Quick abnormal-bid-detection method for construction contract auctions

Article

Published Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

www.reading.ac.uk/centaur

CentAUR

Quick Abnormal-Bid-Detection Method
for Construction Contract Auctions

Pablo Ballesteros-Pérez[1]; Martin Skitmore[2]; Raj Das[3];
and Maria Luisa del Campo-Hitschfeld[4]

Abstract: Noncompetitive bids have recently become a major concern in both public and private sector construction contract auctions. Consequently, several models have been developed to help identify bidders potentially involved in collusive practices. However, most of these models require complex calculations and extensive information that is difficult to obtain. The aim of this paper is to utilize recent developments for detecting abnormal bids in capped auctions (auctions with an upper bid limit set by the auctioner) and extend them to the more conventional uncapped auctions (where no such limits are set). To accomplish this, a new method is developed for estimating the values of bid distribution supports by using the solution to what has become known as the German Tank problem. The model is then demonstrated and tested on a sample of real construction bid data, and shown to detect cover bids with high accuracy. This paper contributes to an improved understanding of abnormal bid behavior as an aid to detecting and monitoring potential collusive bid practices. **DOI: [10.1061/(ASCE)CO.1943-7862.0000978](#).** *This work is made available under the terms of the Creative Commons Attribution 4.0 International license, http://creativecommons.org/licenses/by/4.0/.*

Author keywords: Bidding; Abnormal bid; Bid covering; Collusion; German Tank problem; Construction auctions; Contracting.

Introduction

In the bidding context, collusion (or bid rigging as it is sometimes known) occurs when businesses that would otherwise be expected to be genuinely competing for work secretly conspire to raise prices or sometimes to lower the quality of goods or services for purchasers in a bid process (OECD 2009). Collusive bids can be particularly damaging in public procurement, since in Organization for Economic Cooperation and Development (OCED) countries, for example, public procurement represents about 15% of gross domestic product (GDP; OECD 2007), a value even higher in other countries (Aoyagi and Fréchette 2009). Collusive practices also absorb resources from procurers and taxpayers, since this action usually undermines the advantages of a competitive market and diminishes public confidence in the competitive process (Marshall and Marx 2009; Anderson and Cau 2011). Moreover, these are illegal in many countries, involving considerable resources dedicated to prosecuting those companies involved (Bajari and Summers 2002; Hendricks et al. 2008).

[1]Assistant Professor, Departamento de Ingeniería y Gestión de la Construcción, Faculty of Engineering, Universidad de Talca, Camino los Niches, km 1, Curicó, Chile (corresponding author). E-mail: pballesteros@utalca.cl

[2]Professor, School of Civil Engineering and the Built Environment, Queensland Univ. of Technology, Gardens Point, Brisbane, Australia. E-mail: rm.skitmore@qut.edu.au

[3]Senior Lecturer, Dept. of Mechanical Engineering, Univ. of Auckland, 20 Symonds St., Auckland 1010, New Zealand. E-mail: r.das@auckland.ac.nz

[4]Assistant Professor, Centro de Sistemas de Ingeniería, KIPUS, Faculty of Engineering, Universidad de Talca, Camino los Niches, km 1., Curicó, Chile. E-mail: mdelcampo@kipus.cl, mdelcampo@utalca.cl

So-called bid-covering or cover bidding is the most recurrent form of collusive arrangement in sealed bid auctions (Ishii 2008). It occurs when individuals or firms agree to submit bids in which either (1) a competitor agrees to submit a bid that is higher than the bid of the designated winner, (2) a competitor submits a bid that is known to be insufficiently competitive to be accepted from the technical standpoint, or (3) a competitor submits a bid that contains special contractual terms that are known to be unacceptable to the auctioneer (OECD 2007). Cover bids are also used in the well-known collusive arrangement of bid rotation (Porter and Zona 1993; Ishii 2009), which involves participating firms continuing to bid while taking turns to be the winning bidder.

Thus, classifying bids as abnormal and combating collusion are primary concerns for auctioneers as those bidders who manage to form a viable cartel or bidding ring (i.e., a group of companies planning to restrict the amount of actual competition among the participants in one or several auctions) can seriously affect winning bid values (Blume and Heidhues 2008; Hu et al. 2011). As a result, Klemperer (2002) and Anderson et al. (2012) regard that collusion, as well as other competition policy issues, is being more important in the design of practical auctions than the budget-constraint, affiliation, and risk-aversion issues that are often addressed in the mainstream theory of auctions.

The literature proposes some forms of auction rules to discourage collusion, such as establishing a reserve price that is a function of cartel size (Graham and Marshall 1987), selecting efficient auction mechanisms depending on the amount of correlation between colluders (Laffont and Martimort 2000), exploiting informational asymmetries concerning potential colluders and including a nontrivial probability of not selling the object auctioned (Che and Kim 2006, 2008), and including both effective ceiling and reserve prices (Chowdhury 2008). However, collusion schemes are always difficult to detect as they are typically negotiated in a strictly secretive way and are not usually evident from the results of a single auction, as cover bids have to give the appearance of genuinely competitive bids (Bajari and Summers 2002). An additional problem is that an

effective strategy for avoiding collusion usually requires the auctioneer to be able to predict the distribution function from which bidders are assumed to draw their values and be aware of which bidders belong to which cartel (Bajari and Summers 2002), while obtaining such information is intricate, if not impossible, in practice (Hu et al. 2011).

A collusive arrangement is often revealed only upon the appearance of a steady pattern of distrustful or abnormal behavior from several bidders over a period of time (OECD 2009). However, the existence of such patterns does not necessarily act as evidence of collusion, as there may simply be decreasing returns to scale of bidders' cost functions or a change in market conditions. For example, lower marginal costs occur with firms with idle capacity (low current workload) and hence, when reflected in the bid, are relatively more likely to win the auction (Porter and Zona 1993; Porter 2005). Despite this cautionary note, sets of tools that help identify particular bid behaviors as being collusive, or at least abnormal, can be of use to auctioneers (Rasch and Wambach 2009).

In this paper, we aim at identifying abnormal bids in either public or private sealed-bid auctions, as well as in auctions in which other nonprice criteria in addition to the economic bid value may be involved. These nonprice criteria are increasingly common in both public sector procurement auctions (Perng et al. 2006; Tan et al. 2006; Bergman and Lundberg 2013) and private sector auctions (Bajari and Summers 2002; Gayle and Richard 2008).

The paper makes three major contributions, as follows: (1) it extends a recent model developed by Ballesteros–Pérez et al. (2013b) for detecting abnormal bids in capped auctions (where an upper bid limit is set by the auctioneer) to the more conventional uncapped auctions (where no such limits are set), (2) it presents a new method for estimating the bid distribution supports or bounds of bids by an approach associated with the solution to the German Tank problem (a well-known statistical case study that shares several characteristics with bidding), and (3) a new abnormal-probability metric named P_{abn} for highlighting abnormal bids is proposed for use in combination with the extended Ballesteros–Pérez et al. (2013b) model. The three components of the model are applied and evaluated to a sample of real construction bid data, and shown to detect cover bids with a high level of accuracy, despite being conceived as a rough detection tool.

Background

A large body of economic theory demonstrates that the existence of both competitive and collusive bid strategies depends very much on the cost structure of the bidders (Curtis and Maines 1973; Maskin and Riley 2000), and rules of the auction (Porter and Zona 1993, 1999; Baldwin et al. 1997; Pesendorfer 2000). Theories of collusion in auctions also highlight the importance of preauction meetings among bidders, in which incentives or compensations are generally provided by the winner to the losers. The McAfee and McMillan (1992) static scheme characterizes efficient collusion when no side transfer is possible, and in which the designated winner is independent of history. Subsequently, this analysis was extended by Aoyagi (2003) and Skrzypacz and Hopenhayn (2004) to a repeated framework in which, as opposed to the McAfee and McMillan (1992) static bid rotation (Porter and Zona 1999), bid coordination is based on past history within a dynamic bid rotation scheme.

In contrast to the theoretical literature, although there has been a great deal of empirical work aimed at detecting collusion in procurement auctions (J. E. Harrington, "Detecting cartels," Working Paper, Johns Hopkins University, Baltimore; Paha 2011), little

attention has been paid in the literature relating to the inner working of the bidding rings or collusive bid groups (McAfee and McMillan 1992; Hendricks et al. 2008). Related work includes the Porter and Zona (1993, 1999) modeling of the probability of a bidder winning by assuming a bid function linear in observable cost factors. Subsequent work (P. Bajari and L. Ye, "Competition versus collusion in procurement auctions: Identification and testing," Working Paper, Stanford University, Stanford, California, 2003) consistently observes the violation of the so-called conditional independence and exchangeability conditions, which must always be satisfied by a competitive bid strategy.

Finally, Ballesteros–Pérez et al (2012a, 2013a, 2014) introduce a bid tender forecasting model that is partially reconfigured to detect extreme abnormal bidders in capped auctions (Ballesteros–Pérez et al. 2013b). This method, which the writers term the Ballesteros–González–Cañavate method, is aimed at identifying bidders whose behavior is not conditionally independent and exchangeable, that is, not in accordance with a regular or predictable pattern. In short, this approximate but quick method assumes that individual bids are in accordance with a uniform distribution in the absence of some kind of abnormal behavior among the bidders involved.

In this connection, multiple statistical distributions have been used to analyze bid patterns, the main ones of which in the context of construction contract auctions are the uniform, normal, lognormal, gamma, and Weibull densities (Skitmore 2014).

Therefore, in the absence of any generally agreed distribution, the uniform distribution continues to be used in this paper for three reasons, as follows: (1) several previous researchers consider it to be accurate enough to depict construction bid data, (2) the method is intended to be sufficiently robust for the uniform distribution to generate reasonably approximate results, and (3) any other statistical distribution can be rescaled into a uniform distribution $U(0, 1)$ by using its cumulative distribution probability values if necessary.

However, a problem concerning finite distributions such as the uniform density is to estimate the value of the supports (upper and lower bounds) involved, as these are different for each auction and each auction happens only once. Of the several estimators available for this, one known as the solution to the German Tank problem provides a simple yet relatively accurate method and is presented in more detail in the next section.

In sum, noncompetitive bids have become a major concern in both public and private procurement auctions, and several models have been developed to help identify collusive bidders. However, most of these models require complex calculations and extensive information that is difficult to obtain in real-life situations. Therefore, the implementation of other simpler but less accurate models similar to the Ballesteros–González–Cañavate model should help in highlighting noncompetitive bid behaviors in the large amount of auctions that are handled daily by contracting authorities all around the world.

Thus, the research objectives of this paper are three-fold, as follows: (1) to extend the Ballesteros–González–Cañavate capped auction model to the uncapped auction, since this is the more widespread procurement approach in many countries, the United States included; (2) since the extension of the Ballesteros–González–Cañavate model will require working with an underlying bid distribution then, assuming that this distribution is well-represented by the Uniform density, a method will also be proposed to estimate the value of the supports involved; and (3) a new metric named P_{abn} to identify abnormal bids will also be introduced. This metric will help in focusing the extended Ballesteros–González–Cañavate model to those bids or combinations of bids with higher values of P_{abn}.

Methodology

Analysis of Individual Bid Distributions

This subsection deals with extending the Ballesteros–González–Cañavate capped auction model for use in uncapped auctions. Therefore, here a "bidder's bid" B_i will be the monetary bid made by a given bidder i in an uncapped auction, where $0 < B_i < +\infty$.

First and similarly to the Ballesteros–González–Cañavate model, the method developed needs to quantify how many bidders are involved in the auction to study the bidders' relative economic bid distances from each other (in other words, the expected average bid gap between two consecutive bids) so as to be contrasted with a standard pattern distribution. Aiming to do this, there is an extensive literature focusing on predicting the potential number of bidders in auctions (e.g., Ngai et al. 2002; Carr 2005). However, when the request for proposals deadline is reached, the number of participating bidders is disclosed ex post and then the standard pattern distribution to which the bids will have to be compared can be defined.

Likewise, assuming bids are randomly and uniformly distributed, then the expected difference in probability and value between the ith and $i+1$th of the N ranked bids in an auction is a constant. Therefore, the probability of surpassing the nth bid ($P_{n\text{th}}$) can be easily expressed in terms of the subsequent straightforward linear expression

$$P_{n\text{th}} = 1 - \frac{N - i + 0.5}{N} = \frac{i - 0.5}{N} \tag{1}$$

where $i = 1$ is the most economical bid (lowest bid); and $i = N$ is the most expensive bid (highest bid). Therefore, the variable $P_{n\text{th}}$ also represents the bidder's nth position performance by means of a coefficient that ranges from $1/2N$ to $1 - (1/2N)$, the distance between bid i and bid $i+1$ being always the value $1/N$ (see the y values in Fig. 1).
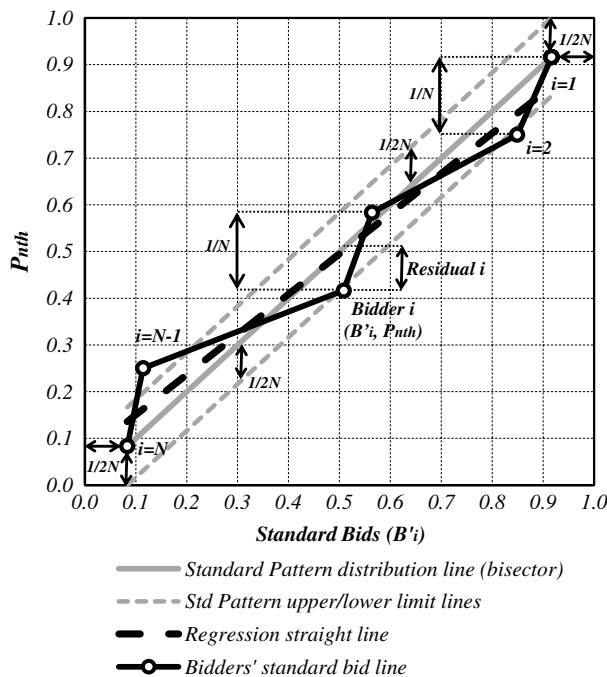

Fig. 1. Elements of a standard bid graph (modified from Ballesteros-Pérez et al. 2013b)

The next step is to correlate every bid (B_i) with its respective probability of being surpassed, $P_{n\text{th}}$. Aiming to simplifying future calculations, using the same interval of variation as $P_{n\text{th}}$ is again preferred, as an alternative to the range [$B_{\min}$ (lowest bid, most economical), $B_{\max}$ (highest bid, most expensive)], since the true bid distribution supports are not known. Therefore, the mathematical expression for rescaling the bids from their natural range [$B_{\min}$, $B_{\max}$], in monetary-unit basis, to the range [$1/2N, 1 - (1/2N)$], in per-unit basis, which will allow calculating what is named standard bids, B_i', is as per

$$B_i' = \frac{1}{2N} + \frac{N-1}{N} \cdot \frac{B_i - B_{\min}}{B_{\max} - B_{\min}} \tag{2}$$

In Fig. 2, Eq. (2) assigns x-axis values ranging from $1/2N$, if $B_i = B_{\min}$, to $1 - (1/2N)$, if $B_i = B_{\max}$, but keeps intact the original relative distances between bids on the x-axis. That is, unlike y-axis $P_{n\text{th}}$ values, the distance between bidder i and bidder $i+1$ will not usually be $1/N$, but proportional to the original relative distance when previously expressed in monetary bid values.

Therefore, beginning with a group of bids which took part in an uncapped auction and whose values have been previously ordered from lowest to highest (B_i', i), a new set of (B_i', $P_{n\text{th}}$) values can be obtained by using Eqs. (1) and (2). If these latter points fall approximately on a straight line, from ($1/2N, 1/2N$) to [$1 - (1/2N)$, $1 - (1/2N)$], this indicates that the bids can be treated as perfectly in accordance with a uniform distribution, and when other complementary conditions are also fulfilled, no abnormal bids should be present.

Hence, after every participating bid has been ordered and converted into a standard bid value (B_i') and its respective $P_{n\text{th}}$ is also calculated, it is necessary to compare this set of standardized bid values to the standard pattern distribution (SPD) whose mathematical expression is just a straight line

$$Y_{\text{pattern}} = B_i' \tag{3}$$

In short, Eq. (3) means a cumulative distribution function whose representation is a bisector line no matter the number of bids, and whose valid range of values will be from $1/2N$ to $1 - (1/2N)$ in both horizontal and vertical axes. This standard line suggests that any two adjacently ranked bids will be placed on average a $1/N$ value from each other, both in their B_i' values and $P_{n\text{th}}$ values.

Nevertheless, as mentioned previously, perfect matching between the SPD and each group of (B_i', $P_{n\text{th}}$) points is difficult to achieve; thus, to delimit a band in which the recently calculated set of (B_i', $P_{n\text{th}}$) points can be classified as close enough to the SPD, a new couple of boundary lines, named standard pattern upper and lower limit lines, have to be defined. The mathematical expressions of these boundary lines are these, and Fig. 1 shows that they are located at a $1/2N$ distance just over and below the SPD

$$Y_{\text{lower}} = B_i' - \frac{1}{2N} \tag{4}$$

$$Y_{\text{upper}} = B_i' + \frac{1}{2N} \tag{5}$$

There is therefore a (B_i', $P_{n\text{th}}$) set of points that lie on a line with $N - 1$ segments and, since this composite line could be partially inside and partially outside the band defined by Eqs. (4) and (5), it is more appropriate to represent the group of (B_i', $P_{n\text{th}}$) values by its regression straight line. This way, whenever the regression line is completely within the boundaries defined by the lower and upper
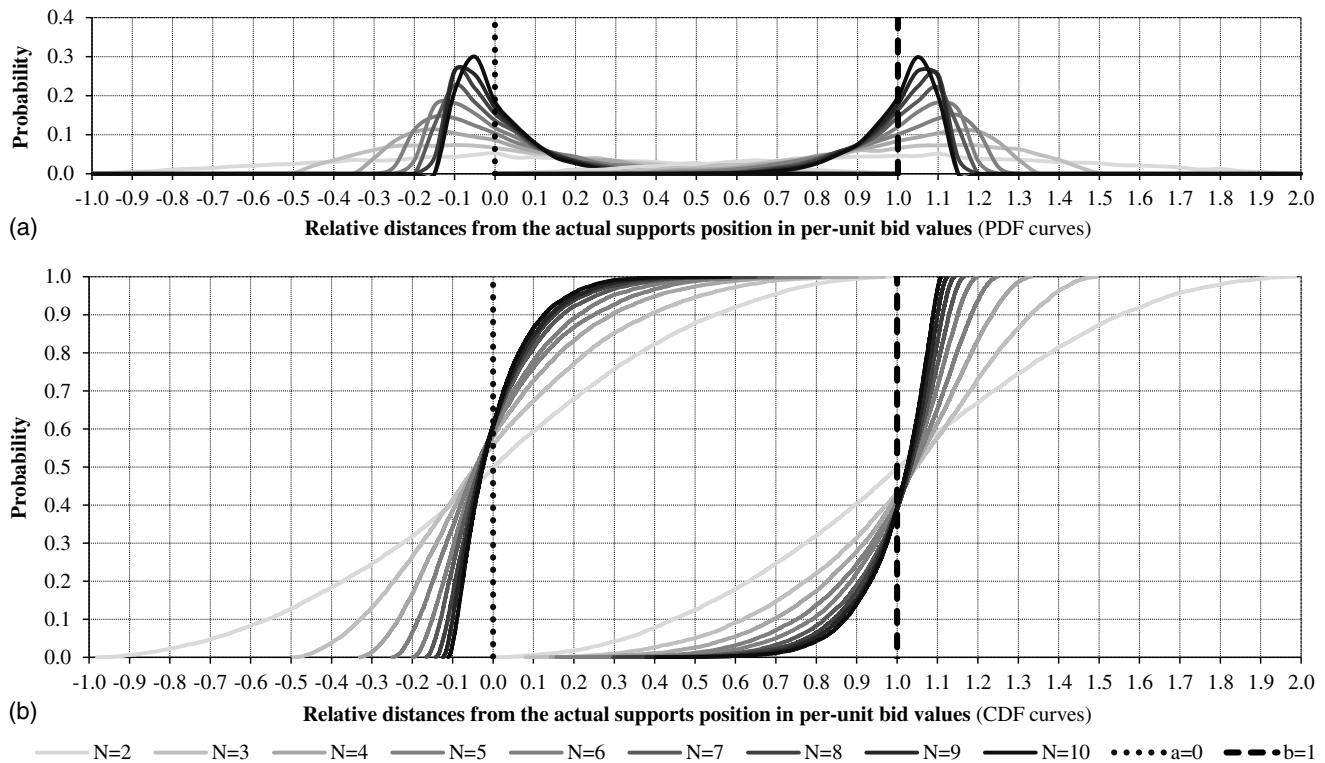
Fig. 2. Curves of the bid distribution support positions a and b: (a) probability density function curve; (b) cumulative distribution function curve

limit lines, it will be possible to assume that the bid distribution is actually close enough to the SPD. The band limit width $1/N$ (from $-1/2N$ to $+1/2N$) coincides with the distorting effect equivalent to one nonexistent bidder that one or several collusive bidders may generate over the bids distribution. The upper and lower limit lines then define a band in which the B_i' regression line (least-squares) should be squeezed in as long as the exchangeability condition is granted. Notwithstanding, there is another condition to fulfill; the coefficient of determination, R^2, of this least-squares line must be close enough to 1, in order to claim that the standard bids set is well-represented by its regression curve. Ballesteros–Pérez et al. (2013b) suggest that R^2 should be above 0.90 (a 0.10 distance from 1.0), but it seems to be more appropriate to set its level according to the number of bids actually analyzed (in this case, a $1/N$ value from 1.0). Therefore, the R^2 will be required to be above $R_{\min}^2 = (N-1)/N$, since, by fulfilling this condition, the least-squares regression line should explain the equivalent percentage of variability, which means that the actual $P_{n\text{th}}$ values should not be separated by more than a $1/N$ value from the regression line y values due to uncontrolled data variability.

In addition, the condition of conditional independence is also to be granted whenever bids represent genuine competition. This condition can be broken down into two more verifications, as follows: (1) the residuals (the difference between B_i''s $P_{n\text{th}}$ and Y_{pattern} values as in Fig. 1) have to be in accordance with a normal distribution, and hence a Student's t-test should be carried out on the residuals dataset; and (2) the mean of the B_i' standard bids, B_m', should be nearly 0.5.

Condition 2, when embedded in Eq. (2), is equivalent to $B_m' = P_{n\text{th}=N/2} = 0.5$ which means that whenever there are no collusive or abnormal bids, the bid distributions are symmetrical around their B_m' value, as also originally stated in Ballesteros–Pérez et al. (2013b). Hence, a new coefficient that monitors B_m' deviations has been created. This coefficient, named $B_{m\,\text{distortion}}'$, is able to

measure the distance of B_m' from 0.5 in multiples of $1/N$. Therefore, the mathematical expression of the standard mean bid distortion is

$$
\begin{aligned}
B_{m\,\text{distortion}}' &= \frac{|B_m' - 0.5|}{0.5} \bigg/ \frac{1}{N} = N\frac{|B_m' - 0.5|}{0.5} = N|2B_m' - 1| \\
&= N\left|2\frac{B_m - B_{\min}}{B_{\max} - B_{\min}} - 1\right|
\end{aligned}
\tag{6}
$$

Measuring this B_m' deviation as Eq. (6) proposes also has the advantage of revealing how many bidder positions the B_m' value has been dislodged in $1/N$ multiples. Values lower than 1.0 are required to grant the conditional independence condition; that is, to ensure there is not a subset of bids that are in accordance with a different distribution and quite probably have a different mean value.

To sum up, to guarantee conditional independence as well as exchangeability, whenever a set of auction bids is analyzed, the following four mathematical conditions must be satisfied:
1. The $(B_i', P_{n\text{th}})$ least-squares line must be completely inside the zone bounded by the lines defined by Eqs. (4) and (5);
2. The regression straight line's coefficient of determination must be above $R_{\min}^2$, i.e., actual $R^2 > R_{\min}^2 = (N-1)/N$;
3. The differences between B_i''s $P_{n\text{th}}$, and Y_{pattern} values (residuals) must be in accordance with a normal distribution, which means checking the condition $t_{\text{student}(\alpha=5\%)} < t_{\text{student}\,B_i'}$; and
4. The mean standard bid B_m' must have been displaced less than a $1/N$ value from 0.5, which is equivalent to $B_{m\,\text{distortion}}' < 1$.

These four conditions are already included in the original Ballesteros–González–Cañavate model, but here they have been refined with two partially reformulated conditions in order to increase their effectiveness in uncapped auctions. Therefore, any group of bidders that do not comply with any of these four conditions means that at least one bidder participated with an abnormal

bid generating an effect that adds up to $1/N$. Nonetheless, in practice, the bidders should always be allowed to justify their respective economic bids (since they may possible have been unintentionally estimated incorrectly). However, when a bidder exhibits a steady abnormal behavior detected in this way, it would qualify as potentially collusive and worthy of further investigation. The method could therefore be used for every auction as soon as the bids are disclosed, since it would act as a quick, but preliminary, collusion-detection mechanism.

Estimation of Distribution Supports

While the previous subsection extended the original Ballesteros–González–Cañavate model for abnormal bid detection to uncapped auctions, this section refines this further by including a second contribution of this paper in estimating the distribution supports.

Estimating the values of the supports of finite bid distributions requires simple but not obvious reasoning that has not been previously applied to auctions to date. Assuming that bids are in accordance with a continuous uniform distribution, $U(a, b)$, it is expected that the true lower boundary, a, will always be a little below the observed $B_{\min}$, whereas the true upper limit, b, will always be little above the observed $B_{\max}$. The distance between the observed boundary values ($B_{\min}$ and $B_{\max}$) and the true values (a and b) can then be calculated by using the German Tank problem solution.

In the statistical theory of estimation, the problem of estimating the maximum of a discrete uniform distribution from sampling without replacement is known as the German Tank problem due to its application in World War 2 to the estimation of the number of German tanks (Goodman 1954). During the course of the war, the Western Allies made sustained efforts to determine the extent of production of German Panther tanks. To do this they made use of the gearbox serial numbers printed on captured or destroyed German tanks (Ruggles and Brodie 1947). This provided the solution to the problem that can be understood intuitively as "the population maximum equals the sample maximum plus the average gap between observations in the sample," the gap between the ranked observations being added to compensate for the negative bias of the sample maximum as an estimator for the population maximum.

Although the individual bid distribution is effectively continuous instead of the discrete German Tank problem version, the same principle can be used to obtain the bounds a and b by translating the German Tank problem into auction bids, as follows:
- The maximum of a uniform distribution (b) equals the sample maximum ($B_{\max}$) plus the average gap (there are $N - 1$ gaps between $B_{\max}$ and $B_{\min}$) between the ranked observations in the sample, and
- the minimum of a uniform distribution (a) equals the sample minimum ($B_{\min}$) minus the average gap ($N - 1$ again) between observations in the sample.

Therefore, a and b are estimated as

$$a = B_{\min} - \frac{B_{\max} - B_{\min}}{N - 1} \tag{7}$$

$$b = B_{\max} + \frac{B_{\max} - B_{\min}}{N - 1} \tag{8}$$

Additionally, two major parameters allow the analysis of future auction bid distributions, as follows: (1) the mean, $(a + b)/2$, and (2) the standard deviation (SD), $(b - a)/\sqrt{12}$, which can be immediately calculated by Eqs. (7) and (8).

However, in order to prove that the boundary estimates of a and b are accurate, it is necessary to check if the mean and the SD really fit the actual auction bid data. In platykurtic distributions, such as the uniform distribution, despite both the sample mean and the sample median are unbiased estimators of the midpoint, neither is as efficient as the sample midrange, i.e., the arithmetic mean of the sample maximum and the sample minimum (which is also the maximum likelihood estimate). Therefore, on this occasion, the mean value is not useful, since by definition $(B_{\max} + B_{\min})/2 \approx (a + b)/2$. Then, if

$$b - a = B_{\max} + \frac{B_{\max} - B_{\min}}{N - 1} - B_{\min} + \frac{B_{\max} - B_{\min}}{N - 1}$$
$$= \left(1 + \frac{2}{N - 1}\right)(B_{\max} - B_{\min}) = \frac{N + 1}{N - 1}(B_{\max} - B_{\min}) \tag{9}$$

the SD (σ) should be equal to

$$\sigma = \frac{b - a}{\sqrt{12}} = \frac{N + 1}{N - 1} \cdot \frac{B_{\max} - B_{\min}}{\sqrt{12}} \tag{10}$$

To date, the bid SD has been a difficult parameter to predict, with highest coefficients of determination around 0.7 (Ballesteros-Pérez et al. 2012a, b), mostly because no researcher has considered including the number of bids N involved in Eq. (10), since it was counterintuitive. Therefore, if Eqs. (7)–(9) constitute a reasonable approximation of the uniform bounds and dispersion parameters representing the bid distribution, the coefficient of determination in the actual bid dataset proposed should be noticeably above 0.7. This issue will be addressed in the Method Validation section.

Metric for Abnormal Bids

Specific bids can be tested for abnormality by checking Conditions 1–4 mentioned previously. However, doing this comprehensively would involve 2^N individual and group combinations of bidders, a value that becomes too high as the number of bids N increases. Hence, the third contribution of this paper is an alternative way for identifying bids that are more likely to be abnormally higher or lower than others, in addition to the lowest and highest ones. This is achieved by a simplified probabilistic analysis, unlike the original Ballesteros–González–Cañavate method, which just checks the four conditions for the lowest and highest bids. This involves considering how likely it is that one bid could have fallen outside the supports limits $[a, b]$. Nevertheless, since the true values of the supports are not known but estimated by Eqs. (7) and (8), it is necessary to obtain the general statistical distributions of a and b, which at the same time are expressed as a function of the variables N, $B_{\max}$, and $B_{\min}$.

The number of bidders, N, is known. However, the lowest ($B_{\min}$) and the highest ($B_{\max}$) bids have to be expressed in terms of the first-order and last-order statistics, respectively. In extreme-value theory, these order statistics are in accordance with a beta distribution for the uniform distribution, as per David and Nagaraja (2003)

$$B_{\min} = \text{beta}(B_i, \alpha = 1, \beta = N) \tag{11}$$

$$B_{\max} = \text{beta}(B_i, \alpha = N, \beta = 1) \tag{12}$$

Introducing Eqs. (11) and (12), into Eqs. (7) and (8), results in the curves represented in Fig. 2, which were obtained by simulation.

These curves (Fig. 2) take on values within the range $[\breve{a}, \hat{a}] = [1/(1 - N), 1]$ in the case of the lower support a and within $[\breve{b}, \hat{b}] = [0, N/(N - 1)]$ in the case of the upper support b, but if they are
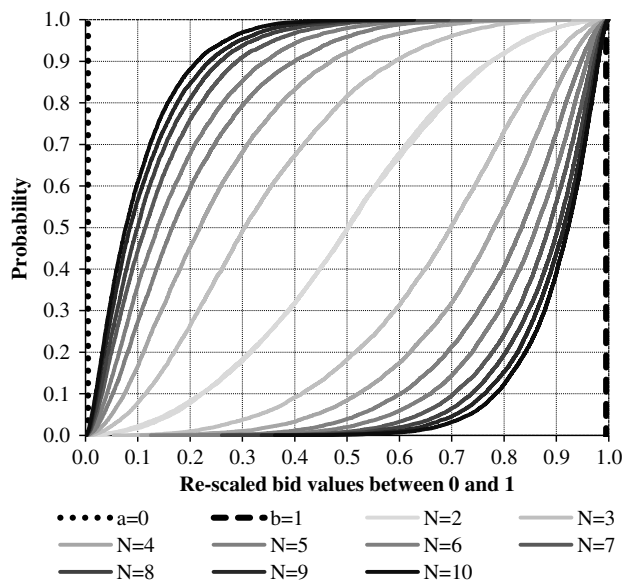
Fig. 3. Rescaled probabilistic curves of the support positions

rescaled within the intervals $[\breve{a}, \hat{a}] = [0, 1]$ and $[\breve{b}, \hat{b}] = [0, 1]$, respectively, the result is as shown in Fig. 3.

This series of curves with the same domain, even though not corresponding to beta distributions, are very accurately approximated by this type of curve as

$$a \approx \mathrm{beta}(B_i, \alpha = A, \beta = B) \qquad (13)$$

$$b \approx \mathrm{beta}(B_i, \alpha = B, \beta = A) \qquad (14)$$

where parameters A and B can be estimated by the method of moments as

$$A = \mu \left[\frac{\mu(1-\mu)}{\upsilon} - 1 \right] \qquad (15)$$

$$B = (1-\mu) \left[\frac{\mu(1-\mu)}{\upsilon} - 1 \right] \qquad (16)$$

where

$$\mu = \frac{1}{N} \qquad (17)$$

$$\upsilon = \frac{N-1}{N(N+1)(N+2)} \qquad (18)$$

where the mean μ is exact; and the variance υ of these distributions was obtained by the simulation, and always results in p-values below 1%, when approximating Eqs. (7) and (8) by Eqs. (13) and (14), respectively.

Now, with the supports a and b distributions closely approximated, the next step is to compare the actual bid values B_i with the probability curves of the supports and obtain the probabilities of falling outside the range $[a, b]$. In order to check this last condition, two probability-related values are defined, as follows: (1) $P_{\mathrm{abn\,low}}$, which is the lower bound of the probability that a B_i value is less than the lower support a; and (2) $P_{\mathrm{abn\,high}}$, which is the lower bound of the probability that a B_i value is more than the upper support b. By means of the cumulative distribution functions [Eqs. (13) and (14)], these probabilities are easily calculated for every B_i as

$$P_{\mathrm{abn\,low}} = 1 - \mathrm{beta}(B_i, \alpha = A, \beta = B) \qquad (19)$$

$$P_{\mathrm{abn\,high}} = \mathrm{beta}(B_i, \alpha = B, \beta = A) \qquad (20)$$

However, these probabilities are only lower bounds, i.e., they will always underestimate the probability since the true values of the supports are only approximated. However, they serve their purpose since high values of these coefficients always helps shed light on those bids worth checking in more detail among the number of 2^N possible combinations of bid groups.

The following two probabilities result: (1) the closer are these probabilities to unity, the higher is the probability that if that bid was removed from the original set of bid values, Conditions 1–4 will be fulfilled; and (2) when the probability is closer to zero, the more unlikely a bid will be capable of complying with the four conditions and therefore the less likely it will qualify as abnormally low or high.

Nevertheless, to provide an unequivocal interpretation, these two probabilities can be merged into one that represents in a single value the lower bound of the probability that every bid falls outside the range $[a, b]$. This probability is

$$P_{\mathrm{abn}} = P_{\mathrm{abn\,low}} + P_{\mathrm{abn\,high}} \qquad (21)$$

The interpretation of Eq. (21) is analogous to the interpretations given for probabilities $P_{\mathrm{abn\,low}}$ and $P_{\mathrm{abn\,high}}$.

However, lower P_{abn} values do not necessarily indicate that a bid would be abnormal if it is located well within the bid distribution instead of near the extremes, this being a task for the Ballesteros–González–Cañavate model presented previously.

Fig. 4 shows the calculations involved for Auction No. 33 taken from the construction bid dataset introduced in the next section. Additionally, the original bid values B_i also have to be rescaled between $[a, b]$ using a per-unit scale (as the third row in the top left of the Fig. 4 shows) to allow their comparisons with Eqs. (13)

| Auction Entry ID = | 33 | | Bidder 304's position (i) = | | ↓ 6 ↓ |
| | | | | | (Cover) |

Bidder (ranked) i	1	2	3	4	5	6
Bids (monetary) Bᵢ	792,966	811,788	819,971	847,621	847,892	853,793
Bᵢ (re-scaled) (Bᵢ-a)/(b-a)	0.1429	0.3639	0.4600	0.7847	0.7878	0.8571

N	Bₘᵢₙ	Bₘₐₓ	a (eq. 7)	b (eq. 8)
6	792,966	853,793	780,801	865,958

a's and b's Beta distribution parameter calculations

μ (eq. 16)	υ (eq. 17)	A (eq. 14)	B (eq. 15)
0.1667	0.0149	1.3889	6.9444

Pₐᵦₙ low (eq. 18)	0.4899	0.0781	0.0268	0.0001	0.0000	0.0000
Pₐᵦₙ high (eq. 19)	0.0000	0.0019	0.0092	0.2915	0.2987	0.4899
Pₐᵦₙ (eq. 20)	**0.4899**	0.0800	0.0361	**0.2916**	**0.2988**	**0.4899**

Fig. 4. Example of Auction No. 33 complete preliminary analysis for detecting abnormal bids

and (14), as performed in the simulation curves, which are used afterwards in Eqs. (19) and (21), and which also range from 0 to 1.

The highest values of the lower bound of probabilities P_{abn} for this auction shows that the lowest bid ($i = 1$), as well as the three higher ones [(1) $i = 4$, (2) $i = 5$, and (3) $i = 6$] are those worth checking for compliance with the four conditions. However, as will be seen subsequently in Table 1, in this auction, only Bids 4–6 satisfy the four conditions when they are individually removed, so they are the ones that become classified as abnormal (abnormally high in this case).

The major feature of this third contribution of the paper, therefore, is that metric P_{abn} is extremely simple and quick to calculate, and is capable of detecting abnormal bids located nearer the extremes of the distribution. This feature, together with the ability to detect abnormal bids located near the bid distribution average by the extended Ballesteros–González–Cañavate model, constitutes a valuable improvement. Nonetheless, broadly speaking, any collusive bidder that aspires to effectively leverage the bid distribution will need to be abnormally high or abnormally low to exert sufficient influence, particularly when cartel bids do not comprise the larger part of the bids, a situation that increases the probability of being detected.

Validation

Method

As described, this paper presents three different but complementary contributions, but they have to be analyzed mostly as a group. The first contribution is the extension of the original Ballesteros–González–Cañavate capped auction model to uncapped auctions. Two of the original four conditions have been refined to increase its accuracy whereas the variables of the four conditions have been transformed for use with bids values, unlike the dimensionless bids in the original model. The manner in which this extended model will be tested requires applying the reformulated four conditions to an uncapped auction dataset, with some known cover bids of one bidder, and observe the ratio of their correct and incorrect detections as abnormal bids.

The second contribution is to propose a new estimator of the bid distribution supports. This involves using Eq. (10) to estimate the bid SD. If this provides a reasonably good approximation of the true bid SDs, then the true supports a and b will also be well-approximated since Eq. (10) is linearly proportional to the difference $b - a$, and therefore the coefficient of determination will be close to unity.

The third contribution constitutes a new metric named P_{abn} to help focus the Ballesteros–González–Cañavate model on bidders located near the extremes being more likely to be abnormally high or low (remember that P_{abn} is almost useless for abnormal bids not located near B_{min} or B_{max}). In this case, metric P_{abn} will also be tested against the bid dataset but only in combination with the refined Ballesteros–González–Cañavate model.

Auction Dataset

To evaluate the practical application of the method, the Skitmore and Pemberton (1994) set of uncapped auction bid data is analyzed. These were donated by a construction company (encoded as Bidder 304) operating in the London area of the United Kingdom (U.K.) and covered this company's building contract bid activities during a 12-month period in the early 1980s for a total of 86 contracts. The 51 resulting auctions for which a full set of bids were available

are given an auction identifier (ID) according to the original bid dataset numbering and are presented in Table 1 along with the number of participating bidders (Column 2) and Bidder 304's position. Column 6 indicates whether Bidder 304's bids are genuine or cover bids according to the information provided by the donating company. The is a series of auctions where the sole awarding criterion was the lowest bid, with no abnormally low bid criterion applied preset in the auction specifications by the auctioner and with no knowledge of whether other bidders entered cover bids. This dataset therefore constitutes a very robust test to the method due to the following:

- The existence of an abnormally low bid criterion greatly conditions the way collusive bidders act (which makes it easier to discriminate between bids that are only abnormal, and those that are both abnormal and also potentially collusive), and
- Other bids might also be cover bids and therefore generate extra noise in the data.

Calculations and Validating Results

Table 1 reflects the application of the method in 50 out of the 51 auctions (Auction No. 16 was not taken into account because, as the original Ballesteros–González–Cañavate model, the minimum number of bidders to be analyzed must be greater than 3). The second block in Table 1 (Columns 4–7) contains the method's predictions concerning whether Bidder 304's bid was genuine or cover. In short, Conditions 1–4 were checked when Bidder 304 (alone or in combination with other bidders) was removed. The resulting percentage of correct/incorrect predictions is noticeably high (86% versus 14%), especially taking into account the absence of information concerning any other cover bidders involved. Furthermore, the model also detected two out of three abnormal bids not located near the extremes (Auction Nos. 9, 15, and 40); this fact is mentioned since detection is very difficult in this situation.

Bidder 304's abnormal bids, being cover bids, are always high, which leaves the ability of the method to detect abnormally low bids untested. However, as deduced from the model's four conditions, the approach to detecting abnormally low bids is equivalent to that of abnormally high bids, suggesting that since abnormally high bids have been successfully detected, abnormally low bids should be equivalently so, since they are symmetrical cases.

Block 3 in Table 1 shows the SD values of the observed bids (Column 8, σ actual) as well as the estimated values obtained by applying Eq. (10) [Column 9, σ, Eq. (10)]. Eq. (10) was expressed as a function of N (Column 2) as well as B_{max} and B_{min} (not presented due to the lack of space). The coefficient of determination (R^2) is 0.988, very close to 1, leading to the conclusion that Eq. (10), as well as the support estimators proposed in Eqs. (7) and (8), constitute a very good approximation of the bid distribution boundaries.

The last block on the right-hand side of Table 1 (last 10 columns) presents the P_{abn} values for all bids in which the positions occupied by Bidder 304 have been underlined. The sequence of calculations performed to obtain these values was identical to the one described in Fig. 4 for Auction No. 33. That is, B_i, N, B_{max}, and B_{min} were used to obtain supports a and b from Eqs. (7) and (8). The original B_i values were then rescaled within the interval [a, b] and with these rescaled values (from 0–1), μ, υ, A, and B were easily obtained from Eqs. (15)–(18). Finally, $P_{abn\ low}$ and $P_{abn\ high}$ were calculated by Eqs. (19) and (20), and the final P_{abn} values obtained by Eq. (21).

Table 1. Uncapped Dataset Used for the Analysis and Prediction Results

Auction Identifier[a]	N	Position of Bidder 304	Results: Inspection of the four conditions	Prediction	Actual	Result	σ, actual	σ, Eq. (9) $R^2 = 0.988$	1	2	3	4	5	6	7	8	9	10
															P_{abn} values[b]			
1	6	5	Remove Bidder 6	Genuine	Genuine	Right	24,156	27,077	0.49	0.30	0.22	0.08	0.04	0.49	—	—	—	—
2	4	4	No abnormal bids detected	Genuine	Genuine	Right	14,951	16,149	0.56	0.17	0.33	0.56	—	—	—	—	—	—
3	7	6	Remove Bidder 7	Genuine	Genuine	Right	213,677	230,422	0.47	0.43	0.42	0.24	0.24	0.16	0.47	—	—	—
4	6	3	Remove Bidders 3 + 4, 3 + 5, or 3 + 6	Cover	Cover	Right	28,791	25,561	0.49	0.21	0.35	0.36	0.38	0.49	—	—	—	—
5	6	6	No abnormal bids detected	Genuine	Genuine	Right	14,118	13,378	0.49	0.45	0.04	0.10	0.23	0.49	—	—	—	—
6	9	4	Remove Bidder 9	Genuine	Genuine	Right	51,544	64,674	0.45	0.21	0.07	0.05	0.03	0.03	0.01	0.00	0.45	—
7	7	1	No abnormal bids detected	Genuine	Genuine	Right	95,999	104,020	0.47	0.12	0.07	0.04	0.05	0.09	0.47	—	—	—
8	4	2	No abnormal bids detected	Genuine	Genuine	Right	419,630	493,277	0.56	0.17	0.16	0.56	—	—	—	—	—	—
9	6	3	Remove Bidders 2, 3, 4, or 6	Cover	Cover	Right	44,384	50,503	0.49	0.14	0.12	0.06	0.05	—	—	—	—	—
10	4	1	No abnormal bids detected	Genuine	Genuine	Right	39,503	43,631	0.56	0.33	0.17	0.56	—	—	—	—	—	—
11	6	6	Remove Bidders 2 or 3	Genuine	Genuine	Right	76,901	76,027	0.49	0.25	0.19	0.04	0.21	0.49	—	—	—	—
12	6	1	Remove Bidders 2 or 3	Genuine	Genuine	Right	106,582	110,193	0.49	0.34	0.21	0.05	0.08	0.49	—	—	—	—
13	4	4	No abnormal bids detected	Genuine	Genuine	Right	33,978	34,884	0.56	0.46	0.17	0.56	—	—	—	—	—	—
14	6	5	No abnormal bids detected	Genuine	Cover	Wrong	45,120	43,098	0.49	0.41	0.03	0.04	0.33	0.49	—	—	—	—
15	6	3	Remove Bidders 2 + 3, 3 + 4, or 3 + 5	Cover	Cover	—	97,034	116,595	0.49	0.14	0.10	0.07	0.07	0.49	—	—	—	—
16	3	2	N is too low to analyze	—	Cover	—	21,150	23,843	0.64	0.43	0.64	—	—	—	—	—	—	—
17	10	10	Remove Bidders 2, 3, or 10	Cover	Cover	Right	20,643	22,589	0.44	0.14	0.10	0.02	0.01	0.00	0.00	0.04	0.04	0.44
18	6	5	No abnormal bids detected	Genuine	Cover	Wrong	18,441	18,751	0.49	0.18	0.08	0.13	0.16	0.49	—	—	—	—
19	9	4	No abnormal bids detected	Genuine	Genuine	Right	177,243	171,618	0.45	0.08	0.08	0.03	0.00	0.00	0.17	0.39	0.45	—
20	8	2	Bidders to be removed are greater than $N/2$	Genuine	Genuine	Right	287,432	306,506	0.46	0.44	0.36	0.24	0.05	0.02	0.01	0.46	—	—
21	7	2	Bidders to be removed are greater than $N/2$	Genuine	Genuine	Right	321,413	382,299	0.47	0.22	0.11	0.10	0.09	0.07	0.47	—	—	—
22	6	6	Remove Bidders 1, 4, or 5	Genuine	Genuine	Right	17,467	19,293	0.49	0.03	0.03	0.21	0.25	0.49	—	—	—	—
23	5	4	No abnormal bids detected	Genuine	Genuine	Right	35,096	38,563	0.52	0.29	0.09	0.08	0.52	—	—	—	—	—
24	8	4	Remove Bidders 1, 2, or 3	Genuine	Genuine	Right	43,277	39,389	0.46	0.37	0.22	0.04	0.02	0.04	0.30	0.46	—	—
25	6	5	Remove Bidders 2 or 5 + 6	Cover	Cover	Right	79,358	84,535	0.49	0.31	0.19	0.13	0.04	0.49	—	—	—	—
26	7	7	Remove Bidders 2 or 3	Genuine	Genuine	Right	27,035	27,386	0.47	0.25	0.10	0.09	0.02	0.17	0.47	—	—	—
27	4	4	Remove Bidders 2 or 3	Genuine	Genuine	Right	165,506	168,422	0.56	0.53	0.16	0.56	—	—	—	—	—	—
28	7	2	No abnormal bids detected	Genuine	Genuine	Right	75,423	86,869	0.47	0.05	0.02	0.02	0.04	0.17	0.47	—	—	—
29	6	5	Remove Bidders 1 + 2, 3 + 4, or 5 + 6	Cover	Cover	Right	44,337	48,165	0.49	0.03	0.14	0.23	0.32	0.49	—	—	—	—
30	6	1	Remove Bidders 4, 5, or 6	Genuine	Genuine	Right	90,354	94,814	0.49	0.04	0.04	0.24	0.38	0.49	—	—	—	—
31	6	5	Remove Bidders 1, 2, or 3	Genuine	Cover	Wrong	64,663	58,562	0.49	0.38	0.20	0.06	0.33	0.49	—	—	—	—
32	6	4	Remove Bidders 1 or 2	Genuine	Genuine	Right	31,029	28,234	0.49	0.45	0.27	0.09	0.26	0.49	—	—	—	—
33	6	6	Remove Bidders 4, 5, or 6	Cover	Cover	Right	24,473	24,583	0.49	0.08	0.04	0.29	0.30	0.49	—	—	—	—
34	7	7	Remove any bidder/s except Bidder 6	Genuine	Cover	Wrong	79,823	92,664	0.47	0.14	0.07	0.02	0.02	0.04	0.47	—	—	—
35	6	6	Remove Bidders 4 + 6	Genuine	Genuine	Right	39,184	46,132	0.49	0.16	0.13	0.04	0.04	0.49	—	—	—	—
36	6	4	No abnormal bids detected	Genuine	Genuine	Right	35,498	37,051	0.49	0.40	0.05	0.03	0.09	0.49	—	—	—	—
37	9	8	Remove Bidders 7 + 9	Genuine	Genuine	Right	292,439	367,558	0.45	0.11	0.01	0.01	0.00	0.00	0.01	0.01	0.45	—
38	5	5	No abnormal bids detected	Genuine	Genuine	Right	35,418	35,988	0.52	0.32	0.08	0.20	0.52	—	—	—	—	—
39	7	7	Remove Bidders 1, 2, or 3	Genuine	Cover	Wrong	193,372	178,076	0.47	0.33	0.15	0.04	0.03	0.42	0.47	—	—	—
40	8	4	Remove Bidders 1, 2, 3, or 8	Genuine	Cover	Wrong	106,729	115,820	0.46	0.32	0.04	0.02	0.01	0.01	0.08	0.46	—	—
41	6	3	Remove Bidders 2, 4, or 6	Genuine	Genuine	Right	69,814	81,337	0.49	0.20	0.06	0.05	0.04	0.49	—	—	—	—
42	8	3	Remove Bidder 8	Genuine	Genuine	Right	156,785	184,833	0.46	0.16	0.08	0.06	0.04	0.01	0.01	0.46	—	—
43	6	5	Remove Bidders 1, 4, or 5	Cover	Cover	Right	48,324	56,186	0.49	0.03	0.07	0.10	0.19	0.49	—	—	—	—
44	5	1	Remove Bidders 2, 3, or 4 + 5	Genuine	Genuine	Right	54,408	57,011	0.52	0.37	0.26	0.07	0.52	—	—	—	—	—
45	4	4	Remove Bidders 3 or 4	Cover	Cover	Right	42,668	44,297	0.56	0.24	0.50	0.56	—	—	—	—	—	—
46	7	5	Remover Bidders 2 or 6 + 7	Genuine	Genuine	Right	81,045	76,812	0.47	0.25	0.09	0.08	0.02	0.43	0.47	—	—	—
47	8	4	Remove Bidders 1 + 2 + 3 or 6 + 7 + 8	Genuine	Genuine	Right	97,091	99,232	0.46	0.42	0.21	0.16	0.09	0.01	0.05	0.46	—	—

43, i.e., 86%, right; 7, i.e., 14%, wrong

Table 1. (*Continued.*)

Auction Identifier[a]	N	Position of Bidder 304	Results Inspection of the four conditions	43, i.e., 86%, right; 7, i.e., 14%, wrong Prediction	Actual	Result	$R^2 = 0.988$ σ, actual	σ, Eq. (9)	P_{abn} values[b] 1	2	3	4	5	6	7	8	9	10
48	7	5	No abnormal bids detected	Genuine	Genuine	Right	50,748	55,627	**0.47**	**0.29**	0.03	0.02	0.03	0.05	**0.47**	—	—	—
49	5	5	Remove Bidders 1 + 2 + 3 or 4	Genuine	Cover	Wrong	7,161	7,015	**0.52**	**0.29**	**0.26**	0.19	**0.52**	—	—	—	—	—
50	5	3	Remove Bidders 1, 2, or 4 + 5	Genuine	Genuine	Right	41,483	42,189	**0.52**	**0.45**	0.14	0.10	**0.52**	—	—	—	—	—
51	6	6	No abnormal bids detected	Genuine	Genuine	Right	57,239	67,333	**0.49**	0.11	0.05	0.03	0.06	**0.49**	—	—	—	—

[a]ID = identifier.
[b]Values are in bold when $P_{abn} \geq 0.25$.

If a tentative threshold is set at $P_{abn} = 0.25$, a quick count from the last 10 columns of Table 1 reveals that whenever Bidder 304's bids were genuine. The metric P_{abn} was above a value of 0.25 on 19 occasions (right predictions, 56%) out of 34 (15 wrong predictions, 44%). However, when Bidder 304's bids were actually cover bids, the results were slightly better, with $P_{abn} > 0.25$ in 10 auctions (62.5% right predictions) out of 16 auctions (37.5% wrong predictions). This is a first approach by setting a P_{abn} value at 0.25, but seems clear that this metric has at least a moderate-to-weak correlation with the predictions generated by the extended Ballesteros–González–Cañavate model at the cost of very simple and quick calculations.

In short, unlike the previous two research components developed previously, metric P_{abn} is not a stand-alone abnormal bid detection component but a complementary coefficient whose main goal is to rank suspect bids for applying other abnormal bid detection methods such as the Ballesteros–González–Cañavate model. This is especially the case when the number of possible combinations of bidders is very high, complicating the application of the model mentioned previously to all possible scenarios, i.e., subgroups of bidders' bids. However, there is still way to go concerning the accuracy of this metric.

Another conclusion concerning these results is that the standard bids are well-represented by their respective regression straight lines once the abnormal bidders have been removed (if there are any) since their regression lines are always within the limit lines representing a good fit with the uniform distribution (otherwise not all the four conditions would have been satisfied). However, the validation carried out in this paper represents only a tentative outcome since more bid datasets such as the one used would be necessary to ensure the wider application of the method proposed, bid datasets that are unfortunately very difficult to obtain in practice.

Discussion

The complete extended Ballesteros–González–Cañavate model is potentially able to detect all abnormal bids in addition to solely abnormally expensive or cheap bids as in the original model. Furthermore, the empirical test in the paper shows the model to be remarkably accurate at detecting abnormal bids in the form of cover bids in a set of real bid data, even in the most difficult situation where competition is purely on bid value. This demonstrates that the extended model developed in this paper is robust to the uniform distribution assumption for uncapped and capped auctions. In addition, using the German Tank solution quite surprisingly results in the estimated uniform distribution supports being expressed solely as a function of the number of bidders involved.

The assumption of uniformly distributed bids is also not necessarily restrictive. Where another distribution is involved, the bid values can be transformed into a uniform distribution by using the probability values of the cumulative distribution without loss of generality.

Previous simpler models have been applied to only relatively simple bid situations, where a smart cartel might avoid being detected by using the very same tests that check exchangeability and conditional independence in reverse, by trying different bid values until they simulate real competition, while still fulfilling their hidden intentions. Therefore, despite cartel bids being generally more highly correlated than truly competitive bids (Porter and Zona 1993), provided the tests can be used by either the auctioneer or the cartel itself, competitive bidding might always be compromised.

Conclusions

This paper extends and substantially refines the Ballesteros–González–Cañavate model for detecting both abnormally high and low bids, provides mathematical expressions to approximate the bid distribution supports involved, and proposes a new metric to focus on potentially collusive bidders in uncapped auctions. This abnormal bid detection model partially avoids several drawbacks that other models suffer since it does not need any information concerning the bidders involved (apart from all the bids entered) nor about the contract, and the necessary statistical procedure is quite straightforward while the data generated allow attention to be drawn to deviations in $1/N$ multiples, which may eventually indicate a potentially abnormal behavior. However, for bid pricing decisions the influence of several other factors are acknowledged, such as market conditions, current workloads, and the relationship between the bidder and the owner or engineer; these might not always be well-represented by the method developed in this paper and may eventually generate unexpected deviations, or false abnormal bid alerts.

In the case study of the research reported in this paper, the application described performed sufficiently well overall, with almost every bid identified as abnormal by the method being an actual cover bid. However, the proposed method has not been extensively tested, since these bid datasets, such as the one analyzed in this paper, are extremely difficult to obtain so that it is unlikely that further validation tests will be possible. In practical terms, therefore, although particularly potentially useful for sifting the great amount of bid data with which contracting authorities have to work, it is unlikely that the method could be implemented beyond a first and quick check. In the case where several contract auctions are found to contain repetitive abnormal bid behavior, the additional use of other more accurate yet complex and time-consuming existing methods will always be needed.

Finally, the discovery of the central role of the number of bidders in the German Tank solution, together with the success of its use in uncapped abnormal bid detection, suggests that it may be reasonable to believe that the bid SD might not be the only parameter influenced by the number of bidders in the auction. The next logical step, therefore, is to find which other parameters may also be expressed as a function of the number of bidders and try to predict the number of bidders itself for future construction contract auctions.

Acknowledgments

References

Anderson, E. J., and Cau, T. D. H. (2011). "Implicit collusion and individual market power in electricity markets." *Eur. J. Oper. Res.*, 211(2), 403–414.

Anderson, L., Freeborn, B., and Hulbert, J. (2012). "Risk aversion and tacit collusion in a Bertrand duopoly experiment." *Rev. Ind. Org.*, 40(1), 37–50.

Aoyagi, M. (2003). "Bid rotation and collusion in repeated auctions." *J. Econ. Theory*, 112(1), 79–105.

Aoyagi, M., and Fréchette, G. (2009). "Collusion as public monitoring becomes noisy: Experimental evidence." *J. Econ. Theory, 144(3), 1135–1165.

Bajari, P., and Summers, G. (2002). "Detecting collusion in procurement auctions." *Antitrust Law J.*, 70(1), 143–170.

Bajari, P., and Ye, P. L. (2003). "Deciding between competition and collusion." *Rev. Econ. Stat.*, 85(4), 971–989.

Baldwin, L., Marshall, R., and Richard, J. F. (1997). "Bidder collusion at Forest Service timber sales." *J. Polit. Econ.*, 105(4), 657–699.

Ballesteros-Pérez, P., González-Cruz, M. C., and Cañavate-Grimal, A. (2012a). "Mathematical relationships between scoring parameters in capped tendering." *Int. J. Proj. Manage.*, 30(7), 850–862.

Ballesteros-Pérez, P., González-Cruz, M. C., and Cañavate-Grimal, A. (2013a). "On competitive bidding: Scoring and position probability graphs." *Int. J. Proj. Manage.*, 31(3), 434–448.

Ballesteros-Pérez, P., González-Cruz, M. C., Cañavate-Grimal, A., and Pellicer, E. (2013b). "Detecting abnormal and collusive bids in capped tendering." *Autom. Constr.*, 31(5), 215–229.

Ballesteros-Pérez, P., González-Cruz, M. C., Pastor-Ferrando, J. P., and Fernández-Diego, M. (2012b). "The iso-score curve graph. A new tool for competitive bidding." *Autom. Constr.*, 22, 481–490.

Ballesteros–Pérez, P., González–Cruz, M. C., Fernández–Diego, M., and Pellicer, E. (2014). "Estimating future bidding performance of competitor bidders in capped tenders." *J. Civ. Eng. Manage.*, 10.3846/13923730.2014.914096, 702–713.

Bergman, M. A., and Lundberg, S. (2013). "Tender evaluation and supplier selection methods in public procurement." *J. Purchas. Supply Manage.*, 19(2), 73–83.

Blume, A., and Heidhues, P. (2008). "Modeling tacit collusion in auctions." *J. Inst. Theor. Econ.*, 164(1), 163–184.

Carr, P. G. (2005). "Investigation of bid price competition measured through pre bid project estimates, actual bid prices and number of bidders." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364(2005)131:11(1165), 1165–1172.

Che, Y.-K., and Kim, J. (2006). "Robustly collusion-proof implementation." *Econometrica*, 74(4), 1063–1107.

Che, Y.-K., and Kim, J. (2008). "Optimal collusion-proof auctions." *J. Econ. Theory*, 144(2), 565–603.

Chowdhury, P. R. (2008). "Controlling collusion in auctions: The role of ceilings and reserve prices." *Econ. Lett.*, 98(3), 240–246.

Curtis, F. J., and Maines, P. W. (1973). "Closed competitive bidding." *Omega*, 1(5), 613–619.

David, H. A., and Nagaraja, H. N. (2003). *Order statistics*, Wiley, Hoboken, NJ.

Gayle, W. R., and Richard, J. (2008). "Numerical solutions of asymmetric, first-price, independent private values auctions." *Comput. Econ.*, 32(3), 245–278.

Goodman, L. A. (1954). "Some practical techniques in serial number analysis." *J. Am. Stat. Assoc.*, 49(265), 97–112.

Graham, D. M., and Marshall, R. C. (1987). "Collusive bidder behavior at single-object second-price and English auctions." *J. Polit. Econ.*, 95(6), 1217–1239.

Hendricks, K., Porter, R., and Tan, G. (2008). "Bidding rings and the winner's curse." *RAND J. Econ.*, 39(4), 1018–1041.

Hu, A., Offerman, T., and Sander, O. (2011). "Fighting collusion in auctions: An experimental investigation." *Int. J. Ind. Org.*, 29(1), 84–96.

Ishii, R. (2008). "Collusion in repeated procurement auction: A study of a paving market in Japan." *Rep. Prepared for the Institute of Social and Economic Research*, Osaka Univ., Suita, Japan.

Ishii, R. (2009). "Favor exchange in collusion: Empirical study of repeated procurement auctions in Japan." *Int. J. Ind. Org.*, 27(2), 137–144.

Klemperer, P. (2002). "What really matters in auction design." *J. Econ. Perspect.*, 16(1), 169–189.

Laffont, J., and Martimort, D. (2000). "Mechanism design with collusion and correlation." *Econometrica*, 68(2), 309–342.

Marshall, R. C., and Marx, L. M. (2009). "The vulnerability of auctions to bidder collusion." *Q. J. Econ.*, 124(2), 883–910.

Maskin, E., and Riley, J. (2000). "Asymmetric auctions." *Rev. Econ. Stud.*, 67(3), 413–438.

McAfee, R. P., and McMillan, J. (1992). "Bidding rings." *Am. Econ. Rev.*, 82(3), 579–599.

Ngai, S. C., Drew, D. S., Lo, H. P., and Skitmore, R. M. (2002). "A theoretical framework for determining the minimum number of bidders in construction bidding competitions." *Constr. Manage. Econ.*, 20(6), 473–482.

OECD (Organization for Economic Cooperation and Development). (2007). *Bribery in procurement, methods, actors and countermeasures*, Paris, France.

OECD (Organization for Economic Cooperation and Development). (2009). *Guidelines for fighting bid rigging in public procurement.*

Paha, J. (2011). "Empirical methods in the analysis of collusion." *Empirica*, 38(3), 389–415.

Perng, Y. H., Juan, Y. K., and Chien, S. F. (2006). "Exploring the bidding situation for economically most advantageous tender projects using a bidding game." *J. Constr. Eng. Manage.*, 10.1061/(ASCE)0733-9364 (2006)132:10(1037), 1037–1042.

Pesendorfer, M. (2000). "A study of collusion in first-price auctions." *Rev. Econ. Stud.*, 67(3), 381–411.

Porter, R. H. (2005). "Detecting collusion." *Rev. Ind. Org.*, 26(2), 147–167.

Porter, R. H., and Zona, J. D. (1993). "Detecting of bid rigging in procurement auctions." *J. Polit. Econ.*, 101(3), 518–538.

Porter, R. H., and Zona, J. D. (1999). "Ohio school milk markets: An analysis of bidding." *RAND J. Econ.*, 30(2), 263–288.

Rasch, A., and Wambach, A. (2009). "Internal decision-making rules and collusion." *J. Econ. Behav. Org.*, 72(2), 703–715.

Ruggles, R., and Brodie, H. (1947). "An empirical approach to economic intelligence in World War II." *J. Am. Stat. Assoc.*, 42(237), 72–91.

Skitmore, R. M. (2014). "Generalised gamma bidding model." *J. Oper. Res. Soc.*, 65(1), 97–107.

Skitmore, R. M., and Pemberton, J. (1994). "A multivariate approach to contract bidding mark-up strategies." *J. Oper. Res. Soc.*, 45(11), 1263–1272.

Skrzypacz, A., and Hopenhayn, H. (2004). "Tacit collusion in repeated auctions." *J. Econ. Theory*, 114(1), 153–169.

Tan, Y. T., Shen, L. Y., Khalid, A. G., and Song, S. C. (2006). "An examination of the factors affecting contractor's competition strategy: A Hong Kong study." *Int. J. Proj. Org. Manage.*, 1(1), 4–23.