Estimating the number of new and repeated bidders in construction auctions

Ballesteros-Pérez, Pablo; Skitmore, Martin

Estimating the number of new and repeated bidders in construction auctions

Authors:

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

[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 7186   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/

Estimating the number of new and repeated bidders in construction auctions

ABSTRACT

The number of new bidders – bidders from whom there is no previous registered participation – is an important variable in most bid tender forecasting models, since the unknown competitive profile of the former strongly limits the predictive accuracy of the latter. Analogously, when a bidder considers entering a bid or when an auctioneer is handling a procurement auction, assessing the likely proportion of experienced bidders is considered an important aspect, as some strategic decisions or even the awarding criteria might differ.

However, estimating the number of bidders in a future auction that have not submitted a single bid yet is difficult, since there is no data at all linking their potential participation, an essential requirement for the implementation of any forecasting or estimation method.

A practical approach is derived for determining the expected proportion of new bidders to frequent bidders as a function of the population of potential bidders. A multinomial model useful for selective and open tendering is proposed and its performance is validated with a dataset of actual construction auctions. Final remarks concern the valuable information provided by the model to an enduring unsolved bidding problem and the prospects for new research continuations.

Keywords: *Modelling; Forecasting; Bidding; Tendering; New bidders; Multinomial.*

1

Introduction

Public bidding constitutes a significant source of work for many contractors in the construction industry and makes a substantial contribution to the GDP of developed countries (8.7 trillion USD globally according to the International Monetary Fund (IMF, 2014) or 9.7 trillion USD according to the World Bank (WB, 2013)). Bidders weigh many factors when making their decision to bid (d2b) and setting an appropriate mark-up. According to studies in the UK (Shash, 1993) and the US (Ahmad & Minkarah, 1988), the number and identity of bidders to be faced are among the most important of these. Similarly, the huge literature on the economic theory of auctions makes the crucial assumption that these are common knowledge to the bidders or that all are independent and identically distributed (*iid)* (see Klemperer, 2004, for a review of the main contributions). It has been known for many years, for example, that the number and competitive profiles of bidders profoundly conditions major tender outcomes (e.g., Dyer et al., 1989; Levin and Ozdenoren, 2004; Hu, 2011, Ishii et al., 2014) with the phenomenon of the *winner's curse* being the most celebrated (Capen et al., 1971).

Also, from the auctioneer's point of view, it is equally important to assess fairly and adequately the potential bidders that will submit bid proposals (Ballesteros-Pérez, et al., 2015; European Union, 1999; Wang et al., 2013). It is known, for instance that inexperienced bidders are more unpredictable and more prone to submit abnormally high or low bids (Ballesteros-Pérez et al., 2013a, 2015b), while experienced, recurrent, bidding is correlated with a higher rate of success (Fu et al., 2002, 2003). This usually requires designing effective awarding criteria in line with both the auctioned item characteristics and specific potential bidders in order to make a sound and effective discriminative ranking of proposals (Ballesteros-Pérez et al., 2016a; Liu et al., 2015).

2

Empirical tests have, however, shown the *iid* assumption to be inappropriate for construction contract bidding (Skitmore, 1991; Oo et al., 2010) – leaving the need for a reliable method of determining the number and identity of bidders in support of both bidding practice and the tenability of theory. In practice, this is done by the bidder's personal experience and/or analysis of competitors' previous participation ratios in projects of similar characteristics (e.g., owner, type and size) (Ballesteros-Pérez et al., 2014) and nearby location (Fu, 2004) [1]. More rigorous Bid Tender Forecasting models (BTFM) (Ballesteros-Pérez et al., 2012, 2013b; Carr, 1982; Friedman, 1956; Gates, 1967; Skitmore & Pemberton, 1994) have been advanced that rely on a database of the past bidding behaviour of individual bidders to provide a statistical base for accurate forecasts and projections for forthcoming bidding encounters.

Such models work well in situations where there is a large proportion of regular bidders. Construction contract auctions, however, invariably involve a large population of potential bidders from which an irregular few bid or are selected to bid – a typical auction-bidder matrix for a database is well over 90% sparse (Skitmore, 2013). This raises the question of how to deal with a situation where there is little or no data concerning potential bidders, either because the number of previous similar auctions is too low, too old or just because the information involved has not been shared or made publicly available.

Friedman (1956) and Gates' (1967) classical recommended treatment for potential bidders not contained (registered) in the database is to be modelled as an average bidder - a convenient imposed assumption of most BTFM (Ballesteros-Pérez et al., 2016b) as it reduces the problem to one of just knowing the number of such bidders and not their identities. Even in this simplified form, however, the number or

---

[1] This is not to mention the known, but dubious, practice of informal intercommunication between bidders and their contacts – the bidders "grapevine".

at least the proportion against other previously identified key competitors, cannot be directly inferred – a difficulty that has been regarded as problematic even from the very first studies in tendering theory (Runeson & Skitmore, 1999).

Indeed, in order to sharpen the accuracy of any BTFM – which is particularly necessary when there is not a large proportion of regular bidders or at the early stages of data gathering - it is necessary to consider the identity of any potential competitor bidder in the model. Friedman and Gates' models allowed for this feature, but their implementation is also problematic due to the one missing crucial piece of information: namely, when identities are considered, three essential pieces of information are required. First, it is necessary to model the competitive profiles of each bidder separately, as some will bid more aggressively or conservatively than others. Second, it is required to model the competitive profile of those bidders from which we know nothing (those who have not yet submitted even a single bid). These can be generally modelled as the average bidder. Third, it is necessary to anticipate the probabilities that each of the known and still unknown bidders will submit a bid for the next auction. However, estimating the probabilities of a single known bidder is relatively easy since it can be achieved by analysing, for instance, how frequently they submitted bids for similar contracts in the past. However, in terms of the probability that an unknown bidder submits a bid, the real problem is not to estimate the probability, but to anticipate the total number of the unknown bidders who might try to submit their first bid (or alternatively extend the problem to anticipate how many of the so far once-, twice, thrice-bidders might submit another bid). This is the essential piece of information that all current BTFM are lacking and the *raison d' être* for the model proposed later.

The problem endures today and, apart from Mercer and Russell's (1969) distant suggestion to derive information from the analysis of other more frequent bidders for which data *is* available, no method has yet been developed for anticipating how

4

many such new bidders might enter a future bid (Skitmore, 1986). In recognition of this, and in view of its importance for both theory and practice, two approximate approaches are examined. First, assessing the number/proportion of bidders who, despite not having submitted a single bid yet, are likely to participate in the next auction, and, second, appraising the population size of all bidders from which such potential bidders are likely to emerge. For this, a well-known statistical distribution not commonly found in the construction management literature is used: the Multinomial distribution. This distribution is shown in the results to provide good approximations of the needed estimates for both open and selective tendering schemes.

The paper is structured as follows. In the next section, a brief but focused *Literature Review* is provided. This is followed in *Materials and Methods* by an introduction of the construction tender dataset used in the study to exemplify the models' performance; a clarification of the terminology and most important variables that are incorporated in the multinomial model; its theoretical basis; and a simple method for estimating the bidders' population size. Next, a *Calculations* section presents two useful conceptual simplifications of the general multinomial model when the number of future participating bidders is known and unknown, and a numerical example of the use of the former. A *Results* section then provides a final performance summary of the model. Finally, the last *Conclusions* section summarises the conceptual framework developed, highlighting potential future ways of improvement and the prospects for new research continuations.

Furthermore, for the sake of clarity it is noted that the terms *tender* and *auction* will be used synonymously from this point.

Literature review

Estimating the number of new bidders not registered yet in a database, N_N, for a forthcoming construction auction is more complex than it seems, particularly in the initial stages collecting historical construction auction data and/or when there is hardly any experience with similar types of contracts or specific contracting authorities (Ballesteros-Pérez et al., 2010). However, it is precisely in these situations that anticipating N_N is the most useful.

Many attempts are reported in the tendering literature to forecast the *total* number of bidders, *N,* for a future auction (Ballesteros-Pérez et al., 2015c). The first proposals by Friedman (1956) contain several alternatives to forecast the expected value of *N*. One of the most popular of these has been to treat *N* as a purely stochastic variable, with values being drawn randomly from a known probability distribution. In experimental settings, for example, *N* is frequently considered to be purely stochastic (McAfee & McMillan, 1987) or even as a fixed value in Games and Auctions theory (Harstad et al., 1990), the main concern being to examine how different assumptions relating to types of auction formats and bidders' valuations of the auctioned items affect the players'/bidders' optimal equilibrium strategies (Klemperer, 2004). Friedman's initial suggestion for modelling this is the Poisson distribution, on the premise that the formation of *N* is due the 'arrival' of bidders into the auction, the Poisson being known to provide a good model in such 'arrival' situations. Indeed the use of this particular distribution triggered a long academic discussion about the suitability of this and other distributions for a wide range of bidding situations. Throughout the subsequent six decades, a long list of candidate distributions involving almost all the classical probability distributions (e.g. Poisson, Normal, Log-Normal, Uniform, Weibull, Gamma, Laplace) have been tested, but with inconclusive results (Stark and Rothkopf 1979, Engelbrecht-Wiggans *et al.* 1986, Skitmore 2013, Ballesteros-Pérez *et al.* 2015a, 2015c). Nowadays, despite an

6

extensive multi distribution fit analysis performed by Ballesteros-Pérez et al. (2015c), the debate concerning the appropriate distribution for modelling the statistical variation of N remains open, although the Poisson distribution is still commonly adopted in the analysis of online auctions (Mohlin et al., 2015) and numerical simulations (Takano et al., 2014).

Another of Friedman's proposals, reiterated by Rubey and Milner (1966) ten years later, consists of just combining the little information available about the potential competitors with the bidder's experience. In this respect, it is known in practice and theory that the value of N in a construction contract auctions depends on project type and size (Azman, 2014; Drew & Skitmore, 2006), client (Ballesteros-Pérez et al., 2012), geographical location (Al-Arjani, 2002; Benjamin, 1969) and market conditions (Ngai et al., 2002; Skitmore, 1981). However, these variables are generally difficult to standardize (Oo et al., 2007, 2010a, 2010b).

As a result, only one partial early refinement has been implemented so far to improve the estimation of N. This refinement utilises the usually existing moderate correlation between N and contract economic size (project budget) (Rickwood, 1972; Wade & Harris, 1976), a result that certainly offers higher accuracy in predicting the value of N than considering it to be purely random (Ballesteros-Pérez et al., 2015c).

However, forecasting the number of new bidders, N_N, involves an extra layer of uncertainty. An obvious approach is to infer its quantity from N-N_F, the difference between N and the expected number of frequent bidders N_F. Forecasting N-N_F, although still quite imprecise, is likely to be far less so than forecasting N_N. Forecasting N-N_F has encountered two main problems though. Firstly, it is usual for the same bidder to bid for different types of work (multi-market scheme) (Morin & Clough, 1969). Secondly, the bidder's decision to bid (d2b) is limited by the amount of work the bidder can carry out concurrently (Oo et al., 2012; Skitmore, 1988). Both

problems clearly act in opposite directions since they make bidders bid more or less competitively than imagined, respectively.

The best approach found so far for forecasting N_N or N_N as N-N_F has been to treat probabilistically the identities of bidders by groups of similar characteristics (Shaffer & Micheau, 1971; Wade & Harris, 1976). For example, Skitmore (1986) classified the potential competitors as "key" (frequent and highly competitive bidders) and "strangers" (bidders either with low competitiveness and/or low frequency of participation) (Dass et al., 2014).

Nevertheless, these approaches do not solve the problem of anticipating the participation of bidders from whom the company does not have any information. As pointed out earlier, bidders from whom there is no previous experience can only be modelled as averagely competitive. However, the present study focuses only on the number of new bidders, not in their competitive behaviour. This latter issue will be left for a separate piece of research. Also, since 1986 there have been no significant contributions in this area with the exception of some adjacent studies concerning new advances on the statistical nature of N (e.g. Athias & Nuñez, 2009; Ballesteros-Pérez et al., 2015c; Costantino et al., 2011; Skitmore, 2008).

Materials and Methods

Tender dataset

In order to develop a thorough explanation of the two methods proposed for forecasting N_N, a real construction tender database is presented. This comprises data in Skitmore's (1986) PhD thesis obtained from the records of a bidding information agency that held details of most bids for most projects in the London area in card form. To collect the data, "a period of one week was spent copying a sample of project data for the period October 1976 to June 1977". The bids and associated bidder's names were recorded and the names later encoded for analysis. The

8

resulting number of projects for which a full set of bids, together with the identity of the bidders, were available for analysis totalled 373" (Skitmore, 1986: 353).

The dataset retrieved contains 1915 sealed economic bids submitted by 354 different bidders' in 373 building-related auctions with an average number of total participating bidders of 5.13 (ranging from a minimum of 2 bidders up to a maximum of 11 bidders per auction). Concerning the contract (budget) size, the 373 auctions ranged from 10,266 GBP to 8.8 million GBP. A complete transcription of the tender dataset can be found as Supplemental online material.

Despite being relatively old, this dataset is chosen because it is large enough to reveal patterns that would not be otherwise revealed. Also, this kind of homogeneous database, including both bids and bidder ID information continuously over such a short-time period is generally extremely difficult to obtain in present day circumstances.

Finally, it is noted that the dataset contains selective auctions only, that is, the maximum number of bidders was determined and restricted by the auctioneer (client/consultant) to a group of pre-qualified bidders. Therefore, the auctioneer knew the identities of the potential bidders but that information was not shared with the bidders themselves. This is in contrast with an open tendering scheme (open auctions), where there is no such upper limit set on number of bidders, and allows a comparison to be made between these two different bidding scenarios.

The contracts (auctions) in the database are ordered in the sequence of the auction opening dates, such that contract $i=1,2,…,c$, where c is the number of contracts in the database. Letting r denote the cumulative number of bids recorded by a bidder by the time of the ith contract, Figure 1 provides two graphs of the number of bidders with r bids as i increases. The top graph shows the number of bidders with exactly r bids submitted, while the bottom graph shows the number of bidders with *at least* r bids (r bids submitted or more). By $i=c=373$ auctions, 123

9

bidders had *r>5* bids submitted but, in order to provide essential information only, no curves representing *r>5* bids are included.

<div align="center">**< Insert Figure 1 here >**</div>

In Figure 1 it is easy to see how the *cumulative r*-graph (top) shows softer curve growths. This is because each cumulative ≥r curve contains the *=r, =r+1, =r+2, …* curves (from the bottom graph).

Conversely, the top graph represents the number of bidders who have submitted exactly *r* bids at a given point in time (expressed in number of auctions completed). This means that, at any moment, any *r* bidder can become an *r+1* bidder once it places another bid, which is the reason why the top curves can increase or decrease or even overlap (since it is possible that, at any auction *i*, there are more *r+1* bidders than *r* bidders).

Finally, both graphs are considered relevant since, once both kinds of curves are modelled, the *mass r* curves (top) will allow us to monitor the evolving experience of a given number of bidders (by bidding more times), while the *cumulative r* curves (bottom), particularly the *r≥1* curve, will allow the number of new bidders to be estimated for the next auction through its decreasing gradient.

Terminology

Before continuing, it is necessary to establish a common nomenclature for the variables that will be used throughout the rest of the analysis. Let:

M(i)   be the population of potential bidders in the market on completion of the ith auction.

c   the number of auctions (contracts) in the database, where *i=1,2,…,c* as shown in the X-axis labels from Figure 1. In our particular dataset *c=373*.

k(i)   the number of participating bidders for the ith auction (previously referred to as *N*)

r the number of bids a bidder has recorded in the database after the ith auction has been completed. That is, $r=1$ for a once-only bidder, $r=2$ for a twice-only bidder, etc., with $r=0$ denoting a bidder who has yet to appear in any of the auctions in the database at this point.

$n(i,r)$ number of r bidders in the database when the ith auction is completed $(0 \leq r \leq i \leq c)$ so that at any time $n(i,r=0)+n(i,r=1)+...+ n(i, r=i)=M(i)$.

x_r is the change (increase or decrease) in the number of r bidders from the ith auction being completed to the $i+1^{th}$ auction being completed, that is, $x_r=n(i+1,r)-n(i,r)$. Obviously, $-k(i+1) \leq x_r \leq k(i+1)$ always. For example, $x_{r=1}>0$ denotes the number of new bidders (now once bidders) from $M(i+1)$, that is once auction $i+1$ is completed (previously referred to as N_N).

$Pr(x_r)$ is the probability of there being a change (increase or decrease) of x_r in the number of r bidders from the ith auction being completed to the $i+1^{th}$ auction being completed.

Multinomial model

Of the $k(i+1)$ participating bidders for the $i+1^{th}$ auction, the bidders can be either:

- from the $n(i,r=0)$ set - potential bidders in the market that are not yet in the database.

- from any $n(i,0<r\leq i)$ set – bidders already identified in the database that have submitted r bids so far.

At the same time, on completion of the $i+1^{th}$ auction, each of these participating $k(i+1)$ bidders will move from being a r-bidder to a $r+1$-bidder and hence they will individually generate a one unit increase in their respective $n(i+1,r+1)$ groups, while they will generate a one unit decrease in their respective $n(i+1,r)$ groups.

Obviously, the more r-bidders for the ith auction, the higher the chances that some will be promoted to the $r+1$ set in the $i+1^{th}$ auction. Namely, if we assume the

promotions happen at random, the probability of a r-bidder bidding for the $i+1^{th}$ auction is the number of r-bidders on completion of the ith auction, $n(i,r)$, divided by the total population of bidders on completion of the ith auction, $M(i)$. This is also the probability of a bidder from a r set being promoted to the $r+1$ set. In other words, the total number of bidders from the predecessor set of r-bidders, when divided by the total number of existing and potential bidders, determines the probability that the $r+1$ set increases on completion of the $i+1^{th}$ auction. Mathematically this simple idea can be expressed as $Pr(x_r)=n(i,r\text{-}1)/M(i)$, which is exactly a multinomial distribution.

The multinomial distribution is a generalization of the binomial distribution but, unlike the binomial where the outcome is either a success or failure (2 possible and complementary outcomes), the multinomial allows for multiple outcomes. The outcomes here are the "promotion" [2] of each of the $k(i+1)$ participating bidders from a $n(i,r)$ group to the $n(i+1,r+1)$. With all this, the multinomial model can be stated as:

---

[2] Until Equation 1 is simplified later into Trinomial and Binomial models, this tentative multinomial model only considers the *promotions*, not the desertions (only positive changes or increases of x_r), of bidders from the r group to the $r+1$ group. Otherwise the sum of outcomes $x_1, x_2, ..., x_i$ would be zero, since the promoted r-bidders would the same as the leavers from the $r\text{-}1$-bidders category; if not for the new bidders ($r=1$ bidders, that is, x_1) who would be the only ones that would add (they would be the only new ones), but not subtract.

$$\text{Pr}\left(\begin{array}{l} \textit{Possible outcomes}: \\ x_1 = \Delta n(i+1, r=1) = n(i+1, r=1) - n(i, r=1), \\ x_2 = \Delta n(i+1, r=2) = n(i+1, r=2) - n(i, r=2), \\ x_3 = \Delta n(i+1, r=3) = n(i+1, r=3) - n(i, r=3), \\ ... \\ x_i = \Delta n(i+1, r=i) = n(i+1, r=i) - n(i, r=i), \\ x_{i+1} = \Delta n(i+1, r=i+1) = \\ \qquad = n(i+1, r=i+1) - n(i, r=i+1), \\ \\ \textit{Number of trials}: \ k(i+1), \\ \\ \textit{Probabilities of each outcome}: \\ \\ p_1 = \dfrac{n(i, r=0)}{M(i)} = \dfrac{M(i) - \sum\limits_{j=1}^{i} n(i, r=j)}{M(i)} = 1 - \sum\limits_{j=2}^{i+1} p_j, \\ \\ p_2 = \dfrac{n(i, r=1)}{M(i)}, \\ \\ p_3 = \dfrac{n(i, r=2)}{M(i)}, \\ ... \\ p_i = \dfrac{n(i, r=i-1)}{M(i)}, \\ \\ p_{i+1} = \dfrac{n(i, r=i)}{M(i)} \end{array}\right) = \frac{k(i+1)!}{x_1! x_2! x_3! ... x_i! x_{i+1}!} p_1^{x_1} p_2^{x_2} p_3^{x_3} ... p_i^{x_i} p_{i+1}^{x_{i+1}}$$

(1)

with

$$0 \le x_1, x_2, x_3, ..., x_i, x_{i+1} \le k(i+1) \quad and \quad \sum_{j=1}^{i+1} x_j = k(i+1)$$

Therefore, for a series of $k(i+1)$ independent trials, each of which leads to a success for exactly one of r categories (sets of $n(i, r=1,2,...i+1)$ bidders at auction $i+1$), with each category having a given fixed success probability p_i (equalling $n(i, r=i-1)/M(i))$, the multinomial distribution gives the probability of any particular combination of numbers of successes for the various $i+1$ categories (success

13

understood here as the *promotion* – desertions not counted – of one of the $k(i+1)$ participating bidders from set $n(i,r)$ to set $n(i+1,r+1)$).

The problem now is that $M(i)$ is not known, and that this multinomial distribution has a plethora of possible outcomes (different combinations of *r*-bidders promotions to the *r*+1 groups) for a higher number i of auctions. Therefore, it is quite difficult to know if the outcomes that have been obtained at each step (as we add one more auctions and the number of $n(i,r)$ sets also increase) are similar to one of the most likely outcomes whose probabilities would be obtained by (1). This forces us to resort to a simplified trinomial (for PMF =*r*-bidder curves as in top Figure 1) or binomial (for CDF ≥*r*-bidder curves as in bottom Figure 1) expressions which are dealt with in the next section.

Bidder population size

As has already been noted, an estimate of the bidder population size $M(i)$ is indispensable for the multinomial model. However, despite being an important figure for any auctioneer or bidder that operates in a particular market, its estimation has yet to be considered in the construction management literature. Particularly, it is important to remember that $M(i)=n(i,r=0)+n(i,r=1)+…+ n(i, r=i)$, from which the number of sets $n(i,r≥0)$ are known as they have been observed (counted) in the previous i auctions, but not the group of $n(i,r=0)$ whose size is not known and also varies as auctions are completed.

A simple, yet powerful, method for estimating $M(i)$ relates to the ratio of new to not new bidders in each auction. This estimation method assumes that every bidder has the same probability of participating, which will not be true for a small group of intensive bidders, but it appears to work well for the overall bidder population, particularly with open auctions, as reported later in the *Results*.

To understand the method, suppose that, as bids are opened and contracts awarded, a register is kept of all the different bidders that have submitted at least one bid. It is expected that during the first auctions, most of the bidders will be "new" (they will have not submitted a bid before). However, as the number of auctions increase, the proportion of non-new (registered) bidders will start to catch up with the proportion of new bidders (who are registered for the first time on completion of the ith auction). Once these proportions become equal, for instance, we can assume that the probability of new and non-new bidders being present in an auction are also equal, which implies that the populations of new and non-new bidders are equal too. Hence, the total population of new and non-new bidders is twice that on the non-new bidders – the ones already in the database! Assume, for example, that on completion of auction i-1=24, 100 different bidders have been identified and that, for auction i=25, there are $k(i=25)=8$ participating bidders, 4 bidders of which are new and 4 have submitted at least one bid in previous encounters. The number of different bidders in the database is now 100, so the estimated total population of potential bidders is 200.

Of course, the higher the number of $k(i)$ bidders in auction i, the higher the accuracy of the estimation, but the number of participating bidders tend to be rather small (especially in selective auctions).

The estimate of $M(i)$ at auction i is therefore

$$M(i) = \frac{n(i-1, r \geq 1)}{\dfrac{k(i) - x_1}{k(i)}} = \frac{k(i) \cdot n(i-1, r \geq 1)}{k(i) - n(i, r \geq 1) + n(i-1, r \geq 1)} \tag{2}$$

where the numerator represents the number of "different" bidders identified until the previous auction (frequent bidders when auction i-1 is completed) divided by $(k(i)-x_1)/k(i)$ which corresponds to the proportion of "frequent" bidders (out of the

k(i) participating bidders in auction *i*). Substituting the values of the numerical example above into (2) gives

$$M_{i=25} = \frac{k(i) \cdot n(i-1, r \geq 1)}{k(i) - n(i, r \geq 1) + n(i-1, r \geq 1)} = \frac{8 \cdot 100}{8 - 104 + 100} = \frac{800}{4} = 200 \: bidders$$

as expected.

Figure 2 represents the evaluations of (2) for all auctions *i* (from 1 to 373) provided $x_1 \neq k(i)$ in our London dataset. Of course, every time the denominator in (2) equals zero (when the number of new bidders x_1 from auction *i*-1 to auction *i* equals the number of participating bidders *k(i)*), M_i goes to +infinity, so these points are not included. Also, shown in Figure 2 is the best regression expression found. This has a coefficient of determination of 0.5611, which is sufficient for a rough approximation like this.

< Insert Figure 2 here >

In addition to providing the first approximation to the number of potential bidders in a given market, Figure 2 provides further interesting information, such the fact that this population keeps expanding over time (since once a bidder has been identified it cannot be a new bidder anymore). That is, it is not a constant value and, despite some bidders dropping out of the market, their cumulative registered number is monotonically increasing. It must be borne in mind, however, that this is not equivalent to say that the number of "new" bidders will keep growing. As noted above, it is quite likely that, sooner or later, the proportion of frequent bidders will grow faster than *M(i)*, but without being able to exceed this value either.

Calculations

Simplified multinomial model when the number of bidders is known

The problem with the multinomial model is the extraordinarily high number of combinations of sets of *n(i,r)* bidders, among which the *k(i)* participating bidders of

16

an auction *i* can be distributed at the same time. Two simplifications of the multinomial expression in (1) are proposed. First, a trinomial expression for the number of bidders that have submitted "exactly" *r* bids; second, a binomial expression for modelling the number of bidders that have submitted "at least" *r* bids.

Therefore, a simplified expression of (1) for modelling the *r*-only bidders is a trinomial distribution (multinomial with just three outcomes) that focuses on one specific *n(i,r)* set at a time. However, whereas in the multinomial expression we only observed the increasing (promotions) numbers of bidders in all *r* sets at auction *i+1* (because it was implicitly assumed that an increment in *n(i+1,r)* meant a decrement in *n(i,r-1)*), now it is necessary to take into account both the inflows (bidders that are promoted from the *r-1* set at the *i*[th] auction to the *r* set at the *i+1*[th] auction) and outflows (bidders that are promoted from the current *r* set at the *i*[th] auction to the *r+1* set at the *i+1*[th] auction) in the trinomial model.

Consider a specific set of *r* bidders. For example if, on completion of the 50th auction in the database there are 15 twice-only bidders, then *n(i,r)=15* for *i=50* and *r=2*. For auction *i+1*, let x_+ represent the number of bidders arriving from the *r-1* set that increase *n(i,r)*, x_- represents the number of bidders departing to the *r+1* set that decrease *n(i,r)*, and $x_=$ represents the number of bidders that cannot cause an increase nor a decrease in *n(i,r)* as they do not belong to either the *r* set or *r-1* set, so that

$x_+ + x_- + x_= = k(i+1)$, the number of bidders for auction *i+1*.

Hence, the trinomial expression for a given *n(i,r)* set is:

$$= Trinomial \begin{pmatrix} x_+ = \Delta^+ n(i+1,r), & x_- = \Delta^- n(i+1,r), & x_= = k(i+1) - x_+ - x_-, \\ n^o\,trials = k(i+1), & p_+ = \dfrac{n(i,r-1)}{M(i)}, & p_- = \dfrac{n(i,r)}{M(i)}, & p_= = 1 - p_+ - p_- \end{pmatrix} =$$

$$= \frac{k(i+1)!}{x_+! x_-! (k(i+1) - x_+ - x_-)!} p_+^{x_+} p_-^{x_-} \left(1 - p_+ - p_-\right)^{k(i+1) - x_+ - x_-}$$

(3)

However, this expression still requires further work, since there are several combinations of x_+, x_- and $x_=$ that produce the same x_r value (increment or decrement of the number of bidders on the $n(i,r)$ set from auction i to auction $i+1$, i.e., $n(i+1,r)=n(i,r)+x_r$, where $x_r=x_+-x_-$ with $-k(i+1) \leq x_r \leq k(i+1)$). For example, $n(i,r)$ can remain constant ($x_r=0$) when:

- all $k(i+1)$ bidders are $x_=$.

- there is one x_+ and one x_-, while the remaining $k(i+1)-2$ bidders are $x_=$.

- there are two x_+ and two x_-, while the remaining $k(i+1)-4$ bidders are $x_=$.

- …

Therefore, what is wanted is to add the "probabilities" obtained by different evaluations of (3) which produce the same "equivalent outcome x_r". Adding these equivalent outcomes to compute the probability $Pr(x_r)$ of a specific value of x_r involves counting all the possible ways this value of x_r can occur as a combination of x_+, x_- and $x_=$ taking into account that each combination corresponds to one evaluation of (3).

Table 1 contains the values of x_r of all possible pairs of x_+ and x_-, since it is assumed that $x_==k(i+1)-x_+-x_-$ for the $i+1$[th] auction.

< **Insert Table 1 here** >

In Table 1, the blank cells represent the impossible combinations of x_+ and x_- since $x_+ + x_- \leq k(i+1)$. Also, as can be observed, the number of ways the same x_r values can occur coincide with the diagonal elements in the Table. By induction, it is not difficult to develop the generic expression that summarises the count of all possible combinations of x_+, x_- and $x_=$ that result in the same x_r value, which is $(k(i+1)+\alpha-|x_r|)/2$ with a parameter α that discriminates when the x_r value is odd ($\alpha=1$) or even ($\alpha=2$).

Therefore, the extension of (3) considering the sum of all the equivalent x_r outcomes from Tables 1 is

$$\Pr\left(x_r = \Delta n(i+1, = r) = n(i+1, = r) - n(i, = r) \quad \text{with} -k(i+1) \le x_r \le k(i+1)\right) =$$

$$= \sum_{s=1}^{(k(i+1)+\alpha-|x_r|)/2} Trinomial \left(\begin{array}{l} x_+ = \delta_x x_r + s - 1, \\ x_- = (\delta_x - 1)x_r + s - 1, \\ x_= = k(i+1) - x_+ - x_- = k(i+1) - |x_r| - 2(s-1), \\ n^\circ\,trials = k(i+1), \\ p_+ = \dfrac{n(i, r-1)}{M(i)}, \\ p_- = \dfrac{n(i, r)}{M(i)}, \\ p_= = 1 - p_+ - p_- \end{array} \right) = \qquad (4)$$

$$\text{with } \delta_x = 1 \text{ if } x_r \ge 0 \text{ and } \delta_x = 0 \text{ if } x_r < 0$$

$$= \sum_{s=1}^{(k(i+1)+\alpha-|x_r|)/2} \frac{k(i+1)!}{x_+! x_-! (k(i+1)-x_+-x_-)!} p_+^{x_+} p_-^{x_-} \left(1 - p_+ - p_-\right)^{k(i+1)-x_+-x_-} =$$

$$= \sum_{s=1}^{(k(i+1)+\alpha-|x_r|)/2} \left\{ \begin{array}{l} \dfrac{k(i+1)!}{(\delta_x x_r + s - 1)!((\delta_x - 1)x_r + s - 1)!(k(i+1)-|x_r|-2(s-1))!} \cdot \\ \cdot \left(\dfrac{n(i, r-1)}{M(i)}\right)^{\delta_x x_r + s - 1} \left(\dfrac{n(i, r)}{M(i)}\right)^{(\delta_x - 1)x_r + s - 1} \left(1 - \dfrac{n(i, r-1)+n(i, r)}{M(i)}\right)^{k(i+1)-|x_r|-2(s-1)} \end{array} \right\}$$

where the index "*s*" in (4) varies the values of x_+, x_- and $x_=$ so that they take the opportune diagonal values according to Table 1. Index *s* always starts by taking on the value 1 and finishes when it reaches the total amount of different *(k(i+1)+α-|x_r|)/2* combinations.

On the other hand, (4) is valid just for the number of bidders that have submitted "exactly" *r* bids, now a binomial expression for modelling the number of bidders that have submitted "at least" *r* bids (the CDF that models the values of *n(i,≥r)* sets over time) is presented by means of a simple binomial expression (multinomial with just two outcomes).

In the cumulative version (which increases every time a "new" *r-1*-bidder is promoted to a *r*-bidder, but it does not decrease every time a bidder leaves the *r*-bidders category to become an *r+1*-bidder), each *n(i,r)* set increases with the same probability p_+ that was used for the trinomial model. Therefore, this time the expression is straightforward

$$\Pr\left(x_r = \Delta n(i+1, \geq r) = n(i+1, \geq r) - n(i, \geq r) \quad \text{with } 0 \leq x_r \leq k(i+1)\right) =$$

$$= Binomial\left(\begin{array}{l} N^o \text{ of "new" r-bidders} = x_r, \\ N^o \text{ trials} = k(i+1), \\ Prob. \text{ of success}: p_+ = \dfrac{n(i, r-1)}{M(i)} \end{array}\right) = \dfrac{k(i+1)!}{x_r!(k(i+1)-x_r)!} p_+^x \left(1-p_+\right)^{k(i+1)-x_r} \quad (5)$$

Numerical examples

Since trinomial expression (4), even after simplification, may appear to be particularly daunting, a brief numerical example is provided for the ease of understating. This example will also include the application of (5).

Assume it is desired to estimate the probability that the current set of twice-bidders (*n(i,r=2)*) will decrease by one bidder (*$x_{r=2}$=-1*) for the next auction given that there will be 8 participating bidders involved. That is, to find *Pr(x_r=-1)* for *r=2* on completion of auction *i+1* with *k(i+1)=8*.

For the case of *k(i+1)=8* bidders, Table 2 particularises Table 1.

< Insert Table 2 here >

Now, *(k(i+1)+α-/x_r/)/2=(8+1-/-1/)/2=4* is first calculated with *α=1* since *x_r=-1* is considered an odd number. Therefore, index *s* from (4) will take on the values 1, 2, 3 and 4. The *x_+*, *x_-* and *$x_=$* trinomial variables depend on:

- *x_r* and /*x_r*/, which are -1 and 1, respectively.
- *δ_x* which equals 0 when *x_r* is negative (as is our example, since *x_2=-1*) and equals 1 when *x_r* is positive (the case for *x_r=0* is irrelevant)
- *k(i+1)* which equals 8.

20

Table 3 now summarises all the values for x_+, x_- and $x_=$ as index s changes:

< Insert Table 3 here >

It is easy now to realise that these results are forcing the x_+ and x_- through the right cells in Table 2 (which are $(x_-, x_+)=(1,0)$, $(2,1)$, $(3,2)$ and $(4,3)$, in bold and underlined text).

The only variables that remain to be calculated are the ones corresponding to the probabilities p_+, p_- and $p_=$. These probability values do not depend on the index s, but on the current number of bidders from the same and immediately previous $n(i,r)$ group. In this case, we are interested in obtaining the probabilities for $n(i+1,r=2)$. Suppose we know that $M(i)=100$, $n(i,r=1)=25$, $n(i,r=2)=15$ and $n(i,r=3)=10$, being the other groups with $r>3$ with 0 bidders at this point. Then, the probability values from (4) are simply

$$p_+ = \frac{n(i,r-1)}{M(i)} = \frac{n(i,r=2-1=1)}{100} = \frac{n(i,r=1)}{100} = \frac{25}{100} = 0.25$$

$$p_- = \frac{n(i,r)}{M(i)} = \frac{n(i,r=2)}{100} = \frac{15}{100} = 0.15$$

$$p_= = 1 - p_+ - p_- = 1 - 0.25 - 0.15 = 0.60$$

Hence, by using equation (4), the probabilities associated with a variation of $x_r=-1$ bidders in the group of $n(i,r=2)$ bidders will be:

$$\Pr\left(\Delta n(i+1,r=2) = n(i+1,r=2) - n(i,r=2) = x_r = x_2 = -1\right) =$$

$$= \frac{8!}{0!1!7!} 0.25^0 0.15^1 0.60^7 + \frac{8!}{1!2!5!} 0.25^1 0.15^2 0.60^5 + \frac{8!}{2!3!3!} 0.25^2 0.15^3 0.60^3 +$$

$$+ \frac{8!}{3!4!1!} 0.25^3 0.15^4 0.60^1 = 0.0336 + 0.0735 + 0.0255 + 0.0013 = 0.1339$$

On the other hand, the binomial expression (5) is easier, however, only positive x_r values are allowed, since, by definition, the number of *r-at least* bidders is monotonically increasing (since once they have submitted r bids, they cannot leave the set of bidders that have submitted "at least" r bids, even by submitting more

bids). Assuming the same variable values so far with the exception of x_r, that will be now considered to equal 3 ($x_r=3$). With this, (5) becomes

$$\Pr\left(x_r = \Delta n(i+1, \geq r = 2) = n(i+1, \geq 2) - n(i, \geq 2) = x_r = x_2 = 3\right) =$$

$$= Binomial \left(\begin{array}{l} N^o \text{ of "new" twice } (r = 2) \text{ -bidders} = x_r = 3, \\ N^o \text{ trials} = k(i+1) = 8, \\ Prob. \text{ of success}: p_+ = \dfrac{n(i, r-1)}{M(i)} = \dfrac{n(i,1)}{M(i)} = \dfrac{25}{100} = 0.25 \end{array} \right) =$$

$$= \frac{8!}{3!(8-3)!} 0.25^3 (1-0.25)^{8-3} = 0.2076$$

As can be seen, calculations are not difficult and can be easily processed with the aid of standard spreadsheet software.

Simplified multinomial model when the number of bidders is unknown

On many occasions, particularly in Open tendering, neither the bidders nor the contracting authority will know (at least not accurately) how many bidders will participate in $i+1$[th] auction, therefore, the number of bidders for auction $i+1$ has to be considered as a stochastic variable.

For simplification purposes, it is assumed here that $k(i+1) \sim Poisson(\lambda)$, where λ is calculated as the average of the participating bidders. However, any other distribution would be feasible whenever it is discrete and standardised within the interval $[0, +\infty]$. Therefore, the trinomial expression from (4), where the new index "j" takes the values of all possible $k(i+1)$ participating bidders combinations, would now be

$$\Pr\left(x_r = \Delta n(i+1,=r) = n(i+1,=r) - n(i,=r) \quad \text{with} -\infty \le x_r \le +\infty\right) =$$

$$= \sum_{j=|x_r|}^{+\infty} \left\{ Poisson(j,\lambda) \cdot \sum_{s=1}^{(j+\alpha-|x_r|)/2} Trinomial \begin{pmatrix} x_+ = \delta_x x_r + s - 1, \\ x_- = (\delta_x - 1)x_r + s - 1, \\ x_= = j - x_+ - x_- = j - |x_r| - 2(s-1), \\ n^\circ\,trials = j, \\ p_+ = \dfrac{n(i,r-1)}{M(i)}, \\ p_- = \dfrac{n(i,r)}{M(i)}, \\ p_= = 1 - p_+ - p_- \end{pmatrix} \right\} \tag{6}$$

$$= \sum_{j=|x_r|}^{+\infty} \left\{ \frac{e^{-\lambda}\lambda^j}{j!} \cdot \sum_{s=1}^{(k(i+1)+\alpha-|x_r|)/2} \left[\frac{j!}{(\delta_x x_r + s - 1)!((\delta_x-1)x_r+s-1)!(j-|x_r|-2(s-1))!} \cdot \right. \right.$$
$$\left. \left. \left(\frac{n(i,r-1)}{M(i)}\right)^{\delta_x x_r + s - 1} \left(\frac{n(i,r)}{M(i)}\right)^{(\delta_x-1)x_r+s-1} \left(1 - \frac{n(i,r-1)+n(i,r)}{M(i)}\right)^{j-|x_r|-2(s-1)} \right] \right\}$$

whereas the binomial expression would be:

$$\Pr\left(x_r = \Delta n(i+1,\ge r) = n(i+1,\ge r) - n(i,\ge r) \quad \text{with } 0 \le x_r \le +\infty\right) =$$

$$= \sum_{j=x_r}^{+\infty} \left\{ Poisson(j,\lambda) \cdot Binomial \begin{pmatrix} N^\circ\ of\ "new"\ r\text{-}bidders = x_r, \\ N^\circ\,trials = j, \\ Prob.of\ success: p_+ = \dfrac{n(i,r-1)}{M(i)} \end{pmatrix} \right\} = \tag{7}$$

$$= \sum_{j=x_r}^{+\infty} \left\{ \frac{e^{-\lambda}\lambda^j}{j!} \cdot \frac{j!}{x_r!(j-x_r)!} p_+^x (1-p_+)^{j-x_r} \right\} = \sum_{j=x_r}^{+\infty} \left\{ \frac{e^{-\lambda}\lambda^j}{x_r!(j-x_r)!} p_+^{x_r} (1-p_+)^{j-x_r} \right\}$$

which are relatively easy expressions to handle, since they just require calculating the right (Trinomial or Binomial) half as in the *Numerical examples* section for a single x_r value but when the number of $k(i+1)=j$ participating bidders varies according to a Poisson distribution, and then multiply all the possible outcomes by their respective probability that that number of bidders will occur (obtained by the Poisson distribution).

Results

As can be easily observed, calculations with (4) and, to a lesser extent with (5), are extensive, yet not difficult, as the number i of auctions grow and the number of sets $n(i,r)$ to be recalculated for auction $i+1$ also grow accordingly. Also, after each auction i is concluded, the size of each $n(i,r)$ set (with $r=1,2,..., i$) has to be updated. This is the reason why here, only calculations for probably the most important group, $r=1$, are shown and a thorough comparative analysis presented. Experiments performed with other $n(i,r)$ sets as well as in other databases, however, indicate that it is precisely the $n(i,r=1)$ group that is the most difficult to forecast accurately. Hence, the results shown here will be representative of the worst predictions the model can offer, and therefore the minimum performance to be expected.

Now, some aspects have to be attended to before the results are presented. The multinomial models (trinomial for the $n(i,r)$ sets and binomial for the $n(i, \geq r)$ sets) are incremental. That is, on completing the ith auction, they require gathering the values of the numbers of bidders belonging to the r sets of $n(i,r)$, and along with the estimated $M(i)$ value (as a function of i according to the regression result in Figure 2), they launch i forecasts (one for each $n(i,r)$ set) for auction $i+1$). These models are probabilistic (they give the probabilities for each possible increment of bidders x_r for auction $i+1$ for the i sets of $n(i+1,r)$ bidders), which is why here, with far more than two dimensions to work with, a graphical representation is not possible. Instead, the expected value of these predictions has been used (in order to provide a single numerical result for each auction i for the group of $n(i,r=1)$ and $n(i,r\geq1)$ bidders). This enables this average value to be compared with the actual $x_{r=1}$ values observed from auction i^{th} to auction $i+1^{th}$.

Furthermore, since the forecasts obtained by this multinomial model are naturally incremental (they oscillate around or near zero most of the time), the chosen metric for comparing performance is the *absolute deviation* from the actual values. This is

24

instead of the quadratic deviations, which usually give biased results when working steadily near zero. The absolute deviations measures for the 373 auctions and for each sub model (trinomial and binomial), and the sum, average and maximum value of these residuals are included in Table 4.

It is noted that the model has been assessed under two schemes: one in which the number of participating bidders $k(i+1)$ was known for auction $i+1$ (as in Selective tendering), and other in which this number was not known (as in Open tendering). It is worth highlighting that all deviations when $k(i+1)$ were known have been divided by the respective $k(i+1)$ value from each auction $i+1$ in order to express them irrespective of the number of participants.

< Insert Table 4 here >

Here, as stated earlier, a Poisson distribution is used to model the number of participating bidders for situations in which $k(i+1)$ is not known and the use of another more suitable distribution may possibly have improved the results of the multinomial models in the last column of Table 4. However, despite all the absolute deviation values being quite satisfactory in general, the best performance values in Table 4 are obtained by the binomial model.

For the binomial model, even when the number of bidders for the next auction $k(i+1)$ is not known, the absolute (error) deviation remains below one bidder (0.80) on average. Furthermore, even the absolute deviation of the trinomial model when $k(i+1)$ is not known (the worst result from the Table) is a relatively small value (1.33); more if we take into account that the average number of participating bidders per auction in this database is 5.13 bidders, that is, the average variation of the $x_{r=1}$ value can range from -5.13 to 5.13 per auction (on average).

The complete set of calculations and results can be found as supplemental online material, as they are too extensive to be included in the main body of the paper or in an Appendix.

25

Conclusions

Estimating the number of "new" bidders for upcoming auctions as well as the approximate size of the population of potential bidders has important competitive implications in real-life bidding scenarios. Similarly, most Bid Tender Forecasting models need to somehow handle the appearance of new bidders from whom there is no registered information concerning both their competitiveness profiles and the number of these "unpredictable" bidders. However, literature references to either of these two topics are very scarce, especially in the Construction Management context.

For the very first time, one model has been developed for anticipating the proportion and number of these new bidders, which are useful for both Open (where the number of participants is unknown) and Selective (where the number of participants is generally known) tendering. A performance analysis of the deviations for the model applied to set of London tenders indicates that a multinomial distribution model will constitute a good alternative for situations in which the number of "new" bidders needs to be predicted, as well as for the number of previously identified bidders who are submitting more bids for the next auctions.

Both outcomes are relevant in the tendering context, as, for example, the second result is useful for describing the variations of the level of experience of potential bidders, assuming, as the literature has commented on many occasions, that those bidders who submit more bids are generally more successful. On the other hand, the multinomial model is also useful to describe how many bidders of whom we know nothing about will probably take part in a future auction, together with how long it will take them to submit their second bid on average, i.e., how often they will participate. In essence, this model provides a better understanding of the frequency of bid submission by bidders.

The implementation of the multinomial model requires a first approximation of the population size of potential bidders. Superficially, it may be thought that this population size should not be difficult to assess, at least in those cases in which there is a categorised pre-qualification system where potential bidders have to be registered beforehand if they want to submit bids for a certain kind of projects. However, this is not entirely true, since there are many aspects that make such pre-qualification systems actually *open lists* of potential bidders. Firstly, most systems are not strictly compulsory and allow the participation of bidders who are not registered provided they prove their technical and economic solvency to carry out the job - otherwise many foreign contractors would not be allowed to take part, which might be considered discriminatory. Of course, the non-enlisted course of action is always more time-consuming and requires more bureaucracy, but it is not unusual for big budget contracts to attract several international competitors. Second, a pre-qualification system cannot always take into account the contractor's proximity to the project location, nor the idle capacity of each contractor over time, both of which are important determinants of the probability of bidder participation in terms of how many might participate and how likely it is that each will participate. Third, other possibilities are that bidders might form groups of companies (diminishing the absolute number of competitors); a project can require the participation of several disciplines or areas of expertise (making the count of potential pre-qualified bidders more difficult); and a common reality is that pre-qualification lists are not always updated.

Furthermore, although the intention here is to develop a method and to illustrate its use by application to sample of data that happen to be obtained from the London area, the empirical results of further application to other similar and more recent databases are also worthy of note. One is to confirm that the graphical trends and shapes of the curves in Figure 1 do not appear to differ significantly for other similar,

more recent databases. Similarly, the highest R^2 values when potential expressions like the one presented in Figure 2 for the estimation of the population size of bidders are used, also continuing to increase as more auctions are added.

Finally, it is to be expected that the models developed here could also be applied to other markets where auctions have larger number of bidders and auctions that are rather similar (the complexity of the product does not vary) such as government sales of used cars for instance. In those auctions bidders behave, at least on average, more systematically and repetitively, which also allows a modelling like the one proposed here.

To sum up, the method developed to assess the bidders population size is the very first one in the bidding literature, and that this indicates that the population keeps increasing over time also sheds some light on other tendering outcomes, such as, for instance, why some bidders bid repeatedly achieve higher participation ratios and why the steady inflow of new bidders is generally never interrupted. However, the model presented is far from perfect. For example, to somehow include the contract economic size in the models is likely to improve their accuracy, as this variable has been positively correlated to the "total" number of participating bidders. In addition, these models rely on rough estimates of bidder population size, which will need to be improved for the accuracy to be increased. In this vein, we have recently started work on a new deterministic Beta Binomial model which will offer a simpler estimation of the number of new and repetitive bidders to be found in future auctions. This alternative is expected to provide the practitioner with a simpler technique to anticipate only the *average number* of new and frequent (compared to the Multinomial model, which is able to calculate the probabilities associated with each potential number of r-bidders), but with the appeal of reduced calculation complexity and with the advantage of not being dependent on an estimation of the

population size of bidders. More work is still required on this alternative model and will probably be part of a future research paper.

Finally, another line of future research will be to identify a way to conveniently model the highly frequent versus sporadic bidders (which is detrimental for the multinomial model since it necessarily assumes that all the bidder population has the same probability of participating) might also lead to more accurate predictive results. The approaches developed, therefore, constitute just the first treatment of a new problem whose future analysis offers considerable promise for understanding these and other important tendering phenomena.

References

Ahmad, I., & Minkarah, I. (1988). Questionnaire Survey on Bidding in Construction. *Journal of Management in Engineering*, 4(3), 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. *International Journal of Project Management*, 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. *Applied Mechanics and Materials*, 567, 595–600. doi:10.4028/www.scientific.net/AMM.567.595

Ballesteros-Pérez, P., González-Cruz, M. C., & Pastor-Ferrando, J. P. (2010). Analysis of construction projects by means of value curves. *International Journal of Project Management*, 28(7), 719–731. doi:10.1016/j.ijproman.2009.11.003

Ballesteros-Pérez, P., González-Cruz, M. C., & Cañavate-Grimal, A. (2012). Mathematical relationships between scoring parameters in capped tendering. *International Journal of Project Management*, 30(7), 850–862. doi:10.1016/j.ijproman.2012.01.008

Ballesteros-Pérez, P., González-Cruz, M. C., Cañavate-Grimal, A., & Pellicer, E. (2013a). Detecting abnormal and collusive bids in capped tendering. *Automation in Construction*, 31, 215–229. doi:10.1016/j.autcon.2012.11.036

Ballesteros-Pérez, P., González-Cruz, M. C., & Cañavate-Grimal, A. (2013b). On competitive bidding: Scoring and position probability graphs. *International*

Journal of Project Management, 31(3), 434–448. doi:10.1016/j.ijproman.2012.09.012

Ballesteros-Pérez, P., González-Cruz, M. C., Fernández-Diego, M., & Pellicer, E. (2014). Estimating future bidding performance of competitor bidders in capped tenders. *Journal of Civil Engineering and Management*, 1–12. doi:10.3846/13923730.2014.914096

Ballesteros-Pérez, P., del Campo-Hitschfeld, M. L., Mora-Melia, D., & Domínguez, D. (2015a). Modeling bidding competitiveness and position performance in multi-attribute construction auctions. *Operations Research Perspectives*, 2, 24–35. doi:10.1016/j.orp.2015.02.001

Ballesteros-Pérez, P., Skitmore, M., Das, R., & del Campo-Hitschfeld, M. L. (2015b). Quick Abnormal-Bid–Detection Method for Construction Contract Auctions. *Journal of Construction Engineering and Management*, 141(7), 04015010. doi:10.1061/(ASCE)CO.1943-7862.0000978

Ballesteros-Pérez, P., González-Cruz, M. C., Fuentes-Bargues, J. L., & Skitmore, M. (2015c). Analysis of the distribution of the number of bidders in construction contract auctions. *Construction Management and Economics*, In press. doi:10.1080/01446193.2015.1090008.

Ballesteros-Pérez, P., Skitmore, M., Pellicer, E., & González-Cruz, M. C. (2015d). Scoring rules and abnormally low bids criteria in construction tenders: a taxonomic review. *Construction Management and Economics*, 33(4), 259–278. doi:10.1080/01446193.2015.1059951

Ballesteros-Pérez, P., Skitmore, M., Pellicer, E., & Zhang, X. (2016a). Scoring Rules and Competitive Behavior in Best Value Construction Auctions. *Journal of Construction Engineering and Management*, In press, http://dx.doi.org/10.1061/(ASCE)CO.1943-7862.0001144.

Ballesteros-Pérez, P., Skitmore, M., Pellicer, E., & Gutiérrez-Bahamondes, J. H. (2016b). Predicting the identity of bidders in procurement auctions. *International Journal of Project Management*, 34 (2). pp. 158-172. ISSN 0263-7863 doi: 10.1016/j.ijproman.2015.11.001.

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

Capen, E. C., Clapp, R. V., & Campbell, W. M. (1971). Competitive Bidding in High-Risk Situations. *Journal of Petroleum Technology*, 23, 641–653.

Carr, R. I. (1982). General bidding model. *Journal of the Construction Division. Proceedings of the American Society of Civil Engineers*, 108(CO4), 639–650.

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

Dass, M., Reddy, S. K., & Iacobucci, D. (2014). Social networks among auction bidders: The role of key bidders and structural properties on auction prices. *Social Networks*, 37(1), 14–28. doi:10.1016/j.socnet.2013.11.002

Drew, D. S., & Skitmore, R. M. (2006). Competitiveness in bidding: a consultant's perspective. *Construction Management and Economics*, 10(3), 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 Journal of Economics*, 20(2), 268–279. doi:10.2307/2555693

Engelbrecht-Wiggans, R., Dougherty, E. L., & Lohrenz, J. (1986). A Model for the Distribution of the Number of Bids on Federal Offshore Oil Leases. *Management Science*, 32(9), 1087–1094. doi:10.1287/mnsc.32.9.1087

European Union (1999) Prevention, Detection and Elimination of Abnormally Low Tenders in the European Construction Industry, reference DG3 alt wg 05, dated 02 May 1999 (modified version of documents 01 to 04 as agreed at the meetings of the ALT WG).

Friedman, L. (1956). A competitive bidding strategy. *Operations Research*, 1(4), 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*.

Fu, W. K., Drew, D. S., & Lo, H. P. (2002). The effect of experience on contractors' competitiveness in recurrent bidding. *Construction Management and Economics*, 20(8), 655–666. doi:10.1080/0144619022000014060

Fu, W. K., Drew, D. S., & Lo, H. P. (2003). Competitiveness of Inexperienced and Experienced Contractors in Bidding. *Journal of Construction Engineering and Management*, 129(4), 388-395, doi: 10.1061/(ASCE)0733-9364(2003)129:4(388)

Gates, M. (1967). Bidding strategies and probabilities. *Journal of the Construction Division, Proceedings of the American Society of Civil Engineers*, 93(CO1), 75–107.

Harstad, R. M., Kagel, J. H., & Levin, D. (1990). Auctions with an uncertain number of bidders.pdf. *Economic Letters*, 33, 35–40.

Hu, A. (2011). How bidder's number affects optimal reserve price in first-price auctions under risk aversion. *Economics Letters*, 113(1), 29–31. doi:10.1016/j.econlet.2011.05.041

IMF, International Monetary Fund (2014). Report for Selected Countries and Subjects.

Ishii, N., Takano, Y. & Muraki, M. (2014) An order acceptance strategy under limited engineering man-hours for cost estimation in Engineering-Procurement-Construction projects. *International Journal of Project Management*, 32(3), 519-528. doi: 10.1016/j.ijproman.2013.07.009.

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

Levin, D., & Ozdenoren, E. (2004). Auctions with uncertain numbers of bidders. *Journal of Economic Theory*, 118, 229–251. doi:10.1016/j.jet.2003.11.007

Liu, X., Wang, D., & Jiang, Z. (2015). Simulation and Analysis of Bid Evaluation Behaviors for Multiattribute Reverse Auction. *IEEE Systems Journal*, 9(1), 165–176. doi:10.1109/JSYST.2013.2258737

McAfee, R. P., & McMillan, J. (1987). Auctions with a Stochastic Number of Bidders. *Journal of Economic Theory*, 19, 1–19.

Mohlin, E., Östling, R., & Wang, J. T. (2015). Lowest unique bid auctions with population uncertainty. *Economics Letters*, 134, 53–57. doi:10.1016/j.econlet.2015.06.009

Morin, T. L., & Clough, L. H. (1969). OPBID: Competitive bidding strategy model. *Journal of the Construction Division 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. *Construction Management and Economics*, 20(6), 473–482. doi:10.1080/01446190210151041

Oo, B.L., Drew, D., & Lo, H.-P. (2007). Modelling contractors' mark-up behaviour in different construction markets. *Engineering, Construction and Architectural Management*, 14(5), 447–462. doi:10.1108/09699980710780755

Oo, B.-L., Drew, D. S., & Lo, H.-P. (2010a). Modeling the Heterogeneity in Contractors' Mark-Up Behavior. *Journal of Construction Engineering and Management*, 136(7), 720–729. doi:10.1061/(ASCE)CO.1943-7862.0000186

Oo, B.-L., Drew, D. S., & Runeson, G. (2010b). Competitor analysis in construction bidding. *Construction Management and Economics*, 28(12), 1321–1329. doi:10.1080/01446193.2010.520721

Oo, B.L., Lo, H.-P., & Teck-Heng Lim, B. (2012). The effect of bidding success in construction bidding. *Engineering, Construction and Architectural Management*, 19(1), 25–39. doi:10.1108/09699981211192553

Rickwood, A. K. (1972). An investigation into the tenability of bidding theory and techniques, and proposal for a bidding game. *MSc Thesis. Loughborough University of Technology*.

Rubey, H., & Milner, W. W. (1966). A statistical approach to bidding. *Ed. McMillan*.

Runeson, G., & Skitmore, M. (1999). Tendering theory revisited. *Construction Management and Economics*, 17(3), 285–296. doi:10.1080/014461999371493

Shaffer, L. R., & Micheau, T. W. (1971). Bidding with competitive strategy models. *Journal of the Construction Division ASCE*, 97, 113–126.

Shash, A. A. (1993). Factors considered in tendering decisions by top UK contractors. *Construction Management and Economics*, 11(2), 111–118. doi:10.1080/01446199300000004

Skitmore, M. (1981). Why do tenders vary? *The Chartered Quantity Surveyor*, 4, 128–129.

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. *Construction Management and Economics*, 26(8), 787–803. doi:10.1080/01446190802175678

Skitmore, M. (2013). Generalised gamma bidding model. *Journal of the Operational Research Society*, 65(1), 97–107. doi:10.1057/jors.2013.18

Skitmore, M., & Pemberton, J. (1994). A Multivariate Approach to Construction Contract Bidding Mark-up Strategies. *Journal of the Operational Research Society*, 45(11), 1263–1272. doi:10.2307/2583854

Stark, R. M., & Rothkopf, M. H. (1979). Discussion of "Bidding strategies and probabilities" by Marvin Gates. *Operations Research*, 27, 364–390. doi:http://www.jstor.org/stable/170135

Takano, Y., Ishii, N., & Muraki, M. (2014) A sequential competitive bidding strategy considering inaccurate cost estimates. *Omega,* 42(1), 132–140. doi:10.1016/j.omega.2013.04.004.

Wade, R. L., & Harris, R. B. (1976). LOMARK: A Bidding Strategy. *Journal of the Construction Division*, 102(1), 197–211.

Wang, D., Liu, X. & Liu, L. (2013) Bid evaluation behavior in online procurement auctions involving technical and business experts. *Electronic Commerce Research and Applications*, 12(5), 328-336. doi:10.1016/j.elerap.2012.10.001

WB, World Bank (2013). Gross domestic product 2012, PPP.

$x_- \blacktriangledown$ $x_+ \blacktriangleright$	0	1	2	3	…	$k(i+1)$-3	$k(i+1)$-2	$k(i+1)$-1	$k(i+1)$
0	0	1	2	3	…	$k(i+1)$-3	$k(i+1)$-2	$k(i+1)$-1	$k(i+1)$
1	-1	0	1	2	…	$k(i+1)$-4	$k(i+1)$-3	$k(i+1)$-2	
2	-2	-1	0	1	…	$k(i+1)$-5	$k(i+1)$-4		
3	-3	-2	-1	0	…	$k(i+1)$-6			
…	…	…	…	…	…				
$k(i+1)$-3	$-k(i+1)$+3	$-k(i+1)$+4	$-k(i+1)$+5	$-k(i+1)$+6					
$k(i+1)$-2	$-k(i+1)$+2	$-k(i+1)$+3	$-k(i+1)$+4						
$k(i+1)$-1	$-k(i+1)$+1	$-k(i+1)$+2							
$k(i+1)$	$-k(i+1)$								

Table 1: Possible combinations of x_r values as a function of the x_+ and x_- values.

x_- ▼ x_+ ►	0	1	2	3	4	5	6	7	8
0	0	1	2	3	4	5	6	7	8
1	*-1*	0	1	2	3	4	5	6	
2	-2	*-1*	0	1	2	3	4		
3	-3	-2	*-1*	0	1	2			
4	-4	-3	-2	*-1*	0				
5	-5	-4	-3	-2					
6	-6	-5	-4						
7	-7	-6							
8	-8								

Table 2: Possible combinations of x_r when $k(i+1)=8$.

| s ►	1	2	3	4
$x_+ = \delta_x x_r + s - 1$	$0\cdot(-1)+1-1=0$	$0\cdot(-1)+2-1=1$	$0\cdot(-1)+3-1=2$	$0\cdot(-1)+4-1=3$
$x_- = (\delta_x - 1)x_r + s - 1$	$(0-1)(-1)+1-1=1$	$(0-1)(-1)+2-1=2$	$(0-1)(-1)+3-1=3$	$(0-1)(-1)+4-1=4$
$x_= = k(i+1) - x_+ - x_-$	$8-0-1=7$	$8-1-2=5$	$8-2-3=3$	$8-3-4=1$

Table 3: Calculation of x_+, x_- and $x_=$ as a function of s, x_r, δ_x and $k(i+1)$ in the

numerical example

r-bidder curve	(Sub) Model	Absolute dev.	k(i+1) known	k(i+1) unknown
n(i,r=1) set (PMF)	Multinomial (Trinomial)	Sum	103.22	497.50
		Average	0.28	1.33
		Maximum	1.28	5.85
n(i,r≥1) set (CDF)	Multinomial (Binomial)	Sum	55.27	297.55
		Average	0.15	0.80
		Maximum	0.82	3.96

Table 4: Summary of the model performance as a function of the absolute deviations

between actual and forecasted values
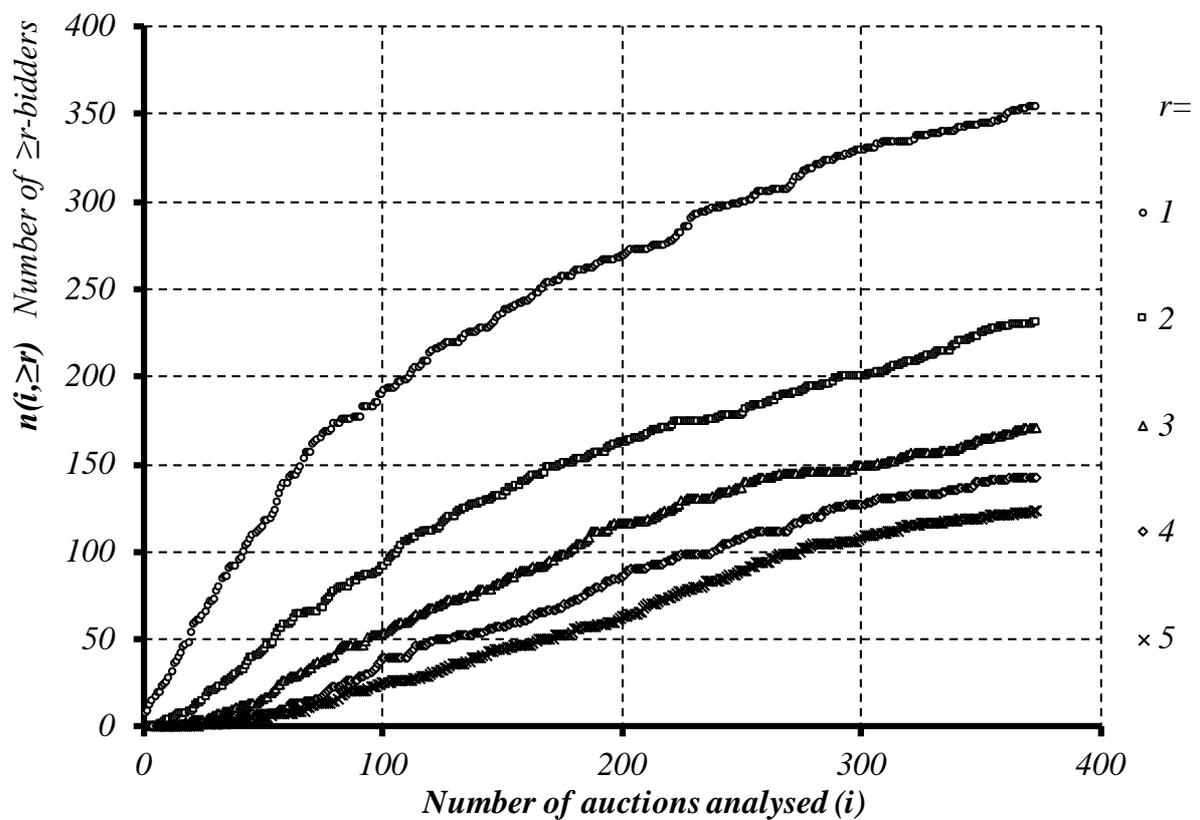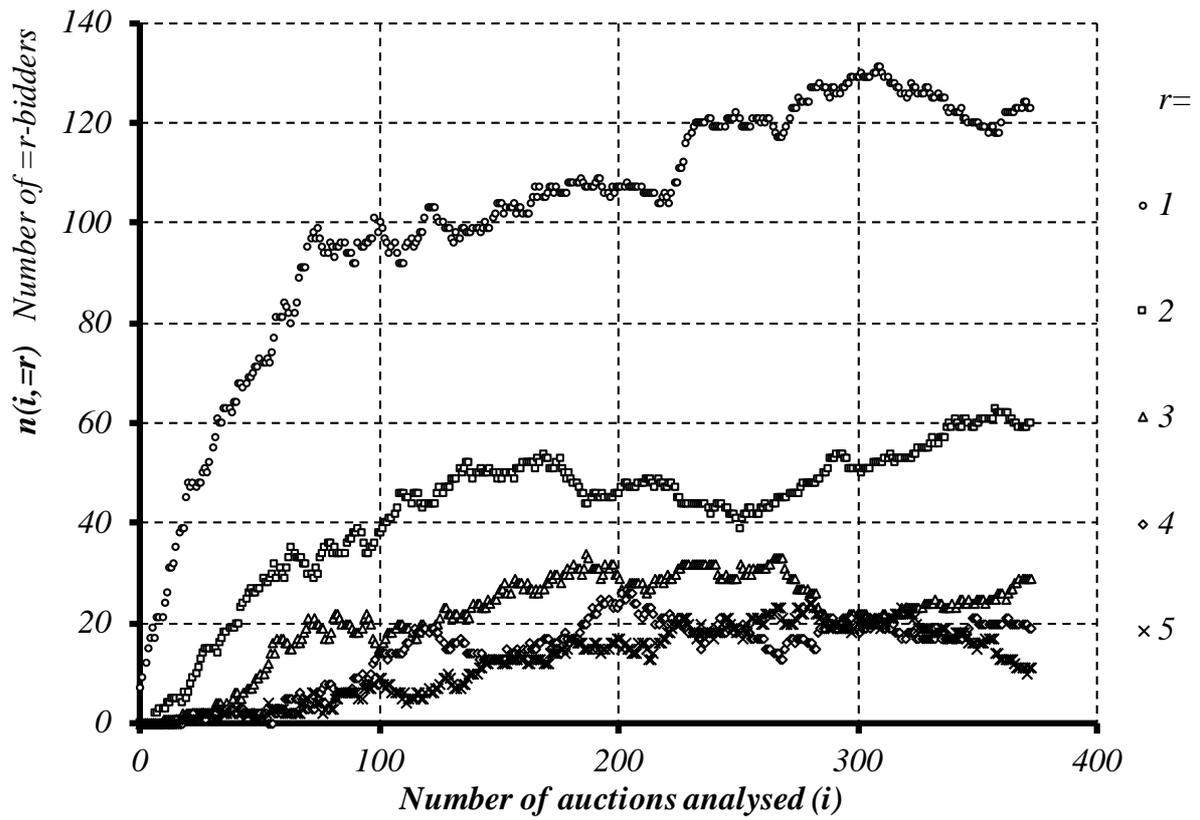
Figure 1: Number of bidders who had submitted "exactly" *r* bids (top) or "at least" *r*

bids (bottom).
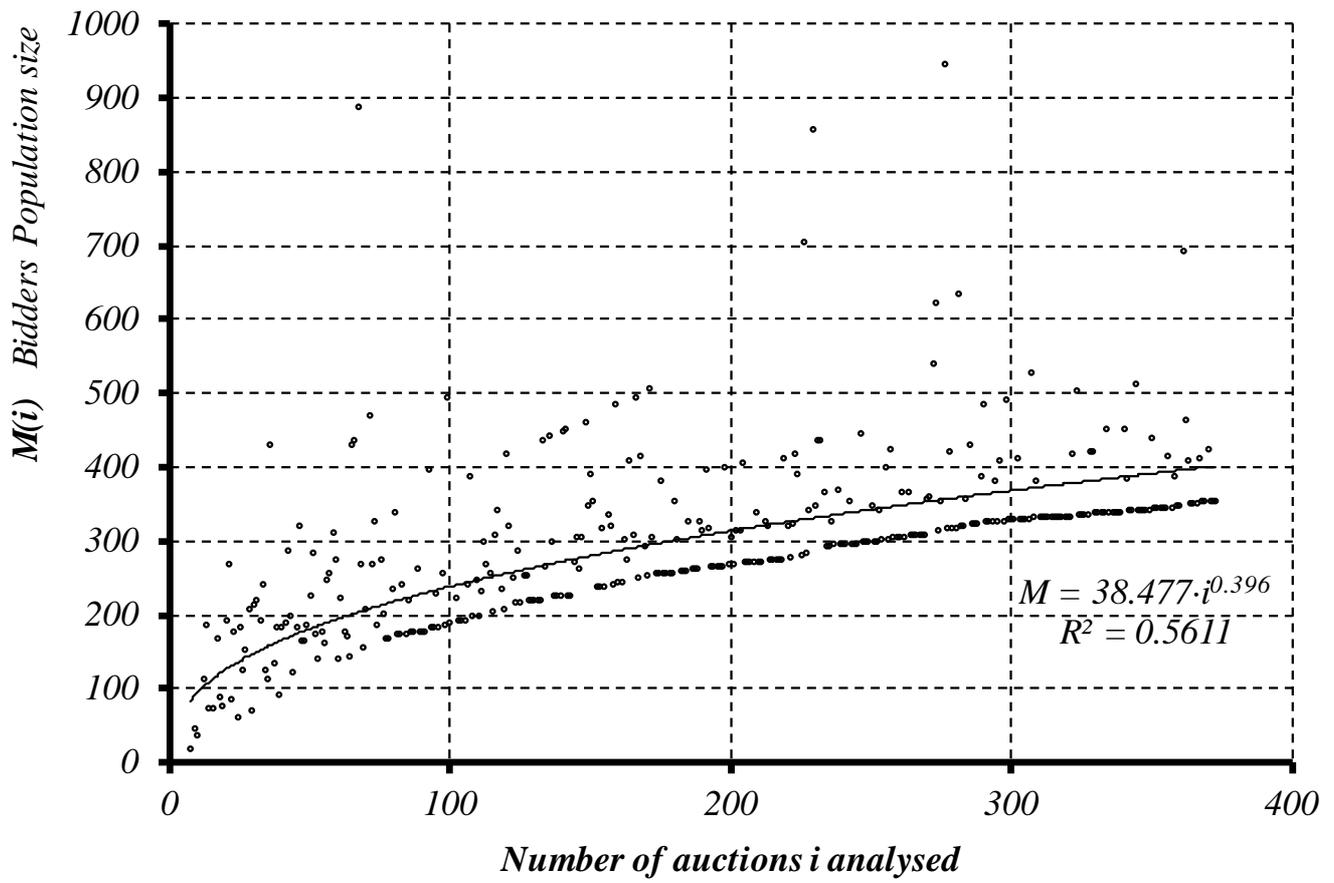
Figure 2: Potential regression depicting the best approximation for the bidders

population size *M(i)*.