Working Paper No. 17-05

Heterogeneous uncertainty in procurement auctions with unknown competitors: a structural estimation of heteroskedasticity*

Seamus X. Li
University of Colorado Boulder

November 23, 2017

Department of Economics

University of Colorado Boulder
Boulder, Colorado 80309

Heterogeneous uncertainty in procurement auctions with unknown competitors: a structural estimation of heteroskedasticity

Seamus X. Li[y]
University of Colorado Boulder

November 23, 2017

Download the latest version at
seamusx.net/research/job-market-paper

Abstract

While economists often treat heteroskedasticity as a statistical technicality, heteroskedastic outcomes in certain market settings can arise out of heterogeneous uncertainty. Procurement auction, in particular, poses two sources of uncertainty because it exhibits both private-value and common-value characteristics. When rms cannot observe the number of bidders, a third dimension of uncertainty is added, and the e ect of asymmetric information readily emerges in such highly uncertain environments. By exploiting the heteroskedasticity of normalized bids with respect to rm size in highway procurement auctions, I estimate the structural parameters of both uncertainty and its heterogeneity within a semiparametric generalized method of moments framework. The estimation results allow further analyses of rm behavior and auction design through calibration and counterfactuals. In addition, the paper shows that structural parameters can be extracted from heteroskedasticity under fairly simple assumptions, and the method may be extended to the study of other market settings with heteroskedastic outcomes.

JEL Classi cations: C57, D44, H74, R42

I would like to thank my dissertation adviser, Professor Daniel Ka ne, for his continued encouragement and guidance, as well as Professors Chrystie Burr, Oleg Baranov, and Scott Savage for their valuable feedback. I am especially grateful to Dr. Ermias Weldemicael for a fruitful internship at the Colorado Department of Transportation, without which this paper would not have been possible.

[y]Contact: Department of Economics, 256 University of Colorado Boulder, Boulder, CO 80309-0256. Email: seamus.li@colorado.edu

1 Introduction

Heteroskedasticity is a well-studied problem in statistics and econometrics. It is often seen as a nuisance in statistical inference, as it can lead to inefficient and biased standard error estimates for parameters and must be accounted for in robust analyses. Estimation models usually assume that the heteroskedasticity exists in the error term, which is a reasonable abstraction as heteroskedasticity is commonly found in variables of differing scales, where groups of outcomes greater in average magnitude also experience greater variation in realized values. There is a large body of literature and many widely adopted methods regarding heteroskedasticity such that it is almost second nature for researchers to apply some type

(CDOT) with relatively rich details, I estimate the structural parameters of both uncertainty and its heterogeneity with respect to firm size, which enable further analyses of both firm behavior and auction design through calibration and counterfactuals.

Clearly, not all heteroskedasticity indicates heterogeneous uncertainty, such as in the case of the magnitude of the outcome variable, and it is important to distinguish between heterogeneous uncertainty and other types of heterogeneity in terms of agents' preferences, costs, and constraints. In the CDOT data, heteroskedastitcity of bidding behavior persists after project value has been normalized and magnitude is no longer a source of conditional variance. However, heteroskedasticity of bidder behavior exhibits no apparent change in the mean with respect to firm size, contrary to the conventional wisdom of increasing returns to scale at firm level. Since heteroskedasticity is a data variation specific in the second moment, incorporating differing beliefs about the variance of private value into a standard first-price auction model lends a plausible explanation for these observations.

I estimate a structural model within a general method of moments (GMM, Hansen, 1982) framework, which affords the ability to specify important assumptions in the second moment to aid identification. To reduce computational complexity, I adopt a two-stage process where bidders' private values are first estimated semiparametrically, and the main parameters of interest are then estimated with nonlinear GMM. As a proxy for firm size, I use the total number of bids by unique firms in the sample period, which is also a good measure of incumbency, another source of asymmetric information. Because firms cannot observe the number of bidders *ex ante*, a problem with both endogeneity and simultaneity arises, which I address with instrumental variables of project value and type that satisfy the exclusion restriction with normalized bids. To account for market factors and potential issues with temporal autocorrelation outside of a panel or time-series framework, I control for additional variations using contemporaneous and lagged construction permit data in Colorado.

I find that small firms face significantly greater uncertainty in private value and, to a lesser extent, in common value as well. Calibration analysis show that that firms generally anticipate the number of bidders well from public signals despite not observing it directly, although smaller firms more often overestimate the amount of competition. Through either analytical or computational intractibility, often due to probabilistically constructed treatment for unobserved variables, such as losing bids. More recent literature makes heavy use of nonparametric methods for identi cation of private value distribution under various types of auction setting and data restrictions, following the seminal work of Guerre et al. (2000)[4]. Within such literature, public project procurement, in particular highway construction procurement with higher value projects and more regulated bidding procedures, has proven fertile ground for auction analysis and provides useful precedent for this paper. As a matter of public records, procurement auction data tend to be more accessible, if not more complete, which partly facilitates the study of bidder heterogeneity both in terms of private value[5] and bidding behavior[6].

Understanding heterogeneous uncertainty can be important to various microeconomic applications. In auctions, speci cally, one can no longer rely on revenue equivalence to expect similar revenue or expenditure outcome when information is asymmetric and uncertainty is heterogeneous and therefore bidding outcome varies based on design. Government agencies spend a signi cant portion of their resources on private contractors to provide a myriad of goods and services. While the methods of procurement vary, open market contract bidding is often a preferred mechanism that has several advantages, such as transparency, avoidance of favoritism and nepotism, competitive pricing, and a selection of quality[7] The government also supports tax payer, citizen, and community interests such as minimizing expenditure and expanding access to disadvantaged businesses[8]. Having a structural understanding of the dispersion of uncertainty among di erent business partners can inform the assesement of performance in achieving these goals. Given the breakdown of revenue equivalence, having a measure of heterogeneous uncertainty can also aid in optimizing procurement design to better achieve both expenditure and a rmative objectives[9].

---

et al. (1995), etc..

[4]Notable additional works and extensions of nonparametric identi cation include Elyakime et al. (1994); Athey and Haile (2002); Fevrier (2008); Henderson et al. (2012); Armstrong (2013), etc..

[5]Krasnokutskaya (2011) and Armstrong (2013) both investigate the identi cation of private value under unobserved heterogeneity with Michigan highway procurement data..

[6]De Silva et al. (2003) nds that incumbent tend to bid more aggressively (lower) in Oklahoma highway procurement auctions.

[7]Bajari et al. (2008) provide some empirical comparison between auction and negotiation in procurement and suggest some drawbacks of procurement auction despite its popularity.

[8]Nakabayashi (2013) investigates the e ect and e cacy of small business set aside in public construction projects in Japan and found that while many business would not participate without the set aside, it also increases government cost due to reduced competition. CDOT does not have a speci c small business carve out; instead, it takes a rmative action toward disadvantaged businesses through its Disadvantaged Business Enterprise Program (Colorado Department of Transportation).

[9]In procurement auction analysis by civil engineers and nancial planners, bid spread is often of particularly interest, though it is often done in a descriptive manner (Skitmore et al.

> Potential contractors submit sealed bids with itemized cost information. The bidders are unable to observe the identities, the number, or the bid amounts of other bidders before the winner is announced.
>
> The bids are compared to an engineer's estimate produced internally with engineering and market assumptions. The engineer's estimate is also sealed at the time of the bid letting. The lowest bidder usually wins, provided that the submission is deemed feasible, adequate, and does not unreasonably deviate from the engineer's estimate in either direction[11].
>
> Once a winner is announced, the engineer's estimate and all bids, including each bidder's itemized cost, are announced publicly.

A few straightforward observations can be made about this bidding procedure. First, the format is a variation of the first-price sealed-bid auction, but with a common value component in the form of engineer's estimate that is opaque to bidders. Second, bidders are shielded from the number and the identity of other bidders, which adds additional uncertainty. Conversely, past bidding and cost statistics are published in great detail as a matter of transparency and public accountability, which means that firms may use this information to reduce uncertainties in this highly uncertain bidding format.

In addition, CDOT takes affirmative action toward small businesses and disadvantaged businesses (those owned by minorities, women, and other socially and economically disadvantaged individuals) through various programs and services, and the agency has an interest in ensuring that these business have access to its projects and are represented.

2.2 Summary statistics and descriptive analysis

Projects range from tens of thousands to tens of millions of dollars and it presents several statistical problems, such as difficulty of comparison, very large heteroskedasticity, and uncertain latent private value estimation. Normalizing bids by the engineers' estimate could solve the problem if bidding behavior in ratio terms is not influenced by project size, and descriptive analysis shows that it appears not. In fact, the bid-to-estimate ratio over the years exhibit a very consistent and well-behaved log-normal distribution (Figure 1):

The log-normal distribution of bid-to-estimate ratio (relative bid) further suggests that the bid generating process for individual bidders follows a Cobb-Douglas form, as log-normal

Table 1: Summary statistics.

	Observations	Average	Median	St.d.	Max	Min
Engineer's estimates (USD)	1439	3,253,806	1,533,720	5,359,410	57,418,152	47,780[a]
Relative bid	6230	1.1042	1.0674	0.229086	4.1111	0.4421
Number of bidders	1439	4.334	4	2.165909	15	1
Firm size[b]	353	17.65	3	32.58243	207	1
Monthly market (USD)[c]	137	454,702	431,882	202,433	913,024	113,926

	All	Valid[d]		Firm Size	# Firms	# Bids
Observations	6252	5763		1 - 25	273	1200
Auctions	1439	1313		26-50	23	821
Unique rms	353	339		51-75	16	921
Months	164	125		76-100	13	1068
Sample period start	05/2004	01/2006		101-150	8	847
Sample period end	05/2018	05/2016		>		

Figure 1: Kernel density estimate of bid distributions by size cohort.

distribution describes the product of random variables of certain attributes. In addition,

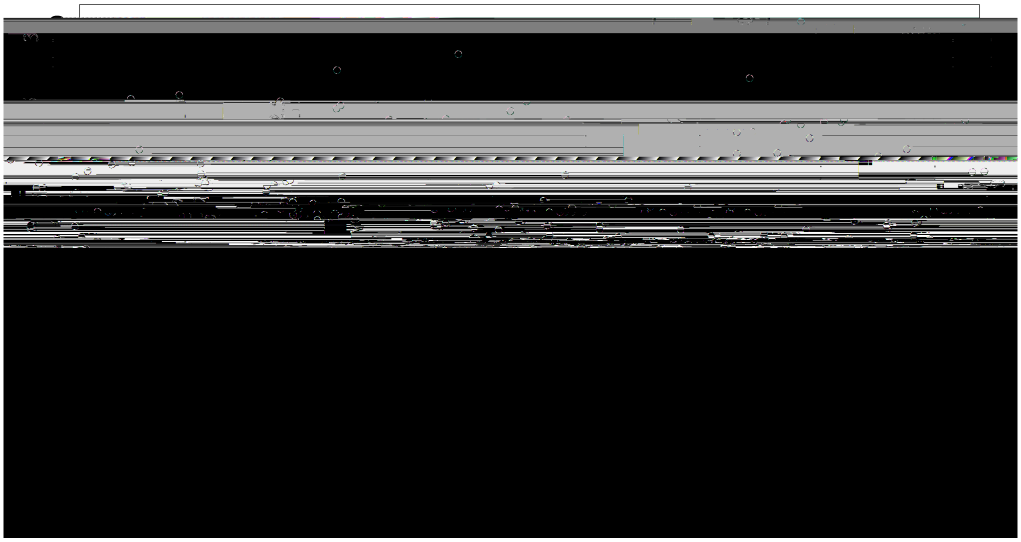


Figure 2: Relative bid spread by annual bidding size cohort. Red line denotes tted value to size. Color spectrum denotes distribution of logged relative bid.

Figure 3: Quarterly average of 1st, 2nd, and 3rd relative bid and number of bidders.

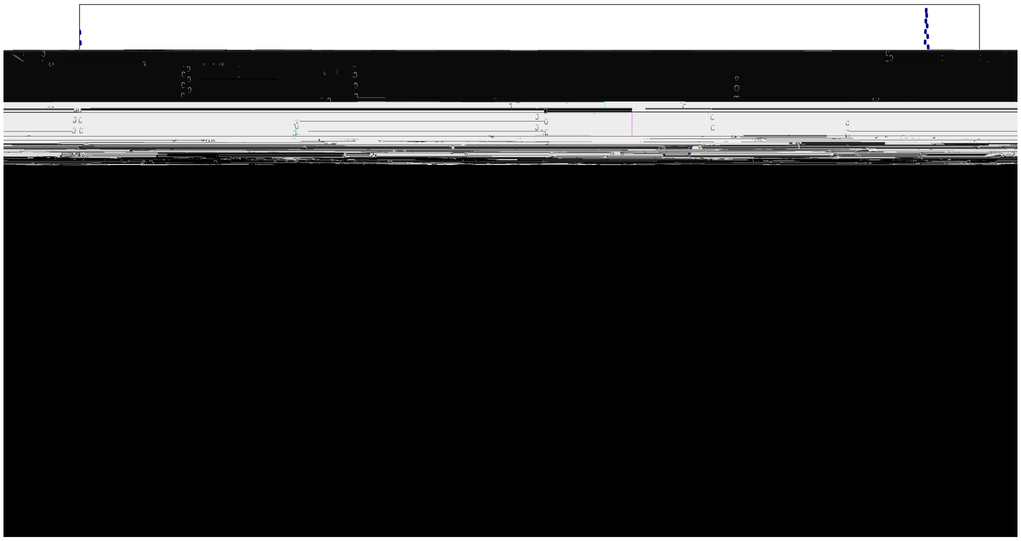
Table 2: Descriptive OLS coefficients of determination (R^2).

Variable	Firm Size	Number of bidders	Relative bid
Firm size	-	0.0000	0.0005
Project type	0.0389	0.1494	0.0226
Project value	0.0080	0.0036	0.0067
Number of bidders	0.0000	-	0.0065

value and project type (Table 2). In addition, there appears to be little linear relationship among firm size, project type, project value, relative bid, and number of bidders, suggesting that entry by firms of different sizes is not particularly predicated on project type and project value. Echoing Figure 2, firm size is a particularly poor linear predictor for bid outcome. However, the inertness of relative bids to seemingly influential factors suggests that the underlying data generating process is stable and well-behaved and the relative bid construction may be a good normalization technique to study bidding behavior between projects of differing scale.

One confounding result is that the relative bids exhibit classical auction theory behavior with regard to number of bidders, despite that the bidders are not able to observe it. The co-movement of quarterly average 1st, 2nd, and 3rd bids (Figure 3) sheds light on this question. If bidding behavior exhibit temporal synchronicity, it suggests that it is influenced by market forces, which affect both entry and private value. If outside market offers good opportunities, a resource-constrained firm faces a higher opportunity cost of entering the highway bidding market, which would raise the firm's private value and inhibit entry.

Indeed, the relative bid proves highly sensitive to market conditions, as shown in figure 4, where the an ARIMA model anticipates shocks well with construction market indicators[12] (lagged monthly values), and the selected market variables prove to be a good predictor of bidding behavior. This offers an important insight to the effect of how exogenous shocks and unknown number of bidders should be treated in the empirical analysis.



distributed among is and js while independently distributed among ks. Assume also that the benchmark value of the project, v_j is generated by a similar process,

$$v_j = \prod_k X_{j;k}^{k} \tag{2}$$

where $X_{j;k}$ is similarly distributed as $X_{i;j;k}$. The Cobb-Douglas function itself is irrelevant to subsequent modeling. However, it has two important implications. First, since $X_{;k}$s are independently distributed, $v_{i;j}$ is log-normally distributed. This is a result of the Central Limit Theorem such that the product of independent random variables has a log-normal distribution. Second, because v_k is similarly distributed as $v_{i;j}$, the relative private value

$$r_{i;j} = \frac{v_{i;j}}{v_j}$$

without necessarily specifying any distributional parameters. They also have a rm-speci c approximation of, or con dence in, the variance, $_i$, where $_i > 0$. Note that because $_j$ is the normal counterpart of the mean of log-normal random variable $r_{i;j}$, and

$$E[r_{i;j}] = e^{\ _j + \frac{^2}{2}} \neq e^{\ _j} \qquad (4)$$

As such, $_i$ can considered the parameter of heterogeneous private value uncertainty, and $_{i;j}$ a measure of common value in,2

Figure 5: Numerical results of $B(r_{i;j})$ response to various changes in parameters.

standard rst-price auction model then yields the optimal bidding function[21, 22]

$$B(r_{i;j}) = r_{i;j} + \frac{\int_{\ln r_{i;j}}^{\infty} [1 - \Phi(\frac{\ln x - \mu_{i;j}}{\sigma_i})]^{n_j - 1} dx}{[1 - \Phi(\frac{\ln r_{i;j} - \mu_{i;j}}{\sigma_i})]^{n_j - 1}} \qquad (7)$$

Figure 5 shows the the responses of $B(r_{i;j})$ to various changes in parameters. Note that $\mu_{i;j}$ and σ_i are parameters of the normal distribution from log relative values and are lower in magnitude compared to $r_{i;j}$. It is a necessary result that $\frac{\partial B()}{\partial r_{i;j}} > 0$ as increasing monotonicity of $B()$ in $r_{i;j}$ is a requirement for 7c127(t)-247(for)-05 -1.7t1.49 0 i707)

Figure 6: Simulated results of bid spread kernel density estimates with di ering uncertainty from the same distribution of private values.

when holding belief in mean relative value constant, a higher spread attens the distribution with a longer right tail and improves the probabilistic standing the rm, hence the rm bids more con dently. As a result, for any given and , higher demonstrate a need to identify, separate, and parameterize these two opposing effects of uncertainty in the estimation strategy through a structural approach.

4 Estimation Strategy

There are several challenges to the identification of the structural model. First, given the highly nonlinear, algebraicly intractable form of the behavioral solution, the estimation equation must be structured in a manner that ensures identification. Second, as discussed in the Data section, there exists a high degree of endogeneity and simultaneity between the relative bids and the number of bidders, which is not observable to the firms *ex ante*. Finally, the same nonlinearity and intractability, along with the number of observations and estimation parameters, imposes a large numerical complexity, and care must be taken to reduce the computational expense. To address these issues, I adopt a generalized method of moments (GMM) framework that incorporates instrumental variables and nonparametric techniques.

4.1 The structural model

4.1.1 Estimation equation

The relative private value $r_{i;j;t}$ is unobserved, but it can be modeled as a latent variable dependent on manifest variables. Following Lafront et al. (1995), I assume that the firm's reservation valuation is determined by the function [24]

$$r_{i;j;t} = e^{\alpha_i + M_t^{0} B_M} \quad (8)$$

Where α_i is the firm fixed effect and M_t is the vector of market factors. Differing from Lafront et al., however, is that the structure does not include any firm characteristics, such as firm size, as explanatory variables of private value. Given the focus of identifying the dep14(size,)sdeniisp48359tuct-423submitt28 Td 438(ide438bid)423(m)78bpproauot obs 438(side438bid)t-4 outside opportunities. In this sense, the estimated fixed effect may not necessarily reflect the firm-specific cost of construction alone. This simplifies the estimation procedure such that unobserved heterogeneity in bidding decision need not be addressed. The private values are estimated apart from the main estimation equation semiparametrically and the method is described in section 4.3 Implementation.

The structural model derives directly from the behavioral framework. Given the construction of the optimal response function, generalized method of moments is used to estimate the structural model below:

$$y_{i;j;t} = r_{i;j;t} + \underbrace{\frac{\int_{\ln r_{i;j;t}}^{1} \left[1 \quad \frac{x \quad s_i \quad {}_{j;t}}{{}_r s_i} \right]^{n_j \quad 1} dx}{\left[1 \quad \frac{\ln r_{i;j;t} \quad s_i \quad {}_{j;t}}{{}_r s_i} \right]^{n_j \quad 1}}}_{b_{i;j;t}} + "_{i;j} \qquad (9)$$

where $_r = [(I \quad 1)^{\ 1} \sum_i \quad _i]^{\frac{1}{2}}$ is the standard deviation, and $_r = I^{\ 1} \sum_i \quad _i$ the mean, of private values calculated from the fixed effect estimates. The structural equation also presents the two other main parameters of interest: private value heterogeneous uncertainty parameter and common value heterogeneous uncertainty parameter with regard to firm size. Contrasting the optimal bidding function, $_r s_i$ substitutes for $_i$ and $s_i \quad _{j;t}$ substitutes for $_{i;j}$. A negative would support the hypothesis that smaller firms have normalization, because the magnitude of relative bid is still affected by project value through the number of bidders, although the project value is no longer correlated with the error term. Second, as discussed in the Data section, there is a strong simultaneity between the number of bidders and bidding behavior based on market conditions, which causes the same issues as endogeneity in estimation. Although it is common in empirical auction studies to assume that the number of bidders is known, such assumption in the presence of both endogeneity and simultaneity will cause the estimators to be biased, and while the number of bidders does not require a parametric estimator itself as the exponent of the survival function, it will attenuate the estimation of other parameters of bidding behavior in the nonlinear model as the observed number of bidders strongly correlates with bid markup beyond its actual effect.

Instrumental variable is an obvious strategy to address this issue. Assuming a Poisson data generating process for the number of bidders with an exponential link function:[25]:

$$E[n|X_{IV}] =$$

The model also partially abstracts from endogenous entry with respect to firm size except for the correlations picked up by the covariates in the structural model. This is justified by the observations from Table 2 that there is little pairwise linear relationship among firm size, project type, and project value. The limitation of this abstraction is briefly discussed in the Conclusion.

4.2 Generalized Method of Moments

The GMM estimator is chosen due to its ability to specify an important second moment assumption that is discussed in subsections 4.2.1 and 4.2.2. At minimum, the GMM estimator requires the first moment conditiosn that $E[\varepsilon_{i;j}|W;\theta_0] = 0$, where $W \in R^{K+1}$ contains 1 containing dependent variable Y, explanatory variables X, and additional instrumental variables, with θ_0 being the vector of estimation parameters θ at their true value. The error term in the nonlinear structural equation is assumed to be additive in y, and the error term is therefore simply $\varepsilon_{i;j} = y_{i;j;t} - b_{i;j;t}$, on which the moment conditions are defined in the following subsection, and because W only fully appear in $\varepsilon_{i;j}$, we define Z as the vector of explanatory and instrumental variables for other constituent expressions in the moment conditions.

4.2.1 Moment conditions

All moment conditions are constructed around the usual assumption that the vector of functions of Z, $h(Z_{i;j;t};\theta)$, is independent from the error term $\varepsilon_{i;j} = y_{i;j;t} - b_{i;j;t}$ such that

$$E[(y_{i;j;t} - b_{i;j;t})^k | h(Z_{i;j;t};\theta);\theta_0] = E[(y_{i;j;t} - b_{i;j;t})^k] = \mu_{\varepsilon;k} \tag{11}$$

where θ_0 is the true value of the the parameters and $\mu_{\varepsilon;k}$ is the kth central moment of $\varepsilon_{i;j}$. This leads to

$$E[h(Z_{i;j;t};\theta)(y_{i;j;t} - b_{i;j;t})^k | \theta_0] = h(Z_{i;j;t};\theta)\mu_{\varepsilon;k} \tag{12}$$

For the estimation, conditions for the first three moments are used:

$$g(W_{i;j;t};\theta_0) = E\begin{bmatrix} h(Z_{i;j;t};\theta)(y_{i;j;t} - b_{i;j;t}) \\ h(Z_{i;j;t};\theta)\{\sigma_y^2(s_i) - (y_{i;j;t} - b_{i;j;t})^2 - [b_{i;j;t} - \mu_y(s_i)]^2\} \\ h(Z_{i;j;t};\theta)(y_{i;j;t} - b_{i;j;t})^3 \end{bmatrix}_{\theta_0} = 0 \tag{13}$$

The first moment conditions are conventionally defined to assume that the error term has zero mean and independent from Z. While the first moment conditions are usually sufficient for many econometric problems, the structural model requires some higher moments be defined as well to achieve identification. Most importantly, the first moment conditions alone do not account for any potential heteroskedasticity with respect to firm size in the error term (*Proposition 2*[26]).

The problem is resolved in the second moment conditions, where $\sigma_y(s_i)$ is the conditional variance of y on firm sized s_i, and it is derived from the fact that $\sigma^2_y(s_i) = \sigma^2_b(s_i) + \sigma^2_\varepsilon$ under the assumption that distributions of private value and error term are independent from each other, which results in additive variance of its constituent variables. It also relies on the assumption that heteroskedasticity exists in the dependent variable through heterogeneous uncertainties in the bidding function, but not in the error term, at least not with regard to firm size. This is a novel assumption based on the structural model, see section 4.2.2 Heteroskedasticity for more discussion.

The third moment condition assumes that that residuals are symmetrically distributed. While b has an appearance of log-normal distribution with a clear skewness, the random noise after the optimal bidding strategy based on the log-normally distributed private value is accounted for is assumed to be symmetrically distributed around 0.

$h(Z_{i;j;t};\theta)$ can be a vector of any functions of $Z_{i;j;t}$ to the extent that the model can still be identified, including simply the vector $Z_{i;j;t}$. Therefore $g(W_{i;j;t};\theta)$ is a $3 \times p$ matrix where p is the number of parameters. To obtain optimal estimators of θ in a nonlinear GMM model, the vector of functions of explanatory and instrumental variables takes a certain form of the gradient of the optimal bidding function b:

$$h(Z_{i;j;t};\theta) = \frac{\nabla_\theta b(\theta \mid Z_{i;j;t})}{\sigma_\varepsilon(Z_{i;j;t})} = \frac{\partial b(Z_{i;j;t};\theta)/\partial\theta}{\sigma_\varepsilon(Z_{i;j;t})} \tag{14}$$

Where $\sigma_\varepsilon(W_{i;j;t})$ is the heteroskedastic error dependent on $W_{i;j;t}$ of an unknown form. While there are methods to approximate $\sigma_\varepsilon(W_{i;j;t})$, it is not necessary as the heteroskedasticity with respect to firm size is specially treated (see the following subsection) while the model abstracts from other sources of heteroskedasticity and, if present, uses a heteroskedasticity-consistent model. The optimal h(9 5 5gs82 1 1 . 9552 Tf244 0 6he nc0:

(Z_{i;j;t}; r b(jZ_{i;j;t} (Z_{i;j;t};

classes of parameters, see section A.3 First-order derivatives[27].

4.2.2 Heteroskedasticity

In addition to the second central moment assumption $E[(y_{i;j;t} - b_{i;j;t})^2 | h(Z_{i;j;t}; \theta); \theta_0] = Z\sigma^2_{\varepsilon}$, the second moment conditions also rely on two additional assumptions that

$$V[y_{i;j;t}|s_i] = E[(y - \mu_y(s_i))^2] = E[(y - \mu_y(s_i))^2 | h(Z_{i;j;t}; \theta); \theta_0] = \sigma^2_y(s_i) \tag{16}$$

where $\mu_y(s_i) = E[y_{i;j;t}|s_i] = E[b^{IV}_{i;j;t}] + E[\varepsilon_{i;j}] = E[b^{IV}_{i;j;t}]$, and

$$E[b_{i;j} | h(Z_{i;j;t}; \theta); \theta_0] = \sigma^2_b(Z) \tag{17}$$

where σ^2_b takes the form of $[\tilde{b}_{i;j;t} - \mu_y(s_i)]^2$ as the fitted value $\hat{b}_{i;j;t}$ is correlated with the error term under instrumental regression. Instead, the fitted $\tilde{b}_{i;j;t}$ uses the fitted values $\tilde{n}_{i;j;t}$ of n_j using Poisson regression against Z as described in section 4.2.2 Heteroskedasticity, similar to the first-stage estimation in 2SLS. The fitted $\tilde{n}_{i;j;t}$ can be considered a combined signal of number of bidders observable to both firms and the investigator.

Because the data is not a random sample, $\sigma^2_y(s_i)$ is assumed to be the sub-population variance of all sub-population observations in the data, and it takes the form of and $\sigma^2_y(s_i) = \frac{1}{N}\sum(y - \bar{y}_{s_i})^2$ (*cf.* sample variance estimator $\hat{\sigma}^2_y(s_i) = \frac{1}{N-1}\sum(y - \bar{y}_{s_i})^2$). The population variance assumption is a strong but defensible one for want of a means of incorporating this estimation into the structural estimation itself. This assumption allows us to estimate sub-population variance directly without changing the structural estimation while still maintaining a higher level of generality than assuming a known private value distribution. Properties of the estimator with a random sample is worth exploring in further studies.

The firm-size-dependent outcome variance $\sigma^2_y(s_i)$ can be estimated in three ways: simple cohort sub-population variance, parametric fit, and non-parametric fit. The simple cohort sub-population variance can be calculated by the equation above. However, due to that sub-population variance depends on the initial random draw from the distribution of $r_{i;j}$, this method is akin to measurement with error and may subject the structural estimation to attenuation bias.

$$ {}^{2}_{y}(s_i) = \_^{2}_{y} s_i + "_{y;i} \tag{18}$$

where $\_{y}$ and can be estimated using non-linear least squared (NLLS) by using the sub-population variance for ${}^{2}_{y}(s_i)$. Alternatively, it can also be estimated by the Method of Moments (MM)[28], using the following rst and second moment conditions by de nition of mean and variance:

$$ g_y(y_{i;j;t}; \quad) = E \left[\begin{array}{cc} (y_{i;j;t} \quad {}_{y} + \quad s_i) & \left[\begin{array}{c} 1 \\ s_i \end{array} \right] \\ \_^{2}_{y} s_i \quad (y_{i;j;t} \quad {}_{s})^2 & \end{array} \right]_0 = 0 \tag{19}$$

where ${}_{y}$, $\_^{2}_{y}$, and

Figure 7: Cohort variance with fitted lines.

to note that the underlying $y_{i;j;t} - \mu_y(s_i)$ used to estimate $\sigma_y^2(s_i)$ is still correlated with $\varepsilon_{i;j}$, without which $\sigma_y^2(s_i) = \sigma_b^2(s_i) + \sigma_\varepsilon^2$ would not stand.

4.2.3 Identification

For the main parameters of interest, the model mainly utilizes the GMM estimator as described by Hansen (1982) and this subsection presents an overview of the identification and properties of the estimator. The moment conditions are estimated by taking its sample

$$G_0 = \text{plim} M^{-1} \sum_{M_{i;j;t}} \frac{\partial g(W_{i;j;t}; \theta)}{\partial \theta'} \Bigg|_{\theta_0} \tag{24}$$

The local identification of the nonlinear model requires the sufficient and necessary rank condition for the estimated $\hat{G} = G(\hat{\theta})$ that

$$\text{rank}(\hat{G}) = p \tag{25}$$

In other words, the estimated $\hat{G}$ must be of full rank for the model to be identified, otherwise the variance-covariance matrix (under optimal weighting matrix)

$$\hat{V}[\hat{\theta}] = M(\hat{G}' w_M \hat{G})^{-1} \tag{26}$$

cannot be calculated as $\hat{G}' w_M G$ would be singular. The optimal weighting matrix is calculated as

$$w_M = M^{-1} \sum_{M_{i;j;t}} [gg'|_{\theta_o}] \tag{27}$$

introduces additional variations to the matrix of instruments. Alternatively, generalized inverse may be used to produce variance-covariance and weighting matrices when numerical private values to be identically distributed.

Krasnokutskaya (2011)[32] proposes a log-decomposition of bids if the e ect of heterogeneity is multiplicative factor. This works well with the structural model, in which the private value is de ned as equation (8

$$\ln r_{i;j;t} = \ _i + M_t^{0} B_M + _r \ "_{i;j;t}$$

The reasons for the refitting are threefold. First, $_i$s still need to be estimated to produce $_r$ and $_r$. Second, the refitting corrects some of the correlation between M_t and $n_{i;j;t}$ that would bias B_M through omitted variables. Third, the observed bids are not assumed to be perfectly in accordance with the equilibrium strategy, and using the pseudosample itself in place of $r_{i;j;t}$ overfits the model; instead, the refitted values, which are the conditional expectation of the private values, accounts for the measurement error in the pseudosample and reduces the likelihood of estimation bias.

Once $r_{i;j;t}$s are estimated, they are plugged back into the structural model. Now the parameters that remain to be estimated are only and [34]. The fixed effect dummies are hereon dropped from Z for the main estimation, while M_t are retained as instruments for n_j.

5 Results and analyses

This section presents the estimation results and a brief discussion on policy implications. Several variables are transformed prior to the analysis. The market factors are converted to 2005 dollars using Construction Pricing Index and scaled to the millions. The engineer's estimates are also converted to 2005 dollars and logged. Number of bids within sample period is normalized to 1 against the firm with the highest number of bids based on the full valid sample before data cleanup and trimming. Auctions with only 1 bidder are removed from the sample prior to estimation. A period variable, measured by month, is included as an additional instrument to account for any unmodeled time trend.

5.1 Estimation results

Figure 8 shows the estimation results for the private value. The top panels compares the estimated private values to observed bids, and the bottom panels visualize the kernel density estimates. The left panels show the results for the pseudosample, and the right panels for the fitted private values.

Given the construction of the pseudosample, the pseudo private value is necessarily less than or equal to the observed bids, while around one third of the fitted private values are greater than the observed bids (below $b = r$ line). This does not pose a problem as the

[34] In future exploration of this working paper, the fitted parameters will serve as starting values for the full GMM estimation with optimal h(Z) within a high-performance computing environment.

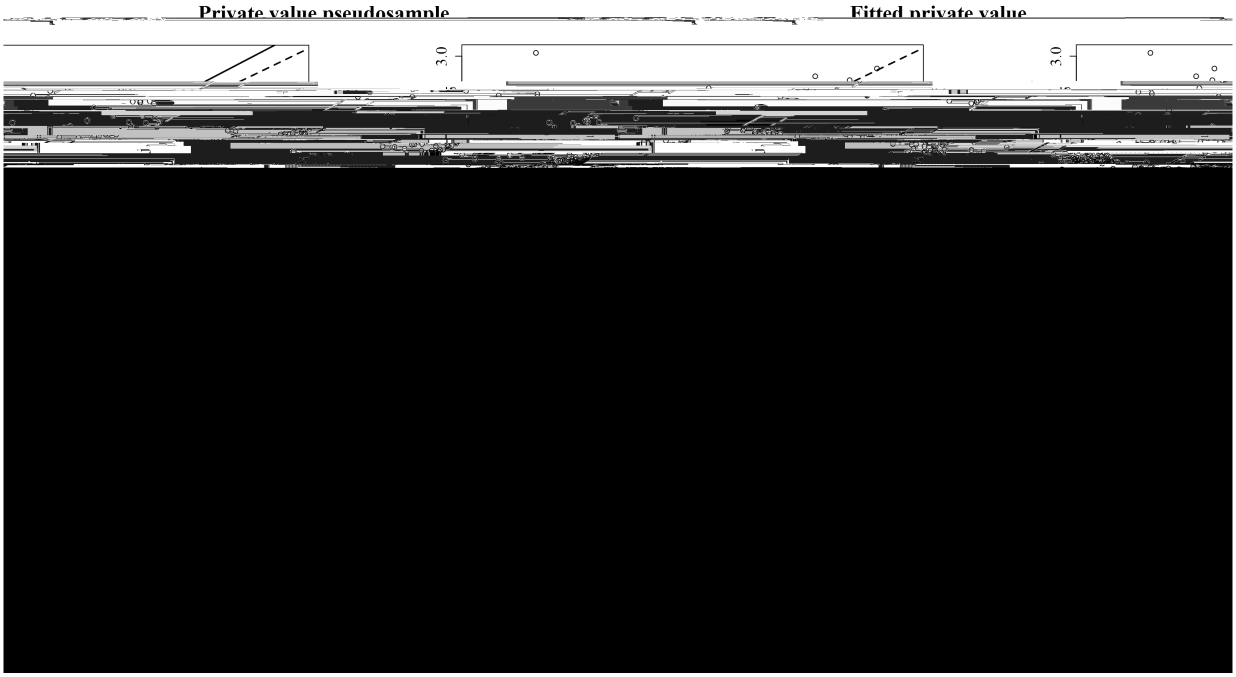


Figure 8: Estimated private value.

di erence is absorbed into the error term, and the linear t shows an average of 8.55% markup using the tted values.

Figure 9 shows the estimated distribution of xed e ects tted from the pseudosample. Counterintuitively, I nd that larger rms tend to have a higher private value despite the assumption of economy of scale. However, as the private value estimate is not limited to accounting cost alone, this nding is not a surprise. Larger rms face more opportunity cost through at least two channels: rst, the greater capacity of large rms bring about more opportunities within multiple markets, some of which may have better value; second, larger rms are also more likely to be closer to or exceeding capacity constraint since they have a revolving inventory of deliverables, whereas smaller rms tend to cycle through growing and lean seasons. In addition, this result is consistent with both theoretical predictions and observations; despite having greater opportunity cost, larger rms bid lower on average due to having less uncertainty in both private and common values.

Tables 3 shows the main parameter results. The models without CV uncertainty assumes $= 0$. As expected, an uninstrumented n_j attenuates the heterogeneity estimates for both private and common values, although not to a great degree. The heterogeneous private value uncertainty estimate, , remains signi cant in all speci cations, and the results from Model 3(4) suggests that one-time bidders face as much as eight times more uncertainty than the most frequent bidders, although the e ect tapers o quickly as rm size increases.

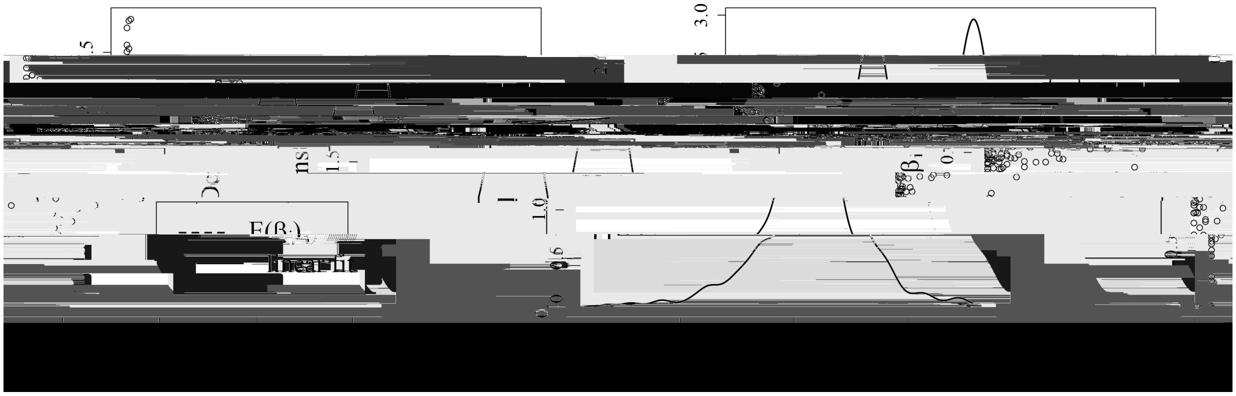


Figure 9: Estimated fixed effects.

Table 3: Estimation Results.

	Model Specification			
IV for n_j	No		Yes	
CV uncertainty	No	Yes	No	Yes
Parameter	3(1)	3(2)	3(3)	3(4)
	-0.34789	-0.34581	-0.38698	-0.39061
	(0.00707)	(0.01201)	(0.00577)	(0.00964)
		0.02389		-0.13324
		(0.33126)		(0.21474)

	First-stage results	
	$_r$	0.20435
	E($_i$)	0.04346
	Total observations	5682

Table 4: Alternative speci cations for common value uncertainty.

	Model Speci cation	
i;j;t	Benchmark	M^0

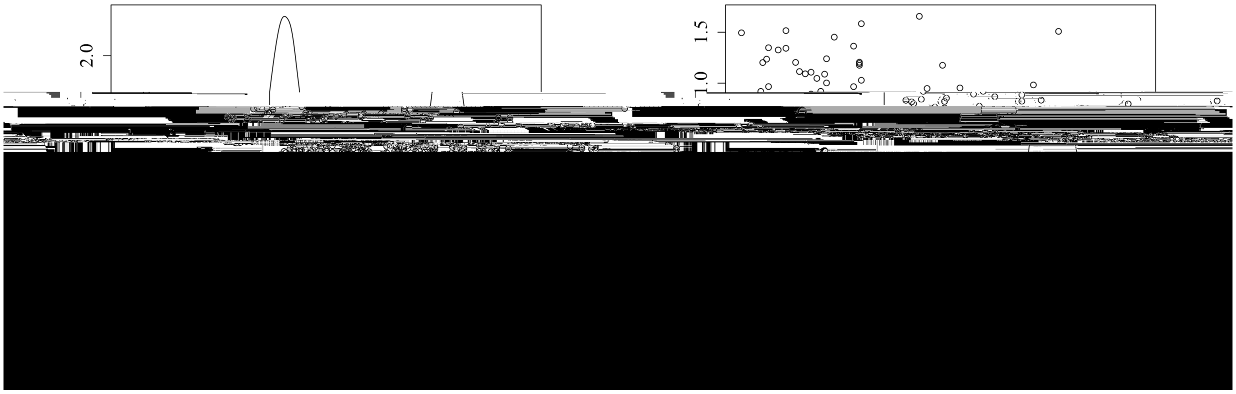


Figure 10: Estimated error term.

those of . While Model 4(2) deviates from the theoretical de nition of common value and Model 4(3) can be too volatile due to the introduction of one more parameter without added covariates, these additional results support the heterogeneous private and common value uncertainty estimate from the benchmark model. The term s_i $_r$ by itself is more di cult to interpret; however, the result from Model 4(1), which only estimates the e ect on rm size on the belief of $_r$, conforms to the casual prediction of Proposition 1 that smaller rms observe a lower $_{i;j}$.

Figure 10 shows the estimated error term. The rst and third moment conditions are well attained (left panel). For the second moment conditions, heteroskedasticity is mostly reduced except for the smallest rms (right panel)[36]. This is likely due to that the bidding behavior of small rms is not as well explained by the structural model as the larger rms, suggesting that smaller rms abide by the optimal bidding strategy to a lesser degree and tend to bid more erratically, giving rise to another source of the heteroskedastic outcome related to uncertainty but not accounted for in the model.abide by the optimal bidding strategy to a lesser degree and tend to bid more erratically, giving rise to another source of the heteroskedastic outcomeabide by the optimal bidding strategy to a lesser degree and tend to bid more erratically, giving rise to another source of the heteroskedastic outcome

5.2 Analysis

The estimation results allow the analysis of rms' behavior facing uncertain number of bidders through a simple calibration exercise, as well as the counterfactuals of potential

[36]Grouping the second moment conditions sample a0 d 00 -25.64856(o)-48r momca lesser degree and

outcomes when the heterogeneity of uncertainty is removed. For the remaining discussion, the benchmark model is used, which contains the most conservative estimates for heterogeneous uncertainty.

5.2.1 Number of bidders

While the model assumes that $\varepsilon_{i;j}$ is independent from Z, the fitted $b_{i;j;t}$ is not in the instrumental variable model. As such, the correlation between the instrumental variables and $y_{i;j;t}$ can be estimated by partially fitting the structural model with the estimated parameters while substituting n_j with

$$n_{i;j;t} = \ (n_j; \tilde{n}_{i;j;t}; s_i | H) \tag{34}$$

where $H = \{ \ _1; \ _2; \ _3\}$ are the the parameters to be calibrated. The calibration uses the method of moments with the following moment conditions:

$$E[y_{i;j;t} \quad b_{i;j;t} | \tilde{n}_{i;j;t}; n_j \quad \tilde{n}_{i;j;t}; s_i]_{H_0} = 0 \tag{35}$$

where $\tilde{n}_{i;j;t}$ is the combined signal for the number of bidders obtained from the Poisson regression used in $\hat{b}_{i;j;t}$ of the second moment condition and in the nonparametric estimation, and $n_j \quad \tilde{n}_{i;j;t}$ is the difference between observed number of bidders and the combined signal. In this sense, $\tilde{n}_{i;j;t}$ is the public signal observable to both firms and the investigator, and $n_j \quad \tilde{n}_{i;j;t}$ is the additional variation in the number of bidders for which the investigator observes no signal, but it may be signaled to bidders. The calibrated parameters would describe how well bidders are able to anticipate both components of the number of bidders.

In Model 5(1), the results show that firms in general anticipate the number of bidders well, particlarly using signals both observable to the investigator, and to a lesser extent the remaining variations. The previous section finds that while uninstrumented n_j attenuates

Table 5: Calibrated firm anticipation of the number of bidders.

$(n_j; \tilde{n}_{i;j;t})$	Model Specification	
	${}_1\tilde{n}_{i;j;t} + {}_2(n_{i;j} \quad \tilde{n}_{i;j;t})$	$[\ {}_1\tilde{n}_{i;j;t} + {}_2(n_{i;j} \quad \tilde{n}_{i;j;t})]s_i^{\ 3}$
Parameter	5(1)	5(2)
1	1.05400	0.80147
	(0.04464)	(0.05348)
2	0.7789	0.5289
	(0.1354)	(0.10642)
3		-0.13753
		(0.02839)

The bid submission process does not conveniently allow firms to simultaneously observe the number of bidders. However, the contingent bid design proposed by Harstad et al. (1990) lets firms submit multiple bids at once, each for a different realized number of bidders, thereby removing this dimension of uncertainty[37]. The effect of removing n_j uncertainty is discussed in the following section.

5.2.2 Expenditure and allocation

Table 6 shows the counterfactuals of average lowest bids grouped by project value, measured by engineer's estimates, under various scenarios. The predicted scenario (Model 6(1)) uses the fitted bids with combined signal $\tilde{n}_{i;j;t}$, the $;$ $= 0$ scenario assumes a hypothetical removal of the heterogeneity in both private-value and common-value uncertainty, and the known n_j scenario uses the fitted bids with observed n_j.

Given that the observed bids have a larger variance than the predicted bids, the predicted average lowest bids are conceivably higher than observed. In the equalized private and common value uncertainty scenario, the average lowest bid in all project value tiers are lower than the predicted. In the known n_j scenario, the opposite is true, which is consistent with the overestimation of competition, especially by smaller firms, discussed in the previous section.

Alternatively, Table 7 shows the average bid by project value. Here in the equalized

[37] Although firms may adopt a different strategy due to increased bidding cost and effort to conceal private value.

Table 6: Average lowest bid by project value.

Scenario		Observed	Predicted	; = 0	Known n_j	Both
Project value	Projects		6(1)	6(2)	6(3)	6(4)
\$50K-\$500K	267	1.0060	1.0437	0.9851	1.0760	1.0048
\$500K-\$1.2M	267	0.9855	1.0348	0.9833	1.0687	1.0055
\$1.2M-\$2.5M	239	0.9834	1.0390	0.9908	1.0657	1.0117
\$2.5M-\$5M	227	0.9662	1.0471	1.0079	1.0707	1.0236
\$5M-100M	231	0.9708	1.0489	1.0137	1.0686	1.0283
Total	1231	0.9832	1.0425	0.9954	1.0700	0.0141

private and common value uncertainty scenario, the average bid is lower than both observed and predicted accounts, with only small di erences between the observed and the predicted. In the known n_j scenario, while the average bid is still mostly higher, the di erence is quite reduced. The average bid counterfactual lends a robust additional support for the cost-saving aspect of equalizing private and common value uncertainty, especially given the conditional expectation nature of regression models.

Under the same scenarios, I also examine the potential allocational outcome with respect to rm size. Table 8 shows the lowest bid share by size cohort[38]

Table 7: Average bid by project value.

Scenario		Observed	Predicted	; = 0	Known n_j	Both
Project value	Projects		7(1)	7(2)	7(3)	7(4)
\$50K-\$500K	267	1.1665	1.1164	1.0583	1.1318	1.0666
\$500K-\$1.2M	267	1.1136	1.0999	1.0476	1.1182	1.0570
\$1.2M-\$2.5M	239	1.1063	1.0984	1.0534	1.1132	1.0626
\$2.5M-\$5M	227	1.0742	1.0916	1.0577	1.1031	1.0648
\$5M-100M	231	1.0713	1.0955	1.0642	1.1048	1.0701
Total	1231	1.1065	1.1005	1.0563	1.1143	1.0643

Table 8: Lowest bid share by size cohort.

Scenario	Observed	Predicted	; = 0	Known n_j	Both
Firm size		8(1)	8(2)	8(3)	8(4)
1 - 25	0.2136	0.1227	0.2900	0.1129	0.2868
26-50	0.1641	0.2071	0.2380	0.1917	0.2429
51-75	0.1795	0.2518	0.2015	0.2299	0.1917
76-100	0.1584	0.2348	0.1795	0.2283	0.1803
101-150	0.1560	0.1129	0.0626	0.1324	0.0626
>150	0.1284	0.0707	0.0284	0.1048	0.0357

still lead to both better overall allocational efficiency and less expenditure uncertainty, with the increased cost and affirmative loss mostly compensated for, if combined with reduced heterogeneity in private and common value uncertainty (Models 6(4), 7(4), and 8(4)).

6 Conclusion

In this paper, I show that smaller firms tend to have greater uncertainty in procurement auctions, and with the identified parameters of heterogeneous uncertainty, I also show that efforts to reduce heterogeneity in uncertainty may lead to both cost savings for the government and better allocations to smaller firms. More generally, I propose, develop, and solve a model to recover structural parameters of heterogeneous uncertainty through heteroskedastic outcomes in procurement auctions, and the described method may be extended to studying the origin of heteroskedastic outcomes in other market settings as well.

A limitation of the paper is the partial abstraction from the selective entry. While neither project value or type are found to be good predictors of entry and bidding behavior, and they are also used as instrumental variables such that the error term cannot be correlated with these factors, the paper assumes that all projects attract bidders from the same distribution of private values, which may not be the case if entry is endogenous. In this sense, if projects more often bid on by smaller firms tend to have a higher dispersion in private values from participating firms, the heterogeneous uncertainty estimate would absorb some of that effect, though the variance itself is still a form of uncertainty even if it is correlated with, but arguably exogenous to, firm size. If the opposite is true, the heterogeneous uncertainty estimate with respect to firm size would be attenuated. Allowing endogenous entry and conditional distribution of private values on project value and type, if feasible either through

D. G. De Silva, T. Dunne, and G. Kosmopoulou. An empirical analysis of entrant and incumbent bidding in road construction auctions. *The Journal of Industrial Economics*, 51(3):295{316, 2003. 3, 11, 31

B. Elyakime, J. J. La ont, P. Loisel, and Q. Vuong. First-price sealed-bid auctions with secret reservation prices. *Annales d'Economie et de Statistique*, pages 115{141, 1994. 3

P. Fevrier. Nonparametric identi cation and estimation of a class of common value auction models. *Journal of applied econometrics*, 23(7):949{964, 2008. 3

V. Flambard and I. Perrigne. Asymmetry in procurement auctions: Evidence from snow removal contracts. *The Economic Journal*, 116(514):1014{1036, 2006.

A. R. Gallant and G. Tauchen. Which moments to match? *Econometric Theory*, 12(4): 657{681, 1996.

M. Gentry and T. Li. Identi cation in auctions with selective entry. *Econometrica*, 82(1): 315{344, 2014.

R. Goncalves. Empirical evidence on the impact of reserve prices in english auctions. *The Journal of Industrial Economics*, 61(1):202{242, 2013.

M. B. Gordy. Computationally convenient distributional assumptions for common-value auctions. *Computational Economics*, 12(1):61{78, 1998.

D. J. Henderson, J. A. List, D. L. Millimet, C. F. Parmeter, and M. K. Price. Empirical implementation of nonparametric rst-price auction models. *Journal of Econometrics*, 168 (1):17{28, 2012. 3, 11

K. Hendricks and R. H. Porter. An empirical study of an auction with asymmetric information. *The American Economic Review*, pages 865{883, 1988. 10

K. Hendricks and R. H. Porter. An empirical perspective on auctions. Technical report, CSIO working paper, 2006. 2, 11

K. Hendricks, J. Pinkse, and R. H. Porter. Empirical implications of equilibrium bidding in rst-price, symmetric, common value auctions. *The Review of Economic Studies*, 70(1): 115{145, 2003.

K. Jurado, S. C. Ludvigson, and S. Ng. Measuring uncertainty. *The American Economic Review*, 105(3):1177{1216, 2015. 1

J. H. Kagel and D. Levin. The winner's curse and public information in common value auctions. *The American economic review*, pages 894{920, 1986.

J. H. Kagel, D. Levin, R. C. Battalio, and D. J. Meyer. First-price common value auctions: Bidder behavior and the "winner's curse". *Economic Inquiry*, 27(2):241{258, 1989.

T. Li, I. Perrigne, and Q. Vuong. Conditionally independent private information in ocs wildcat auctions. *Journal of Econometrics*, 98(1):129{161, 2000.

T. Li, I. Perrigne, and Q. Vuong. Structural estimation of the a liated private value auction model. *RAND Journal of Economics*, pages 171{193, 2002. 11

J. Marion. Are bid preferences benign? the e ect of small business subsidies in highway procurement auctions. *Journal of Public Economics*, 91(7):1591{1624, 2007.

D. McFadden. A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica: Journal of the Econometric Society*, pages 995{1026, 1989. 2

P. Milgrom and I. Segal. Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2): 583{601, 2002. 41

L. Moretti and P. Valbonesi. Firms' quali cations and subcontracting in public procurement: an empirical investigation. *The Journal of Law, Economics, and Organization*, 31(3): 568{598, 2015.

J. Nakabayashi. Small business set-asides in procurement auctions: An empirical analysis. *Journal of Public Economics*, 100:28{44, 2013. 3

T. H. Noe, M. Rebello, and J. Wang. Learning to bid: The design of auctions under uncertainty and adaptation. *Games and Economic Behavior*, 74(2):620{636, 2012.

H. J. Paarsch. Deciding between the common and private value paradigms in empirical models of auctions. *Journal of econometrics*, 51(1-2):191{215, 1992.

H. J. Paarsch, H. Hong, et al. An introduction to the structural econometrics of auction data. *MIT Press Books*, 1, 2006. 2

A. Pakes and D. Pollard. Simulation and the asymptotics of optimization estimators. *Econometrica: Journal of the Econometric Society*, pages 1027{1057, 1989. 2

R. Schmalensee, M. Armstrong, R. D. Willig, and R. H. Porter. *Handbook of industrial organization*, volume 3. Elsevier, 1989. 2

M. Skitmore, D. Drew, and S. Ngai. Bid-spread. *Journal of construction engineering and management*, 127(2):149{153, 2001. 3, 11

H. Vazquez-Leal, R. Castaneda-Sheissa, U. Filobello-Nino, A. Sarmiento-Reyes, and J. Sanchez Orea. High accurate simple approximation of normal distribution integral. *Mathematical problems in engineering*, 2012, 2012.

P.-L. Yin. Information dispersion and auction prices. 2006.

R. Zeithammer and C. Adams. The sealed-bid abstraction in online auctions. *Marketing Science*, 29(6):964{987, 2010.

Y. Zhu. Essays on empirical auctions and related econometrics, 2014.

A Proofs and Solutions

A.1 Proof of Proposition 1

Given the conversion from log-normal mean m and variance v to normal mean and variance ^{2}

$$= \ln \frac{m}{\sqrt{1+\frac{v}{m^2}}} \tag{36}$$

$$^2 = \ln \left(1+\frac{v}{m^2}\right) \tag{37}$$

can be rewritten as

$$= \ln m \quad \frac{1}{2}\ln\left(1+\frac{v}{m^2}\right)$$

$$= \ln m \quad \frac{1}{2}\,{}^2 \tag{38}$$

Even if the bidder correctly observes m, a misoberved $_i$ results in

$$_{i;j} = \ln m_j \quad \frac{1}{2}$$

reservation value $v_i \overset{iid}{\sim} F(v_i)$, where $F(\cdot)$ is the cumulative distribution function of private values. The probability that bidder i has the lowest private value among n bidders is therefore $S^{n-1}(v_i)$, where $S(v_i) = 1 - F(v_i)$ is the survival function of v_i. The expected payoff from any monotonic bidding strategy b_i is

$$\pi(v_i; b_i) = (b_i - v_i)S(v_i)^{n-1} \tag{41}$$

Let $B(v_i)$ be the optimal bidding function that is monotonically increasing in v_i and symmetric under the same belief and $B^{-1}(v_i)$ be its inverse, the payoff can be rewritten as

$$\begin{aligned} \pi_i(v_i; b_i) &= (b_i - v_i)S(B^{-1}(b_i))^{n-1} && (42) \\ \pi_i(v_i) &= (B(v_i) - v_i)S(v_i)^{n-1} && (43) \end{aligned}$$

By Envelope Theorem,

$$\begin{aligned} \frac{d\pi_i(v_i)}{dv_i} &= \left.\frac{\partial \pi_i(v_i; b_i)}{\partial v_i}\right|_{b_i = B(v_i)} \\ &= -\left. S(B^{-1}(b_i))^{n-1}\right|_{b_i = B(v_i)} \\ &= -S(v_i) \end{aligned} \tag{44}$$

Integrating the expression above from bidder i's private value to the upper bound, we obtain

$$\int_{v_i}^{\bar{v}} \frac{d\pi_i(x)}{dx} dx = -\int_{v_i}^{\bar{v}} S(x)^{n-1} dx \tag{45}$$

By the fundamental theorem of calculus, the same integral is also equal to

$$\begin{aligned} \int_{v_i}^{\bar{v}} \frac{d\pi_i(x)}{dx} dx &= \underbrace{\pi_i(\bar{v})}_{=0} - \pi_i(v_i) \\ &= -\pi_i(v_i) \end{aligned} \tag{46}$$

Because the bidder with the highest reservation has a zero probability of winning, $\pi_i(\bar{v}) = 0$. Settting the two representations of the integral equal, we obtain the optimal bidding function

$$\beta_i(v_i) = \int^{v_i} S(x)^{n-1}dx \quad (47)$$

$$(B(v_i) - v_i)S(v_i)^{n-1} = \int_{v_i}^{\bar{v}} S(x)^{n-1}dx \quad (48)$$

$$B(v_i) = v_i + \frac{\int_{v_i}^{\bar{v}} S(x)^{n-1}dx}{S(v_i)^{n-1}} \quad (49)$$

Replace $S(\cdot)$ with the survival function of log-normal distribution expressed in terms of normal CDF, we arrive at

$$B(r_{i;j}) = r_{i;j} + \frac{\int_{\ln r_{i;j}}^{\infty}[1-\Phi(\frac{\ln x-\mu_{i;j}}{\sigma_i})]^{n_j-1}dx}{[1-\Phi(\frac{\ln r_{i;j}-\mu_{i;j}}{\sigma_i})]^{n_j-1}} \quad (50)$$

A.3 First-order derivatives

Given the estimation equation

$$y_{i;j;t} = r_{i;j;t} + \underbrace{\frac{\int_{\ln r_{i;j;t}}^{\infty}\left(1-\Phi\left(\frac{x - s_i\theta_{j;t}}{\rho s_i}\right)\right)^{n_j-1}dx}{\left(1-\Phi\left(\frac{\ln r_{i;j;t} - s_i\theta_{j;t}}{\rho s_i}\right)\right)^{n_j-1}}}_{b_{i;j;t}} + \varepsilon_{i;j} \quad (51)$$

Let $\Phi(\cdot)$ represent the normal CDF including all relevant variables. By Leibniz's rule of integral differentiation, the first-order partial derivatives of parameters of the structural equation are calculated as below with simplification steps omitted.

Given $n = e^{X_{IV}'B_{IV}}$, we have for each $\beta_{iv} \in B_{IV}$

$$\begin{aligned}\frac{\partial b}{\partial \beta_{iv}} &= \frac{\int_{\ln r}^{\infty}\frac{\partial}{\partial \beta_{iv}}[1-\Phi(x;\cdot)]^{n(\beta_{iv})-1}dx}{[1-\Phi(\cdot)]^{n-1}} - \frac{\int_{\ln r}^{\infty}[1-\Phi(x;\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{2(n-1)}}\frac{\partial}{\partial \beta_{iv}}[1-\Phi(\cdot)]^{n(\beta_{iv})-1} \\ &= \frac{\int_{\ln r}^{\infty} x_{iv}n\ln[1-\Phi(x;\cdot)][1-\Phi(x;\cdot)]^{n-1}(x;\cdot)dx}{[1-\Phi(\cdot)]^{n-1}} \\ &\quad - \frac{\int_{\ln r}^{\infty}[1-\Phi(x;\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n-1}}x_{iv}n\ln[1-\Phi(x;\cdot)]\end{aligned} \quad (52)$$

Let $\phi(\cdot)$ represent the corresponding normal PDF to $\Phi(\cdot)$. Given $r = e^{\alpha_i + M_t'B_M}$, for each

$$
\begin{aligned}
\frac{\partial}{\partial s_i} V[\varepsilon | s_i; \ _0] &= \frac{\partial}{\partial s_i} \int_G [g^2(\ (s_i); \ (s_i) | W; \ _0) \quad _G] f(\varepsilon | s_i) dg \\
&= \int_G \frac{\partial}{\partial s_i} \ g^2(\ (s_i); \ (s_i) | W; \ _0) f(\varepsilon | s_i) \ dg \\
&= \int_G f(\varepsilon | s_i) \frac{\partial}{\partial s_i} g^2(\ (s_i); \ (s_i) | W; \ _0) dg \\
&\quad + \int_G g^2(\ (s_i); \ (s_i) | W; \ _0) \frac{\partial}{\partial s_i} f(\varepsilon | s_i) dg
\end{aligned}
\tag{62}
$$

Note that $\text{het}(s_i) = \int_G g^2(\ ; \ | W; \ _0) \frac{\partial}{\partial s_i} f(\varepsilon | s_i) dg$, then

$$
\frac{\partial}{\partial s_i} V[\varepsilon | s_i; \ _0] = \int_G f(\varepsilon | s_i) \frac{\partial}{\partial s_i} g^2(\ (s_i); \ (s_i) | W; \ _0) dg + \text{het}(s_i) \tag{63}
$$

Now,

$$
\begin{aligned}
& \frac{\partial}{\partial s_i} g^2(\ (s_i); \ (s_i) | W; \ _0) \\
= \ & 2 g(\ (s_i); \ (s_i) | W; \ _0) \frac{\partial}{\partial s_i} g(\ (s_i); \ (s_i))
\end{aligned}
\tag{64}
$$

Since $g(\ (s_i); \ (s_i) | W; \ _0) = 0$, we have

$$
\frac{\partial}{\partial s_i} V[\varepsilon | s_i; \ _0] = \text{het}(s_i) \tag{65}
$$

A.5 Proof of inverse bidding function pseudosample estimator

This is a sketch of proof of the pseudosample estimator following Guerre et al. (2000) with modifications for reverse auctions. Rewrite the objective function 41 as

$$
(v_i; b_i) = (b_i - v_i) S(B^{-1}(b_i))^{n-1} \tag{66}
$$

where $B_i^{-1}(b_i) = v_i$ is the inverse optimal bidding function. The first-order conditions become

$$
\begin{aligned}
\frac{d}{db_i} (v_i; b_i) &= (b_i - v_i)(n-1) S(B^{-1}(b_i))^{n-2} S'(B^{-1}(b_i))(B'(B^{-1}(b_i))^{-1} + S(B^{-1}(b_i))^{n-1} \\
&= [(b_i - v_i) S(B^{-1}(b_i))^{-1} S'(B^{-1}(b_i))(B'(B^{-1}(b_i))^{-1} + 1] S(B^{-1}(b_i))^{n-1} \\
&= 0
\end{aligned}
\tag{67}
$$

Given that $S(v_i) = 1 - F(v_i)$ and $B_i^{-1}(b_i) = v_i$, simplify to yield the first-order differential equation

$$1 - (b_i - v_i)(n - 1)\frac{f(v_i)}{S(v_i)B'(v_i)} = 0 \tag{68}$$

The solution to equation 68 is the same as the solution to the optimal bidding function in reverse auctions (equation 49). Let $S_b(\cdot)$ and $f_b(\cdot)$ denote the survival and density function of b_i. Since $B(v_i)$ is monotonically increasing in v_i, it must be the case that $S_b(b_i) = \Pr(b_{-i} > b_i) = \Pr(v_{-i} > b_i) = S(v_i)$ and that $B'(v_i) > 0$ such that $f_b(b_i) = |\frac{d}{db_i}B^{-1}(b_i)|f(B^{-1}(b_i)) = f(v_i)/B'(v_i)$, therefore

$$\frac{f_b(b_i)}{S_b(b_i)} = \frac{f(v_i)}{S(v_i)B'(v_i)} \tag{69}$$

Substitute into 68 we obtain

$$1 - (b_i - v_i)(n - 1)\frac{f_b(b_i)}{S_b(b_i)} = 0 \tag{70}$$

which solves to yield the structural form of the inverse bidding function pseudosample estimator

$$v_i = b_i - \frac{1}{n - 1}\frac{S_b(b_i)}{f_b(b_i)} \tag{71}$$