DISCUSSION PAPERS IN ECONOMICS

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†
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 firms cannot observe the number of bidders, a third dimension of uncertainty is added, and the effect of asymmetric information readily emerges in such highly uncertain environments. By exploiting the heteroskedasticity of normalized bids with respect to firm 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 firm 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 Classifications: C57, D44, H74, R42

*I would like to thank my dissertation adviser, Professor Daniel Kaffine, 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.

†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 of treatment in its presence. However, few studies consider the underlying mechanism that gives rise to heteroskedasticity, and understanding the origin of heteroskedasticity in a given market setting may shed light on certain structures and dynamics of the associated economic activities.

Instead of treating heteroskedasticity as an issue to be dealt with, this study investigates how heteroskedastic outcomes can yield useful information about agent heterogeneity, specifically heterogeneous uncertainty. Recent macroeconomic literature has been fruitful in studying uncertainty induced by economic shocks by examining second moment constructions[1] of empirical data, but little recent work has been done in applied microeconomics, even though uncertainty is incorporated in many microeconomic models. Uncertainty in microeconomic theory is a necessary result of incomplete information about market participants, but it can also stem from other sources, such as the degree of adherence to rational expectations, responsiveness to macroeconomic changes, and ability to decipher noisy signals, *etc.*. Given the differing levels of resources available to heterogeneous agents to extract information rent and the general presence of information asymmetry, it reasons that the degree of uncertainty among agents is heterogeneous as well.

Extending the practice in the macroeconomic literature to measure uncertainty, heteroskedastic outcome in certain market settings provide convenient data variation in the second moment for the understanding of heterogeneous uncertainty. In highly uncertain market settings such as procurement auctions, where uncertainty exists in both private value and common value, the effect of asymmetric information can be particularly pronounced, and if the number of bidders cannot be observed by firms *ex ante*, there is an added layer of uncertainty that can further augment the manifestation of said heterogeneity. Using over 10 years of procurement auctions records from the Colorado Department of Transportation

[1] Jurado et al. (2015) summarizes the existing body of literature on shock-induced uncertainty in addition to proposing a method of directly measuring time-varying uncertainty through business cycles.

(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 counterfactuals, I also find that reduced heterogeneity in uncertainty may result in both better cost savings to the government and improved allocation of projects to smaller firms.

The paper is informed by the literature on the econometrics of industrial organization and auction theory[2]. Earlier papers often rely on simulated methods[3] to deal with

[2] Laffont (1997), Athey and Haile (2006), Hendricks and Porter (2006) survey various empirical strategies in auction theory. Schmalensee et al. (1989), Cameron and Trivedi (2005), and Paarsch et al. (2006) provide useful instructional text on implementation.

[3] Summarized in Pakes and Pollard (1989); McFadden (1989) and notably applied to auctions in Laffont

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 identification 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, specifically, 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 significant 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 different business partners can inform the assesment 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 affirmative objectives[9].

et al. (1995), *etc.*.

[4]Notable additional works and extensions of nonparametric identification 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 identification of private value under unobserved heterogeneity with Michigan highway procurement data..

[6]De Silva et al. (2003) finds 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 effect and efficacy 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 specific small business carve out; instead, it takes affirmative action toward disadvantaged businesses through its Disadvantaged Business Enterprise Program (Colorado Department of Transportation).

[9]In procurement auction analysis by civil engineers and financial planners, bid spread is often of particularly interest, though it is often done in a descriptive manner (Skitmore et al., 2001). Identifying private value and heterogeneous uncertainty provides a robust economic basis for variations in bid spread.

The study makes several contributions to distinct areas of economic literature. First and most broadly, it puts forward a structural use of heteroskedasticity to extrapolate underlying economic parameters. Second, it accomplishes the direct estimation of both uncertainty in procurement auctions and its heterogeneity parameters with a standard model under fairly simple assumptions, a method that may be extended to other areas of microeconomic studies in different market settings where heteroskedastic agent response is observed. Third and more narrowly, it provides a new means of revisiting auction deisgn based on empirical data.

The remainder of this paper is organized as follows. Section 2 motivates the study approach with a discussion about the background and descriptive characteristics of the data. Section 3 develops a theoretical model of sealed-bid auction that incorporates heterogeneous uncertainty with in the private-value and common-value dimensions and investigates its properties. Section 4 describes a structural GMM estimation framework based on the theoretical results. Section 5 presents the estimation results and a few calibration and counterfactual analyses. Section 6 concludes the study with some thoughts about its implications and limitations.

2 Data

The bids data is obtained from the Colorado Department of Transportation (CDOT) on various types of highway projects that required open market contract bidding[10]. The data spans an 11.5-year period from January 2015 to June 2016. The market index data is obtained from the US Census Bureau and contains the monthly valuation of newly created private residence construction permit in Colorado for the same period. The private residence construction value data is chosen as the market indicator variable because it is assumed to correlate both with macroeconomic variables and with supply and demand specifically in the construction industry, either of which can exert an influence on firm's entry decisions and bidding behavior.

2.1 CDOT contract bidding

The Colorado Department of Transportation (CDOT) procures some of the construction and maintenance activities of its physical assets, as well as consulting and professional services, through open market contract bidding. The bidding process has the following general properties and procedures:

[10] See Colorado Department of Transportation (2014) for complete bidding rules, which include bidder prequalification, bidding procedure, selection criteria, and anti-collusion enforcement.

- 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

[11]The bids are rated on several factors and the "best value" is selected. In this sense, CDOT procurement differs from a standard first-price reverse auction. I assume that bidders do not behave differently from a first-price auction in this paper.

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 firms	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		>150	6	903

[a]Projects less than $50,000 typically need not go to competitive bid.
[b]Total number of bids within sample period.
[c]Colorado private construction permit value (Source: US Census) including current to 11 months lagged.
[d]Without missing value for requisite variables.

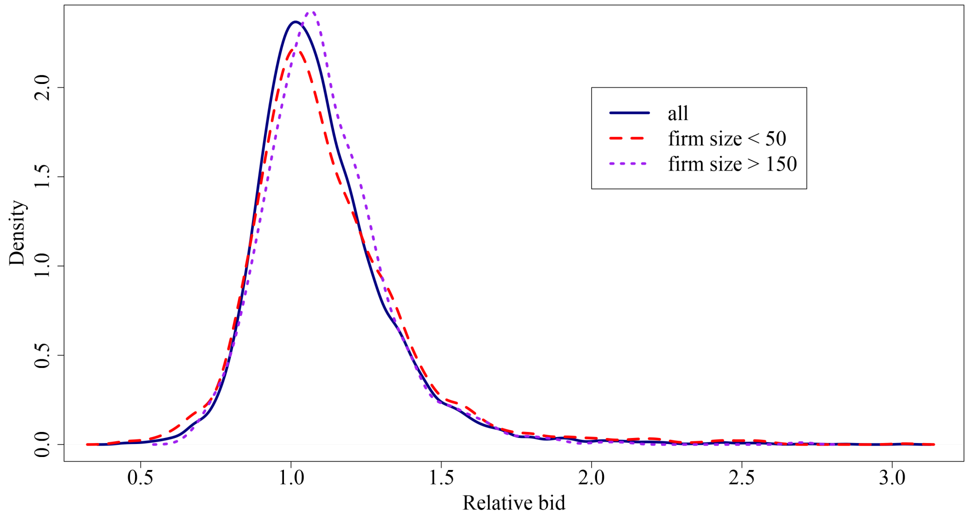

Figure 1: Kernel density estimate of bid distributions by size cohort.

distribution describes the product of random variables of certain attributes. In addition, the engineer's estimates can also follow a similar distribution and preserve the overall log-normality, as the quotient of two log-normal random variables is also log-normally distributed. This conjecture may meet some limitation as engineer's estimate varies between projects, but not within. However, because the bidders cannot observe the engineer's estimates, there are additional stochastic processes at play to the relative bid distribution.

The data shows that on average, firms bid around 10% above the engineer's estimate. This is consistent with the predictions of standard auction theory. However, if we examine the spread of the relative bids broken down by firm sized (Figure 2), proxied by the number of bids submitted over the entire sample, we can see a tightening of spread as firm size grows. The heteroskedasticity of bidding behavior over firm size is on clear display, and it raises the question of whether it results from heterogeneous uncertainty related to firm size.

Curiously, though, the figure does not show any discernible change in average relative bid, and the fitted line has only a negligible negative slope, *i.e.* large firms do not appear to be more or less likely to cut low on average, conditional on firm size alone. There are two important implications of this observation. First, the marginal distribution of the private values of larger firms may be tighter around the mean. Second, firm's bidding behavior with respect to size may not be monotonic. These observations provide an unique incision point to answering the research question, and it informs the suitable construction of theoretical and empirical models.

The relative bid is surprisingly unresponsive to a variety of factors, such as project

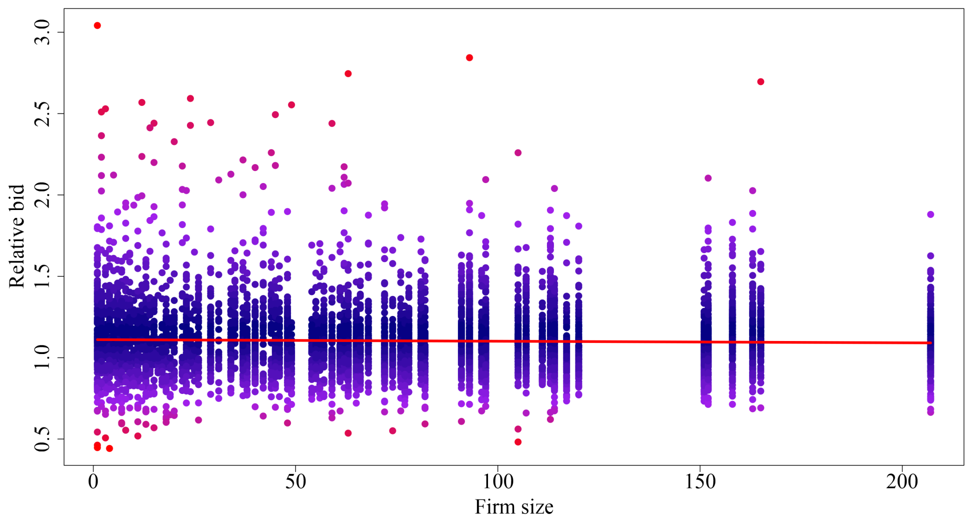

Figure 2: Relative bid spread by annual bidding size cohort. Red line denotes fitted value to size. Color spectrum denotes distribution of logged relative bid.

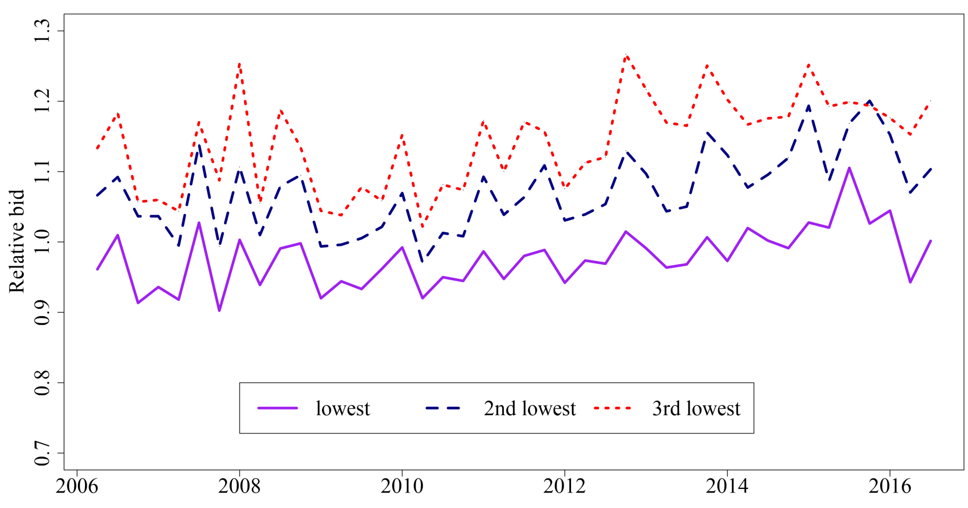

Figure 3: Quarterly average of 1^{st}, 2^{nd}, and 3^{rd} relative bid and number of bidders.

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 1$^{\text{st}}$, 2$^{\text{nd}}$, and 3$^{\text{rd}}$ 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.

[12] The model is descriptive and is not cross-validated to ensure best fit or minimized autocorrelation, though it can be shown to yield good predictive power based within a certain future time frame based on the training set and lagged explanatory variables used.

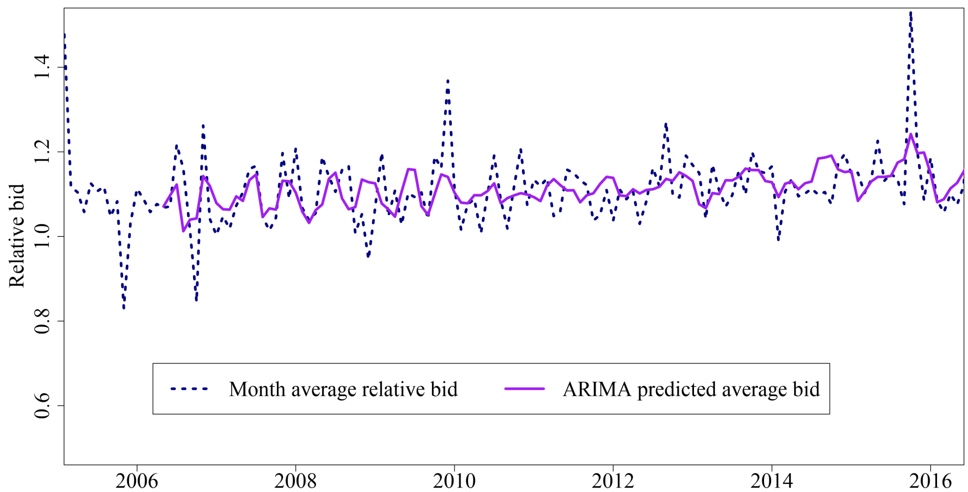

Figure 4: Observed and predicted monthly average bid using construction market data.

3 Behavioral Framework

The theoretical model borrows from the standard sealed-bid first-price auction model adjusted for procurement auctions. The standard model assumes symmetric information, which goes against the central premise of heterogeneous uncertainty assumed in the paper. However, the use of the standard model is justified because firms are assumed to formulate their bidding function as if it were the symmetric equilibrium strategy based on their heterogeneous beliefs, and Bayesian-Nash equilibrium is not assumed to have been attained in the bidding outcome[13]. In addition, the standard model provides a convenient framework in which to implement the estimation of the parameters of uncertainty in both private and common value.

3.1 Assumptions

Motivated by the empirical distribution of relative bids discussed in Section 2 and above, I assume that firm i's private value, or opportunity cost, for project j, $v_{i,j}$, is determined by a Cobb-Douglass process:

$$v_{i,j} = \prod_k X_{i,j,k}^{\alpha_k} \tag{1}$$

where $X_{i,j,k}$ is the kth factor of firm i's cost input and is independently and identically

[13] Lebrun (1996) shows that asymmetric equilibrium does not exist in first-price auctions, while Hendricks and Porter (1988) shows that firms' bidding strategies are still consistent with Bayesian-Nash equilibrium with asymmetric information.

distributed among is and js while independently distributed among ks. Assume also that the benchmark value of the project, $\bar{v}_j$ is generated by a similar process,

$$\bar{v}_j = \prod_k \bar{X}_{j,k}^{\alpha_k} \tag{2}$$

where $\bar{X}_{j,k}$ 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 $X_{\cdot,k}$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 $\bar{v}_k$ is similarly distributed as $v_{i,j}$, the relative private value

$$r_{i,j} = \frac{v_{i,j}}{\bar{v}_j} \tag{3}$$

is also log-normally distributed, here we assume with $E[\ln(r_{i,j})] = \mu_j$ (mean log relative value) and $Var[\ln(r_{i,j})] = \sigma$, where u_j and σ are parameters of the normal distribution that results from logrithmizing $r_{i,j}$.

In addition to the descriptive findings of this study, the log normality assumption of auction value distribution is well-supported by empirical auction literature [14] as well as recent civil engineering literature on highway project contract bidding data[15]. The relative value, as well as the subsequent relative bid, construction is attractive because it normalizes the bids on projects of differing monetary sizes, preserves the log normality of the original distributions, and combines both the private value and common value aspects of these auctions with observed and unobserved value components[16]. In addition, it provides for a convenient alternative to the affiliated private value model (Li et al., 2002) and simplifies the estimation strategy.

An important assumption is that μ_j is project-specific but σ is not[17]. This stems from the reasoning that some benchmark value is close to $E[v_{i,j}]$, but fleeting market forces outside of the benchmark value result in time-specific deviations in valuation that affect all firms the same way, while the spread of $v_{i,j}$ is unaffected.

The firms do not observe mean log relative value μ_j, but they have a derived belief $\mu_{i,j}$[18]

[14] Haile and Tamer (2003); Henderson et al. (2012); Guerre et al. (2000); Laffont et al. (1995); Hendricks and Porter (2006); *etc.*.

[15] Skitmore et al. (2001); Ballesteros-Pérez and Skitmore (2017).

[16] De Silva et al. (2003) uses a similar construction for procurement auctions with both private and common value components.

[17] See Conclusion for a brief discussion about this assumption.

[18] The belief $\mu_{i,j}$ is "derived" because it is the log ratio between bidders' belief about μ_j and the benchmark value $\bar{v}_j$ unobservable to bidders, the latter of which may not factor into firm behavior.

without necessarily specifying any distributional parameters. They also have a firm-specific approximation of, or confidence in, the variance, $\lambda_i \sigma$, where $\lambda_i > 0$. Note that because μ_j is the normal counterpart of the mean of log-normal random variable $r_{i,j}$, and

$$E[r_{i,j}] = e^{\mu_j + \frac{\sigma^2}{2}} \neq e^{\mu_j} \tag{4}$$

As such, λ_i can considered the parameter of heterogeneous private value uncertainty, and $\mu_{i,j}$ a measure of common value uncertainty.

Proposition 1. If the firm overestimates σ by $\lambda_i > 1$, the firm also underestimates $\mu_{i,j}$ if it correctly observes $E[r_{i,j}]$.[19]

This is due to the increasing right-skewedness of log-normal distributions as variance increase. More precisely, the belief can be expressed analytically as

$$\mu_{i,j} = \mu_j + \frac{1}{2}(1 - \lambda_i^2)\sigma^2 \tag{5}$$

which is strictly decreasing in λ_i given that $\lambda_i > 0$. This is a casual prediction assuming that σ were the only source of mean private value, but it shows that common value uncertainty can arise out of private value uncertainty alone, in addition to other factors that may influence a firm's belief in μ_j and further confound firms' bidder behavior.

3.2 Bidder's problem

Based on its observations and beliefs, the firm then solves the following expected profit maximization problem, given the log-normal distribution of private values:

$$\max_{b_{i,j}} \mathbf{E}[\pi(b_{i,j})] = (b_{i,j} - r_{i,j}) \left[1 - \Phi\left(\frac{\ln r_{i,j} - \mu_{i,j}}{\lambda_i \sigma} \right) \right]^{n_j - 1} \tag{6}$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, $b_{i,j}$ is the relative bid compared to the engineer's estimate, and n_j is the number of bidders in project j[20]. The

[19] See Proof of Proposition 1.

[20] Harstad et al. (1990) and Levin and Ozdenoren (2004) provide some theoretical treatment of auctions with uncertain number of bidders, which is the case of CDOT procurement auctions. However, considering the principle question of heterogeneous uncertainty, unknown number of bidders is abstracted from the theoretical model, but it will be treated in the empirical strategy.

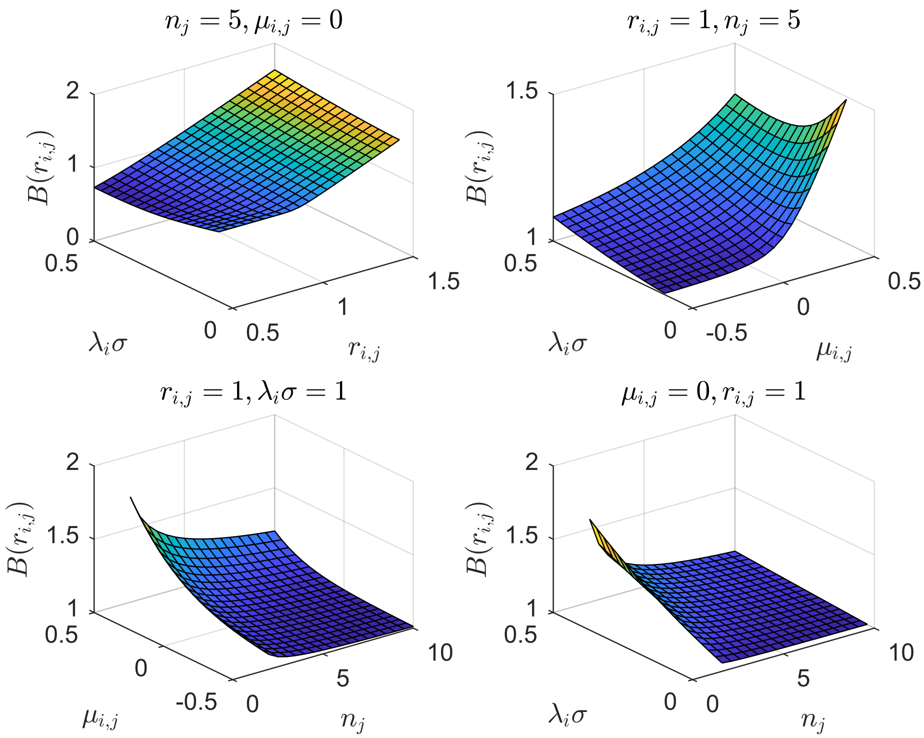

Figure 5: Numerical results of $B(r_{i,j})$ response to various changes in parameters.

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

$$B(r_{i,j}) = r_{i,j} + \frac{\int_{\ln r_{i,j}}^{\infty} [1 - \Phi(\frac{ln x - \mu_{i,j}}{\gamma_i \sigma})]^{n_j - 1} dx}{[1 - \Phi(\frac{\ln r_{i,j} - \mu_{i,j}}{\gamma_i \sigma})]^{n_j - 1}} \tag{7}$$

Figure 5 shows the the responses of $B(r_{i,j})$ to various changes in parameters. Note that $\mu_{i,j}$ and $\lambda_i \sigma$ 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 $\frac{\partial B(\cdot)}{\partial r_{i,j}} > 0$ as increasing monotonicity of $B(\cdot)$ in $r_{i,j}$ is a requirement for the solution, and unsurprisingly bids decrease with increasing number of bidders. Neither standard results are changed by the additions to the model, *ceteris paribus.*

Of particular interests are the bidding behavior in response to $\lambda_i \sigma$ and $\mu_{i,j}$. Bids generally increase in response to increases in the uncertainty spread, $\lambda_i \sigma$. This can be explained by

[21] See section A.2 Solution of the bidder's problem.

[22] The equilibrium strategy would be different if there is a binding reservation price. While CDOT has certain policies pertaining to reservation price expressed as a percentage above the engineer's estimate, the reservation price is not always binding and, more importantly, unknown to bidders, who do not observe either the engineer's estimate or the percentage threshold. As a result, reservation price is abstracted from the theoretical and empirical models.

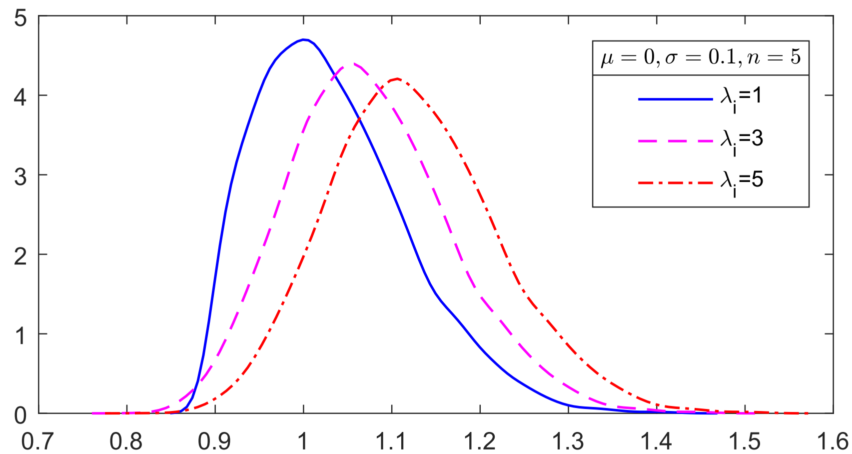

Figure 6: Simulated results of bid spread kernel density estimates with differing uncertainty from the same distribution of private values.

when holding belief in mean relative value constant, a higher spread flattens the distribution with a longer right tail and improves the probabilistic standing the firm, hence the firm bids more confidently. As a result, for any given μ and σ, higher λ_i leads to greater bid spread (Figure 6[23]).

Similarly, bids increase in response to increases in the belief about the mean relative value, $\mu_{i,j}$. This can be explained by that a higher belief in the mean relative value shifts the distribution to the right and increases the probabilistic standing of the firm, hence the firm bids more confidently as well. Because the model predicts that firms with higher belief in $\lambda_i\sigma$ also have lower belief in μ_j, the behavior of the firm can become confounded as the effect of uncertainty goes both ways.

One curious result is when $\lambda_i\sigma$ interacts with $\mu_{i,j}$ for a given realization of private value (Figure 5 top right panel), lower uncertainty can lead to higher bids. Note that for the particular outcome, the parameters are set $r_{i,j} = 1$, which is equivalent to $\ln(r_{i,j}) = 0$, and this result occurs only when $\mu_{i,j} > \ln r_{i,j}$. Intuitively, if a firm's private value is below the perceived the $\mu_{i,j}$, a smaller $\lambda_i\sigma$ shortens the spread around $\mu_{i,j}$ and increases the probabilistic standing of the firm, thereby increasing the firm's bid.

This result provides an interesting dynamic as to how private and common value uncertainties interact with firms' bidding behavior. The model suggests that uncertainly may both increase and decrease a firm's strategic bid through several channels, which can easily confound the empirical investigation of firm's bidding behavior. The model results

[23]The results are simulated since an analytical solution to the inverse bid function $v_i = B^{-1}(b_i)$ does not exist.

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]

$$r_{i,j,t} = e^{\beta_i + \mathbf{M}_t' \mathbf{B}_\mathbf{M}} \tag{8}$$

Where β_i is the firm fixed effect and $\mathbf{M}_t$ 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 role of firm size on uncertainty, the determinants of heterogeneity in private value is not of interest, and if present, firm-level scale and other idiosyncratic effects on private value should be absorbed by the firm fixed effect. The logged private value is therefore simply the logged private value variations with market factors plus firm fixed effect.

The private value is generously defined and may encompass any opportunity cost associated with submitting a bid, such as bid preparation cost, capacity constraint, and

[24] Period t is positively identified by project ID j and time subscript is omitted for all variables except the dependent variables and the market indicator.

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:

$$y_{i,j,t} = \underbrace{r_{i,j,t} + \frac{\int_{\ln r_{i,j,t}}^{\infty} \left\{1 - \Phi\left[\frac{x - s_i^{\beta_\mu}\mu_{j,t}}{\sigma_r s_i^{\beta_\sigma}}\right]\right\}^{n_j - 1} dx}{\left\{1 - \Phi\left[\frac{\ln r_{i,j,t} - s_i^{\beta_\mu}\mu_{j,t}}{\sigma_r s_i^{\beta_\sigma}}\right]\right\}^{n_j - 1}}}_{b_{i,j,t}} + \varepsilon_{i,j} \tag{9}$$

where $\sigma_r = [(I-1)^{-1}\sum_i \beta_i]^{\frac{1}{2}}$ is the standard deviation, and $\mu_r = I^{-1}\sum_i \beta_i$ 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 β_σ and common value heterogeneous uncertainty parameter β_μ with regard to firm size. Contrasting the optimal bidding function,$\sigma_r s_i^{\beta_\sigma}$ substitutes for $\lambda_i\sigma$ and $s_i^{\beta_\mu}\mu_{j,t}$ substitutes for $\mu_{i,j}$. A negative β_σ would support the hypothesis that smaller firms have greater private value uncertainty, and a non-zero β_μ would indicate heterogeneous common value uncertainty.

The structural model as presented cannot be directly identified with the conventional first moment conditions alone. There are special considerations for n_j, discussed in the following subsection, and for σ_r , discussed in section 4.2.1 Moment conditions.

4.1.2 Number of bidders

The data used in this study comes from a bidding process where bidders are blind to the number of other bidders and this issue is not addressed in the behavioral framework. While the number of bidders is known to the investigator, the fact that firms at the time of bidding cannot observe anything about their competitors poses an estimation challenge for two reasons. First, even though entry is perhaps always endogenous with the value of projects, with known number of bidders, firm's bidding behavior would be fully accounted for by the structural model, but because in this setting firms cannot reliably use the number of bidders to form their optimal strategy, the problem of endogeneity arises. While using the relative bid partially addresses this issue, endogeneity cannot be assumed away even with 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]:

$$\mathbf{E}[n|\mathbf{X_{IV}}] = e^{\mathbf{X}'_{\mathbf{IV}}\mathbf{B_{IV}}} \tag{10}$$

where $\mathbf{X}_{\text{IV}}$ is a vector of the determinants of entry including an intercept. The determinants are grouped into two categories: endogenous determinants and simultaneous determinants. Endogenous determinants include project value (in terms of engineer's estimate) and project type, and simultaneous determinants are market indicators and other covariates excluding the number of bidders. In effect, the determinants are the included variables and additional instruments similar to that of $\mathbf{Z}$ in a two-stage linear least squares model. A discussion of implementing instrumental variables in a nonlinear GMM model can be found in section 4.2.1 Moment conditions.

4.1.3 Additional considerations

Despite having a time subscript to include time-dependent variables, such as market conditions and seasonal dummies, the model is assumed to be cross-sectional and abstracts from any autoregressive processes such that $\varepsilon_{i,j}$ is time invariant. While it is difficult to model temporal autocorrelation due to the lack of a panel structure of the data, this is also a fair stylization given that the temporal variations in relative bids are well explained by market conditions, as well as the competitive nature of procurement auctions, it can be reasonably assumed that there is little endogenous mechanism to cause firm's bidding behavior to be autocorrelated outside of firm fixed effects and market conditions, both of which are accounted for by the model.

[25] The Poisson process is a reasonable assumption given the distribution of the number of bidders, which is a count variable with similar mean and variance. Negative binomial distribution may provide a slightly better fit given the small dispersion difference, but it requires additional parameters to be estimated.

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 $\mathbf{E}[\varepsilon_{i,j}|\mathbf{W}, \mathbf{\Theta_0}] = 0$, where $\mathbf{W} \in \mathbb{R}^{K+1}$ contains 1 containing dependent variable $\mathbf{Y}$, explanatory variables $\mathbf{X}$, 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 $\mathbf{W}$ only fully appear in $\varepsilon_{i,j}$, we define $\mathbf{Z}$ 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 $\mathbf{Z}$, $\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)$, is independent from the error term $\varepsilon_{i,j} = y_{i,j,t} - b_{i,j,t}$ such that

$$\mathbf{E}[(y_{i,j,t} - b_{i,j,t})^k|\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta), \Theta_0] = E[(y_{i,j,t} - b_{i,j,t})^k] = \mu_{\varepsilon,k} \tag{11}$$

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

$$\mathbf{E}[\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)(y_{i,j,t} - b_{i,j,t})^k|\Theta_0] = \mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)\mu_{\varepsilon,k} \tag{12}$$

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

$$g(\mathbf{W}_{i,j,t}, \Theta_0) = \mathbf{E}\left[\begin{array}{c} \mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)(y_{i,j,t} - b_{i,j,t}) \\ \mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)\{\sigma_y^2(s_i) - (y_{i,j,t} - b_{i,j,t})^2 - [\tilde{b}_{i,j,t} - \mu_y(s_i)]^2\} \\ \mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)(y_{i,j,t} - b_{i,j,t})^3 \end{array}\right]_{\Theta_0} = 0 \tag{13}$$

The first moment conditions are conventionally defined to assume that the error term has zero mean and independent from $\mathbf{Z}$. 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 $\sigma_y^2(s_i) = \sigma_b^2(s_i) + \sigma_\varepsilon^2$ 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.

$\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)$ can be a vector of any functions of $\mathbf{Z}_{i,j,t}$ to the extent that the model can still be identified, including simply the vector $\mathbf{Z}_{i,j,t}$. Therefore $g(\mathbf{W}_{i,j,t}, \Theta)$ 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:

$$\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta) = \frac{\nabla_\Theta b(\Theta|\mathbf{Z}_{i,j,t})}{\sigma_\varepsilon(\mathbf{Z}_{i,j,t})} = \frac{\partial b(\mathbf{Z}_{i,j,t}, \Theta)/\partial\theta}{\sigma_\varepsilon(\mathbf{Z}_{i,j,t})} \tag{14}$$

Where $\sigma_\varepsilon(\mathbf{W}_{i,j,t})$ is the heteroskedastic error dependent on $\mathbf{W}_{i,j,t}$ of an unknown form. While there are methods to approximate $\sigma_\varepsilon(\mathbf{W}_{i,j,t})$, 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 $\mathbf{h}(\cdot)$ therefore becomes

$$\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta) = \nabla_\Theta b(\Theta|\mathbf{Z}_{i,j,t}) = \frac{\partial b(\mathbf{Z}_{i,j,t}, \Theta)}{\partial\theta} \tag{15}$$

This formulation mirrors the first-order condition of nonlinear least squared (NLLS) regression in the first moment condition. The approach is computationally much more expensive than using $\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta) = \mathbf{Z}_{i,j,t}$, but it allows a much more flexible configuration. For the first order partial derivatives of $b(\cdot)$ that constitute $\mathbf{h}(\mathbf{Z}_{i,j,t}, \Theta)$ with respect to several

[26]See section A.4 Proof of proposition 2 for a sketch of proof.

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

4.2.2 Heteroskedasticity

In addition to the second central moment assumption $\mathbf{E}[(y_{i,j,t}-b_{i,j,t})^2|\mathbf{h}(\mathbf{Z}_{i,j,t},\Theta),\Theta_0]=Z\sigma_\varepsilon^2$, the second moment conditions also rely on two additional assumptions that

$$\mathbf{V}[y_{i,j,t}|s_i]=\mathbf{E}[(y-\mu_y(s_i))^2]=\mathbf{E}[(y-\mu_y(s_i))^2|\mathbf{h}(\mathbf{Z}_{i,j,t},\Theta),\Theta_0]=\sigma_y^2(s_i) \tag{16}$$

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

$$\mathbf{E}[b_{i,j}|\mathbf{h}(\mathbf{Z}_{i,j,t},\Theta),\Theta_0]=\sigma_b^2(\mathbf{Z}) \tag{17}$$

where σ_b^2 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 $\mathbf{Z}$ 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, $\sigma_y^2(s_i)$ is assumed to be the sub-population variance of all sub-population observations in the data, and it takes the form of and $\sigma_y^2(s_i)=\frac{1}{N}\sum(y-\bar{y}_{s_i})^2$ (*cf.* sample variance estimator $\hat{\sigma}_y^2(s_i)=\frac{1}{N-1}\sum(y-\bar{y}_{s_i})^2$). 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 $\sigma_y^2(s_i)$ 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.

For the parametric fit, following a similar formulation as the structural model, the nonlinear parametric estimation takes the form of

[27] In the current iteration of the working paper, only $\mathbf{h}=\mathbf{Z}$ is estimated. $\mathbf{h}=\nabla_\Theta b$ will be estimated in future explorations with access to high-performance computing environment. This can be made computationally feasible by subsample bootstrapping with distributed computing.

$$\sigma_y^2(s_i) = \underline{\sigma}_y^2 s_i^\delta + \varepsilon_{y,i} \tag{18}$$

where $\underline{\sigma}_y$ and δ can be estimated using non-linear least squared (NLLS) by using the sub-population variance for $\sigma_y^2(s_i)$. Alternatively, it can also be estimated by the Method of Moments (MM)[28], using the following first and second moment conditions by definition of mean and variance:

$$g_y(y_{i,j,t}, \Delta) = \mathbf{E}\left[\begin{array}{c} (y_{i,j,t} - \mu_y + \delta_\mu s_i) \\ \underline{\sigma}_y^2 s_i^{\delta_\sigma} - (y_{i,j,t} - \mu_s)^2 \end{array} \middle| \begin{bmatrix} 1 \\ s_i \end{bmatrix}\right]_{\Delta_0} = 0 \tag{19}$$

where μ_y, $\underline{\sigma}_y^2$, δ_μ and δ_σ are to be estimated, with $\underline{\sigma}_y^2$ and δ_σ being the same as the NLLS model. Both methods impose some assumptions about the structure of the heteroskedasticity, which is not known, and while in the parametric estimates they mimic that of the structural model, they cannot be reliably assumed to be similar. The third method to estimate non-parametrically[29] does not assume any structure to the heteroskedasticity of y:

$$\sigma_y(s_i) = m(s_i) + \varepsilon_{y,i} \tag{20}$$

However, because the underlying distribution of y is unknown[30], there is no clear criteria to bandwidth selection, which can affect the behavior of the fit. For the purpose of this study, the smallest bandwidth that yields a monotonically decreasing fit with respect to firm size is selected and compared to sub-population variance and the parametric fit.

Figure figure 7 shows the calculated and fitted cohort variance to these specifications. The fitted values yield similar results. For the main estimation results, the MM fit is used due to its evaluation from full sample, and other specifications are evaluated to check for robustness.

The model investigates the heteroskedasticity of the outcome variable in the dimension of firm size. While the moment conditions include heteroskedastic private values and assumes the error term is not heteroskedastic to firm size, it is entirely conceivable that heteroskedasticity still exists in other dimensions. Therefore, heteroskedasticity-consistent standard errors should still be used for inference. By the same construction, it is important

[28]Method of Moments is used because the model is just-identified with 4 parameters and 4 moment conditions.

[29]The Nadaraya-Watson estimator is easily implemented without incurring the curse of dimensionality since the regression is univariate (*c.f.* spline).

[30]The distribution of y is assumed to be that of the sum of two independent variables, $B(r, \cdot)$ and ε, where r is log-normally distributed but $B(r, \cdot)$ is not, and ε is only assumed to have zero mean, finite variance, and zero skewness from the moment conditions.

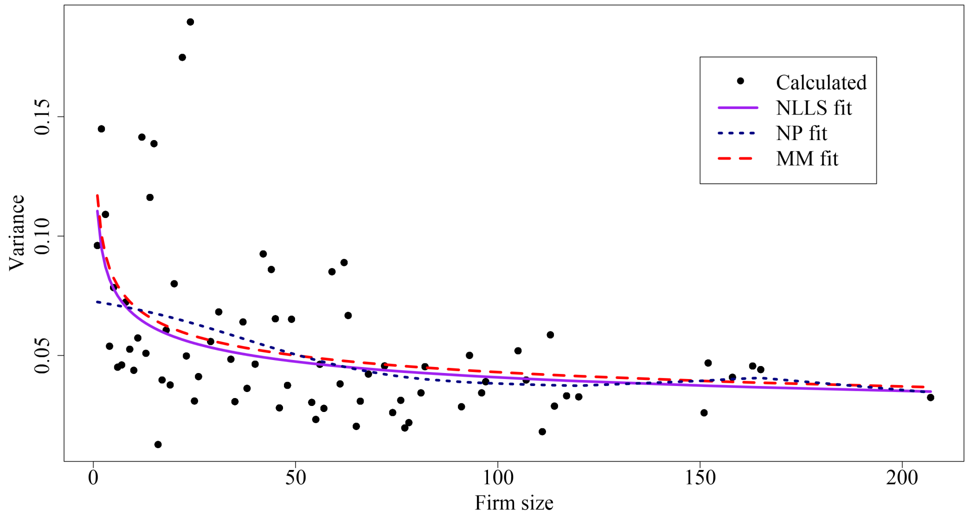

Figure 7: Cohort variance with fitted lines.

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 average:

$$\hat{g}(\mathbf{W}_{i,j,t}, \Theta) = M^{-1} \sum_{m_{i,j,t}} g(\mathbf{W}_{i,j,t}, \Theta) = 0 \tag{21}$$

where M is the number of observations and m indexes each observation. By minimizing a certain norm of the objective function

$$Q_M(\Theta) = \hat{g}(\mathbf{W}_{i,j,t}, \Theta)' \times w_M \times \hat{g}(\mathbf{W}_{i,j,t}, \Theta) \tag{22}$$

where w_M is a weighting matrix based on the sample, the GMM estimator is obtained as

$$\hat{\Theta} = \arg\min_{\Theta} ||Q_M(\Theta)|| \tag{23}$$

By convention, the Frobenius norm is chosen. At Θ_0, there exists a $3 \times p$ matrix G such that

$$G_0 = \text{plim} M^{-1} \sum_{M_{i,j,t}} \left[\frac{\partial g(\mathbf{W}_{i,j,t}, \Theta)}{\partial \theta'} \right]_{\Theta_0} \tag{24}$$

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)

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

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

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

While iterative estimator and continuously updating estimators (CUE) can afford better properties, due to the computational complexity of the estimation, the two-step estimator is used, which estimates the optimal w_L by using the identity weighting matrix in the first step and is asymptotically consistent and efficient.

4.3 Implementation

There are two obstacles to estimating the structural model directly. The obvious first is a computational one, where the large parameter space coupled with the large sample in a highly nonlinear objective function imposes not only a great computational cost, but also makes the estimation result too sensitive to inital values such that there is little guarantee of global identification being reached. The second obstacle comes with the matrix of fixed effects, in which many firms have sparse observations, giving rise to large swaths of zeros, and the fixed effects themselves have a narrow support near zero as the logged relative private value estimates. As a result, the issue of numerical invertibility arises. To exacerbate the issue, any time-invariant measure of firm size is necessarily perfectly collinear with the matrix of fixed effect dummies. Yet, the estimation of fixed effects cannot be circumvented because it is crucial to the estimation of private value uncertainty σ_r and to the minimization of heteroskedasticity with respect of firm size in the error term.

The second obstacle may be resolved using the optimal instrument, $\mathbf{h}(\mathbf{Z}) = \nabla_\theta b$, which introduces additional variations to the matrix of instruments. Alternatively, generalized inverse may be used to produce variance-covariance and weighting matrices when numerical singularity is encountered. However, neither method addresses the first obstacle of computational cost and global identification.

Instead of estimating private values directly through the structural equation, I use a semiparametric method based on the nonparametric identification of first-price auction proposed by Guerre et al. (2000) and subsequent papers. In a standard first-price sealed-bid auction, assuming equilibrium strategy and no reservation price, such that $b_i^* = \arg\max(v_i - b_i)\Pr(v_i < v_{-i}|n)$, a pseudosample of private value can be estimated using the following inverse bidding function:

$$\hat{v} = b + \frac{1}{n-1} \cdot \frac{\hat{F}(b)}{\hat{f}(b)} \tag{28}$$

where $\hat{F}(b)$ and $\hat{f}(b)$ are respectively the empirical cumulative distribution and estimated density of observed bids. The estimation equation does not require an analytical solution to the bidding function or even any specific assumptions on the underlying distribution of the private values provided that it is not degenerate. In the case of differing number of bidders and heterogeneous factors X, the joint distribution and density estimates of observed bids with the additional variables are used. Adjusted for reverse auction, the inverse bidding function becomes[31]

$$\hat{v} = b - \frac{1}{n-1} \cdot \frac{\hat{S}(b,n,X)}{\hat{f}(b,n,X)} \tag{29}$$

where $S(\cdot)$ is the joint survival function of bids, number of bidders, and other covariates. With the relative bid construction in the structural model, the additional covariates are the exogenous variables, namely market factors. Athey and Haile (2002) further propose that in the case of uncertain bidders, the number of bidders can be substituted with a public signal instead. Because of the multiple simultaneous factors that affect the number of bidders, I use the Poisson-fitted $\tilde{n}_{i,j,t}$ against the vector of instruments as a combined signal, as used in the second moment conditions, in place of the public signal.

The Guerre et al. (2000) paper also establishes the properties of the method using kernel estimators. However, given the sample size and the number of market factors, the direct application of equation (29) with kernel estimators is not feasible due to the curse of dimensionality. At the same time, the market factors cannot be omitted, as they change the distribution of private value in each period, whereas the nonparametric estimator requires

[31] See section A.5 for a sketch of proof.

private values to be identically distributed.

Krasnokutskaya (2011)[32] proposes a log-decomposition of bids if the effect of heterogeneity is multiplicative factor. This works well with the structural model, in which the private value is defined as equation (8). For any private value $r_{i,j,t}$, the bid can be expressed as $b_{i,j,t} = \rho(\tilde{n}_{i,j,t}, s_i) r_{i,j,t} = \rho(\tilde{n}_{i,j,t}, s_i) e^{\beta_i + \mathbf{M}_t' \mathbf{B_M}}$ such that

$$\ln b_{i,j,t} = \ln \rho(\tilde{n}_{i,j,t}, s_i) + \beta_i + \mathbf{M}_t' \mathbf{B_M} \tag{30}$$

Let $u_{i,j,t} = \ln \rho(\tilde{n}_{i,j,t}, s_i) + \beta_i$, $\mathbf{B_M}$ can be fitted as

$$\ln y_{i,j,t} = \mathbf{M}_t' \tilde{\mathbf{B}}_\mathbf{M} + \tilde{u}_{i,j,t} \tag{31}$$

which removes the correlation with the market factors from the bids. Conveniently, $\mathbf{M}_t' \tilde{\mathbf{B}}_\mathbf{M}$ is already the market variation in private value since the multiplicative factor for bidding markup, $\rho(\cdot)$, is additive in logged form, and $\mathbf{M}_t' \tilde{\mathbf{B}}_\mathbf{M}$ can be added directly back to the estimated $\tilde{\beta}_{i,j,t}$. The pseudosample of private values, including market variations, is therefore produced according to equation (29):

$$\begin{aligned}
\ln \hat{r}_{i,j,t} &= \tilde{\beta}_{i,j,t} + \mathbf{M}_t' \tilde{\mathbf{B}}_\mathbf{M} \\
&= \tilde{u}_{i,j,t} - \frac{\hat{S}(\tilde{u}_{i,j,t}, \tilde{n}_{i,j,t}, s_i)}{(1 - \tilde{n}_{i,j,t}) \hat{f}(\tilde{u}_{i,j,t}, \tilde{n}_{i,j,t}, s_i)} + \mathbf{M}_t' \tilde{\mathbf{B}}_\mathbf{M} \\
&= \ln y_{i,j,t} - \frac{\hat{S}(\tilde{u}_{i,j,t}, \tilde{n}_{i,j,t}, s_i)}{(1 - \tilde{n}_{i,j,t}) \hat{f}(\tilde{u}_{i,j,t}, \tilde{n}_{i,j,t}, s_i)}
\end{aligned} \tag{32}$$

The pseudosample is trimmed according to the method described in Guerre et al. (2000) due to bias of the kernel estimator at the boundaries of distributional support[33]. While the pseudosample is conventionally used to estimate the density of private values, I also use it to estimate the fixed effects as well as to reëstimate the market variation through a second fitting according to

[32] The paper specifically deals with unobserved heterogeneity, and the properties of the estimator are further explored in Armstrong (2013). This paper assumes that heterogeneity exists in the dimension of firm size and is therefore observed.

[33] Bids that fall within a hypercube within the conditional maximum and minimum bids are trimmed, with the side length of the hypercube defined by the estimation kernel support and bandwidth. In addition, under the symmetry assumption of logged private value, estimated private values lower than $\max(b)$ are also trimmed, with $\max(r) = \max(b)$ since the bidder with the highest private value is assumed to never bid with a markup under the standard model. In the structural equation, the upper limit, representing the upper bound of the value distribution, is still ∞ for simplicity, since at the highest bid they are numerically equivalent.

$$\ln \bar{r}_{i,j,t} = \bar{\beta}_i + \mathbf{M}_t^{\prime}\bar{\mathbf{B}}_{\mathbf{M}} +_r \bar{\varepsilon}_{i,j,t}$$

The reasons for the refitting are threefold. First, β_is still need to be estimated to produce μ_r and σ_r. Second, the refitting corrects some of the correlation between $\mathbf{M}_t$ and $\tilde{n}_{i,j,t}$ that would bias $\tilde{\mathbf{B}}_{\mathbf{M}}$ 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 $\bar{r}_{i,j,t}$s are estimated, they are plugged back into the structural model. Now the parameters that remain to be estimated are only β_σ and β_μ[34]. The fixed effect dummies are hereon dropped from $\mathbf{Z}$ for the main estimation, while $\mathbf{M}_t$ 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 $\mathbf{h}(\mathbf{Z})$ within a high-performance computing environment.

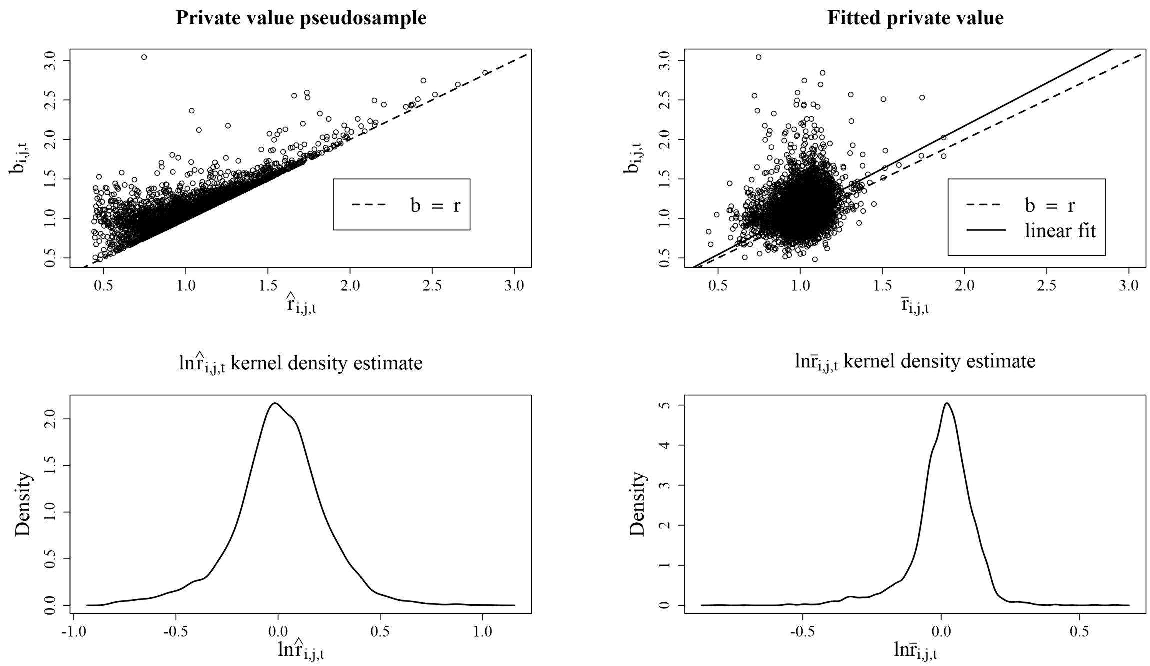

Figure 8: Estimated private value.

difference is absorbed into the error term, and the linear fit shows an average of 8.55% markup using the fitted values.

Figure 9 shows the estimated distribution of fixed effects fitted from the pseudosample. Counterintuitively, I find that larger firms 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 finding is not a surprise. Larger firms face more opportunity cost through at least two channels: first, the greater capacity of large firms bring about more opportunities within multiple markets, some of which may have better value; second, larger firms are also more likely to be closer to or exceeding capacity constraint since they have a revolving inventory of deliverables, whereas smaller firms 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 firms 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 $\beta_\mu = 0$. 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, β_σ, remains significant in all specifications, 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 effect tapers off quickly as firm size increases.

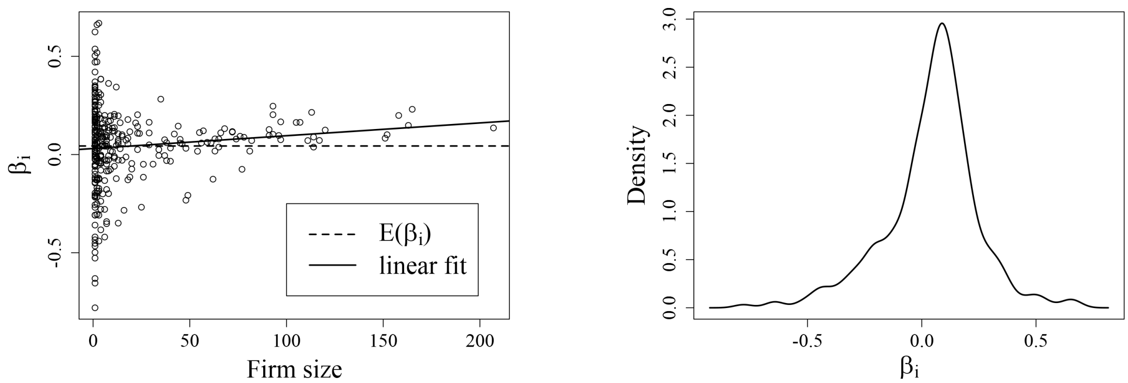

Figure 9: Estimated fixed effects.

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)
β_σ	-0.34789	-0.34581	-0.38698	-0.39061
	(0.00707)	(0.01201)	(0.00577)	(0.00964)
β_μ		0.02389		-0.13324
		(0.33126)		(0.21474)
	First-stage results			
	σ_r		0.20435	
	$\mathbf{E}(\beta_i)$		0.04346	
	Total observations		5682	
	Trimmed observations		340	

Table 4: Alternative specifications for common value uncertainty.

$\mu_{i,j,t}$ / Parameter	Model Specification: Benchmark 3(4)	$\mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M} + \mu_r s_i^{\beta_\mu}$ 4(1)	$\mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M} s_i^{\beta_\mu} + \mu_r$ 4(2)	$(\mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M} + \mu_r^E) s_i^{\beta_\mu}$ 4(3)
β_σ	-0.39061	-0.40044	-0.41830	-0.48726
	(0.00964)	(0.00578)	(0.02008)	(0.05631)
β_μ	-0.13324	2.39738	-0.34821	-0.39956
	(0.21474)	(4.69272)	(0.13468)	(0.21531)
μ_r^E				0.16712
				(0.03261)

The benchmark model 3(4) also shows greater common value uncertainty for small firms, especially with respect to fluctuations in market factors[35], although the result is not statistically significant and has an opposite sign from the estimate without IV. There is an inherent difficulty in identifying and interpreting common value uncertainty. Note that in the the construction from the structural model

$$\begin{aligned}\ln r_{i,j,t} - \mu_{i,j,t} &= \ln r_{i,j,t} - s_i^{\beta_\mu}\mu_{j,t} \\ &= \beta_i + \mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M} - s_i^{\beta_\mu}(\mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M} + \mu_r) \\ &= \beta_i + (1 - s_i^{\beta_\mu})\mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M} - s_i^{\beta_\mu}\mu_r \end{aligned} \tag{33}$$

A negative β_μ can therefore be interpreted both as more uncertainty about the effect of market fluctuation on mean private value and as less responsiveness to market fluctuations in determining the firm's own private value. Either interpretation suggest that smaller firms process market signals less reliably than bigger firms. Given the lack of statistical power of $\hat{\beta}_\mu$, a few alternative specifications of common value uncertainty are tested, shown in Table 4.

In the alternative specifications, both limiting β_μ to market factors only (Model 4(2)) and allowing the direct estimation of the private value fixed effect mean μ_r^E from the structural equation (Model 4(3)) not only increase the magnitude and significance of β_μ, but also

[35] Note that $\mathbf{M}_t'\bar{\mathbf{B}}_\mathbf{M}$ varies in sign.

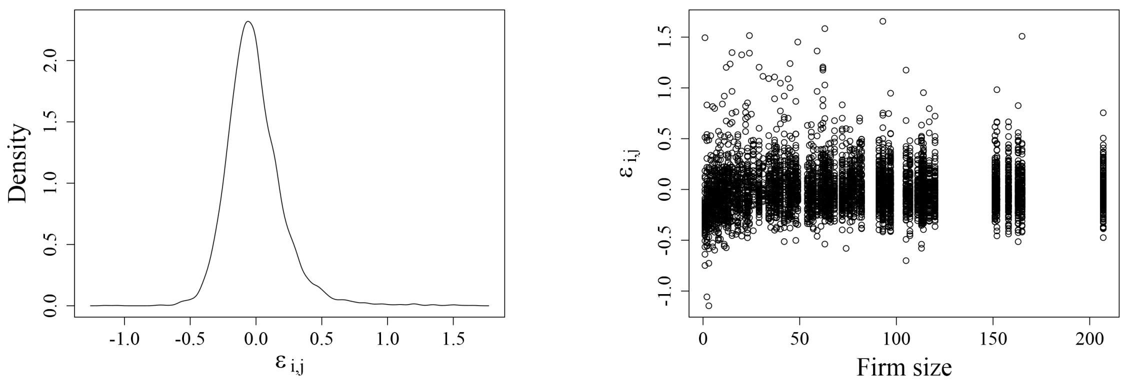

Figure 10: Estimated error term.

those of μ_σ. While Model 4(2) deviates from the theoretical definition 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 $s_i^{\beta_\mu}\mu_r$ by itself is more difficult to interpret; however, the result from Model 4(1), which only estimates the effect on firm size on the belief of μ_r, conforms to the casual prediction of Proposition 1 that smaller firms observe a lower $\mu_{i,j}$.

Figure 10 shows the estimated error term. The first and third moment conditions are well attained (left panel). For the second moment conditions, heteroskedasticity is mostly reduced except for the smallest firms (right panel)[36]. This is likely due to that the bidding behavior of small firms is not as well explained by the structural model as the larger firms, suggesting that smaller firms 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 firms' 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 average by firm size, instead of the average of the entire sample, may improve this result.

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 $\mathbf{Z}$, 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

$$\bar{n}_{i,j,t} = \xi(n_j, \tilde{n}_{i,j,t}, s_i|\mathbf{H}) \tag{34}$$

where $\mathbf{H} = \{\eta_1, \eta_2, \eta_3\}$ are the the parameters to be calibrated. The calibration uses the method of moments with the following moment conditions:

$$\mathbf{E}[y_{i,j,t} - \bar{b}_{i,j,t}|\tilde{n}_{i,j,t}, n_j - \tilde{n}_{i,j,t}, s_i]_{\mathbf{H}_0} = 0 \tag{35}$$

where $\tilde{n}_{i,j,t}$ is the combined signal for the number of bidders obtained from the Poisson regression used in $\tilde{b}_{i,j,t}$ of the second moment condition and in the nonparametric estimation, and $n_j - \tilde{n}_{i,j,t}$ 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 $n_j - \tilde{n}_{i,j,t}$ 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, particularly 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 the estimates of heterogeneous uncertainty, the effect is not large, which can be explained by the firms' good ability to anticipate bids. The effect of firm size is calibrated in Model 5(2), which finds that smaller firms interpret signals at a greater magnitude than larger firms. This does not necessarily mean that smaller firms anticipate the number of bidders better; rather, small firms tend to overestimate the expected number of bidders based on signals available to the investigator, and large firms tend to interpret more cautiously any additional signals that the investigator cannot observe. Coupled with the finding in the previous section that large firms tend to have higher private value yet bid about the same on average, this result is consistent with the conclusion of De Silva et al. (2003) that entrants tend to bid more aggressively (with less markup) than incumbents.

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

$\xi(n_j, \tilde{n}_{i,j,t})$ / Parameter	Model Specification $\eta_1\tilde{n}_{i,j,t} + \eta_2(n_{i,j} - \tilde{n}_{i,j,t})$ 5(1)	$[\eta_1\tilde{n}_{i,j,t} + \eta_2(n_{i,j} - \tilde{n}_{i,j,t})]s_i^{\eta_3}$ 5(2)
η_1	1.05400	0.80147
	(0.04464)	(0.05348)
η_2	0.7789	0.5289
	(0.1354)	(0.10642)
η_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 $\beta_\sigma, \beta_\mu = 0$ 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: Project value	Projects	Observed	Predicted 6(1)	$\beta_\sigma, \beta_\mu = 0$ 6(2)	Known n_j 6(3)	Both 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 differences between the observed and the predicted. In the known n_j scenario, while the average bid is still mostly higher, the difference 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 firm size. Table 8 shows the lowest bid share by size cohort[38]. Though there is a larger discrepancy between the observed and predicted shares for these estimates, compared to both scenarios, equalizing private and common value uncertainty leads to better allocation of projects to smaller firms, while removing n_j uncertainty has more mixed results.

In both the expenditure and the allocation scenarios, it is clear that equalizing private and common value uncertainty leads to better outcomes with respect to the government objectives of cost saving and affirmative policy. While this paper refrain from discussing any concrete and actionable strategy as to how this can be achieved, efforts undertaken by the government to reduce private and common value uncertainty for all firms will be in service of these objectives. Finally, although removing n_j uncertainty attains the opposite effect, it may

[38]Recall that the lowest bid is not necessarily the winning bid under the "best value" selection criteria of CDOT; however, it provides a good approximation of winner share.

Table 7: Average bid by project value.

Scenario: Project value	Projects	Observed	Predicted 7(1)	$\beta_\sigma, \beta_\mu = 0$ 7(2)	Known n_j 7(3)	Both 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: Firm size	Observed	Predicted 8(1)	$\beta_\sigma, \beta_\mu = 0$ 8(2)	Known n_j 8(3)	Both 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 structural modeling or econometric treatment, may yield more refinement to the paper's findings.

References

Y. An, Y. Hu, and M. Shum. Estimating first-price auctions with an unknown number of bidders: A misclassification approach. *Journal of Econometrics*, 157(2):328–341, 2010.

T. B. Armstrong. Bounds in auctions with unobserved heterogeneity. *Quantitative Economics*, 4(3):377–415, 2013. 3, 25

S. Athey and P. A. Haile. Identification of standard auction models. *Econometrica*, 70(6): 2107–2140, 2002. 3, 24

S. Athey and P. A. Haile. Empirical models of auctions. Technical report, National Bureau of Economic Research, 2006. 2

P. Bajari and A. Hortacsu. The winner's curse, reserve prices, and endogenous entry: Empirical insights from ebay auctions. *RAND Journal of Economics*, pages 329–355, 2003.

P. Bajari, R. McMillan, and S. Tadelis. Auctions versus negotiations in procurement: an empirical analysis. *The Journal of Law, Economics, & Organization*, 25(2):372–399, 2008. 3

J. Balat. Highway procurement and the stimulus package: Identification and estimation of dynamic auctions with unobserved heterogeneity. *Johns Hopkins University Mimeo*, 2013.

P. Ballesteros-Pérez and M. Skitmore. On the distribution of bids for construction contract auctions. *Construction Management and Economics*, 35(3):106–121, 2017. 11

S. Bikhchandani and J. G. Riley. Equilibria in open common value auctions. *Journal of Economic Theory*, 53(1):101–130, 1991.

L. Brannman, J. D. Klein, and L. W. Weiss. The price effects of increased competition in auction markets. *The review of Economics and Statistics*, pages 24–32, 1987.

A. C. Cameron and P. K. Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005. 2

Colorado Department of Transportation. Disadvantaged business enterprise program. URL `https://www.codot.gov/business/civilrights/dbe`. 3

Colorado Department of Transportation. Rules governing construction bidding for cdot public projects, 2014. URL `https://www.codot.gov/business/bidding/documents/rules-governing-construction-bidding-2-ccr-601-10`. 4

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. Laffont, 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 identification 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. Identification in auctions with selective entry. *Econometrica*, 82(1):315–344, 2014.

R. Gonçalves. 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.

E. Guerre, I. Perrigne, and Q. Vuong. Optimal nonparametric estimation of first-price auctions. *Econometrica*, 68(3):525–574, 2000. 3, 11, 24, 25, 46

P. A. Haile and E. Tamer. Inference with an incomplete model of english auctions. *Journal of Political Economy*, 111(1):1–51, 2003. 11

L. P. Hansen. Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, pages 1029–1054, 1982. 2, 22

R. M. Harstad. Alternative common-value auction procedures: Revenue comparisons with free entry. *Journal of Political Economy*, 98(2):421–429, 1990.

R. M. Harstad, J. H. Kagel, and D. Levin. Equilibrium bid functions for auctions with an uncertain number of bidders. *Economics Letters*, 33(1):35–40, 1990. 12, 32

T. Heinrich. Communication and reputation in procurement auctions - some empirical evidence. *Economics Letters*, 114(2):164–167, 2012.

D. J. Henderson, J. A. List, D. L. Millimet, C. F. Parmeter, and M. K. Price. Empirical implementation of nonparametric first-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 first-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.

D.-H. Kim. Flexible bayesian analysis of first price auctions using a simulated likelihood. *Quantitative Economics*, 6(2):429–461, 2015.

E. Krasnokutskaya. Identification and estimation of auction models with unobserved heterogeneity. *The Review of Economic Studies*, 78(1):293–327, 2011. 3, 25

J.-J. Laffont. Game theory and empirical economics: The case of auction data. *European Economic Review*, 41(1):1–35, 1997. 2

J.-J. Laffont, H. Ossard, and Q. Vuong. Econometrics of first-price auctions. *Econometrica: Journal of the Econometric Society*, pages 953–980, 1995. 2, 11

L. Lamy. The econometrics of auctions with asymmetric anonymous bidders. *Journal of Econometrics*, 167(1):113–132, 2012.

B. Lebrun. Existence of an equilibrium in first price auctions. *Economic Theory*, 7(3): 421–443, 1996. 10

D. Levin and E. Ozdenoren. Auctions with uncertain numbers of bidders. *Journal of Economic Theory*, 118(2):229–251, 2004. 12

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 affiliated private value auction model. *RAND Journal of Economics*, pages 171–193, 2002. 11

J. Marion. Are bid preferences benign? the effect 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' qualifications 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 μ and variance σ^2

$$\mu = \ln\left(\frac{m}{\sqrt{1+\frac{v}{m^2}}}\right) \tag{36}$$

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

μ can be rewritten as

$$\begin{aligned}\mu &= \ln m - \frac{1}{2}\ln(1+\frac{v}{m^2}) \\ &= \ln m - \frac{1}{2}\sigma^2\end{aligned} \tag{38}$$

Even if the bidder correctly observes m, a misoberved $\lambda_i\sigma$ results in

$$\mu_{i,j} = \ln m_j - \frac{1}{2}(\lambda_i\sigma)^2 \tag{39}$$

Given the reverse conversions from normal mean to log-normal mean

$$m = e^{\mu+\frac{\sigma^2}{2}}$$

we obtain

$$\begin{aligned}\mu_{i,j} &= \mu_j + \frac{1}{2}\sigma^2 - \frac{1}{2}(\lambda_i\sigma)^2 \\ &= \mu_j + \frac{1}{2}(1-\lambda_i^2)\sigma^2\end{aligned} \tag{40}$$

Given the functional form, it can be easily be shown that $\mu_{i,j}$ decreases in λ_i. ■

A.2 Solution of the bidder's problem

This is a brief description of the standard Envelope Theorem approach (Milgrom and Segal, 2002) for solving first-price auctions applied to reverse auctions. Each bidder i has a private 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,

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

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

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

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

$$\Pi_i(v_i) = \int_{\underline{v}}^{v_i} S(x)^{n-1}dx \tag{47}$$

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

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

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

$$B(r_{i,j}) = r_{i,j} + \frac{\int_{\ln r_{i,j}}^{\infty} [1 - \Phi(\frac{lnx - \mu_{i,j}}{\gamma_i \sigma})]^{n_j - 1} dx}{[1 - \Phi(\frac{\ln r_{i,j} - \mu_{i,j}}{\gamma_i \sigma})]^{n_j - 1}} \tag{50}$$

A.3 First-order derivatives

Given the estimation equation

$$y_{i,j,t} = \underbrace{r_{i,j,t} + \frac{\int_{\ln r_{i,j,t}}^{\infty} \left\{1 - \Phi\left[\frac{x - s_i^{\beta_\mu}\mu_{j,t}}{\sigma_r s_i^{\beta_\sigma}}\right]\right\}^{n_j - 1} dx}{\left\{1 - \Phi\left[\frac{\ln r_{i,j,t} - s_i^{\beta_\mu}\mu_{j,t}}{\sigma_r s_i^{\beta_\sigma}}\right]\right\}^{n_j - 1}}}_{b_{i,j,t}} + \varepsilon_{i,j} \tag{51}$$

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

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

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

β_i , we have

$$
\begin{aligned}
\frac{\partial b}{\partial \beta_i} &= \frac{\partial b}{\partial \beta_i} r(\beta_i) + \frac{\frac{\partial b}{\partial \beta_i}\int_{\ln r}^{\infty}[1-\Phi(x,\beta_i,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n-1}} - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{2(n-1)}}\frac{\partial b}{\partial \beta_i}[1-\Phi(\beta_i,\cdot)]^{n-1} \\
&= x_i r - \frac{\int_{\ln r}^{\infty}(n-1)\frac{x_i}{\sigma_r s^{\beta_\sigma}}\phi(x,\cdot)[1-\Phi(x,\cdot)]^{n-2}dx + x_i[1-\Phi(\cdot)]^{n-1}}{\Phi^{n-1}(\cdot)} \\
&\quad + \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n}}(n-1)\frac{x_i}{\sigma_r s^{\beta_\sigma}}\phi(\cdot) \qquad (53)
\end{aligned}
$$

For each β_m in $\mathbf{M}_t'\mathbf{B_M}$, we have

$$
\begin{aligned}
\frac{\partial b}{\partial \beta_m} &= \frac{\partial b}{\partial \beta_m} r(\beta_m) + \frac{\frac{\partial b}{\partial \beta_i}\int_{\ln r}^{\infty}[1-\Phi(x,\beta_m,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n-1}} - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{2(n-1)}}\frac{\partial b}{\partial \beta_m}[1-\Phi(\beta_m,\cdot)]^{n-1} \\
&= x_m r - \frac{\int_{\ln r}^{\infty}(n-1)(1-s^{\beta_\mu})\frac{x_m}{\sigma_r s^{\beta_\sigma}}\phi(x,\cdot)[1-\Phi(x,\cdot)]^{n-2}dx + x_m[1-\Phi(\cdot)]^{n-1}}{\Phi^{n-1}(\cdot)} \qquad (54) \\
&= + \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n}}(n-1)(1-s^{\beta_\mu})\frac{x_m}{\sigma_r s^{\beta_\sigma}}\phi(\cdot) \qquad (55)
\end{aligned}
$$

For the main parameters of interest,

$$
\begin{aligned}
\frac{\partial b}{\partial \beta_\sigma} &= \frac{\frac{\partial b}{\partial \beta_\sigma}\int_{\ln r}^{\infty}[1-\Phi(x,\beta_\sigma,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n-1}} - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{2(n-1)}}\frac{\partial b}{\partial \beta_\sigma}[1-\Phi(\beta_\sigma,\cdot)]^{n-1} \\
&= \frac{\int_{\ln r}^{\infty}(n-1)\sigma_r\beta_\sigma s^{\beta_\sigma-1}\frac{\ln r_i - s^{\beta_\mu}(\mathbf{M}_t'\mathbf{B_M}+\mu_r)}{\sigma_r^2 s^{2\beta_\sigma}}\phi(x,\cdot)[1-\Phi(x,\cdot)]^{n-2}dx}{\Phi^{N}(\cdot)} \\
&\quad - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n}}(n-1)\sigma_r\beta_\sigma s^{\beta_\sigma-1}\frac{\ln r_i - s^{\beta_\mu}(\mathbf{M}_t'\mathbf{B_M}+\mu_r)}{\sigma_r^2 s^{2\beta_\sigma}}\phi(\cdot) \qquad (56)
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial b}{\partial \beta_\mu} &= \frac{\frac{\partial b}{\partial \beta_\mu}\int_{\ln r}^{\infty}[1-\Phi(x,\beta_\mu,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n-1}} - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{2(n-1)}}\frac{\partial b}{\partial \beta_\mu}[1-\Phi(\beta_\mu,\cdot)]^{n-1} \\
&= \frac{\int_{\ln r}^{\infty}(n-1)\beta_\mu s^{\beta_\mu-1}\frac{\mathbf{M}_t'\mathbf{B_M}+\mu_r}{\sigma_r s^{\beta_\sigma}}\phi(x,\cdot)[1-\Phi(x,\cdot)]^{n-2}dx}{\Phi^{N}(\cdot)} \\
&\quad - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1}dx}{[1-\Phi(\cdot)]^{n}}(n-1)\beta_\mu s^{\beta_\mu-1}\frac{\mathbf{M}_t'\mathbf{B_M}+\mu_r}{\sigma_r s^{\beta_\sigma}}\phi(x,\cdot) \qquad (57)
\end{aligned}
$$

Finally, for the number of bidders decomposition estimation

$$
\begin{aligned}
\frac{\partial b}{\partial \eta} &= \frac{\int_{\ln r}^{\infty} \frac{\partial}{\partial \eta}[1-\Phi(x,\cdot)]^{n(\eta)-1} dx}{[1-\Phi(\cdot)]^{n-1}} - \frac{\int_{r}^{\infty}[1-\Phi(x,\cdot)]^{n-1} dx}{[1-\Phi(\cdot)]^{2(n-1)}} \frac{\partial}{\partial \eta}[1-\Phi(\cdot)]^{n(\eta)-1} \\
&= \frac{\int_{\ln r}^{\infty} x_{\eta} \ln[1-\Phi(x,\cdot)][1-\Phi]^{n-1}(x,\cdot) dx}{[1-\Phi(\cdot)]^{n-1}} \\
&\quad - \frac{\int_{\ln r}^{\infty}[1-\Phi(x,\cdot)]^{n-1} dx}{[1-\Phi(\cdot)]^{n-1}} x_{\eta} \ln[1-\Phi(\cdot)] \qquad (58)
\end{aligned}
$$

A.4 Proof of proposition 2

This sketch of proof shows that incorporating heterogeneous private value uncertainty in firm size alone in the first moment conditions does not remove heteroskedasticity from the error term. Let

$$
g(\gamma, \mu | W, \Theta_0) = \mathbf{E}[r_{i,j,t} + \frac{\int_{\ln r_{i,j,t}}^{\infty}[1-\Phi(\frac{lnx-\mu}{\sigma})]^{n_j-1} dx}{[1-\Phi(\frac{\ln r_{i,j,t}-\mu}{\sigma})]^{n_j-1}} - b_{i,j,t}] \qquad (59)
$$

and G represent the random variable generated by $g(\cdot)$. Assuming that $\mathbf{V}[\varepsilon|s_i, \Theta_0]$ is differentiable in s_i, $\mathbf{het}(s_i)$ defines a continuous measure of heteroskedasticity

$$
\begin{aligned}
\mathbf{het}(s_i) &= \frac{\partial}{\partial s_i} \mathbf{V}[\varepsilon|s_i, \Theta_0] \\
&= \frac{\partial}{\partial s_i} \int_G g^2(\gamma, \mu|W, \Theta_0) f(\varepsilon|s_i) dg \\
&= \int_G g^2(\gamma, \mu|\mathbf{W}, \Theta_0) \frac{\partial}{\partial s_i} f(\varepsilon|s_i) dg \qquad (60) \\
&\neq 0
\end{aligned}
$$

Now incorporate firm size within the moment condition

$$
g(\sigma(s_i), \mu(s_i)|\mathbf{W}, \Theta_0) = \mathbf{E}[r_{i,j,t} + \frac{\int_{\ln r_{i,j,t}}^{\infty}[1-\Phi(\frac{lnx-\mu(s_i)}{\sigma(s_i)})]^{n_j-1} dx}{[1-\Phi(\frac{\ln r_{i,j,t}-\mu(s_i)}{\sigma(s_i)})]^{n_j-1}} - b_{i,j,t}] \qquad (61)
$$

then,

$$\begin{aligned}
\frac{\partial}{\partial s_i}\mathbf{V}[\varepsilon|s_i,\Theta_0] &= \frac{\partial}{\partial s_i}\int_G [g^2(\sigma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)-\mu_G]f(\varepsilon|s_i)dg \\
&= \int_G \frac{\partial}{\partial s_i}\Big[g^2(\sigma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)f(\varepsilon|s_i)\Big]dg \\
&= \int_G f(\varepsilon|s_i)\frac{\partial}{\partial s_i}g^2(\sigma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)dg \\
&\quad + \int_G g^2(\sigma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)\frac{\partial}{\partial s_i}f(\varepsilon|s_i)dg \qquad (62)
\end{aligned}$$

Note that $\mathbf{het}(s_i) = \int_G g^2(\gamma,\mu|\mathbf{W},\Theta_0)\frac{\partial}{\partial s_i}f(\varepsilon|s_i)dg$, then

$$\frac{\partial}{\partial s_i}\mathbf{V}[\varepsilon|s_i,\Theta_0] = \int_G f(\varepsilon|s_i)\frac{\partial}{\partial s_i}g^2(\sigma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)dg + \mathbf{het}(s_i) \qquad (63)$$

Now,

$$\begin{aligned}
&\frac{\partial}{\partial s_i}g^2(\gamma(s_i),\mu(s_i)|\mathbf{W},\Theta_0) \\
=\ & 2g(\gamma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)\frac{\partial}{\partial s_i}g(\gamma(s_i),\mu(s_i)) \qquad (64)
\end{aligned}$$

Since $g(\gamma(s_i),\mu(s_i)|\mathbf{W},\Theta_0)=0$, we have

$$\frac{\partial}{\partial s_i}\mathbf{V}[\varepsilon|s_i,\Theta_0] = \mathbf{het}(s_i)\ \blacksquare \qquad (65)$$

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

$$\pi(v_i,b_i) = (b_i - v_i)S(B^{-1}(b_i))^{n-1} \qquad (66)$$

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

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

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

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