Samuel Spreen Group 9

ECON4020

Professor Aryal

04/30/2020

Abstract:

This paper analyzes the effects of small business bid preference programs in California Department of Transportation (Caltrans) procurement auctions with respect to the procurement cost of the auctioneer and bidder competition within each auction. Utilizing a dataset provided by Caltrans, we were able to recover the cost distributions of both small and large business bidders and use these distributions to run a counterfactual scenario to determine the auction outcomes where the small business bid preference program was not present. Analyzing the differences in outcomes between our observed and counterfactual cases, we found that the small business bid preference program decreased the overall procurement cost of Caltrans by 11.60 percent. This shows that the implementation of a small business bid preference program changed the bidding behavior of participating firms in such a way that auction participants bid more aggressively and drove down the value of the overall winning bid benefiting Caltrans with a lower procurement cost.

An Analysis of Bid Preference Programs:

Procurement auctions are frequently used by governments and businesses to license out contracts with unknown underlying valuations. In a procurement auction, bidders submit bids that represent the amount of money they receive for supplying a good or service to the auctioneer. Because the auctioneer wants the contract completed at the lowest possible price, procurement auctions are reverse auctions in which the firm with the lowest submitted bid wins the contract and is paid the amount they bid to complete it. Additionally, many procurement auctions utilize bid preference programs in order to assist favored firms. Bid preference programs improve the bids of favored firms when determining the winner of the auction but do not affect the actual payout received by the winning bidder when a contract is awarded. For example, The California Department of Transportations (Caltrans) implements a bid preference program when they award highway construction contracts; this bid preference program takes the form of a 5 percent discount applied to all bids submitted by small business bidders. Given this procurement auction framework, the 5 percent discount is only applied to small business bids when determining the winner of the auction and these bidders receive the full value of their bid if they win an auction. The goal of the Caltrans bid preference program is to increase competition between small and large business bidders in their highway procurement auctions. The aim of this paper is to identify how the Caltrans bid preference program changes the total procurement costs faced by Caltrans as well as to quantify how the bid preference program changes the competitiveness of small business bidders within these auctions.

In order to qualify as a small business, bidders participating in Caltrans highway procurement auctions must have less than 100 employees, be independently owned and or operated, and have average annual gross receipts of 10 million dollars or less over their past 3 years of operation. If a business does not meet one of these requirements, they are classified as a large business and do not receive the 5 percent bid preference when participating in the procurement auctions. The rationale for providing bid preference programs stems from two factors: small businesses are often run by and employ a greater percentage of underrepresented groups and by awarding contracts to small businesses, the procurement dollars are more likely to circulate within the local economy than if a large business bidder won the contract. Thus, from the viewpoint of Caltrans, there is a positive externality generated by awarding a greater percentage of these auctions to small businesses. Despite this positive externality, it is important to analyze the effects of Caltrans' bid preference program for several reasons. To begin, by implementing bid discounts that impact the competitive auction environment, Caltrans is directly altering the payout awarded by each procurement auctions and without further analysis, it is unclear how the bid preference program is altering total procurement costs. For example, this bid preference program could be increasing the overall procurement cost faced by Caltrans resulting in the misappropriation of millions of taxpayer dollars. Additionally, without identifying how the results of these highway procurement auctions change under the bid preference program, we cannot be certain that the bid preference program is actually fulfilling its original purpose: to increase competition between small and large business bidders. If further analysis proves that the Caltrans bid preference program does not generate the desired outcome, then a new bid preference system should be implemented in its place. For example, a lump-sum subsidy provided to small businesses or an adjustable discount with a percentage that changes based on the value of the specific contract could prove effective. Although this paper will only explore the effectiveness of the current 5 percent bid discount program, it is important to make note of these alternative bid preference programs in the event that the current Caltrans system is ineffective.

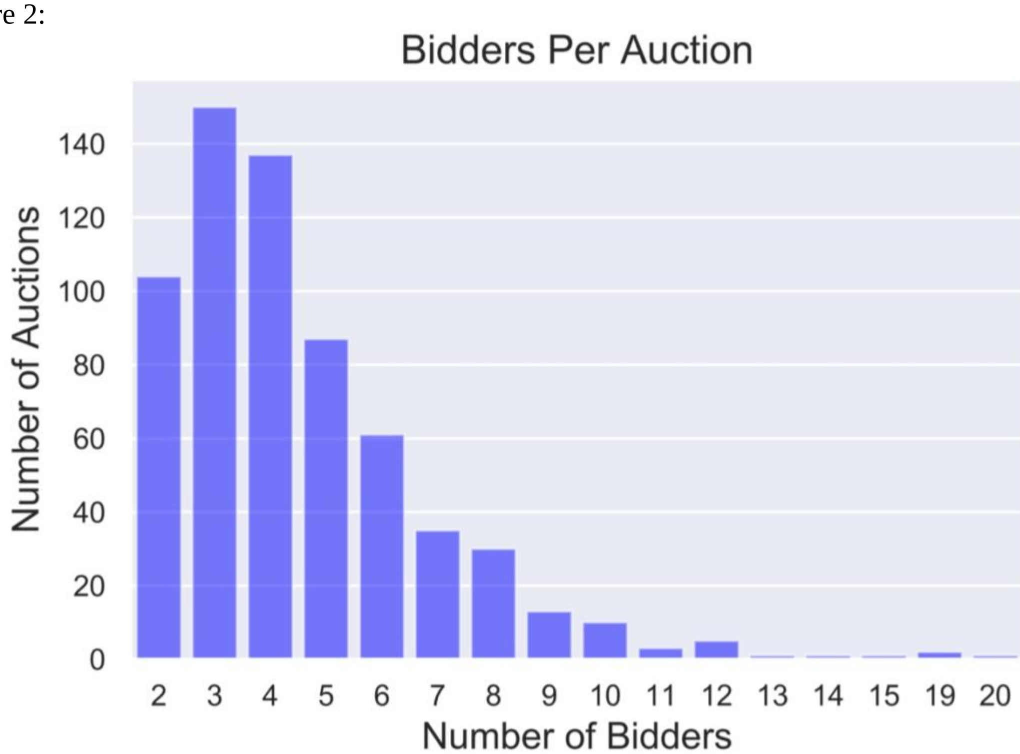
Exploratory Data Analysis:

The data set we will be working with throughout this paper consists of 705 different procurement auctions -identified by unique project IDs- for highway construction projects run by the California Department of Transportation. In our exploratory data analysis, we removed several auctions from the data set due to various data inconsistencies. These inconsistencies included auctions with a single bidder, auctions with repeated bidder entries, and auctions with outlier bids we classified as being either below the 1 percent bid quantile or above the 99 percent bid quantile. It is important that the entire auction is removed from the dataset and not just the erroneous bid entry because by removing a single bid, we are changing the competitive format of the auction which could lead to incorrect results. After removing these entries, there were 641 remaining auctions with 501 unique companies bidding of which 243 were classified as small businesses and received preference from Caltrans in the form of a 5 percent discount on their bids. Each auction also had an associated cost estimate assessed by Caltrans engineers as well as an estimated number of work days the project would take to complete. On average, there were 4.53 bidders per auction and the following five-number summary and bar graph reveal the distribution of bidders per auction.

Figure 1:

Number of Bidders				
Min	Quartile 1	Median	Quartile 3	Max
2	4	5	7	20

Figure 2:



Given figure 1 and figure 2, we can see that the distribution of the number of bidders is right skewed with a majority of auctions having between 2 and 5 bidders while there are very few auctions that have in excess of 9 bidders. It is important to note the most frequent number of bidders in figure 2 because in the estimation and counterfactual analysis sections of this paper, we will subset the data based on the number of auction participants and we will want to choose a bidder combination that provides a sufficiently large sample of auctions for estimating their cost distributions.

Looking at the bid distribution of all firms, figure 3 indicates that there is an incredibly wide range between bids across all auctions with a maximum bid of over 11 million and a minimum of just over 100 thousand. As a result, depending on the combination of bidders chosen for estimation, we may need to remove more auctions due to their outlying bid values in comparison to the overall distribution.

Figure 3:

Distribution of All Bids				
Min	Quartile 1	Median	Quartile 3	Max
103,845.00	288,362.25	467,452.50	738,929.13	11,777,770.00

Separating the bids by the business designations reveals that the frequency of bids submitted by both large and small businesses is comparable up until approximately log 5.5 (base 10) where the frequency of small business bids drops off significantly in comparison with the bids submitted by large businesses (figure 4). Additionally, both of these histograms are right skewed with the distribution of large business bids having more extreme values than the distribution of small business bids. Given that the Caltrans procurement auctions are reverse auctions where the lowest bid wins, the disparity between the frequencies of small and large business bids beyond log 5.5 does not indicate that large businesses are bidding higher than small businesses in a specific auction. Instead, it is reflective of the fact that larger firms can bid on projects with higher cost estimates because, unlike smaller businesses, they have the capital and resources to undertake and complete these projects. Furthermore, the five number summaries of the bids broken out by small businesses in figure 5 and large businesses in figure 6 support this observation; we can see that the maximum bid for small businesses is nearly half the value of the maximum bid for large businesses indicating that small businesses aren't bidding in auctions with higher cost estimates.

Figure 4:

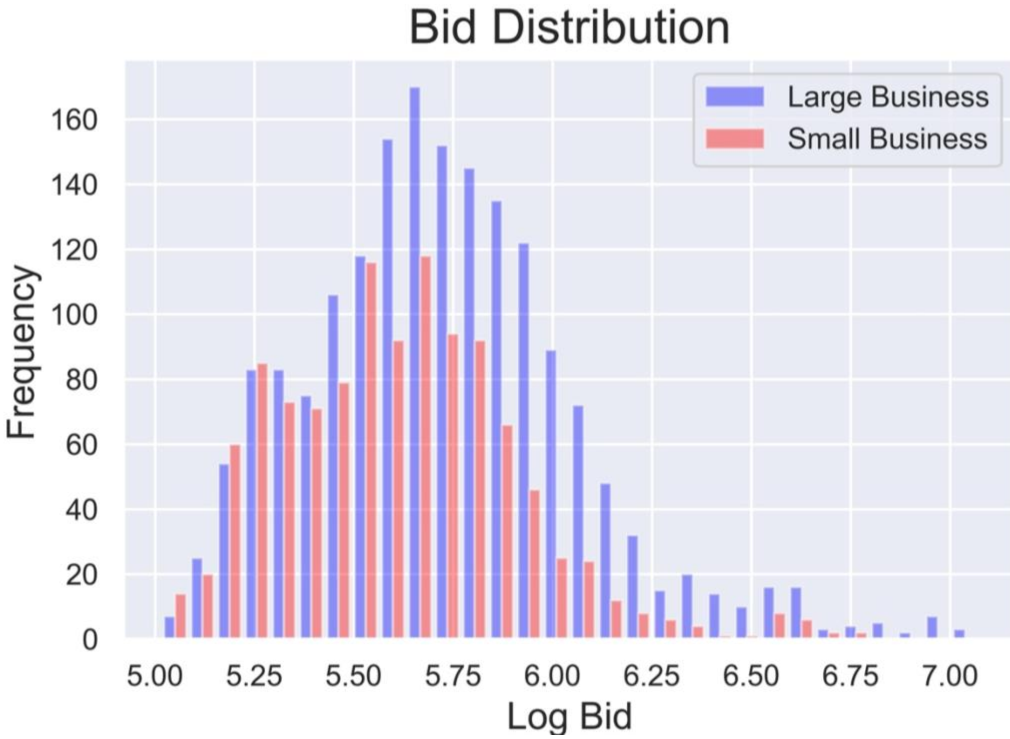


Figure 5:

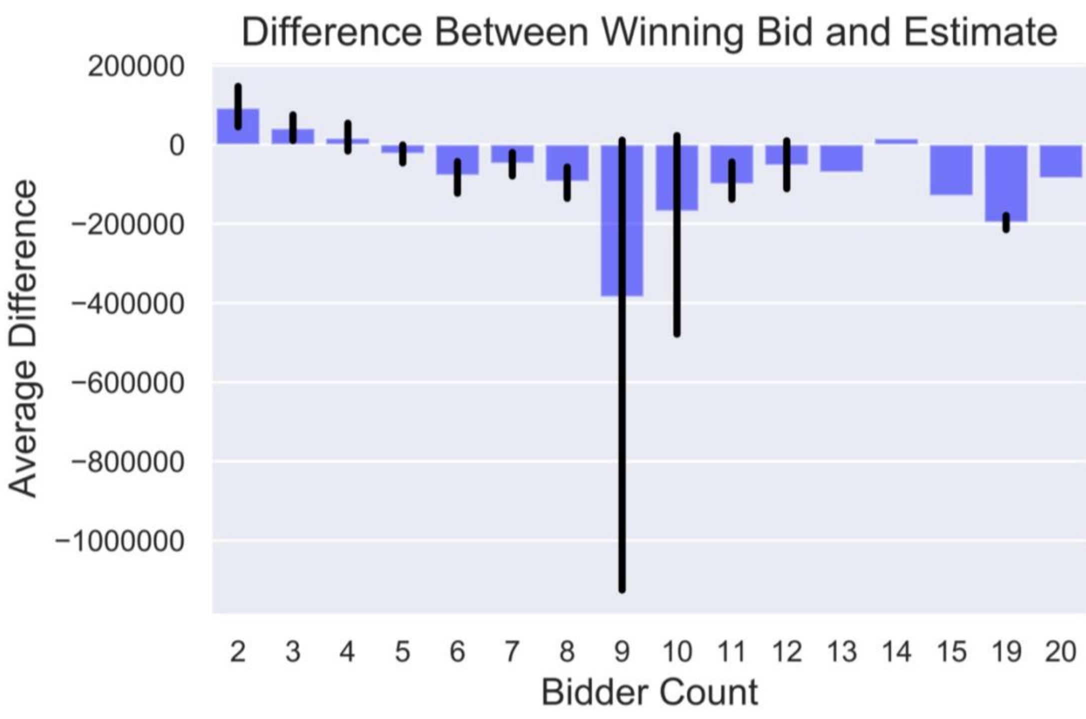
Distribution of Small Business Bids				
Min	Quartile 1	Median	Quartile 3	Max
103,845.00	245,786.00	395,526.00	601,299.00	5,965,853.00

Figure 6:

Distribution of Large Business Bids				
Min	Quartile 1	Median	Quartile 3	Max
106,210.00	318,227.00	514,209.00	846,282.00	11,777,770.00

Another important relationship to visualize is that between the value of bids submitted and the number of bidders in an auction, figure 7 shows the average difference between the winning bid and the engineers' cost estimate for each auction with the same number of bidders. The black lines on this bar chart indicate the range of the aggregated data and help show the variation of the differences between the winning bid of each auction and its respective cost estimate; the bars without the black lines indicate that only one auction occurred at that number of bidders. This bar chart shows that as the number of bidders increases, the average difference between the winning bid amount and the cost estimate decreases and eventually becomes negative which indicates that the winning bidders are -on average- bidding less than the estimated cost. Figure 7 reflects an important function of procurement auctions: price discovery. For example, of all the auctions with 9 bidders, the average difference between the winning bid and the cost estimate is -$384,830.77 indicating that the winning bidders were able to complete these projects at a lower cost than the value that Caltrans initially assessed. Without using these auctions to award construction contracts, Caltrans could have easily assigned contracts to companies with a payout equivalent to the engineers' estimate which would have resulted in the winners of these 9 bidder auctions being overpaid by $384,830.77 on average and thus increasing Caltrans' procurement costs.

Figure 7:



Furthermore, the output from an ordinary least square's regression of the winning bids on the number of small business bidders, the number of large business bidders, and the project cost estimate in figure 8 reveals two negative and statistically significant coefficients on the number of small and large business bidders. The regression summary output indicates that as the number of bidders in an auction increases, holding the estimated cost constant, the value of the winning bid submitted decreases. These negative regression coefficients are reflective of competition that occurs in first price auctions because as the number of bidders increases, the value of bids approaches the true private value of the good or service and thus drives down the procurement cost. Additionally, the negative coefficient on the number of large business bidders of -42,680 is larger in absolute value than the coefficient on the number of small business bidders of -18,730. This disparity is reflective of the advantages of returns to scale and other factors that allow larger firms to bid less and deliver the same goods and services at a lower price than smaller firms.

Figure 8:

```
                            OLS Regression Results
==============================================================================
Dep. Variable:                    Bid   R-squared:                       0.892
Model:                            OLS   Adj. R-squared:                  0.892
Method:                 Least Squares   F-statistic:                     1755.
Date:                Mon, 10 Feb 2020   Prob (F-statistic):          2.08e-307
Time:                        22:49:20   Log-Likelihood:                -8902.0
No. Observations:                 641   AIC:                         1.781e+04
Df Residuals:                     637   BIC:                         1.783e+04
Df Model:                           3
Covariance Type:            nonrobust
==============================================================================
                 coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
Intercept   2.003e+05   2.44e+04      8.204      0.000    1.52e+05    2.48e+05
SmallBiz   -1.873e+04   5570.464     -3.362      0.001   -2.97e+04   -7791.878
LargeBiz   -4.268e+04   6749.947     -6.323      0.000   -5.59e+04   -2.94e+04
Estimate       0.9199      0.013     70.151      0.000       0.894       0.946
==============================================================================
Omnibus:                      472.899   Durbin-Watson:                   1.506
Prob(Omnibus):                  0.000   Jarque-Bera (JB):           150790.360
Skew:                          -2.128   Prob(JB):                         0.00
Kurtosis:                      78.018   Cond. No.                     2.51e+06
==============================================================================
```

Theoretical Model:

Considering an auction with $n \geq 2$ bidders where $n = n_s + n_l$ with n_s being the number of small business bidders and n_l being the number of large business bidders, we observe two distinct cost distributions such that n_s small business bidders follow the distribution $C \sim F_s(\cdot)$ and n_l large business bidders follow the distribution $C \sim F_l(\cdot)$. Given that these bidders exhibit asymmetric preferences, by rule, $F_s(\cdot) \neq F_l(\cdot)$ which can be explained by the capabilities and resources available to small and large businesses that differ and impact their cost functions. For any bidder i that bids b_i with cost C_i , that bidder makes a profit equivalent to $(b_i - C_i)$ if she wins the auction. If bidder i is a large business bidder than she wins only if all other $n_l - 1$ bidders bid more than her bid b_i and all n_s bidders bid more than her bid multiplied by 1.05 to account for the discount applied to the bids of small businesses. Utilizing this notation, we can show that the probability that bidder i wins with bid b_i is as follows:

$$\Pr(i \text{ wins with bid } b_i) = \Pr(\,(n_l - 1) \text{ bid} \geq b_i) * \Pr(\,n_s \text{ bid} \geq 1.05 * b_i) \tag{1}$$

$$= [1 - F_l\left(\beta_l^{-1}(b_i)\right)]^{(n_l-1)} * [1 - F_s\left(\beta_s^{-1}(1.05 * b_i)\right)]^{(n_s)} \tag{2}$$

If bidder i is a small business bidder than she wins only if all other $n_s - 1$ bidders bid more than her bid b_i and all n_l large bidders bid more than her bid divided by 1.05 to account for the small business discount. The probability of winning for a small business bidder is as follows:

$$\Pr(i \text{ wins with bid } b_i) = \Pr(\,(n_s - 1) \text{ bid} \geq b_i) * \Pr(\,n_l \text{ bid} \geq (b_i\,/\,1.05)\,) \tag{3}$$

$$= [1 - F_l\left(\beta_l^{-1}(b_i/1.05)\right)]^{(n_l)} * \; [1 - F_s\left(\beta_s^{-1}(b_i)\right)]^{(n_s-1)} \tag{4}$$

We can now combine the probability of a win seen in equations 2 and 4 with the payoff of a win which is the profit the business makes $(b_i - C_i)$ and create two maximization problems seen below in equations 5 and 6.

$$max_{b_i}\{\,(b_i - C_i) * [1 - F_l\left(\beta_l^{-1}(b_i)\right)]^{(n_l-1)} * [1 - F_s\left(\beta_s^{-1}(1.05 * b_i)\right)]^{(n_s)}\,\} \tag{5}$$

$$max_{b_i}\{\,(b_i - C_i) * [1 - F_l\left(\beta_l^{-1}(b_i\,/\,1.05)\right)]^{(n_l)} * [1 - F_s\left(\beta_s^{-1}(b_i)\right)]^{(n_s-1)}\,\} \tag{6}$$

Taking the derivative of these equations with respect to b_i brings us to the first order conditions in equations 7 and 8 seen below.

$$1 = (b_i - C_i)\left(\frac{(\,n_l - 1) * f_l\left(\beta_l^{-1}(b_i)\right)}{\left(1 - F_l\left(\beta_l^{-1}(b_i)\right)\right) * \beta_l'\left(\beta_l^{-1}(b_i)\right)} + \frac{1.05 * n_s * f_s\left(\beta_s^{-1}(1.05 * b_i)\right)}{\left(1 - F_s(\beta_s^{-1}(1.05 * b_i)\right) * \beta_s'\left(\beta_s^{-1}(1.05 * b_i)\right)}\right) \tag{7}$$

$$1 = (b_i - C_i)\left(\frac{n_l * f_l\left(\beta_l^{-1}(b_i/1.05)\right)}{1.05\left(1 - F_l\left(\beta_l^{-1}(b_i/1.05)\right)\right) * \beta_l'\left(\beta_l^{-1}(b_i/1.05)\right)} + \frac{(n_s - 1) * f_s\left(\beta_s^{-1}(b_i)\right)}{\left(1 - F_s\left(\beta_s^{-1}(b_i)\right)\right) * \beta_s'\left(\beta_s^{-1}(b_i)\right)}\right) \tag{8}$$

Here, equation 7 represents the evaluated first order condition for large business bidders and equation 8 represents the evaluated first order condition for small business bidders.

In order to derive the cost distributions of the small and large business bidders, we can introduce the random variable G with the property that:

$$(1 - F_s\left(\beta_l^{-1}(b)\right)) = (1 - G_l(b|n_l, n_s) \qquad (9)$$

$$(1 - F_s(\beta_s^{-1}(b))) = (1 - G_s(b|n_l, n_s) \qquad (10)$$

The random variable $G_{l,s}(b|n_l, n_s)$ is the cumulative density function (CDF) of the bids for either large or small businesses given an auction with n_l large business bidders and n_s small business bidders. Furthermore, in order to calculate probability density function (PDF) of the bids $g_{l,s}(b|n_l, n_s)$ we take the derivative of the CDF of $G_{l,s}(b|n_l, n_s)$ in order to reach to reach the following equations for g:

$$g_l(b|n_l, n_s) = \ f_l\left(\beta_l^{-1}(b)\right) * \frac{1}{\beta_l'(\beta_l^{-1}(b))} \qquad (11)$$

$$g_s(b|n_l, n_s) = \ f_s\left(\beta_s^{-1}(b)\right) * \frac{1}{\beta_s'(\beta_s^{-1}(b))} \qquad (12)$$

Utilizing $G_l(b|n_l, n_s)$ and $g_l(b|n_l, n_s)$ estimated from the bid distributions, we can now substitute these densities into equation 7 for all terms that use F_l, f_l and β_l likewise using $G_s(b|n_l, n_s)$ and $g_s(b|n_l, n_s)$ in equation 8 for all terms that use F_s, f_s and β_s. After completing these substitutions and solving for C_i, we can arrive at the following two equations to recover a bidder's estimated cost:

$$C_i = b_i - \frac{1}{\left(\frac{(n_l - 1) * g_l(b_i|n_l, n_s)}{\left(1 - G_l(b_i|n_l, n_s)\right)} + \frac{1.05(n_s) * g_s(1.05b_i|n_l, n_s)}{\left(1 - G_s(1.05b_i|n_l, n_s)\right)}\right)} \qquad (13)$$

$$C_i = b_i - \frac{1}{\left(\frac{(n_l) * g_l(b_i/1.05|n_l, n_s)}{1.05\left(1 - G_l(b_i/1.05|n_l, n_s)\right)} + \frac{(n_s - 1) * g_s(b_i|n_l, n_s)}{\left(1 - G_s(b_i|n_l, n_s)\right)}\right)} \qquad (14)$$

Using equation 13 for large business bidders and equation 14 for small business bidders, we can apply these equations to the bids of each business of each respective type in order to calculate the cost C_i for that business and compile the individual costs for all small and large business bidders in order to arrive at the respective cost distributions $C \sim F_l(\cdot)$ and $C \sim F_s(\cdot)$.

Empirical Model:

When deriving the cost distributions $C \sim F_l(\cdot)$ and $C \sim F_s(\cdot)$ from the Caltrans data, we will need to make several adjustments to equations 13 and 14 of the theoretical auction model in order to facilitate accurate estimations. To begin, as shown in the exploratory data analysis section of the paper, the associated engineers' estimate for the projects within the Caltrans auctions varies wildly and we must condition our densities on this engineers' estimate x in order to accurately evaluate the density function at any respective bid. As a result, the random variables $G_{l,s}$ and $g_{l,s}$ will be conditioned on the engineers' estimate x so that now their densities are represented as $G_{l,s}(b|x, n_l, n_s)$ and $g_{l,s}(b|x, n_l, n_s)$. After implementing these changes to equations 13 and 14, the following equations will now be used to estimate the costs within our dataset.

$$C_i = b_i - \frac{1}{\left(\frac{(n_l - 1) * g_l(b_i|x, n_l, n_s)}{\left(1 - G_l(b_i|x, n_l, n_s)\right)} + \frac{1.05(n_s) * g_s(1.05b_i|x, n_l, n_s)}{\left(1 - G_s(1.05b_i|x, n_l, n_s)\right)}\right)} \qquad (15)$$

$$C_i = b_i - \frac{1}{\left(\frac{(n_l) * g_l(b_i/1.05|x, n_l, n_s)}{1.05\left(1 - G_l(b_i/1.05|x, n_l, n_s)\right)} + \frac{(n_s - 1) * g_s(b_i|x, n_l, n_s)}{\left(1 - G_s(b_i|x, n_l, n_s)\right)}\right)} \qquad (16)$$

Estimation Strategy:

In order to carry out our estimations of the individual bidder costs C_i and the distributions by bidder type $C \sim F_l(\cdot)$ and $C \sim F_s(\cdot)$, we will need to estimate our random variables $G_{l,s}(b|x, n_l, n_s)$ $g_{l,s}(b|x, n_l, n_s)$ using kernel density estimation. In kernel density estimation, we can use a programming language of choice in order to estimate the probability density function and cumulative density function of a random variable in a non-parametric manner; this means that we are using the sample data of our random variable in order to make inferences about the overall population density functions for that random variable. In the case of the Caltrans dataset, we will be estimating the kernel density of the conditional distribution of bids with respect to the engineers estimate x as well as the number of small bidders n_s and the number of large bidders n_l. After creating these densities for both large and small business bidders, we will estimate a bidder's cost by using the following procedure:

1) Identify a combination of n_s small business and n_l large business bidders and subset the data to that specific combination

2) Remove the outliers from this subset of auctions in order to achieve a more normal distribution that we estimate using kernel density estimation. Outliers are generally classified as being the upper quantiles or the right tales of the bid densities.

3) Take the bid of a business b_i, the engineers' estimate of the project being bid on x, the number of small business bidders n_s and the number of large business bidders n_l as inputs and evaluate equation 15 if the bidder is a large business and equation 16 if the bidder is a small business. Our result from this step is the cost C_i of that bidder within the evaluated auctions.

4) Compile the costs for large business bidders and the costs for small business bidders in order to achieve the cost distributions $C \sim F_l(\cdot)$ and $C \sim F_s(\cdot)$.

Estimation Results:

When estimating the cost distributions $C \sim F_l(\cdot)$ and $C \sim F_s(\cdot)$ we controlled for the engineers' estimate by focusing on all auctions around the median engineers estimate x. Following the procedure outlined in the previous section, the cost distributions $C \sim F_l(\cdot)$ and $C \sim F_s(\cdot)$ conditioned on the median estimate x were estimated and their associated PDFs and CDFs can be found in figures 9, 10 and 11. One issue that was encountered during estimation was that we recovered several negative costs and decided to drop the corresponding auctions associated with these negative cost estimates. Looking at the overlaid CDFs of the small and large business recovered cost densities in figure 11, we can see that the CDF of the large business bidders' cost density lies to the left of the CDF of the small business bidders' cost density function. This means that for any evaluated cost C_i within the cost densities $F_l(\cdot)$ and $F_s(\cdot)$, the probability that the cost is less than or equal to that evaluated cost C_i is greater for large businesses than for small businesses. Theoretically, this makes sense because we expect large businesses to be more efficient and deliver the same projects at lower costs meaning for any left tail probability at an evaluated cost θ, the probability of $F_l(\theta) > F_s(\theta)$.

Figure 9:

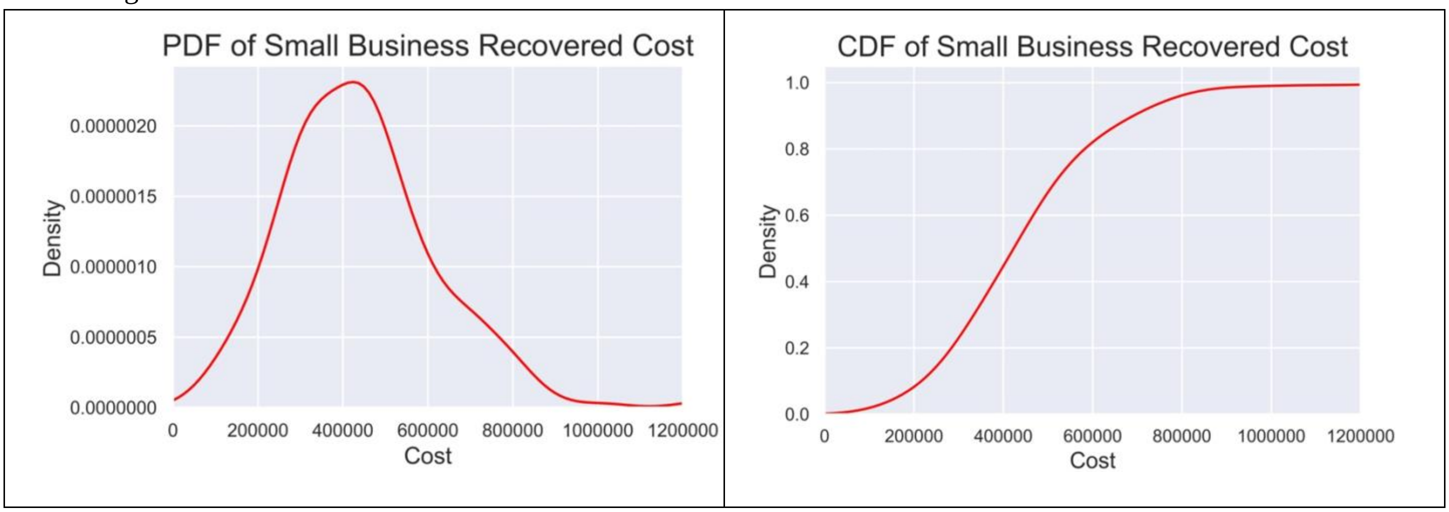


Figure 10:

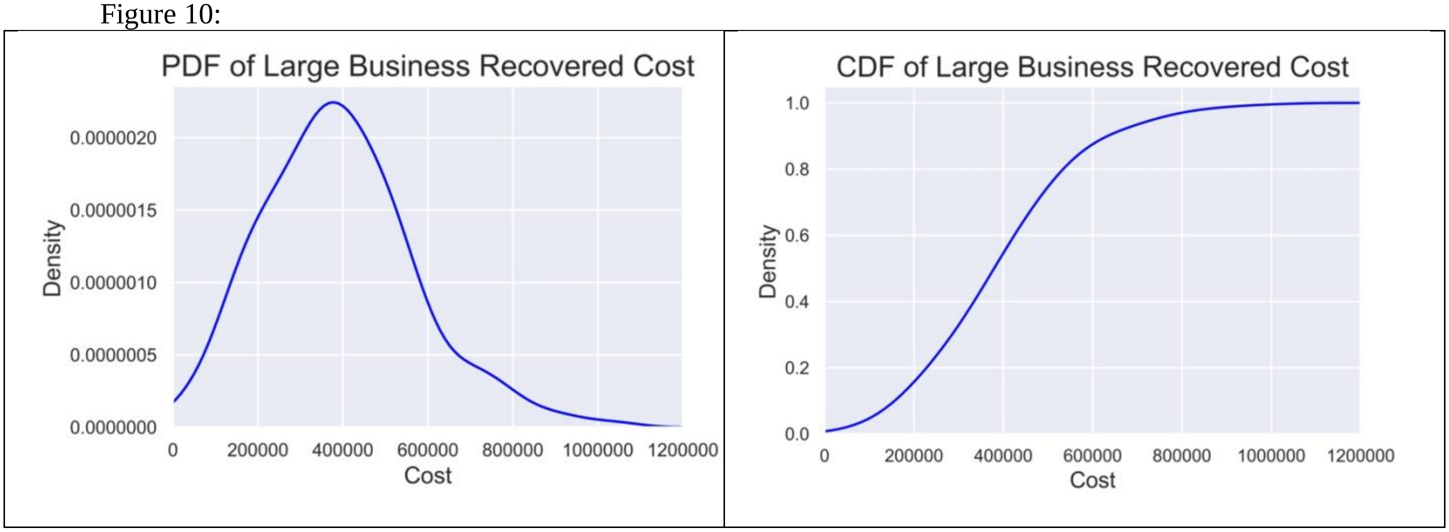
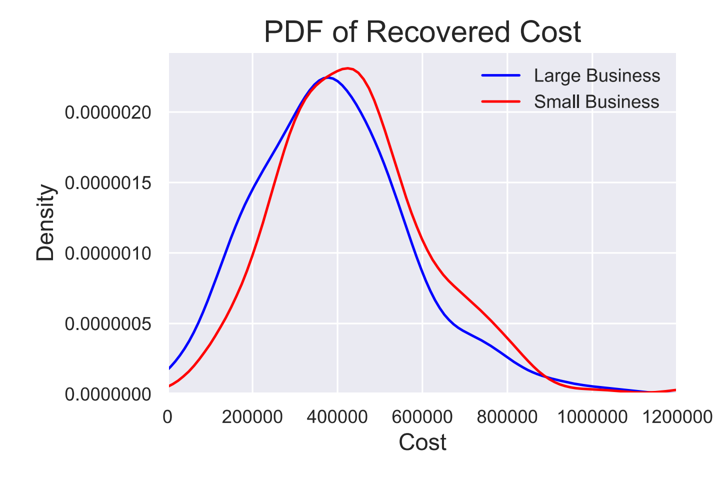
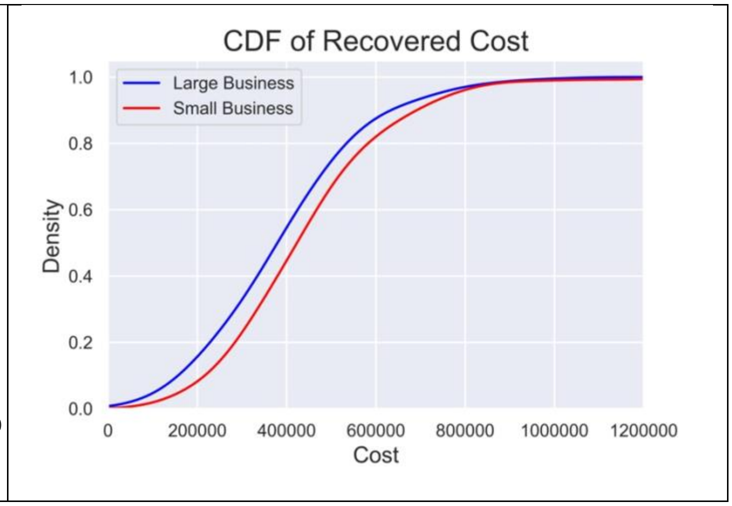


Figure 11:





Counterfactual Analysis:

In order to determine the effect of the small business bid preference program on the overall procurement cost faced by Caltrans, we need to map the recovered cost of each firm back to a new bid, b_i', where b_i' is the bid of the same businesses in the same auction if there was not a bid discount. In order to do so, we will evaluate the symmetric auction cases where one of either of n_s or n_l is equal to zero and estimate their respective large and small bidder cost distributions like we did for our asymmetric auctions. However, because there is only on type of bidder within the symmetric case, the bid preference program does not apply because all bidders are of the same type. This means that in order to recover the cost of a bidder we use the equation 17 which is the symmetric bidder cost recovery equation.

$$C_i = b_i - \frac{1}{(n-1)} * \frac{1 - G(b|n)}{g(b|n)} \qquad (17)$$

It is also worth noting that equation 17 can be viewed as a reduced form -or special case- of equation 15 or 16 where one of either n_s or n_l is equivalent to zero because we are looking at auctions with only one type of bidder. Additionally, the bid discount is no longer applied meaning $\delta = 0$ instead of $\delta = 5$ which changes the scaling coefficient of 1.05 to 1.00 because without a discount the bids are not scaled when ordering the auction results.

After applying equation 17 to our symmetric auctions and retrieving the cost estimates for all bidders within these auctions, we now have a group of several hundred pairs of bids b_i and costs C_i. We then use this symmetric auction data to construct two multiple linear regressions, one for auctions with only small bidders and one for auctions with only large bidders. Both regressions follow the form:

$$b_i = \beta_1(C_i) + \beta_2\left(C_i^2\right) + \beta_3(n) + \beta_4(x) \quad (18)$$

In this regression equation, we are regressing a business's bid b_i on its recovered cost C_i, its recovered cost squared C_i^2, the total number of bidders within the auction n and the engineers estimate x in order to map the relationship between these variables within a symmetric auction. The inclusion of the cost squared variable is used to reconcile the fact that many cost equations are nonlinear and this coefficient helps us capture the effects of a business's changing marginal cost. After running the regression seen in equation 18 for both small and large business bidders, we now have four β coefficients for small business bidders and four β coefficients for large business bidders. We treat these symmetric cases with their fitted models as the training data in a machine learning problem where we train a model on a training set of data before making predictions on a test set of data that the model has not encountered while being trained. Having trained our models for the symmetric bidding behavior of small and large business bidders, we then use the values of C_i, C_i^2, n, and x from our asymmetric auctions to predict a bid $b_i{}'$ where this bid represents what business i would have bid given that there was no preference program. In summary, within our two regression models, the symmetric auctions are used as the training data for these models while the asymmetric auctions are used as the test data. We have now predicted each firm's bid without the preference program and we observe their bid with the preference program within the data which means we can evaluate how the results of the Caltrans procurement auctions change in the presence of the preference program.

Counterfactual Results:

To conduct the counterfactual analysis, we focused on two specific bidder combinations: auctions with $n_s = 2$ and $n_l = 2$ as well as auctions with $n_s = 2$ and $n_l = 3$. We decided to focus on these combinations of bidders because we wanted a combination of n_s and n_l where the removal of outliers would not result in a sample that was too small to estimate using kernel density estimation. After the removal of outlying auctions from these two bidder combinations, there were a total of 26 auctions and 117 bids to use for counterfactual analysis. For each of the 26 auctions used in our counterfactual analysis, we compared the winning bids under the bid preference program with the counterfactual winning bid generated when there was not a discount given to small businesses and reach the results seen in the figure 12.

Figure 12:

	Total Procurement Cost	Small Business Wins	Average Small Business Bid	Average Large Business Bid
Without Bid Preference	9,746,417	12	409,597	401,206
With Bid Preference	8,615,837	12	404,838	380,0595
Percentage Difference	-11.60%	0%	-1.16%	-5.14%

Looking at the results in figure 12, we can see that the total procurement cost faced by Caltrans -which is defined as the summation of the winning bids- decreased by 11.60 percent when the bid preference program was introduced. Given this significant decrease in the total procurement cost for these auctions, it appears that the bid preference program is functioning as intended with respect to a decrease in procurement cost. When considering the bid preference program as a technique to increase competition between small and large business bidders, within the evaluated auctions, we can see that the number of small business wins for these contracts did not change meaning that the bid preference program did not change any procurement outcomes. The fact that the number of small business wins did not increase within this dataset could be attributed to the fact that the average large business bid decreased by 5.14 percent in the presence of the bid preference program which could have rendered the effects of the bid discount negligible at best.

Limitations:

Although our counterfactual results indicate that the total procurement cost of Caltrans decreased due to the implementation of the bid preference program and that this program did not assist small businesses in winning more procurement contracts, the generalization of these results raises a few issues. For example, the counterfactual analysis used only looked at two bidding combinations with a sample size of 26 auctions meaning that the results could change in relation to the 641 auctions within the entire dataset. Additionally, the accuracy of kernel density estimation varies drastically based on the smoothing parameter, also referred to as *bandwidth*, that is used in the estimation. Due to time constraints, the bandwidth parameter used was “normal reference” which assumes that the underlying density being estimated is approximately normal. It is possible that a more precise bandwidth parameter exists and that the use of the “normal reference” bandwidth parameter oversmoothed the data reducing its variance while simultaneously increasing its bias which would result in inaccurate estimations. Lastly, auctions with negative cost estimates were dropped from the analysis; however, the utilization of a log transformation could have helped reduce the number of negative recovered costs allowing for more data to be used in our estimation and counterfactual results.

Works Cited:

1. Duong, Tarn. *An Introduction to Kernel Density Estimation*, www.mvstat.net/tduong/research/seminars/seminar-2001-05/.
2. Krasnokutskaya, Elena, and Katja Seim. "Bid Preference Programs and Participation in Highway Procurement Auctions." *American Economic Review*, vol. 101, no. 6, 2011, pp. 2653–2686., doi:10.1257/aer.101.6.2653.
3. Marion, Justin. "Are Bid Preferences Benign? The Effect of Small Business Subsidies in Highway Procurement Auctions." *Journal of Public Economics*, vol. 91, no. 7-8, 2007, pp. 1591–1624., doi:10.1016/j.jpubeco.2006.12.005.
4. "Statsmodels.nonparametric.kernel_density.KDEMultivariateConditional¶." *Statsmodels*, www.statsmodels.org/stable/generated/statsmodels.nonparametric.kernel_density.KDEMultivariateConditional.html.

Appendix:

All of the code used for this project can be found at: https://github.com/samspreen/ECON4020.git