w o r k i n g
p a p e r

11 02

Disadvantaged Business Enterprise Goals in Government Procurement Contracting: An Analysis of Bidding Behavior and Costs

Dakshina G. De Silva, Timothy Dunne, Georgia Kosmopoulou, Carlos Lamarche

FEDERAL RESERVE BANK OF CLEVELAND

Working Paper 11-02 January 2011

Disadvantaged Business Enterprise Goals in Government Procurement Contracting: An Analysis of Bidding Behavior and Costs

Dakshina G. De Silva, Timothy Dunne, Georgia Kosmopoulou, Carlos Lamarche

Programs that encourage the participation of disadvantaged business enterprises (DBE) as subcontractors have been a part of government procurement auctions for over three decades. In this paper, we examine the impact of a program that requires prime contractors to subcontract out a portion of a highway procurement project to DBE firms. We study how DBE subcontracting requirements affect bidding behavior in federally funded projects. Within a symmetric independent private value framework, we use the equilibrium bidding function to obtain the cost distribution of firms undertaking projects either with or without subcontracting goals. We then use nonparametric estimation methodsto uncover and compare the cost of firms bidding on a class of asphalt projects related to surface treatment in Texas. The analysis shows little differences in the cost structure between auctions that have subcontracting goals and those that do not.

JEL codes: H4, H57, D44.

Keywords: Government procurement, auctions.

Dakshina G. De Silva is at Texas Tech University and can be reached at dakshina.de-silva@ttu.edu. Timothy Dunne is at the Federal Reserve Bank of Cleveland and can be reached at tim.dunne@clev.frb.org. Georgia Kosmopoulou (corresponding author) is at the University of Oklahoma and can be reached by mail at University of Oklahoma, Department of Economics, 203 Hester Hall, 729 Elm Avenue, Norman, OK 73019; by phone at 1-405-325-3083; or e-mail at georgiak@ou.edu. Carlos Lamarche is at the University of Oklahoma and can be reached at lamarche@ou.edu. The authors would like to thank Andrés Aradillas-López, Han Hong, Joakim G. Laguros, Lawrence J.White, and conference participants at the 2009 International Industrial Organization Conference and the Stanford Institute for Theoretical Economics summer 2009 workshop. We are indebted to staff at the Texas Department of Transportation for providing useful information.

1 Introduction

Minority preference policies have been a part of government procurement programs in the United States since the late 1960's. Their goal is to enhance the opportunities of minority businesses and counter the effects of past discrimination. Critics of these policies claim that they result in reverse discrimination, limit competition, and raise project costs.[1] Two incentive schemes have been used widely thus far – rules requiring participation of minority firms as subcontractors and bid preference programs.[2] Our analysis focuses on the former, policies that set minority firm participation goals. Specifically, these rules require that prime contractors subcontract out a set percentage of the overall value of a project to minority firms. Such a requirement could affect the prime contractor's make-or-buy decisions in two ways. First, it may influence the overall level of subcontracting a firm uses on a project and, second, it may influence who the firm subcontracts with on a project. Both instances impose constraints on the prime contractors, potentially raising projects costs.

This paper examines whether project costs differ between auctions that have subcontracting goals and auctions without such goals. The paper employs a structural auction model to infer contractors' costs from observed bids in order to compare the costs across project types. Nonparametric methods developed by Guerre, Perrigne, and Vuong (2000) and Haile, Hong and Shum (2006) are used to estimate the distribution of latent costs, allowing us to control for project heterogeneity and selection. Papers by Marion (2007) and Krasnokutskaya and Seim (2007) also use a structural auction approach to examine bid preference schemes for small businesses; however, these papers do not examine environments where subcontracting goals are implemented. To be sure, our empirical analysis is not an evaluation of the program itself, as the program has been in place for several decades. Rather the structural model

[1] See Holzer and Neumark (2000) for an overview of affirmative action programs and how they affect small and disadvantaged businesses.

[2] Bid preference programs give explicit advantage to small and minority bidders in auctions. For example, in the case of California state highway contracts with bid preferences analyzed by Krasnokutskaya and Seim (2007), small business are awarded a contract if they are within 5% of the low bid.

is employed to quantify differences in costs across projects with and without subcontracting goals.

The empirical setting is first-price sealed bid auctions for highway construction projects let in Texas over the period 1998-2007.[3] Our findings show that, once project heterogeneity and bidder participation are controlled for, there is little difference in costs between projects that are assigned subcontracting goals and projects that are not assigned such goals. When we examine an even more homogeneous sample of projects, we find even greater similarity in costs between the two project groups. We also construct estimates of the markup of the bid above the cost and find that the magnitude of the markup is consistent with that reported in the literature and varies little between auctions with and without subcontracting requirements.

The effect of minority preference policies on bidding and costs have been examined in recent studies. Several papers deal with bid preference schemes. Denes (1997) compares bids submitted between solicitations restricted to small businesses and unrestricted solicitations. He finds that bids are no higher in restricted solicitations.[4] Krasnokutskaya and Seim (2007) analyze bid preference programs in California highway procurement by examining how bidding and participation decisions are affected by a program that provides preferential treatment to small firms. They find that the preferential treatment of small businesses creates losses in efficiency (since the small firms have higher costs on average) but no change in the overall cost of procurement. In a related study of the California state procurement auctions, Marion (2007) found that the distortion in participation patterns in bid preference programs is responsible for a 3.8 percent increase in the cost of the winning firm. Despite this evidence, the effect of such programs on the state's cost is ambiguous. By invoking bid preferences the state gives an advantage to minority bidders and compels the non-minority bidders to bid more aggressively

[3] Our structural analysis only includes asphalt paving projects, as we focus on a relatively homogeneous set of projects and include those that best match the assumptions of the independent private value environment. Papers such as Bajari and Ye (2003) also focus on subsets of construction projects in order to achieve greater homogeneity in the items under study.

[4] Other studies that have been done focus on whether companies that benefit from affirmative action in procurement continue to succeed after the programs are no longer in effect (Holzer and Neumark, 2000).

and win contracts at a lower bid. At the same time, since the competitive pressure is reduced for minority bidders they bid less aggressively than otherwise; and when the item is awarded to them, they impose additional cost on the state (McAfee and McMillan, (1989) and Maskin and Riley, (2000)).

The potential for efficiency distortions is different for programs setting minority subcontracting goals. These programs are widely used in federal procurement contracts and may constrain the make-or-buy decision of prime contractors. Efficiency distortions could be introduced due to potentially less efficient production of tasks by subcontractors compared to the prime contractor, to the use of less efficient subcontractors on subcontracted tasks, or to changes in competition intensity in the subcontracting market. Marion (2009) using data from the California Department of Transportation spanning the period between 1996 and 1999 shows that the subcontracting goals set for highway construction contracts in California raise disadvantaged business enterprise usage significantly, so that the constraints appear to bind.

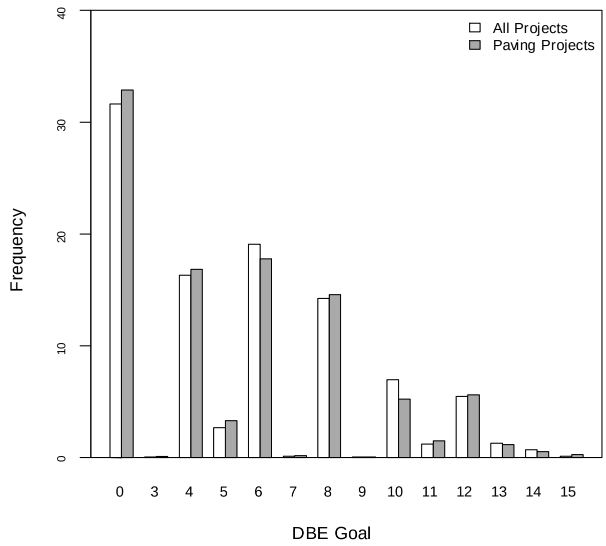
The paper proceeds as follows. The next section describes the disadvantaged business enterprise program and provides an overview of the data. Section 3 presents the model and structural empirical analysis. Section 4 concludes.

2 Texas Auctions and Bidding Patterns

2.1 Data Description

Our analysis utilizes data on auctions and bidding from the state of Texas. The Texas Department of Transportation (TxDOT) holds regularly scheduled highway procurement auctions that incorporate goals for the awarding of subcontracts to disadvantaged business enterprises (DBEs). DBEs are small businesses that are owned and controlled by members of a minority group including women-owned businesses. For selected federally-funded projects, TxDOT assigns a proportion of the contract value that must be performed by DBEs. Figure 2.1 presents the distribution of the DBE goals for two groups of federally-funded projects let between 1998

Figure 2.1: Distribution of DBE goals for All Projects and Paving Projects.

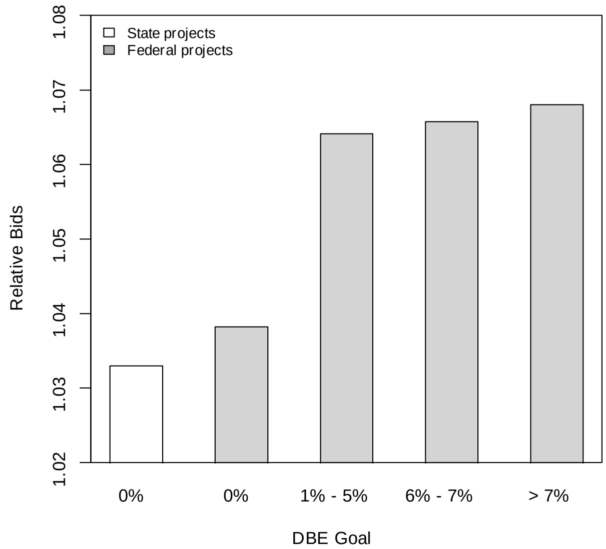


and 2007 - all projects and paving projects.[5] Since paving projects are the focus of the empirical analysis that follows, we provide a separate breakout for this group of projects. Across all projects, the DBE goals range from zero to 15 percent with about two-thirds of projects having DBE goals above zero. Paving projects make up about one half of the overall number of projects.

As in other states, the Department of Transportation in Texas chooses which projects to assign DBE status and the level of the DBE goal for each project. The state makes its decisions by considering a number of factors including – the type of project (asphalt, bridgework, etc.), the geographic location of the project, and the availability of pre-qualified DBE subcontractors in locations that can do specific tasks. TxDOT has a separate office that manages these assignments, which is distinct from the offices that design, cost out, and let

[5] The sample is restricted to federally-funded projects that are estimated to cost in excess of $400,000, as TxDOT only considers projects that are estimated to cost at least $400,000 for assignment of DBE goals. State funded projects do not have DBE goals.

Figure 2.2: Average Relative Bid for Paving Projects by DBE intensity



the projects.

The TxDOT bid data that we have access to contain information on all road construction projects offered for bid letting in Texas for the period from September, 1998 through August, 2007.[6] Our empirical analysis focuses on paving projects, a relatively homogeneous group of projects that previous studies have shown fit the independent private values framework well. Projects are auctioned off on a monthly basis using a first-price, sealed-bid format. For each project, we have the date of the bid letting, information on the location of the project, an overall description of the project, a detailed list of the tasks involved, the estimated length of the project (in calendar days), the state's engineering estimate of the project's total cost, whether the project is federally or state funded, and the DBE participation requirement. State

[6] The structure of the DBE program changed during our sample period, as the U.S. Federal Highway Administration (FHWA) moved toward more race-neutral approaches to meet DBE subcontracting objectives. In an earlier version of the paper, we tested for differences in bidding behavior associated with changes in how the DBE program was administered. We found no evidence of a change in bidding in DBE vs. non-DBE auctions in response to such program changes.

Variable	Without DBE Goals	With DBE Goals
Number of projects	1839	1220
Number of state projects	1241	-
Average number of bidders	3.805	3.892
	(1.774)	(1.922)
Average engineer's cost estimate	2.969	4.240
(in millions of dollars)	(2.824)	(3.913)
Average relative bid	1.035	1.066
	(0.192)	(0.176)
Average number of bid components	40.506	81.560
	(30.181)	(49.509)

Table 2.1: Summary statistics. DBE denotes disadvantaged business enterprises. Standard deviations are in parentheses.

projects and federally funded projects of less than \$400,000 do not have DBE goals. From the bid letting, we know the identity of the plan holders - the firms that purchase the plans for a project, which plan holders submit bids, and the dollar value of each bid submitted. Our paving contract data contain 11,745 bids from 3,059 auctions.[7]

Figure 2.2 shows bid statistics for state and federal paving projects of more than \$400,000, breaking out the federal projects into four DBE utilization categories: 0%, 4-5%, 6-7%, or greater than 7%. The average relative bids – the bid submitted by a firm divided by TxDOT's engineering cost estimate (ECE) for a project – is higher in projects with DBE goals. However, there are marked differences in the characteristics of state, DBE and non-DBE projects as shown in Table 2.1. Projects without DBE goals are generally smaller than DBE projects and contain a smaller number of bid components.

2.2 Bidding Regression Results

To further explore the patterns related to DBE status, we present a set of descriptive regressions where the dependent variable is the log of the bid submitted by an individual bidder. All models include a common set of basic project characteristics, including controls for the

[7] We also restrict our sample to projects under \$20 million dollars. This excludes 69 projects and we make these restrictions to be consistent with the sample employed in the structural model. The results in Table 3.3 are similar with or without this restriction.

DBE status of the auction, project size (measured as the log of the engineering estimate), project location and time effects.[8] In some specifications, we also control for project characteristics including – project length (log of the calendar days to complete a project), project type (shares of specific material components), and the number of project components. The last variable has been used in a number of studies to proxy for the complexity of the project. To control for bidder cost heterogeneity, variables that measure a bidder's backlog of projects and the distance to a project are included. A list of variable definitions is presented in Table A.1 in the appendix. Two alternative DBE specifications are presented for the bid regressions – one version includes a simple indicator variable that identifies auctions with positive DBE goals, while the second version includes a set of dummy variables that control for DBE goal percentage.

The first column of Table 2.2 provides the estimates from a model that includes an indicator variable for whether a federal project has a DBE goal or not and controls for engineering cost estimate, whether the project is a state or federal project, location and time effects. The coefficient on the DBE variable in this parsimonious specification is positive and statistically significant indicating that DBE auctions had higher average bids than those observed in non-DBE auctions (including both federal and state projects). There is no difference in the average bid between state projects and federal projects that do not have DBE goals. The second column adds one additional control variable – the log of the number of bid components – the complexity variable. Adding this variable to the regression moves the coefficient on the DBE variable close to zero and it is no longer statistically significant, while the complexity variable enters the regression with a positive and statistically significant coefficient. This indicates that projects with more tasks generally have higher bids, even after controlling for project size. Column 3 incorporates additional controls for project and bidder characteristics. Project length, bidder backlog and bidder distance to project location all increase bids, while

[8]Project location is modeled using 24 indicator variables that identify which construction district a project is in. Time effects are controlled for by a set on 119 indicators variables that identify the month and year of the project letting.

Variable	**Log of Bids**				
	(1)	(2)	(3)	(4)	(5)
DBE projects	0.051*	0.003	0.001		
	(0.008)	(0.008)	(0.008)		
State projects	-0.010	-0.002	0.004		
	(0.008)	(0.008)	(0.008)		
DBE: 0% (Fed projects)				-0.003	0.003
				(0.008)	(0.036)
DBE : 1% - 5%				0.001	0.008
				(0.009)	(0.038)
DBE : 6% - 7%				-0.004	0.003
				(0.010)	(0.039)
DBE : $> 7\%$				-0.009	-0.001
				(0.010)	(0.039)
Log ECE	0.973*	0.954*	0.943*	0.944*	0.944*
	(0.004)	(0.003)	(0.005)	(0.005)	(0.005)
Log complexity		0.073*	0.054*	0.055*	0.054*
		(0.005)	(0.008)	(0.008)	(0.009)
Log complexity $\times$ state projects					0.002
					(0.010)
Log days to complete the project			0.022*	0.022*	0.022*
			(0.006)	(0.006)	(0.006)
Log backlog			0.001*	0.001	0.001
			(0.000)	(0.000)	(0.000)
Log distance			0.010*	0.010*	0.010*
			(0.002)	(0.002)	(0.002)
Division effects (24)	Yes	Yes	Yes	Yes	Yes
Time effects (119)	Yes	Yes	Yes	Yes	Yes
Material shares (11)	No	No	Yes	Yes	Yes
Number of observations	11745	11745	11745	11745	11745

Table 2.2: Descriptive bid regressions. While DBE here is an indicator variable denoting disadvantaged business enterprises, ECE denotes the value of the engineer's cost estimate. The symbol * denotes statistical significance at the 5 percent level. Standard errors (in parentheses) are clustered at the auction level.

the coefficient on the DBE variable is close to zero.

Column 4 includes all the regressors in Column 3 but replaces the zero-one indicator variable for DBE with a set of dummy variables that capture differences in the level of DBE goals across projects. The model includes four dummy variables for the DBE groupings reported in Figure 2.2 and the omitted category represents the state projects. None of the DBE coefficients is statistically significant and all are close to zero in magnitude. Moreover, there is no rise in the coefficients as the DBE goal increases, as one might expect if higher DBE goals were forcing prime contractors to subcontract more activity to less efficient producers of a task.[9]

How should we interpret the results on complexity and DBE status of projects? TXDOT assigns DBE goals based, in part, on the tasks involved in a project. Projects with a large number of tasks are more likely to have tasks appropriate for DBE subcontracting. Thus, if the state always assigned DBE status to complex projects, then our regressions would not be able to distinguish the differences in bidding due to complexity from differences in bidding due to DBE requirements. One way to see if the positive effect of complexity is merely proxying for DBE status is to examine complexity's correlation with bids in state projects. State projects are not assigned DBE goals but we can still measure the number of tasks for these projects. Column (5) of Table 2.2 presents the results of a model that includes an interaction term between a state project indicator variable and the log of complexity. This interaction term tests whether there is any difference in the relationship between bids and complexity for state versus federal projects. The coefficient on the interaction term is essentially zero (0.002) and not statistically significant. This shows that the relationship between complexity and bids is similar in state and federal projects, suggesting that the correlation between bids and complexity is not being driven by the DBE assignment process. Rather the correlation likely reflects the increase in costs associated with doing more complicated projects.

[9] We also estimated a set of models where the dependent variables were the number of bidders and the winning bids. There was no difference in participation patterns or winning bids between auctions with and without DBE goals.

One potential problem with the above analysis is that it requires that the engineer's cost estimate not be influenced by DBE assignment. In Texas, the office that assigns DBE goals to auctions is clearly distinct from the parties responsible for designing and costing out a project. In addition, the setting of a project's DBE goal occurs only after a project's cost is estimated. So, DBE assignment of a project does not influence the engineer's cost estimate for the project. However, in estimating project costs, TXDOT can use information from prior bid submissions to estimate the costs of specific project components.[10] A problem with our analysis would arise if some project tasks are only performed by DBE firms and only occur in DBE auctions. Effectively there would be no cost differential to estimate in this circumstance, as differences in costs due to DBE subcontracting would simply be reflected in the engineer's cost estimate in DBE projects. Moreover, we do know that the office that makes DBE assignments does consider the task list in making DBE assignments. The tasks considered most suited to DBEs in Texas are listed in Table A.2 in the appendix. Table A.2 shows the frequency of each task broken out by federal auctions with DBE goals, federal auctions without DBE goals, and state projects. The specific bid items presented center around landscape, traffic control and miscellaneous construction activities.[11] Importantly, the frequency data show that these DBE tasks are not limited to DBE auctions. State projects that are not assigned DBE goals, also incorporate these tasks, and so do federal non-DBE projects. In general, we see that roughly 57% of these DBE tasks are in DBE projects while 43% are in non-DBE projects. These project percentages are similar to the overall percentage of bid items across DBE and non-DBE projects (60% vs. 40%). Thus, there does not appear to be a specialized group of tasks that only occur in DBE projects.

[10] For example, many projects require the use landscaping services. TXDOT may use past seed, planting and fertilizer prices to generate an estimate of landscape costs for a project to be let.

[11] Most projects also incorporate a Mobilization bid component. Mobilization tasks "include establishing and removal of offices, plants, facilities and moving personnel, equipment, and supplies to (or from) the project area to begin (or complete) work." Mobilization is given as a lump sum figure for a project and averages around 8 percent of a project's total estimated cost. The interquartile range is from 7% to 9%. DBE subcontractors can also be paid to perform these tasks and are; however, since mobilization tasks do not have specific price components (they are really an overhead type cost), they are not subject to the potential bias discussed here.

Overall, the bidding patterns suggest little difference in bids submitted in auctions with DBE goals compared to auctions without DBE goals. However, these descriptive regressions do not control for the competitive environment or for such features as selection into the auctions. In the next section, we employ a structural approach that will allow us to control for competition, entry into the auctions and generate estimates of the latent cost distributions for bidders participating both in DBE and non-DBE auctions.

3 Structural Analysis

This section uses nonparametric estimation methods to uncover the cost of firms bidding in procurement auctions. Before proceeding to the empirical analysis, we outline a simple bidding model.

3.1 Model

There are n risk neutral bidders who compete for a government contract in a first price sealed bid auction where the low bidder is awarded the contract. There are two types of projects, indexed by j, those that have no subcontracting goals and those that do (i.e., $j = \{0, 1\}$). The cost of contract j to a bidder i, is private and denoted by c_{ij}. The density of the private cost c_{ij} is f_j and is strictly positive on the support $[c_{L_j}, c_{H_j}]$. In a procurement auction, a bidder who is awarded contract j at a bid of b_{ij} receives a net profit of $b_{ij} - c_{ij}$. Each bidder is maximizing expected profit given by:

$$E[\pi_{ij}(b_{1j}, b_{2j}, ..., b_{nj}, c_{ij})] = (b_{ij} - c_{ij})\left(1 - F_j\left(\varphi(b_{ij})\right)\right)^{n-1}.$$

In the symmetric independent private value (IPV) case, the equilibrium bid function is

$$\beta(c_{ij}|F_j, n) = c_{ij} + \frac{\beta'(c_{ij})[1 - F_j\left(c_{ij}\right)]}{(n-1)f_j\left(c_{ij}\right)} \tag{3.1}$$

where $b_{ij} = \beta(c_{ij})$ and $\varphi(b_{ij}) = c_{ij}$.

Notice that the cost of the contract consists of the sum of the cost of various tasks comprising the project, some or potentially all of which may be undertaken by the primary contractor.

In projects having subcontractor participation goals, a number of tasks representing a minimum percentage of the estimated cost, have to be undertaken by DBE subcontractors. We ask if there is a difference in bidding distributions between projects that have subcontracting goals in place and those that do not and whether the combined cost of the project is different across $j's$. It is obvious that, if the minority subcontractors are less efficient they will impose a cost to the state agency.

Within the symmetric independent private value framework, we use the equilibrium bidding function (3.1) to obtain the cost distribution of firms undertaking projects either with subcontracting goals or without subcontracting goals. Let $G_0(b)$ be the distribution function of bids in projects without subcontracting goals and $G_1(b)$ the distribution function of bids in projects with subcontracting goals. Let $g_0(b)$ and $g_1(b)$ be the associated densities. Considering the standard monotonicity condition imposed on the equilibrium bid function $\beta(c)$, we write $F(c) = F(\beta^{-1}(b)) = G(b)$, and $f(c) = g(b)\ \beta'(c)$. If we substitute these expressions into the equilibrium bidding function, we find that the latent cost of undertaking a project without subcontracting goals can be written as,

$$c_0 = b_0 - \frac{1}{n_0 - 1} \frac{1 - G_0(b_0)}{g_0(b_0)}, \tag{3.2}$$

where n_0 is the number of firms bidding in projects without subcontracting goals. Similarly, the latent cost associated with a project that has subcontracting goals is,

$$c_1 = b_1 - \frac{1}{n_1 - 1} \frac{1 - G_1(b_1)}{g_1(b_1)}, \tag{3.3}$$

where n_1 is the number of firms bidding in projects with subcontracting goals. The right hand side of these equations can be estimated with nonparametric methods using the observed vector of bids.

3.2 Asphalt project data

The identification and estimation of equations (3.2) and (3.3) rely on the assumptions associated to the IPV framework, which are tested in Section 3.4. We require a sample of projects that are relatively homogeneous and fit the IPV framework. From related literature (see Bajari and Ye (2003), De Silva, Dunne, Kankanamge, and Kosmopoulou (2008)) and our discussions with state highway and civil engineers, we believe that asphalt projects appear to best match these requirements. Asphalt projects rely more on the individual firm's state of equipment and internal efficiency to determine the cost and are relatively homogeneous.

Although asphalt projects are less heterogeneous than the full sample of paving projects used in Section 2, they may include work on non-asphalt components such as bridge, subgrade, etc. We made two adjustments to the sample to obtain an even more homogeneous set of projects. First, we restrict attention to asphalt paving projects with an estimated cost between 1 million and 20 millions, with an asphalt material share higher than 50 percent of the engineer's cost estimate, and with bridge and earthwork components of less than 5 percent. Bridge and earthwork components introduce uncertainty in the cost that is likely more common to all bidders. We also restricted the sample to projects with no subgrade and base course tasks.[12] Those tasks introduce common uncertainty in costs and appear most often in the construction of new roads. We present descriptive statistics for this sample in the first four columns of Table 3.1 (we call this sample Asphalt Projects). We consider the sample for all levels of participation in the first two columns and a subsample with 3 and 4 bidders in the next two columns. In the empirical analysis that follows, we will focus on samples with similar number of bidders. Second, we construct a more selected sample of contracts that relates exclusively to surface treatment.[13] The descriptive statistics for this sample are presented in the last four columns of Table 3.1. Notice that the size and number of tasks are much more similar across DBE and non-DBE projects in these subsamples compared to the

[12] Subgrade tasks are associated with the top surface of a roadbed upon which the pavement structure, shoulders, and curbs are constructed. Base course tasks are associated with the layers of specified material placed on a subgrade to support a surface course.

[13] Bajari and Ye (2003) analyze a model of independent private cost estimates using seal coating projects, which is class of surface treatment projects. Surface treatment may be used for primary and secondary roads that carry light traffic, as part of the original construction, or to rejuvenate old roads. Examples include overlay of asphalt, seal coats, single and multiple surface treatments. Surface treatments could be applied to concrete roads or bituminous asphalt roads.

	Asphalt projects				Surface treatment projects			
	All Bidders		3 and 4 Bidders		All Bidders		3 and 4 Bidders	
	Non DBE	DBE	Non DBE	DBE	Non DBE	DBE	Non DBE	DBE
Bid (millions	3.202	3.159	3.121	3.298	2.820	2.907	2.840	3.631
of dollars)	(2.111)	(2.009)	(2.010)	(1.874)	(2.447)	(2.590)	(2.226)	(3.275)
Engineer's cost	3.203	3.169	3.075	3.332	2.822	2.875	2.772	3.620
estimate	(1.920)	(1.965)	(1.793)	(1.902)	(2.423)	(2.457)	(2.030)	(3.207)
Bridge work	0.006	0.007	0.005	0.009	0.004	0.005	0.005	0.009
	(0.011)	(0.013)	(0.010)	(0.013)	(0.009)	(0.011)	(0.010)	(0.014)
Earth work	0.011	0.011	0.011	0.012	0.012	0.008	0.012	0.011
	(0.013)	(0.014)	(0.013)	(0.015)	(0.015)	(0.013)	(0.014)	(0.016)
Pavement	0.058	0.036	0.049	0.025	0.006	0.014	0.007	0.025
	(0.179)	(0.131)	(0.166)	(0.095)	(0.021)	(0.061)	(0.023)	(0.092)
Concrete	0.006	0.006	0.005	0.007	0.004	0.004	0.005	0.007
	(0.011)	(0.011)	(0.010)	(0.012)	(0.009)	(0.011)	(0.010)	(0.013)
Subcontracting	-	6.584	-	6.286	-	6.543	-	6.578
goals		(2.358)		(2.373)		(2.436)		(2.819)
Complexity	32.451	37.766	32.377	42.327	30.723	33.174	31.350	37.513
of the project	(14.788)	(18.504)	(15.498)	(19.603)	(14.797)	(17.274)	(15.530)	(18.674)
Days to complete	84.820	89.026	81.013	88.822	70.981	71.659	67.040	77.766
the project	(61.082)	(49.377)	(61.029)	(42.346)	(62.592)	(34.544)	(49.821)	(35.568)
Number of	4.230	4.644	3.413	3.443	4.207	4.851	3.469	3.416
bidders	(1.821)	(2.099)	(0.493)	(0.498)	(1.757)	(2.135)	(0.500)	(0.494)
Number of:								
Auctions	206	175	112	76	134	126	68	50
Observations	751	664	368	248	473	475	226	126

Table 3.1: Summary statistics for the samples of asphalt projects and surface treatment projects. The data set also includes five indicator variables for the geographic location of the projects. DBE stands for disadvantage business enterprises. Standard deviations are in parentheses.

differences observed in the overall paving sample used in Section 2.

3.3 Nonparametric estimation and auction heterogeneity

Standard non-parametric methods can be used to estimate $(1 - G(\mathbf{b}|\mathbf{x}))/g(\mathbf{b}|\mathbf{x})$, where the vector $\mathbf{x} \in \boldsymbol{\mathcal{X}} \subset \mathbf{R}^p$ includes variables capturing observed project heterogeneity (e.g., Guerre, Perrigne, and Vuong 2000). We incorporate auction specific characteristics replacing the unconditional distribution functions $G_j(b)$ and $g_j(b)$ in equations (3.2) and (3.3) by conditional distributions of a form $G_j(\mathbf{b}|\mathbf{x})$ and $g_j(\mathbf{b}|\mathbf{x})$, where $\mathbf{x}$ includes the engineer's cost estimate as in Marion (2007). These conditional functions can be estimated by considering the empirical version of standard definitions, $\hat{g}_j(\mathbf{b}_j|\mathbf{x}_j) = \hat{g}_j(\mathbf{b}_j, \mathbf{x}_j)/\hat{f}_j(\mathbf{x}_j)$ and $\hat{G}_j(\mathbf{b}_j|\mathbf{x}_j) = \hat{G}_j(\mathbf{b}_j, \mathbf{x}_j)/\hat{f}_j(\mathbf{x}_j)$, and the following estimators defined in Guerre, Perrigne, and Vuong (2000):

$$
\begin{aligned}
\hat{g}_j(\mathbf{b}_j, \mathbf{x}_j) &= \frac{1}{nL_j h_{jg}^2} \sum_{l=1}^{L_j} \sum_{i=1}^{n} K_g\left(\frac{b - b_{jil}}{h_{jg}}, \frac{\mathbf{x} - \mathbf{x}_{jl}}{h_{jg}}\right), \\
\hat{G}_j(\mathbf{b}_j, \mathbf{x}_j) &= \frac{1}{nL_j h_{jG}} \sum_{l=1}^{L_j} \sum_{i=1}^{n} K_G\left(\frac{\mathbf{x} - \mathbf{x}_{jl}}{h_{jG}}\right) 1\left\{b_{jil} \leq b\right\}, \\
\hat{f}_j(\mathbf{x}_j) &= \frac{1}{L_j h_{jf}} \sum_{l=1}^{L_j} K_f\left(\frac{\mathbf{x} - \mathbf{x}_{jl}}{h_{jf}}\right),
\end{aligned}
$$

where $1\{\cdot\}$ is an indicator function, $K_g(\cdot)$, $K_G(\cdot)$, and $K_f(\cdot)$ are continuously differentiable kernel functions defined over a compact support, and h_g, h_G, and h_f are the associated bandwidths. Several kernels satisfy these conditions, including the triweight kernel,

$$
K(u) = \frac{35}{32}\left(1 - u^2\right)^3 1\left\{|u| \leq 1\right\}.
$$

We use this triweight kernel to estimate the density $f_j(\mathbf{x}_j)$ and the distribution function $G_j(\mathbf{b}_j, \mathbf{x}_j)$. Moreover, we consider the product of two triweight kernels for estimating the density $g_j(\mathbf{b}_j, \mathbf{x}_j)$. Both the rates in Guerre, Perrigne and Vuong (2000) and the factors associated with the choice of the triweight kernel (see, e.g, Härdle 1991) suggest employing bandwidths of the form $h_{jG} = c\hat{\sigma}(b_j)(nL_j)^{-1/5}$, $h_{jg} = c\hat{\sigma}(b_j)(nL_j)^{-1/6}$, and $h_{jf} = c\hat{\sigma}(\mathbf{x}_j)(nL_j)^{-1/5}$, where $\sigma(b)$ is defined as the standard deviation of b and $c = 2.978 \times 1.06$.

Given the potential benefits of using the logarithm of bids rather than bids, we consider the logarithmic transformation for the variable of interest $\mathbf{c}_j$ (see Li, Perrigne and Vuong (2000), and Marion (2007)). We define the pseudo cost $\hat{\mathbf{c}}$ as follows:

$$\hat{\mathbf{c}}_j = \begin{cases} \exp(\mathbf{a}_j)(1 - m_j(\mathbf{a}_j, \mathbf{z}_j)) & \text{if } \max\{h_{jG}, h_{jg}\} \leq a_{jil} \leq a_{j\max} - \max\{h_{jG}, h_{jg}\} \\ +\infty & \text{otherwise,} \end{cases} \tag{3.4}$$

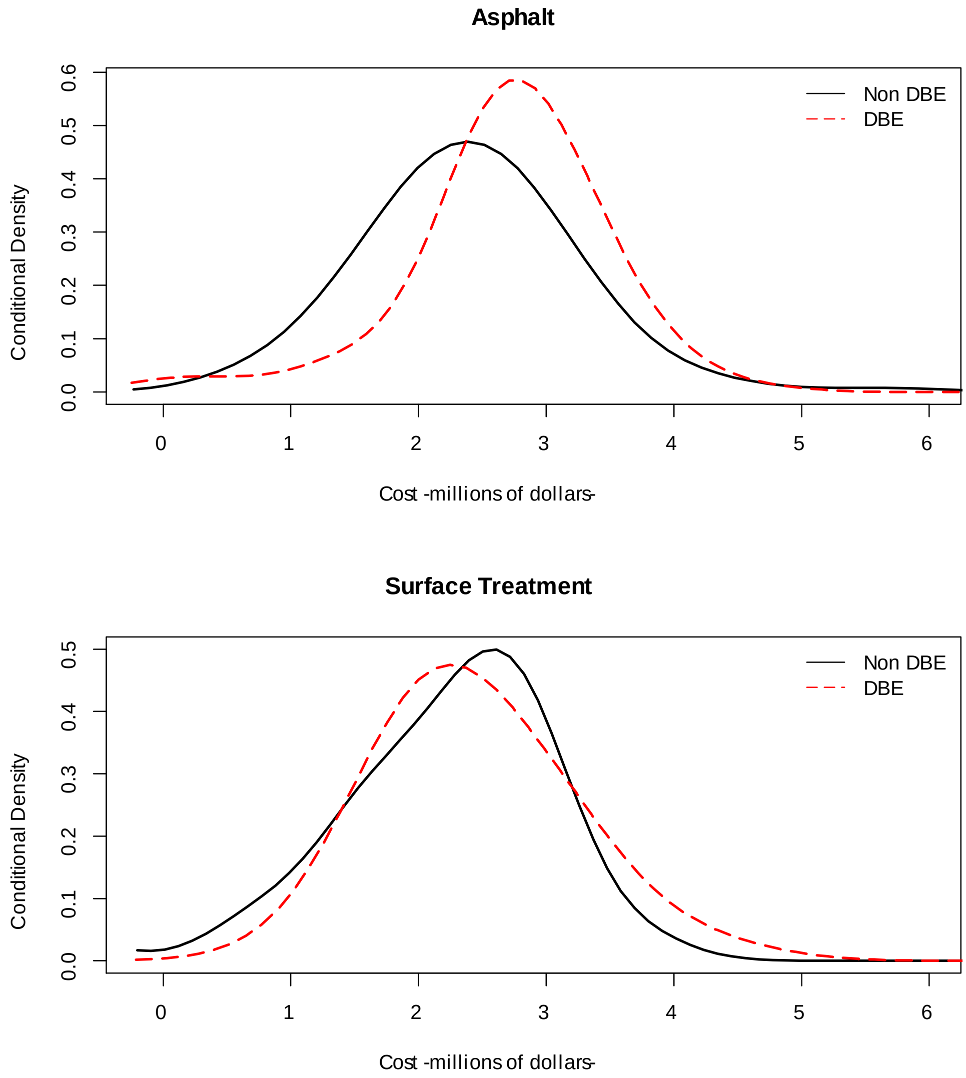
where the variables $\mathbf{a}_j = \log(\mathbf{b}_j)$, $\mathbf{z}_j = \log(\mathbf{x}_j)$, and

$$m_j(\mathbf{a}_j, \mathbf{z}_j) = \frac{1}{n-1} \frac{1 - \hat{G}_j(\mathbf{a}_j|\mathbf{z}_j)}{\hat{g}_j(\mathbf{a}_j|\mathbf{z}_j)}.$$

The upper bound of the support includes a variable $a_{j\max}$ defined as $\max\{a_{j1}, ...a_{jnL_j}\}$. In the first stage, we now use equation (3.4) to obtain $\hat{\mathbf{c}}_0$ and $\hat{\mathbf{c}}_1$, and in the second stage, we use these pseudo costs and the engineer's cost estimate to estimate the conditional distributions $\hat{g}_0(\hat{\mathbf{c}}_0|\mathbf{x}_0)$ and $\hat{g}_1(\hat{\mathbf{c}}_1|\mathbf{x}_1)$.

Figure 3.1 presents the conditional densities evaluated at the median of the engineer's cost estimate. These empirical distributions were obtained considering the samples described in Table 3.1. The continuous lines show kernel density estimates for the cost of firms bidding in projects without subcontracting goals (non-DBE), and the dashed line present estimates for the cost of firms bidding in projects with subcontracting goals (DBE). Because the bid's distributions are not comparable in cases of different number of bidders, we estimate the vector of pseudo cost $\hat{\mathbf{c}}_j$ separately for 3 and 4 bidders. Then, we pool the values for different number of bidders to estimate the conditional density of cost $\hat{g}_j(\hat{\mathbf{c}}_j|\mathbf{x}_j)$. The upper panel in Figure 3.1 shows that the cost distributions of firms bidding in asphalt projects when small bridge and earthwork components are present in the project, and the lower panel presents results from the sample of surface treatment projects. The distribution of firms undertaking DBE projects is shifted to the right, suggesting the possibility that the program generated inefficiencies. However, when we consider the more homogeneous sample of asphalt surface treatment projects, the differences in the cost distributions tend to disappear.

Figure 3.1: Cost densities for DBE and non-DBE projects. The densities were obtained considering a non-parametric method that uses trimming and the engineer's cost estimate as a conditioning variable. While the chart at the top is obtained using the sample of Asphalt projects, the chart at the bottom is obtained considering the sample of Surface Treatment projects.



In our application, one needs to control for many auction-specific characteristics. Recall that the effects of project size, project complexity, and project length are statistically significant in all variants of the model estimated in Table 2.2. It is natural then to use the estimation method proposed by Haile, Hong and Shum (2006). The advantage of their approach relative to the approach developed by Guerre, Perrigne, and Vuong (2000) is that it enables one to control for many auction-specific characteristics without increasing the sample size. The basic idea is to impose an additively separable structure on how observable factors $\mathbf{x}$ and latent auction heterogeneity w affect costs.

Consider the function Γ: $\boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{W}} \rightarrow \mathbb{R}$, and $\exists(\mathbf{x_0}, w_0) \in \boldsymbol{\mathcal{X}} \times \boldsymbol{\mathcal{W}} \subset \mathbf{R}^p \times \mathbf{R}$ such that $E(\Gamma(\mathbf{x},w)) = \Gamma(\mathbf{x_0}, w_0)$. Under assumptions of separability, the equilibrium bid function can be written as,

$$\begin{aligned} \beta(c|n, \mathbf{x},w) &= \beta(c|n, \mathbf{x_0}, w_0) + \Gamma(\mathbf{x},w) \\ &= \alpha(n) + \Gamma(\mathbf{x_0},w_0) + \tilde{\Gamma}(\mathbf{x},w) + \tilde{\beta}(c|n, \mathbf{x_0},w_0) \end{aligned}$$

where $\tilde{\Gamma}(\mathbf{x},w) = \Gamma(\mathbf{x},w) - \Gamma(\mathbf{x_0},w_0)$, $\alpha(n) = E(\beta(c|n, \mathbf{x_0},w_0))$, and $\tilde{\beta}(c|n, \mathbf{x_0},w_0)$ is a conditional zero mean term. Because in equilibrium we have that $b = \beta(\cdot)$, $b^0 \equiv \alpha(n) + \Gamma(\mathbf{x_0},w_0) + \tilde{\beta}(c|n, \mathbf{x_0},w_0) = b - \tilde{\Gamma}(\mathbf{x},w)$ is interpreted as the bid a firm would have submitted in equilibrium to an auction with observables characteristics $\Gamma(\mathbf{x},w) = \mathbf{\Gamma}(\mathbf{x_0},w_0)$. Notice that we need to control directly for the effect of w. Assuming that $\zeta(\mathbf{x}, \mathbf{z}) = \min\{n{\in}N\colon \Pr(N{\leq}n|\mathbf{x}, \mathbf{z}) \geq\tau\}$ for a quantile $\tau{\in}(0, 1)$, we write,

$$n = \zeta(\mathbf{x}, \mathbf{z}) + w, \tag{3.5}$$

where $\mathbf{z}$ is a vector of instruments and w is an index that includes unobserved factors independent of $\mathbf{x}$. In this paper, we take a control variate approach, estimating $w = n - \zeta(x, z)$ as suggested in Haile, Hong and Shum (2006).

We use a non-parametric approach to estimate $(1 - G(\hat{b}^0))/g(\hat{b}^0)$, where $\hat{b}^0 = b - \hat{\Gamma}(\mathbf{x}, \hat{w})$. We obtain $\hat{w}$ after estimating (3.5) by censored quantile regression. We use the number of plan holders as an instrument. The vector $\mathbf{x}$ includes controls for project's size (engineer's cost estimate and a quadratic term on the engineer's estimate), variables associated with the complexity of the project (number of project's component and a quadratic term on the number of project's component), controls for length of the project (calendar days to finish the project, and an interaction term between calendar days and engineer's cost estimate), controls for project's type (percentage of earthwork, percentage of bridge work, percentage of asphalt pavement work, and percentage of concrete work), and four variables indicating the location of the projects.[14] To control for asymmetries among bidders, it is standard to include the distance to the project location and the capacity commitment of the firm (backlog) in the vector $\mathbf{x}$ (Jofre-Bonet and Pesendrofer 2003, Bajari and Ye 2003). Finally, we estimate (3.2) and (3.3) using the homogenized bids, and the following estimators,

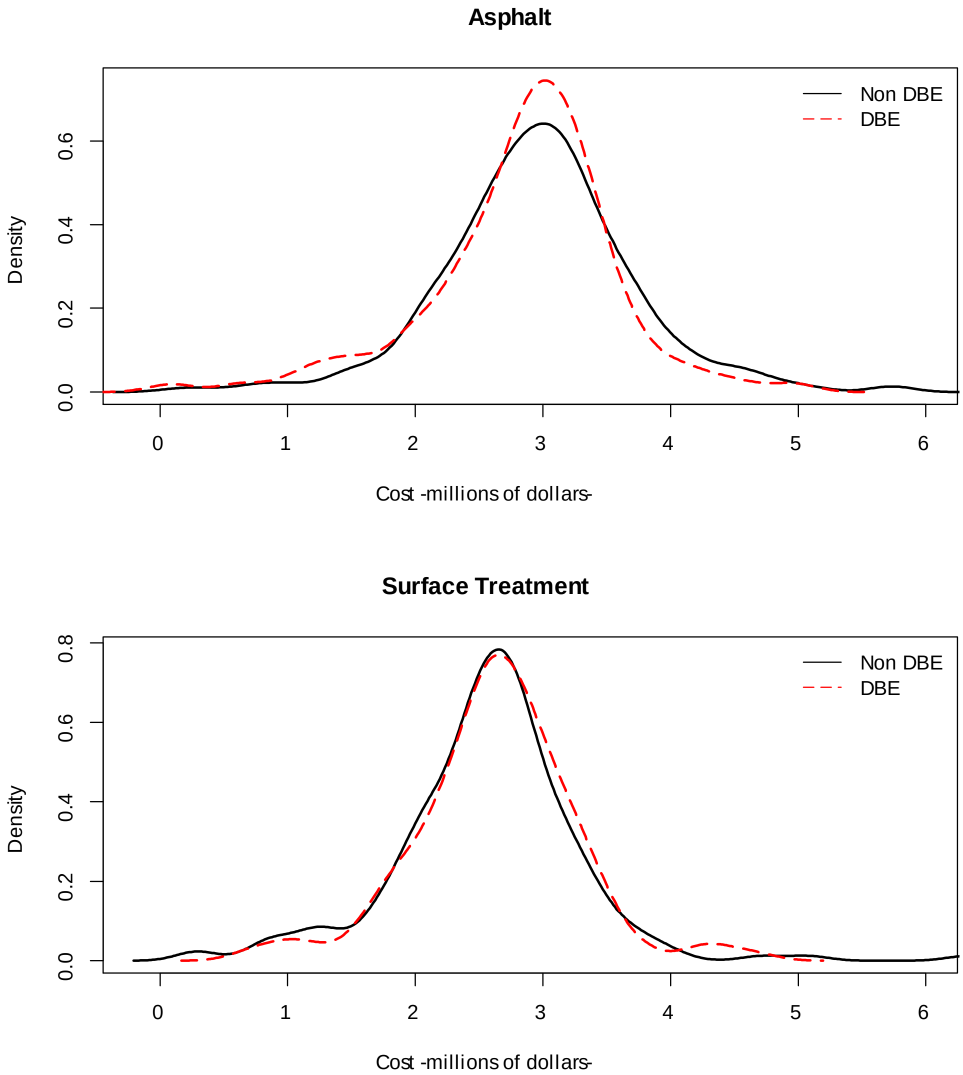
$$\begin{aligned}
\hat{g}_j &= \frac{1}{nL_j h_j} \sum_{l=1}^{L_j} \sum_{i=1}^{n} K\left(\frac{\hat{b}^0 - \hat{b}^0_{jil}}{h_j}\right), \\
\hat{G}_j &= \frac{1}{nL_j} \sum_{l=1}^{L_j} \sum_{i=1}^{n} 1\left\{\hat{b}^0_{jil} \leq \hat{b}^0\right\},
\end{aligned}$$

where, as before, $1\{\cdot\}$ is an indicator function, L denotes the number of auctions, $K(\cdot)$ is a continuously differentiable kernel function defined over a compact support, and h is a properly chosen bandwidth. We use the triweight kernel defined above.

Figure 3.2 presents a comparison between the cost distributions of projects with subcontracting goals and without subcontracting goals. We present the results for the sample of asphalt projects in the top panel, and the results for the sample of surface treatment projects in the bottom panel. The comparison of the cost distributions for asphalt projects presented at the top of Figure 3.2 suggests slightly different locations and different scales. These differences tend to disappear when we consider the sample of surface treatment projects. The evidence presented in the bottom panel indicates that the differences in the cost distributions

[14] We evaluated the sensitivity of our results to the choice of the mean function, letting $\Gamma(\cdot)$ to be a smooth function. We estimated the function considering standard local polynomial regression and generalized additive methods. In our application, the evidence suggests that the conclusions are not sensitive to the choice of the conditional mean function.

Figure 3.2: Cost densities for DBE and non-DBE projects. The densities were obtained using non-parametric methods employed on a sample of homogenized bids. These bids were obtained considering a parametric model that include auction specific characteristics as covariates and a control variate function. While the chart at the top is obtained using the sample of asphalt projects, the chart at the bottom is obtained using the sample of surface treatment projects.



	Quantiles of the Cost Distribution				
	1.000	2.000	3.000	4.000	5.000
			Asphalt projects		
Non-DBE	[0.003, 0.068]	[0.085, 0.309]	[0.487, 0.753]	[0.059, 0.285]	[0.001, 0.052]
DBE	[0.004, 0.094]	[0.077, 0.323]	[0.496, 0.867]	[0.025, 0.274]	[0.000, 0.034]
			Surface treatment projects		
Non-DBE	[0.009, 0.118]	[0.136, 0.726]	[0.190, 0.806]	[0.000, 0.148]	[0.000, 0.028]
DBE	[0.001, 0.119]	[0.088, 0.656]	[0.178, 0.908]	[0.001, 0.197]	[0.000, 0.029]

Table 3.2: Variability bands for the estimated densities in Figure 3.2. The intervals were constructed considering a block bootstrap procedure. The quantiles are in millions, and we considered 10,000 bootstrap repetitions. DBE stands for disadvantage business enterprises.

are negligible.[15]

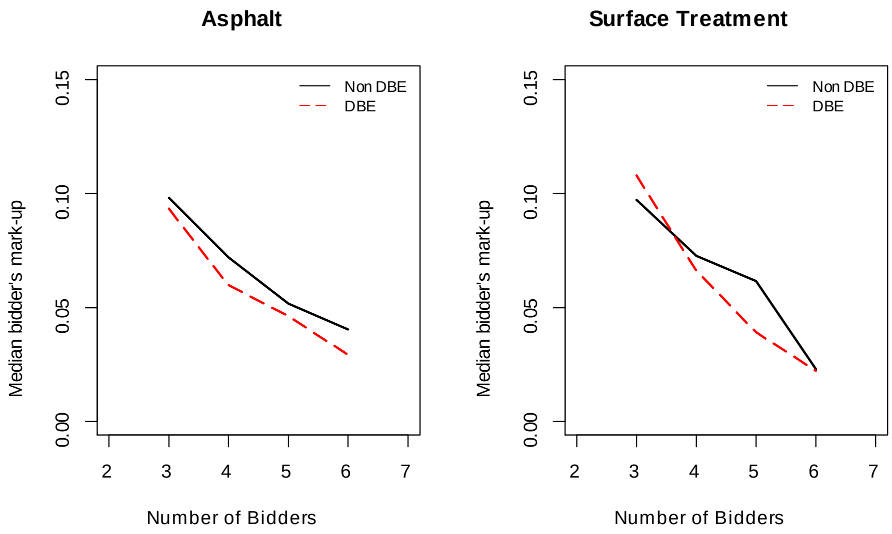
At first glance, the results presented in Figure 3.2 indicate that the cost distributions may not be significantly different. To examine this further, we provide 95 percent variability bands for several quantiles of the cost distributions in Table 3.2. Because the homogenized bids are based on estimates obtained in a first stage, standard pointwise confidence intervals for the densities cannot be used. Alternatively, we can use the bootstrap to provide a measure of the variability of the estimates (see, e.g., Pagan and Ullah 1999). Specifically, a bootstrap procedure is implemented as follows.[16,17] We draw an auction from a sample of projects and we include all bidders for that project. We continue sampling projects with replacement until we obtain a sample of L projects. Using this sample, we estimate $\Gamma(\mathbf{x},w)$. We then construct the homogenized bids $\hat{\mathbf{b}}^0$, separately for 3 and 4 bidders. Using these samples of bids, we compute the estimates of the DBE and non-DBE densities. We iterate the procedure 10,000 times. We finally construct pointwise 95 percent variability bands from the quantiles of the empirical distributions. The results of the table suggest that the DBE and non-DBE distributions are

[15] The motivation of controlling for endogenous participation is associated with projects with small bridge and earthwork components. We also compared the pseudo-cost distributions assuming that unobserved heterogeneity does not affect the identification of equations (3.2) and (3.3). After controlling for observed heterogeneity by estimating $\Gamma(\mathbf{x})$, we obtained costs distributions that were similar to the ones presented in Figure 3.2.

[16] Other bootstrap procedures have been implemented in the literature (see, e.g., Hendricks, Pinkse, and Porter (2003) and Krasnokutskaya (2010)). Hendricks, Pinkse, and Porter (2003) use a slightly different bootstrap procedure.

[17] In principle, we can use a simpler bootstrap procedure, because we might not need to account for dependence of any form and/or heterocedastic errors. We evaluate the framework's assumptions in Section 4.4.

Figure 3.3: Bidder's mark-up in projects with 3, 4, 5, and 6 bidders. The panels present the median mark-up, defined as $(b - \hat{c})/b$. While the left panel is obtained using the sample of asphalt projects, the right panel is obtained using the sample of surface treatment projects.



not statistically different, since the variability bands overlap at different quantiles of the cost distributions.[18]

To investigate the performance of our empirical strategy, we construct estimates of the median bidders' markup $(b - \hat{c})/b$ in auctions with 3, 4, 5, and 6 bidders. Extending the previous analysis to include auctions with 5 and 6 bidders allows us to examine in more detail the markups generated by the approach. Figure 3.3 shows continuous lines representing the median markup in non-DBE projects, and dashed lines denoting the median markup in the DBE projects. The panels show small differences between the continuous line and the dashed line, suggesting that the program did not generate considerable differences in the levels of the markups during our sample period. Moreover, these markups seem to be similar with

[18] It is important to note that inference is based on estimated costs, and therefore, the limiting distribution of the test may be affected. To the best of our knowledge, the literature offers one approach. An earlier version of Haile, Hong and Shum (2006) investigated a Kolmogorov-Smirnov test based on a resampling approach. According to the authors, the test performance was poor in small samples. Issues associated with inference using estimated costs are out of the scope of the paper.

n	Asphalt projects			Tests		Surface treatment projects			Tests	
	Median	Mean	SD	(1)	(2)	Median	Mean	SD	(1)	(2)
	non-DBE					non-DBE				
3	2.926	2.916	0.771	0.126	0.554	2.577	2.542	0.674	0.677	0.935
4	3.016	3.054	0.799			2.634	2.593	0.760		
	DBE					DBE				
3	2.946	2.797	0.813	0.787	0.540	2.579	2.601	0.656	0.259	0.286
4	2.911	2.903	0.638			2.735	2.666	0.598		

Table 3.3: Tests for invariance to number of bidders. Columns (1) present p-values of Wilcoxon tests. Columns (2) present the p-values of the standard Kolmogorov-Smirnov test.

the levels documented in the literature (see, e.g., Bajari and Ye (2003)), varying between 2 and 11 percent. Lastly, the downslopes show the effect of competition on markups in these procurement auctions.

3.4 Testing the IPV assumptions

Our analysis was performed using the symmetric independent private value framework, which essentially implies that exchangeability of marginal distributions and independence (Athey and Haile (2007)). Under exogenous variation of bidders, this framework suggests that the marginal distributions for $n = 3$ and $n = 4$ must be equal, because the costs are invariant to n (Lemma 1, Haile, Hong, and Shum (2006)).

In Table 3.3, we present evidence on estimates for the marginal distributions of projects with different number of bidders. While the columns marked as (1) provide p-values corresponding to Wilcoxon tests, the columns marked as (2) provide p-values corresponding to Kolmogorov-Smirnov tests.[19] The first statistic is a common test suggested in the literature to evaluate difference in location, and it is applied to evaluate if the costs distributions have same locations. The second statistic is a test for independence, and it is applied in this case to evaluate if the cost distributions in auctions with 3 and 4 bidders are significantly different. The testing procedures are described in the Appendix.

[19] The tests are based on estimated costs, which might lead to incorrect inference. The estimation of costs may influence the limiting distribution of our tests. To the best of our knowledge, there is not well-established correction to this difficult problem. This issue is out of the scope of our paper.

The tests seem to suggest that the sample of projects exclusively related to surface treatment fits the framework better. The costs distributions obtained using the surface treatment projects vary less in the number of bidders, and therefore, they appear to satisfy the condition on the marginal distributions associated to the IPV framework.

The IPV framework also relies on the independence of the bids submitted to an auction. We employed three testing procedures to evaluate conditional independence on pairs of bids (i, j) in auctions with 3 and 4 bidders: (i) the non-parametric test proposed by Blum et al. (1961); (ii) Kendall τ rank correlation coefficient; (iii) a Kolmogorov-Smirnov test for independence. The findings can be summarized as follows: (a) we failed to reject the null hypothesis at 5 percent in 101 tests out of the 108 performed tests, and (b) surface treatment projects provided a better fit to the assumptions of the model relative to the sample of asphalt projects.

3.5 Additional Considerations on Selection Issues

The previous analysis shows that there is little difference in project costs between projects that are assigned subcontracting goals versus projects that are not assigned goals. This section discusses a few additional issues associated with bidder's participation and project heterogeneity. We previously addressed endogenous participation influenced by project unobserved heterogeneity using the method proposed by Haile, Hong and Shum (2006). In the analysis of the DBE program however, one needs to consider that the program may affect costs, and therefore, the participation of bidders in an auction. The DBE program might be affecting participation in auctions with and without DBE subcontracting goals. In order to address this issue, we restricted attention to bidders participating in both DBE and non-DBE auctions.

Therefore, our previous results might not be affected by bidder's selection issues.[20,21]

A more important issue seems to be associated with DBE assignments. As we discussed earlier, it is likely that the state would assign DBE status to a project with a large number of tasks involved.[22,23] The possibility of this type of selection bias could be incorporated into the analysis by replacing the selection probability by a non parametric function (Das, Newey, and Vella 2003). A more convenient approach for this setting with a relatively large number of covariates, is to estimate the selection probability by the propensity score. The propensity score s is the conditional probability of selection estimated by standard parametric models (e.g., probit). We use the total number of bid items in a project, the number of days to complete the project and indicators for the location of the project to estimate the conditional probability of selection. We observe that these first two effects have the expected sign and are significant at 5 percent. To obtain the homogenized bid, we now condition on $\hat{s}$, and therefore the first stage regression is now $\mathbf{b} = \tilde{\Gamma}(\mathbf{x}, \hat{w}, \hat{s}) + \mathbf{u}$. The panels in Figure 3.4 present estimates of 3.2 and 3.3 that use these samples of bids $\hat{\mathbf{b}}^0$. After controlling for observed and unobserved heterogeneity and the possibility of selection bias, we again find that the costs

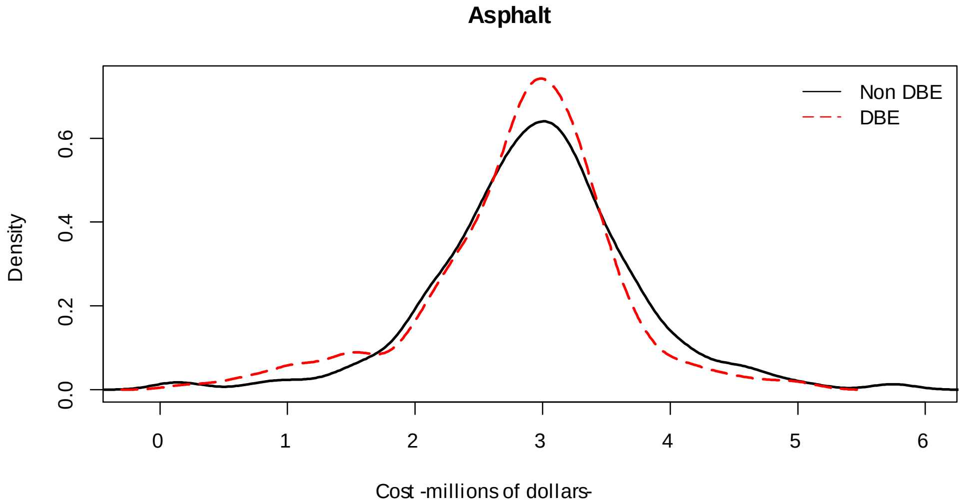
[20] The samples of asphalt projects and surface treatment projects presented in Table 3.1 exclude bidders who participated in DBE or non-DBE auctions alone. The vast majority of bidders participate in both auctions. The asphalt sample includes 84 percent of all bidders participating in the auctions throughout the period of analysis, and the surface treatment sample includes 82 percent of all bidders. Although the sample sizes are reduced, these sample refinements minimize and potentially eliminate issues associated with bidder's participation.

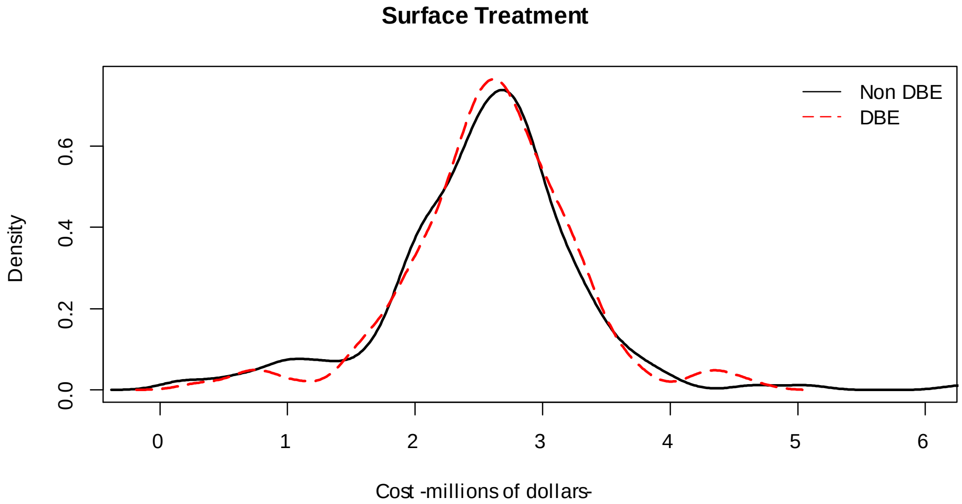
[21] As a robustness check, we also estimated the models including all bidders. Our findings revealed that the results presented in this paper are not affected dramatically. We continued to find small differences in costs in the asphalt sample and no apparent differences in costs in the surface treatment sample. The mark-ups ranged from 2 to 10 percent, as in Figure 3.3. We do not present the results to avoid repetition, but they are available upon request.

[22] A natural concern in the first stage regression is the suspected endogenous indicator for DBE assignment (see, e.g., Marion 2010). It is important to note that by the nature of the exercise, the first stage regression does not include a suspected endogenous variable, but of course, the non-random assignment $j = \{0, 1\}$ may create biases.

[23] More formally, the state would assign DBE status to a project if $1\{\mathbf{d}'\boldsymbol{\delta} + \eta > 0\}$, where $1\{\cdot\}$ is an indicator function. The vector $\mathbf{d}$ includes the total number of bid items (project components) in a project and the availability of minority firms given by the geographic location of the project. The variable η is assumed to be an error term that could be correlated with the error term in the model for $\mathbf{b}$ but it is independent of $\mathbf{d}$. Because in the first stage $\mathbf{b} = \tilde{\Gamma}(\mathbf{x}, \hat{w}) + \mathbf{u}$, it is then possible that $E\{\mathbf{u}|\mathbf{x}, w, \mathbf{d}'\boldsymbol{\delta} > \eta\} \neq 0$ generates selection bias even in the case that $\mathbf{u} \perp \mathbf{x}$. Although they seem to represent two different issues, addressing observed heterogeneity is related in our case to correcting for selection bias. For identical projects with characteristics $(\mathbf{x_0}', w_0)'$, one would expect $\tilde{\Gamma}(\mathbf{x}, w) = \boldsymbol{\Gamma}(\mathbf{x}, w) - \boldsymbol{\Gamma}(\mathbf{x_0}, w_0) = 0$, and also $E\{\mathbf{u}|\mathbf{x}, w, \mathbf{d}'\boldsymbol{\delta} > \eta\} = 0$, simply because $\mathbf{d}'\boldsymbol{\delta}$ would tend to be constant.

Figure 3.4: Cost densities for DBE and non-DBE projects. The densities were obtained considering non parametric methods employed on a sample of homogenized bids. The panels were obtained from a model that address observed projects heterogeneity, endogenous participation, and the possibility of selection bias. The upper panel is obtained using the sample of asphalt projects and the lower panel is obtained using the sample of surface treatment projects.





distributions present small differences, which turn out to be negligible when we consider the sample of surface treatment projects.

4 Conclusion

This paper examines the differences in bidding and costs between projects that have subcontracting goals and projects that do not. The analysis uses the nonparametric structural approach developed by Haile, Hong and Shum (2006) that allows one to control for many auction-specific characteristics and endogenous participation without increasing sample size. This is particularly important in our setting as project size, complexity, materials use and other characteristics vary markedly across projects. Our empirical results show little difference in the level of bids submitted or in the estimated costs between projects with subcontracting goals and projects without such goals. When we utilize an even more homogeneous sample of projects, the differences are even less. Finally, we show that the implied markups generated from the Haile, Hong and Shum approach are consistent with those reported in the literature and do not differ substantially for auctions with and without subcontracting goals.

A simple interpretation of the result is that the supply and quality of DBE subcontractors was sufficient during our period of analysis so that prime contractors were effectively unrestricted in their bidding due to the presence of DBE requirements. The Census Bureau's 2002 Survey of Business Owners indicates that Texas has a relatively large number of minority-owned construction firms in comparison to the average state, reflecting, at least in part, the large minority population of the state. Moreover, our findings do not necessarily mean the program has had no effects on contracting. The program may have encouraged the formation and success of minority and women-owned businesses increasing the supply of DBE subcontractors, something that we cannot test with our data. Alternatively, the program may have affected project costs but the effects may have occurred outside our window of observation. Specifically, they may have occurred when the program was introduced – several decades before our period of analysis. That said, our results suggest that during the period under study DBE subcontracting requirements did not substantially raise the bids or costs of prime contractors.

References

ATHEY, S., AND P. A. HAILE (2007): "Nonparametric Approaches to Auctions," *Handbook of Econometrics, Edited by J. Heckman and E. Leamer*, 6A, Elsevier.

BAJARI, P., AND L. YE (2003): "Deciding Between Competition and Collusion," *Review of Economics and Statistics*, 85(4), 971–989.

BLUM, J. R., J. KIEFER, AND M. ROSENBLATT (1961): "Distribution Free Tests for Independence based on the Sample Distribution Function," *Annals of Mathematical Statistics*, 32, 485–498.

CAMPO, S., I. PERRIGNE, AND Q. VUONG (2003): "Asymmetry in First-Price Auctions with Affiliated Private Values," *Journal of Applied Econometrics*, 18(2), 179–207.

DAS, M., W. K. NEWEY, AND F. VELLA (2003): "Nonparametric Estimation of Sample Selection Models," *Review of Economic Studies*, 70, 33–58.

DE SILVA, D. G., T. DUNNE, A. KANKANAMGE, AND G. KOSMOPOULOU (2008): "The Impact of Public Information on Bidding in Highway Procurement Auctions," *European Economic Review*, 52, 150–181.

DENES, T. A. (1997): "Do Small Business set-Asides Increase the Cost of Government Contracting?," *Public Administration Review*, 57(5), 441–44.

GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

Haile, P. A., H. Hong, and M. Shum (2006): "Nonparametric Tests for Common Values In First Price Sealed-Bid Auctions," working paper.

Hájek, J., Z. Šidak, and P. K. Sen (1999): *Theory of Rank Tests*. Academic Press.

Härdle, W. (1991): *Smoothing Techniques with Implementation in S*. Springer-Verlag, New York.

Hendricks, K., J. Pinkse, and R. H. Porter (2003): "Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions," *Review of Economic Studies*, 70(1), 115–45.

Holzer, H., and D. Neumark (2000): "Assessing Affirmative Action," *Journal of Economic Literature*, 38, 483–568.

Jofre-Bonet, M., and M. Pesendorfer (2003): "Estimation of a Dynamic Auction Game," *Econometrica*, 71(5), 1443–89.

Krasnokutskaya, E. (2010): "Identification and Estimation in Highway Procurement Auctions Under Unobserved Auction Heterogeneity," forthcoming in Review of Economic Studies.

Krasnokutskaya, E., and K. Seim (2007): "Preferential treatment Program and Participation Decisions in Highway Procurement," working paper, Department of Economics, Univesity of Pennsylvania.

Li, T., I. Perrigne, and Q. Vuong (2000): "Conditionally independent private information in OCS wildcat auctions," *Journal of Econometrics*, 98, 129–161.

Marion, J. (2007): "Are Bid Preferences Benign? The Effect of Small Business Subsidies in Highway Procurement Auctions," *Journal of Public Economics*, 91, 1591–1624.

——— (2009): "How Costly is Affirmative Action? Government Contracting and California's Proposition 209," *Review of Economics and Statistics*, 91(3), 503–522.

——— (2010): "Affirmative Action and the Utilization of Minority- and Women-Owned Businesses in Highway Procurement," forthcoming in Economic Inquiry.

Maskin, E., and J. Riley (2000): "Asymmetric Auctions," *Review of Economic Studies*, 67(3), 413–438.

McAfee, R. P., and J. McMillan (1989): "Government Procurement and International Trade," *Journal of International Economics*, 26, 291–308.

Pagan, A., and A. Ullah (1999): *Nonparametric Econometrics*. Cambridge University Press.

A Appendix

A.1 Testing Procedures

Let $\{(Z_i, V_i)\}_{i=1}^n$ be random samples with densities g_Z, g_V, and joint density $f(z, v)$. Hájek, Šidák, and Sen (1999) suggest the following tests for independence and location.

1. B test: This test, which was employed in Campo, Perrigne, and Vuong (2003), was proposed by Blum et al. (1961). The test statistic is equal to,

$$B_n = \iint (F(z,v) - G_Z(z)G_V(v))^2 dF(z,v),$$

   where $F(z, v) = n^{-1} \sum\limits_{i=1}^{n} I(Z_i \leq z, V_i \leq v)$, $G_Z(z) = n^{-1} \sum\limits_{i=1}^{n} I(Z_i \leq z)$, and $G_V(v) = n^{-1} \sum\limits_{i=1}^{n} I(V_i \leq v)$. This test is consistent and distribution free.

2. Kolmogorov-Smirnov test: Using the previous definitions, we can write this test as $KS_n = \sup |F(z, v) - G_Z(z)G_V(v)|$.

3. Wilcoxon rank test: We set $Z_{n+j} = V_j$ for $j = 1, ...n$ and $N = 2n$. Let R_i $(i = 1, ...N)$ be the rank of the observation Z_i in the ordered sequence $Z^{(1)} < Z^{(2)} < ... < Z^{(N)}$. This test is based on the statistic $S = \sum\limits_{i=1}^{n} R_i$. Another form of the test is called Mann-Whitney statistic, which is based on the number of pairs (Z_i, V_i) such that $Z_i < V_j$. Under the null hypothesis of no differences in location, the standardized version of S is asymptotically normal as $n \to \infty$.

4. Kendall τ rank correlation test: Let R_i $(i = 1, ...n)$ be the rank of the observation Z_i in the ordered sequence $Z^{(1)} < Z^{(2)} < ... < Z^{(n)}$ and Q_i $(i = 1, ...n)$ be the rank of the observation V_i in the ordered sequence $V^{(1)} < V^{(2)} < ... < V^{(n)\cdot}$. This test is based on the following statistic: $\tau = (n(n-1))^{-1} \sum\limits_{i \neq j}\sum sgn(R_i - R_j) sgn(Q_i - Q_j)$, where $sgn(A)$ denotes the sign of A. Under H_0, the standardized version of τ tends to a Gaussian distribution.

Variable	Definition
Log of bids	Log value of bids
Log of ECE	The log value of the engineer's cost estimate (ECE).
Bridge work percentage	The value of the bridge work bid items relative to the ECE.
Earth work percentage	The value of the earth work bid items relative to the ECE.
Pavement work percentage	The value of the pavement work bid items relative to the ECE.
Concrete work percentage	The value of the concrete work bid items relative to the ECE.
DBE projects	Projects with DBE goals
State projects	State funded projects
DBE: 0% (Fed projects)	DBE goal dummy for goals: 0% (for federal projects.)
DBE : 1% - 5%	DBE goal dummy for goals: 1% - 5%
DBE : 6% - 7%	DBE goal dummy for goals: 6% - 7%
DBE : > 7%	DBE goal dummy for goals: > 7%
Complexity	The total amount of bid items (project components) in a project described by TxDOT.
Calendar days	Number of days to complete the project assigned by TxDOT
Division dummies	TxDOT has 25 divisions, which are identified by division dummies
Zone dummies	TxDOT divides Texas into five major geographic zones. We identify these zones using zone dummies
Material shares for asphalt projects	They include mainly surfacing, earthwork, miscellaneous construction, drainage and foundation structures, subgrade and base, and traffic.
Distance to the project location	The distance between the county the project is located in and the distance to the county of the firm's location [log(distance+1)]
Backlog	It is constructed by summing across the non-completed value of the contract of existing contracts. The variable is similar to the variables used by Bajari and Ye (2003) and Jofre-Bonet and Pesendorfer (2003).

Table A.1: Variable definitions.

Item description	Federal projects		State Projects
	With DBE goals	Without DBE goals	
Preparing right of way	2,249 [59.01]	789 [20.70]	773 [20.28]
Embankment	2,835 [59.81]	970 [20.46]	935 [19.73]
Topsoil	888 [57.33]	287 [18.53]	374 [24.14]
Compost	244 [49.90]	114 [23.31]	131 [26.79]
Sodding for erosion control	1,064 [59.91]	316 [17.79]	396 [22.30]
Seeding for erosion control	5,020 [59.64]	1,718 [20.41]	1,679 [19.95]
Fertilizer	466 [61.48]	139 [18.34]	153 [20.18]
Vegetative watering	2,528 [56.62]	986 [22.08]	951 [21.30]
Soil retention blankets	851 [66.07]	218 [16.93]	219 [17.00]
Irrigation system	174 [40.47]	73 [16.98]	183 [42.56]
Wildflower seeding	43 [44.79]	21 [21.88]	32 [33.33]
Landscape planting	1,791 [44.95]	521 [13.08]	1,672 [41.97]
Landscape establishment	142 [31.91]	33 [7.42]	270 [60.67]
Salvaging, stockpiling asphalt pavement	817 [72.36]	144 [12.75]	168 [14.88]
Barricades, signs, and traffic handling	3,106 [39.83]	2,073 [26.58]	2,619 [33.59]
Erosion & environmental controls	6,664 [47.95]	3,300 [23.75]	3,933 [28.30]
Constructing detours	1,061 [78.59]	149 [11.04]	140 [10.37]
One-way traffic control	148 [45.40]	42 [12.88]	136 [41.72]
Portable concrete traffic barrier	3,996 [77.55]	576 [11.18]	581 [11.27]
Permanent concrete traffic barrier	1,024 [81.66]	111 [8.85]	119 [9.49]
Textured concrete & landscape pavers	203 [61.70]	34 [10.33]	92 [27.96]
Concrete curb and gutte	2,366 [67.75]	440 [12.60]	686 [19.64]
Right of way markers	203 [66.78]	86 [28.29]	15 [4.93]
Crash cushion attenuators	1,181 [76.69]	188 [12.21]	171 [11.10]
Chain link fence	376 [79.49]	50 [10.57]	47 [9.94]
Wire fence	423 [60.52]	186 [26.61]	90 [12.88]
All DBE items	39,863 [56.95]	13,564 [19.38]	16,565 [23.67]
All items	300,680 [59.76]	88,459 [17.58]	114,035 [22.66]

Table A.2: Bid items in federal and state projects. Percentages are in brackets.