	10,858	(0.7)	313	(2.8)
Architecture	63,880	(3.8)	607	(5.5)	60,817	(4.0)	674	(6.1)
Painting	7,089	(0.4)	219	(2.0)	6,552	(0.4)	223	(2.0)
Maintenance	172,592	(10.2)	2,396	(21.9)	151,119	(10.2)	2,535	(22.5)
Dredging	5,885	(0.3)	24	(0.2)	5,929	(0.4)	25	(0.2)
Machinery & equipment	32,108	(1.9)	473	(4.3)	27,355	(1.8)	434	(3.9)
Info & telecom facility	41,644	(2.5)	733	(6.7)	43,824	(3.0)	731	(6.5)
Others	154,412	(9.2)	987	(9.0)	139,645	(9.4)	1,209	(10.7)
Total	1,686,658	(100.0)	10,964	(100.0)	1,484,949	(100.0)	11,243	(100.0)

	FY2006				FY2005			
Category	Amount	(%)	Count	(%)	Amount	(%)	Count	(%)
General & Heavy	606,795	(47.9)	4,068	(34.7)	626,493	(53.3)	3,752	(32.2)
Paving	132,088	(10.4)	862	(7.3)	123,043	(10.5)	799	(6.9)
Bridge	96,592	(7.6)	208	(1.8)	30,407	(2.6)	148	(1.3)
Landscaping	10,301	(0.8)	311	(2.7)	13,524	(1.2)	356	(3.1)
Architecture	51,404	(4.1)	692	(5.9)	42,670	(3.6)	717	(6.1)
Painting	6,800	(0.5)	231	(2.0)	6,009	(0.5)	241	(2.1)
Maintenance	141,751	(11.2)	2,718	(23.2)	138,855	(11.8)	2,919	(25.0)
Dredging	4,792	(0.4)	19	(0.2)	4,050	(0.3)	21	(0.2)
Machinery & equipment	37,102	(2.9)	560	(4.8)	37,221	(3.2)	594	(5.1)
Info & telecom facility	37,102	(2.9)	848	(7.2)	43,657	(3.7)	998	(8.6)
Others	140,766	(11.1)	1,216	(10.4)	108,786	(9.3)	1,118	(9.6)
Total	1,265,492	(100.0)	11,733	(100.0)	1,174,715	(100.0)	11,663	(100.0)

Table 2: Projects yearly (¥million)

		FY2005-2008			
Category	Amount	(%)	Count	(%)	
General & Heavy	3,003,265	(53.5)	16,131	(35.4)	
Paving	508,461	(9.1)	3,296	(7.2)	
Bridge	341,096	(6.1)	758	(1.7)	
Landscaping	45,168	(0.8)	1,256	(2.8)	
Architecture	218,771	(3.9)	2,690	(5.9)	
Painting	26,460	(0.5)	914	(2.0)	
Maintenance	604,318	(10.8)	10,568	(23.2)	
Dredging	20,656	(0.4)	89	(0.2)	
Machinery & equipment	133,786	(2.4)	2,061	(4.5)	
Info & telecom facility	166,226	(3.0)	3,310	(7.3)	
Others	543,610	(9.7)	4,530	(9.9)	
Total	5,611,817	(100.0)	45,603	(100.0)	

Table 3: Projects from FY2005 through FY2008 (¥million)

		No. Obs.	Mean	Std.Dev.	Coeff. of Var.	Max	Min
BE Score	: $X_{i,k}$	130,050	1,020.04	155.27	0.132	1,859	475
Normalized Score	: $x_{i,k}$	130,050	-0.001	0.0778	-	0.57	-0.42

Table 4: Summary statistics : The BE score of actual bidders

		¥ Million*			
Project	No.	Engineer's Estimated Costs			
Size	Observation	Mean	Std.Dev.	Max	Min
730 or more	333	2,127.56	1,718.66	17,600	737
300 - 730	654	474.24	117.19	725	300
60 - 300	12,203	144.78	63.99	299	60.01
less than 60	1,830	38.191	16.95	59.99	.0075
Total	15,020	190.10	400.60	17,600	.0075

*The amount of money is based on the engineer's estimate.

Table 5: Summary statistics on project size

		Bids			Costs		
		OLS	FE	RE	OLS	FE	RE
$x_{i,k}$		-0.027	-0.023	-0.023	-0.033	-0.029	-0.029
		(4.55)**	(6.15)**	(6.31)**	(5.04)**	(6.57)**	(6.73)**
Auction date		-0.000	-	-	-0.000	-	-
		(121.38)**			(118.99)**		
Scoring auction dummy		-0.067	-	-	-0.077	-	-
		(32.32)**			(33.74)**		
Auction form dummy 2		0.019	-	-	0.021	-	-
		(6.16)**			(6.07)**		
Auction form dummy 3		0.044	-	-	0.042	-	-
		(22.08)**			(19.01)**		
Auction form dummy 4		0.020	-	-	0.020	-	-
		(7.95)**			(7.95)**		
Constant		129.877	0.856	0.841	141.618	0.813	0.783
		(122.22)**	(3004.12)**	(613.72)**	(119.72)**	(2436.29)**	(496.88)**
Observations		110416	110324	110324	110416	110324	110324
R-squared		0.36	0.00	-	0.35	0.00	-
Number of auctions		-	14857	14857	-	14857	14857

Absolute value of t statistics in parentheses; * significant at 5%; ** significant at 1%
$x_{i,k}$: % difference from average bidders in auction k

Table 6: Regression results of normalized bids and estimated costs

Letter grade		Obs	BE scores			
			Mean	Std. Dev.	Min	Max
Large firms	Actual	155	1333.4	173.25	997	1859
	Registered	197	1327.4	160.85	997	1859
Small firms	Actual	7,986	915.0	112.03	475	1460
	Registered	31,535	823.1	125.53	329	1460

Actual large (small) firms are the bidders in high-(low-)end projects. Registered large (small) firms are the firms with letter grades A and B (C and D).

Table 7: BE scores of actual and registered bidders

Letter grade		Obs	BE scores			
			Mean	Std. Dev.	Min	Max
A	Actual	34	1,580.97	138.49	1239	1859
	Registered	37	1,580.51	132.71	1239	1859
B	Actual	118	1,264.09	104.99	997	1526
	Registered	160	1,268.88	97.74	997	1526
C	Actual	3,829	987.52	93.30	668	1460
	Registered	5,148	997.55	97.48	668	1460
D	Actual	3,766	846.62	79.12	475	1086
	Registered	26,387	789.09	99.42	329	1118

Table 8: BE scores of actual and registered bidders (by letter grade)

	OLS	Robust OLS	FE
x_{ij}	0.342	0.342	0.292
	(5.23)**	(4.90)**	(4.81)**
No. Bidders	-0.096	-0.096	-0.091
	(83.91)**	(45.70)**	(77.90)**
Auction date	0.256	0.256	
	(25.57)**	(24.82)**	
EST_k	0.006	0.006	
	(8.89)**	(4.11)**	
$logEST_k$	-0.01	-0.01	
	(1.14)	(0.76)	
Scoring auction dummy	0.098	0.098	
	(2.76)**	(2.03)*	
Auction format dummy 1	0.245	0.245	
	(6.41)**	(4.97)**	
Constant	-2.264	-2.264	-1.406
	(13.86)**	(9.27)**	(8.56)**
Observations	10709	10709	10709
R-squared	0.48	0.48	
No. auctions			204

Absolute value of t statistics in parentheses

* significant at 5%; ** significant at 1%

FE controls auction-specific effects. Invited bidders are excluded.

Table 9: Regression result for bid margins

Project Category	Mean	No. obs	Std. dev.	Max.	Min.
Low-end < ¥300 mn	990.3	121,046	100.98	1,750	475
High-end ≥ ¥300 mn	1419.3	8,977	199.65	1,859	799
Total	1,019.78	131,552	151.63	1,859	475

Table 10: Project category

		OLS	Robust OLS	FE
x_{ij}	(δ_1)	0.128	0.128	0.128
		(8.48)**	(8.48)**	(8.48)**
(No. Bidders)$^{-1}$	(δ_2)	1.019	1.019	1.014
		(104.46)**	(111.68)**	(94.60)**
Auction date		0.002	0.002	0.004
		(0.89)	(0.91)	(1.63)
EST_k		-0	-0	0
		(0.14)	(0.15)	(0.25)
logEST_k		-0	-0	-0.001
		(0.26)	(0.27)	(0.57)
Scoring auction dummy		0.005	0.005	0.006
		(0.63)	(0.68)	(0.74)
Auction format dummy 1		-0.006	-0.006	-0.011
		(0.71)	(0.75)	(1.26)
Constant				0.026
				(0.57)
Observations		79327	79327	79327
R-squared		0.25	0.25	0.11

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%
Except invited bidders

Table 11: Regression result for the linear probability model

	low-end: s = L		*high-end*: s = H	
	OLS	Robust OLS	OLS	Robust OLS
x_{ij} (α_1^s)	1.083	1.083	2.835	2.835
	(12.20)**	(12.14)**	(9.22)**	(9.59)**
No. Bidders (α_2^s)	-0.238	-0.238	-0.155	-0.155
	(152.00)**	(65.88)**	(31.92)**	(19.83)**
Scoring auction dummy $(\alpha_{3,1}^s)$	0.716	0.716	0.601	0.601
	(22.73)**	(17.51)**	(5.26)**	(5.29)**
Estimated cost $(\alpha_{3,2}^s)$	-0.029	-0.029	0.006	0.006
	(3.00)**	(2.86)**	(2.80)**	(2.88)**
Constant (α_0^s)	-3.103	-3.103	-3.577	-3.577
	(81.93)**	(56.97)**	(30.32)**	(27.55)**
Observations	9782	9782	927	927
R-squared	0.72	0.72	0.54	0.54

Absolute value of t statistics in parentheses, * significant at 5%; ** significant at 1%
Invited bidders are excluded

Table 12: Regression result for expected payoffs

	OLS	Robust OLS
x_{ij} (θ_1)	-0.07	-0.063
	(5.28)**	(5.53)**
No. bidders (θ_2)	-0.003	-0.004
	(14.74)**	(18.67)**
Auction date	-0.138	-0.143
	(68.07)**	(68.07)**
EST_k	0.001	0.001
	(5.74)**	(5.74)**
$logEST_k$	-0.066	-0.062
	(39.29)**	(39.29)**
Scoring auction dummy	-0.065	-0.051
	(8.91)**	(8.91)**
Auction format dummy 1	0.002	0.01
	(0.25)	(0.25)
Constant	2.227	2.155
	(70.21)**	(70.21)**
Observations	11453	11453
R-squared	0.46	0.53

Absolute value of t statistics in parentheses
* significant at 5%; ** significant at 1%
Except invited bidders

Table 13: Regression result for lowest bids

| | Set-asides | | Unrestricted | |
Project category	Low	High*	Low	High*
Mean no. small bidders	7.46	0	5.29	0
Mean no. large bidders	0	8.49	1.72	5.98
Mean no. total bidders	7.46	8.49	7.01	5.98
Mean Scores	990.3	1419.3	1095.7	1419.3
Procurement cost change	-	-	-0.40%	0.99%
Overall effect	-	-	0.17%	
Project volume (¥ bn.) Sum of engineer's estimates (Share %)	1854.97 (59.0)	1273.03 (41.0)	1854.97 (59.0)	1273.03 (41.0)
Entry costs (% of engineer's estimates)	1.29%	1.14%	1.29%	1.14%
Profits (large firms) (% of engineer's estimates)	-	1.41%	1.96%	2.12%

*High-end projects are those in which the engineer-estimated cost is no less than ¥300 million.

Table 14: Estimation for the effect of set-asides