Measuring Learning-by-Doing in an Auction Setting with Heterogeneous Bidders

Przemyslaw Jeziorski* Elena Krasnokutskaya† Lucia Tiererova‡

August 23, 2017

Preliminary and Incomplete. Do not circulate.

Abstract

This paper utilizes fifteen years of data reflecting operation of California highway procurement market to study the impact of experience on new bidders' costs while accounting for unobserved bidder heterogeneity.

Keywords: highway procurement, learning, unobserved agent heterogeneity

JEL Classification: C73, C63, D44, H57, L20

*Haas School of Business, email: przemekj@haas.berkeley.edu

†Johns Hopkins University, email: ekrasno1@jhu.edu

‡Johns Hopkins University, Department of Economics. Email: ltierer1@jhu.edu

1 Introduction

Learning-by-doing is an important phenomena which motivates many government policies. New firms may be reluctant to enter markets where they need to acquire experience in order to effectively compete with incumbents; and where the experience acquisition in turn is likely to be impeded by entrants' inferior productivity. These considerations may significantly limit entry or endanger survival of new firms in the markets where substantial learning is required.

We utilize data from California Department of Transportation, which is responsible for the maintenance and construction of roads in the state of California, to study the impact of learning-by-doing on new entrants' success in the market place. The jobs in this market are organized in the form of projects which are allocated through first price sealed bid auctions. We have access to all bids for all projects auctioned in this market between January 2002 and December 2016. This allows us to follow progress of firms entering the market, and account for the circumstances when they are or they are not successful. California Department of Transportation is the main source of work for the firms specializing in the road repairs. While some private jobs are also available, the market for these jobs is either very limited in size or, in the case of large corporations, is difficult to break in for a new firm in the market. We are thus able to account for the predominant majority of work performed by new firms.

California procurement market provides a rich setting in which to study the effects of learning-by-doing. Indeed, we document that on average two hundred new firms enter this market every year. The turnover of new firms is very high with 40% of firms exiting the market after one year and 60% after two years of entering it. Since we observe the firms' participation and prices at the project level (given the competitive structure of the auction) we are able to relate their behavior and perfomance to the successes that had in the past. Due to the percieved importance of learning-by-doing in this market, several goverment policies are put in place which aim to facilitate growth of young firms by enabling them to win despite somewhat less competitive prices.[1] At the same time, very few formal studies of learning in this market exist so that benefits of such programs in terms of new firms' survival and improved productivity remain largely unknown.[2]

The analysis of learning-by-doing is complicated by a classical problem of potential unobserved heterogeneity of firms. Indeed, a firm may win early in the career and can go on winning

[1] These programs essentially permit young firms to include a "surcharge" in their bids which is discounted at the allocation stage. More details on the bid preference programs can be found in Krasnokutskaya and Seim (2011).

[2] The exeptions are Tirerova (2014) and .

because it was lucky initially and then was able to sustain its success by improving its productivity through learning-by-doing or, alternatively, because it is more productive relative to its competitors which helped it to win early in its career and helps it to keep winning as time progresses. When both effects are present, unobserved bidder heterogeneity, if not taken into account, would bias researchers' assessment of the importance of learning.

In our market the prices are set at the project level which allows us to observed prices charged by the firm at different levels of experience while controlling for competitive incentives. This could, in principle, inform us about the changes in the firms' costs arising from accumulated experience. The difficulty with this approach is that in procurement setting firm's costs to some degree stochastically vary across projects. In order to assess a trend in costs associated with changing experience levels, a researcher has to pool the observations associated with different firms since the number of firms-specific observations at a given level of experience is rather small. In the presence of unobserved bidder heterogeneity such pooling would bias the estimates of the effect of experience on costs.

The challenge therefore is to decompose the observed distribution of bids, which is a mixture of bid distributions associated with different types of new entrants, while allowing that the distributions of costs conditional on type (and thus the distributions of bids conditional on type) may vary across experience levels. The important difficulties are that the initial distribution of types is not observed and that the transition across experience levels is also likely to be type-specific. These issues present substantial challenges to the analysis. While some very general results on the nonparametric identification of mixture models have recently become available (see, for example, Bonhomme, Jochman, and Robin (2015), Bonhomme, Jochman, and Robin (2015) and Compiane and Kitamura (2016)), practical implementation of these methodologies in our setting is difficult. Indeed, these methodologies require that we should observe in the data the joint distribution of a vector of observation which are drawn from the same mixture distribution. In our setting, due to the importance of competition, we would need to observe for each set of firms, that could plausible compete in the same auction, a collection of bids these firms submit while competing in the same auction on several occasions (the set of competitors in each case has to be restricted to these firms and the experience levels of all new firms should remain unchanged). These restrictions impose impossibly high data requirements; these requirements are not met even in our dataset which has a relatively long time dimension and a very large number of projects auctioned over time.

Our solution to this challenge is to utilize the fact that our data document behavior of many firms at different levels in their growth progression. A key assumption that enables our analysis is that for the firms that succeed and stay in the market the impact of experience eventually flattens out. In other words, the cost distribution of firms that has been in the market for some time and completed "many" projects is stable and is invariant to new successes (we can allow for stationary variations in the distribution of costs associated with changes in backlog – see on this below). For such firms the mixture decomposition of the distributions of bids (and costs) is feasible since many observations are available. Once the distributions of costs conditional on type are recovered, a researcher can then tackle the subpopulation of new firms recovering the adjustment to cost distributions associated with experience, as well as the mixing weights associated with different unobserved bidder types in the population characterized by a given level of experience.

Several comments are in order.

(a) Within this framework we may allow that the effect of experience on costs depends on bidder unobserved type. Importantly, even if the impact of experience is restricted to be type-invariant, the estimates of the impact of experience on costs which ignore unobserved heterogeneity would be biased dur to selection into winning on type.

(b) Notice, that this approach allows us to recover mixing probabilities conditional on experience level in unrestricted way.

(c) This approach does impose that at least for some early levels of experience the set of unobserved bidder types should be the same as in the subpopulation of established firms. This assumption does not appear to be particularly restrictive given that an element of stochasticity is present in bidders' costs which implies a degree of randomness inherent in winning a project. Also, once other model primitives are recovered, it should be feasible to allow for the presence of extra type(s) (which disappear in subpopulation of successful firms) at the initial levels of experience.

(d) In the current draft we distinguish between established large and established small firms. For established large firms we take into account the impact of backlog on costs. Essentially, the assumption here is that these firms are large enough that they continue to operate even when carrying some backlog (but their costs go up) whereas established small firms work on one project at a time. We allow for the distribution of costs to differ between these two sets of firms. In other words we allow the scale effects to exist on the upper end of age spectrum.

We assume that the distribution of costs of new firms approaches the distribution of small established firms when they accumulate sufficient experience.

Once the primitives of the model are recovered, we would use the estimates to evaluate the effectiveness of the bid preference programs utilized in government procurement. Specifically, we plan to assess the impact on firms' survival and on the long-term cost of procurement.

The paper is organized as follows. Section two summarizes the features of California highway procurement market. Section three discusses the data and provides summary statistics and some descriptive regression results. Section four outlines the model. Section five discusses estimation approach and Section six presents some preliminary results.

2 Highway Procurement Market

Our analysis focuses on California market for highway procurement which is supervised by California Department of Transportation (CalTrans). The firms participating in this market specialize in highway repairs and highway-related construction. The work is organized in the form of projects which are allocated through first-price sealed bid auctions. When an auction for the project is announced the firms that are potentially interested in performing this work have to purchase bid documents which outline the detailed list of steps involved in project execution prepared by CalTrans engineers. Firms then have to decide whether to prepare a bid for this project or not. Bid preparation involves finding out the (firm-specific) costs of completing this project. This step involves negotiation with suppliers and subcontractors as well as developing a realistic plan of project execution. During this stage the set of firms planning to submit a bid for a given project becomes generally known. Before submitting a bid the firms also have to figure out a mark-up they would like to charge on top of their costs which takes into account the realized competition for the project.

It is a common belief in this market that new firms tend to be less competitive (have higher costs) relative the more experienced firms. However, their efficiency (productivity) improves with experience. If new firm fails to win sufficient number of projects in the beginning it may have to exit before it had a chance to improve its productivity. The potential importance of learning motivates several government bid preference programs which help new firms to win projects despite their relatively low productivity. On the other hand, it is possible that the role of learning is overestimated if the firms who win projects in the beginning and thus survive to stay in the market are the firms that have higher (unobserved) productivity whereas those that drop out from the market due to "failing to learn" are, in fact, very low productivity firms. In other words, selection into winning may be mistaken for the effect of learning. The goal of this paper is to disentangle these two effects.

3 Data

We have access to the data on 15 years of operation of the California highway procurement market (2002-2016). The data contain information on all projects offered for bid in this market, all firms that purchased bid documents for each project, the firms that eventually submitted bids, the amounts of the bids, and the identities of the winners. This allows us to trace the progress of the firms entering the market, and relate their survival to the success in winning projects.

3.1 Market Structure

We define a firm to be new entrant in a given year if we do not observe this firm submitting a bid in a previous two years.[3] We further subdivide the set of incumbent firms into large and small firms using the definition of a small firm adopted by the CalTrans. Specifically, the firm is considered small if the total receipts from market for the last three years do not exceed $10mln.

Table 1 summarizes the distribution, behavior and the performance of firms in the market. The Table indicates that this market attracts a large number of new entrants. Indeed, during the six years that are reflected in the Table, 1249 new firms entered the market (on average 208 per year). These firms, however, submit relatively small number of bids during their tenure in the market (7 bids on average) and win only one project on average before exiting the market. Table 2 sheds some light on this regularity by reporting the exit behavior of new firms by cohort. As the findings in this table indicate, the set of new firms is characterized by extremely high turnover. Indeed, on average 40% of the entering cohort exits after one year in the market, and the further 20% exit after two years in the market. On the other hand, around 30% of entering cohort stays longer than five years.

Returning to Table 1, we find that the set of large incumbents contains only 20 firms whereas the set of small incumbents consists of 166 firms. However, large firms tend to submit five times as many bids and win roughly five times as many projects as small incumbents do.

[3] This approach requires us to discard the first two years of data.

Table 1: Summary Statistics: Firms

	Firms			
			Experienced	
	All	New	Large	Small
Number	1,535	1,249	20	166
Bids submitted:				
Total	26,714	8,812	6,881	11,421
Average (across firms)	17.4	6.7	344	69
Std.Dev.(across firms)	57.6	17.0	49.3	41.2
Projects won:				
Total	4,562	1,282	1,233	2,047
Average (across firms)	3.0	1.0	60.1	12.3
Std.Dev.(across firms)	12.3	3.0	16.1	10.3

Notes: This table reports statistics describing the set of firms participating in California Highway Procurement market. The results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

Table 2: Summary Statistics: Turnover

Entering	Remaining in the market:						
Market	2005	2006	2007	2008	2009	2010	2011
2005	103	73(71%)	72(70%)	66(64%)	59(57%)	49(48%)	24(23%)
2006	–	131	94(72%)	81(62%)	72(55%)	55(42%)	45(34%)
2007	–	–	185	124(67%)	103(56%)	74(40%)	46(25%)
2008	–	–	–	218	93(43%)	63(29%)	43(20%)
2009	–	–	–	–	262	96(37%)	49(19%)
2010	–	–	–	–	–	193	87(41%)
2011	–	–	–	–	–	–	150

Notes: This table reports statistics describing the turnover of new firms in California Highway Procurement market. Each row characterizes the set of firms entering the market in a particular year. The results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

3.2 Auction-related Summary Statistics

Table 3 provides summary of auction-related statistics. In this market around 700 projects are sold every year. The Table indicates that the projects vary in size quite substantially: from $\$190,000$ on a lower end to $\$10,670,000$ on the upper end of the size distribution with the median project's value equal to $\$700,000$. Projects further differ in the allowed duration which ranges from one to eight months with the median duration equal to 50 days. Engineering description breaks the project into items with the number of items ranging from 14 to 57 and the median

Table 3: General Summary Statistics

		Project Size			
		All	Small	Medium	Large
Engineer's estimate (in mln)	mean	4.83	0.19	0.54	10.67
	median	0.69	0.19	0.53	3.25
	s.d.	16.57	0.05	0.16	24.16
Working days	mean	128.25	32.41	57.10	231.27
	median	50.00	30.00	35.00	125.00
	s.d.	192.94	19.16	69.01	251.97
Number of items	mean	43.18	15.51	21.83	73.75
	median	22.00	14.00	17.00	57.00
	s.d.	50.05	7.81	12.84	63.52
Number of bidders	mean	4.94	5.18	4.83	4.97
	median	4.00	4.00	4.00	5.00
	s.d.	2.68	3.18	2.68	2.52
Money-left-on-the-table (%)	mean	8.94	14.05	10.29	5.96
	median	6.04	9.31	7.05	4.59
	s.d.	12.11	18.56	10.91	9.74
(estimate-bid1)/estimate (%)	mean	−6.93	−10.86	−5.94	−6.69
	median	−7.85	−12.47	−6.25	−8.15
	s.d.	20.64	23.98	21.28	18.68

Notes: This table reports statistics summarizing auction-related features of the data. The results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

number of items given by 22. Our algorithm for aggregating items into tasks which is described above indicates that the project on average consists of four tasks (the average number of tasks is 3.77 and the median number is 4). Also, an auction on average attracts 5 bidders (the median number of bidders is 4 with smaller projects attracting higher number of participants).

The table further reports statistics traditionally considered in auction markets. The variable "money-left-on-the table" is constructed as the difference between the lowest and second lowest bid divided by the lowest bid: $\frac{(\text{rank2}-\text{rank1})}{\text{rank1}}$. It reflects the level of uncertainty about contractors' costs or informational asymmetries in the market. In our data, the second lowest bid is on average 9% higher than the winning bid. This is comparable to other datasets used to study highway procurement market. Information asymmetries can also be seen from the relative difference of the winning bid to the engineer's estimate. The winning bid is on average 7% below the engineer's estimate.

Table 4: The Effect of Experience on Bid Levels

Variable	(I)	(II)	(III)
Cum.Win(6mo)	-0.011***	-0.013***	-0.005*
	(0.003)	(0.003)	(0.003)
Cum.Win(6to12mo)	-0.002	-0.001	-0.003
	(0.003)	(0.003)	(0.003)
Cum.Win(12to18mo)	-0.013***	-0.013***	-0.005***
	(0.003)	(0.003)	(0.003)
Cum.Win(18mo)	-0.006***	-0.013***	-0.005***
	(0.003)	(0.003)	(0.003)
Federal Aid	-0.020	-0.030**	-0.051***
	(0.012)	(0.014)	(0.013)
Log(Estimate)	1.017***	1.017***	1.002***
	(0.006)	(0.006)	(0.005)
Potential Bidders			
Total number	-0.007***	-0.007***	-0.002**
	(0.001)	(0.001)	(0.001)
Number of new	-0.001	-0.001	-0.006***
	(0.002)	(0.002)	(0.002)
Backlog*Large	–	0.012***	0.007*
		(0.004)	(0.004)
Constant	-0.164**	-0.149**	-0.208**
	(0.077)	(0.077)	(0.083)
Year FE	N	N	Y
Month FE	N	N	Y
District FE	N	N	Y

The standard errors are in parenthesis. he results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

3.3 Descriptive Regressions

Table 4 reports the results of a descriptive regression relating the bids to the firm's type and experience. In this regression experience is measured by the number of projects the firm won by the time it submits a bid in question. We find that that experience has a statistically significant impact on bids even after controlling for project characteristics and potential competition. However, it is difficult to interpret these estimates in the context of the descriptive regression. Indeed, we expect two effects to be present: (a) the cost may decline with experience which should allow firm to submit more competitive (lower) bids. On the other hand, if firm anticipates benefits of learning it should submit lower bids in the beginning of the career to increase probability of winning which should facilitate learning. Thus we would expect bids to increase with experience as the need for learning diminishes. The magnitude and the direction of the impact associated with experience thus depends on the relative magnitudes of these two effects. Indeed, we observe that impact of the experience varies over time while remaining mostly negative in the end.

Table 5: The Effect of Experience on Probability of Entry

	Large	Small	New
Constant	0.671*** (0.005)	0.418*** (0.003)	0.265*** (0.006)
Estimate	-0.047*** (0.004)	-0.125*** (0.003)	-0.125*** (0.004)
Working Days	0.006 (0.003)	0.003 (0.002)	0.0005 (0.001)
Potential Bidders:			
# of Large	-0.067*** (0.002)	-0.041*** (0.002)	-0.039*** (0.001)
# of Small	-0.029*** (0.003)	-0.021*** (0.003)	-0.003 (0.002)
# of New (exp1-exp3)	-0.013 (0.010)	-0.019** (0.010)	-0.007 (0.010)
# of New (exp4-exp5)	-0.018*** (0.002)	-0.022*** (0.003)	-0.014*** (0.003)
Cum.Wins	–	–	0.024 (0.004)

The standard errors are in parenthesis. he results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

Table 5 further investigates the impact of experience on the probability of participation (the likelihood that firm submits a bid conditional on purchasing bid documents). We find that the likelihood of participation increases with the number of projects won. Again, this result reflects two effects. On one hand, experience may result in higher expected profit for any given projects (lower costs and higher probability of winning) or it may lower participation costs. It is not possible to distinguish between these two effects on the basis of descriptive regressions.

This analysis, of course, does not allow us to take into account the fact that new firms may potentially be heterogeneous in the ways that are not reflected in our data. We now turn to the structural analysis of this setting where we attempt to disentangle different channels through which experience impacts firms while allowing for the unobserved heterogeneity of firms participating in this market.

4 Model

I this section we summarize the model which incorporates learning in the presence of unobserved bidder heterogeneity.

4.1 Market Setting

1. Each period government seeks to procure services for a single project. A project is characterized by a vector of observable characteristics, x_l. For now we assume that all project characteristics are observable. The case with unobserved project heterogeneity can be accommodated but estimation is more complicated.

2. We assume that three types of the firms serve the market: old large firms (OL), old small firms (OS), and new firms (Y).

3. Old large firms have large capacity and, therefore, may concurrently work on several projects. Their costs for a given project, however, depend on capacity utilization (% of capacity which is engaged with other projects), $z_{i,l}$. Old small firm have small capacity and tend to stay out of the market during the period when they are engaged on another project.

4. The distribution of costs for OL firms further depends on firm's productivity, w_i: $F_{C,OL}(.|z_{i,l}, w_i)$. The distribution of costs for OS firms depend on productivity but are influenced by a 'lack of scale advantage' effect: $F_{C,OS}(.|w_i)$.

5. The distribution of costs is similar to the distribution of costs of the old small firms but is further scaled up by the experience effect which in the reasonable long horizon disappears so that the distribution of costs of small firms coincides with the distribution of costs of old small firms. For example, the effect of experience on costs may be exponential, $\mu_{C,Y,k,s} = (1 + a_1 \exp(-a_2 \times s)\mu_{C,OS,k}$.

6. Productivity w_i takes values from a discrete set $\{\bar{w}_1, ..., \bar{w}_K\}$. We assume that the support of the productivity distribution is the same across firms' groups. However, we allow for the frequency of types to differ across groups. The assumption of common support may possibly be extended. One way to deal with this assumption is to focus on the population of new firms that successfully completed at least one project. This should weed out really bad types that had no hope of surviving in the market.

7. We do not track the evolution of the subpopulation of small firms. Specifically, we assume that this population is stable: the distribution of experience and of the productivity conditional on experience is constant and known to all firms participating in this market. The type of the new firm is not known to other firms. At the time of the auction all potential participants observe experience of new firms.

8. Time line for each period (one auction per period):

   i. at the beginning of the period the composition of the set of potential sellers, $M_t = M_t^{OL} \cup M_t^{OS} \cup M_t^{Y}$, is revealed. Specifically, for old firms the composition in terms of productivity is known: $\{\bar{w}_k, |M_t^{OL,k}|\}$, $\{\bar{w}_k, |M_t^{OS,k}|\}$; for new firms the composition in terms of experience is known: $\{s_{i,t} = m, |M_t^{old,m}|\}$.

   ii. Each potential seller draws participation costs, $\varphi_{i,t} \propto F_\varphi$; decides whether to participate or not taking entry cost and information about M_t into account. Participation decision: $d_{i,t} = \{0, 1\}$. Assume that the distribution of entry costs does not explicitly depend on the productivity but it may depend on experience in the case of new firms (!!!)

   iii. Those who decide to participate draw project costs, $c_{i,t}$, from corresponding distributions and decide on the bid level, $b_{i,t}$; the set of participants observed at the time of bidding.

4.2 Equilibrium Characterization

1. Strategies are type specific and depend on a state. We use $\kappa(i)$ to denote the productivity group of firm i and A_t to denote the set of actual bidders participating in auction t.

   a. Old large firms: the participation strategy: $d_{i,t} = \sigma^{P,OL}_{\kappa(i)}(\varphi_{i,t}, z_{i,t}, M_t)$; the bidding strategy: $b_{i,t} = \sigma^{B,OL}_{\kappa(i)}(c_{i,t}, z_{i,t}, A_t)$.

   b. Old small firms: the participation strategy: $d_{i,t} = \sigma^{P,OS}_{\kappa(i)}(\varphi_{i,t}, M_t)$; the bidding strategy: $b_{i,t} = \sigma^{B,OS}_{\kappa(i)}(c_{i,t}, A_t)$.

   c. New firms: $d_{i,t} = \sigma^{P,Y}_{\kappa(i)}(\varphi_{i,t}, M_t; s_{i,t})$, $b_{i,t} = \sigma^{B,Y}_{\kappa(i)}(c_{i,t}, A_t; s_{i,t})$.

4.3 Old Large Firms

2. Notice that the problem of old large firms is stationary. Indeed, since they do not track evolution of individual new firms (Old firms consider individual new firms as a draw from the population of new firms) then their environment remains stationary throughout the sample. [What if is there is seasonality, yearly fixed effects – could we somehow extract them prior to tackling the dynamic part?] We use σ^*, $W^{OL}_{1,\kappa(i)}(.; z_i, z)$, and $W^{OL}_{0,\kappa(i)}(.; z_i, z)$ to denote strategies of other firms, the value function conditional on participation, and the value function conditional on non-participation, correspondingly. The value function

depends on the vector of the backlog for all OL in the population and on the firm's own backlog. This information is in principle subsumed in the description of M. We include it separately to emphasize its role.

$$W^{OL}_{\kappa(i)}(\varphi_i, z_i, z, M; \sigma^*) = \max_{d\in\{0,1\}} E_{A|\{M/i\}}\Bigg\{ d \times (W^{OL}_{1,\kappa(i)}(z_i, z, A; \sigma^*) - \varphi_i) + (1-d) \times W^{OL}_{0,\kappa(i)}(z_i, z, A; \sigma^*)\Bigg\}.$$

Further, Specifically,

$$\begin{aligned}
W^{OL}_{1,\kappa(i)}(z_i, z, A; \sigma^*) &= \int_c \max_b \Bigg\{ (b-c)\Pr(i \text{ wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*) \\
&+\delta \Pr(i \text{ wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*) V^{OL}_{\kappa(i)}(z_i + x, z') \Bigg\} dF^{OL}_{C,\kappa(i)}(c; z_i) \\
&+\delta \sum_{j\in A^{OL}/i} \Pr(j \text{ wins}|b, z_i, z, A, \bar{w}_{\kappa(j)}; \sigma^*) V^{OL}_{\kappa(i)}(z_i, z') \\
&+\delta \sum_{j\in A^{OS}} \Pr(j \text{ wins}|b, z, A, \bar{w}_k; \sigma^*) V^{OL}_{\kappa(i)}(z_i, z) \\
&+\delta \sum_{j\in A^{Y}} \sum_k \Pr(j \text{ wins}|b, s_j, A, \bar{w}_k; \sigma^*) \Pr(\bar{w}_k|s_j, d_j = 1) V^{OL}_{\kappa(i)}(z_i, z);
\end{aligned}$$

and

$$\begin{aligned}
W^{OL}_{0,\kappa(i)}(z_i, z, A; \sigma^*) &= 0 + \delta \sum_{j\in A^{OL}} \Pr(j \text{ wins}|z, A, \bar{w}_{\kappa(j)}; \sigma^*) V^{OL}_{\kappa(i)}(z_i, z') \\
&+\delta \sum_{j\in A^{OS}} \Pr(j \text{ wins}|z, A, \bar{w}_{\kappa(j)}; \sigma^*) V^{OL}_{\kappa(i)}(z_i, z) \\
&+\delta \sum_{j\in A^{Y}} \sum_k \Pr(j \text{ wins}|z, A, \bar{w}_k, s_j; \sigma^*) \Pr(\bar{w}_k|s_j, d_j = 1) V^{OL}_{\kappa(i)}(z_i, z).
\end{aligned}$$

3. **Participation decisions.** Old large firm enters if

$$E_{A|\{M/i\}}\left[W^{OL}_{1,\kappa(i)}(z_i, z, A; \sigma^*)\right] - \varphi_i \geq E_{A|\{M/i\}}\left[W^{OL}_{0,\kappa(i)}(z_i, z, A; \sigma^*)\right].$$

Denote

$$\varsigma_{\kappa(i)}^{OL}(z_i,z,M) = E_{A|\{M/i\}}\left[W_{1,\kappa(i)}^{OL}(z_i,z,A;\sigma^*)\right] - E_{A|\{M/i\}}\left[W_{0,\kappa(i)}^{OL}(z_i,z,A;\sigma^*)\right].$$

Then such a firm enters with probability

$$p_{\kappa(i)}^{new}(z_i,z,M) = \int 1(\varphi_i \leq \varsigma_{\kappa(i)}^{OL}(z_i,z,M))dF_{\varphi}(\varphi_i).$$

The ex-ante value function can then be expressed as (after integrating φ out):

$$V_{\kappa(i)}^{OL}(z_i,z;\sigma^*) = E_M\Bigg\{p_{\kappa(i)}^{OL}(z_i,z,M)(E_{A|\{M/i\}}[W_{1,\kappa(i)}^{OL}(z_i,z,A;\sigma^*)] -$$
$$E[\varphi_i|\varphi_i \leq \varsigma_{\kappa(i)}^{OL}(z_i,z,M)]) + (1-p_{\kappa(i)}^{OL}(z_i,z,M))E_{A|\{M/i\}}[W_{0,\kappa(i)}^{OL}(z_i,z,A;\sigma^*)]\Bigg\}.$$

$E_{\varphi}[\varphi_i|\varphi_i \leq \varsigma_{\kappa(i)}^{OL}(z_i,z,M)]$. Here we exploit distributional assumption that F_{φ} follows exponential distribution.

$$E[\varphi_i|\varphi_i \leq \varsigma_{\kappa(i)}^{OL}] = -\frac{\varsigma_{\kappa(i)}^{OL}}{p_{\kappa(i)}^{OL}}\left[\frac{p_{\kappa(i)}^{OL}}{\log(1-p_{\kappa(i)}^{OL})} + (1-p_{\kappa(i)}^{OL})\right].$$

Here $p_{\kappa(i)}^{OL}$ and $\varsigma_{\kappa(i)}^{OL}$ are short cuts for $p_{\kappa(i)}^{OL}(z_i,z,M)$ and $\varsigma_{\kappa(i)}^{OL}(z_i,z,M)$ correspondingly.

Substituting the expression above and those for W_1 and W_0 into the expression for the value function obtain:

$$V_{\kappa(i)}^{OL}(z_i,z;\sigma^*) = E_M\left\{\left[1+\frac{p_{\kappa(i)}^{OL}}{\log(1-p_{\kappa(i)}^{OL})}\right]E_{A|\{M/i\}}[W_{1,\kappa(i)}^{OL}(z_i,z,A;\sigma^*)] -\right.$$
$$\left.\frac{p_{\kappa(i)}^{OL}}{\log(1-p_{\kappa(i)}^{OL})}E_{A|\{M/i\}}[W_{0,\kappa(i)}^{OL}(z_i,z,A;\sigma^*)]\right\}.$$

4. **Probabilities of Winning.** Three groups of probabilities of winning enter Bellman equation: (a) own probability of winning conditional on own bid, state, and the set of actual participants; (b) other contractor's probability of winning in the auction where the bidder himself participates conditional on own bid, state, and the set of actual participants; (c) other contractor's probability of winning in the auction where the bidder himself does not participate conditional on state, and the set of actual participants. Notice that the proba-

bilities of winning can be expressed using corresponding distributions of bids. Specifically,

$$
\begin{aligned}
&(a)\ \Pr(i \,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}) = \\
&\Pr(b \leq B_j,\, \forall j \in A^{OL}/i) \times \Pr(b \leq B_j,\, \forall j \in A^{OS}) \times \Pr(b \leq B_j,\, \forall j \in A^{Y}) = \\
&\prod_{j \in A/i)^{OL}} (1 - F^{OL}_{C,\kappa(j)}(\beta^{-1}(b; \bar{w}_{\kappa(j)}, z_j, z, A); z_j)) \times \\
&\prod_{j \in A^{OS}} (1 - F^{OS}_{C,\kappa(j)}(\beta^{-1}(b; \bar{w}_{\kappa(j)}, z, A))) \prod_{j \in A^{Y}} (1 - F^{Y}_{C,\kappa(j)}(\beta^{-1}(b; \bar{w}_{\kappa(j)}, s_j, A); s_j));
\end{aligned}
$$

$$
\begin{aligned}
&(b)\ \Pr(j \,\text{wins}|b, z_i, z, A; \sigma^*) = \\
&\int_{\underline{b}}^{b} f^{OL}_{B,\kappa(j)}(x; z_j, z, A) \prod_{p \in A^{OL}/\{i,j\}} (1 - F^{OL}_{C,\kappa(p)}(\beta^{-1}(x; \bar{w}_{\kappa(p)}, z_p, z, A); z_p)) \times \\
&\prod_{p \in A^{OS}} (1 - F^{OS}_{C,\kappa(p)}(\beta^{-1}(x; \bar{w}_{\kappa(p)}, z, A))) \times \\
&\prod_{p \in A^{Y}} (1 - F^{Y}_{C,\kappa(p)}(\beta^{-1}(x; \bar{w}_{\kappa(p)}, s_p, z, A); s_p) dx;
\end{aligned}
$$

$$
\begin{aligned}
&(c)\ \Pr(j \,\text{wins}|z, A; \sigma^*) = \\
&\int_{\underline{b}}^{\bar{b}} f^{OL}_{B,\kappa(j)}(x; z, A) \prod_{p \in A^{OL}/\{i,j\}} (1 - F^{OL}_{C,\kappa(p)}(\beta^{-1}(x; \bar{w}_{\kappa(p)}, z_p, z, A); z_p)) \times \\
&\prod_{p \in A^{OS}} (1 - F^{OS}_{C,\kappa(p)}(\beta^{-1}(x; \bar{w}_{\kappa(p)}, z, A))) \times \\
&\prod_{p \in A^{Y}} (1 - F^{Y}_{C,\kappa(p)}(\beta^{-1}(x; \bar{w}_{\kappa(p)}, s_p, z, A); s_p) dx.
\end{aligned}
$$

The probabilities of winning for the cases when the winners are OS or Y firms should be adjusted in an obvious way.

5. To integrate out the set of actual participants, A,:

$$\Pr(A|M) = \prod_{k,p} C^{|A^{OL}_{k,p}|}_{|M^{OL}_{k,p}|} \left[(p^{OL}_{k,p})^{|A^{OL}_{k,p}|} (1 - p^{OL}_{k,p})^{|M^{OL}_{k,p}| - |A^{OL}_{k,p}|} \right] \times$$
$$\prod_{k} C^{|A^{OS}_{k}|}_{|M^{OS}_{k}|} \left[(p^{OS}_{k})^{|A^{OS}_{k}|} (1 - p^{OS}_{k})^{|M^{OS}_{k}| - |A^{OS}_{k}|} \right] \times$$
$$\prod_{k} \prod_{m} C^{|A^{Y}_{k,m}|}_{|M^{Y}_{k,m}|} \left[(p^{Y}_{k,m})^{|A^{Y}_{k,m}|} (1 - p^{Y}_{k,m})^{|M^{Y}_{k,m}| - |A^{Y}_{k,m}|} \right].$$

Here $|M^{OL}_{k,p}|$ and $|A^{OL}_{k,p}|$ denote cardinalities of the sets of potential and actual bidders with productivity level $\bar{w}_k$ and the backlog level z_p (assume that backlog is discretized) in the case of of old large bidders; whereas $|M^{Y}_{k,m}|$ and $|A^{Y}_{k,m}|$ denote cardinalities of the sets of potential and actual bidders with productivity level $\bar{w}_k$ and the experience level of s_m in the case of new firms.

6. **Bidding Strategies.** F.O.C wrt bid:

$$\frac{d}{db} \Bigg[(b - c) \Pr(i \,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*)$$
$$+\delta \Pr(i \,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*) V^{OL}_{\kappa(i)}(z_i + x, z'); \sigma^*)$$
$$+\delta \sum_{j \in A^{OL}/i} \Pr(j \,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(j)}; \sigma^*) V^{OL}_{\kappa(i)}(z_i, z')$$
$$+\delta \sum_{j \in A^{OS}} \Pr(j \,\text{wins}|b, z, A, \bar{w}_k; \sigma^*) V^{OL}_{\kappa(i)}(z_i, z)$$
$$+\delta \sum_{j \in A^{Y}} \sum_{k} \Pr(j \,\text{wins}|b, s_j, A, \bar{w}_k; \sigma^*) \Pr(\bar{w}_k | s_j, d_j = 1) V^{OL}_{\kappa(i)}(z_i, z) \Bigg] = 0.$$

Completing differentiation:

$$\begin{aligned}
&\Pr(i\,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*) + \\
&(b - c + \delta V^{OL}_{\kappa(i)}(z_i + x, z'); \sigma^*))\frac{d}{db}[\Pr(i\,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*)] + \\
&+\delta \sum_{j\in A^{OL}/i} \frac{d}{db}[\Pr(j\,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(j)}; \sigma^*)]V^{OL}_{\kappa(i)}(z_i, z') \\
&+\delta \sum_{j\in A^{OS}} \frac{d}{db}[\Pr(j\,\text{wins}|b, z, A, \bar{w}_{\kappa(j)}; \sigma^*)]V^{OL}_{\kappa(i)}(z_i, z) \\
&+\delta \sum_{j\in A^{Y}} \frac{d}{db}[\Pr(j\,\text{wins}|b, s_j, A; \sigma^*)]V^{OL}_{\kappa(i)}(z_i, z) = 0.
\end{aligned}$$

Inverse Bid function:

$$\begin{aligned}
&c = (b + \delta V^{OL}_{\kappa(i)}(z_i + x, z'); \sigma^*)) + \\
&\frac{1}{\frac{d}{db}[\Pr(i\,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*)]}\Bigg[\Pr(i\,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(i)}; \sigma^*)] \\
&+\delta \sum_{j\in A^{OL}/i} \frac{d}{db}[\Pr(j\,\text{wins}|b, z_i, z, A, \bar{w}_{\kappa(j)}; \sigma^*)]V^{OL}_{\kappa(i)}(z_i, z') \\
&+\delta \sum_{j\in A^{OS}} \frac{d}{db}[\Pr(j\,\text{wins}|b, z, A, \bar{w}_{\kappa(j)}; \sigma^*)]V^{OL}_{\kappa(i)}(z_i, z) \\
&+\delta \sum_{j\in A^{Y}} \frac{d}{db}[\Pr(j\,\text{wins}|b, s_j, A; \sigma^*)]V^{OL}_{\kappa(i)}(z_i, z)\Bigg].
\end{aligned}$$

7. Value function. Eliminate optimization by substituting an inverse bid function. Notice that the Bellman equation depends on the distributions of bids. The distribution of bids conditional on new firms' type does not enter Bellman equation because old large firm does not observe the type of new firm.

4.4 Old Small Firms

TBA

4.5 New Firms

TBA

5 Estimation Strategy

This section outlines our estimation strategy.

Notice that if the distribution of bids conditional on type and the backlog (for old large) or conditional on type (for old small firms) are known than the value function for each type of OL and OS bidders can be recovered by solving the corresponding Bellman equation. Then, the value function and corresponding bid distribution could be combined to recover the inverse bid function and the distribution of costs conditional on type (and backlog for OL firms). After that the experience-specific multiplier could be recovered by jointly recovering the mixture probabilities and the Bellman equation for the new firms.

1. Decompose the distribution of bids submitted by large old firms and small old firms into the component distributions of bids and mixing probabilities. This could be achieved by using the finite mixtures approach in the following way. We assume that the composition of the set of actual participants is observed when contractors decide on their bids. Therefore, the observed distribution of bids submitted by a large old firm conditional on backlog are given by $F_B^{OL}(.|z_i, A)$. For clarity, assume that A include the large old firm itself and a small old firm. Then, if each of the firms can be one of two types ($\bar{w}_1$ or $\bar{w}_2$) with probabilities ($\pi_{1,L}$, $1-p_{1,L}$) for OL firms and ($\pi_{1,S}$, $1-p_{1,S}$) for small firm then the observed distribution can be represented as

$$F_B^{OL}(b|z_i, A) = \pi_{1,L}\pi_{1,S}F_B^{OL}(b|z_i, A = [1,1]) + \pi_{1,L}(1-\pi_{1,S})F_B^{OL}(b|z_i, A = [1,2]) +$$
$$(1-\pi_{1,L})\pi_{1,S}F_B^{OL}(b|z_i, A = [2,1]) + (1-\pi_{1,L})(1-\pi_{1,S})F_B^{OL}(b|z_i, A = [2,2]).$$

   Here, $A = [k_1, k_2]$ denotes the composition of the set of active bidders in terms of types. Notice that we have a similar equation for the old small firm included in A where we also condition on the backlog of the competitor.

$$F_B^{OS}(b|A) = \pi_{1,L}\pi_{1,S}F_B^{OL}(b|A = [(1,z),1]) + \pi_{1,L}(1-\pi_{1,S})F_B^{OL}(b|A = [(1,z),2]) +$$
$$(1-\pi_{1,L})\pi_{1,S}F_B^{OL}(b|A = [(2,z),1]) + (1-\pi_{1,L})(1-\pi_{1,S})F_B^{OL}(b|A = [(2,z),2]).$$

Utilizing these equations we need to solve for the component distributions $\{F_B^{OL}(b|z_i, A = [k_1, k_2])\}_{k_1,k_2}$, $\{F_B^{OS}(b|A = [(k_1, z), k_2])\}_{k_1,k_2}$, and the mixing probabilities $\pi_{1,L}$ and $\pi_{1,S}$. Recent literature on identification of mixture models (see Bonhomme, Jochman, and Robin (2015), Bonhomme, Jochman, and Robin (2016) and Compiane and Kitamura (2016)) indicate that the component distributions as well as the mixing probabilities can be recovered if the distribution of a vector of repeated observations (given the structure of A) is observed. The identification can additionally be strengthened in several ways.

(a.) The method utilizing pairwise comparisons outlined in Krasnokutskaya, Song, and Tang (2017) can be used to recover the underlying grouping (according to unobserved types) of the set of large old firms. This would eliminate the need to integrate out the type of the old large firm and would reduce the number of objects that have to be identified. That is the equations become

$$F_{B,1}^{OL}(b|z_i, A) = \pi_{1,S} F_B^{OL}(b|z_i, A = [1, 1]) + (1 - \pi_{1,S}) F_{B,1}^{OL}(b|z_i, A = [1, 2])$$
$$F_B^{OS}(b|A) = \pi_{1,S} F_B^{OL}(b|A = [(1, z), 1]) + (1 - \pi_{1,S}) F_B^{OL}(b|A = [(1, z), 2]).$$

In this case only the component distributions $\{F_B^{OL}(b|z_i, A = [\kappa(i), k_2])\}_{k_2}$, $\{F_B^{OS}(b|A = [(\kappa(i), z), k_2])\}_{k_2}$, and the mixing probability $\pi_{1,S}$ would have to be obtained from the equations above.

(b.) The impact of the backlog may be parameterized, so that $F_{C,k}^{OL}(c|z) = F_{C,k}^{OL}(c + m(z))$. In this case, the mixture equation would have to be solved simultaneously with Bellman equation while imposing that $\{F_B^{OL}(b|z_i, A = [\kappa(i), k_2]) = F_{C,\kappa(i)}^{OL}(\beta_{OL}^{-1}(b; z_i, A = [\kappa(i), k_2]))$.

2. Once the component bid distributions for OL and OS firms are obtained then, the Bellman equations obtained above can now be solved for the type-specific value functions of old large and old small bidders. This, in turn, allows us to obtain type-specific inverse bid functions, and thus recover type-specific distributions of costs for the old large and old small bidders, $F_{C,k}^{OL}(c|z)$ and $F_{C,k}^{OS}(c)$.

3. In the last step, we jointly solve for value functions (and thus inverse bid functions which combined with the cost distributions result in the corresponding bid distributions), and the corresponding cost adjustment associated with accumulated experience, as well as mixing probabilities. This step relies on the assumption that the types of firms in the set of the

new firms and the set of the old small firms is the same, and that the distribution of costs of the new firm conditional on type is the same as the distribution of costs of old small firm conditional on type subject to parameterized impact of experience on the mean of the cost distribution. However, the mixing probabilities maybe different among these two sets of bidders. This assumption could be weakened by perhaps assuming the new firms who won and successfully completed one project have the same set of types as old small firms. Then, denoting by $V^{(n)}_{Y,k}(s)$ and $\{F^{(n)}_{Y,B,k}(b|s,A)\}_{s,A}$ the nth iteration of the value function and of the set of possible bid distributions, we solve $F_{Y,B}(b|s,A) = \sum_k F_{Y,B,k}(b|s,A)\pi^Y_k(s)$ while iterating on value function and inverse bid function:

$$
\begin{aligned}
&V^{(n+1)}_{Y,k}(s) = \sum_A g_1(F^{(n)}_{Y,B,k}(b|s,A))\,\Pr(A) \\
&\beta_Y^{-1,(n+1)}(b;A) = g_2(V^{(n)}_{Y,k}(s), F^{(n)}_{Y,B,k}(b|s,A)) \\
&F^{(n+1)}_{Y,B,k}(b|s,A) = F^{OS}_{C,\kappa(i)}((1 + a_1 \exp(-a_2 s) \times \beta_Y^{-1,(n)}(b;A)).
\end{aligned}
$$

Here we need to solve for the parameterized impact of experience on the cost distribution and the mixing weights which are allowed to change with experience level. Since the form of the cost distribution (up to parameterized scaling) is known we can form separate equations for different A as well as by aggregating across A while holding the experience level fixed. The approach used above is not feasible since we do not have sufficient number of observations to estimate the distribution of the vector of bids conditional on (s,A) for different A. The type of new firms is thus integrated out so that only the distribution of bids conditional on experience enters the Bellman equation.

6 Preliminary Findings

In this section we present the results of estimation which allow the cost distribution to change with accumulated experience while ignoring potential unobserved bidder heterogeneity. When there is no unobserved bidder heterogeneity then the distributions of bids (and the probabilities of participation) conditional on state (backlog for OL bidders, and experience for Y bidders) can be directly estimated from the data. These objects can then be pluged into the Bellman equation which allows us to recover the value function conditional on state, inverse bid function and thus the distributions of costs conditional on state following standard methodology similar to the one presented in Jofre-Bonet and Pesendorfer (2003) and Jesiorski and Krasnokutskaya (2016).

The result of this estimation are summarized in Tables 6 and 7. An interesting finding which emerges is that bidders costs indeed change with experience and eventually decline. Further, the average costs of new firms become similar to the costs of OS firms after approximately five projects.

Table 6: Bid Density: no unobserved bidder heterogeneity

	Large	Small	New
Constant (shape)	1.245*** (0.015)	1.236*** (0.013)	1.285*** (0.016)
Constant (scale)	0.240*** (0.005)	0.176*** (0.003)	0.315*** (0.006)
Estimate	-0.017*** (0.004)	-0.008*** (0.003)	-0.022*** (0.004)
Backlog	0.018*** (0.003)	–	–
Potential Bidders:			
# of Large (backlog=0)	-0.047*** (0.002)	-0.023*** (0.002)	-0.038*** (0.001)
# of Large (backlog=1)	-0.031*** (0.002)	-0.018*** (0.002)	-0.019*** (0.001)
# of Small	-0.034*** (0.003)	-0.020*** (0.003)	-0.022*** (0.002)
# of New (exp1-exp3)	-0.009 (0.010)	0.014 (0.010)	-0.019** (0.010)
# of New (exp4-exp5)	-0.019*** (0.002)	-0.010*** (0.003)	-0.034*** (0.003)
Cum.Wins	–	–	-0.007*** (0.002)

Assume bids distributed according to Weibull distribution. Dependent variable is a bid scaled by engineer's estimate. The standard errors are in parenthesis. In the description of potential bidders 'backlog=1' refers to large bidders with backlog above average, 'backlog=0' refers to large bidders with backlog below average. The results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

Table 7: Implied Average Costs by project size and firm types

Engineer's			New				
Estimate	Large	Small	exp=1	exp=2	exp=3	exp=4	exp=5
1	0.763	0.875	0.952	0.971	0.986	0.915	0.848
2	0.752	0.864	0.931	0.944	0.913	0.890	0.875
3	0.754	0.854	0.923	0.975	0.932	0.890	0.866
4	0.743	0.864	0.949	0.979	0.915	0.893	0.860
5	0.732	0.848	0.934	0.932	0.912	0.876	0.837

The cost is normalized by engineer's estimate. The projects are categorized according to the quintiles of engineer's estimate. For large firms the reported average costs are for the mean backlog level. The results in the Table are based of the data from 2002-2011. We will update the statistics to reflect the full dataset at a later date.

References

BONHOMME, S., K. JOCHMAN, AND J.-M. ROBIN (2015): "Estimating Multivariate Latent-Structure Models," *Annals of Statistics*, 44, 540–563.

——— (2016): "Nonparametric Estimation of Finite Mixtures from Repeated Measurements," *Journal of Royal Statistical Society, Series B*, 78, 211–229.

COMPIANE, G., AND Y. KITAMURA (2016): "Using Mixtures in econometric models: a brief review and some new results," *The Econometrics Journal*, 19, C95–C127.

JESIORSKI, P., AND E. KRASNOKUTSKAYA (2016): "Dynamic Procurement Auctions with Subcontracting," *Rand Journal of Economics*, 47(4), 749–899.

JOFRE-BONET, M., AND M. PESENDORFER (2003): "Estimation of a Dynamic Auction Game," *Econometrica*, 71(5), 1443–1489.

KRASNOKUTSKAYA, E., AND K. SEIM (2011): "Bid Preference Programs and Participation in Highway Procurement," *American Economic Review*, 101(6), 2653–2686.

KRASNOKUTSKAYA, E., K. SONG, AND X. TANG (2017): "Uncovering Unobserved Agent Heterogeneity through Pairwise Comparisons," Manuscript.

TIREROVA, L. (2014): "Dynamic Model of Bidder Learning in Procurement Auctions," Dissertation, Johns Hopkins University.