Online Appendix for

Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs

Patrick Bajari* Stephanie Houghton † Steven Tadelis‡

December 2012

The following Appendix provides additional details about the data and estimation procedures used in the paper. Tables A1-A3 summarize the bidding behavior of the top 20 firms in our sample, the number of participants in each auction, and the distribution of these auctions over time. The next two tables present information about our use of instruments in the reduced form regressions of bids on contract characteristics and ex post changes. Specifically, in Table A4, we present F-statistics from first-stage regressions of the instruments. Table A5 demonstrates the robustness of our results by comparing model specifications that instrument for different subsets of the potentially endogenous variables. A final section describes in detail the procedure used to obtain the structural estimates. This section also includes Table A6, which presents results for alternative specifications of the first-stage recovery of the bid distributions using fixed effects and random effects to control for unobserved auction heterogeneity.

*University of Minnesota and NBER
†Texas A&M University
‡UC Berkeley

Table A1: Bidding Activities of Top 20 Firms

ID	No. of Wins	Total Bid for Contracts Awarded	Final Payments on Contracts Awarded	No. of Bids Entered	Participation Rate	Conditional on Bidding for a Contract		
						Average Bid	Average Engineer's Estimate	Average Distance (Miles)
104	160	554,232,998	616,115,118	484	59.1%	3,395,548	3,379,604	133.3
75	15	233,245,265	267,245,145	42	5.1%	10,607,480	10,894,612	81.9
135	7	121,048,703	102,084,697	54	6.6%	13,769,467	13,414,968	441.0
244	23	87,147,853	94,787,509	73	8.9%	3,932,621	3,721,600	58.3
12	34	72,495,433	72,215,257	73	8.9%	2,201,216	2,377,883	31.8
262	24	72,088,982	76,124,748	101	12.3%	2,806,625	2,830,655	215.4
125	16	57,970,813	62,164,914	74	9.0%	3,236,709	3,030,180	81.5
147	1	52,666,668	53,890,666	5	0.6%	19,007,185	19,037,046	86.0
251	23	48,605,745	51,533,241	38	4.6%	1,990,136	2,126,489	46.3
107	21	43,852,728	45,655,279	59	7.2%	2,609,335	2,688,572	53.3
23	17	41,695,376	46,204,955	67	8.2%	3,123,777	2,886,551	67.1
410	1	33,092,725	36,268,057	1	0.1%	33,092,725	28,181,000	141.0
237	22	31,916,930	31,053,539	80	9.8%	2,094,049	2,065,371	69.7
265	4	26,786,493	26,426,965	9	1.1%	7,283,186	7,406,581	234.5
186	17	26,566,823	27,995,110	53	6.5%	1,621,933	1,630,168	48.2
234	6	24,883,692	27,841,209	24	2.9%	2,189,430	2,001,743	166.2
162	17	23,556,856	25,487,495	39	4.8%	1,358,393	1,427,103	61.9
126	8	23,454,933	23,719,853	46	5.6%	1,597,387	1,633,259	69.7
25	2	23,118,363	25,627,033	13	1.6%	4,954,998	4,913,823	44.5
141	13	22,904,644	24,262,589	57	7.0%	2,644,021	2,515,985	61.4

Table A2: Bid Concentration Among Contracts Awarded to Lowest Bidder

Number of Bidders	2	3	4	5	6	7	8	9	10	11+	Total
Contracts in 1999	21	47	36	30	11	8	4	2	3	0	162
Contracts in 2000	30	45	49	43	30	21	6	12	6	7	249
Contracts in 2002	13	13	12	24	14	21	5	4	2	2	110
Contracts in 2003	2	9	6	5	2	1	1	0	0	1	27
Contracts in 2004	21	32	31	19	9	7	4	2	2	0	127
Contracts in 2005	46	38	34	7	8	6	5	0	0	0	144

Table A3: Project Distribution throughout the Year

Month	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
Contracts in 1999	13	11	19	12	18	18	24	20	13	4	8	2
Contracts in 2000	12	14	23	36	16	26	10	39	24	21	20	8
Contracts in 2002	4	8	11	19	24	11	7	2	14	3	7	0
Contracts in 2003	0	0	0	0	0	0	2	8	5	4	2	6
Contracts in 2004	2	8	15	29	33	6	6	10	7	7	3	1
Contracts in 2005	4	10	24	26	23	17	5	6	10	10	5	4

Table A4: First-Stage Results Testing Instrument Quality

Endogenous Variable	First Stage F-Stat
Positive Adjustments	23.38
Negative Adjustments	9.22
Extra Work	4.79
Deductions	5.53
CCDBOverrun	31.96
Number of Observations	3661

The five categories of ex post changes are each normalized by a measure of project size, $\bar{b} \cdot q^{act}$. These suggest that the resident engineer's identity is only strongly correlated with positive adjustments and the dollar overrun on itemized tasks, but weakly correlated with extra work and deductions.

Table A5: Bid Function Regressions Using Actual Quantities Instead of Estimates

Variable	IV.	VII.		V.	VIII.		VI.	IX.	
$DIST_i$	0.0221	0.0217	0.0220	0.0089	0.0089	0.0089	0.0123	0.0119	0.0121
	(5.94)	(5.85)	(5.94)	(3.61)	(3.61)	(3.61)	(4.98)	(5.32)	(5.31)
$RDIST_i$	0.0353	0.0347	0.0355	-0.0013	-0.0013	-0.0013	0.0131	0.0116	0.0123
	(3.41)	(3.24)	(3.43)	(-0.15)	(-0.15)	(-0.15)	(1.83)	(1.57)	(1.69)
$FRINGE_i$	-0.00004	0.0001	0.0004	0.034	0.034	0.034	0.0293	0.0297	0.0297
	(-0.00)	(0.02)	(0.04)	(6.46)	(6.46)	(6.46)	(5.59)	(5.32)	(5.32)
Number of Bidders	0.0058	0.0062	0.0062	0.0024	0.0030	0.0028	-0.0032	-0.0026	-0.0028
	(1.11)	(1.21)	(1.19)	(0.47)	(0.58)	(0.55)	(-0.80)	(-0.73)	(-0.78)
NPosAdj	0.8032	0.8758	0.8815	0.8190	0.9044	0.9166	0.7712	0.9194	0.9319
	(5.89)	(3.87)	(3.92)	(5.85)	(3.82)	(3.89)	(6.39)	(4.67)	(4.76)
NNegAdj	-1.7988	-1.6894	-1.8365	-1.7367	-1.5647	-1.7848	-1.8894	-1.5473	-1.9697
	(-2.23)	(-1.78)	(-2.25)	(-2.24)	(-1.72)	(-2.27)	(-2.25)	(-1.62)	(-2.87)
NEX	0.1644	0.2234	0.1644	0.1647	0.2261	0.1647	0.1559	0.2089	0.1558
	(1.74)	(1.64)	(1.76)	(1.77)	(1.66)	(1.79)	(1.79)	(2.49)	(3.52)
NDED	-1.0231	-1.7460	-0.9932	-1.3268	-2.4838	-1.2893	-0.9580	-2.2187	-0.9337
	(-1.31)	(-0.99)	(-1.28)	(-1.84)	(-1.30)	(-1.80)	(-1.46)	(-1.58)	(-1.22)
NOverrun	0.0057	0.0065	0.0066	0.0059	0.0068	0.0069	0.0054	0.0070	0.0071
	(5.46)	(3.72)	(3.79)	(5.46)	(3.70)	(3.78)	(5.76)	(3.94)	(4.04)
Constant	0.9054	0.8967	0.9015	-0.0518	-0.0628	-0.0564	0.9556	0.9443	0.9480
	(31.74)	(29.73)	(31.30)	(-1.91)	(-2.17)	(-2.06)	(44.80)	(44.59)	(46.82)
Project Effects	None	None	None	Fixed	Fixed	Fixed	Random	Random	Random
Instruments	None	Resident Engineer	Resident Engineer, only for NPosAdj, NOverrun	None	Resident Engineer	Resident Engineer, only for NPosAdj, NOverrun	None	Resident Engineer	Resident Engineer, only for NPosAdj, NOverrun
R^2	0.0738	0.0712	0.0732	0.7621	0.7621		0.0599	0.0577	0.0581
Num. of Obs.	3661	3661	3661	3661	3661	3661	3661	3661	3661

This table reproduces six of the columns from Table 7 in the body of the paper. An additional column has been added for each of the no/fixed/random effects specifications, showing the estimates when we only instrument for NPosAdj and NOverrun where instrument strength is not an issue. Note the similarity in our estimates across specifications. As with Table 7, the dependent variable for all nine regressions is the vector product of the unit price bids and the actual quantities, divided by a measure of the project size ($q^{act} \cdot \bar{b}$). Cluster-robust standard errors are used to compute t-Statistics, shown in parentheses. NOverrun is a measure of the quantity-related overrun on standard contract items, calculated as the vector product of the CCDB prices (where available) and the difference between actual and estimated quantities.

A Details on the Structural Estimation

Our structural approach uses a two-step semiparametric estimator that builds on those discussed in Elyakime, Laffont, Loisel, and Vuong (1994) and Guerre, Perrigne, and Vuong (2000).[1] In the first step, we estimate the density and the CDF of the bid distribution for project n, denoted by $h_j^{(n)}$ and $H_j^{(n)}$ respectively. In the second step, we use those estimates in a GMM estimator based on the first-order conditions in Equation (4). This allows us to recover the adjustment cost coefficients, $\tau_{a+}, \tau_{a-}, \tau_d$, and τ_x, along with a specific form of the penalty from skewed bidding captured by the parameter σ.

Step 1: Estimating Bid Distributions

Because the bidder's payoff function contains expectations of the probability that his bid is the lowest, the first-order conditions will contain the density and CDF of the bid distributions. Specifically, we are interested in an estimate for

$$\left(\sum_{j \neq i} \frac{h_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})}{1 - H_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})} \right)^{-1} \tag{1}$$

for each contract n and each bidder i. As we note in the paper, we cannot recover fully nonparametric estimates of these distributions while still controlling for important measures of firm-specific and auction-specific heterogeneity. Instead we use the following semiparametric approach that "homogenizes" the submitted bids from all firms and all contracts, and uses the homogenized bids to consistently estimate the underlying distribution of firm valuations.

First, we regress the normalized bid on the firm's distance and a fringe indicator, allowing for project-specific random effects:

$$\frac{\mathbf{b}_j^{(n)} \cdot \mathbf{q}^{e,(n)}}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^e} = x_j^{(n)\prime} \mu + u^{(n)} + \varepsilon_j^{(n)}$$

where $x_j^{(n)}$ includes the bidder's distance to the job site and an indicator for whether the bidder is a fringe firm (with less than 1% of the value of contracts awarded). Let $\hat{\varepsilon}_j^{(n)}$ denote the fitted residual from this regression:

$$\hat{\varepsilon}_j^{(n)} = \frac{\mathbf{b}_j^{(n)} \cdot \mathbf{q}^{e,(n)}}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^e} - x_j^{(n)\prime} \hat{\mu} - \hat{u}^{(n)}$$

[1] This approach is detailed in the Athey and Haile (2007) chapter of the *Handbook of Econometrics*, and similar versions are applied in Krasnokutskaya (2011) and Shneyerov (2006).

These residuals are assumed to be iid with distribution $G_{\mathcal{N}}(\cdot)$, where $\mathcal{N}$ indexes the distribution by the number of bidders in contract (n). We can use the empirical distribution of these residuals to recover an estimate for the distribution of bids, since as we show in the paper,

$$H_j^{(n)}(b) \equiv G_{\mathcal{N}}\left(\frac{b}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{e,(n)}} - x_j^{(n)\prime}\mu - u^{(n)}\right)$$

Specifically, in order to construct $\hat{H}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})$, we first compute the empirical CDF of the residuals, $\hat{G_{\mathcal{N}}}(\cdot)$, by pooling all residuals from bids on contracts with the same number of bidders,$\mathcal{N}$, as in contract (n).[2] Then we evaluate this distribution at $\frac{\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)}}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{e,(n)}} - x_j^{(n)\prime}\hat{\mu} - \hat{u}^{(n)}$ to determine the probability that bidder i's bid of $\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)}$, normalized, would be less than his rival bidder j's, normalized bid. Put simply, we count the fraction of the fitted residuals that are less than $\frac{\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)}}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{e,(n)}} - x_j^{(n)\prime}\hat{\mu} - \hat{u}^{(n)}$. This is done for each contract n and each bidder i, for each of bidder i's rivals, indexed by j.

Next, in order to recover the empirical density of the bids, $\hat{H}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})$, we need an estimate of the empirical density of the residuals, $g_{\mathcal{N}}(\cdot)$, where again, $\mathcal{N}$ indexes the density by the number of bidders in contract (n). We use a kernel density estimator, with a normal kernel and a bandwidth determined by Silverman's rule of thumb (a value of 0.0255 for our data).[3] Using a change of variables, we convert this estimated residual density to an estimate of the bid density:

$$\hat{h}_j^{(n)}(b) = \frac{1}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{e,(n)}} \cdot \hat{g}_{\mathcal{N}}\left(\frac{b}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{e,(n)}} - x_j^{(n)\prime}\hat{\mu} - \hat{u}^{(n)}\right)$$

We use the above to calculate $\hat{h}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})$ for each contract n (with number of bidders, $\mathcal{N}$) and each bidder i, for each of bidder i's rivals, indexed by j. Each of the resulting values are then combined with the estimates of the CDF to form the expression in 1.

Note that the estimates of both $\hat{H}_j^{(n)}(b)$ and $\hat{h}_j^{(n)}(b)$ make use of bidder- and project-specific information, as they are evaluated at values that depend on the project's size, $\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{e,(n)}$, rival bidder j's characteristics $x_j^{(n)}$, and the unobserved project heterogeneity, $\hat{u}^{(n)}$. Furthermore, a separate distribution is estimated for each set of $\mathcal{N}$ bidders to account for the fact that, in equilibrium, the distribution of bids will be different in a 2-firm auction

[2]We thank an anonymous referee for reminding us to emphasize that in a first-price auction, the distribution of the mean zero $\varepsilon_j^{(n)}$ will vary by the number of bidders in equilibrium. That is, there is a separate empirical distribution for contracts where n=2 (which we estimate using residuals from 266 bids), n=3 (estimated using residuals from 552 bids), n=4 (671 bids), n=5 (639 bids), n=6 (444 bids), n=7 (448 bids), n=8 (200 bids), n=9 (180 bids), n>=10 (261 bids). We pool contracts with over 10 bidders as there are a limited number of contracts with such large sets of bidders.

[3]Varying this bandwidth slightly did not significantly alter the results.

as compared to a 5-firm auction, since bidders know the number of participants at the time of bidding. These estimated distributions are reasonably precise, drawing upon anywhere from 180 to 671 observed bids. The Matlab code to construct these estimates is available as a supplement to this online appendix.

Step 2: GMM Estimation of the First-Order Conditions

We use the estimates from Step 1 to construct $\left(\sum_{j\neq i} \frac{\hat{h}_j^{(n)}(\mathbf{b}_i^{(n)}\cdot \mathbf{q}^{e,(n)})}{1-\hat{H}_j^{(n)}(\mathbf{b}_i^{(n)}\cdot \mathbf{q}^{e,(n)})}\right)^{-1}$. This term is found in the bidder's first-order condition given in equation (10) of the paper. Following that equation, we can construct the composite error, $\widetilde{e}_i^{(n)}$, as:

$$\begin{aligned}
\widetilde{e}_i^{(n)} = \quad & \frac{1}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{a,(n)}} \left(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{a,(n)} - \sum_{t=1}^{T} \frac{db_t^i(s^i)}{ds^i} q_t^{a,(n)} \left(\sum_{j\neq i} \frac{\hat{h}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})}{1-\hat{H}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})} \right)^{-1} \right) \\
& + \frac{1}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{a,(n)}} \left[(1-\tau_{a+})A_+^{(n)} + (1+\tau_{a-})A_-^{(n)} + (1-\tau_x)X^{(n)} + (1+\tau_d)D^{(n)} \right] \\
& - \frac{1}{\overline{\mathbf{b}}^{(n)} \cdot \mathbf{q}^{a,(n)}} \left[P(\mathbf{b}^{i(n)}) - \sum_{t=1}^{T} \frac{db_t^i(s^i)}{ds^i} \frac{\partial P(\mathbf{b}^{i(n)})}{\partial b_t^i} \left(\sum_{j\neq i} \frac{\hat{h}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})}{1-\hat{H}_j^{(n)}(\mathbf{b}_i^{(n)} \cdot \mathbf{q}^{e,(n)})} \right)^{-1} \right]
\end{aligned}$$

where we parameterize $P(\mathbf{b}^{i(n)})$ as

$$P(\mathbf{b}^{i(n)}) = \sigma \sum_{t=1}^{T} \left(b_t^i - \overline{b}_t\right)^2 \left| \frac{q_t^e - q_t^a}{q_t^e} \right|$$

We form the moment condition

$$m_N(\sigma, \tau_{a+}, \tau_{a-}, \tau_d, \tau_x, \hat{h}, \hat{H}) = \frac{1}{N} \sum_n \sum_i \widetilde{e}_i^{(n)}(\sigma, \tau_{a+}, \tau_{a-}, \tau_d, \tau_x, \hat{h}, \hat{H})(z_i^{(n)} - \overline{z}_i^{(n)})$$

where the instruments, $z_i^{(n)}$ include the engineer's estimate, a full set of dummy variables for the resident engineer assigned to the project, and (in some specifications) month and district dummy variables. We use a non-linear least squares optimization algorithm (Matlab's `lsqnonlin`) to minimize the objective function $m_N' W m_N$, where W is a positive semi-definite weighting matrix. We first estimate $(\widehat{\sigma}, \widehat{\tau_{a+}}, \widehat{\tau_{a-}}, \widehat{\tau}_d, \widehat{\tau}_x)$ using the identity weighting matrix, then use those estimates to construct the optimal weighting matrix as the inverse of the sample variance of m_N.

Table A6: Structural Estimation, using Alternative Specifications for the First-Stage Recovery of the Bid Distributions

	I-A	I-B	II-A	II-B	III-A	III-B	IV-A	IV-B
Implied Marginal Transaction Costs								
Positive Adjustments *(τ^{A+})*	4.759	2.203	4.192	2.224	4.523	2.132	4.557	2.236
	(4.032)	(0.409)	(2.121)	(0.373)	(4.015)	(0.401)	(2.113)	(0.379)
Negative Adjustments *(τ^{A+})*	-0.994	0.305	-15.145	4.802	-0.268	0.743	-1.863	2.758
	(53.680)	(3.156)	(33.565)	(4.323)	(53.686)	(3.08)	(33.572)	(3.015)
Extra Work *(τ^{X})* *	1.091	1.084	2.449	1.233	1.079	1.076	2.209	1.227
	(2.708)	(0.152)	(1.470)	(0.203)	(2.708)	(0.154)	(1.469)	(0.198)
Deductions *(τ^{D})*	9.069	0.556	10.574	2.881	7.200	0.033	4.246	1.478
	(77.845)	(4.878)	(39.989)	(3.860)	(77.904)	(4.844)	(40.035)	(3.601)
Skewing Parameter								
Penalty *(σ)*	-4.699E-05	-1.225E-05	-4.235E-05	-1.309E-05	-4.35E-05	-1.11E-05	-4.84E-05	-1.22E-05
	(5.69E-05)	(1.00E-05)	(3.38E-05)	(9.10E-06)	(5.56E-05)	(9.49E-06)	(3.31E-05)	(1.04E-05)
Number of Obs	3661	3661	3661	3661	3661	3661	3661	3661
First-Stage Bid Distribution**	Contract Fixed Effects	Contract Fixed Effects	Contract Fixed Effects	Contract Fixed Effects	Contract Random Effects	Contract Random Effects	Contract Random Effects	Contract Random Effects
Weighting Matrix***	Identity	Optimal	Identity	Optimal	Identity	Optimal	Identity	Optimal
Instruments Used in Second Stage GMM		Resident Engineer, Engineer's Estimate		Resident Engineer, Engineer's Estimate, Month and District Dummies		Resident Engineer, Engineer's Estimate		Resident Engineer, Engineer's Estimate, Month and District Dummies

* These estimates represent an upper bound on transaction costs associated with changes in scope. They do not account for marginal costs associated with performing the extra work, which for a reasonable profit margin of 20 percent would lower our estimate by $0.80.

** To recover the bid distribution from which the moment conditions (based on the first-order conditions) are formed, we obtain residuals from a first stage regression of bids on contract and bidder characteristics. This "homogenizes" the data by controlling for contract- and bidder-specific characteristics. In Columns I-A, I-B, II-A, and II-B, the first-stage regression includes bidder distance, fringe status, and a contract fixed effect. In Columns III-A, III-B, IV-A, and IV-B, the first-stage regression includes bidder distance, fringe status, the number of bidders, and a contract random effect (the fixed effect approach did not include the number of bidders as it would have been fully absorbed by the fixed effect. In both cases, the residuals for all 3661 bids (819 contracts) were then pooled in order to recover the bidding distribution from which bidders would form their expectations of winning. This differs from the approach in the paper – where separate distributions are recovered for each set of contracts with the same number of bidders – but the resulting estimates are very similar. We prefer the indexing approach used in the paper, as it trades off a higher variance (fewer observations used to construct each distribution) in favor of unbiasedness.

*** Consistent GMM estimates were computed using the identity matrix as the weighting matrix. In a second step, efficient GMM estimates were computed using the optimal weighting matrix derived from the variance of the sample moments in the first step. Standard errors appear in parentheses.