Detecting Auctioneer’s Corruption: Evidence from Russian Procurement Auctions*

Pasha Andreyanov† Alec Davidson‡ Vasily Korovkin §

March 27, 2017

Abstract

This paper documents evidence of large-scale auctioneer’s corruption in procurement using a novel data set of first-price sealed-bid auctions in Russia. We develop a set of new methods for detecting the leakage of bids by the auctioneer to one of the bidders in a sealed-bid auctions setting. The data set consists of more than three million auctions over the last five years. Our first group of methods employs both the information on timings when the bids were placed as well as the bids’ data itself. Our second group of methods only uses the information on bids, without timings. The estimates from bids and timings’ method show that 8-9% of the auctions are affected by auctioneer’s corruption in the form of bid leakage. The estimates from the bids’ method show that 16% of the auctions are affected. Both types of methods allow us to explore heterogeneity of bid leakage corruption across industries and regions.

Keywords: corruption, procurement auctions, detection. JEL Classification: D44, D73, C57, L40, H57.

*We are grateful to John Asker, Leonardo Bursztyn, Ernesto Dal Bo, Lorenzo Casaburi, Christian Dippel, Georgy Egorov, Gaston Illanes, Seema Jayachandran, Kei Kawai, Lidia Kosenkova, Alexey Makarin, Nicola Persico, Robert Porter, Mar Reguant, Romain Wacziarg, Noam Yuchtman and the participants of the Brown Bag at the Anderson School of Management, DEVPEC2016 conference, GEM-BPP workshop, RSSIA2016, NEUDC2016, and Northwestern IO lunch for helpful comments and suggestions. We would also like to thank Olga Anchishkina, Ivan Begtin, and Andrei Yakovlev for providing the institutional details. We are also grateful to the UCLA Anderson Center for Global Management for financial support. This paper was previously circulated as “Corruption vs. Collusion: Evidence from Russian Procurement Auctions”.

†pandreyanov@gmail.com
‡awdavids@ucla.edu
§vaskorovkin@gmail.com, to whom correspondence should be addressed.
1 Introduction

Extensive empirical and theoretical literature is dedicated to detecting collusion in auctions.\footnote{See Marshall and Marx (2012), Whinston et al. (2008), and Porter (2005) for cases of collusion and methods of detection.} However, the literature on detecting corruption in auctions, in other words, agreements between bidders and an auctioneer, is practically non-existent. Corruption can be as important as collusion, especially in the context of public procurement and in emerging economies.\footnote{See Kenny (2007), Olken and Pande (2012), and Lewis-Faupel et al. (2014) and references therein for discussion of corruption in public procurement projects in developing countries.} If one can improve the ability of the government to detect and punish corruption in auctions in the same spirit as collusion, it could lead to significant savings of taxpayers’ money.

This paper examines a very natural type of corruption – bid leakage in sealed-bid auctions. Under bid leakage, an auctioneer forms an agreement with one of the bidders and reveals the information on all other bids to this preferred bidder. The preferred bidder can then decide whether to place her bid. In the public procurement setting, the auctioneer is a government official who represents a public body, and the bidder is one of the firms bidding for the contract.

We develop two groups of methods to detect bid leakage. The first one uses data on both the bids and their timings. The second group of methods abstracts from the rarely available information on timings and only uses the bidding data. We apply our method to the universe of first-price sealed-bid procurement auctions in Russia spanning the years 2014-2016. This data set is extremely rich. It includes detailed records of the bids of each contractor, together with the time at which each bid was placed. The data spans a broad range of industries, locations, and public bodies.

Depending on the method we use, from 8% to 16% of the auctions are affected by this type of corruption. The absolute number of contracts affected is thus either 60 or 130 thousand over five years. Scaling up the estimates to the whole universe of first-price sealed-bid auctions yields a result that 480 thousand contracts were affected by bid leakage corruption. In monetary terms, it is more than $1.3 billion in five years. We find that bid leakage corruption is present in all regions and contracts for both goods and services, but the estimates experience a substantial amount of heterogeneity.

To illustrate our approach, we start with two real examples from Russian procurement. In the first example, Firm A participated in five auctions organized by the same public body. Each auction had four bidders and a reserve price of $6,000 in ruble equivalent. In each of the five auctions, the complaining firm was the second best after the winner, by
the margins of 50 cents to $12. Surprisingly, the winner was always the same company. In an another case, Firm B faced a similar situation and noted that the winner was also always the last one to bid.\footnote{We found this and other case studies on a major online forum that discusses issues of public procurement in Russia. Link to the source (in Russian). We found many other examples in Russia that describe patterns suggestive of bid leakage. For instance, here, and here. In an even more drastic example, one of the participants complains that his firm was always the second in 60 auctions in a row. He always lost by less than a half percent of the reserve price and by around half a percent of his bid, and he always lost to the same firm. This behavior is highly suggestive that the winner knew the bid of the complaining firm Here is the source in Russian.}

We base our methods on the insights from the aforementioned examples. Our first approach exploits the fact that the preferred bidder wants to collect \textit{all} information on the bids of other participants. Thus, ideally, she wants to be the last one to bid in the auction. We calculate the probability of winning conditional on being the last one to bid. Basic independent private value model suggests that this probability should be equal to the inverse of the number of participants. Comparing the two numbers, we infer the share of corrupt auctions.

The previous approach is intuitive; however, it does not fully use the richness of the timings’ data. In a more general approach, we focus on the abnormal mass of the auctions in which winners place their bids close to the deadline disproportionately more often than other bidders. The share of corrupt auctions is then estimated from a normalized difference of the CDFs of timings of the winner and of the runner-up. Reassuringly, the placebo share of “corrupt” auctions based on the difference between the runner-up and the third bids is a statistical zero.

Most of the data sets on sealed-bid auctions lack the information on the timings of the bids. Therefore, we build a method that only employs the bidding data. The upper bound on the share of corrupt auctions is given by the share of auctions in which the winner bids slightly below the runner-up. Next, we compare the share of auctions in which the winner bids slightly below the runner-up with the share of auctions in which the runner-up bids slightly below the third strongest bidder. The normalized difference between the two gives us a lower bound on the share of corrupt auctions.

There are two caveats that we need to address in order to interpret our estimates as valid measures of corruption. The first caveat is that honest bidders might anticipate corruption and thus bid closer to the deadline and bid more aggressively. We argue that, if anything, this behavior will make our estimates from the timings method the lower bound estimates. In contrast, more aggressive bidding should not affect our bids’ method since it is based on the difference in bids and not their absolute values.
The second caveat is that even without corruption there could be inherent cost differences across bidders that lead to more efficient bidders bidding closer to the deadline. However, we show that these differences must be extremely asymmetric to explain our results. That is, we find that the distribution of timings observed in the data is the same for all of the bidders apart from the winner. Likewise, the distribution of bids’ differences is the same for all of the bidders’ pairs apart from the winner - runner-up pair.

We build on the literature on detecting collusion in auctions (see Porter 2005 and Asker et al. 2010 for a review). The main contribution of our paper is that we detect corruption – an agreement between an auctioneer and one of the bidders as opposed to collusion – an agreement between two or more bidders as in the previous literature (Hendricks and Porter 1988; Porter and Zona 1993; Baldwin et al. 1997; Porter and Zona 1999; Pesendorfer 2000; Bajari and Ye 2003; Ishii 2009; Asker 2010; Athey et al. 2011; Conley and Decarolis 2011; Haile et al. 2012; Kawai and Nakabayashi 2014; Schurter 2016). This is the first paper to build a method of corruption detection for the first-price sealed-bid auctions and one of the first papers to empirically study corruption in auctions. At the same time, we draw insights on the behavior of the bidder and the auctioneer in the auction with bid leakage corruption from a large theoretical literature. Our parsimonious model of strategic behavior in procurement auctions with bid leakage is very much in line with existing papers (Compte et al. 2005; Menezes and Monteiro 2006; Lengwiler and Wolfstetter 2006; Burguet and Perry 2007; Arozamena and Weinschelbaum 2009; Lengwiler and Wolfstetter 2010). None of them provide any empirical evidence on this type of corruption.

Our paper is also the first to exploit the timing as a strategic variable in sealed-bid auctions as opposed to open auctions such as eBay. We show that timing can matter even in sealed-bid auctions, without any considerations that are usually relevant to open auctions (see for example Bajari and Hortacsu 2003; Ockenfels and Roth 2002, 2006; Hendricks et al. 2012; Hopenhayn and Saeedi 2015). The papers mainly explain two patterns – *sniping*, i.e bidding at the very last moment and – *incremental bidding*, undercutting the bids by a little bit. We observe similar behavior in our data on first-price sealed-bid auctions. In our setting, this behavior is impossible to justify without bid leakage being in action.

From the policy side this paper contributes to the literature on electronic procurement (see Elbahnasawy 2014; Lewis-Faupel et al. 2014; Best et al. 2016) and e-governance in general (Banerjee et al. 2014; Muralidharan et al. 2014). Our findings show that public bodies

---

4 With a notable exception of Cai et al. (2013) paper that studies corruption in auction choice in land markets in China.
5 We complement the literature that studies the effects of auctioneers on entry decisions Best et al. 2016; Lacetera et al. 2016.
can still be involved in corrupt activities even despite transparent and detailed procedure of recording their activity, so e-procurement does not have to eradicate corruption in our setting. One policy lesson from our paper is that recording timing of bids is helpful for detecting various types of illegal behavior and measuring their magnitudes.

The rest of the paper is organized as follows. Section 2 describes the institutional background and the details of the procurement procedure that we analyze. Section 3 describes the data and illustrates the ideas behind our methods using simple graphical analysis. Section 4 sets up our first groups of methods and shows the results of the estimation. Section 5 abstracts from the information on timings and introduces a method that is solely based on the bids data. It shows the results from this method and compares it to bids and timings method. Section 6 presents a stylized model of bidding behavior under bid leakage, and examines the welfare implications. Section 7 explores how corruption varies across auctioneers and tests some results from the model. Section 8 concludes.

2 Institutional Background

The three most widely used procurement auctions’ types in Russia are requests for quotations, open auctions, and open tender. A request for quotation is a first-price sealed-bid reverse auction. An open auction is a reverse English auction, which means that the participants observe the behavior of each other while participating in a bidding process online. The third type of the auctions – the open tender, resembles a first-price sealed-bid reverse auction, but it has a quality component that enters the decision process. In other words, it is a scoring auction.

We focus on the requests for quotations or, as we call them throughout the paper – the first-price sealed-bid auctions. This type of procedure is used for small contracts. The reserve price of such a contract can be at most 500,000 rubles (≈$8,400), and at most 10% of annual expenditures of a public body can be assigned this way, but not more than 100 million rubles per year (≈1.6m). Typical examples of these contracts are purchases of office supplies for a municipality, books for a school, or medical supplies for a hospital. Small repairs, street cleaning and other types of services can be purchased through this procedure as well. Requests for quotations require less preparation, but they are also less

---

6 There are other variations of these auction procedures, but they cover an insignificant share of the contracts. Sources for this Section: “RusTenders” Website, Federal Law #44, Hramkin and Balsevich et al. (in Russian).

7 Please click here to update the exchange rate.

8 Please click here to update the exchange rate.
transparent for the controlling agencies to oversee.

A request for quotations proceeds as follows. The public body first posts an announcement of the auction on a special website. This public notice is standardized, and it has exhaustive information about the contract. It includes the deadline to post the bid and the requirements to qualify as a bidder. The announcement has to be uploaded to the website more than seven business days before the deadline of the auction for larger contracts and more than four business days for smaller contracts.\(^9\)

During the time of the auction, anyone can submit an application, which has to include a contract price (we call it a “bid”), together with the documentation that confirms the eligibility of the firm to participate in the request for quotation. Applications are accepted in sealed envelopes, by email or online through a special website.\(^10\) Sealed envelopes usually have to be submitted on business days from 9 AM to 1 PM or from 2 PM to 5 PM, except for the state holidays.

After the auction has ended, the applications are opened by the auctioneer. Applications can be rejected if the bidders did not meet the posted requirements. The winner is determined by the lowest bid. If the bids are equal, the earliest application submitted wins. The committee then writes a protocol with the results of the auction, which is also stored on the public website.\(^11\) Figure 1 shows a timeline for a typical auction.

---

**Figure 1: Timeline of a typical auction**

- **August 1st**: Public notice is posted; Reserve price $1,430
- **August 11th**: Sealed Bid 1: $999
- **August 12th**: Sealed Bid 2: $1,085
- **August 13th**: Deadline: Bid 1 Wins

Most of the potential participants monitor the Internet for the announcements from a given public body and after they submit a bid they receive a letter that notifies when

---

\(^9\)The size of the contract is determined by a publicly known reserve price, and the threshold level for the four-days length auctions is 250,000 rubles or \(\approx 4,200\) (Please click here to update the exchange rate).

\(^10\)Our results are similar for both online and sealed-bid auctions, so we do not examine them separately.

\(^11\)The bidders or their representatives have a right to participate in the opening procedure, can request to disclose any information from the bidding envelopes and can make a recording of the procedure.
and how the bid were submitted. So, it is relatively easy for them to trace whether the information was entered correctly.

3 Data and Graphical Analysis

We use a digital archive of all procurement auctions that took place starting January 2011. The archive consists of announcements and protocols, in a standardized .xml format. A typical announcement has the information on the terms of the contract, reserve price in the auction, the deadline of the auction and the deadline at which the work specified in the contract should be delivered. Protocols include a bid for each application, a time when the bid was submitted and also the information on whether the bid was accepted or rejected.

After matching the announcements to the protocols, and extracting all of the necessary information we obtain a data set of more than 3.3 million requests for quotations. We drop the auctions with any quote being rejected and the auctions with less than three quotes and we are left with around 900,000 auctions.

Table 1 of Appendix A shows summary statistics for the whole sample and for the new law FZ#44.

The reserve price in requests for quotations is a maximum price that needs to be hit by the winning bid in order for the auction to be considered valid. This reserve price has to be below or equal to 500,000 rubles, or $8,400 as of the current exchange rate. In our sample, the mean reserve price is 213,507 rubles, which is less than a half of a maximum allowed reserve price. Mean winning bid is 157,850, while the mean ratio of the winning bid to the reserve price is 73.3%. Those numbers are lower for the new law – FZ#44.

Table 1 of Appendix A shows summary statistics for the whole sample and for the new law FZ#44.

The new regulation, FZ#44 came in action in January 2014. The regulation abolished certain requirements that used to create red tape for requests for quotations. It also declared a transition to electronic procurement.

For more details on the reserve price distribution and the distribution of the winning bid one can resort to Appendix B. Panel A of Figure A1 shows that the distribution of reserve prices has almost a full support from 0 to 500,000 rubles. The empirical density of the reserve price is monotonically decreasing apart from the spikes at round numbers and especially on the maximum level of 500,000 rubles, and except for the neighborhood of 0. Panel B of Figure A1 shows that the winning bids have a similar distribution.

---

12 Available in FTP or HTML formats from the official website zakupki.gov.ru.
13 The auctions without quotes do not seem to provide any insights for our analysis. We assume for now that rejecting a bid is orthogonal to bid leakage that we are studying. We do not use the auctions with one or two bidders bidder in our main analysis since we can not employ our methods without observing the third bid.
14 The new regulation, FZ#44 came in action in January 2014. The regulation abolished certain requirements that used to create red tape for requests for quotations. It also declared a transition to electronic procurement.
15 Please click here to update the exchange rate.
16 For more details on the reserve price distribution and the distribution of the winning bid one can resort to Appendix B. Panel A of Figure A1 shows that the distribution of reserve prices has almost a full support from 0 to 500,000 rubles. The empirical density of the reserve price is monotonically decreasing apart from the spikes at round numbers and especially on the maximum level of 500,000 rubles, and except for the neighborhood of 0. Panel B of Figure A1 shows that the winning bids have a similar distribution.
indirectly suggesting that the level of competition might have indeed went up after the introduction of the new law.

"Table 1 goes here"

Bid leakage implies that an auctioneer agrees to cooperate with one of the bidders in an auction. Since the auctioneer is able to observe other bidders’ bids before the conclusion of the auction she informs the preferred firm of each bid as it is received. Then, once the bidder is confident that all other bids have been placed she places a bid so as to win the auction by as little as possible.

The corrupt bidder needs to collect all of the information on other bids. While the bidder has access to all information submitted to the auctioneer he cannot know the exact number of bidders beforehand. Thus, the only way for him to ensure that he observes all other bids is to place his bid as late as possible.

Figure 2. Hours to the deadline for the winner

Figure 2 illustrates this point. It shows the distribution of timings for the winner. A large share of bids occur the day before the deadline and a large share of those bids occur
in the last hour before the deadline. The share of bids is decreasing with the time to the
deadline. In every day, except for the day of the deadline, the distribution is hump-shaped
with one small gap in the middle, when there is no bidding – lunch break.

The second pattern that will imply bid leakage is that on average the winner is more
likely to be the last one to bid. Of course this will underestimate corruption by ignoring
the auctions, where there was a bid placed after the winning (corrupt) one and it was
either lower than the winning one, or the winning bidder did not have an opportunity to
undercut it.

Figure 3 shows that the number of bids placed in the last three hours before the dead-
line is large for all of the bidders, however the winner tends to bid later more than the
other bidders. Most of the extra density for the winner comes from the last 30 minutes
of the auction. The densities for the second and the third bidders are practically indistin-
guishable from each other.

Figure 3. Minutes to the deadline for the winner

Notes: Kernel density estimates for the timing of bids. Distance in minutes to the deadline. Last
three hours. The densities are drawn separately for the winning bids, runner-up bids and the third
best bid. Auctions with three and more bidders.
We quantify those differences in Table 2 of Appendix A. We are doing it by reporting the joint probabilities of having rank $i$ in time and rank $j$ in bid. Then we report the conditional probabilities of winning conditional one being the last one to bid. We do it separately for all auctions and for the law FZ#44. The excess probability of winning the auction conditional on being the last one to bid is 5.5% for the whole sample and 9.6% for the new law.

"Table 2 goes here"

We also show that the differences are negative and smaller in magnitude if we repeat the exercise by throwing away the winning and the last bid and running a placebo test with the second and the third ranks of bids and times.\textsuperscript{17}

The third pattern that might arise as a consequence of bid leakage is a small difference between bids. If the lowest honest bid is 1,000 rubles we would expect the corrupt bidder to place a bid of 999.9 to undercut the current winner.\textsuperscript{18} Figure 4 concentrates on the normalized differences in bids.\textsuperscript{19} There is a spike at zero – with the difference in bids being very close to each other or equal. Figure A2 in Appendix A show that if we zoom in even closer near zero the pattern stays the same.

\textsuperscript{17}Our interpretation of the result that corruption goes up with the introduction of the new transparency requirement is that bid leakage substituted other, more blatant types of corruption, such as directly manipulating with envelopes and/or rejecting the bids. We are working on a formal test for this.

\textsuperscript{18}Alternatively if the current winning bid is below the marginal cost of the corrupt bidder will simply decline to bid or he will place a suitably high bid.

\textsuperscript{19}We report the difference of log-bids by pairs of order statistics, however the absolute differences or differences normalized by the reserve price will yield the same results.
Figure 4. Difference in bids between the winner and the runner-up, the runner-up and the third bidder, the third bidder and the fourth bidder

Notes: Kernel density estimates and histogram for the difference of the log runner-up bid and log winning bid; difference between log third bid and log runner-up bid, and the same for the third and the fourth bid. Auctions with three and more bidders.

4 Detection using Bids and Timings

We start by formalizing the intuition that if there is no corruption winner should be as likely to be the last one to bid as any other bidder and that the distribution of timings for the winner and for other bidders should be the same. We start by setting up assumptions that are needed for this intuition to hold.

Our sample consists of \( N \) auctions indexed by \( i \). Each auction is characterized by a tuple \((b_i, t_i, K_i, X_i)\): vector of bids, vector of timings, number of bidders and other auction level covariates.\(^{20}\) For now we abstract from auction level covariates, and we treat the number of bidders as fixed, and drop the \( i \) index.

We start with a following assumption:

\[^{20}\text{b}_i = (b_{i1}, ..., b_{iK_i}), \ \text{t}_i = (t_{i1}, ..., t_{iK_i}), \ \text{and} \ K \text{ – number of bidders in the auction.} \]
**Assumption 1. Independence and Symmetry of Timings.**

1a. \( b \perp t \),

1b. \( g(..., t^k, ..., t^j, ...) = g(..., t^j, ..., t^k, ...), \ \forall k, j \),

where \( g \) is the joint p.d.f of timings.

The first part of the assumption holds when the timing does not enter the maximization problem of a bidder, and it is not a strategic variable. When can the assumption fail to hold? One violation of this assumption happens, when there is an unobserved type of a bidder that is correlated with the timing decision. If more qualified firms take more time to prepare their bids, they might place the bid at the conclusion of the auction. Although this seems highly implausible for the type of the auctions in our data we address these concerns later in the section. The second part of the assumption is a technical one.

In the absence of corruption Assumption 1 will imply that the winner will be the last one to bid with a probability of \( 1/K \), where \( K \) is the number of bidders in the auction.

Assume that an auction is corrupt with probability \( \alpha \). Assume in addition that in corrupt auctions the winner is the last one to bid with probability \( \theta (1/K < \theta \leq 1) \). In the limiting case, the winner is always the last one to bid after discovering the bids of all other participants.

We can decompose the observed probability of winning conditional on being the last one into two components:

\[
P[bidder wins | he bids last] = \alpha \cdot \theta + \frac{1}{K} (1 - \alpha).
\]

Solving for \( \alpha \) one gets,

\[
\alpha = \frac{P[bidder wins | he bids last] - 1/K}{\theta - 1/K}.
\]

We use a lower bound estimate for the share of corrupt auctions, \( \alpha \). We plug in \( \theta = 1 \) and an empirical analog of \( P[bidder wins | he bids last] \).

We report the resulting estimates \( \hat{\alpha}(K) \) conditional on the number of bidders in Table 3 (Appendix A).\(^{21}\)

\(^{21}\)In this subsection we only use the new law FZ#44.

The share of corrupt auctions estimated by this method varies from 7% to 10%. Since we want to argue that this is indeed driven by the mechanism that we discussed, we run two other procedures.

First, we concentrate on two subsamples. The first one has only the auctions with all of the bids placed at least five hours to the deadline. The second subsample has all of...
the other auctions. Bid leakage implies that the results from Table 3 are not driven by the early bidding subsample. Indeed, Table 4, Panel A show that early bidding does not account for much of the pattern, while the rest of the sample does.

"Table 4 goes here"

The second exercise that we run is a placebo, where we drop all of the last and winning bids and test whether the artificial "winner" tend to bid later. Table 5 shows that even though the estimates are sometimes statistically significant, they are small in magnitude.

"Table 5 goes here"

4.1 Method based on conditional CDFs of timings

Denote order statistics for bids and times as $b_{p(k)}$ and $t_{p(k)}$. For instance, the minimal bid is $b_{p(1)}$, while the time that is closest to the deadline is $t_{p(1)}$. We will also need a following notation for the CDF of timings conditional on the corresponding bid being the $k$th lowest

$$G_{b_{p(k)}}(t) := \mathbb{P}[t_j < t | b_j = b_{p(k)}].$$

Assume again that an auction is corrupt with probability $\alpha$. The observed CDF of timings of the winner can be decomposed into a corrupt and a fair component. We write it as,

$$G_{b_{p(1)}}(t) = \alpha \cdot G_{b_{p(1)}}^{\text{corrupt}}(t) + (1 - \alpha) \cdot G_{b_{p(1)}}^{\text{fair}}(t).$$

We need to choose a parametrization for $G_{b_{p(1)}}^{\text{corrupt}}(t)$ keeping in mind that with corruption the winner will bid close to the deadline. For now we use a CDF that is discrete at 0 $G_{b_{p(1)}}^{\text{corrupt}}(t) = 1_{t=0}$, but we can also use a CDF of a uniform distribution on $[0, \epsilon]$, where $\epsilon$ is relatively small compared to an average distance to the deadline. The results will be similar if we use uniform distribution instead of discrete. The interpretation of it that if corruption happens the bid of the winner is by necessity placed at the deadline.

The assumption of the equality of the CDFs gives us that under no corruption: $G_{b_{p(1)}}^{\text{fair}}(t) = G_{b_{p(2)}}^{\text{fair}}(t) = G_{b_{p(2)}}^{\text{obs}}(t)$. In this case $\alpha$ is identified. In the model it is equal to

$$\alpha = \frac{G_{b_{p(1)}}(t) - G_{b_{p(2)}}(t)}{1_{t=0} - G_{b_{p(2)}}(t)}.$$

A sample analog of it depends on $t$.

$$\hat{\alpha}(t) = \frac{\hat{G}_{b_{p(1)}}(t) - \hat{G}_{b_{p(2)}}(t)}{H(t) - \hat{G}_{b_{p(2)}}(t)} = \frac{1}{N} \sum 1\{t_{b_{p(1)}} < t\} - 1\{t_{b_{p(2)}} < t\} \frac{1}{N} \sum NH(t) - 1\{t_{b_{p(2)}} < t\},$$

22 The results stay the same if we replace five hours by ten hours and 30 minutes by 10 or 60 minutes.
23 One can impose some other parametric distribution and $\epsilon$ with a maximum likelihood procedure.
where $\hat{G}(t)$ is an empirical CDF. Another estimate of $\alpha$ can be derived by choosing a grid of times $t$ and averaging $\hat{\alpha}(t)$ over the grid.

$$\hat{\alpha} = \frac{1}{|T|} \sum_{t \in T} \hat{\alpha}(t) = \frac{1}{|T|} \sum_{t \in T} \frac{\hat{G}_{b(1)}(t) - \hat{G}_{b(2)}(t)}{1_{t=0} - \hat{G}_{b(2)}(t)},$$

where $T$ is some grid. We show that $\hat{\alpha}(t)$ and $\tilde{\alpha}$ are consistent for $\alpha$ and derive their asymptotic properties in Appendix C.

In practice we implement the method in a following way. We cut the sample to the last three days before the deadline, in other words 72 hours. We are doing it to avoid finite sample bias arising from the skewness of the distribution of timings towards the deadline (see for example Figure 2). We also drop auctions where two bids were submitted together (within fifteen minutes from each other). These two constraints leave us with around 2/3 of the initial sample. Next we get the estimates of $\hat{\alpha}(t)$ for $t = 0.5$, $t = 1$ and $\tilde{\alpha}$ for a grid of points on $[0, 5]$ with a step of 0.5.

We use an average of $G_{(2)}$ and $G_{(3)}$ together to build another counterfactual. Specifically, instead of using $G_{b(1)}^{fair}(t) = G_{b(2)}(t)$, we use $G_{b(1)}^{fair}(t) = (G_{b(2)}(t) + G_{b(3)}(t))/2$.

In addition we set up two placebos $\hat{\alpha}$ estimated from only the difference of $G_{(2)}$ and $G_{(3)}$ and the difference of $G_{(3)}$ and $G_{(4)}$

$$\hat{\alpha}_{placebo}(t) = \frac{\hat{G}_{b(2)}(t) - \hat{G}_{b(3)}(t)}{1_{t=0} - \hat{G}_{b(3)}(t)}.$$ 

Table 3 shows the results for the general method. All of the main estimates are significant on 5% (and 1%) level. The overall share of corrupt auctions varies from 6 to 9% depending on the cutoff level, with an average over cutoffs being 8%. Placebos are either insignificant or small in magnitude.

"Table 6 goes here"

We made two implicit assumptions to estimate $\alpha$.

We assume that without corruption

$$G_{b(1)}^{fair}(t) = G_{b(2)}^{fair}(t) = ... = G_{b(K)}^{fair}(t).$$

This assumption directly follows from the independence of bids and timings, but it is weaker.

We can not test it, but what we can test directly

$$G_{b(2)}^{obs.}(t) = ... = G_{b(K)}^{obs.}(t).$$
The second assumption is that \( G_{b(2)}^{\text{obs.}}(t) = \ldots = G_{b(K)}^{\text{obs.}}(t) = G_{b(2)}^{\text{fair}}(t) \). The former set of equalities holds if the bidders react to corruption in a similar way, while the last one states that they do not change their equilibrium choice of timing with corruption. What is more likely to happen in practice is that even honest bidders bid later in a presence of corruption. Namely, that

\[
G_{b(2)}^{\text{obs.}}(t) \geq G_{b(k)}^{\text{fair}}(t), \; \forall k > 1.
\]

In this case our estimate will be attenuated.

One can test whether the winners’ distribution of timing is different from the runner-up, while the latter is the same as for the third bidder. Figure 5 shows the CDFs of timings for the winner, runner-up, third and fourth bidders and shows the test-statistics for the Kolmogorov-Smirnov test. While all the differences are significant, the test statistics is ten times larger for the difference of the winner and the runner-up. Visually there is no difference between the CDFs of all the bidders apart from the winner.

Figure 5. CDFs by ranking of bidder.

5 Detection: Bids Only

Assume again that the share of corrupt auction is \( \alpha \), but now we are studying bids’ margin – the difference between bids. In order to control for auction level heterogeneity (both
observable for an econometrician and unobservable) we assume separability of log-bids.\textsuperscript{24}

The bid of bidder $k$ in auction $i$ is modeled as

$$\log b_{ki} = \eta_i + \theta_{ki}$$

where $\eta_i$ is auction characteristic that can be unobserved for an econometrician and $\theta_{ki}$ is an idiosyncratic component of the bid. We assume that $\theta_{ki}$ is i.i.d. For any pair of bids in the auction it is true that

$$\log b_{ki} - \log b_{ji} = \theta_{ki} - \theta_{ji}.$$ 

So taking the difference of logs allows us to control for auction level heterogeneity under our assumptions.

Specifically it will hold for order statistics. Define $\delta_{jk} = \theta_{(k)i} - \theta_{(j)i}$

So the difference in log-bids between the winner and the runner-up under our assumptions can be represented as a difference between two order statistics of i.i.d random variables.

Next we look at the probability of $\delta_{12}$ getting into an interval of 0 to $\epsilon$. We drop auction level index $i$.

For instance,

$$\Pr[\delta_{12} < \epsilon] = \alpha \cdot 1 + (1 - \alpha) \cdot \Pr^{fair}[\delta_{12} < \epsilon],$$

or

$$\alpha = \frac{\Pr[\delta_{12} < \epsilon] - \Pr^{fair}[\delta_{12} < \epsilon]}{1 - \Pr^{fair}[\delta_{12} < \epsilon]},$$

The main problem here is building a good estimate for $\Pr^{fair}[\delta_{12} < \epsilon]$.

If the bids are never close in non-corrupt auctions, then $\Pr^{fair}$ is zero and this yields an upper bound estimate on the share of corrupt auctions $- \bar{\alpha}$.\textsuperscript{25}

If in addition the distribution of $\theta$ is symmetric around its mean,\textsuperscript{26}

$$\Pr^{fair}[\delta_{12} < \epsilon] \leq \Pr^{fair}[\delta_{23} < \epsilon]$$

This yields a lower bound for the share of corrupt auctions $- \alpha$. In other words

$$\alpha \in [\bar{\alpha}, \alpha] = \left[\frac{\Pr[\delta_{12} < \epsilon] - \Pr[\delta_{23} < \epsilon]}{1 - \Pr[\delta_{23} < \epsilon]}, \Pr[\delta_{12} < \epsilon]\right].$$

\textsuperscript{24}although this assumption is restrictive since it does not allow bidders’ effect to vary across the auctions, we will argue that it might hold in our case.\textsuperscript{25}We set $\epsilon = 1\%$ of the reserve price for now. In Panel A of Table 8 of Appendix A we also study the case of $\epsilon = 0.5\%, 2\%, 5\%$.\textsuperscript{26}We prove it in Appendix C.
Table 4 shows the results of estimation for $\bar{\alpha}$ and $\alpha$. The estimates are higher than the ones derived from the bids and timings method. The pooled estimate from the conservative part of the baseline method gives the share of corrupt auctions being almost 16%. As reported in Table 5 the estimate is somewhat sensitive to a choice of $\epsilon$ and varies from 10 to 25%. An optimal choice of $\epsilon$ is work in progress.

In addition to controlling for all auction level variables we also only control for the reserve price by normalizing the absolute bid amounts by it, without taking logs. The results are reported in Table 5 Panel B. They are somewhat similar larger than the results for homogenized bids, but are qualitatively similar. This alternative normalization is motivated by Kawai and Nakabayashi (2014).

6 The model

We start with a standard first-price sealed-bid procurement auction with independent private costs. The goal of this section is to address two questions: (a) does bid leakage make the bidders bid more aggressively by reducing their bid closer to their true costs? (b) What are the revenue and efficiency implications of that? The answers to both of the questions depend on the underlying distribution of costs. Thus, it is theoretically possible that corruption makes honest bidders bid more aggressively, reducing the expected costs of the public body.

The public body is buying a good or a service from $N$ buyers, indexed by $j$. The cost of bidder $j$ is drawn from a CDF $F$ on the interval $[0, \bar{c}]$. All the players are risk-neutral. First, we describe the case without corruption. We call this case a fair auction. We drop the bidder’s index for convenience.

In the case of a fair auction bidder maximizes her expected utility,

$$\Pi^f(b) = (b - c)(1 - F(\phi(b)))^{N-1},$$

where $b$ is the bid, $c$ are the costs, and $\phi^f(\cdot)$ is an inverse of the bidding function without corruption. Since we are looking for a symmetric equilibrium we assume that $\phi^f(\cdot)$ is the same for all bidders. This equilibrium is characterized by a following first order condition:

$$b - \phi^f(b) = \frac{1 - F(\phi^f(b))}{\phi^f(b)f(\phi^f(b))(N - 1)}. \quad (1)$$

---

27 It is equivalent to the independent private values assumption for the value auctions or IPV.

28 Expected cost minimization is equivalent to revenue maximization in value auctions. The results for value auctions are derived in Arozamena and Weinschelbaum (2009). Conditions for procurement auctions have minor differences from their setting, but those differences are of a technical character.
Corruption in our setting takes a particular form. An auctioneer reveals the bids to the preferred bidder and the preferred bidder can place a bid \( b^c = \min \{-c\} - \epsilon \). The expected profit of an honest bidder is,

\[
\Pi^c(b) = (b - c)(1 - F(b))(1 - F(\phi^c(b)))^{N-2}.
\]

And the first order condition is,

\[
b - \phi^c(b) = \frac{(1 - F(\phi^c(b)))(1 - F(b))}{(N - 2)\phi^c(b)f(\phi^c(b))(1 - F(b)) + (1 - F(\phi^c(b)))f(b)}, \tag{2}
\]

We assume that virtual valuation is weakly-increasing as is commonly assumed in the literature (Myerson (1981)).

**Assumption 1. Regularity.** \( w(x) = x - \frac{1-F(x)}{f(x)} \) is weakly-increasing.

One can show that Assumption 1 is sufficient, but not necessary for the existence of equilibrium with corruption, with bidding functions derived from (2).\(^{29}\)

If corruption makes the bidders more aggressive, for the same value of the costs with corruption the bid will be lower. Or the costs for the corruption case have to be higher to achieve the same bid level. More aggression implies \( \phi^c(b) > \phi^f(b) \) \( \forall b \). Using the F.O.C.s we get,

\[
(N - 2)\phi^c(b)H(\phi^c(b)) + H(b) - (N - 1)\phi^f(b)H(\phi^f(b)) > 0, \tag{3}
\]

where \( H(x) = f(x)/(1 - F(x)) \) is a hazard function.

**Assumption 2. Concavity of the virtual valuation.** \( w(x) = x - \frac{1-F(x)}{f(x)} \) is concave.

One can show that Assumption 2 will make (3) hold and thus the honest bidders will bid more aggressively in case of corruption and will reduce their bids compared to a fair auction. Again this assumption is sufficient and it can be relaxed, but again it is true for quite a general choice of the underlying costs distribution.

As a matter of fact Assumptions 1 and 2 are sufficient for corruption to be inefficient and not optimal (see Arozamena and Weinschelbaum (2009)).

However, if Assumption 1 holds, while Assumption 2 is violated it can happen that honest bidders bid more aggressively with corruption, and it leads to a reduction of expected costs, thus making corruption optimal from a point of the government.\(^{30}\)

\(^{29}\)The case of two bidders is trivial. The case with more than two bidders is something that we establish in Appendix C (in progress). There is a broad class of distributions that satisfy Assumption 1, e.g uniform, normal, logistic, extreme values, exponential, Laplace, Pareto, chi-squared and chi. Power function distribution satisfies it for a value of the main parameter larger than 1.

\(^{30}\)We are working on the example of this “optimality of corruption” that involves Pareto distribution (similarly to Arozamena and Weinschelbaum (2009)).
Now we switch to a more realistic setting, when bidders do not know ex-ante, whether the auction is corrupt or not, but they correctly estimate the probability of an auction being corrupt – $\alpha$. The payoff function for an honest bidder becomes

$$\Pi(b) = (b - c)(\alpha (1 - F(b))(1 - F(\phi(b)))^{N-2} + (1 - \alpha)(1 - F(\phi(b)))^{N-1}).$$

This function is a weighted average of the function for honest auctions and auctions, where there is always a corrupt bidder:

$$\Pi(b) = \alpha \Pi^c(b) + (1 - \alpha)\Pi^f(b).$$

The bidding strategy will be bounded between the bidding strategy for the fair auctions and the bidding strategy for corrupt auctions. If the bidders bid more aggressively in a corrupt auction, so will they do in the case where corruption happens with probability $\alpha$ and the reverse is true. Thus corruption can be socially optimal in a narrow sense defined by the model. We discuss it in the next subsection.

However, corruption does not introduce any discontinuities in the bidding functions of honest bidders. Hence, on average, the difference between the winner and the runner-up in an auction could not shrink to zero.

7 Heterogeneity of Auctioneers (in progress)

We concentrate on the bids & timings method outlined in the first part of section 4. Namely we estimate $\alpha$ from excessive probabilities. Instead of estimating an average $\alpha$ we concentrate on a subset of auctioneers that participated in at least 200 auctions. There are 158 of those.

For each of the auctioneers we derive $\hat{\alpha}(3)$ – the estimate for auctions with three bidders. The estimate for the pdf of $\hat{\alpha}$ is reported in Figure 6.
For most of the auctioneers the estimate of $\alpha$ is positive. Some of them have $\hat{\alpha}$ close to zero or even negative (albeit insignificant). We pool auctioneers with high and with low $\hat{\alpha}$ together. We define the former by $\hat{\alpha} \geq 0.175$, while the latter is defined by $\hat{\alpha} \in [0, 0.03]$. Finally we study the distribution of log-bids for all the bids pooled apart from the winning one. We compare it to the distribution of all log-bids.
It is clear from Figure 7 that bidders in auctions with corrupt auctioneers do not bid more aggressively. If anything the opposite is true. However, the differences can be confounded by region, type of good and services and other auction-level covariates that we need to control for. (in progress)

8 Discussion

This paper is the first one to detect statistically incidence of bid leakage auctioneers’ corruption in the first-price sealed-bid auctions. We develop two methods. The first uses the fact that the winners are bidding abnormally close to the deadline, while the second one exploits the information on the bids’ margins of the bidders. We find evidence of bid leakage corruption of a magnitude of 8 to 16%. These numbers translate to 60 or 130 thousand contracts with corrupt bidders. If we scale it up to include the auctions with only two bidders, the number goes up to 480 thousand contracts and up to $1.4 billion dollars involved.

This paper suggests, among other things, that recording the timings of bids, even when the timing is irrelevant to the outcome of the auction, could help curtailing non-
competitive behavior in the future.

According to our numbers bid leakage is quite a widespread phenomenon in Russia, but it also appears in other settings. For instance, the government of Singapore banned Siemens for five years from participating in any public procurement after the company bribed an official to learn the bids of their competitors in an auction. A similar case happened when the rights for building a new airport in Berlin were sold and the rights for building a tunnel in Slovenia (Lengwiler and Wolfstetter 2006). Finally, in Italy, corrupt bureaucrats invented to use a laparoscope – a surgical device with a camera – to get inside the envelopes with the bids without damaging them. Arguably, bid leakage is the most fundamental violation stemming from the agreement between an auctioneer and a bidder in a sealed-bid auction setting.

References


---

31 The link that discusses the case (in Italian), more details.
Best, M., J. Hjort, and D. Szakonyi (2016). Individuals and organizations as sources of state effectiveness, and consequences for policy design.


## Appendix A: Tables

Table 1: Summary statistics

<table>
<thead>
<tr>
<th>Panel A. Whole Sample</th>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of bidders</td>
<td>$K$</td>
<td>4.0</td>
<td>3</td>
<td>1.6</td>
</tr>
<tr>
<td>Reserve price, rubles</td>
<td>$R$</td>
<td>213,507</td>
<td>180,000</td>
<td>153,604</td>
</tr>
<tr>
<td>Winner’s bid, rubles</td>
<td>$b_{(1)}$</td>
<td>157,850</td>
<td>125,412</td>
<td>124,287</td>
</tr>
<tr>
<td>Winner’s time to the deadline, hours</td>
<td>$t_{b_{(1)}}$</td>
<td>27.8</td>
<td>4.4</td>
<td>47.1</td>
</tr>
<tr>
<td>Winning bid to reserve price</td>
<td>$r_{(1)} = \frac{b_{(1)}}{R}$</td>
<td>73.3%</td>
<td>76.5%</td>
<td>18.0%</td>
</tr>
<tr>
<td>Winning to second best bid distance</td>
<td>$db_{(12)} = \frac{b_{(2)} - b_{(1)}}{b_{(2)}}$</td>
<td>5.6%</td>
<td>2.3%</td>
<td>8.1%</td>
</tr>
</tbody>
</table>

*Notes: The sample includes only auctions with three or more bidders and no rejected bids. $N = 841,552$*

<table>
<thead>
<tr>
<th>Panel B. Only FZ#44</th>
<th>Variable</th>
<th>Mean</th>
<th>Median</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of bidders</td>
<td>$K$</td>
<td>3.9</td>
<td>3</td>
<td>1.5</td>
</tr>
<tr>
<td>Reserve price, rubles</td>
<td>$R$</td>
<td>193,233</td>
<td>149,919</td>
<td>150,017</td>
</tr>
<tr>
<td>Winner’s bid, rubles</td>
<td>$b_{(1)}$</td>
<td>134,112</td>
<td>95,460</td>
<td>116,138</td>
</tr>
<tr>
<td>Winner’s time to the deadline, hours</td>
<td>$t_{b_{(1)}}$</td>
<td>32.3</td>
<td>17.9</td>
<td>48.2</td>
</tr>
<tr>
<td>Winning bid to reserve price</td>
<td>$r_{(1)} = \frac{b_{(1)}}{R}$</td>
<td>68.8%</td>
<td>72.0%</td>
<td>19.5%</td>
</tr>
<tr>
<td>Winning to second best bid distance</td>
<td>$db_{(12)} = \frac{b_{(2)} - b_{(1)}}{b_{(2)}}$</td>
<td>7.9%</td>
<td>4.6%</td>
<td>9.3%</td>
</tr>
</tbody>
</table>

*Notes: The sample includes only auctions with three or more bidders and no rejected bids. $N = 171,539$*
Table 2: Rank of the Bid by Rank of the Price

Panel A1. Ranks, all auctions

<table>
<thead>
<tr>
<th>Time Rank</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>12.33%</td>
<td>9.76%</td>
<td>11.24%</td>
</tr>
<tr>
<td>Bid Rank 2</td>
<td>11.32%</td>
<td>11.12%</td>
<td>10.89%</td>
</tr>
<tr>
<td>3</td>
<td>9.68%</td>
<td>12.46%</td>
<td>11.20%</td>
</tr>
</tbody>
</table>

Panel A2. Conditional Probabilities, all auctions

|                      | P(win|last) | P(win|not last) | Difference |
|----------------------|---------|-------------|------------|
|                      | 36.99%  | 31.50%      | 5.49%      |

|                      | P(runner-up|second or third time) | P(loser|second or third time) | Difference |
|----------------------|--------------------------|---------------------------|------------|
|                      | 47.16%                   | 49.31%                    | -2.15%     |

Panel B1. Ranks, law FZ#44

<table>
<thead>
<tr>
<th>Time Rank</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>13.24%</td>
<td>9.68%</td>
<td>10.41%</td>
</tr>
<tr>
<td>Bid Rank 2</td>
<td>10.61%</td>
<td>11.39%</td>
<td>11.33%</td>
</tr>
<tr>
<td>3</td>
<td>9.48%</td>
<td>12.26%</td>
<td>11.59%</td>
</tr>
</tbody>
</table>

Panel B2. Conditional Probabilities, law FZ#44

|                      | P(win|last) | P(win|not last) | Difference |
|----------------------|---------|-------------|------------|
|                      | 39.72%  | 30.14%      | 9.58%      |

|                      | P(runner-up|second or third time) | P(loser|second or third time) | Difference |
|----------------------|--------------------------|---------------------------|------------|
|                      | 48.16%                   | 49.44%                    | -1.28%     |

Notes: The sample includes only auctions with three bidders, however the results for more bidders will be similar. The (i,j) element of the matrix denotes the probability that a bidder submits the j-th lowest bid in the second round conditional on submitting the i-th lowest bid in the first round. Conditional probabilities of winning when being the last one and not the last one, the same for the runner-ups. Panel A: the number of auctions is 457,014. Panel B: the number of auctions is 95,417.
Table 3: Bids & Timings. Baseline Method. Shares of Corrupt Auctions.

<table>
<thead>
<tr>
<th>OLS Estimate</th>
<th>(1) K=3</th>
<th>(2) K=4</th>
<th>(3) K=5</th>
<th>(4) K&gt;5</th>
</tr>
</thead>
<tbody>
<tr>
<td>(OLS Estimate)</td>
<td>0.0958***</td>
<td>0.0864***</td>
<td>0.0646***</td>
<td>0.0688***</td>
</tr>
<tr>
<td>(Standard errors in parentheses)</td>
<td>(0.00186)</td>
<td>(0.00252)</td>
<td>(0.00332)</td>
<td>(0.00262)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.301***</td>
<td>0.228***</td>
<td>0.187***</td>
<td>0.127***</td>
</tr>
<tr>
<td>(Standard errors in parentheses)</td>
<td>(0.00107)</td>
<td>(0.00126)</td>
<td>(0.00148)</td>
<td>(0.000971)</td>
</tr>
<tr>
<td>N</td>
<td>95,417</td>
<td>39,187</td>
<td>18,089</td>
<td>103,732</td>
</tr>
<tr>
<td>Implied share</td>
<td>0.095</td>
<td>0.086</td>
<td>0.065</td>
<td>~0.06</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.009</td>
<td>0.007</td>
<td>0.004</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Notes: All shares are significant at 1% level. Columns present the results by the number of bidders. Only the new law. The OLS estimates are from the regression on dummy for being the winner on the dummy for being the last one ("bid-level panel").
Table 4: Bids & Timings. Early bidders and Late bidders.

Panel A. Early bidders only

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K&gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLS Estimate</td>
<td>0.0244***</td>
<td>0.0459***</td>
<td>0.0200**</td>
<td>0.0343***</td>
</tr>
<tr>
<td></td>
<td>(0.00342)</td>
<td>(0.00562)</td>
<td>(0.00869)</td>
<td>(0.00860)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.325***</td>
<td>0.239***</td>
<td>0.196***</td>
<td>0.138***</td>
</tr>
<tr>
<td></td>
<td>(0.00197)</td>
<td>(0.00281)</td>
<td>(0.00389)</td>
<td>(0.00326)</td>
</tr>
<tr>
<td>N</td>
<td>28,484</td>
<td>7,907</td>
<td>2,648</td>
<td>6,257</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.009</td>
<td>0.007</td>
<td>0.004</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Panel B. No early bidders

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>K=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K&gt;5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLS Estimate</td>
<td>0.126***</td>
<td>0.0966***</td>
<td>0.0722***</td>
<td>0.0724***</td>
</tr>
<tr>
<td></td>
<td>(0.00221)</td>
<td>(0.00281)</td>
<td>(0.00359)</td>
<td>(0.00276)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.291***</td>
<td>0.226***</td>
<td>0.186***</td>
<td>0.126***</td>
</tr>
<tr>
<td></td>
<td>(0.00128)</td>
<td>(0.00141)</td>
<td>(0.00161)</td>
<td>(0.00102)</td>
</tr>
<tr>
<td>N</td>
<td>66,933</td>
<td>31,280</td>
<td>15,441</td>
<td>97,475</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.016</td>
<td>0.009</td>
<td>0.005</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Notes: The results are for FZ#44 only. Early bidders – auctions with all bids placed at least 5 hours before the deadline.
Table 5: Bids & Timings. Placebos.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>I(runner up)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=4</td>
<td>-0.013***</td>
<td>0.026***</td>
<td>0.022***</td>
<td>0.019***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K=5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K&gt;5</td>
<td>0.008*</td>
<td>0.014**</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.006)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.494***</td>
<td>0.313***</td>
<td>0.238***</td>
<td>0.193***</td>
<td>0.481***</td>
<td>0.318***</td>
<td>0.239***</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>I(third)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>65038</td>
<td>29199</td>
<td>14298</td>
<td>16864</td>
<td>20326</td>
<td>10914</td>
<td>13972</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.000</td>
<td>0.001</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Notes: Columns (1) to (4) shows the results for auctions with three and more bidders, with winner and last bid being dropped. Columns (5) to (7) repeat the same exercise for auctions with four and more bidders with winner and runner up as well as last and second-last bids being dropped.
Table 6: Bids & Timings, CDFs. Shares of Corrupt Auctions.

<table>
<thead>
<tr>
<th></th>
<th>t=0.5</th>
<th>t=1</th>
<th>t=2</th>
<th>t in [0,3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_{12})</td>
<td>0.064</td>
<td>0.084</td>
<td>0.091</td>
<td>0.081</td>
</tr>
<tr>
<td>CI</td>
<td>[0.052, 0.075]•</td>
<td>[0.082, 0.087]•</td>
<td>[0.089, 0.094]•</td>
<td>[0.079, 0.082]•</td>
</tr>
<tr>
<td>(\alpha_{123})</td>
<td>0.063</td>
<td>0.085</td>
<td>0.096</td>
<td>0.092</td>
</tr>
<tr>
<td>CI</td>
<td>[0.052, 0.074]•</td>
<td>[0.083, 0.087]•</td>
<td>[0.093, 0.099]•</td>
<td>[0.091, 0.093]•</td>
</tr>
<tr>
<td>N</td>
<td>513,577</td>
<td>513,577</td>
<td>513,577</td>
<td>513,577</td>
</tr>
</tbody>
</table>

* indicates significance on 5% level.

<table>
<thead>
<tr>
<th></th>
<th>Placebos</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_{23})</td>
<td>-0.001</td>
</tr>
<tr>
<td>CI</td>
<td>[-0.011, 0.009]</td>
</tr>
<tr>
<td>(\alpha_{34})</td>
<td>-0.006</td>
</tr>
<tr>
<td>CI</td>
<td>[-0.015, 0.004]</td>
</tr>
<tr>
<td>N</td>
<td>513,577</td>
</tr>
</tbody>
</table>

* indicates significance on 5% level.

Notes: Auctions with simultaneous bidding are excluded. Columns show \(\hat{\alpha}\) for different choices of \(t\).
Table 7: Bids’ margins results

<table>
<thead>
<tr>
<th>K</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Pooled</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline:</td>
<td>0.32***</td>
<td>0.297***</td>
<td>0.287***</td>
<td>0.275***</td>
<td>0.304***</td>
</tr>
<tr>
<td>Upper Bound</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Baseline:</td>
<td>0.152***</td>
<td>0.178***</td>
<td>0.163***</td>
<td>0.147***</td>
<td>0.157***</td>
</tr>
<tr>
<td>Lower Bound</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0)</td>
</tr>
<tr>
<td>N</td>
<td>431,555</td>
<td>191,121</td>
<td>94,215</td>
<td>49,178</td>
<td>833,017</td>
</tr>
</tbody>
</table>

Notes: *** p<0.01. Columns present the results by the number of bidders. The last column presents pooled estimates. Auctions without simultaneous bidding.
Table 8: Bids. Baseline Method. Shares of Corrupt Auctions. Robustness.

### Panel A. Homogenized Bids

<table>
<thead>
<tr>
<th>K</th>
<th>ε</th>
<th>0.005</th>
<th>0.01</th>
<th>0.02</th>
<th>0.05</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.206</td>
<td>0.320</td>
<td>0.449</td>
<td>0.619</td>
<td>431,185</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.087</td>
<td>0.152</td>
<td>0.220</td>
<td>0.273</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.185</td>
<td>0.297</td>
<td>0.425</td>
<td>0.607</td>
<td>191,046</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.113</td>
<td>0.178</td>
<td>0.234</td>
<td>0.264</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.178</td>
<td>0.287</td>
<td>0.414</td>
<td>0.597</td>
<td>94,164</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.105</td>
<td>0.163</td>
<td>0.207</td>
<td>0.208</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.171</td>
<td>0.274</td>
<td>0.402</td>
<td>0.590</td>
<td>49,159</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.098</td>
<td>0.147</td>
<td>0.184</td>
<td>0.176</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled</td>
<td>0.192</td>
<td>0.304</td>
<td>0.432</td>
<td>0.609</td>
<td>832,478</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.096</td>
<td>0.157</td>
<td>0.215</td>
<td>0.248</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Panel B. Bid/Reserve Price

<table>
<thead>
<tr>
<th>K</th>
<th>ε</th>
<th>0.50%</th>
<th>1%</th>
<th>2%</th>
<th>5%</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>0.238</td>
<td>0.359</td>
<td>0.487</td>
<td>0.661</td>
<td>431,185</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.116</td>
<td>0.188</td>
<td>0.255</td>
<td>0.308</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.231</td>
<td>0.350</td>
<td>0.480</td>
<td>0.670</td>
<td>191,046</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.150</td>
<td>0.218</td>
<td>0.271</td>
<td>0.299</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.234</td>
<td>0.352</td>
<td>0.484</td>
<td>0.679</td>
<td>94,164</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.147</td>
<td>0.206</td>
<td>0.244</td>
<td>0.234</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>0.234</td>
<td>0.354</td>
<td>0.485</td>
<td>0.687</td>
<td>49,159</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.142</td>
<td>0.196</td>
<td>0.220</td>
<td>0.187</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pooled</td>
<td>0.236</td>
<td>0.356</td>
<td>0.486</td>
<td>0.670</td>
<td>832,478</td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.131</td>
<td>0.197</td>
<td>0.251</td>
<td>0.282</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: All shares are significant at 1% level. Columns represent the number of bidders. The last column pools the estimates.
Appendix B: Extra Graphs

Figure A1: Distribution of the Reserve Prices and Winning Bids

Note: Histograms for Reserve Price and Winning Bid distributions, fraction of all auctions on the y–axis. Left Panel: Reserve Price, Right Panel: Winning Bid. All numbers are in rubles.

Figure A2. Difference in bids between the winner and the runner-up, and the runner-up and the third bidder, neighborhood of zero

Notes: Kernel density estimates and histogram for the difference of the runner-up bid and the winning bid, normalized by the reserve price; difference between the third bidder and the runner-up, normalized by the reserve price. Auctions with three and more bidders.
Appendix C: Proofs

Proof of Lemma 1.

We wish to proof that $G_{b_{(k)}}(t) = G_{b_{(j)}}(t) = G(t), \forall k, j$. The first part is trivial. By definition and independence $G_{b_{(k)}}(t) = \mathbb{P}[t^j < t | b^j = b_{(k)}] = \mathbb{P}[t^j < t] = G(t)$.

We show the second statement for the case of two bidders, but it is easily generalizable for the case of $K$ bidders. We want to show that $\mathbb{P}[t^j < t^k | b^j = \min b, b^k > \min b] = 1/2$. By independence we just need to find $\mathbb{P}[t^j < t^k]$

$$\mathbb{P}[t^j < t^k] = \int_{-\infty}^{t^k} \int_{-\infty}^{t^k} g(t^j, t^k) dt^j dt^k$$

At the same time

$$\mathbb{P}[t^j < t^k] = 1 - \mathbb{P}[t^j \geq t^k] = 1 - \int_{t^k}^{+\infty} \int_{t^k}^{+\infty} g(t^j, t^k) dt^j dt^k.$$

Changing the order of integration and applying symmetry in the last integral yields

$$\mathbb{P}[t^j < t^k] = 1 - \mathbb{P}[t^j < t^k] \tag*{□}$$

B& T. Consistency and normality

$$\hat{\alpha} = \frac{1}{|T|} \sum_{t \in T} \frac{\hat{G}_{b_{(1)}}(t) - \hat{G}_{b_{(2)}}(t)}{H(t) - \hat{G}_{b_{(2)}}(t)} = \frac{1}{|T|} \sum_{t \in T} \frac{\frac{1}{N} \sum 1\{b_{(1)} < t\} - 1\{b_{(2)} < t\}}{\frac{1}{N} \sum NH(t) - 1\{b_{(2)} < t\}}.$$

First let’s analyze

$$\hat{\alpha}(t) = \frac{\frac{1}{N} \sum 1\{b_{(1)} < t\} - 1\{b_{(2)} < t\}}{\frac{1}{N} \sum NH(t) - 1\{b_{(2)} < t\}}.$$

By uniform consistency of the empirical CDF and Slutsky theorem,

$$\operatorname{plim}_{N \to \infty} \hat{\alpha}(t) = \frac{G_{b_{(1)}}(t) - G_{b_{(2)}}(t)}{H(t) - G_{b_{(2)}}(t)} = \alpha.$$

Since $\hat{\alpha}$ is an average of consistent estimators it is itself consistent\(^\text{32}\) Now let’s consider

$$\sqrt{N}(\hat{\alpha}(t) - \alpha) = \frac{\frac{1}{\sqrt{N}} \sum 1\{b_{(1)} < t\} - 1\{b_{(2)} < t\}}{\frac{1}{\sqrt{N}} \sum NH(t) - 1\{b_{(2)} < t\}} \tag{C}$$

$1\{b_{(1)} < t\} - 1\{b_{(2)} < t\}$ is not correlated across auction. Moreover, since we assumed $b \perp t$.

$$\operatorname{Var}[1\{b_{(1)} < t\} - 1\{b_{(2)} < t\}] = G_{b_{(1)}}(t)(1 - G_{b_{(1)}}(t)) + G_{b_{(2)}}(t)(1 - G_{b_{(2)}}(t)).$$

\(^{32}\)For two estimators $\alpha_n$ and $\beta_n$ such that $\operatorname{plim} \alpha_n = \alpha$ and $\operatorname{plim} \beta_n = \beta$ one can write a triangular inequality $|(\alpha_n + \beta_n) - (\alpha + \beta)| \leq |\alpha_n + \beta_n| + |\alpha + \beta|$ to derive the result.
The numerator of (C) converges to \( H(t) - G_{b(2)}(t) \). By C.L.T.,
\[
\sqrt{N} (\hat{a}(t) - \alpha) \overset{d}{\rightarrow} N(0, \mathbb{V}_\alpha)
\]
\[
\mathbb{V}_\alpha = \frac{G_{b(1)}(t)(1 - G_{b(1)}(t)) + G_{b(2)}(t)(1 - G_{b(2)}(t))}{(H(t) - G_{b(2)}(t))^2}
\]

A confidence interval for a fixed \( t \) is
\[
CI_{0.05}(t) = \hat{a}(t) \pm \frac{\mathbb{V}_\alpha(t) \cdot z_{N(0,1)}}{\sqrt{N}}.
\]

Hence we built a confidence interval for \( \hat{a}(t) \), for a fixed \( t \). The two functions that we are using for estimation are \( H(t) = t/\epsilon \) and \( H(t) = 1 \).

Now we need to find the asymptotic variance of \( \hat{a}(t) \). Consider \( \sqrt{N} (\frac{1}{|T|} \sum_{t \in T} \hat{a}(t) - \alpha) \).

This can be rewritten as a sum of asymptotically normal estimators. Now the question is whether they are asymptotically independent.

\[
\lim_{N \rightarrow \infty} \text{cov}(\hat{a}(t), \hat{a}(s)) = \lim_{N \rightarrow \infty} \mathbb{E}[\hat{a}(t) \cdot \hat{a}(s)] - \alpha^2 =
\]
\[
\lim_{N \rightarrow \infty} \mathbb{E}\left[\frac{(\hat{G}_{b(1)}(t) - \hat{G}_{b(2)}(t))(\hat{G}_{b(1)}(s) - \hat{G}_{b(2)}(s))}{(1 - \hat{G}_{b(2)}(t))(1 - \hat{G}_{b(2)}(s))}\right] =
\]
\[
\lim_{N \rightarrow \infty} \mathbb{E}\left[\frac{(\hat{G}_{b(1)}(t) - \hat{G}_{b(2)}(t))(\hat{G}_{b(1)}(s) - \hat{G}_{b(2)}(s))}{(1 - \hat{G}_{b(2)}(t))(1 - \hat{G}_{b(2)}(s))}\right] = \alpha^2
\]
where the last equality follows from independence of \( t \)'s. Applying Slutsky theorem several times,

\[
(\hat{G}_{b(1)}(t) - \hat{G}_{b(2)}(t))(\hat{G}_{b(1)}(s) - \hat{G}_{b(2)}(s))
\]

Hence asymptotic variance of \( \hat{a} \) is weighted average of each of the variances as a function of \( t \).

**Baseline Bids’ Method**

**Assumption 3. IPV'** Homogenized valuations \( v_{ik} \) are i.i.d. within an auction. Bidders’ homogenized strategies \( \beta \cdot () \) are monotonic transformations of \( v \).

**Assumption 4. Symmetry** Density of \( v_{ik} \) is symmetric around mean.

\[
f(\mathbb{E}v_k + \delta) = f(\mathbb{E}v_k - \delta).
\]

**Corollary**

\[
\mathbb{P}^{\text{fair}}[0 < \xi(2) - \xi(1) < \epsilon] \leq \mathbb{P}^{\text{fair}}[0 < \xi(3) - \xi(2) < \epsilon].
\]
Proof

Define order statistics spacings

\[ w_1 = \xi_{(2)} - \xi_{(1)} \]
\[ w_2 = \xi_{(3)} - \xi_{(2)} \]

We can explicitly derive the p.d.f.s of these two quantities. By Assumption 3, the joint p.d.f of the first and the second order statistic is,

\[ f_{12}(x, y) = \frac{n!}{(n-2)!} (1 - F(y))^{n-2} f(x) f(y) \]

and same for the second and the third order statistic,

\[ f_{23}(x, y) = \frac{n!}{(n-3)!} F(x) (1 - F(y))^{n-3} f(x) f(y) \]

Doing a substitution of variables \( u = x, w = y - x \) and noticing that the Jacobian is equal to one,

\[ f_{w_1}(w) = \int_{\Omega} n(n-1)(1 - F(u + w))^{n-2} f(u + w) f(u) du, \]
\[ f_{w_2}(w) = \int_{\Omega} n(n-1)(n-2) F(u)(1 - F(u + w))^{n-3} f(u + w) f(u) du. \]

where \( \Omega \) is support of the bids distribution.

Now,

\[ \mathbb{P}^{fair}[0 < w_2 < \epsilon] - \mathbb{P}^{fair}[0 < w_1 < \epsilon] = \]
\[ n(n-1) \int_{0}^{\epsilon} dw \int_{\Omega} [(n-2)F(u) + F(u + w) - 1](1 - F(u + w))^{n-3} f(u + w) f(u) du. \]

By mean value theorem it is equal to

\[ \epsilon \cdot n(n-1) \int_{\Omega} [(n-2)F(u) + F(u + \epsilon') - 1](1 - F(u + \epsilon'))^{n-3} f(u + \epsilon') f(u) du \]

where \( \epsilon' \) is some value in \( [0, \epsilon] \).

Essentially we need to show that

\[ \int_{\Omega} [(n-2)F(u) + F(u + \epsilon') - 1](1 - F(u + \epsilon'))^{n-3} f(u + \epsilon') f(u) du \geq 0 \]
First, we show that it holds for the case of three bidders. Doing a Taylor expansion for $n = 3$ one can get

$$\int_{\Omega} [(2F(u) - 1)f(u)^2 + \epsilon'(f(u)^3 + f(u)f'(u)(2F(u) - 1))]du,$$

The first part is zero by symmetry, so one needs to check whether a following is true for each symmetric distribution:

$$\epsilon' A := \epsilon' \int_{\Omega} f(u)^3 + f(u)f'(u)(2F(u) - 1)du \geq 0.$$

Consider

$$T := \int_{\Omega} f(u)f'(u)(2F(u) - 1)du,$$

Integration by part gives,

$$T = f(u)^2(2F(u) - 1)|_{\Omega} - \int_{\Omega} [f'(t)(2F(t) - 1) + 2f(t)^2]f(t)dt$$

or

$$T = f(\sup(\Omega))^2 - \int_{\Omega} f(t)^3dt$$

Hence,

$$\epsilon' A = \epsilon' f(\sup(\Omega))^2 \geq 0$$

Now we can use a Taylor expansion on the LHS of the inequality for more than three bidders and rewrite it into parts:

$$\int_{\Omega} [(n - 3)F(u)(1 - F(u))^{n-3}f(u)^2du + \int_{\Omega} [(2F(u) - 1)(1 - F(u))^{n-3}f(u)^2du$$

Using integration by parts and iterating one can get

$$(n - 3)!(n - 3) \int_{\Omega} F(u)f(u)^{n-1}du + \int_{\Omega} (2F(u) - 1)f(u)^{n-1}du.$$.

The first part is positive, while the second part is zero by symmetry.

Another representation of it is

$$(n - 3)!(n - 1)F(u) - 1)F(u)f(u)^{n-1}du.$$.

When $n$ is growing, this also tend to grow. So we either need symmetry or large enough number of bidders for the inequality of interest to hold.