Loughborough University

This item was submitted to Loughborough's Research Repository by the author.
Items in Figshare are protected by copyright, with all rights reserved, unless otherwise indicated.

Improving the estimation of probability of bidder participation in procurement auctions

PLEASE CITE THE PUBLISHED VERSION

https://doi.org/10.1016/j.ijproman.2015.11.001

PUBLISHER

© 2015 Elsevier Ltd. APM and IPMA.

VERSION

AM (Accepted Manuscript)

PUBLISHER STATEMENT

This work is made available according to the conditions of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence. Full details of this licence are available at: https://creativecommons.org/licenses/by-nc-nd/4.0/

LICENCE

CC BY-NC-ND 4.0

REPOSITORY RECORD

Ballesteros-Perez, Pablo, Martin Skitmore, Eugenio Pellicer, and Jimmy H. Gutierrez. 2019. “Improving the Estimation of Probability of Bidder Participation in Procurement Auctions”. figshare. https://hdl.handle.net/2134/32755.

THANKS FOR DOWNLOADING THIS PAPER

This is a post-refereeing version of a manuscript published by Elsevier

Please, proper citation of the paper is:

Ballesteros-Pérez, P., Skitmore, M., Pellicer, E., Gutiérrez-Bahamondes, J.H. (2016). "Improving the estimation of probability of bidder participation in procurement auctions". International Journal of Project Management, 34(2), 158–172

http://dx.doi.org/10.1016/j.ijproman.2015.11.001

The authors recommend going to the publisher's website in order to access the full paper.

If this paper helped you somehow in your research, feel free to cite it.

This authors' version of the manuscript was downloaded totally free from:

https://www.researchgate.net/publication/267276284_Estimating_future_bidding_performance_of_competitor_bidders_in_capped_tenders/upload

IMPROVING THE ESTIMATION OF PROBABILITY OF BIDDER PARTICIPATION IN PROCUREMENT AUCTIONS

Authors:

Ballesteros-Pérez, Pablo [a] ; Martin Skitmore [b] ;

Pellicer, Eugenio [c] ; Gutiérrez-Bahamondes, Jimmy H. [d]

[a] Lecturer

School of Construction Management and Engineering

Whiteknights

Reading

RG6 6AW

United Kingdom

Email: p.ballesteros@reading.ac.uk; pabbalpe@hotmail.com

Phone: +44 (0) 118 378 8201 Fax: +44 (0) 118 931 3856

Corresponding author

[b] Professor of Construction Economics and Management

Room S711

School of Civil Engineering and the Built Environment

Queensland University of Technology

Gardens Point. Brisbane Q4001 Australia

Tel: +61 7 31381059 (w); +61 7 38933170 (A/H); 0450673028 (mob)

Email: rm.skitmore@qut.edu.au

http://staff.qut.edu.au/staff/skitmore/

[c] Associate Professor

Civil Engineer, M.Sc., Ph.D.

School of Civil Engineering. Universitat Politècnica de València

Camino de Vera s/n, 46022, Valencia, Spain

Phone: +34 963 879 562 Fax: +34 963 877 569

E-mail: pellicer@upv.es

[d] Industrial Engineer, M.Sc.

Facultad de Ingeniería. Universidad de Talca

Camino los Niches, km 1. Curicó (Chile)

Tel.: (+56) 9 99384535

E-mail: jimmygtrrz@gmail.com

IMPROVING THE ESTIMATION OF PROBABILITY OF BIDDER PARTICIPATION IN PROCUREMENT AUCTIONS

ABSTRACT

Anticipating the number and identity of bidders has significant influence in many theoretical results of the auction itself and bidders' bidding behaviour. This is because when a bidder knows in advance which specific bidders are likely competitors, this knowledge gives a company a head start when setting the bid price. However, despite these competitive implications, most previous studies have focused almost entirely on forecasting the number of bidders and only a few authors have dealt with the identity dimension qualitatively. Using a case study with immediate real-life applications, this paper develops a method for estimating every potential bidder's probability of participating in a future auction as a function of the tender economic size removing the bias caused by the contract size opportunities distribution. This way, a bidder or auctioner will be able to estimate the likelihood of a specific group of key, previously identified bidders in a future tender.

KEYWORDS: Bidding; Bidder identity; Auction; Tendering; Construction; Competitiveness.

CLASSIFICATIONS: Bidding (Contractors); Managing Proposal; Contract Procurement and Tendering (Clients); Engineering and Construction

1. Introduction

Studies of construction companies in the U.S. (Ahmad and Minkarah, 1988) and the UK (Shash, 1993) have shown one of the three most important factors that influence bidding decisions to be the likely number of bidders involved. This way, previous experience of the bidders' involvement in a series of past tenders for the same type of work and for the same Contracting Authority[1] provides qualitative knowledge about which firms regularly take part in those contracts, as well as the degree of competitiveness each regular bidder has demonstrated historically (Fu, 2004). Consequently, quantifying who and how many bidders are likely submit a future tender for a specific contract provides valuable information for a contractor when making the decision-to-bid (d2b); this is also a key factor in strategically setting a bid price to optimise the likelihood of winning the contract. There is also an economic impact at a business level, because being awarded more contracts decreases the overhead of the company in the short term, and increases its production and turnover in the long run, as well.

Likewise, there are several contributions that evidence interesting connections between the number of bidders N and tender outcomes. For instance, the degree of correlation between the average bid and the highest and lowest bids in collective bid tender forecasting models is higher on average when N is higher (Ballesteros-Pérez et al., 2012a, 2012b; Skitmore, 1981a). Similarly, the amplitude of the bid standard deviation was recently demonstrated to be also proportional to N (Ballesteros-Pérez et al., 2015a).

Most of the vast literature on the economic theory auctions also assumes the value of N to be known in symmetrical models, and at least the identity of all the bidders known in asymmetric models (e.g., Maskin and Riley, 2000). This is especially crucial for the theory of

[1] The terms Contracting Authority, client and owner are taken here to be synonymous.

common value auctions, where the value of *N* is a major determinant of the extent of the *winner's curse* (e.g., Capen et al., 1971).

Finally, another approach to this issue (Fu, 2004) tries to foresee the number and identities of bidders in practice through personal experience of the past appearance rate of bidders, mostly as a function of project features (client, type, size, location, etc).

Nonetheless, forecasting the number and identity of bidders is challenging, since no conclusive solution has yet been found for its accurate prediction, nor exists a suitable quantitative model to forecast the identities of a single or a group of specific key competitors likely to submit a future tender (Ballesteros-Pérez et al., 2015b). This is the gap of knowledge identified and, therefore, the point of departure of this research. Hence, the goal of this paper is to propose a useful quantitative model for forecasting, first, the identities of likely competing bidders in procurement auctions by means of calculating their respective probabilities of participation; and, second, the number of total contributing bidders involved. It builds on the previous work of Skitmore (1986), who introduced a simple model that assimilated probabilities of the bidders assuming them constant from tender to tender. However, the model proposed in this paper will take into consideration the contract economic size, removing the bias caused by uneven number of contract opportunities. The use of the model is explained and demonstrated by a case study with an actual construction tender dataset.

The paper is structured as follows. In the next section a brief *Literature Review* is provided first, which is followed by a *Model Outline* in the next section. A *Case Study* is then described in the fourth section, introducing an example tender dataset contained in the construction bidding literature, together with the calculations required to implement the model. Next, a *Results* section provides the case study bidders' probabilities of participating; and a validation subsection shows how closely the proposed approach depicts reality. Finally,

two *Discussion* and *Conclusions* sections comment on these results, highlighting where improvements are possible, posing unsolved research questions and presenting options for future work. For the sake of clarity, the American term "procurement (or reverse) auction" and the European "tender" are used here synonymously.

2. Literature review

As far back as 1956, Lawrence Friedman proposed several methods for estimating the average number of bidders in a future tender observing that, on many occasions, there is little information available to a firm about their competitors' intentions, but that it should still be possible to obtain a good estimation of the identity and number of future participating bidders if this firm is shrewd enough to combine such scarce information with its managers' experience (Friedman, 1956). This approach was restated by Rubey and Milner (1966) in a more broad tendering scenario with a specific emphasis on contract type.

Afterwards, another step taken by Friedman concerning the number of bidders was to suggest the existence of a relationship between the contract size (the complete budget to carry out the project) and the number of bidders N (Runeson and Skitmore, 1999). This was followed through by Gates (1967) and Wade and Harris (1976) but the results were weak. Other researchers had a similar experience, with Sugrue (1977), for example, failing to find a noteworthy relationship between the contract size and both the number of bidders and the number of suppliers and subcontractors involved. Also of interest at this point, is that Park (1966) had suggested ten years earlier the possibility of a non-linear relationship between the contract size and the number of bidders, but without any empirical support at the time.

Later, Skitmore (1981b) performed several tests with some international tender datasets from different time periods that helped identify a correlation between market conditions and the number of bidders but did not develop a mathematical model. Skitmore (1986) analysed

this phenomenon again with different data, confirming that the correlation between the market conditions and the number of bidders really existed. This relationship was weak or moderate in most cases; however, the correlation was significantly better when contract size was considered. More recently, Ballesteros-Pérez et al. (2015b) backed up this statement, underlining the importance of including the variable contract size in modelling *N* instead of simply treating *N* as a purely random variable.

In this regard, there have been many statistical distributions used so far to model the number of bidders *N*. Friedman (1956) suggested that *N* might follow a Poisson distribution, reasoning that if related individuals independently choose whether or not to bid for a particular object, this would be equivalent to the number of bids following a binomial distribution - which is known to be well approximated by the Poisson when the average of the number of bids is a small fraction of the total possible.

Several researchers have tried to test Friedman's conjecture. For example, Keller and Bor's (1978) analysis of the bidding configurations of a significant number of analogous construction contracts and their outcomes concurred with the Poisson assumption. However, others making use of U.S. Outer Continental Shelf Statistical summary data from 1976 to 1978, indicate that *N* might follow not only a distribution different to the Poisson, but may even be bimodal.

Friedman later hypothesized that the distribution of the residuals resulting from a regression of number of bidders with contract size would follow a Poisson distribution (Engelbrecht-Wiggans et al., 1986). Tests by Skitmore (1986) indicated that the normal distribution was a better choice no matter contract size is measured on either a natural or logarithmic scale.

Since then, many other distributions have been used to model *N*, such as: normal (Ballesteros-Pérez et al., 2014, 2013a), uniform (Ballesteros-Pérez et al., 2013b, 2015c), power law (Bolton and Ockenfels, 2014), gamma (Engelbrecht-Wiggans, 1980), Gaussian (Costantino et al., 2011), Laplace (Ballesteros-Pérez et al., 2015c) and Weibull (Ballesteros-Pérez and Skitmore, 2014). A compromise solution has also been to consider the number of bidders as a purely stochastic variable in experimental settings (McAfee and McMillan, 1987) or as a fixed value in game and auction theory (Harstad et al., 1990). The Poisson model, however, has continued to endure since the very first and celebrated compilations on auctions and bidding models from Stark and Rothkopf (1979) and Engelbrecht-Wiggans (1980) to modern and current online auctions (Bajari and Hortacsu, 2003). All in all though, the choice of model continues to be inconclusive.

Concerning the bidder identity dimension, over the years, bidders have been generally found to vary with project type (Drew and Skitmore, 2006), size (Al-Arjani, 2002; Benjamin, 1969; Drew and Skitmore, 2006), client (Bageis and Fortune, 2009), specific location (Azman, 2014) and, as already mentioned, with market conditions (Ngai et al., 2002; Skitmore, 1981b; Skitmore 1986). Nonetheless, Morin and Clough (1969) pointed out that it is usual that the same small bidders – and sometimes large companies – submit proposals for diverse types of work, taking into consideration that there are limitations to the number of contracts that can be managed by one bidder at any one time (Skitmore, 1988); this accounts for circumstances in which several bidders agree to bid or not in very similar tenders.

In this sense, Wade and Harris (1976) proposed to treat probabilistically the identities of bidders and their groups, whereas Shaffer and Micheau (1971) had briefly mentioned a predictive technique named the "*multidistribution model*" that allowed any contractor to foresee, with a high level of confidence, which competitors would participate in a specific tender. The particulars of this model seemed to be written down previously in Casey and

Shaffer (1964). Unfortunately, even though the authors read through this publication, no mention whatsoever was found about a model for forecasting bidders' identities.

Taking this into account, the tendering theory literature classifies potential competitors as "key" (e.g., Gates, 1967; Oo et al., 2010) and "strangers" (Skitmore, 1986), when trying to forecast the identities of several competitors, particularly the identities of those from whom little information is available. Oo et al. (2010) particularly, put forward a model to categorise "key" competitors considering four factors: project size, market, type of work and number of bidders. According to them, key contractors exhibit a higher level of competitiveness than other more sporadic competitors, the problem still being how to foresee which of these will take part in the next tender.

It is also necessary to mention the importance of the number and identity of bidders in auction and game theory (Klemperer, 2004). The interest in these variables has never been developed from a forecasting point of view, but driven by the diverse theoretical results it produces in several types of auctions formats and under different types of valuations. In auction and game theory, the main concern is the study of how different quantities or proportions of bidders with homogeneous or heterogeneous "personalities" (e.g. risk-averse, risk-neutral or risk-seeking bidders) will condition the results in the auction, while it is considered irrelevant exactly who they are. The underlying assumptions in most of these theoretical models, however, have seldom been checked empirically; some exceptions are: Athias and Nuñez (2009), Skitmore (2010) and Costantino et al. (2011). They will not be considered in this paper due to their wide-ranging different foci and interests; however, the interested reader can find an in-depth review of the effect of N over auction outcomes in Dyer et al. (1989), Levin and Ozdenoren (2004) and Hu (2011).

To sum up, as can be directly inferred by omission from the wide range of studies related to N, there is to date no quantitative model to specifically forecast the number and identities

of a single or a group of specific key competitors likely to enter a future tender. This was the knowledge gap highlighted in the Introduction and the main concern of the remainder of this paper. This study takes advantage of the previous work by Skitmore (1986) and develops a probabilistic framework for estimating and combining the bidders' probabilities of participation as a function of the future contract economic size.

3. Model outline

As stated in the Literature Review, a model for forecasting the identity of bidders in the first instance, and their total number in the second instance, is proposed based on Skitmore (1986). This suggests that the probability $Pr_j(i)$ that a bidder i bids for a tender j ($\boldsymbol{I}$ is a set of bidders with $i=1, 2, ..., n$; $\boldsymbol{J}$ is a set of analysed tenders with $j=1, 2, ..., C_d$) is given by:

$$\Pr_j(i) = \Pr(i) = \frac{C_i}{C_d} = R_i \tag{1}$$

where C_i is the *count* of the number of tenders in which bidder i has bid, and C_d is the *count* of the total number of tenders contained in an analysed *dataset*. As can be seen, there is no difference between the probabilities of bidding in different tenders j ($Pr_j(i) = Pr(i)$), since variables C_i and C_d remain constant for the same bidder i within the same dataset $\boldsymbol{J}$. From here on though, we will name this probability as the Participation Ratio of bidder i (R_i) to follow the more conventional nomenclature found in recent competitive bidding literature (Ballesteros-Pérez et al., 2014, 2015c).

Although we will challenge this initial assumption later, for the time being, we will assume that $Pr(i)$ represents an accurate approximation of $Pr_j(i)$ as equation (1) hypothesises, and that R_i equals $Pr(i)$. If this was true, many interesting and straightforward probabilistic results could be inferred whenever there is independency among bidders' R_i values. For

example, the probability that a subset of bidders $\boldsymbol{K \in I}$ (a given set of specific participating bidders) submits a bid in a future tender would be:

$$\Pr(K) = \prod_{i \in K} R_i \tag{2}$$

The probability that a subset of bidders $\boldsymbol{K \in I}$ submits a bid while other subset of bidders $\boldsymbol{K' \in I}$ do not (with $\boldsymbol{K \cap K'} = \emptyset$, because the same bidder *i* can either participate or not in one tender) would be:

$$\Pr(K, \overline{K}') = \prod_{i \in K} R_i \prod_{i \in K'} (1 - R_i) = \Pr(K) \cdot \prod_{i \in K'} (1 - R_i) \tag{3}$$

and, if the number of potential bidders (those identified and counted, because they have submitted at least one bid) is *n*, the number of possible combinations of bidders participating in a tender C_n would be:

$$C_n = \sum_{N=0}^{n} \binom{n}{N} = \sum_{N=0}^{n} \frac{n!}{N!(n-N)!} = 2^n \tag{4}$$

where $\binom{n}{N}$ is the combinatorial number of *n* elements (total number of bidders identified) taken in subsets of *N* elements (participating bidders). For the combinations (scenarios) calculated in (4) it is worth noting that their respective probabilities can be easily calculated one by one by means of expression (3), that is, it is possible to know in advance how likely it is that each combination of participating (***K***) and not participating (***K'***) bidders takes place. Of these, two of the most interesting scenarios would be, perhaps: the most likely combination of bidders, i.e., the highest probable combination of participating bidders:

$$\begin{aligned} &\text{Maximum } \Pr(K, \overline{K}' \mid K, K' \in I) = \prod_{i \in I} \max(R_i, 1 - R_i) \\ &\forall i \in K \text{ if } \max(R_i, 1 - R_i) = R_i \\ &\forall i \in K' \text{ if } \max(R_i, 1 - R_i) = 1 - R_i \end{aligned} \tag{5}$$

and the less likely combination of participating bidders, i.e., the scenario with lowest probability of happening:

$$\begin{aligned}&\text{Minimum } \Pr\left(K,\overline{K}'\middle|\, K,K'\in I\right)=\prod_{i\in I}\min\left(R_i,1-R_i\right)\\&\forall i\in K \ \text{ if } \ \min\left(R_i,1-R_i\right)=R_i\\&\forall i\in K' \ \text{ if } \ \min\left(R_i,1-R_i\right)=1-R_i\end{aligned} \tag{6}$$

And, finally, it is also possible to determine the probabilities that the total number of participating bidders $\boldsymbol{K}\in\boldsymbol{I}$, whose sum of elements $|\boldsymbol{K}|$, equals a given number N, as:

$$\Pr\left(\left|K\right|=N\right)=\sum_{|Q|=1}^{\binom{n}{N}}\left[\prod_{i\in K_Q}R_i\prod_{i\in K_Q{}^c}\left(1-R_q\right)\right] \tag{7}$$

$\boldsymbol{Q}$ being the set of all possible $\boldsymbol{K_Q}$ subsets with size N (with $\boldsymbol{K_Q}\cup\boldsymbol{K_Q}{}^c=I$ and $\boldsymbol{K_Q}\cap\boldsymbol{K_Q}{}^c=\varnothing$). Expression (7) is useful for determining how likely it is that N bidders submit a bid (with $N=1,2,...n$) in a future tender.

To sum up, if each bidder i's individual probabilities of submitting a bid for tender j, that is $Pr_j(i)$, were known, forecasting how likely it is that any combination of known and potential bidders happens by means of Equation 2 would be straightforward, along with other statistical information offered by equations (3)-(7). For the sake of clarity, a numerical example is proposed to show the implementation of these equations into real practice. Normally, there are many potential bidders, but for the sake of simplicity we will assume a situation in which there are just four bidders ($i=1,2,3,4$) that participated one, two, five and ten times respectively ($C_1=1$, $C_2=2$, $C_3=5$ and $C_4=10$) in our database of tenders which comprises 10 tenders in total ($C_d=10$).

With these data and equation (1), it is possible to calculate the participation ratios from these four bidders as follows:

$$R_1=C_1/C_d=1/10=0.1 \qquad R_2=C_2/C_d=2/10=0.2$$

$$R_3=C_3/C_d=5/10=0.5 \qquad R_4=C_4/C_d=10/10=1.0$$

Now, it would be useful to estimate, for example, the probabilities of having a tender in which bidders 1, 2 and 4 participated. By means of expression (2) this calculation is very

simple, since it is the product of the participation ratios of those bidders who participate (subset ***K***). In our case Pr*(K=1, 2, 4)=$R_1 \cdot R_2 \cdot R_4$=0.1·0.2·1.0=0.02*.

As can be seen, the probability of finding a tender with bidders 1, 2 and 4 submitting a bid would be of 2%, but, what about bidder 3? The former probability encompasses the situations in which bidder 3 enters a bid and when he does not. If we wanted to estimate the probabilities that bidders 1,2 and 4 participated (subset ***K****=1, 2, 4*) whereas bidder 3 explicitly did "not" participate (subset ***K'****=bidder 3*) then, we should resort to equation (3), which is Pr*(K=1, 2, 4, K'=3)=($R_1 \cdot R_2 \cdot R_4$)·(1-R_3)=(0.1·0.2·1.0)·(1-0.5)=0.01.*

Of course, the situation in which bidder 3 participates along with bidders 1, 2 and 4 amounts to 1% as well, since it is complementary to the situation calculated above.

However, as can be seen, the number of different combinations of bidders participating and not participating might be quite large, even when the number of bidders is low, such as in our example where there are just four bidders (*n=4*). This generally forces a series of calculations in which the probabilities associated to each scenario are quite small. In this connection, expression (4) is straightforward and allows us to calculate the different number of participating bidder combinations when *n* is set by means of $C_{n=4}=2^n=2^4=16$.

But, out of these 16 scenarios or combinations of bidders, probably the most significant might be the most likely and the most unlikely combination of participating versus not participating bidders. For this, we have expressions (5) and (6) respectively. Expression (5) requires the maximum values between the pair "R_i" and "*1-*R_i" for each bidder *i*, whereas expression (6) takes advantage of the minimum values of these pairs. Both expressions multiply those maxima, or minima, respectively.

In our example, the maximum values for bidders 1, 2, 3 and 4 are, respectively: *1-0.1=0.9* (bidder 1 not participates); *1-0.2=0.8* (bidder 2 not participates); *0.5* or *1-0.5* indistinctly (bidder 3 has the same probability of participating than not participating), and 1.0 (bidder 4

does participate). Therefore, the most likely combination[2], according to expression (5) would correspond to this probability $0.9 \cdot 0.8 \cdot 0.5 \cdot 1.0 = 0.36$.

On the other hand, the least likely tender requires the minima of the pairs "R_i" and "$1\text{-}R_i$" for each bidder, which in this case are the following: *0.1* (bidder 1 participates); *0.2* (bidder 2 participates); *0.5* or *1-0.5* indistinctly (bidder 3 has again the same probability of participating than not participating), and 0.0 (bidder 4 not participates). And by means of expression (6) the least likely scenario corresponds to $0.1 \cdot 0.2 \cdot 0.5 \cdot 0.0 = 0.0$, which actually cannot occur since bidder 4 always participates. It must be emphasised nevertheless, that despite that there are other combinations with a probability of zero whenever bidder 4 does not participate, the real value of equations (5) and (6) is the identification of what is the more likely outcome for each single bidder. For example, thanks to equation (5) we can say that the most likely outcome is that bidders 1 and 2 will not participate, and vice versa. In other words, the most unlikely scenario is that bidders 1 and 2 will submit a bid for the next tender, independent of the global probabilities, due to the presence of one specific bidder with a R_i or $1\text{-}R_i$ value of 0 or 1, conditioning the overall probability result.

Finally, expression (7) calculates the probabilities that any total given number of participating bidders (no matter who they are exactly) takes place. For example, if we wanted to obtain the probabilities of all the combinations (tenders) with $N=2$ participating bidders, this calculation would entail the sum of multiple products in which two bidders participate (R_i values) and the other two does not ($1\text{-}R_i$ values). The combinations would be in this case:

$$\Pr(N=2) = (R_1 \cdot R_2 \cdot (1-R_3) \cdot (1-R_4)) + (R_1 \cdot (1-R_2) \cdot R_3 \cdot (1-R_4)) + (R_1 \cdot (1-R_2) \cdot (1-R_3) \cdot R_4) +$$

$$+ ((1-R_1) \cdot R_2 \cdot R_3 \cdot (1-R_4)) + ((1-R_1) \cdot R_2 \cdot (1-R_3) \cdot R_4) + ((1-R_1) \cdot (1-R_2) \cdot R_3 \cdot R_4) = 0.49$$

[2] In this case there are two equally likely scenarios actually, each of which has a probability of occurrence of 0.36, and their difference relies just in whether bidder 3 participates or not.

As can be seen, although the calculation burden would be higher (but not more complex) as the number of bidders increases, the equations proposed allows obtaining very useful information for real-life bidding. The problem is, however, that these expressions rely on a series of individual bidders' probability of participation values (the R_i values) whose accuracy is critical for providing reliable estimations afterwards. Hence, this is why the proposed model will focus on improving the estimation of the $Pr_j(i)$ values (as Figure 2 will show later as a function of the contract size), as they allow the identity of future participating bidders to be forecasted in multiple facets like the ones enabled by equations (1) to (7).

A bigger problem, however, concerns the hypothesis regarding the independence assumption of the R_i values, which subsequently involves the independence of the $Pr_j(i)$ values. In general, the independence assumption is, of course, unlikely to hold within a tender dataset as it is generally believed that the same bidders frequently bid for the same type, size, client, location, etc. of tenders. An approach would be to estimate the covariance matrices of all combinations of R_i values taken by groups of two, three, … , n. However, it is not likely that sufficient data will be available for any database due to the extreme sparseness (generally over 90%) of the bidder-bidder matrix (Skitmore, 2013).

An alternative approach is to estimate the R_i values as a function of the tender j size (in this case $R_i \neq R_{ij}$). This can be done because, as it is generally considered that certain bidders are associated with certain tender or project characteristics (Fu, 2004), this should go some way towards removing the interdependencies amongst bidders. Therefore, this approach is the one addressed by the model proposed in this paper as it aims to calculate all the R_{ij} values for the n known and identified bidders, i.e. not assume them as constant ($R_i = R_{ij}$ as expression (1) proposed), but varying with the next tender economic size. This improvement will also require, of course, changing the R_i values in equations (2), (3), (5)-(7) with the new R_{ij} values

as a function of the tender *j* size. A full explanation of this procedure is provided in the following section, using a case study as an example, in order to clarify its implementation.

We acknowledge, however, that the model neglects other tender attributes, such as the nature of the work (type), client and location. This can be addressed by restricting the dataset to pooled tenders with similar characteristics to the target tender along the lines of Yeung and Skitmore (2012).

4. Case study

4.1. Tender dataset

A case study is used to illustrate the calculations of the bidders' Participation Ratios as a function of the tender *j* economic size (R_{ij}) from Fu's (2004) real dataset of 47 Hong Kong construction contracts awarded by open tender during the period 1991-1996 reproduced in Table 1. It is worth noting that, despite this complete dataset encompasses a larger number of contracts, only 47 out of the total 265 comprising the same type of works have been used. Besides, despite being nearly twenty years old, its positive attributes for this case study are that: (a) it is a small size, that allows the complete calculations to be shown within the usual space limitations of a journal paper; (b) it has already been published and still fully accessible, which removes any doubts of its veracity; (c) it contains the engineer's cost estimates for the 47 tenders. Finally, it should be noted that such tender databases are very rare in the literature, much less containing the bidders' identities and cost estimates, leaving few other options available.

<Table 1>

Table 1 contains several blocks of information. The first two columns provide the tender identification (*j*) and the number of bidders (*N*) that took part in each tender *j*; this can also be obtained by summing the number of bidders whose numeric ID is given in the last seventeen

columns on the right. The second block (third and fourth columns) provides the average bid of all bids submitted by the bidders who participated (whose bids are not presented due to lack of space) in natural (B_m) and log ($LN\,B_m$) scales. The third block (fifth column) contains the cost estimates of the engineer which serve as pre-tender estimates and which are noted as B_o.

It is also noted that, even though it is not strictly necessary, the original bidders' IDs have been recoded giving them a number in decreasing order according to the total number of bids each of them eventually submitted. For example, bidder 1 participated in 27 of the 47 tenders, while the 39 bidders from 62 to 100, participated only once (once-bidders). This allows a clear separation between frequent bidders (who bid at least twice) from once-bidders and the more intuitive arrangement of subsequent calculations and tables.

4.2. Calculations

The objective is to reach an expression of the R_{ij} values that takes into account the tender j economic size because with these, it is possible to calculate useful statistical information about a single or a group of bidders' participation as exemplified by equations (1) to (7). In order to do so, we begin by analysing expression (1), whose ratio C_i/C_d has been named *biased Participation Ratio*, R_i.

The problem or limitation with the biased Participation Ratio is that expression (1) does not take this into account the actual bidders' C_i values being conditioned by the contract size (budget) opportunities distribution for which those bidders can submit bids. For example, imagine two bidders - A and B. Bidder A prefers to take part in tenders with big budgets with R_A=0.5 (participates in half of those big contracts), whereas bidder B shows preference for smaller contracts with the same R_B=0.5. Now, assume that a series of tenders are open for bidding, being mostly of these of small economic size. What will be observed is that bidder B participates more frequently than bidder A ($R_B>R_A$) because $C_B>C_A$ for the same C_d. If there

were a majority of big contracts on the other hand, the perception of the bidders A and B's participation ratios would be exactly the opposite, that is, we would think that $R_A > R_B$. This is the reason the ratios are called *biased*, since they strongly depend on the distribution of contract size opportunities. There is only one exception: when the number of contract size opportunities follows a perfect uniform distribution, that is, when there are the same opportunities for every size of contracts for the bidders to submit their bids. However, the distribution of contract size opportunities has been widely studied and has long been known to resemble a log-normal distribution (Skitmore, 1986), distorting the bidders' participation ratios by augmenting the R_i values of bidders' that prefer to participate in tenders whose economic size is near the average of the log-normal distribution, and diminishing the R_i ratios of bidders that prefer to take part in tenders located far from the distribution average.

The first step is to correct the R_i values, that is to obtain the actual (*unbiased*) R_{ij} values, which requires resorting nonetheless to the biased R_i values, whose calculations are presented on the upper block of Table 2. As can be easily seen, this basically retrieves all the bidders' mean bid values (in log scale) and presents them ordered from lower to higher. On the left, the same calculations are reproduced for the complete dataset. The counts of all bidders' bids (C_i), the ~~50~~ 47 tenders (C_d), as well as the *biased participation ratios* (R_i) are presented on the lowest two lines of the upper block of Table 2.

<Table 2>

The intermediate lower block in Table 2 is devoted to calculating the mean (μ_i) and standard deviation (σ_i) of the series of each bidder *i*'s $LN\ B_m$ values by which, every bidder *i*'s log-normal distribution can be defined. Also, on the left, the same calculations are performed for the whole database (μ_d and σ_d values).

Finally, the lowest block in Table 2 shows a series of Anderson-Darling tests to check whether the log-normal distributions initially assumed represent a good fit. The results show

that, with just one exception out of the 62 cases (which can be attributed to a type 1 error), the log-normal distribution cannot be rejected at the 5% standard level of significance.

Figure 1 represents the actual cumulative distribution functions (CDF) of the five most frequent bidders in the database along with the complete dataset CDF values. The X values in Figure 1 correspond to the $LN\,B_m$ values shown in the upper block of Table 2, whereas the Y values (probabilities) have been calculated by the expression $(Order\text{-}0.5)/Order_{max}$ with $Order_{max} = C_i$ (for the five bidders) or C_d (for the complete dataset), respectively.

<Figure 1>

It has been commented that the distribution of the contract size opportunities (in this case, the second column in Table 2 and the thickest curve in Figure 1) usually follows a log-normal distribution. However, by looking at all the bidders' log-normal distributions in Figure 1 and Table 1, a reasonable question that arises is whether the log-normal distributions are the consequence of either the database log-normal filter or that the bidders' distributions are genuinely log-normal. This question cannot be answered directly, because it is not possible to remove the underlying effect of the complete dataset distribution of contract sizes from each bidder's tender participation curve.

One way to investigate this is by analysing other databases. In this regard, Table 3 shows a series of twelve international construction and services tender databases that have been divided into two vertical divisions. The upper three are databases in which at least one bidder participated in almost all tenders. The last eight databases (ID from 4 to 12, the seventh of which is the one used partially in this case study) are all databases gathered from a single client or contracting authority in which all the tenders (not only those in which the same bidder took part) are retrieved.

<Table 3>

The aim of Table 3 is to determine if, no matter the database is complete or conditioned by the participation preference of one bidder, the distribution of every bidder's contract size participation can be well modelled by a log-normal distribution. The series of pass (✓) values of the Kolmogorov-Smirnov tests performed on the twelve databases evidences this to be indeed the case. Therefore, the biased values behave as log-normal no matter whether the underlying distribution is originally log-normal or not. Hence, the *unbiased* participation ratios R_{ij}, as a function of the tender j size, can be calculated for every bidder i as:

$$R_{ij} = f\left(B_{mj}, e^{\mu_i}, e^{\sigma_i}, C_i, e^{\mu_d}, e^{\sigma_d}, C_d\right) = \frac{LogNormal\,(B_{mj}, e^{\mu_i}, e^{\sigma_i})}{LogNormal\,(B_{mj}, e^{\mu_d}, e^{\sigma_d})} \cdot \frac{C_i}{C_d} \tag{8}$$

Unlike the *biased* R_i ratios, these are free from distortion by the contract size opportunities distribution, since the bidder i's log-normal distribution in absolute values has been taken (this is why it is multiplied by C_i in the numerator, that is, the number of tenders in which bidder i is willing to participate depending on the contract size) and divided into the number of tenders (not only probability density, and that is why it is also multiplied by C_d in the denominator) that are available in each tender size range. A similar result could have been obtained by dividing the tenders into ranges of different tender sizes and calculating an R_i value for each range, but that approach would have left the domain outside the minimum and maximum contract size with no information.

Expression (8) can be simplified by working with log values to

$$R_{ij} = f\left(LN\ B_{mj}, \mu_i, \sigma_i, C_i, \mu_d, \sigma_d, C_d\right) = \frac{Normal\,(LN\ B_{mj}, \mu_i, \sigma_i)}{Normal\,(LN\ B_{mj}, \mu_d, \sigma_d)} \cdot \frac{C_i}{C_d} \tag{9}$$

this being the expression that has been used in the case study. For the unaware reader, the ratio of two normal distributions cannot be reduced further (and not always is another log-normal distribution), unless the distributions in both the numerator and the denominator are standard normal distributions, which would result in the Cauchy distribution.

5. Results

5.1. Bidders' unbiased participation ratios

If equation (9) is applied to all bidders' values, since variables μ_i, σ_i, C_i, μ_d, σ_d, C_d were all calculated in Table 2 for all bidders who submitted at least five bids, for the sake of brevity, and the complete database respectively, the result would be a ratio of two normal distributions that can be transformed into a figure as soon as variable B_m takes on a particular value. Therefore, Figure 2 represents the particularisation of expression (9) for the five most frequent bidders.

<Figure 2>

As can be seen, curves represented in Figure 2 differ greatly from a constant value as equation (1) initially proposed, and that is why these ratios are expected to be far more close to reality. However, there is one exception to the rule: when a bidder *i* has submitted just one bid, the standard deviation σ_i of their log-normal distributions cannot be calculated, therefore equation (9) for calculating the unbiased participation ratios (R_{ij}) cannot be evaluated due to insufficient data. In those cases, the biased participation ratios ($R_{i,}$), which are constant irrespective of the contract size by means of expression (1), constitute the only valid alternative. As noted, this happens when a bidder has only submitted one bid (once-bidder), which tends to be a big proportion of bidders within most databases. Hence, in the case study, bidders 62 to 100 were once-bidders and in those cases all of them are assumed to have the same participation ratio as all of them participated indistinctly once out of the all auctions of the dataset, which numerically is equivalent to R_{62j}= ...=$R_{100\,j}$= R_{62}=...= R_{100}=1/C_d=1/47= =0.0213.

Finally, another important matter that has to be addressed is how to accurately forecast a future tender B_m value so as to be able to work with equation (9). The answer is quite straightforward: there is usually a strong linear correlation between the engineer's or the

same bidder's cost estimates and the tender mean bids (Ballesteros-Pérez et al., 2013a, 2012a). Regarding this correlation, Figure 3 shows the best regression straight line between B_o (X) and B_m (Y) whose coefficient of determination is close to one (R^2=0.9343).

<Figure 3>

This regression line is generally assumed to have a zero intercept (it crosses the origin of coordinates) so just a single pair of (B_o, B_m) values is needed to provide a first estimate of the line gradient (which in this case equals 1.114). In cases in which the forecaster does not even have single prior cost estimate available, he/she can assume that B_o=B_m, which is expected to be better than using expression (1).

5.2. Method validation

As can be observed, the method of obtaining the *unbiased* participation ratios R_{ij} is quite simple. These latter values can be directly implemented instead of the current R_i values in equations (2) to (7) depending on which statistical results are wanted.

However, a validation procedure is needed to check the assumptions made in the model. Initially, the alert reader might think that Discriminant Analysis could be a logical option, but the generally low participation ratio values as well as the high number of scenarios (possible combinations of bidders) would make the direct comparison between actual and predicted group membership extremely difficult.

An alternative validation procedure is required and to do this, 10 000 artificial tenders were generated by Monte-Carlo simulation. Each simulation *j* selected randomly a B_{mj} value as the inverse of a log-normal distribution whose average and standard deviation were the same as the μ_d and σ_d values, respectively, in Table 2. Then, all not once-bidders' (bidders from 1 to 61) R_{ij} values were calculated by means of expression (9), while the R_{ij} values of once-bidders (bidders 62 to 100) were considered as constant (R_{62j}= …=$R_{100\,j}$= R_{62}=…= R_{100}=1/C_d=1/47 =0.0213).

Having calculated all the R_{ij} values, the participation of each bidder was decided as a function of this dichotomic option: if a 0-to-1 randomly generated number was below that bidder *i*'s R_{ij} value, then that bidder participates in the tender *j*, otherwise ($random>R_{ij}$) the bidder is considered not to submit a bid. This approach is equivalent to numerically obtaining the probabilities of a Poisson Binomial distribution whose different Bernouilli experiments are associated to the R_{ij} probability values. Therefore, for all the 10 000 artificial tenders, the number of bidders eventually participating was counted (*N* values) and their probability density and cumulative functions (*N* model curves shown in Figure 4) were easily obtained.

The comparison of the actual distribution of *N* and the curve obtained by means of the 10,000 simulations is presented in Figure 4. The overlap between both curves is remarkable, with the maximum deviation of below 0.057 between the Y's *N* actual and model curves occurring at $X=N=11$, being well within the Kolmogorov-Smirnov critical value of 0.194 for $p<0.05$. The results therefore confirm empirically that the method is very precise, despite the large amount of once-bidders and despite using a log-normal distribution for modelling the contract size opportunities instead of a multiple repetition of the 47 actual B_m values.

<Figure 4>

6. Discussion

Despite the method developed above being a simple, yet powerful, approach to forecasting individual bidder participation, some issues deserve further comment:

1. The question arises as to whether the same assumptions hold for datasets in which there is a shared bidder (for example tender databases 1 to 3 in Table 3). This was tested by using the same method in that kind of databases (not included here due to space limitations), with quite similar results. However, it is clear that when the proposed method is implemented in a database gathered by the same bidder whose participation is constant in

all (or almost all) contracts, the forecasting results will only become useful for that specific bidder. This is because the database will quite probably have lost some information concerning the participation of other bidders against which the shared bidder does not usually compete.

2. It can be questioned whether the method could also apply to other tenders with a larger number of bidders, with more complex characteristics or even with a selective tendering scheme, instead of the open tendering approach like the example shown here. Concerning the first two variants (tenders with larger N and multi-attribute tenders in which other factors besides the economic bid are taken into consideration for awarding the contract), the model developed is equally valid, since it relies on two main variables: the average bid price of a tender and a series of bidders that either participate or not, and those variables are shared with the latter two kinds of tenders as well. Regarding selective tenders, unlike open tendering where bidders can freely submit their bids whenever they are qualified to do so, only those bidders invited by the contracting authority can take part. Perhaps surprisingly, results obtained with selective tendering databases (also not presented here due to lack of space) indicate that, unless the contracting authority chooses the invited subset of bidders without following coherent criteria (for instance, according to the contract size or type of work), the method continues to be applicable.
3. The question also arises of how to handle the situation when there are N=1 tenders or even N=0 tenders. With all the databases examined, neither of these situations occur. However, it is quite likely that the datasets themselves were created on this premise (even avoiding tenders with $N=2$ bidders such as in the present dataset) probably to avoid problems related to under competitiveness. This may be because, when this happens in practice, tenders are usually re-opened and the first encounter generally forgotten and not registered. To accommodate this situation, a suitable approach to obtain the corrected N

model curve from Figure 4 would be to just calculate the sum of probabilities Pr(N=0) + Pr(N=1) = $P_{N=0,1}$, and divide the rest of probability values Pr(N=2, 3, …) by "1–$P_{N=0,1}$" while setting Pr(N=0)=0 and Pr(N=1)=0.

4. Another interesting issue is whether the method could gain in accuracy by using the actual bid values for each bidder i instead of the B_m values when calculating, for example, the fit of each log-normal distribution. Empirical results from other researchers with other databases (Ballesteros-Pérez and Skitmore, 2014) indicate that the differences found in the μ_i and σ_i values are very small, especially after the transformation into log values. Moreover, the fact that a forecasting bidder might need only the B_m values instead of all the bidders' bids makes the method more attractive as the number of calculations is even lower.
5. The treatment of once bidders is a possible cause for concern, with the simplification assuming that their R_{ij} values are assumed to be constant and equal to $1/C_d$. Once-bidders, or even those bidders who might potentially participate in future but are not yet present in the database, constitute one of the most recurring problems in the empirical competitive bidding literature. The best that can be said here is that this simplification did not affect the accuracy shown in Figure 4 because the series of 10 000 B_m values generated in the simulation verification procedure responded to the same pattern as the original data, lending support to the expectation that once-bidders react analogously according to the participation ratio of 1/47. However, how could we check whether these bidders would have had a preference for some tender sizes when we only a single point is available from each them? One approach to this is to simultaneously arrange the tenders in increasing and decreasing order of size and represent the cumulative ratio of appearance of total number of once-bidders per tender for both orders. The results of this analysis can be found in Figure 5, which, at least with the case study data, shows that treating all once-

bidders as the same bidder with $R_{ij}=R_i=1/C_d$ constitutes a reasonable approximation. This is because both curves evidence very similar gradients most of the time, which means that the ratio of appearance of once-bidders per new tender is quite similar no matter whether the tenders' average size is big or small. Repeating this with other databases indicates this to be generally the case, but with occasional exceptions in which the ratio of appearance of once-bidders for large contracts was higher compared to the smaller contracts. This may be due to bidders for larger contacts being less local than those for smaller contracts and therefore less prone to figure in local databases, but further research is needed to establish this point.

<Figure 5>

6. A last note is worth including about the ethical issues that may arise when implementing the method described here. It is well known that disclosing the identity of participating bidders by the auctioner to a subset of the future participants, or the action of sharing inner and privileged information among bidders with the intention of favouring some while being detrimental to the rest (namely, collusion) is considered an illegal practice. However, the developed method makes use of data that are publicly available or that can be complemented with a company's personal information (for example, the cost estimates of the previous contracts). In this sense, the method only requires processing and taking advantage of personal and public information, therefore its implementation can be considered absolutely valid, legal and advantageous, whenever the information derived is not used in agreement with other bidders.

7. Conclusions

Predicting the number of potential bidders competing in a future tender was a highly debated topic for many years in the construction management literature, and equally affects

other areas in the context of tendering and procurement. After over 30 years of debate, however, researchers from many fields apparently ran out of ideas and enthusiasm and satisfactory solutions were never reached.

That is the point of departure of this work in which a new method for forecasting the identity and number of bidders is proposed, which is expected to reignite the debate whose real-life applications would be immediate and transcendent in competitive bidding, since these variables partially condition the bidders' decision to bid, and their competitive behaviour by means of the final bid price setting, as well as some other important tender results, such as the concentration or dispersion of the bid values or the degree of correlation between the average and the lowest bids, being these key assumptions of many bid tender forecasting models (Ballesteros-Pérez et al., 2015b).

To achieve this goal, the paper starts with a brief but thorough review of the most representative literature on forecasting the number and identity of bidders. A new model is then developed by extending Skitmore's (1986) work by assimilating participation ratios into the probability of each bidder taking part in a future tender. The participation ratios are refined by removing the bias caused by the uneven number of contract size opportunities, and expressing them as a function of the contract economic size through the ratio of two simple log-normal distributions. This provides a model that does not require complex calculations, is quite accurate, and is ready-to-use for real-life applications in which a bidder or contracting authority wants to estimate the likelihood of a specific single individual, or group of key previously identified bidders, submitting a bid in a future tender. The validation process shows that the model is also capable of indirectly calculating the total number of participating bidders' distribution by means of a numerical implementation of a Poisson Binomial distribution whose Bernouilli probabilities of success are assimilated to the Participation Ratios obtained earlier.

The model presented is not complete yet, since there are other tender characteristics (e.g. client, type, location) that have been intentionally neglected. Future work will either embed these characteristics in the model or using a subset of the database of projects with similar characteristics identified by an appropriate method of pooling such as developed by Yeung and Skitmore (2012). More work is also needed in analysing the implications and loss of accuracy due to the assumption of independence of the participation ratios (probability of participation). Although the current model has taken a significant step in that direction, it is well known by practitioners that bidders are often affected by who they are, or likely to be, bidding against – suggesting a possible application of game theory in future work.

REFERENCES

Ahmad, I., Minkarah, I., 1988. Questionnaire Survey on Bidding in Construction. J. Manag. Eng. 4, 229–243. doi:10.1061/(ASCE)9742-597X(1988)4:3(229)

Al-Arjani, A.H., 2002. Type and size of project influences on number of bidders for maintenance and operation projects in Saudi Arabia. Int. J. Proj. Manag. 20, 279–287. doi:10.1016/S0263-7863(01)00015-1

Athias, L., Nuñez, A., 2009. The More the Merrier? Number of Bidders , Information Dispersion , Renegotiation and Winner's Curse in Toll Road Concessions, in: Discussion Paper Series EPPP DP No . 2009-7. pp. 1–2.

Azman, M.A., 2014. Number of Bidders in Small and Medium Public Construction Procurement in Malaysia. Appl. Mech. Mater. 567, 595–600. doi:10.4028/www.scientific.net/AMM.567.595

Bageis, A.S., Fortune, C., 2009. Factors affecting the bid/no bid decision in the Saudi Arabian construction contractors. Constr. Manag. Econ. 27, 53–71. doi:10.1080/01446190802596220

Bajari, P., Hortacsu, A., 2003. The Winner's Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions. RAND J. Econ. 34, 329–355. doi:10.2307/1593721

Ballesteros-Pérez, P., González-Cruz, M.C., Cañavate-Grimal, A., 2012a. Mathematical relationships between scoring parameters in capped tendering. Int. J. Proj. Manag. 30, 850–862. doi:10.1016/j.ijproman.2012.01.008

Ballesteros-Pérez, P., González-Cruz, M.C., Pastor-Ferrando, J.P., Fernández-Diego, M., 2012b. The iso-Score Curve Graph. A new tool for competitive bidding. Autom. Constr. 22, 481–490. doi:10.1016/j.autcon.2011.11.007

Ballesteros-Pérez, P., González-Cruz, M.C., Cañavate-Grimal, A., 2013a. On competitive bidding: Scoring and position probability graphs. Int. J. Proj. Manag. 31, 434–448. doi:10.1016/j.ijproman.2012.09.012

Ballesteros-Pérez, P., González-Cruz, M.C., Cañavate-Grimal, A., Pellicer, E., 2013b. Detecting abnormal and collusive bids in capped tendering. Autom. Constr. 31, 215–229. doi:10.1016/j.autcon.2012.11.036

Ballesteros-Pérez, P., Skitmore, M., 2014. Licitación cuantitativa aplicada al sector de la Construcción (in Spanish) Retrieved on February 10th from: https://www.researchgate.net/publication/263226807_Licitacin_Cuantitativa_aplicada_al_sector_de_la_Construccin.

Ballesteros-Pérez, P., González-Cruz, M.C., Fernández-Diego, M., Pellicer, E., 2014. Estimating future bidding performance of competitor bidders in capped tenders. J. Civ. Eng. Manag. 1–12. doi:10.3846/13923730.2014.914096

Ballesteros-Pérez, P., Skitmore, M., Das, R., del Campo-Hitschfeld, M.L., 2015a. Quick Abnormal-Bid–Detection Method for Construction Contract Auctions. J. Constr. Eng. Manag. 141(7), 04015010. doi:10.1061/(ASCE)CO.1943-7862.0000978

Ballesteros-Pérez, P., González-Cruz, M.C., Fuentes-Bargues, J.L. and Skitmore, M. 2015b. Analysis of the distribution of the number of bidders in construction contract auctions. Constr. Manag. Econ. In press. doi:10.1080/01446193.2015.1090008

Ballesteros-Pérez, P., del Campo-Hitschfeld, M.L., Mora-Melia, D., Domínguez, D., 2015c. Modeling bidding competitiveness and position performance in multi-attribute construction auctions. Oper. Res. Perspect. 2, 24–35. doi:10.1016/j.orp.2015.02.001

Benjamin, N.B.H., 1969. Competitive bidding for building construction contracts. PhD dissertation. Stanford University.

Bolton, G.E., Ockenfels, A., 2014. Does laboratory trading mirror behavior in real world markets? Fair bargaining and competitive bidding on eBay. J. Econ. Behav. Organ. 97, 143–154. doi:10.1016/j.jebo.2013.11.003

Brown, K.C., 1986. A theoretical and statistical study of decision making under uncertainty – competitive bidding for leases of offshore petroleum tracts. PhD dissertation, Southern Methodist University.

Capen, E.C., Clapp, R.V., Campbell, W.M., 1971. Competitive Bidding in High-Risk Situations. J. Pet. Technol. 23, 641–653.

Casey, B.J., Shaffer, L.R., 1964. An evaluation of some competitive bid strategy models for contractors. Report n^{o}4 Department of Civil engineering. University of Illinois.

Costantino, N., Falagario, M., Pietroforte, R., Sciancalepore, F., 2011. A model for predicting the bid distribution in public tenders. Ippc 1–15.

Drew, D.S., 1995. The effect of contract type and size on competitiveness in construction contract bidding. PhD thesis, Department of Surveying, The University of Salford.

Drew, D.S., Skitmore, R.M., 2006. Competitiveness in bidding: a consultant's perspective. Constr. Manag. Econ. 10, 227–247. doi:10.1080/01446199200000020

Dyer, D., Kagel, J.H., Levin, D., 1989. Resolving uncertainty about the number of bidders in independent private-value auctions : an experimental analysis. RAND J. Econ. 20, 268–279. doi:10.2307/2555693

Engelbrecht-Wiggans, R., 1980. Auctions and bidding models: a survey. Manage. Sci. 26, 119–142. doi:10.1287/mnsc.26.2.119

Engelbrecht-Wiggans, R., Dougherty, E.L., Lohrenz, J., 1986. A Model for the Distribution of the Number of Bids on Federal Offshore Oil Leases. Manage. Sci. 32, 1087–1094. doi:10.1287/mnsc.32.9.1087

Friedman, L., 1956. A competitive bidding strategy. Oper. Res. 1, 104–12. doi:10.1287/opre.4.1.104

Fu, W., 2004. The effect of experience in construction contract bidding. PhD thesis, Department of Building and Real Estate, Hong Kong Polytechnic University.

Fuentes-Bargues, J.L., González-Cruz, M.C., González-Gaya, C., 2015. La contratación pública de obras: situación actual y puntos de mejora (in Spanish). Inf. la Construcción 1–9. doi:10.3989/ic.12.130

Gates, M., 1967. Bidding strategies and probabilities. J. Constr. Div. Proc. Am. Soc. Civ. Eng. 93, 75–107.

Harstad, R.M., Kagel, J.H., Levin, D., 1990. Auctions with an uncertain number of bidders.pdf. Econ. Lett. 33, 35–40.

Hu, A., 2011. How bidder's number affects optimal reserve price in first-price auctions under risk aversion. Econ. Lett. 113, 29–31. doi:10.1016/j.econlet.2011.05.041

Keller, A.Z., Bor, R.H., 1978. Strategic aspects of bidding against an unknown number of bidders, in: TIMS/ORSA Meeting. May, New York.

Klemperer, P., 2004. Auctions: Theory and Practice. Princeton University Press.

Levin, D., Ozdenoren, E., 2004. Auctions with uncertain numbers of bidders. J. Econ. Theory 118, 229–251. doi:10.1016/j.jet.2003.11.007

Maskin, E., Riley, J., 2000. Asymmetric Auctions. Rev. Econ. Stud. 67, 413–438.

McAfee, R.P., McMillan, J., 1987. Auctions with a Stochastic Number of Bidders 19, 1–19.

Morin, T.L., Clough, L.H., 1969. OPBID: Competitive bidding strategy model. J. Constr. Div. ASCE 95, 85–106.

Ngai, S.C., Drew, D.S., Lo, H.P., Skitmore, M., 2002. A theoretical framework for determining the minimum number of bidders in construction bidding competitions. Constr. Manag. Econ. 20, 473–482. doi:10.1080/01446190210151041

Oo, B.-L., Drew, D.S., Runeson, G., 2010. Competitor analysis in construction bidding. Constr. Manag. Econ. 28, 1321–1329. doi:10.1080/01446193.2010.520721

Park, W.R., 1966. The strategy of contracting for profit. Ed. Prentice-Hall.

Rubey, H., Milner, W.W., 1966. A Statistical approach to bidding. Ed. McMillan.

Runeson, K.G., 1987. Analysis of building price estimates. MSc thesis, School of Building, University of New South Wales.

Runeson, G., Skitmore, M., 1999. Tendering theory revisited. Constr. Manag. Econ. 17, 285–296. doi:10.1080/014461999371493

Shaffer, L.R., Micheau, T.W., 1971. Bidding with competitive strategy models. J. Constr. Div. ASCE 97, 113–126.

Shash, A.A., 1993. Factors considered in tendering decisions by top UK contractors. Constr. Manag. Econ. 11, 111–118. doi:10.1080/01446199300000004

Skitmore, M., 1981a. Why do tenders vary? Chart. Quant. Surv. 4, 128–129.

Skitmore, M., 1981b. Bidding dispersion – an investigation into method of measuring the accuracy of building costs estimates. Msc Thesis. University of Salford.

Skitmore, M., 1986. A model for the construction project selection and bidding decision. PhD thesis, Department of Civil Engineering, University of Salford.

Skitmore, M., 1988. Fundamental Research in Bidding and Estimating, in: Proceedings British/Israeli Seminar on Building Economics; The International Council for Building Research Studies and Documentation CIB W-55.

Skitmore, M., 2008. First and second price independent values sealed bid procurement auctions: some scalar equilibrium results. Constr. Manag. Econ. 26, 787–803. doi:10.1080/01446190802175678

Skitmore, M., 2013. Generalised gamma bidding model. J. Oper. Res. Soc. 65, 97–107. doi:10.1057/jors.2013.18

Skitmore, M., Pemberton, J., 1994. A Multivariate Approach to Construction Contract Bidding Mark-up Strategies. J. Oper. Res. Soc. 45, 1263–1272.

Stark, R.M., Rothkopf, M.H., 1979. Discussion of "Bidding strategies and probabilities" by Marvin Gates. Oper. Res. 27, 364–390. doi:http://www.jstor.org/stable/170135

Sugrue, P.K., 1977. The design and evaluation of three competitive bidding models for application in the construction industry.

Wade, R.L., Harris, B., 1976. LOMARK: a bidding strategy. J. Constr. Div. ASCE 102, 197–211.

Yeung, D., Skitmore, M., 2012. A method for systematically pooling data in very early stage construction price forecasting. Constr. Manag. Econ. 30, 929–939. doi:10.1080/01446193.2012.733402

Tender (j)	*Nº bidders (N)*	*Mean bid (B_m)*	*Log Mean bid (LN B_m)*	*Engineer's Cost est. (B_o)*	*Bidder$_{(i)}$ (identities of bidders from ↓ to ↑ bid)►*																
					(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*	*(8)*	*(9)*	*(10)*	*(11)*	*(12)*	*(13)*	*(14)*	*(15)*	*(16)*	*(17)*
1	16	9,686,030	16.086	7,946,837	10	35	22	19	4	20	17	87	9	61	83	93	7	100	99	92	
2	9	28,415,299	17.162	32,717,190	61	90	45	88	91	59	19	44	36								
3	10	26,950,453	17.110	22,085,527	97	86	60	59	4	85	44	3	6	45							
4	4	22,905,464	16.947	20,452,058	96	94	45	13													
5	12	15,132,109	16.532	12,818,817	17	1	16	19	18	30	2	25	77	68	13	21					
6	17	20,412,549	16.832	20,166,667	5	1	4	54	6	32	47	9	55	10	13	72	53	52	7	70	41
7	11	37,397,035	17.437	41,197,688	4	14	17	19	12	1	33	22	16	8	28						
8	13	29,557,430	17.202	24,773,359	13	31	16	8	17	12	33	4	28	7	22	41	30				
9	8	29,270,416	17.192	28,409,004	17	1	12	30	40	16	37	39									
10	9	23,288,908	16.963	21,445,392	41	5	12	36	8	52	4	7	38								
11	8	15,837,168	16.578	15,033,081	17	14	43	36	32	7	8	22									
12	12	30,634,246	17.238	30,075,610	28	43	2	8	3	39	4	14	56	22	13	7					
13	7	73,933,651	18.119	59,041,433	6	1	3	31	46	21	7										
14	5	66,871,052	18.018	64,510,534	1	6	31	3	11												
15	5	50,643,989	17.740	35,163,834	4	5	1	6	3												
16	7	54,947,923	17.822	42,532,702	1	58	3	4	16	5	11										
17	6	44,609,273	17.613	41,783,933	3	5	1	4	16	2											
18	7	49,620,702	17.720	49,403,385	58	3	4	16	2	1	5										
19	6	44,257,298	17.606	38,121,405	29	2	56	3	5	69											
20	5	98,482,309	18.405	65,801,501	5	2	12	3	4												
21	5	66,803,823	18.017	55,493,122	31	5	3	1	2												
22	5	56,972,579	17.858	32,969,570	12	5	35	1	2												
23	3	61,792,701	17.939	50,419,645	36	18	24														
24	12	54,063,116	17.806	52,179,813	5	10	60	4	7	6	20	1	3	2	23	8					
25	13	52,632,198	17.779	34,949,799	14	7	23	10	4	1	43	5	57	8	21	11	18				
26	6	81,034,245	18.210	73,785,292	10	6	2	1	4	8											
27	9	113,779,940	18.550	83,713,655	10	20	1	2	4	24	8	5	44								
28	10	82,222,571	18.225	75,087,435	24	5	23	6	35	1	2	9	4	8							
29	9	108,478,803	18.502	78,376,384	23	6	1	2	12	9	32	10	24								
30	9	92,950,923	18.348	95,160,427	10	15	1	23	2	9	24	35	6								
31	8	32,798,912	17.306	31,426,753	28	3	34	6	1	2	5	4									
32	6	89,469,514	18.309	86,982,546	1	6	9	2	15	12											
33	7	98,579,660	18.406	89,757,288	34	15	6	1	2	3	9										
34	6	103,983,088	18.460	88,809,151	6	1	42	2	9	3											
35	9	68,563,040	18.043	53,824,220	30	34	13	18	32	21	33	98	57								
36	15	61,345,823	17.932	50,966,816	3	2	1	24	9	6	37	42	3	6	2	9	1	24	37		
37	8	79,826,297	18.195	65,706,624	15	37	9	2	3	1	42	74									
38	7	58,034,490	17.877	51,677,529	15	9	2	3	37	26	1										
39	14	91,374,772	18.330	82,549,634	40	53	50	27	38	14	6	63	49	11	3	25	21	29			
40	6	108,676,000	18.504	93,097,653	3	18	29	27	67	26											
41	7	155,298,815	18.861	139,121,526	14	2	1	49	29	51	65										
42	6	79,827,787	18.195	73,685,775	75	6	27	48	54	11											
43	6	137,137,793	18.736	133,207,196	51	79	34	3	14	11											
44	15	143,181,832	18.780	151,459,732	38	27	47	15	78	20	11	19	89	48	71	26	84	95	80		
45	11	12,117,616	16.310	10,394,295	20	17	18	1	10	5	81	55	2	33	13						
46	8	25,298,293	17.046	24,352,349	10	82	73	5	7	25	13	39									
47	12	63,922,601	17.973	65,877,433	25	26	50	62	64	15	46	40	11	66	76	19					

Table 1. Example of construction tender dataset used for the analysis

Order (from↓to↑►)	*Dataset (LN B_m)*	*Order (◄from↓to↑)*	*Bidder i's ordered LN B_m values ►*																												
			1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	*20*	*21*	*22*	*23*	*24*	*25-36*	*37-45*	*46-61*	*62-100*
1	16.086	***1***	16.310	16.310	17.110	16.086	16.310	16.832	16.086	16.578	16.086	16.086	17.779	16.963	16.310	16.578	17.877	16.532	16.086	16.310	16.086	16.086	16.532	16.086	17.779	17.939	…	…	…	…
2	16.310	***2***	16.532	16.532	17.238	16.832	16.832	17.110	16.578	16.963	16.832	16.310	17.822	17.192	16.532	17.238	17.973	17.192	16.310	16.532	16.532	16.310	17.779	16.578	17.806	18.225	…	…	…	
3	16.532	***3***	16.832	17.238	17.306	16.963	16.963	17.306	16.832	17.202	17.877	16.832	17.973	17.202	16.832	17.437	18.195	17.202	16.532	17.779	17.162	17.806	18.043	17.202	18.225	18.348	…	…		
4	16.578	***4***	17.192	17.306	17.606	17.110	17.046	17.740	16.963	17.238	18.195	17.046	18.018	17.437	16.947	17.779	18.309	17.437	16.578	17.939	17.437	18.550	18.119	17.238	18.348	18.502	…			
5	16.832	***5***	17.306	17.606	17.613	17.202	17.306	17.806	17.046	17.437	18.225	17.779	18.195	17.858	17.046	18.330	18.348	17.613	17.192	18.043	17.973	18.780	18.330	17.437	18.502	18.550				
6	16.947	***6***	17.437	17.613	17.720	17.238	17.606	18.018	17.202	17.779	18.309	17.806	18.330	18.309	17.202	18.736	18.406	17.720	17.202	18.504	18.780									
7	16.963	***7***	17.613	17.720	17.740	17.306	17.613	18.119	17.238	17.806	18.348	18.210	18.736	18.405	17.238	18.861	18.780	17.822	17.437											
8	17.046	***8***	17.720	17.806	17.806	17.437	17.720	18.195	17.779	18.210	18.406	18.348	18.780	18.502	18.043															
9	17.110	***9***	17.740	17.858	17.822	17.613	17.740	18.210	17.806	18.225	18.460	18.502																		
10	17.162	***10***	17.779	17.877	17.877	17.720	17.779	18.225	18.119	18.550	18.502	18.550																		
11	17.192	***11***	17.806	18.017	18.017	17.740	17.806	18.309																						
12	17.202	***12***	17.822	18.195	18.018	17.779	17.822	18.330																						
13	17.238	***13***	17.858	18.210	18.119	17.806	17.858	18.348																						
14	17.306	***14***	17.877	18.225	18.195	17.822	18.017	18.406																						
15	17.437	***15***	18.017	18.309	18.330	18.210	18.225	18.460																						
16	17.606	***16***	18.018	18.348	18.405	18.225	18.405	18.502																						
17	17.613	***17***	18.119	18.405	18.406	18.405	18.550																							
18	17.720	***18***	18.195	18.406	18.460	18.550																								
19	17.740	***19***	18.210	18.460	18.504																									
20	17.779	***20***	18.225	18.502	18.736																									
21	17.806	***21***	18.309	18.550																										
22	17.822	***22***	18.348	18.861																										
23	17.858	***23***	18.406																											
24	17.877	***24***	18.460																											
25	17.932	***25***	18.502																											
26	17.939	***26***	18.550																											
27	17.973	***27***	18.861																											
28	18.017	***28***																												
29	18.018	***29***																												
30	18.043	***30***																												
31	18.119	***31***																												
32	18.195	***32***																												
33	18.195	***33***																												
34	18.210	***34***																												
35	18.225	***35***																												
36	18.309	***36***																												
37	18.330	***37***																												
38	18.348	***38***																												
39	18.405	***39***																												
40	18.406	***40***																												
41	18.460	***41***																												
42	18.502	***42***																												
43	18.504	***43***																												
44	18.550	***44***																												
45	18.736	***45***																												
46	18.780	***46***																												
47	18.861	***51***																												
Count (C_d)	**47**	***Count (C_i)***	27	22	20	18	17	16	10	10	10	10	8	8	8	7	7	7	7	6	6	5	5	5	5	5	**4**	**3**	**2**	**1**
(Biased) Particip. Ratios $R_i=C_i/C_d$			0.5745	0.4681	0.4255	0.3830	0.3617	0.3404	0.2128	0.2128	0.2128	0.2128	0.1702	0.1702	0.1702	0.1489	0.1489	0.1489	0.1489	0.1277	0.1277	0.1064	0.1064	0.1064	0.1064	0.1064	**0.0851**	**0.0638**	**0.0426**	**0.0213**

Mean (μ_d)	17.762	***Mean (μ_i)***	17.854	17.925	17.951	17.558	17.623	17.995	17.165	17.599	17.924	17.547	18.204	17.734	17.019	17.851	18.270	17.360	16.763	17.518	17.328	17.506	17.761	16.908	18.132	18.313	…	…	…	…
Std. Dev.(σ_d)	0.676	***Std.Dev.(σ_i)***	0.612	0.640	0.450	0.612	0.580	0.506	0.613	0.624	0.811	0.916	0.386	0.615	0.522	0.837	0.298	0.438	0.513	0.886	0.973	1.250	0.714	0.561	0.325	0.245	…	…	…	-

A^2_d	0.657	***A^2_i***	0.639	0.733	0.220	0.240	0.432	0.749	0.220	0.200	1.302	0.456	0.413	0.469	0.318	0.229	0.263	0.386	0.388	0.528	0.165	0.443	0.694	0.449	0.422	0.309	…	…	…	-
A^2crit_d	0.752	***A^2crit_i***	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	0.752	-
$A^2_d<A^2crit_d$?	✓	***$A^2_i<A^2crit_i$?***	✓	✓	✓	✓	✓	✓	✓	✓	×	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓x12	✓x9	✓x16	-

Table 2. Analysis of lognormal participation distributions of all bidders with at least five bids submitted

ID	Database Alias	Type of work	Country	Period	Source
1	UK51	Buildings	England	1981-1982	(Skitmore & Pemberton, 1994)
2	US50	Buildings	USA	1965-1969	(Shaffer & Micheau, 1971)
3	SP45	WWTP	Spain	2007-2008	(Ballesteros-Pérez et al. , 2012)
4	UK272	Civil and Buildings	UK	1969-1979	(Skitmore, 1981)
5	UK218	Civil	England	1979-1982	(Skitmore, 1986)
6	UK373	Buildings	London	1976-1977	(Skitmore, 1986)
7	US64	Buildings	USA	1976-1984	(Brown, 1986)
8	HK199	Buildings	Hong Kong	1981-1990	(Drew, 1995)
9	HK261A	Services	Hong Kong	1991-1996	(Fu, 2004)
10	HK261B	Services	Hong Kong	1991-1996	(Fu, 2004)
11	AU152	Civil	Australia	1972-1982	(Runeson, 1987)
12	AU161	Civil	Australia	1972-1982	(Runeson, 1987)

ID	One shared bidder	Number of auctions	Kolmogorov-Smirnov goodness of fit test (Lognormal distr.)						
			D	$D_{\alpha=0.01}$	$D_{\alpha=0.05}$	$D_{\alpha=0.10}$	$D<D_{\alpha=0.01}$?	$D<D_{\alpha=0.05}$?	$D<D_{\alpha=0.10}$?
1	Yes	51	0.074	0.143	0.123	0.113	✓	✓	✓
2	Yes	50	0.077	0.144	0.125	0.114	✓	✓	✓
3	Yes	45	0.060	0.152	0.131	0.120	✓	✓	✓
4	No	272	0.039	0.063	0.054	0.050	✓	✓	✓
5	No	218	0.051	0.070	0.060	0.055	✓	✓	✓
6	No	373	0.022	0.053	0.046	0.042	✓	✓	✓
7	No	64	0.081	0.128	0.111	0.101	✓	✓	✓
8	No	199	0.032	0.073	0.063	0.058	✓	✓	✓
9	No	202	0.065	0.073	0.063	0.057	✓	×	×
10	No	59	0.089	0.129	0.111	0.102	✓	✓	✓
11	No	152	0.070	0.084	0.072	0.066	✓	✓	×
12	No	161	0.050	0.081	0.070	0.064	✓	✓	✓

Table 3. Lognormal distribution goodness of fit tests for twelve construction and services tender databases found in the project management literature

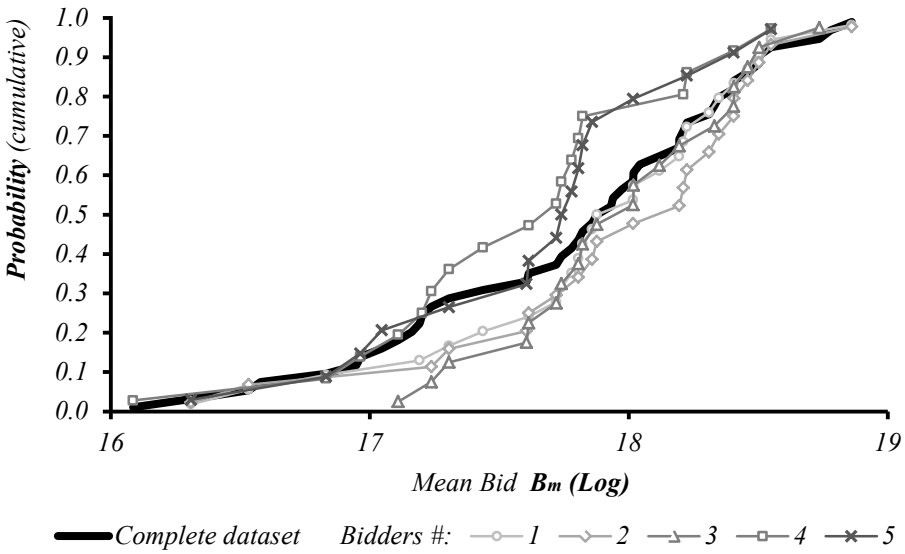


Figure 1. Mean bid cumulative distribution of the complete dataset and the five bidders that had higher participation

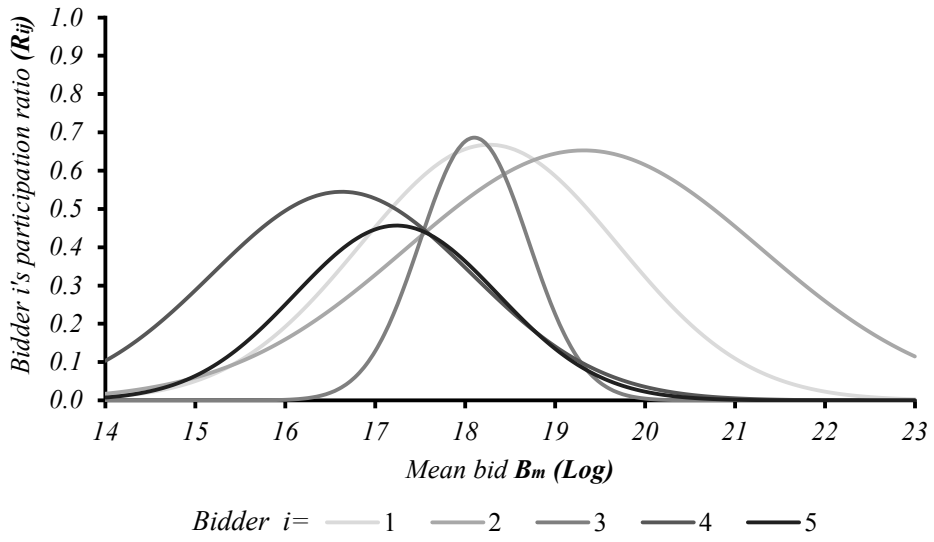


Figure 2. Unbiased Participation Ratio distributions of the five bidders that had higher participation

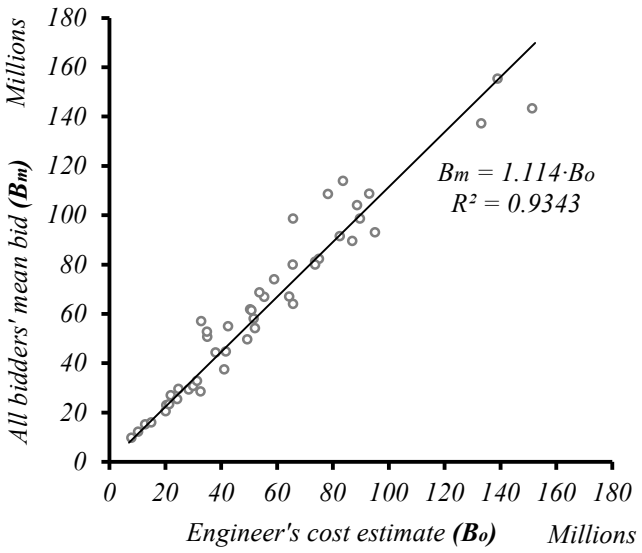


Figure 3. Regression straight line between the engineer's cost estimate (B_o) and all bidders' mean bid (B_m) for all the auctions contained in the example dataset

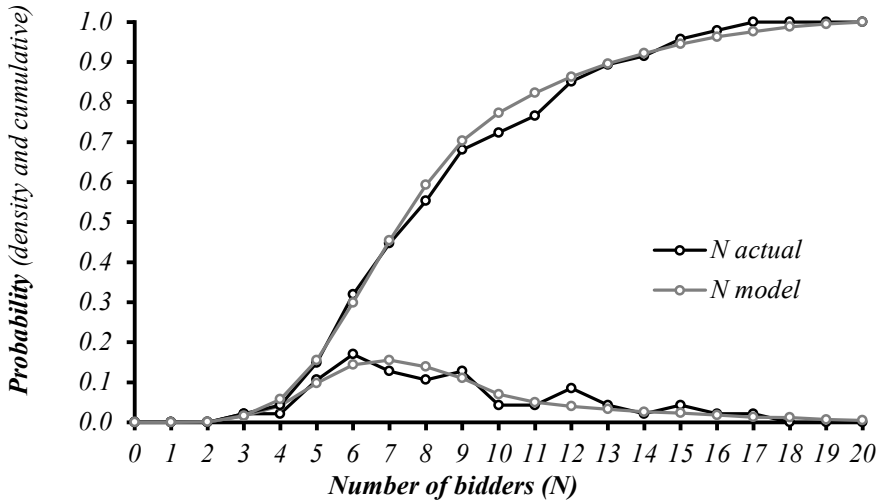


Figure 4. Comparison of the Actual and Model Number of bidders distributions

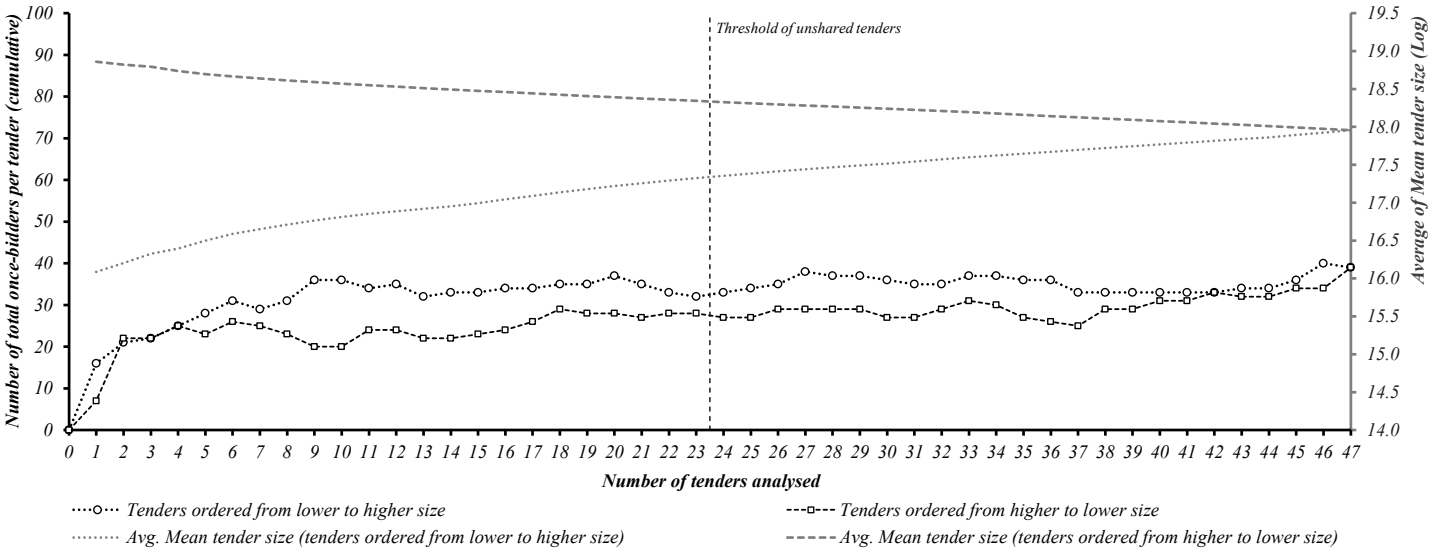


Figure 5. Evolution of once-bidders appearance as a function of the tender size order

View publication stats